Asia-Pacific Language Variation

VOLUME 8 NUMBER 2 2022

John Benjamins Publishing Company
Amsterdam/Philadelphia

Asia-Pacific Language Variation

SPECIAL ISSUE

What's so standard about standards? Variationist principles and debates

Edited by Jonathan R. Kasstan
University of Westminster

Table of contents

Introduction

What's so standard about standards?

Jonathan R. Kasstan
University of Westminster

This special issue of *Asia Pacific Language Variation* is concerned with the role that standard languages have played in the focusing of a socially accountable theory of language change. It has been acknowledged for some time now that variationist theory, like linguistic theory more broadly, has been heavily dependent on, and modelled on the properties of, uniform standardized varieties (Milroy, 2001, p. 539). Indeed, in spite of some important exceptions (e.g., Sankoff, 1980[1976]; a number of case studies in Stanford & Preston, 2009 and Hildebrandt, Jany, & Silva, 2017; as well as papers that continue to emerge in this venue) most variationist studies are (and have been) carried out in what Milroy called 'standard-language cultures', that is, 'cultures [in which] virtually everyone subscribes to the ideology of the standard language' (2001, p. 530).[1,2]

As Sharma (2017, p. 233) among others has pointed out, the motivating factor behind this narrow sampling has been the need for careful hypothesis testing and study replication from which the postulation of principles of language variation and change could emerge. Yet, the extent to which sociolinguistic theorizations have been tested on data from representative populations remains an ongoing debate in the field, as recent *états des lieux* such as Coupland (2016); Labov (2016), and Bell (2017) make clear. Owing to the fact that variationists have worked (and continue to work) largely on languages that have been heavily standardized, the potential influence of standard-language ideology on the selection

1. Nagy & Meyerhoff (2008, 2019) and Adli and Guy (2022) all offer ample quantitative evidence for this claim.

2. Standard-language ideology can be understood as a bias towards an abstracted, idealized, homogeneous spoken language which is imposed from above, and which takes as its model the written language (Lippi-Green, 1994, p. 166). This ideology, which gives rise to shared, community-wide notions of 'correct' (often interpreted as standard) and 'incorrect' (often interpreted as 'dialectal', or 'non-standard' or 'vernacular') language is found in most (if not all) European language communities and elsewhere (Kroskrity, 2010; Schieffelin, Woolard, & Kroskrity, 1998).

https://doi.org/10.1075/aplv.00016.int
Asia-Pacific Language Variation 8:2 (2022), pp. 139–149 | ISSN 2215-1354 | E-ISSN 2215-1362

of variables at the very least should give pause for thought (see Cheshire, 2005, p. 87). Yet, given the success of the earliest variationist studies, the extent to which sociolinguistic theorizations in this paradigm remain ideologically embedded in a standard-language mindset continues to go largely unquestioned.

In their seminal manifesto, Weinreich, Labov and Herzog (1968) reasoned that arriving at a socially accountable theory of language variation and change required a consideration of a set of empirical 'problems' (1968, p. 183). Among the most widely addressed of these in the variationist literature are *embedding* and *evaluation*. Since researchers want to understand how fluctuating variants are selected and propagated, it is important to understand how they are socially embedded. Therefore, as this requires understanding how communities evaluate features, Labov (1966) first proposed that methods targeting *attention to speech* can tap into evaluations when speakers style shift. While the field today acknowledges different approaches to style shifting and the evaluation of variation (see e.g., Eckert & Rickford, 2001; Eckert, 2012), Labov's observations have been so widely replicated that patterns such as those depicted in his classic New York City study have come to be anticipated as likely distributions for macro-level social categories everywhere. There have, however, been dissenting voices.[3]

Nearly thirty years ago, Dorian (1994a, p. 694) had noted that 'the idea that linguistic variation within a speech community will correlate with social differences [has come to gain] unquestioned acceptance'. Yet, her broader contention was that, in *not* questioning this idea, we miss an opportunity to establish what other kinds of sociolinguistic variation might be possible. Dorian's thinking stemmed from her own work on minoritized (and moribund at the time of writing) Embo, a variety of East Sutherland Gaelic (ESG, see e.g., Dorian, 1981). Unlike other, more prestigious varieties of Gaelic, Embo did not benefit from a written standard, and so had: 'no school teachers holding up preferred ESG models to be emulated' (Dorian, 1994a, p. 674); no prestige norm that was readily available for the purpose of standardization; and no local Gaelic-speaking professional class or elites to model. At the same time, Dorian demonstrates in her data no clearly discernible correlations between 'a profusion of variant forms' and classic variationist social factors (Dorian, 1994a, p. 633), a phenomenon that she referred to as *personal-pattern variation*.[4] She further argued that this phenomenon would not be unique to Embo, and invokes Hill and Hill's classic (1986) study of Malinche Mexicano as a comparable example, where the same sort of linguis-

3. It is noteworthy that Labov (2016) himself has warned of the risks inherent in this.

4. Dorian is also clear that 'personal-pattern' variables differ from what linguists have called 'free variation', since they can show linguistic conditioning and also speaker-specific discourse-patterning (see discussion in Dorian, 1994).

tic variability can be observed. This, Dorian contends, invites certain generalizations about the sorts of linguistic ecologies in which such variation might also be observed:

> The Malinche Mexicano speakers live in a region which is rural and isolated, as the ESG speakers do. Many members of both groups are illiterate in their home language, which is not taught in the schools. In both cases isolation plus low or absent literacy, with the insulation the latter provides from linguistic authority external to the community, may play a part in allowing more leeway for variation. (Dorian, 1994a, p. 685)

Dorian's conclusions are principally twofold. First, that there was not at the time of writing a 'movement toward reduction of variants' nor 'the development of differences in the social evaluation of most variants' (Dorian, 1994a, p. 685) in the community she was studying. Second, 'the role of external and internal prestige norming in the degree and kind of variability that a language exhibits may merit further exploration' (1994a, p. 693). More broadly, she suggests that a lack of a prestige norm may encourage the persistence of the sorts of variation she describes, which may also be recalcitrant to modelling in apparent-time, and which may carry little or no social loading. Most recently, Dorian has argued that this concern still has not been adequately addressed by a sociolinguistic theory that remains focused on standard-language cultures (Dorian, 2010, p. 277).

Dorian's warnings underscore the hegemonic domination of the standard in variationist research, which raises a number of pressing issues for linguists to overcome. For example, as Lüpke (2021) has suggested, multilingual non-standard-language communities still 'far outnumber officially monolingual or regulated multilingual nation-states' (2021, p. 139), notwithstanding the colonial violence that brought with it the transplanting of such western ideologies to many societies and cultures. Indeed, outside of northern universalist thinking, southern scholars have long drawn attention to the inapplicability or inappropriateness of the notion of discrete languages and of standards to, for example, complex African multilingual settings (Lüpke, 2021, p. 140).[5] Other issues stem from this concern: in many non-western communities, language use 'is not tied to prestige associated with particular languages' in the same way as has been theorized in the West (*ibid*). Ergo, while the functional, consensual model adopted into sociolinguistics is central to the modelling of sociolinguistic variation, in the sorts of communities that Lüpke describes, the relationship undergirding linguistic variation is

5. Milroy has also suggested that 'different methods and different underlying assumptions about the concept of "a language"' may be needed for the study of variation in language communities that are without standardization of the native language (1999, p. 36).

more likely to be non-hierarchical, and tied instead to speakers' own heteroglossic ideologies. Similar commentary has been offered by Satyanath (2015, 2021) in Indian sociolinguistics, where she claims that the relationship that best characterizes state languages vs. vernaculars, literary vs. spoken varieties, or standard vs. non-standard varieties is decidedly non-hierarchical, even if these relationships might seem on the surface to be hierarchical. Therefore, Satyanath argues, 'it does not come as a surprise that [e.g.] style shifts reported between standard vs. non-standard as experienced in [western variationist studies] are conspicuously absent across diverse settings of India' (2021, p. 15). In these language ecologies, then, the nature of the relationships that exist among languages and speakers are not best characterized by western notions of standardness or power and prestige, nor are hierarchical views of variation that have been influenced by classic works such as Bourdieu (1977) necessarily applicable (cf. Rickford, 1986). Lastly, Dorian (2010, p. 4) outlines two conditions fundamental to the variationist paradigm that are not – in so far as we are aware – universally present in contemporary societies, which might give rise to other forms of variation. One of the conditions in question is socioeconomic stratification, and the other is a set of linguistic variants that can be construed as standard variants. To paraphrase Dorian, if no variant is established as standard then this should allow for the possibility that linguistic variants might well take on no social significance relative to one another.

Given the above, it seems pertinent to continue to consider how and why language varies in societies and cultures where there are no standards or standard-language ideologies as traditionally understood in the more widely reported western contexts. At the very least, there appears to be consensus that all groups will have ideologies: Silverstein for instance has remarked that 'there is no possible *absolute* pre-ideological, i.e. zero-order, social semiotic' (1992, p. 315). Yet, it is clear from the above discussion of e.g. Embo that some communities (perhaps even a majority, cf. Lüpke, 2021) will not put the same premium on prestige forms. Owing to the centrality of notions of prestige for canonical variationist modelling, it is likely that a lack of awareness of the ways in which disparate types of language ideologies shape encounters between speakers has the potential to lead to problems (Milroy, 2001; Cheshire, 2005). However, at the same time, many such communities that are likely to harbor more heteroglossic language ideologies are also increasingly under threat. As with variationism, the field of language endangerment has also grown considerably in recent years, which points to the escalating severity of this concern. Lee (2020) cites Campbell et al. (2013) in estimating that one language ceases to be spoken every three months. Bromham et al. (2021) suggest that this figure could rise to one language every month without intervention. Therefore, while seeking out a more representative pool of language data on which to build sociolinguistic theorizations should be the ideal irrespec-

tive of the vitality of the language, this is not happening at a fast enough pace. At time of unprecedented decline in global linguistic diversity (Loh & Harmon, 2014), and as speakers of many languages age without transmitting them to future generations, the calls for more diverse sources of data should be a pressing concern for variationism in particular (Stanford, 2016). As Trudgill has suggested:

> If we are keen to learn more about the inherent nature of linguistic systems, we must urgently focus more of our attention on linguistic structures and changes of the types that occur in the ever-dwindling number of low-contact, dense social network varieties of languages in the modern world. If we want to [...] built up an accurate picture of the nature of human languages [...] we had better hurry [...] not only because most of the world's languages are in danger, but also because those that are going to be left behind will increasingly tend to be of a single historically atypical type. (2011, pp. 187–8)

The four cases studies included in this issue and summarized below all engage with the open questions and theoretical debates sketched out above. Each paper presents a quantitative analysis of variation in production or perception at different levels of linguistic description, and across disparate language sites, including Ende (Strong, Lindsay, & Drager), Guyanese (Satyanath), Raga (Duhamel), and Chanka Quechua (Povilonis & Guy).

Strong, Lindsey & Drager undertake a sociophonetic study of variable retroflex production among Ende speakers in Limol (Western Province, Papa New Guinea), where these obstruents are realized variably as stops or affricates. In and around Limol, there is what the authors describe as a regional egalitarian multilingualism, common among Ende speakers, where no local language holds more prestige than another. At the same time, there is no explicit standard Ende, and the authors report that speakers do not offer metalinguistic commentary about the use of any given form. However, the study results indicate that there is an interaction between classic age and sex variables when one further locally relevant social factor is modelled – that of orator status – which does predict stop-variant selection. The authors reason that, in this particular context, many positions of power are reserved exclusively for men. Conversely, leadership roles and positions that confer privilege are restricted, but women who do accrue sufficient prestige over time can become orators who practice *kawa*, a long-standing tradition in which select individuals perform regular public orations. The dynamics of this ecology are therefore reflected in the production of retroflex stopping, in that a speaker's gender, age and orator status play a role in speech production, which is linked to the distribution of power in the community, in the absence of an explicit standard. However, given the prestige associated with being an elder orator, it stands to reason that the forms that orators use may come to carry

some level of overt prestige over time. The Ende case study also emphasizes the importance of uncovering locally meaningful social categories that can only be understood through long-term engagement and collaboration with the community under study (see also Stanford, 2009).

Much like Dorian's work described above, Satyanath problematizes the very labelling of variants as 'standard' or 'non-standard' in sociolinguistics, and contends that where such labels are adopted by researchers outside of the West, outcomes do not necessarily map onto attitudes about language, nor do they necessarily result in the same sorts of correlations that are typically observed in western anglophone variationist studies (that are taken for granted). To demonstrate this, she considers one English-speaking setting in particular – that of Guyanese speakers in Guyana. Her study has three specific aims: to explore (a) the usefulness of the notions of standard-ness and social hierarchy as heuristics in understanding language variation in Guyanese; (b) the possible social meanings associated with variants and language variation, more generally; and (c) local language ideologies associated with English and other varieties in the ecology, namely Creole(se). In revisiting and offering an augmented analysis of production and perception data that she gathered in Georgetown (Guyana) in 1989, Satyanath offers an analysis of variation in the use of a number of variants that mark habitual aspect (and that are derived from the English periphrastic *do*), including /daz, az, a iz, z/ and Ø. Findings from her perception task suggest that participants do not assign high status scores to the most English-like variants, and that the variants in the round do not index socio-economic status, *contra* most studies of morpho-syntactic features in English contexts. At the same time, Satyanath notes a loosening of the association of traditionally creole-like variants as belonging to Creole(se), which researchers might otherwise label as 'vernacular' or 'non-standard' but which clearly have a much more neutral association for speakers themselves. Therefore, Satyanath observes a feature pool in which English variants are not more highly ranked than other variants on socio-economic status scores, and, further, that these observations in perception track findings in production too, in that speakers do not style shift to more English-like variants as greater attention is paid to speech (see also Satyanath, 1990). On the balance of evidence, it is more likely, Satyanath contends, that the findings replicated in western sociolinguistics are unique to those linguistic ecologies. Her findings raise the possibility, then, that the orientation to prestige variants as observed in standard-language cultures may be a reflex of a more general, underlying principle: that speakers orient themselves linguistically to the most important structures in their social worlds. After all, language ideologies that might drive orientation to particular features are ultimately 'socioculturally motivated ideas, perceptions and expectations of language' (Blommaert, 1999, p. 1).

Duhamel's study of variable /ɤ/-deletion in Raga (Vanuatu) takes aim at the sociolinguistic gender pattern in particular, a classic distribution reproduced in many anglophone (i.e., standard language) communities that is well-represented in most if not all variationist textbooks. In her study, Duhamel observes that younger and older men deleted /ɤ/ more in apparent-time when compared with middle-aged speakers, be they men or women. As with Strong et al.'s consideration of locally relevant social factors, the distribution of variation in Raga makes sense only when considering the activities of social groups in the community in which older men carry greatest prestige, as well as the broader change in social-economic conditions that the community is currently enjoying in North Pentecost, in which younger men are playing a leading role. In this particular context, elder men are highly revered members of their society, whose input is sought and valued by the rest of the community. Conversely, young men are unestablished members of Raga society. They are however much more likely to serve elder chiefs as part of an evening ritual of *kava* drinking, in which young men prepare the traditional drink for consumption. This offers them opportunity to silently observe the practices and conversations of their elder reference group in an hours-long ritual that is less likely to involve middle-aged men and women. As Duhamel shows, the younger men move on the variation that they identify among their higher-status elders, and adopt more /ɤ/-deletion which they see as a prestige marker. It is however noteworthy that a rise in *kava* trade has coincided with a rise in apparent-time in the deletion of variable /ɤ/. The findings from the study thus augur with Coupland's recent call for variationists to consider the reflexes of – what he terms – 'sociolinguistic change', that is, 'a consequential change over time in language~society relations' (2016, p. 433). In the case of this Raga community, the prestige associated with /ɤ/-deletion is not conferred by a standard, and it is not an overtly prestigious, community-wide norm. It is however argued to be a change, observed among elder men who enjoy significant reverence in their society, that is being adopted by the younger men in the community.

The final case study in this issue focuses on rural and urban Chanka Quechua speakers in the Peruvian city of Andahuaylas, on the Eastern side of the Chanka dialect region. Quechua is reserved for the most intimate domains of usage in and around Andahuaylas, where Spanish is the dominant, superordinate variety. While many speakers are bilingual, employing Quechua in public is often interpreted through the lens of social-class status, and so is avoided, in spite of a renaissance of authentic indigeneity in the region – in which language plays a prominent role – as well as language-planning strategies at governmental level. These efforts have led to a standard for Quechua that was introduced by the Ministry of Education. However, in spite of their efforts to augment literacy rates, a large majority of adult speakers are illiterate in Quechua. Povilonis and Guy

argue that Quechua norms are instead arbitrated at a local level, which can differ from village to village and town to town. As a result, there is no overarching standard-language ideology among Quechua speakers as typically described for standard-language communities more commonly studied in variationism (i.e., where standard-language ideology might be propagated through the widespread adoption of a written standard, in the media, or among high-status speakers). Speakers do however share in an ideology of authenticity that carries prestige as a component of indigeneity. Located in a rural imaginary of the past, this ideology reifies and totemizes an ideal monolingual Quechua speaker *quechua neto* ('pure Quechua'), 'uncontaminated' by Spanish. In this sense it is a type of sociolinguistic authenticity that is commonly observed in many other language settings (see e.g., Dorian, 1994b). To understand how these sociolinguistic dynamics play out in speech production, Povilonis and Guy consider the variable realization of the uvular /q/ phoneme in the past tense /–rqa/ morpheme. Their findings, taken from sociolinguistic interviews and a novel sentence-correction task, suggest that speakers do not offer any conscious evaluation or overt preference for either of the two forms [rqa] vs. [ra], even when explicitly asked about them. However, production evidence from their elicitation task indicates an overall implicit preference for the uvular-full pronunciation of /-rqa/ (the variant used more in conversation by speakers who are older, rural, less-educated and less Spanish-dominant), even if urbanites showed a growing acceptance and use of the uvular-less. The study in particular highlights the need to adopt a more varied understanding of standard-language ideology (as Lüpke, 2021 indicates, above), where doing so requires seeking out alternative community norms and priorities.

To conclude, while strides are being made in exporting revised variationist methods to lesser-studied communities, absent standards or standard-language ideologies (not least in the present venue), gaps in our knowledge remain, and this issue assembles the pieces of further local understandings of such notions to provide an augmented picture of how linguistic variation can manifest.

Acknowledgements

This paper emerged from a panel co-organized by James N. Stanford for NWAV48 (University of Oregon). I am grateful to him for his contributions in the development of this project from the start, as well as those contributions offered by the panel's discussants, Miriam Meyerhoff and Naomi Nagy, whose thoughts have also strongly influenced this issue and the present paper. The contents of this paper have also been informed by the research undertaken by the issue's contributors, who have worked diligently in addressing the research themes and open questions posed in this paper. Lastly, I thank the reviewers who worked to make this issue possible, and the journal's editor, Shobha Satyanath, for her own insightful comments and suggestions as this project has developed.

References

Adli, Aria, & Guy, Gregory (2022). Globalising the study of language variation and change: A manifesto on cross-cultural sociolinguistics. *Language and Linguistics Compass 16*(5). e12452. https://doi.org/10.1111/lnc3.12452

Bell, Allan (2017). Giving voice: A personal essay on the shape of sociolinguistics. *Journal of Sociolinguistics*, *21*(5), 587–602. https://doi.org/10.1111/josl.12268

Bourdieu, Pierre (1977). The economics of linguistic exchange. *Social Science Information 16*, 645–668. https://doi.org/10.1177/053901847701600601

Bradley, David, & Bradley, Maya (2019). *Language Endangerment*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781139644570

Bromham, Lindell, Dinnage, Russel, Skirgård, Hedvig, Ritchie, Andrew, Cardillo, Marcel, Meakins, Felicity, Greenhill, Simon, & Xia Hua (2021). Global predictors of language endangerment and the future of linguistic diversity. *Nature: Ecology & Evolution*, *6*, 163–173.

Campbell, Lyle, Lee, Nala Huiying, Okura, Eve, Simpson, Sean, Ueki, Kaori, & Van Way, John (2013). New knowledge: findings from the Catalogue of Endangered Languages (ELCat). Paper presented at the 3rd International Conference on Language Documentation and Conservation, Honolulu, Hawai'i, Feb. 28.

Cheshire, Jenny (2005). Syntactic variation and spoken language. In: L. Cornips & K. Corrigan (Eds.), *Syntax and variation: Reconciling the biological and the social*. Amsterdam: John Benjamins, 81–106. https://doi.org/10.1075/cilt.265.05che

Coupland, Nikolas (2016). Five Ms for sociolinguistic change. In: Nikolas Coupland (Eds.), *Sociolinguistics: Theoretical debates*. Cambridge: Cambridge University Press, 433–455. https://doi.org/10.1017/CBO9781107449787.021

Dorian, Nancy C. (1981). *Language Death: The life cycle of a Scottish Gaelic dialect*. Pennsylvania: University of Pennsylvania Press. https://doi.org/10.9783/9781512815580

Dorian, Nancy C. (1994a). Varieties and variation in a very small place: Social homogeneity, prestige norms, and linguistic variation. *Language*, *70*(4), 631–698. https://doi.org/10.2307/416324

Dorian, Nancy C. (1994b). Purism vs. compromise in language revitalization and language revival. *Language in Society*, *23*(4), 479–494. https://doi.org/10.1017/S0047404500018169

Dorian, Nancy C. (2010). *Investigating Variation: The Effects of Social Organization and Social Setting*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195385939.001.0001

Eckert, P. (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annu. Rev. Anthropol*, *41*, 87–100. https://doi.org/10.1146/annurev-anthro-092611-145828

Eckert, Penelope, & Rickford, John R. (Ed.) (2001). *Style and sociolinguistic variation*. Cambridge: Cambridge University Press.

Hildebrandt, Kristine A., Carmen, Jany, & Wilson, Silva (Eds.) (2017). Documenting variation in endangered languages. *Language Documentation & Conservation Special Publication*, *13*, 1–5.

Hill, Jane. H. & Hill, Kenneth (1986). Variable developments of -āškā 'possession' in Modern Mexicano (Nahualt). *International Journal of American Linguistics*, *52*, 404–410. https://doi.org/10.1086/466031

Kroskrity, Paul V. (2010). Language Ideologies--Evolving Perspectives. In Jürgen Jaspers, Jan-Ola. Östman, & Jeff Verschueren (Eds.), *Society and Language Use*. Amsterdam: John Benjamins, 192–211. https://doi.org/10.1075/hoph.7.13kro

Labov, William ([1966] 2006). *The social stratification of English in New York City* (2nd edition). Cambridge: Cambridge University Press.

Labov, William (2016). Afterword: Where are we now? *Journal of Sociolinguistics*, *20*(4), 581–602. https://doi.org/10.1111/josl.12200

Lee, Nala H. (2020). The status of endangered contact languages of the world. *Annu. Rec. Linguist. 6*, 301–318. https://doi.org/10.1146/annurev-linguistics-011619-030427

Lippi-Green, Rosina (1994). Accent, standard language ideology, and discriminatory pretext in courts. *Language in Society*, *23*(2), 163–198. https://doi.org/10.1017/S0047404500017826

Loh, Jonathan & Harmon, David (2014). *Biocultural diversity: Threatened species, endangered languages*. Zeist: WWF Netherlands.

Lüpke, Friederike (2021). Standardization in highly multilingual national contexts. In: Wendy Ayres-Bennett & John Bellamy (Eds.), *The Cambridge handbook of language standardization*. Cambridge: Cambridge University Press, 139–169. https://doi.org/10.1017/9781108559249.006

Milroy, James (2001). Language ideologies and the consequences of standardization. *Journal of Sociolinguistics*, *5*(4), 530–555. https://doi.org/10.1111/1467-9481.00163

Nagy, Naogy, & Meyerhoff, Meyerhoff (2008). The social lives of linguistics. In M. Meyerhoff & N. Nagy (Eds.), *Social lives in language – Sociolinguistics and multilingual speech communities* (pp. 1–7). Amsterdam: John Benjamins. https://doi.org/10.1075/impact.24.02nag

Nagy, Naogy, & Meyerhoff, Meyerhoff (2019). The role of standards in the field of variation. Paper presented at New Ways of Analyzing Variation (NWAV) 48, University of Oregon.

Sankoff, Gillian (1980[1976]). *The social life of language*. Philadelphia: The University of Pennsylvania Press.

Satyanath, Shobha (1990). The use of pronouns and its implications for the implicational model of decreolozation in Guyanese. *The Penn Review of Linguistics*, *14*, 129–142.

Satyanath, Shobha (2015). Language variation and change: The Indian experience. In: Dick Smakman, & Patrick Heinrich (Eds.), *Globalising Sociolinguistics: Challenging and expanding theory* (pp.107–122.). London: Routledge.

Satyanath, Shobha (2021). Geneologies of sociolinguistics in India. *Journal of Sociolinguistics*, *25*(5), 762–784. https://doi.org/10.1111/josl.12496

Schieffelin, Bambi, Woolard, Kathryn & Kroskrity, Paul V. (Eds.) (1998). *Language ideologies, theory and practice*. New York: Oxford University Press.

Sharma, Devyani (2017). World Englishes and Sociolinguistic Theory. In M. Filppula, J. Klemola, & D. Sharma (Eds.), *The Oxford handbook of World Englishes* (pp. 232–251). Oxford: Oxford University Press.

Silverstein, M. (1992). The uses and utility of ideology: some reflections. *Pragmatis*, *2*(3), 311–323.

Stanford, James N. (2009). Clan as a sociolinguistic variable: Three approaches to Sui clans. In: J. N. Stanford & D. R. Preston (Eds.), *Variation in indigenous minority languages*. Amsterdam: John Benjamins, 463–484. https://doi.org/10.1075/impact.25.23sta

Stanford, James N. (2016). A call for more diverse sources of data: Variationist approaches in non-English contexts. *Journal of Sociolinguistics*, *20*(4),525–541. https://doi.org/10.1111/josl.12190

Trudgill, Peter (2011). *Sociolinguistic typology: Social determinants of linguistic typology*. Oxford: University Press.
Weinreich, Uriel, Labov, William, & Herzog, Marvin (1968). Empirical foundations for a theory of language change. In: Winfred P. Lehmann & Yakov Malkiel (Eds), *Directions for Historical linguistics* (pp. 95–195). Austin: University of Texas Press.

Address for correspondence

Jonathan R. Kasstan
School of Humanities
University of Westminster
309 Regent Street
London W1B 2HZ
United Kingdom
j.kasstan@westminster.ac.uk

Author Queries

- Please provide a complete reference for the citation '(Stanford & Preston, 2009), (Rickford, 1986), (Dorian, 1994)' in this article.
- Please provide a citation for the reference id "CIT0004 (Bradley, David, & Bradley, Maya (2019)), CIT0036 (Trudgill, Peter. (2011))" since citation is missing in the article.

Kawa and the variable stopping of obstruents in Ende

Katherine Strong[1], Kate L. Lindsey[2], and Katie Drager[1]
[1] University of Hawai'i at Mānoa | [2] Boston University

This paper details a study investigating sociophonetic variation in Ende, a language spoken in Papua New Guinea. The study examines speech produced by 30 individuals, investigating what social and linguistic factors are linked with the variable alternation between stopped and affricated realizations of Ende retroflex obstruents [t͡ʂ~ʈ] and [d͡ʐ~ɖ].

Our analysis provides evidence that the obstruents in question are more likely to be realized as stops when they are voiced and when the speakers are orators. Orators are people who practice *kawa*, a long-standing practice whereby select individuals perform regular public orations. Among orators, the speaker's age also appears to play a role in retroflex stopping. The link between the stop variant and social factors can be understood within the context of the distribution of power in the community, even in the absence of any explicit standard.

1. Introduction

This paper presents patterns of sociophonetic variation in Ende, a Pahoturi River language spoken by the Ende *tän mit* in Western Province, Papua New Guinea. This study explores the distribution of Ende retroflex obstruents [t͡ʂ~ʈ] and [d͡ʐ~ɖ] and what social and linguistic factors may be linked with their variable affrication. This work also highlights the locally relevant social factor of community orator, a prestigious role through which individuals in the Ende community perform daily public orations in a long-standing practice called *kawa* (Lindsey, 2019a, pp. 242–243). Much like how the locally relevant class categories of Estate Class and non-Estate Class in Rickford (1986)'s study of Guyanese creole use in Cane Walk revealed ideological patterns in social class and linguistic practice, the locally relevant distinction of orator in Limol allows us to consider how gender and power operate within the community and map onto linguistic variation. This work speaks to the dual necessities of (1) studying variation in understudied lin-

https://doi.org/10.1075/aplv.21003.str
Asia-Pacific Language Variation 8:2 (2022), pp. 150–173. ISSN 2215-1354 | E-ISSN 2215-1362

guistic settings, including indigenous, minority, endangered, under-documented, or otherwise underrepresented languages, and (2) ensuring that, in doing so, locally meaningful social categories and community-specific social structures are considered alongside more commonly studied demographics. With these goals in mind, we continue the long tradition of ethnographic sociolinguistics (pioneered by Milroy [1980] and Rickford [1986] and notably by Sankoff [1980(1976)] in Papua New Guinea) following a second-wave variationist approach in that our analysis draws on a local practice (i.e., *kawa*) identified through ethnographic methods. Further, we consider how other aspects of an individual speaker's identity might help inform patterns of variation observed in their speech. This study contributes to a new but growing body of work looking at variation in southern New Guinea (see Ellison, Evans, Kashima, Schokkin, & Williams, 2016; Schokkin, 2017, 2018, 2021; Barth, Schokkin, Travis, Lindsey, & Stanford, 2019; Kashima, Lindsey, Schokkin, & Strong, 2019; Lindsey, 2021b) and emerges from the mutually beneficial intersection of variationist sociolinguistics and language documentation (see Hildebrandt, Jany, & Silva, 2017; Meyerhoff, 2015, 2017; Nagy, 2017).

Since the inception of the variationist framework in the 1960s (Labov, 1966, 1969), sociolinguistic studies have centered around particular languages that are predominant in Western contexts, focusing primarily on the speech of monolingual English communities (Hildebrandt et al., 2017; Nagy & Meyerhoff, 2008). Therefore, there is understandable concern about the generalizability of dominant sociolinguistic theories that have been developed based on such a small subset of the world's languages (Rickford, 1986; Smakman & Heinrich, 2015, p. 270). Indigenous, minority, endangered, and otherwise under-represented languages are frequently overlooked. However, the potential for differences in social stratification is high in the contexts where these languages are spoken (see e.g., Barth et al., 2019; Clarke, 2009; Skilton, 2017; Suokhrie, 2016). Mansfield and Stanford (2017, p. 117) call this the *Principle of Sociolinguistic Distance*, arguing that the greater the distance between a lesser-studied community and traditional sociolinguistic settings, the higher the likelihood that such research will pose challenges to current sociolinguistic theories. Further, it is ideal for variationist analyses to be conducted on understudied languages not only to contrast and compare with existing models, but also to arrive at theories of language that are representative of other populations than only those that are easily studied by scholars at Western academic institutions.

While predetermined categories of age, gender, ethnicity, and socioeconomic status have consistently been found to be closely linked with socially-conditioned phonetic variation (e.g., Eckert, 1989; Labov, 1990; Wassink, 2015), additional or alternative social categories that are meaningful to the speech community

can be identified through ethnographic research. As mentioned above, Rickford's pioneering work on Guyanese Creole complicated previous notions of class as defined for Western communities, while more recent work on Sui (Stanford, 2009) and languages of Vanuatu (Meyerhoff, 2015) have shown the importance of complicating traditional social variables such as class and bringing in new ones, such as clan. Of course, this second-wave approach is not only applied to lesser-studied languages (see for example, studies on English-speaking Jocks and Burnouts at a Detroit high school [Eckert, 2000], Mandarin-speaking yuppies and government employees in Beijing [Zhang, 2005], and English-speaking Mobile Black Professionals, Hood Kids, and Bikers in the African American community of Rochester, New York [King, 2018], but examples from lesser-studied languages highlight the benefits of ethnographic documentation in understudied speech communities and demonstrates how social categories identified through this approach can be incorporated into variationist analyses (see, e.g., Skilton [2017], who outlines phonological and morphological variation in Máíhɨki [Peru] through the lens of speaker membership in early life communities of practice, and Clarke [2009], who investigates phonological variation in the aboriginal Sheshatshiu community [Canada]).

The current study draws on a previous analysis of phonetic variation in Ende, in which Lindsey (2019b, 2021b) demonstrates how final /n/-elision in Ende can be best understood through strong associations with the local practice of community oration. To meet the need for more work that is "pattern-driven, rather than variable-driven" (Hay & Drager, 2007, p. 88), the present study examines the relationship between this practice of oration and a second variable: retroflex affrication. This study furthers our understanding of the variation of retroflex obstruents in the context of Ende more broadly and examines how this particular variable aligns with, or contradicts, greater patterns of variation in the language. If community orator status is linked with patterns of realization for multiple variables, this may allude to a relationship between style and social categorization, necessitating a third wave examination of orator speech style in further research (Eckert, 2012).

2. Ende

2.1 Language background

Ende (ISO 639-3 code: kit) is spoken by the Ende *tän mit* ('tribe') primarily in Limol, Malam, and Kinkin in Western Province, Papua New Guinea. The Ende Language Committee was established in 2003 by Warama Kurupel (Suwede) with

Wagiba Geser and Tonny (Tonzah) Warama[1] for the primary aim of Bible translation and community literacy.[2] The committee currently engages in on-going community-emergent linguistic work to document the Ende language, among other programs such as orthography development, technology workshops, and literacy-building efforts. All of the data analyzed in this study are drawn from the Language Corpus of Ende and other Pahoturi River Languages (Lindsey, 2015), which contains over 100 hours of recorded material in Ende collected between 2015 and 2018. There are between 600 and 1,000 speakers of Ende (Lindsey, 2019, p. 123). A type of regional egalitarian multilingualism is commonplace among Ende speakers where no local language holds more prestige than another. Along with English as the primary medium of education and religion, surrounding local languages Gogodala, Kawam, Taeme, Em, and Bitur are also commonly spoken by the Ende population due to long-standing traditions of marriage exchange among clans, subclans, and neighboring settlements (Lindsey, 2019, p. 241). Likewise, there is no power differential between clans and, even if there were, there is no evidence of clan-based variation in Ende and the relevance of clan affiliation appears to be decreasing.

2.2 Gender among the Ende *tän mit*

In the Ende patriarchal social system, men hold primary power. Community defined leadership roles, such as chief, community secretary, and community treasurer, as well as paid government positions, such as judge (magistrate), police officers, and health workers, are positions restricted to men. Leadership roles for women are limited to pastors, teachers, and positions in local women's committees.

Daily work is heavily segregated, with women doing most of the heavy lifting including raising the children, obtaining food and water, maintaining the home, and caring for neighbors and the elderly. Daughters are coveted for the amount of labor they perform for the family. Men begin hunting and cutting grass as teenagers and learn to build houses, shape canoes, fall large trees, and build garden fences when they start their own families. A husband's employment may bring in money or influence that can offset the women's labor.

1. In using parentheses, we are keeping with community practices, whereby names in parentheses after the second name indicate their grandfather's name, and names in parentheses after the first name indicate nicknames.

2. While Ende has no explicit standard in that there appears to be no metalinguistic commentary about the use of any given form, the Ende Language Committee has established an orthography. The orthography is based largely on the forms produced by the committee members, and it is possible that this could affect future perceptions of standardness.

A traditional marital practice of extra-local sister exchange, by which two men exchange their sisters (usually across village lines) in order to marry, is an impediment for women seeking upward economic and social mobility. Because of this practice, many married women find themselves in communities where they are cultural and linguistic outsiders with extremely limited avenues to power. Even Ende women who do marry within the Ende *tän mit* struggle to gain leadership roles. From an early age, daily chores take precedence over their education, which is a requirement for a paid government position. In adulthood, these same chores often prohibit them from engaging in leadership activities. Thus, the main avenue to power for women is to reduce one's duties through marriage to an employed man or by having an abundance of children and gardens. An abundance of children provides daughters, or sons who marry women, who can then assist in the daily labor. Likewise, a surplus of food from a flourishing garden can be used to barter for other women's work. Once women achieve a high enough social status through such avenues, they can begin to influence the community through the practice of *kawa*, thereby gaining even higher status within the community.

2.3 Kawa public oration

In the Ende community, there is a traditional practice of public oration called *kawa* (see Lindsey, 2019: Appendix B.5.1 Public Oration for more details). Practitioners of *kawa* make daily speeches, typically in the early morning, about local news, teachings, goings-on, and other matters of importance in the village. These speeches can be performed while walking, in a public square, or outside of specific houses whose inhabitants are the intended recipients of the particular speech. *Kawa* performances are highly prestigious. As described by *kawa* practitioner, Wagiba Geser, "members of the community, particularly those in positions of respect and leadership, instruct other members of the community in best practices for living a good life" (Lindsey, 2019, p. 242). In this way, speakers use *kawa* to command respect and act on the responsibility of instructing, informing, admonishing, and persuading the broader community. Being a practitioner of *kawa* is not an official status within the community, unlike, for example, clan affiliation. Instead, orator status is based on who Lindsey observed orating between 2015 and 2018. Throughout this paper, we follow Lindsey (2021b) in using the category 'orator status' as a way to distinguish those who have been observed practicing *kawa* and those who have not.

Given that previous work demonstrates a link between orator status and n-elision (Lindsey, 2021b), it is possible that other linguistic variables may also be linked with whether or not the speaker practices *kawa*. Thus, in the current study we examine a second variable: retroflex obstruents.

2.4 Retroflex obstruents in Ende

The variable examined in this study is the Ende retroflex affricate, of which there is a voiceless /t͡ʂ/ and a voiced phoneme /d͡ʐ/ (Lindsey, 2021a). These obstruents are variably realized as affricates [t͡ʂ d͡ʐ] and stops [ʈ ɖ] in production. For example, compare the stop in Figure 1 with the affricate in Figure 2.

Variation between retroflex stops and affricates is attested for other languages but is described as free variation (see, e.g., Echeverría & Contreras, 1965, p. 133 for Araucanian). We hypothesize that the variation among retroflex obstruents in Ende is not free, but is conditioned by some of the linguistic factors outlined in Section 3.4 or some of the social factors examined: gender (Section 3.2), orator status (Section 3.3), and age. Age was included to explore the possibility that the observed variation reflects a change in progress. If Lindsey's (2019, p. 140) suggestion that retroflex obstruents in Pahoturi River languages can be reconstructed as stops in the proto-language is correct, then variable retroflex affrication in Ende could be a sound change in progress. In that case, we would expect older speakers to produce a higher proportion of tokens as stops and younger speakers to produce a higher proportion as affricates.

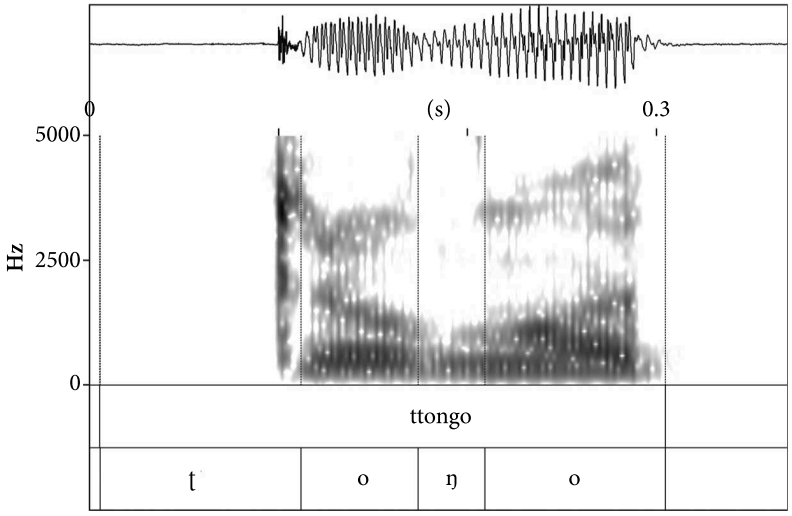


Figure 1. Waveform and spectrogram of [ʈ] from the word *ttongo* 'one, another, a, next', with realization of the variable as a stop. The characteristics of a voiceless stop visible in the waveform include a release burst (a sudden high-intensity pulse after a period of silence) followed by limited frication (irregular pulses) before quickly transitioning into the vowel (regular pulses)

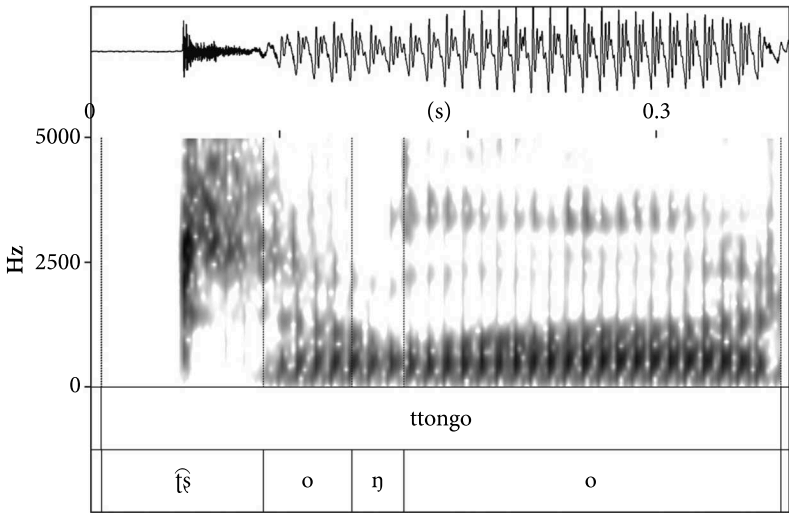


Figure 2. Waveform and spectrogram of [ʈ͡ʂ] from the word *ttongo* 'one, another, a, next', with realization of the variable as an affricate. The characteristics of a voiceless affricate visible in the waveform include a release burst (a sudden high-intensity pulse after a period of silence) followed by extended unvoiced frication (irregular pulses with no voicing bar visible in the spectrogram) before quickly transitioning into the vowel (regular pulses and visible voicing bar)

3. Methods

3.1 Data

To address whether Ende retroflex obstruents show variable affrication, this study used auditory analysis to examine the spontaneous speech data of 30 speakers from Limol village. The data are drawn from sociolinguistic questionnaires recorded in 2018 for the Ende Language Corpus (Lindsey, 2015), with durations of about 20–40 minutes per speaker. The questionnaire, which was conducted monolingually in Ende by the second author, was tailored to investigate particular social and linguistic practices of the community, drawing on both instruction from Wagiba Geser and Warama Kurupel (Suwede) and eleven months of ethnographic fieldwork by the second author in Limol. Moreover, the questionnaire was adjusted during collection to reflect new insights gathered from speakers within the community. In this way, we learned that the numerical measure of someone's age is not culturally significant; instead, speakers group themselves into four generations or cohorts based on speaker home village, education, and shared life history (Lindsey, 2021b). In the last seventy-five years, the geographic location of Limol village changed twice (once 60 years ago, and again 45 years ago), and the

educational opportunities changed twice (secondary education was introduced 45 years ago and elementary education in Limol was introduced 30 years ago). Speakers thus group themselves into approximately 15-year age cohorts based on this shared history. In our analysis, we grouped speakers into groups of approximate ages 15–29, 30–45, 46–61, and older than 62. Just as numerical age was not as important a metric as we may have predicted, we also found that clan affiliation seems to be losing relevance within the community with many younger speakers unaware of their mothers' clan group or even their own.

3.2 Speakers

Of the 73 individuals who participated in the interviews, 30 were selected for the current study to provide as balanced a sample as possible in terms of gender, age, and orator status. The following social variables were annotated for each speaker: age, gender, orator status, hometown, marital status, and (sub-)clan affiliations. The demographic information is provided by Lindsey (2019, p. 146) and is summarized for age, gender, and orator status in Table 1. Age as a factor is categorized into the four emic groups described by Lindsey (2019b). Gender is categorized as "man" and "woman" based on self-reporting and social presentation; we do not have data from non-binary or transgender people that we are aware of at this time.

Table 1. Summary of demographic information from the 30 speakers selected for this study, shown across age, gender, and orator status

	Orator					Non-orator				
	15–29	**30–45**	**46–61**	**62+**	**Total**	**15–29**	**30–45**	**46–61**	**62+**	**Total**
Men	2	2	2	2	**8**	2	2	2	1	7
Women	1	2	2	2	7	2	2	2	2	**8**
Total	**3**	**4**	**4**	**4**	**15**	**4**	**4**	**4**	**3**	**15**

Among the non-orators included in Table 1 is one speaker, Jubli Sowati, who has a social status that requires elaboration. Jubli is categorized as a non-orator because he was not observed by the second author engaging in *kawa*. However, during the 52 weeks that the second author was in Limol, Jubli was only in Limol for holidays. He lives outside of Limol in nearby Upiara, where he works as an English teacher, a job that carries a great deal of prestige in Limol. His father holds

a prestigious position within the community called the recorder.[3] Thus, while Jubli is not an orator, as defined, he holds a different and more prestigious position than other non-orators, especially when compared with others from his age group (15–29). Despite having a very different social profile from the other non-orators in the youngest age group, his data was analyzed for this study as a way to shed light on the extent to which the variable being investigated is indexed directly with the social role of being a *kawa* practitioner or if, instead, the stopped variant is used by practitioners of *kawa* because it is linked with power and prestige more generally. As a result of his particular social profile, we do not include his data in the statistical analysis presented herein but instead mark it separately in the data plots so that patterns in his speech can be compared with others in his cohort.

3.3 Extraction and auditory analysis

Tokens containing retroflex obstruents were extracted from recordings of the sociolinguistic interviews in ELAN using a structured search with regular expressions. Ffmpeg audio converter commands were then utilized to process the extractions into individual audio files. We treated the variable as categorical and binary (affricate or stop), and the first author coded all tokens by listening to each individual audio file over headphones (SONY dynamic stereo MDR-7506). 73 tokens were removed prior to analysis due to ambiguity or elision of the variable of interest.

This study examined 2,686 tokens in total, with 67% of tokens coded as an affricate ($n = 1{,}791$) and 33% coded as a stop ($n = 895$). Each token was annotated for multiple linguistic and social factors, outlined below. The third author coded 10% of the tokens to check for intercoder reliability. The coders agreed on 69% of the total tokens, with the second coder identifying more of the tokens as stops. Despite the relatively low intercoder reliability, the results and discussion presented in Sections 4–5 are valid because of the consistency and direction of the difference; all of the tokens which were marked differently by the two coders were perceived as an affricate by the first author and as a stop or flap by the third author, indicating that they have different perceptual boundaries but their coding was consistent. While consistency was similar across men and women, intercoder consistency was higher for non-orators (77%) than orators (59%). This means that the reported difference between the orators is, if anything, smaller than the difference would be if all of the tokens were coded by the second coder. Therefore, the results presented below are from the analysis of the full dataset coded by the

3. A recorder is someone who takes notes from meetings and prepares documents for official business.

first author. Further acoustic analysis is being planned in order to investigate the variable in a way that more accurately reflects its gradience.

3.4 Linguistic factors

In addition to each speaker's age, gender, and orator status, the following linguistic factors were annotated for each token: voicing of the segment, prenasalization, the type and quality of the preceding and following segments, position within the syllable, position within the word, preceding retroflex in the same utterance, lexical frequency, and grammatical category. Lexical frequency was calculated as the count of a given lexical form from a spoken corpus of Ende containing 210,000 words.

4. Results

4.1 Quantitative patterns

Using the R package ggplot2 (Wickham, 2016), the raw data were plotted to examine patterns. The plot in Figure 3 provides evidence that the likelihood of producing a stop is linked with three social factors: speaker age (four age groups), gender (women vs. men), and orator status (orator vs. non-orator). There is a marked difference in behavior across orators and non-orators, whereby orators, as a whole, produce a larger number of their tokens as stop, and exhibit much more regular patterning across age and gender. Among the orators, when comparing each consecutive age group, the older group produces a higher percentage of tokens as stops compared to the younger group, and, within each age group, the women produce a higher percentage of tokens as stops than the men do, except for the youngest group.

The statistical analysis presented herein focuses on two logistic regression models with mixed effects. Models were fit using the package lme4 (Bates, Maechler, Bolker, & Walker, 2019; R Core Team, 2019). Both are the maximal models that converged and were justified through model comparisons using ANOVA, in line with recommendations by Barr and colleagues (2013). Prior to statistical analysis, data from one of the speakers, Jubli Sowati (JSS), was removed due to his distinct social profile (see Section 3.2). The first model was fit to data from the remaining 29 speakers, with obstruent realization (stop vs. affricate) as the dependent variable. Orator status and voicing were the independent variables, and the model included random intercepts of word and speaker. Also included was a by-speaker random slope for voice since the relationship between voice

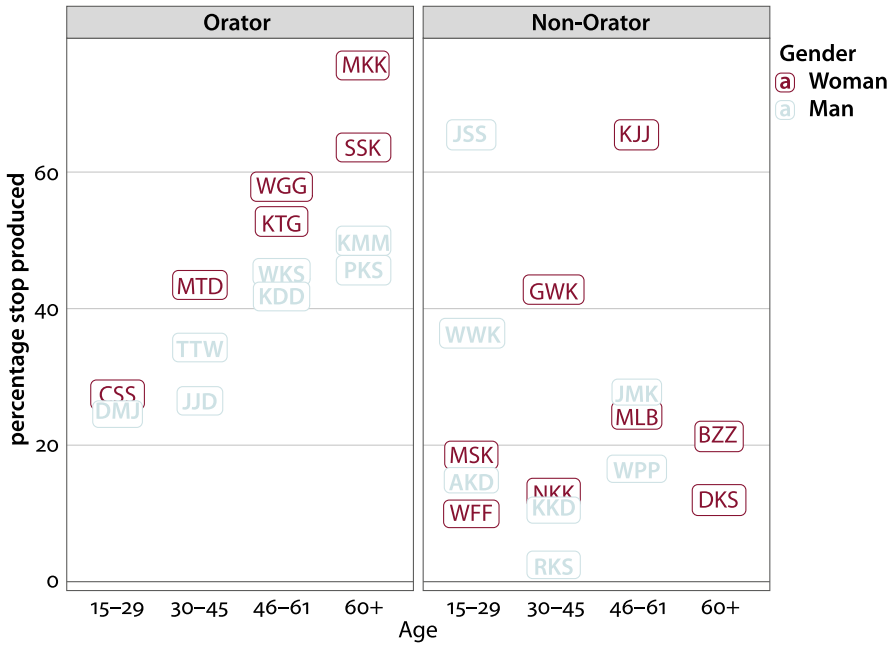


Figure 3. By-speaker percent of tokens realized as stops by orators (left panel) versus non-orators (right) across age, with women shown in dark purple and men in light green. Two speakers, BMK and DKS, produce a similar percentage of stops and are therefore indistinguishable from each other on the plot

and the dependent variable was not necessarily consistent across all speakers. The structure of the model is shown in the caption for Table 2. Syllable position was originally considered as a fixed effect, following prior work with a smaller number of data points (Strong, Lindsey, & Drager, 2020), but including it in the model fit to the current dataset did not significantly improve model fit so it was removed. Convergence was not reached when both age and gender were included as fixed effects. Gender alone did not improve the fit of the model so was removed. Likewise, word frequency (logged) and word category are not included in the model since they only improve model fit (for word frequency) and allow for model convergence (for word category) when word is not included as a random intercept in the model. The model output is presented below in Table 2.

There is a significant effect of orator status ($p < .0001$) and voicing ($p < .0001$) on the likelihood of affrication. Tokens are more likely to be realized as a stop by speakers who are orators compared with non-orators, as evident in Figure 3, and when voiced. These effects are discussed further in Sections 5.1 and 5.3, respectively.

Table 2. Output of model fit to data from 29 speakers, orators and non-orators: Glmer(Affricate.vs.Stop ~ Orator.Status + Voicing + (1 | Word) + (1+Voicing | Speaker), data=Endedata, family=binomial)

Fixed effects	Estimate	Std. error	Z value	Pr(>|z|)
(Intercept)	−0.9542	0.2777	−3.436	0.0006
Orator = yes	1.5233	0.3402	4.477	<.0001
Voicing = voiceless	−1.22	0.2426	−5.029	<.0001

In order to test whether the observed trends among the orators reach statistical significance, a second model was fit to only the tokens produced by the fourteen orators. In this more restricted model of 1,187 tokens, obstruent realization was the dependent variable, and age and voicing were included as independent variables. Because the model did not converge with the random slope included in the model fit to the full dataset, the analysis of the more restricted dataset used a simpler model with speaker and word as random intercepts.

Table 3. Output of model fit to data from the 14 orators: Glmer(Affricate.vs.Stop ~ Age + Voicing + (1 | Word) + (1 | Speaker), data=Orator.data, family=binomial)

	Estimate	Std. error	Z value	Pr(>|z|)
(Intercept)	−0.3693	0.3584	−1.030	0.3028
Age: 30–45	0.3854	0.3526	1.093	0.2744
Age: 46–61	1.1878	0.3462	3.431	0.0006
Age: 62+	1.4998	0.3527	4.253	<.0001
Voicing = voiceless	−1.4344	0.2411	−5.950	<.0001

The restricted model indicates that orators in the second to oldest age group ($p < .001$) and the oldest group ($p < .0001$) are significantly more likely to produce stops than those in the youngest group, and that the voiced variable was significantly more likely to be realized as a stop compared to the voiceless variable ($p < .0001$). The age-based difference may be indicative of a change in apparent time or may result from age grading, where orators increasingly adopt the stop variant over the course of their lifetimes. The tendency for the voiced phoneme to be realized as a stop more often than the voiceless phoneme is discussed further in Section 5.3.

As is evident in Figure 3, women orators in the three oldest age groups produce a larger percentage of tokens as a stop compared to men in the same age groups. However, including speaker gender in the model results in an overfit

model. Speaker gender reaches significance in the overfit model and also in a model that was not overfit but that did not include a random intercept for speaker. No age- or gender-related effects were observed among the non-orators when evaluated with a similar model.

4.2 Voice, stops, and affricates

Across both analyses, the results demonstrate a strong tendency for a larger number of the voiceless tokens to be realized as an affricate compared to the voiced tokens. This finding is consistent with a cross-linguistic tendency for voiceless stops to be more likely to undergo affrication than voiced stops. For example, in a typological study of 63 languages, Hall and Hamann (2006) find that voiced stops cannot undergo assibilation (e.g., di → dzi) unless voiceless stops also do (e.g., ti → tsi). The authors motivate this implication by appealing to the phonetic properties of assibilation contexts. The release of coronal stops (t, d) co-occurs with a turbulent friction phase that is significantly longer when released into a high front vocoid (i, j) than if the same stop were released into a non-high or non-front vocoid (Clements, 1999; Kim, 2001; Ohala, 1983). Crucially, this friction phase is significantly longer after voiceless stops than after voiced ones (Hall, Hamann, & Zygis, 2006), an observation that also holds for velar stops (Guion, 1998).

The results presented herein differ from previous work in that the variation in manner is not categorical, nor is it restricted to only some phonological contexts (e.g., when preceding /i/). While the largest percentage of affricates is observed for tokens preceding /i/ compared to those that precede other vowels, the number of tokens that precede /i/ is relatively small ($n = 20$),[4] the difference is subtle, and the observation is only true when considering the voiced and voiceless tokens together; when considering the voiceless tokens alone, the smallest percentage of stop tokens is observed when the following segment is /o/ (see Figure 4), the most frequent environment for retroflex obstruents.

If we assume retroflex obstruents in Pahoturi River languages can be reconstructed as stops (see Lindsey, 2019, p. 140), then the age-based variation described herein likely stems from a change in progress observed in apparent time. However, the phonetic account provided by Hall and colleagues cannot serve as the impetus for the difference between voiced and voiceless tokens that we observed.

4. The small number of tokens with following /i/ is unsurprising given that retroflex consonants typically condition backing of preceding and following vowels (Flemming, 2003).

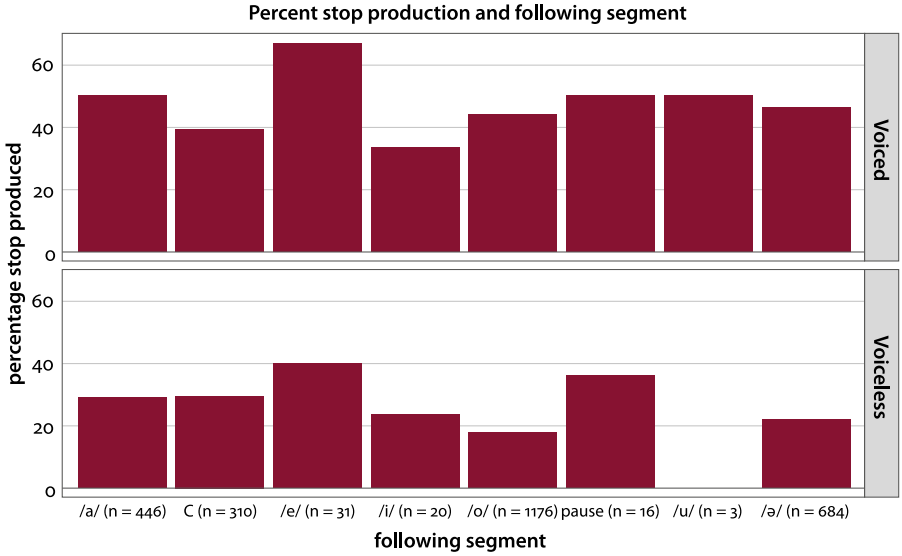


Figure 4. The percentage of tokens within each following segment category that were realized as a stop. The number of tokens within each category is shown in parentheses

5. General discussion

5.1 Group patterns

The quantitative analysis presented in Section 4 provides evidence that the likelihood of affrication is linked with a combination of social factors: speaker age, participation as a community orator, and, to a lesser extent, speaker gender. Compared with non-orators, orators are more likely to realize retroflex obstruents as stops, and, among the orators, older speakers are more likely to realize tokens as stops compared with younger speakers. Given the prestige associated with being an orator and being older, it stands to reason that the forms orators use may carry some level of prestige. While gender is not included as a fixed effect in the models presented herein because it leads to a near singular fit, we believe that the effect of gender is legitimate given that most people who are known to orate in Limol are included in this study, and we observe an effect of gender within every age group except the youngest. Thus, in the remainder of this subsection, we consider the results within the context of gender roles within the community.

In the Ende community, women have less power than men. These highly fixed gender roles empower men and assign them the highest status positions in the village, and orators who are women have less power than orators who are men. We argue that the higher rates of stopping observed among the woman orators stem

from the women harnessing the prestige associated with the variants in an effort to access symbolic power. This interpretation is consistent with work on widely spoken languages in the West (Labov, 1966, 1990; Eckert, 1989), which demonstrates that, in linguistic changes from above, women tend to use standard variants more often than men. Eckert (1989, p. 265) argues that the tendency arises in cases when women use prestige variants as a way to access symbolic power when real power (e.g., socioeconomic, political) is not available to them. Given this link between gender and power, we argue that women orators in Ende may produce the stop variant as a way of accessing symbolic power.

That the stop variant, like n-elision (Lindsey, 2021b), is linked with orator status raises questions about the extent to which the stop variant directly indexes being an orator, or whether instead its greater use by orators stems from its indexation with other social meanings, meanings that orators frequently index. In the next section, we explore the extent to which there is a direct index between the stopped variant of the retroflex obstruents and the researcher-observed category of being a practitioner of *kawa*. To do this, we discuss the proportion of the stopped variant in the speech of select individuals within the context of their social characteristics and position within the community.

5.2 Individual patterns

In order to gain a better understanding of the ways in which individual speakers' realizations pattern in regard to who they are, in this section, we consider the variants produced by select individuals alongside descriptions of their lifestyles, occupations, and community roles.

We begin by focusing on Jubli Sowati (i.e., JSS in Figure 1), a "big man" in the village who works as a teacher outside of Limol. Jubli is of particular interest because, despite being in the youngest age group and not being observed doing *kawa*, Jubli works in a prestigious position and comes from an influential family. None of the other speakers in the sample have a social profile that is similar to Jubli's; among the non-orators, there is no one other than Jubli whose job holds a similar level of prestige. This offers the opportunity to see whether the tendency for orators to use the stopped variant of the retroflex obstruents directly indexes orator status or if it is instead imbued with meanings associated with prestige more generally. If the stopped variant is directly indexed with orator status, we would expect that a small proportion of Jubli's tokens would be realized as a stop since he is not an orator. If, on the other hand, the stopped variant is indirectly indexed to orator status via an association with power, prestige, or other stylistic choices frequently made by orators, then we might expect Jubli to produce a large number of tokens as the stop variant. As shown in Figure 3, Jubli produces

a high proportion of tokens as the stop variant, much higher than others in his age cohort, including those who practice *kawa*. In fact, there are only three speakers for whom analysis revealed a similarly high proportion of stops, two of whom are orators in the oldest age group. Thus, Jubli's distinct social profile is consistent with the especially high number of stopped tokens in his speech. This suggests that the stop variant is linked with power and prestige more generally, and it is accordingly used by orators.

Other non-orators who produce a large number of stops are Kalamato Joannang (KJJ), Gloria Warama (Kurupel) (GWK), and Winson Warama (Kurupel) (WWK). We did not anticipate that these three speakers would behave differently than the other non-orators. We present their profiles here so that they may inform our interpretation of the stopped variant. Kalamato is the treasurer of the Ende Language Committee, and she comes from a prestigious family. Her father, Biku Madura, is the oldest male member of the *Limollang* clan, the clan whose origin place is in Limol. In addition, her husband, Geoff Rowak, holds a very important role in the community, the community chairman, who is tasked with managing the community and helping everyone to live properly (Rowak, 2018, p. 89). These connections may afford Kalamato the sort of position where she might feel inclined to use prestigious forms. The other two speakers may also produce a relatively large number of stops as a result of their family connections. Gloria and Winson are siblings, and they are the children of Warama Kurupel (Suwede) and Wagiba Geser, both of whom are orators and in positions of power in the village. While neither Gloria or Winson use the stopped variant as much as Jubli or Kalamato, they use higher rates than the other non-orators and similarly high rates to one another. Taken together, the high use of the stopped variant by Jubli, Kalamato, Gloria, and Winson may suggest that the variant is indexed with social meanings associated with power and prestige rather than directly indexed to being an orator.

This interpretation, however, is complicated by the production patterns observed in the speech of the late Rex Kurupel (Suwede) (RKS). Rex was Warama Kurupel (Suwede)'s youngest brother, so not only is he related to Gloria and Winson but he is the child of Kurupel (Suwede), who was the founder of the current location of Limol village, and thus was a very "big man" in the village. While such family connections may lead us to expect high rates of the stop variant, Rex produced the lowest proportion of stops of all of the speakers analyzed. Indeed, of his 100 tokens, only 3 were stops. Rex was well respected, and was also a quiet man who worked hard and kept to himself, using his time to build fences or hunt. His brother, Warama, and his nephew, Winson, are similar in nature.

Thus, while the patterns of stop production may suggest that the stopped variant is linked with being an orator via its connection with power and prestige,

it is not the case that all Ende speakers who have power frequently use this variant. A more nuanced understanding of the relationship between phonetic realization and social meaning could be gained in future work that uses discourse analysis to examine how the linguistic forms may be employed for stance taking. In addition, since orator status is based on ethnographic observation, it is possible that some of the individuals coded as non-orators for this analysis have practiced *kawa* when Lindsey is not present. Likewise, orator status is not a stagnant category. Confirmation from community members regarding whether certain individuals practice kawa and additional ethnographic observations in the future may help better understand the relationship between linguistic variation and power in Ende.

6. Conclusion

In this study, we have demonstrated that retroflex obstruents in Ende are realized variably as stops and affricates and that this variation is linked with voicing, speaker age, and whether the speaker is someone who practices *kawa*. The observed effect of age among the orators could easily have gone unnoticed if orator status was not previously identified as a relevant local social factor, highlighting the importance of including locally relevant social factors in quantitative variationist analysis. Our ability to explore the relationships between *kawa* oration and linguistic behavior in this paper has been reliant on the diligent collection of a range of ethnographic metadata related to the community-specific cultural context, and that these materials have been made accessible. This work continues a long tradition of incorporating ethnographic fieldwork into sociolinguistic analysis for uncovering locally relevant social categories that inform our analyses of social ideologies and linguistic practice.

Further, the current study suggests that the social meanings associated with the stopped variant are those having to do with power and prestige, rather than there being a direct link between the variant and an individual's status as a community orator. To understand this, it was vital that we examine variation in the speech of select individuals whose social profiles vastly differed from other individuals whose speech was analyzed. Doing so allowed us to better understand the social meanings underlying the variation we observed, and it crucially relied on the careful observations made by the second author during fieldwork.

Future work that draws on ethnographic documentation to explore broader patterns of variation in understudied languages will enhance, and likely challenge, our understanding of dominant sociolinguistic trends.

Funding

Data collection for this paper would not have been possible without generous funding from the Firebird Foundation and the American Philosophical Society.

Acknowledgements

An earlier version of this paper was published in 2020 as Article 15 in the *University of Pennsylvania Working Papers in Linguistics: Selected papers from NWAV48* 26(2). The updated analysis presented here includes the 16 speakers analyzed for the 2020 paper with the addition of 14 more speakers whose speech was not previously analyzed. We wish to thank all of the Ende speakers who agreed to be interviewed for this study: Maki Bewag, Jerry (Jeks) Dareda, Kenny Dareda, Mecklyn (Takeya) Dieb, Kaoga Dobola, Wendy Frank, Kwale Geser, Wagiba Geser, Kalamato Joannang, Biku Kangge, Andrew Kaoga (Dobola), Jugu Karea, Namaya Karea, Merol Koe†, Donae Kurupel, Paine Kurupel, Rex Kurupel†, Sarbi Kurupel, Warama Kurupel (Suwede), Dugal (Dukes) Mado, Karea Mado, Warani Pewe, Erabal Sali, Christina Soma, Jubli Sowati, Maryanne Sowati (Kurupel), Gloria Warama, Tonny (Tonzah) Warama, Winson Warama, and Bibiae Zakae. We would like to especially thank Warama, Wagiba, and Tonny, who hosted the project and assisted with transcribing the interviews, and Jack Dipa, who translated the abstract into Ende. Data collection was supported by the Firebird Foundation, the American Philosophical Society, and the Stanford Linguistics Department. We are also grateful for the feedback given to us by Amy Schafer, Andrea Berez-Kroeker, Gary Holton, Alexander Mawyer, the audience at the NWAV 48 session "What's so standard about standards?", and the University of Hawai'i at Mānoa Linguistics Department Sociolinguistics Reading Group.

References

Barr, Dale J., Levy, R., Scheepers, C., & Tily, Harry J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barth, Daniel, Schokkin, Dineke, Travis, Catherine, Lindsey, Kate L., & Stanford, James (2019). *Variation off the Beaten Path: Expanding our understanding of social structures*. Workshop at NWAV 48, University of Oregon.

Bates, Douglas, Maechler, Martin, Bolker, Benjamin, & Walker, Steven C. (2019). Package 'lme4'. R package version 1.1–21. https://cran.r-project.org/web/packages/lme4/lme4.pdf

Bewag, Maki (Lynette) (2018). Sociolinguistic Questionnaire – Maki (Lynette) Bewag. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI095

Clarke, Sandra (2009). Sociolinguistic stratification and new dialect formation in a Canadian aboriginal community: Not so different after all?. In James N. Stanford & D.R. Preston (Eds.), *Variation in indigenous minority languages* (pp. 109–128). Amsterdam: John Benjamins. https://doi.org/10.1075/impact.25.06cla

Clements, George. N. (1999). Affricates as noncontoured stops. In Osamu Fujimura, Brian D. Joseph, & Bohumil Palek (Eds.), *Proceedings of LP '98: Item order in language and speech* (pp. 271–299). Prague: Charles University Press.

Dareda, Jerry (Jeks) (2018). Sociolinguistic Questionnaire – Jerry Dareda. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI059

Dareda, Kenny (2018). Sociolinguistic Questionnaire – Kenny Dareda. In Lindsey, Kate L. (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI112

Dieb, Mecklyn (Takeya) (2018). Sociolinguistic Questionnaire – Mecklyn (Takeya) Dieb. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI087

Dobola, Kaoga (2018). Sociolinguistic Questionnaire – Kaoga Dobola. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI110

Echeverría, Max S., & Contreras, H. (1965). Araucanian Phonemics. *International Journal of American Linguistics*, *31*, 132–135. https://doi.org/10.1086/464825

Eckert, Penelope (1989). The whole woman: Sex and gender differences in variation. *Language Variation and Change*, *1*, 245–267. https://doi.org/10.1017/S095439450000017X

Eckert, Penelope (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology*, *41*, 87–100. https://doi.org/10.1146/annurev-anthro-092611-145828

Eckert, Penelope (2000). *Linguistic variation as social practice*. Wiley-Blackwell.

Ellison, T. Mark., Evans, Nicholas, Kashima, Eri, Schokkin, Dineke, & Williams, D. (2016). Do shared norms cross language boundaries? A case study from multilingual Southern New Guinea. *Paper presented at NWAV Asia-Pacific 4, National Chung Cheng University*.

Flemming, Edward (2003). The relationship between coronal place and vowel backness. *Phonology*, *20*(3), 335–373. https://doi.org/10.1017/S0952675704000041

Frank, Wendy (2018). Sociolinguistic questionnaire – Wendy Frank. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI058

Geser, Kwale (Tutu) (2018). Sociolinguistic questionnaire – Kwale Geser. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI118

Geser, Wagiba (2018). Sociolinguistic questionnaire – Wagiba Geser. In Kate L. Lindsey (Ed.), *The Language Corpus of Ende and Other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI047

Guion, Susan G. (1998). The role of perception in the sound change of velar palatalization. *Phonetica*, *55*, 18–52. https://doi.org/10.1159/000028423

Hall, T.A., & Hamann, Silke (2006). Towards a typology of stop assibilation. *Linguistics*, *44*, 1195–1236. https://doi.org/10.1515/LING.2006.039

Hall, T.A., Hamann, Silke, & Zygis, Marzena (2006). The phonetic motivation for phonological stop assibilation. *Journal of the International Phonetic Association*, *36*(1), 59–81. https://doi.org/10.1017/S0025100306002453

Hay, Jennifer, & Drager, Katie (2007). Sociophonetics. *Annual Review of Anthropology*, *36*, 89–103. https://doi.org/10.1146/annurev.anthro.34.081804.120633

Hildebrandt, Kristine A., Jany, Carmen, & Silva, Wilson (Eds.) (2017). *Documenting Variation in Endangered Languages* [Special Issue 13]. http://hdl.handle.net/24754

Joanang, Kalamato (2018). Sociolinguistic questionnaire – Kalamato Joanang. In Kate L. Lindsey (Ed.), *The Language Corpus of Ende and Other Pahoturi River Languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI092

Kangge, Biku (Madura) (2018). Sociolinguistic questionnaire – Biku (Madura) Kangge. In Lindsey, Kate L. (Ed.), *The language corpus of Ende and Other Pahoturi River Languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI091

Kaoga (Dobola), A. (2018). Sociolinguistic questionnaire – Andrew Kaoga (Dobola). In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI109

Karea, Jugu (Mado) (2018). Sociolinguistic questionnaire – Jugu (Mado) Karea. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI089

Karea, Namaya (2018). Sociolinguistic questionnaire – Namaya Karea. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI097

Kashima, Eri, Lindsey, Kate L., Schokkin, Dineke, & Strong, Katherine A. (2019). *Variationist studies in southern New Guinea*. Workshop presented to Current Trends in Papuan Linguistics, Langues et civilisations à tradition orale.

Kim, Hyunsoon (2001). A phonetically based account of phonological stop assibilation. *Phonology*, *18*(1), 81–108. https://doi.org/10.1017/S095267570100402X

King, Sharese (2018). Exploring social and linguistic diversity across African Americans from Rochester, New York. Doctoral dissertation, Stanford University. http://purl.stanford.edu/gj305vh0690

Koe, Merol (2018). Sociolinguistic questionnaire – Merol Koe. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI078

Kurupel, Donae (2018). Sociolinguistic questionnaire – Donae Kurupel. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI051

Kurupel (Suwede), Paine (2018). Sociolinguistic questionnaire – Paine Kurupel. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI053

Kurupel, Rex (2018). Sociolinguistic questionnaire – Rex Kurupel. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI100

Kurupel, Sarbi (2018). Sociolinguistic questionnaire – Sarbi Kurupel. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI114

Kurupel (Suwede), Warama (2018). Sociolinguistic questionnaire – Warama Kurupel. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI052

Labov, William ([1966] 2006). *The social stratification of English in New York City* (2nd edition). Cambridge: Cambridge University Press.

Labov, William (1969). Contraction, deletion, and inherent variability of the English copula. *Language*, *45*, 715–62. https://doi.org/10.2307/412333

Labov, William (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
Labov, William (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation and Change*, *2*, 205–254. https://doi.org/10.1017/S0954394500000338
Lindsey, Kate L. (2015). *Language corpus of Ende and other Pahoturi River languages (LSNG08)*. Digital collection managed by PARADISEC. [Open Access]. http://catalog.paradisec.org.au/collections/LSNG08
Lindsey, Kate L. (2019). Ghost elements in Ende Phonology. Doctoral dissertation, Stanford University. https://purl.stanford.edu/ys194fp6634
Lindsey, Kate L. (2021a). Ende (Illustration of the International Phonetic Alphabet). *Journal of the International Phonetic Association*, 1–21. https://doi.org/10.1017/S0025100320000389
Lindsey, Kate L. (2021b). Variable /n/-realisation in Ende. In E. Kashima & M. Meyerhoff (Eds.), *Asia-Pacific Language Variation*, *7*(1), 30–61. https://doi.org/10.1075/aplv.19013.lin
Mado, Dugal (2018). Sociolinguistic questionnaire – Dugal Mado. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI075
Mado, Karea (2018). Sociolinguistic questionnaire – Karea Mado. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI106
Mansfield, John, & Stanford, James (2017). Documenting sociolinguistic variation in lesser-studied indigenous communities: Challenges and practical solutions. In Kristane A. Hildebrandt, Carmen Jany, & Wilson Silva (Eds.), *Language Documentation & Conservation: Documenting Variation in Endangered Languages* [Special Publication No. 13], 116–136.
Meyerhoff, Miriam (2015). Turning variation on its head: Analyzing subject prefixes in Nkep (Vanuatu) for language documentation. *Asia-Pacific Language Variation*, *1*, 79–109. https://doi.org/10.1075/aplv.1.1.04mey
Meyerhoff, Miriam (2017). Writing a linguistic symphony: Analyzing variation while doing language documentation. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, *62*, 525–549. https://doi.org/10.1017/cnj.2017.28
Milroy, Lesley (1980). *Language and social networks*. Oxford: Blackwell.
Nagy, Nagy, & Meyerhoff, Meyerhoff (2008). *Social lives in language: Sociolinguistics and multilingual speech communities*. John Benjamins: Amsterdam.https://doi.org/10.1075/impact.24.02nag
Nagy, Nagy (2017). Documenting variation in (endangered) heritage languages: How and why? *Language Documentation and Conservation*, *13*, 33–64.
Ohala, John J. (1983). The origin of sound patterns in vocal tract constraints. In P. F. MacNeilage (Ed.), *The Production of Speech* (pp. 189–216). New York: Springer-Verlag. https://doi.org/10.1007/978-1-4613-8202-7_9
Pewe, Warani (2018). Sociolinguistic questionnaire – Warani Pewe. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI073
R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/
Rickford, John (1986). The need for new approaches to class analysis in Sociolinguistics. *Language and Communication*, *6*, 215–221. https://doi.org/10.1016/0271-5309(86)90024-8

Rowak, Geoff (2018). Sociolinguistic questionnaire – Geoff Rowak. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI098

Sali, Erabal (2018). Sociolinguistic questionnaire – Erabal Sali. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI064

Sankoff, Gillian (1980[1976]). *The social life of language*. Philadelphia: The University of Pennsylvania Press.

Schokkin, Dineke (2017). Capturing inherent variability while doing language description: The case of the Idi auxiliary. Paper presented at Association for Linguistic Typology 12, Australian National University.

Schokkin, Dineke (2018). Towards a sociogrammar of Idi: Exploring final nasal deletion in verbs. Paper presented at NWAV Asia-Pacific 5, University of Queensland.

Schokkin, Dineke (2021). Variable realisation of verb-final /n/ in Idi. In Eri Kashima & Miriam Meyerhoff (Eds.), *Asia-Pacific Language Variation*, *7*(1), 1–29. https://doi.org/10.1075/aplv.19011.sch

Skilton, Amalia (2017). Three speakers, four dialects: Documenting variation in an endangered Amazonian language. In K.A. Hildebrandt, C. Jany, & W. Silva (Eds.), *Language Documentation & Conservation: Documenting Variation in Endangered Languages* [Special Publication No. 13], 94–115.

Smakman, Dick, & Heinrich, Patrick (Eds.) (2015). *Globalising Sociolinguistics: Challenging and expanding theory*. New York: Routledge. https://doi.org/10.4324/9781315697826

Soma, C. Christina (2018). Sociolinguistic questionnaire – Christina Soma. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI046

Sowati, Jubli (Joe) (2018). Sociolinguistic questionnaire – Jubli Sowati. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI103

Sowati (Kurupel), Maryanne (2018). Sociolinguistic questionnaire – Maryanne Sowati (Kurupel). In Lindsey, Kate L. (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI107

Stanford, James N. (2009). "Clan as a sociolinguistic variable: Three approaches to Sui clans." *Variation in indigenous minority languages*, 463–484. https://doi.org/10.1075/impact.25.23sta

Strong, Katherine, Lindsey, Kate Lindsay, & Drager, Katie (2020). Gender, oration, and variable affrication in Ende. *University of Pennsylvania Working Papers in Linguistics*, *26*(2).

Suokhrie, Kelhouvinuo (2016). Clans and clanlectal contact: Variation and change in Angami, *Asia-Pacific Language Variation*, *2*, 188–214. https://doi.org/10.1075/aplv.2.2.04suo

Warama (Kurupel), Gloria (2018). Sociolinguistic questionnaire – Gloria Warama (Kurupel). In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI111

Warama, Tonny (Tonzah) (2018). Sociolinguistic questionnaire – Tonny (Tonzah) Warama. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI049

Warama, W. (2018). Sociolinguistic questionnaire – Winson Warama. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI108

Wassink, Alicia Beckford (2015). Sociolinguistic patterns in Seattle English. *Language Variation and Change*, *27*, 31–58. https://doi.org/10.1017/S0954394514000234
Wickham, Hadley (2016). *ggplot 2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org. https://doi.org/10.1007/978-3-319-24277-4
Zakae, Bibiae (2018). Sociolinguistic Questionnaire – Bibiae Zakae. In Kate L. Lindsey (Ed.), *The language corpus of Ende and other Pahoturi River languages*. http://catalog.paradisec.org.au/collections/LSNG08/items/SE_PI050
Zhang, Qing (2005). A Chinese yuppie in Beijing: Phonological variation and the construction of a new professional identity. *Language in Society*, *34*(3), 431–466. https://doi.org/10.1017/S0047404505050153

Abstrakt

Ge peba da Ende ekong llayaba eka kutt eka tän pallall dan. Ngämi mermerae intngemenyamalla 30 Ende Llimoll att llayaba eka panynenatt de, umllang ede ada ngaseka oba eka panynen ngallen a ttongdae dan o ddone. Ede ngämi eka däntamenyaemnalla eka tän komlla dändär ede, ge <tt> a <dd> alle darbnenang a o ttaemang a erag. Ge eka kutt ngallen a *ttatta* wa *ddaddu* eke kutt me dag. Ddob eka panynenang lla da ddone buddobuddog ttaem amallo ge eka kutt de (pällämpälläm eka walle: *stop*) be ddobagabi ge eka kutt eka tän de buddobuddog (llällam peyang) ttaem amallo (pällämpälläm eka walle: *affricate*).

Ngämi umllang gogmam ada eka panynenang lla da ami kawawang dag ttängäm makäp me ede ubi eka kutt eka tän de ddone buddobuddog dänttaemaemneyo. Kawawang llabatt lla da ge eka kutt eka tän de ddone buddobuddog (llällam peyang) dänttaemaemneyo. Eka kutt eka tän <dd> da ddone budddog (llällam peyang) ttaemang dan ge <tt> da alla. Ngämi mit de däbbleya ada llabatt a kawawang a eka kutt eka tän de ddone buddobuddog (llällam peyang) ttaem amallo ge sisor lla da wa kawameny lla da alla, adawatta kawawang a wa llabatt lla da ttängäm me nyamnenma lla dag.

Address for correspondence

Katherine Strong
Department of Linguistics
University of Hawai'i at Mānoa
1890 East-West Road
Moore 569
Honolulu, HI 96822
USA
strongka@hawaii.edu

Co-author information

Kate L. Lindsey
Boston University
klindsey@bu.edu

Katie Drager
Department of Linguistics
University of Hawai'i at Mānoa
kdrager@hawaii.edu

Publication history

Date received: 1 December 2021

Author Queries

- Please provide a complete reference for the citation '(Lindsey, 2019a), (Lindsey (2019b)), (Barr and colleagues (2013))' in this article.
- Please provide a citation for the reference id "CIT0001 (Barr, Dale J., Levy, R, Scheepers, C., & Tily, Harry J. (2013)), CIT0004 (Bewag, Maki (Lynette). (2018)), CIT0007 (Dareda, Jerry (Jeks). (2018)), CIT0008 (Dareda, Kenny. (2018)), CIT0009 (Dieb, Mecklyn (Takeya). (2018)), CIT0010 (Dobola, Kaoga. (2018)), CIT0017 (Frank, Wendy. (2018)), CIT0018 (Geser, Kwale (Tutu). (2018)), CIT0019 (Geser, Wagiba. (2018)), CIT0025 (Joanang, Kalamato. (2018)), CIT0026 (Kangge, Biku (Madura). (2018)), CIT0027 (Kaoga (Dobola), A. (2018)), CIT0028 (Karea, Jugu (Mado). (2018)), CIT0029 (Karea, Namaya. (2018)), CIT0033 (Koe, Merol. (2018)), CIT0034 (Kurupel, Donae. (2018)), CIT0035 (Kurupel (Suwede), Paine. (2018)), CIT0036 (Kurupel, Rex. (2018)), CIT0037 (Kurupel, Sarbi. (2018)), CIT0038 (Kurupel (Suwede), Warama. (2018)), CIT0041 (Labov, William. (1972)), CIT0047 (Mado, Dugal. (2018)), CIT0048 (Mado, Karea. (2018)), CIT0056 (Pewe, Warani. (2018)), CIT0060 (Sali, Erabal. (2018)), CIT0067 (Soma, C. Christina. (2018)), CIT0068 (Sowati, Jubli (Joe). (2018)), CIT0069 (Sowati (Kurupel), Maryanne. (2018)), CIT0073 (Warama (Kurupel), Gloria. (2018)), CIT0074 (Warama, Tonny (Tonzah). (2018)), CIT0075 (Warama, W. (2018)), CIT0078 (Zakae, Bibiae. (2018))" since citation is missing in the article.
- Please provide an acceptance date for this article.
- Please provide keywords for this article.

Standard-ness, national ideologies and their embedding in 'sociolinguistic theory'

Shobha Satyanath
University of Delhi

In spite of the centrality of the vernacular in sociolinguistics, we find, it is constantly pitched against the standard, failing to make meaning on its own. It is common practice in western sociolinguistics to identify variants as standard or non-standard, thus marking a distinction between the two and their social meanings. However, when researching English in non-western contexts, such a distinction may not always be tenable. This study reports on variation in the use of a pre-verbal auxiliary – derived from English periphrastic Do – in the speech of Indo-Guyanese speakers, a community of Indian descent in Guyana. A series of matched guise tests were conducted which suggest confusion in the minds of speakers in unambiguously labeling the variants as English, creole or mixed. Further, auxiliary variants are evaluated neutrally without attaching greater or lesser value or prestige in spite of an on-going change at an advanced stage. This has consequences for the very notions of standard and standard-ness, which continue to dominate sociolinguistic modelling of variation in western contexts.

Keywords: Anglophone creoles, standard, attitudes, national ideology, nonwestern, language ideologies, vernacular, Guyana

1. Overview of theoretical issues

In this paper, I address issues relating to the embedding of western nation-state ideology, standard-ness, and social class in sociolinguistics. The three are interconnected and are responsible for the replication of the kinds of findings that sociolinguistics has come to identify itself with.

To lay bare this relationship the present study draws upon data gathered among speakers of an Anglophone creole spoken in Guyana in 1989 as an example of English in a nonwestern context. I propose that despite Guyana being a country of diasporic communities, and an English speaking one today, it does not draw upon the monocultural ideology associated with white Englishes of the

https://doi.org/10.1075/aplv.22007.sat
Asia-Pacific Language Variation 8:2 (2022), pp. 174–205 | ISSN 2215-1354 | E-ISSN 2215-1362

western nations. To be specific 'English' does not constitute a resource for nation building or social cohesion in Guyana. This is despite English being its national/official language. Instead, it is the multiculturalism of their home countries and 'Creole', the new subaltern identity that serves to shape local language ideologies. This ideology is embedded more in a '*We* versus *Them*' model along a local horizontal dimension, and it challenges the hierarchical relationship between English and Creole, which has been long assumed by linguists to account for variation in such Anglophone communities.

Despite the fact that a great majority of the Guyanese have had no direct associations with their ancestral homelands, the precolonial multiculturalism lives on in their collective memory, in daily cultural practices such as cuisine, beliefs, festivals, music, rituals, kinship, marriage practices, Hindu churches, village lives (where they still see English as "White man's language"), non-English lexical insertions in speech, narratives about the past, including the *jahaaj* 'ship' travel narratives etc. Schlesinger's observations on the practice of village exogamy in marriages in rural Indo-Guyana (1968)[1] is a reminder of a much deeper association with the past (also Sharma, 1986; Roopnarine & Brown, 1997; Edwards, 1978; Despres, 1967).

The present study further argues against the positioning of Anglophone creoles as a continuum of learner varieties with a high degree of substratal influence. It sees this as an extension of European national ideology(ies), and the influences these exert globally on local scholars and laymen alike. Such positions reflect racial stances (a mocktail of class, color and race) towards 'ethnic' varieties of Englishes, be they creole, Afro-American, or non-white, more generally.

The racial exclusivity of western sociolinguistics is well known. One wonders whether the existing findings (more generally) would not have been different had the studies been more inclusive considering the multi-ethnic character of western cities (Holliday, 2019; King, 2021; but see occasional studies such as Thomas, 2001 cited in Labov, 2016, & Baranowski, 2017).[2] This racial exclusivity is one of the major reasons as to why sociolinguistics faces challenges when it is extended to

1. 'Village exogamy is preferred by the Indo-Caribbeans, and girls are usually given in marriage to boys who live in other villages as it prevents marrying among close kin, increases choice of suitors, and strengthens kinship ties between villages. On rare occasions the family seeks for potential spouses from their mother country, India, as well, but given the distance this is usually rare' (Schlesinger, 1968, p. 10).

2. It is possible to argue that this was perhaps overtly guided by the search for native, local speakers (Labov, 1972), but seems hard to justify when studying Afro-American Englishes separately. However, it is difficult to overlook the connection between race and nativity. The US census also lists residents by race. This facilitates the placing of all those of European descent into a single "white" category.

more complex models of speech communities where diversity is the norm, where additional factors other than class and race are at play, and which provide an opportunity for a more inclusive research sampling (see, e.g., Suokhrie, 2016 on a multi-clan Kohima village, Satyanath, 2018 on a multi-ethnic Kohima town, and Kumar, 2021 on a multi-caste village).

Despite the centrality of the vernacular in sociolinguistics and despite sociolinguistics being a progressive and more democratic discipline (see Labov, 1982), western sociolinguistics has been an uncanny victim of its own nationalistic ideology, and has been oblivious to the ideological roots of linguistics and anthropology in the nation and language project of the 19th century (see Heller & McElhinny, 2017; Hutton, 2020; Satyanath, 2021). Western nation-state ideology further tends to devalue the vernacular and accord higher status to standard linguistic forms by embedding them within a socio-economic hierarchy, thus producing both social difference and social inequality at the same time. Standard forms are often assumed to be normative, historically unchanged and unmarked which is clearly problematic as it blurs the diachronic dimension, and may cause other analytical problems. This is even more conspicuous in studies of Anglophone creoles where present-day western standard Englishes are routinely used as norms against which negative value can be assigned to variation in creole structures and their evolution.[3] Therefore, in such a speech community one needs to exercise caution when determining underlying forms. However, this was perhaps not intended, at least initially in sociolinguistics: 'It [the vernacular] is of particular interest to linguists as the form that gives us the most direct connection with the past and current history of the language' (Labov, 2016, p. 585).

Whether or not intended, the non-consideration of the ideological underpinning of western standards and their conscious or unconscious role in sociolinguistic theory and methodology is noteworthy. It has hurt the intellectual growth of the discipline by further widening the divide between northern and southern scholarship (see, for instance, Smakman, 2015, Nagy & Meyerhoff, 2008).[4] It also unnecessarily complicates the understanding of non-western societies and contributes to their exoticization.

---

3. Widespread use of the term 'English lexified creole' is an example of the harm that is being done to Anglophone creoles. It is a highly loaded term implying 'Not English' in grammar and 'English' in lexicon, that is, a Grammar-Lexicon split: a notion often referred to in literature on bilingual mixed languages and implicit in borrowability scales (e.g., Baker, 2013).

4. It is not just the imbalances in the publication rate, and marginalization of southern scholarship in various ways, the terms North and South are now being increasingly used in in the titles of Sociolinguistics conferences and publications. Such a divide has always been implicit in publications on world Englishes and on Anglophone creoles.

1.1 What does standard-ness tell us about the nature of society and the relationship between variants and their users?

Western sociolinguistic experiences reveal that standard variants are typically those associated with higher social classes, education and written forms, which in turn, among other factors, contribute to their overt prestige. Even though the concepts of social meaning and style have expanded over the years, their classical underpinnings remain strong primarily because of the perceived salience of the opposition – whether binary or scalar – between standard and vernacular (Eckert &Labov, 2017) and also their associations with social status, gender and style:

> Linguistic variation has consistently been found to have social meaning in its association with the **status** [*emphasis mine*] and stance of speakers in the context of interaction. This indexical function of variation can contribute to the advancement of ongoing linguistic change. (Eckert & Labov 2017, p. 467)

If such relationships were to be so straightforward and an outcome of a general theory of sociolinguistics, we would expect such a relationship to be far more general in nature and more widely replicable beyond the western world (more discussion in Section 2.2). Yet evidence to be found beyond the West may be thin. On the other hand, paradoxically but predictably, languages constantly keep moving away from the standards or the existing norms by innovating and bringing in newer variants that typically lack any social awareness in the minds of speakers, at least initially (Labov, 1994, 2003). This is a rather general principle which has been widely replicated in studies of ongoing changes worldwide. However, there is also a possibility that the innovative variants even in advanced stages of change may not become the object of overt commentary and discussion. This has been shown in studies in Indian settings (e.g., Bhattacharya, 2017; Dey, 2010; Suokhrie, 2016).

Standard-ness has been a useful heuristic in exploring and identifying vernacular speech in everyday interactions; it acts as a useful guiding principle in separating everyday casual speech from the more normative forms of speech, however those are defined. Nonetheless, the guiding principles which reveal them and the resultant forms may not necessarily be those associated with status, class and normativity in the same sense they have been in the western contexts. A standard may perform different functions in different nations; it may be confined to certain domains which could be as limiting as 'writing' (as is the case with Dravidian languages) or it may be conceptualized as 'lingua franca' (e.g., Western Bengali). These relationships only get further complicated where more than one language forms part of everyday verbal repertoire (as is the case with India, more generally). There may also be other ideologies with different kinds of indexicality at

work (including 'Localness' as seen in Labov, 1963). More importantly, the idea of a hegemonic nation may or may not be a crucial ingredient of those ideologies. Likewise, 'status' is not a monolithic category and it may not invoke the same meaning everywhere. For instance, reverence may not invoke style imitation or act as a social motivation for change. Finally, 'standard' is not an objective concept either (Foley, 1997; Milroy, 2001; Smakman, 2012). The notion of standard is a problem when it is taken too literally to imply earlier, unmarked or even normative forms. The phenomenal growth of English education in Delhi has gone hand in hand with the strengthening of local languages and their increased 'vernacularisation' across various domains of use. The relationship between English and other languages cannot be understood in terms of High versus Low; rather English has become one among the various regional languages present in Delhi. Association of English with education does not translate into its higher value indiscriminately. Whether it is newspapers and popular magazines readership, choice of entertainment, news on television and other electronic media, English comes second to local languages (Satyanath & Sharma, 2016). In this respect, the changing sociolinguistic position of English is reminiscent of a historically well-trodden path. One finds the presence of multilingual inscriptions in India despite Sanskrit emerging as the preferred language of inscriptions (Satyanath, T., 2018).

1.2 Standard-ness, social meaning and advancement of language change

One can see the embedding of standard-ness and social class in theory of language change. In changes from below, incoming (that is, usually non-standard) variants lack any social awareness initially and may acquire newer positive social meanings eventually (if at all), with the new norm replacing the more conservative one(s). It is less clear, however, whether the conservative forms in ongoing changes are ever perceived negatively by speakers. Moreover, subjective reaction tests are seldom administered to an age graded sample so we do not know if different age groups hold similar or different attitudes towards the on-going changes (however, see Sharma, Levon, & Ye, 2022). In many on-going vowel shifts in Philadelphia and in other North American cities, a majority carry no overt social meaning. On the other hand, the raising and tensing of vowels generally have been shown to have negative connotations (Labov, 1972, 2003; Labov, Rosenfelder, & Fruehwald, 2013). If negative connotations are strong enough, these can even reverse a change (Preston, 1996). Likewise, a positive meaning can contribute to the advancement of linguistic change (Eckert & Labov, 2017). Thus change from below, which accounts for the bulk of natural changes, acts as a force that constantly drives a language away from the standard or existing norms through innovation. Stigma or prestige hardly play a role here as long as the innovating forms continue to oper-

ate below the level of consciousness. Only in the advanced stages might the newer forms acquire social meaning as speakers become more aware of the new forms, and this then may contribute to the advancement of change. However, these may not be the necessary outcomes everywhere. In the Indian settings we typically find an absence of social evaluation of the new and older variants as positive or negative even at advanced stages and therefore no evidence of their role in advancement or curtailment of change (e.g., Dey, 2010, Sharma, 2017).

The display of overt positive attitudes for standard variants is then most clearly evident among stable variables (e.g., ing, th/dh, negative concord, t-voicing/flapping) and in changes from above as in (r) in New York city (see, e.g., Labov, 1972, 2018). In changes from above, the newly introduced variants carry prestige and speakers hear themselves using the prestige variants (as revealed through matched guised tasks) even though their speech production may not evidence this (Labov, 1972; Trudgill, 1974).

It follows then that the clearest instances of social meaning and standard-ness or social status are those confined to instances of stable variation and changes from above. The apparent mismatch between attitudes and actual usage is often explained by postulating two different kinds of prestige: overt prestige as reflected in positive attitudes and style shifts, and covert prestige which produces vernacular (Trudgill, 1974).

Most changes from below are shown to originate among the centrally embedded groups, that is, those central in a class hierarchy (Labov, 2001). However, it is less clear whether such groups necessarily command local prestige and hence work as a catalyst for the advancement or outward spread of ongoing changes, while operating from below the level of awareness. This leads to the expectation that perhaps these form the baseline sociolinguistic findings that are likely to be replicated elsewhere too. But what happens when such patterns fail to be replicated in non-western settings? Should the observations gleaned from western speech communities continue to be taken as prototypes for research elsewhere as is currently the case? Satyanath (2015) reported style shifts in several diverse settings in India. However, these patterns of style variation in Indian settings produce disparate social meanings not easily mapped onto a monolingual and a vertical style continuum. She also did not find explanations of variation in terms of prestige associated with standard variants in instances of stable variation (Bhattacharya, 2017; Sunny, 2013; Suokhrie, 2016). These contrast with western settings where most approaches to style and style-shift still assume that speakers' metalinguistic evaluations include standard language ideology.

What do we make out of this? The above discussion suggests non-convergence between the overt positive ratings of standard variants, stigmatization of the nonstandard variants, and speakers' actual usage. The positive ratings

do not necessarily result in greater use of the standard variants, nor does overt negative evaluation necessarily curtail the use of the nonstandard variants (e.g., see Rodríguez-Ordóñez, 2021 in the case of Basque) except temporarily under the effect of the monitor in western contexts (e.g., see Labov et al., 2011). On the contrary, the nonstandard variants account for the bulk of actual everyday verbal behavior. It is possible to resolve this contradiction without using resorting to notions of covert and overt prestige. In style shifts, the steepest differences in the use of standard and nonstandard variants have been shown to typically occur in reading tasks where the role of the monitor guided by ideology is maximal. The shifts within spontaneous speech are less conspicuous. For both (r) and (th/dh), the shifts between casual and careful speech in Labov (1972, Figure 4.2, 4.1) are in fact small (even if significant). These shifts occur at a maximum rate of 10% for any class, and slightly higher for SEC 6–8 in the case of (r) and SEC 0–1 in the case of (th, dh). Conversely, shifts are between 30% and 70% when compared with wordlists and minimal pairs. The alignment of shifts with socioeconomic class underscores the hierarchical relationship between the standard and local linguistic diversity.

With this theoretical backdrop in mind, I now turn to the case study of variation in the use of habitual auxiliary in the speech of Indo-Guyanese community as an example of English in non-western context.

2. Objectives, data and the research design

2.1 Study overview

The specific goals of the paper are to explore: (a) the usefulness of standard-ness and social hierarchy as heuristics in understanding language variation in (daz) and other habitual variants in Guyanese; (b) possible social meanings associated with these variants and language variation, more generally; and (c) local language ideologies associated with English versus Creole(se) as inferable from style shifting in Guyana as well as metalinguistic commentaries that appear during interviews. The study makes use of following kinds of data: (i) speech data to reflect on the nature of variation in (daz) and the sociolinguistic profile of the ongoing changes in the speech community; (ii) the results of perception tasks designed to uncover various possible social meanings that listeners attribute to the auxiliary variants and their users; (iii) two sociolinguistic questionnaires on the use of different language varieties and the qualities participants attached to them.

The study relies on data gathered in 1989 in Georgetown, the capital of Guyana.[5] The post-colonial society of Guyana is organized along race, class, gender and rural-urban dimensions and perhaps also age and status groups (such as priest or those with ritual and specialized skills). These differences, however, are not sharp. Of these the two most talked about dimensions among speakers are rural-urban and race as reflected in references such as Berbice people, Tong 'town' people, the Indian people, the Black people, the Portuguese, the Chiny 'Chinese', the mixed (also terms for specific racial mixing such as Dagla 'mixed'). However, the association between linguistic variation and these dimensions is not straightforward despite being topics of overt discussion and commentary.

The speech data were drawn from two urban neighborhoods of Georgetown and from one racial group- the Indo-Guyanese. The data were gathered through sociolinguistic interviews in the family settings in the presence of family members.

The perception tests were administered to a racially mixed group of 35 university students (many of whom also work(ed) as school teachers) among whom we would expect the ideologies of standard language to be maximally operative, if relevant. Further, in a significant departure from the notion of 'cultural' (Drummond, 1980) and linguistic 'continuum' (Lepage & Decamp, pp. 135, 116–117; Bickeron, 1975; Rickford, 1987), and without attaching values to the terms themselves, this study adopts the emic categories that speakers subscribe to and use themselves. These include, Creolese, English and Mixed, implying a mix of English and Creole.

The rest of the paper is organized as follows: Section 2 provides a comprehensive discussion of the study objectives, datasets and the overall research design. The results of the perception tasks are discussed in Section 3, and the findings of the perceptions tasks are triangulated with additional evidence from language use, Gutman scaling and metalinguistic commentaries in Section 4.

5. It remains to be seen whether or not the distribution of (daz) in the field site under study has shifted since 1989. It might have reached completion, or it might have spread more evenly across rural areas. However, it is unlikely that the rural connections of those living in the Georgetown area have changed. The overall rural-urban composition of Guyana has remained largely stable. As a matter of fact there is a small decline in its urban population from 30% in the 1980s to 26% at the time of writing. Likewise, the proportion of the ethnic groups living in Georgetown has remained stable. Afro-Guyanese in Georgetown were 51% in the 1980 census whereas now that figure is closer to 53%. There is a decline, however in the percentage of the Indo-Guyanese (from earlier 31% to 20%), which might have been caused by their outward migration to other countries (https://statisticstimes.com/demographics/country/guyana-demographics.php).

2.2 Variables and the sociolinguistic profile of change

The linguistic variable examined is (daz) which is used to express habitual aspect in Guyanese. The variants of (daz) include /daz, az, iz, z/; additional habitual variants include /a/, zero, and verbal-s, but these are not phonologically related to (daz).[6] These forms are illustrated in (1a–g).

The medial vowel in /daz/ and /az/ is realized as [ə/ʌ] and the final segment is either a voiceless or a voiced sibilant but usually a voiced segment.[7] The vowel in /iz/ is realized as [ɪ]. (daz) functions as a pre-verbal habitual auxiliary and is derived from the periphrastic *do*, the use of which in non-emphatic contexts is well documented in the history of early modern English (Tieken-Boon vann Ostade, 1987; Nevalainen & Rissanen, 1985). Guyanese contains many such links with earlier vernacular Englishes as evident in its phonology (e.g., palatalized velars in *kyan* 'cannot', *kyab* 'cab', *gyan* 'gone', *gyarden*, 'garden', and velarization of final alveolars as in *grɔng* 'ground', *tɔng* 'town', open bath vowel as in d*a*ddy, m*a*n etc.). Morphosyntactic retentions in Guyanese include the variable use of oblique forms in the nominative and genitive as well as the (daz) variable discussed here (see Satyanath, 2006 for further details of the morphosyntax of Guyanese).

(1) Examples of (daz) variants:[8]

a. Only now an then *we'z* go Berbice (Neelam)

b. An this one, *daz* go Happy Hearts, half day school, nursery. He'z (the other one) go to Ramakrishna (School) (Rupa)

c. She *iz* know*s* about it (Neelam)

d. The show start*s* eight thirty (Devi)

e. Tell she wha' you *daz* do in school (Shanti)

f. Yes, they *daz* make a stage an spread bag and so. The place, everybody Ø sitdown. Indian people *daz* come. Sometimes some Black people Ø come to see, you know. Anybody who want to come, they Ø allow them to come (Poonam)

g. an sometime when *I*'z work (at) Camp Street, the Supervisor suppose(d) to show the right way, *them iz* jus' grinning the teeth the whole day an' makin' jokes at you. *they*'z play bully an *them* in…don' tell nothin'. (Avinash)

6. An additional variant /uz to/ 'used to' is not discussed here.

7. The quality of the vowel in [daz] is the same as in Rickford (1987, p. 9), though Rickford chooses to represent it as [doz]; Bickerton (1975) represents it as does/doz. This may be due to typesetting constraints and avoidance of the use of IPA symbols.

8. The examples are in broad phonemic transcription as much as possible to ensure intelligibility and to avoid glossing. Ø indicates zeros, that is, the eligible context where (daz) could be expected but fails to show up. Overall, (daz) occurs only in about a quarter of the expected cases.

Variation in (daz) is much more structural than stylistic or social in any direct sense, even though the changes intersect with age, class, and gender. Phonological reduction is the outcome of a number of processes such as d-deletion, vowel contraction, vowel alternation, though more than one underlying form is possible. Each of the processes is variably conditioned by a host of factors such as the subject type, properties of preceding consonants, quality of preceding vowels, verbal categories, clause types and grammatical category. A full discussion of these is beyond the scope of the present paper.[9]

More than 50 hours of speech samples from 44 speakers yielded a total of 1350 tokens of (daz). The overall distribution of the variants under consideration here is shown in Table 1. Note that the non-daz habitual forms such as /a/, verbal-s and zero are not included in the table as those are not phonologically related to (daz). These, however form part of the perception tests (3.3).

Table 1. Overall distribution of (daz)

daz	iz	az	z	Total
469	165	44	672	1350
35%	12%	3%	50%	

The overall distribution of (daz) in Table 1 is misleading as usage of the individual variants vary significantly by age (Figure 1) which is indicative of a change in progress. Coefficients in Figure 1 are derived from Goldvarb 3.03b (Sankoff, Tagliamonte, & Smith, 2005).

Figure 1 shows a continuous decline over time in the use of /daz/ accompanied with rise in the use of /z/ and /iz/. /z/ is acquired early on as it is attested even in the speech of speakers as young as 4. The change is therefore highly advanced. /iz/ is a relatively newer variant that is most advanced in the 16–20 age group. The usage of /az/ is highly infrequent though fairly stable over time. The other relevant information about the change includes the role of social class and gender. It is important to mention that the two changes – /z/ and /iz/ – were being led by middle-class younger speakers and operating from below. This pattern augurs

9. As the derivation of (daz) is not the primary objective of discussion, readers are guided to Satyanath (1991) for full details. It is sufficient to mention that variants of (daz) are not the result of progressive phonological weakening as suggested in Rickford (1980). Instead there exists multiple possibilities for generating and relating multiple variants. Therefore, instead of subjecting the entire data as a single set to multivariate analysis, individual processes that produce different variants of (daz) were individually subjected to multivariate analysis. This was necessary given the various possibilities of deriving and relating the multiple habitual variants and the statistical evidence for rule ordering was also tested.

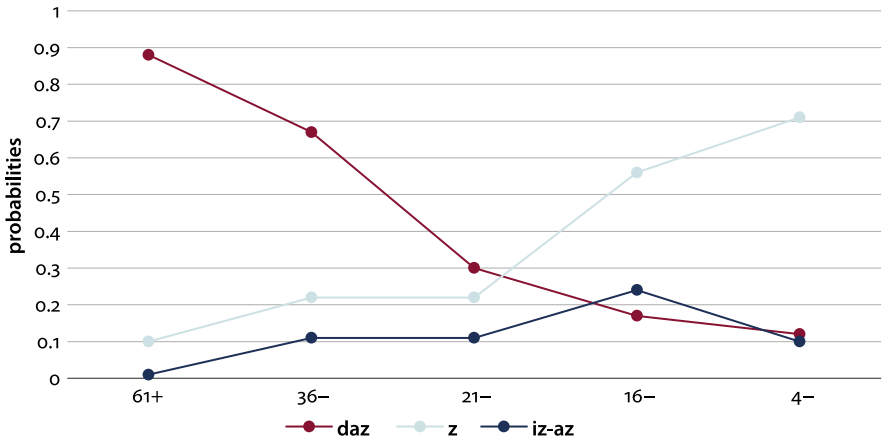


Figure 1. Age coefficients of (daz) [daz and z: $p = 0.003$; iz: $p = 0.004$] in 1989 sample

with the widely observed patterns across western sociolinguistic settings. In contrast, the first change in favour of /z/ is/was being led by young male speakers, whereas the next change in favour of /iz/ shows neutralization of gender-based differences. Leaving out the older speakers who use primarily /daz/ and only peripherally other variants, the four variants are attested in all speakers across the speech community (Satyanath, 1991, 2001).

2.2.1 *Social meaning of (daz) and variants*

At this point it is important to mention that there is no prior objective information available on the possible social meanings of the habitual variants or which forms are considered more or less standard. The only information that is available from the existing literature is cast in the creole-continuum framework which places [a] at the bottom of the continuum (basilectal), [daz] at the midlevel (mesolectal), z, verbal –s, and zeros converge towards the top end of the continuum (acrolectal) (see fn. 7; Section 4). For the remaining reduced variants (/az, iz/), nothing much has been said.[10] Historically, [daz] is the older and the more conservative form but is not present in western Modern Standard Englishes. Going by Standard English, verbal –s with 3rd person singular and zero (where appropriate) would approximate a more standard usage.

10. However, going by the synchronic analysis of various habitual variants, zeros and verbal-s are not the result of decreolization or an indication of acrolectal speech. The two variants are equally present in the speech of categorical [daz] or heavy [daz] users. Likewise verbal -s coexists with [daz] as shown in (1). Much of the existing literature on creoles, including Guyanese creole, has overestimated the value descriptions of various categories under the influence of creole continuum model which essentially views them in terms of relative distance from the modern Standard English without overtly saying so (see Bickeron, 1973, 1975; Rickford, 1979).

Therefore, it was important to consider as many social parameters as possible to test the meanings of (daz) and other habituals such as /a/, zero and verbal -s, which we return to in Section 3.3.

2.3 Perception tests

A variety of subjective reaction tests have been used in variationist studies since Labov (1966). These tests have used modified versions of matched guise techniques developed by Lambert and colleagues (Lambert, Hodgson, Gardner, & Fillenbaum, 1960) and various scales to elicit listener's biases, beliefs and attitudes towards specific linguistic variants. The results are used to infer presence of social meanings that speakers attach to the variants and their users. Labov (1966) provides evidence of positive meanings attached with the use of (r) and negative meaning with stop variants (th/dh) using both style shifts and subjective reaction tests. Outcomes of such tests are taken to reflect community wide perceptions of the variables and their users.

Table 2. Details of the 35 respondents. Two additional Indian female respondents, 45 and 50 years old from Georgetown were included for the task based on Gutman scale (Figure 2)

Age	**Males (12)**	**Females (23)**	**Total (35)**
16–20	3	5	8
22–29	6	5	11
31–35	1	10	11
36–46	2	3	5
Other attributes of the speakers			
Residence	Georgetown (19)	Countryside (16)	
Occupational history	Students (10)	Teachers (19)	Other (6)
Ethnicity	Indo-Guyanese (6)	Afro-Guyanese (24)	Mixed (5)

For the purposes of the present study a total of 13 stimuli representing spontaneous speech were extracted from sociolinguistic interviews with 5 different speakers. Spontaneous speech was used to ensure naturalness and authenticity of the vernacular. Each of the speech stimuli contained the test variants which included the individual variants of (daz): /daz, z, iz/ but also verbal -s, zero and /a/. Non-daz variants were included to assess the extent to which the meanings of (daz) varied in the overall range of habitual forms. Considering the extremely low frequency of /az/ (Table 1), it was not considered. All the /a/ tokens were

extracted from a single speaker.[11] The stimuli were all taken from recordings of Georgetown speakers aged 21–26 years representing three socioeconomic classes. The test sentences were selected in such a way that the same variants were produced by different speakers and the same speaker produced different variants in different sentences. The results showed that different sentences produced by the same speaker were rated differently. This suggests we achieved the desired goal of matched guises and avoided individual speaker biases. Two subjective reaction tests were conducted.

The first test comprised all 13 test stimuli. The listeners were presented with a choice of 10 occupational categories and were asked to assign each speaker (speech stimulus) to one of the ten suitable occupations. As no prior information was available on occupational rankings in Guyana, locally relevant occupations were selected ($n = 10$) bearing in mind the overall levels of education and the local socio-economic structure. The choices exercised by the respondents should allow us to infer the social ranking, if any associated with the habitual variants.

A second test was based on 11 stimuli (reducing one sentence each for z/iz and zero). This time, the listeners were asked to judge each speech excerpt for the following five attributes: language variety- standard English, creole, mix; social class: upper middle class, lower middle class and working class; education level: primary, secondary, university; ethnicity: East Indian- Negro, and residence: town versus country (urban-rural). The stimuli were randomly ordered. They were a multi-racial group comprising Indo-Guyanese, Afro-Guyanese and racially mixed judges (Table 2). However, these personal traits appear to have no effect on the results. For the purposes of computing the results the various stimuli were grouped in terms of the variants it contained (Table 3).

In addition, participants were asked to fill out two questionnaires. In one, participants were asked to report which of the three varieties (creole, mixed, and English) they use, with whom and where (based on Fishman's 1964 domain analysis). On another questionnaire (using modified 5-point Gutman scale), participants were asked to rate the three varieties for five attributes (sweet, musical, intimate, refined, prestigious, and formal).

Given the non-availability of prior objective information of the social ranking of the variants, it was important to test the habitual variants for as many social parameters as possible. Although occupational category would subsume education and socio-economic class parameters, research on non-western societies has shown the utility of testing individual components separately (Sharma, 2017;

11. [a] is more of a rural variant, but it is not entirely absent from urban speech. It can be heard during participant observations- such as when introducing me to a neighbor or an acquaintance. [a] was however not analyzed for variation analysis. The speaker is a heavy [z] user belonging to and is from a middle class family.

Table 3. Perception tests: Number of responses received Stimuli wise

	Total no. of Test stimuli	Break up of no. of test stimuli variant-wise				Total no. of responses
		daz	z, iz	S & zero	A	
Occupational suitability test	13	4	3	4	2	
Responses per stimulus: 33		132	99	132	66	429
Other perception tests*	11	4	2	3	2	
Responses per stimulus: 34		136	68	102	68	374

Note:
Table showing breakup of the test stimuli and the responses received on occupational suitability and other perception tests.
* 'Other' refers to a set of five tasks designed to test attributes such as language variety, social class, education, ethnicity, and demography. The tests were administered to a total of 35 participants. Responses received per test stimulus are less than 35 as one or two participants failed to respond.

Bhattacharya, 2017). Likewise, given the stereotypes based on residence and ethnicity, the same were also tested. Overall, various parameters fail to differentiate among the habitual variants in terms of social ranking or a meaning that could be understood in terms of high versus low or positive versus negative.

3. Research findings

In this section I will discuss the research findings and explore the social meaning(s) associated with the habitual variants. These are divided into three subsections. Section 4 discusses the outcomes of the five perceptions tasks which judged the variants on six parameters: occupation, social class, levels of education, creole-mix-standard, demography and ethnicity. This is furthered with a discussion of language ideologies as revealed through reported language use in different domains with different interlocutors (5.1), metalinguistic commentaries (5.2), and values associated with creole, mixed and English as revealed through Gutman scaling (5.3). Overall, the results are noteworthy as these counter the much hyped linguistic continuum which places variants ranging from creole to English. The continuum tends to assign greater value to the English-like categories. However, these are not matched by the social evaluations assigned to them, as shown in this section.

3.1 Occupational suitability

Listeners were presented with 10 choices representing 10 locally relevant occupational categories. They were asked to assign each speaker (based on test stimuli) to the most suitable of the given occupational categories. Listeners could indicate 'not sure' in the event of uncertainty. Two respondents (out of 35) failed to provide any response so $n = 33$ responses were obtained for each test excerpt resulting in a total of 429 responses (33×13). Figure 2 displays the categories assigned to each of the variants by the respondents. The results were computed in terms of the number of specific responses received for each of the variants and converted into percentages. Generally assignment to a higher ranking profession would imply greater status. The status ordering of the occupations is impressionistic.

The results suggest that the respondents failed to distinguish among the different variants as they could not unambiguously assign them to a given occupational category. The responses for each variant are widely distributed across the 10 choices. A majority of the variants are concentrated in the middle level ranking occupations (circled in Figure 2) followed by the lower or the higher ones. None of the variants received consistently high occupational rankings. Though /z/ overall seems to have lower rankings than /daz/; /daz/ also has its share of lower ranking. Moreover the results fail to shed light on the overall low rankings of more English-like variants. The findings seem to suggest that the variants are used across the social spectrum and do not index one's socio-economic status (However, see Rickford, 1985 who tested speech samples pre-identified as basilectal, mesolectal and acrolectal, discussed in Section 5).

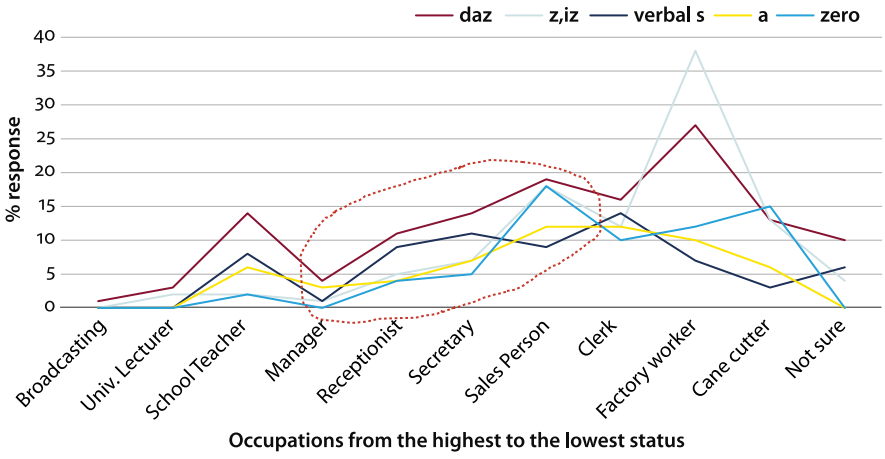


Figure 2. Occupational rankings of variants

3.2 Socio-economic class and education

The results for socio-economic class (Figure 3) and levels of education (Figure 4) also paint a similar picture. It is important to note that a very small percentage of respondents assigned upper middle class (UMC) to any of the variants under consideration and no one assigned upper class (UC) to any one of the variants. Overall, variants received similar ratings suggesting once again that there is no clear status hierarchy and no social meanings that can be assigned unambiguously to any of the variants. In short, there is no evidence that the different variants are socially stratified in any straightforward manner.

As for education, only a fraction of the respondents assigned university education to any of the variants (Figure 3). This may not be surprising considering that despite a high literacy rate in Guyana at that time, very few people had university education. The low assignment of university education to any of the variants seems more a reflection of this reality. Although one of the speakers who contributed the test stimuli had university education at that time, the rankings do not reveal a distinction among those who have higher levels of education versus those who do not. The fieldwork in Guyana suggested that even those with foreign education comfortably used various habitual forms. About half or more than half of the respondents associated high school and primary school education with each of the variants. In contrast, only 7–12% of the respondents attributed university education to /daz/ and verbal –s respectively; 3% to /z, iz/; and no one assigned university education to zero. Despite the apparent asymmetry between university and non-university education the results fail to provide an association between a particular variant and one's level of education. Therefore, the variants do not appear to index level of education either.

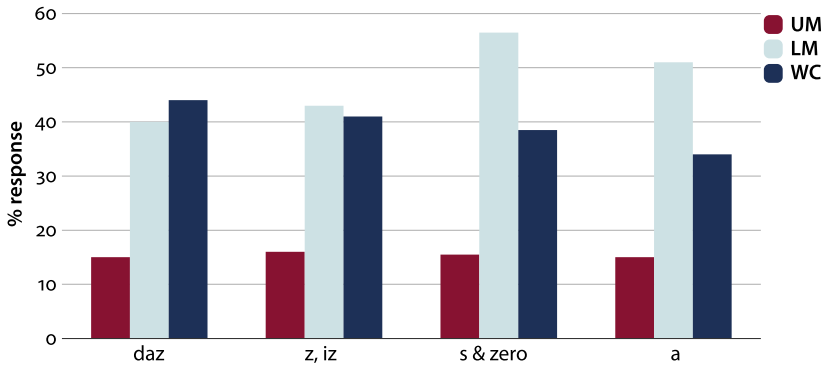


Figure 3. Socioeconmic class ranking of variants

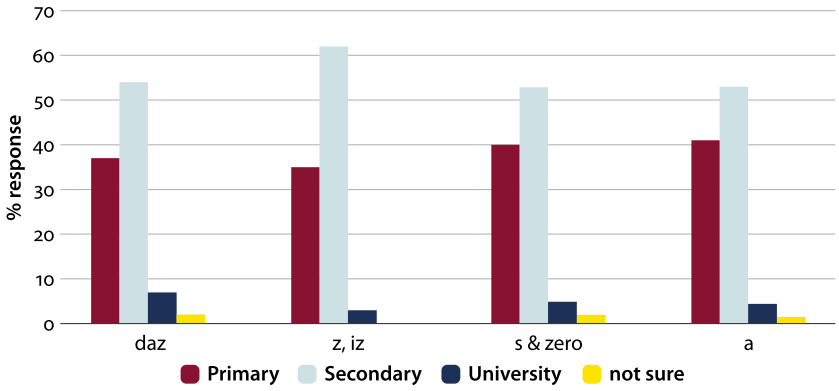


Figure 4. Education levels

3.4 Standard English, mixed and creole

As mentioned earlier, the study relied on the locally used terms such as standard (meaning English or Standard English), mixed and creole that the speakers themselves provided during the overt discussion on language, which generally formed the last part of the structured sociolinguistic interview. On this test, the listeners judged each test excerpt as standard, creole or a mix of creole and standard, thus using a three-way distinction instead of a binary distinction between standard and nonstandard usually used in sociolinguistics. This is the last piece of relevant evidence which could show if any clear social meaning is associated with the different variants, assuming that different variants index different social meanings at all. The findings are presented in Figure 5.

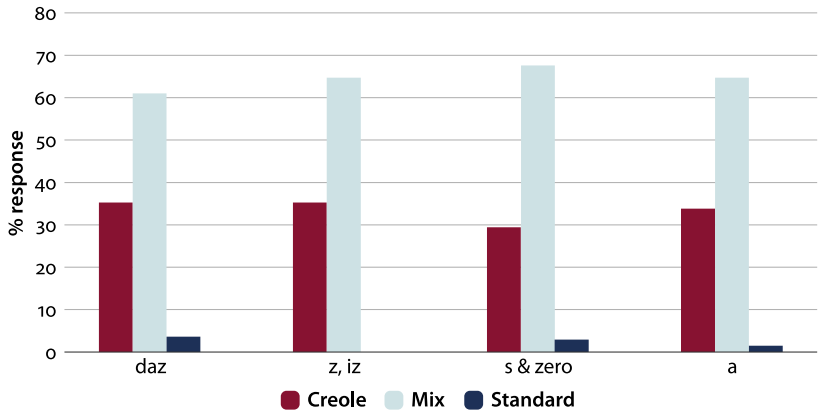


Figure 5. Creole, mixed and standard

The findings suggest that none of the variants under discussion are considered as belonging to the standard. This confirms the local self belief that the local language, Guyanese widely use is Creole or a mix of Standard and Creole. This, in turn, suggests that these carry the maximum functional load in everday interactions and that English or standard English is socially inconsequential. For the purposes of the central theme of the discussion, all the variants belong to both Mixed and Creole. On average, the variants are more mixed than creole providing evidence against the utility of the concept of standardness as well as the creole continuum. The differences among variants in terms of more or less mixed are not significant either. I will return to this point later.

3.5 Demography and ethnicity

Finally, respondents were asked to judge whether the speakers were from urban or rural areas. They were also asked to judge the ethnic background of the speakers as Indo-Guyanese or Afro-Guyanese.[12] The results are presented in Figures 3e and 3f respectively. As shown in Table 2, among the respondents, 24 were Afro-Guyanese, 6 were Indo-Guyanese and five reported themselves as of mixed ethnicity. Likewise, 20 of the 35 respondents were from Georgetown and the remaining reported themselves of rural origin from East Demerara, West Demerara, Berbice[13] and elsewhere. The respondents, though originally from rural areas had spent many years in Georgetown, so were familiar with both urban and rural speech.

As for ethnicity, it is interesting to note that the variants were judged as more Indo-Guyanese than Afro-Guyanese by the respondents regardless of their own ethnicity. There is no correlation between a respondent's own ethnicity and the judgement. The greater uniformity in labelling /a/ as distinctly Indian perhaps confirms the stereotype that /a/ is more rural and that more Indians live in the rural areas. However, considering that the test samples were judged as Afro-Guyanese at all confirms that the racial negative stereotypes the two ethnic groups hold towards each other represent broad social images for in-group consumptions, and language does not index ethnicity/race. The responses are relevant for the question under discussion as more Afro-Guyanese live in the town than Indo-Guyanese. However, results in Figure 6 fail to provide strong evidence of a

12. In Guyana in 1989 when the fieldwork was carried out people used the terms such as 'Negro', 'Black', 'Mixed' and 'Indian'. Indo-Guyanese and Afro-Guyanese are the terms used by the author.

13. and not from the rural areas like Industry which is located adjacent to the Georgetown. This provides a clearer contrast between rural and urban.

link between ethnicity and demography, as the responses were more evenly split over labeling both /a/ and /z, iz/ as rural or urban. This is despite /z, iz/ being involved in ongoing changes, -s and zeros are perceived as more urban but so is /daz/. These findings suggest that the variants are perceived as being used across Guyana in both rural and urban areas irrespective of one's socio-economic status and level of education. This is not surprising considering that a strong rural-urban divide may not exist in Guyana as most families have their roots in the rural areas, where they also have strong family and kinship networks. The relationship between rural and urban Guyana is more bidirectional than unidirectional. In 1989, many people who were studying or working in Guyana either lived in the adjoining rural areas or commuted on a weekly basis to their homes in rural areas. There is no strict racial segregation in Georgetown either.

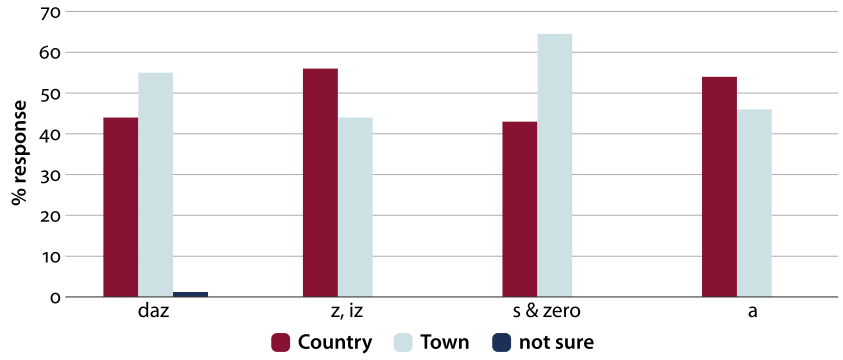


Figure 6. Demography

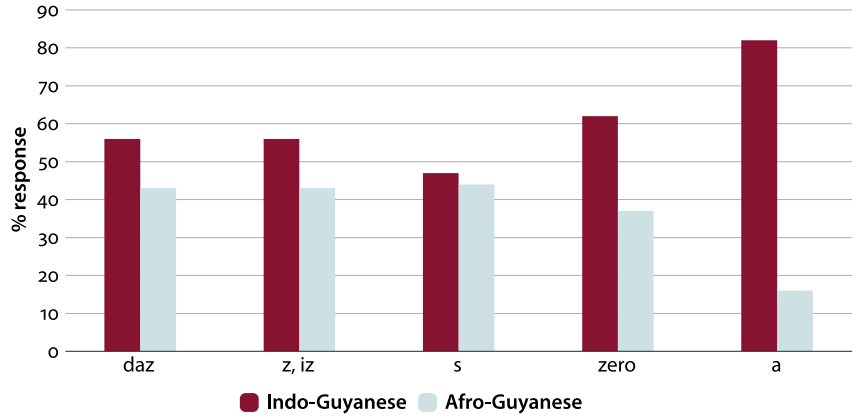


Figure 7. Ethnicity

In summary, the variants under consideration fail to index attributes such as prestige or status that are usually expected to be associated with social class and level of education. The variants, also do not index Creole, Mixed or Standard English suggesting that such indexical relations between linguistic variation and the usual social attributes might not be relevant everywhere. A lot has been written on standard-language ideologies, the class and status nexus of standard and vernaculars, standard versus vernacular cultures (e. g., Koopman, 2016; Milroy 2001), and the social meanings of standards in lesser-studied languages (Lindsay, 2021; King, Maclagan, Harlow, Keegan, & Watson, 2020). However, their historical and social epistemologies that are embedded in the ideology of the modern nation states have not been recognized in sociolinguistics. This realization is central to an understanding of how standard-ness produces meaning and the nature of relationships it holds with the locally attested diversity.

4. Local language ideologies

We have seen that the findings of the various subjective reaction tests in the study fail to reveal any clear associations between the various variants of (daz) and other habitual variants and their possible social meaning(s) in ways that we are used to seeing in studies on English in western contexts. However, an interesting observation that does emerge from the findings discussed above (Section 3) is the broadening of the vernacular: that is, the so-called mixed and creole varieties together refer to two layers (with distinct labels) of one broad variety and not necessarily two or more distinct varieties as much of the early creole literature would like us to believe. The variants are used across the social spectrum and do not correlate with education, occupation etc. The Creole and Mixed, together, represent the vernacular and therefore non-standard in the broader sense, but something that is not devalued and may not even warrant the label 'non-standard'. This is the key finding that emerges from the various subjective reaction tests, which fail to reveal evidence of lower evaluation of the vernacular. The variation in Guyanese, then, can be best understood in terms of a division not between English and Creole, but between standard (English) and vernacular (creole and mix) which is in greater consonance with the variationist studies elsewhere. However, the relationship between the two is not the same as in the western context, as any hierarchical evaluation in terms of prestige ranking remains absent. This constitutes an important difference. I now consider three further pieces of independent evidence (discussed in Sections 4.1–4.3) to show why the findings are different in the case of English in Guyana, and what produces such a difference.

4.1 Language use: Reported and observed

Twenty respondents filled out a questionnaire on what variety (creole, mixed and standard) they would use in different domains with different interlocutors. They included 10 Indo-Guyanese and 10 Afro-Guyanese respondents. The ten Indo-Guyanese also included five who were interviewed earlier. The results are provided in Table 4.

The reported language use is more of an indicator of how language use is governed by the nature of the relationship and the degree of familiarity with the interlocutor and the place, and not necessarily the result of the monitor guided by standard-language ideology. As one moves from home to work domains, the use of Creole and Mixed reportedly decrease and the use of Standard increases. Standard is not the primary choice with close friends, family (mother), neighbors, or with vendors and sellers. Instead, Standard is reportedly used more outside and in the workplace and especially with foreigners. The choice of Standard with foreigners clearly reveals a 'We' versus 'Them' relationship. Though respondents were asked to provide information for each of the three varieties, the findings suggest that Creole and Mixed perform similar functions, and Mixed is used even more widely and is suitable for all domains. The two together represent the colloquial or the everyday vernacular that forms the bulk of language use and carries maximum functional load. My field notes on participant observations, and rapid and anonymous surveys suggest that the vernacular is used across formal and casual domains of usage including in the airport and inside the foreign Embassies irrespective of one's education, class and ethnic background or residence.

I do not attribute Creole and Mixed to lectal variation, nor is there an evidence that Mixed represents a higher level of competence in or is closer to Standard English. Guyanese creole and the mixed variety together define the vernacular, the two together also reveal the presence of an additional layer of vernacular, which emerges only when a speaker is observed in a variety of social settings. That is why sociolinguists are often surprised to find an additional deeper layer of vernacular than what they usually are able to capture during the interviews (see Eckert & Labov, 2017). The need to observe the same speakers in a variety of different social settings is even more relevant in diverse communities. It is not surprising that /a/, though labeled as basilectal in much of the existing literature, was often heard precisely on occasions such as introducing the author to a neighbor or to a vendor in the market (/a/ does not mark social class or competence). It was also evident in the speech of preschoolers (fieldwork observations) who obviously received the input from the family suggesting its continued use though at a lower frequency. Therefore, it is understandable that respondents failed to unequivo-

cally judge /a/ as rural or more Creole than Mixed or as indexing lesser years of schooling.

Space constraints prevent me from indulging in a discussion on style shifting. However, readers are guided to Satyanath (1990) which suggested that in group interactions, a greater shift towards vernacular tended to occur in previous work whenever the conversation was directed at the local participants as opposed to the interviewers, suggesting a further 'We versus Them' dimension.

Table 4. Reported language use. All values are in percentages (based on Satyanath, 1990)

	Layers of vernacular						Monitored speech		
	Creole			Mixed			Standard		
	Home	Outside	Work	Home	Outside	Work	Home	Outside	Work
Father	30	15	15	55	45	30	15	40	55
Mother	42	26	21	58	42	37	0	32	42
Spouse	20	7	0	67	67	33	13	27	67
Younger siblings	28	6	0	61	67	50	11	28	50
Elder siblings	21	0	0	74	63	42	5	37	58
Relatives	26	11	11	58	63	50	16	26	39
Neighbours	40	15	5	55	75	60	5	10	35
Close friends	30	15	10	65	75	45	5	10	45
Friends (opposite sex)	25	15	10	55	60	40	20	25	50
Vendors/ sellers	53	45	53	47	55	42	0	0	5
Bus/taxi	21	10	11	63	70	74	16	20	16
Strangers	5	0	0	32	30	20	63	70	80
Seniors	10	5	5	35	20	5	55	75	90
Equals	20	0	0	75	90	65	5	10	35
Juniors	16	5	5	53	53	35	32	42	60
Foreigners	0	0	0	20	0	0	80	100	100

The discussion so far on English in Guyana demonstrates that there is nothing standard about standards, as the meaning, functions and values that speakers

attach to Standard are not universal. It is more likely that the findings replicated in western sociolinguistics are unique to those linguistic ecologies and ideologies, and are therefore marked.

4.2 Local language ideologies and overt metalinguistic commentaries

I now draw upon the overtly expressed views (by Junior and High School students including one younger school teacher) that provide clear and vivid account of local language ideologies at work. The following commentaries extracted from the interviews further elucidate the relationship between Creole and Mixed vis-à-vis English. Examples in (3a–h) show that deliberate overt display of English devoid of local vernacular features are subject of ridicule, which invites reactions such as:

(3a) "why are you cutting your English on us?." (Gopi)

(3b) "… an there is another gyil, the gyil say, she like to cut up her English on all the children…. we don' tell. we jus talk ordinary talk… we playing in school, we always shout an talk hard." (Gopi)

(3c) "usually girls try to put on (English) and others would tell them not to cut their English on them" (Hemant)

(3d) "There are certain girls in my class. You can't get a word of Creolese from their mouth. She tell me, her father gets angry if she doesn't speak proper English" (Vishal)

Such comments are highly prevalent among the junior and high school students. However, not all stereotypes are aimed at girls (e) and nor are all such usages considered contemptuous (3f refers to the daughter of a Pandit).

(3e) "I notice the ol' man[14] was trying to show his English on you" (Gobind, after I had recorded a conversation with his father in the country side)[15]

(3f) "she is more like Hindus from India and not like Kuli Indians, you know. I often slip back into Creolese when talking to friends but she doesn't understand a word of Creolese" (VK)

VK is justifying the use of English in (3f) as refined by leaning upon personality and other (perhaps tacitly valued) features (e.g. refined, appealing to caste) of in-group membership. The very need to provide a justification underscores that the

14. The use of terms such as 'Old man' to refer to father or an elderly person marks respect meaning 'wise man'. Age reflects accumulated wisdom.

15. Gobind was part of my fieldwork team and assisted me with the fieldwork. He is a University student and lives on the adjoining country side.

non-use of Creolese is exceptional. Likewise, children from rural areas attending school in the city are often the target of ridicule:

(3g) "we make fun of them not because they speak Creolese, but because they try to speak too proper English and this results in making a lot of grammatical errors." (Vishal)

Such children are also made fun of if their casual speech is more conservative than that of an urban child. Vishal further pointed out that 'he and his family speaks 'semi creolese' at home and so do students amongst themselves.

However, the value of English is acknowledged in educational contexts where teachers are expected to use 'proper English' in the classroom, and where failing to do so may invite ridicule:

(3h) "yes I am a teacher and I have to watch what I say in the class, but it doesn't mean that at times I cannot be myself, after all I am a human being" (Nagpal)

The comment in (3h) was made by a school teacher during an informal interaction in the school corridor. His speech consistently included instance of the adjunct 'them' as in *children them, parents them, boys them* when he spoke. These comments suggest that one of the reasons why Creole and Mixed did not appear to be devalued from the subjective reaction tests is that they characterize everyday 'ordinary talk': something of their own (local) as opposed to English which may be seen as contemptuous rather than as an achievement more generally. The results of the Gutman scale provide further evidence in favour of the above metalinguistic commentaries.

4.3 Gutman scales: Qualities associated with linguistic varieties

When using subjective reactions tests, it is often recommended that one also identifies the emotional values (or likeness/approval) that speakers attach to the language variety they use. Accordingly, 37 respondents were asked to rate the three varieties for five qualities – sweet, musical, intimate, refined, prestigious, and formal – each on a scale of 1 (lowest) to 5 (highest). Respondents included two additional Indo-Guyanese (see note in Table 2). The results are displayed in Figure 8.

Figure 8 provides the overall values for each of the five qualities judged. Despite some overlap, one can see that speakers perceive Creole as highly likeable, as they find it musical, intimate and sweet; this is closely followed by the Mixed variety. In contrast, Standard ranks very high on the other three qualities which include refined, prestigious and formal.

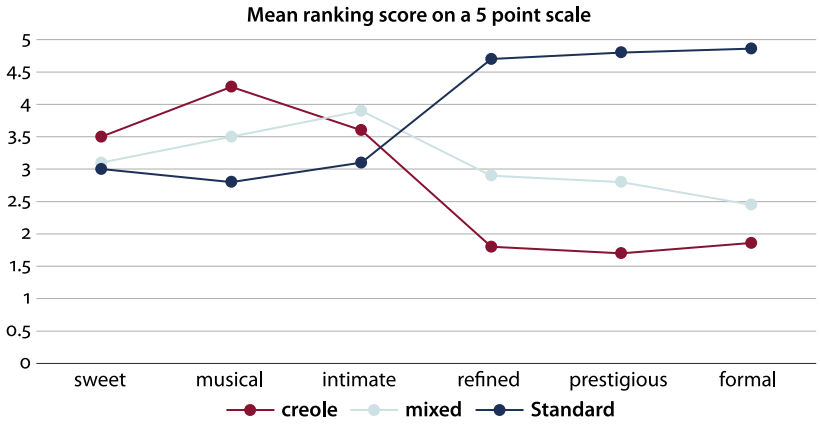


Figure 8. Overall mean values associated with three varieties of English: Creole, mixed and standard

The first three values can be taken to indicate a feeling of warmth; the next three values are more suggestive of distance and formality. This finding is consistent with the reported style shifts and metalinguistic commentaries above and with findings in Table 4 that Standard is the choice with foreigners, outsiders and in specific contexts. Despite respondents ranking Standard highly on prestige and refinedness (Figure 8), this does not translate into other expected social attributes. For example, none of the habitual variants (including /z, iz/ at an advanced stage of change) were judged as Standard. That none of the variants received extreme rankings on any of the parameters suggests the absence of prestige (or, conversely stigma) attached to any of the variants. All in all, the ratings appear to be neutral. The results of the domain analysis and the Gutman scale help illuminate why this might be so.

How do these compare with the existing studies on attitudes in similar creole settings more generally? The results presented in the study auger well with the findings of JMLU (2005). This is one of the largest surveys carried out in Jamaica to test attitudes towards Patwa (the local vernacular) and English by recruiting 1000 respondents from across Jamaica representing three age groups, gender and demography. The tests comprised a number of questionnaires to elicit language use and other attributes of the speakers of Patwa and English. One of the salient findings is that an overwhelming percentage of respondents reported the use of both Patwa and English across age groups, gender, in both rural and urban settings, across occupational backgrounds and from across Jamaica. In fact, use of both is on the rise among the younger speakers suggesting a further increase in positive orientation towards Patwa (JMLU, 2005, Table 4). Further, very little English is used in family settings and English is instead the chosen

code with strangers. The results are overwhelmingly the same regardless of the respondents' demography, age and gender (JMLU 2005, Table 5). Respondents also thought that those who speak English are more educated than those who speak Patwa. A majority of the respondents (71%) thought that an English and Patwa school would be better for their children than an English school alone. (JMLU 2005, Table 16). Overall, the findings suggest an overwhelmingly positive attitudes towards Patwa, and respondents are favourable to declaring Patwa an official language (Table 18).[16] Such results contradict the limited findings of Rickford (1985) for his occupational ranking scale, where higher ranking jobs were assigned to acrolectal speech samples and lower lobs to besilectal samples in the rural area. Apart from the bias in the sample and in the tests stimuli, Rickford himself does not fully endorse the findings as he also finds evidence of positive attitudes. Also as mentioned earlier, the variation in Guyanese speech is much more structural than social or competence based as assumed in earlier studies.

The salient outcomes of the experiments reported here and the major findings of the study can be summed up as follows:

- Overall inconsistency in rating of individual variants across multiple parameters, as these do not converge. Most importantly, the results of occupational ranking scale, social class and education do not coincide for any of the variants as they typically do in western sociolinguistic studies.
- The only convergence that happens across variants is that none of the variants were perceived as belonging to either the highest social class or the lowest social class, suggesting that there is no straightforward relationship between language variation and these parameters, as has been repeatedly observed in the western sociolinguistic studies focused on prestige and social status.
- The responses to a Gutman scale and domain analysis point towards a convergence in values and functions, such that CM (creole, mixed), which together serve as 'our tongue', a language that evokes 'we' feelings, localness, and positive emotional values as melodious, sweet and of intimacy.
- Despite Standard being assigned values such as refined, prestigious, formal, these do not stem from nationalistic ideology as it is operative in the West. Standard may evoke aspirational values for some, but it is also contemptuous, it is foreign, it is colonial and a symbol of an oppressive past. While it is not impossible to imagine similar positive values to be recognized with vernacular in western settings, the difference lies somewhere else at a completely

16. The findings are not comparable with the present study in terms of methods and the focus on the specific variants. Rickford used samples of speech which he termed as basilectal, mesolectal and acrolectal instead of focusing of specific variables. Also see Wassink (1999); Westphal (2015) on Jamaica.

different level, which guides and monitors their (whether creole, mix or standard) use; even condemns overuse of Standard in more familiar contexts as in Guyana. In western contexts the overt prestige of Standard and covert prestige of vernacular show the ambivalence and a conflicting relationship between the two operating at private and public levels.

5. Conclusions

Over the years, sociolinguistic research has consistently produced findings that are not consistent with the realities outside of the West. The findings relate to the relationship between quantitative variation and its social and stylistic stratification, which can be interpreted by invoking notions of standard-ness, the associated prestige that a Standard enjoys, and which in turn, confers certain positive or negative attributes to the linguistic variants. The present study fails to replicate similar findings in the case of English in a non-western context such as Guyana. The reasons are attributed to the absence of a monocultural national ideology, typical of the West that underlies the construction of modern European nation states. Such an ideology subordinates and marginalizes the local diversity in western settings producing the kinds of findings with reference to standard-ness and social meaning. Multicultural ideologies, in contrast produce horizontal relations among languages as well their speech varieties. Whilst terms such as standard and non-standard might be used everywhere, these do not invoke similar attitudes or relationships as seen in more widely studied western societies.

The variants under consideration fail to index attributes such as prestige, status that are usually expected to be associated with social class and level of education. Though a lot has been written on standard language ideologies, class and status nexus of standard and vernaculars, standard and vernacular cultures, the historical and social epistemologies of Standard and standard-ness in the west are embedded in the ideology of modern nation states. This fact has not been recognized in sociolinguistics thus far. This realization is central to the understanding of how standard-ness produces meaning and the nature of relationships it holds with the locally attested diversity. Therefore, standard-ness, the specific social meanings that it indexes for stable variables and variables at advanced stages of change are not universal and can be understood only within the specific context of the rise of the modern European nation states and their national ideologies. The absence of such ideologies fails to replicate western findings. The western findings are a unique outcome of their own unique national ideology and are therefore marked. The overt imbalance in the quantum of the research produced and pub-

lished from multicultural societies and the power imbalances have allowed the west to mark itself as unmarked.

The question is not whether standards usually have higher prestige and non-standards have less of it. Under normal circumstances, it is natural to expect that speakers would have positive attitudes towards vernaculars they speak. The question is under what conditions or circumstances this equilibrium gets disrupted (if it does) and that requires investigation.

Acknowledgements

I am sincerely grateful to Miriam Meyerhoff, Nick Palfreyman, and Jonathan Kaastan for their close reading of the article, challenging comments, useful suggestions all of which have been incorporated. Thank you to Shashi Jain (Gail, India) for providing the Hindi translation of the abstract.

References

Bakker, Peter (2003). Mixed languages as autonomous systems. In Y. Matras & P. Bakker (Eds.), *The mixed language debate: Theoretical and empirical advances* (pp. 107–150). Berlin: Mouton de Gruyter.

Baranowski, Maciey (2017). Class matters: The sociolinguistics of goose and goat in Manchester English. *Language Variation and Change*, *29*(03), 301–339. https://doi.org/10.1017/S0954394517000217

Bhattacharya, Pratibha (2017). Variation and change: A case study of Calcutta Bengali. Unpublished doctoral dissertation. University of Delhi, Delhi.

Bickerton, Derek (1973). On the nature of a creole continuum. *Language*, *49*, 640–69. https://doi.org/10.2307/412355

Bickerton, Derek (1975). *Dynamics o f a Creole System*. London: Cambridge University Press.

DeCamp, David (1960). Four Jamaican creole texts. In Le Page (Ed.), *Jamaican Creole* (pp. 121–180). London: Macmillan.

DeCamp, David (1971). Toward a generative analysis of a post-creole continuum. In Dell Hymes (Ed.), *Pidginization and creolization of languages* (pp. 349–70). Cambridge: Cambridge University Press.

Despres, Leo A. (1967). *Cultural pluralism and national politics in British Guyana*. Chicago: Rand McNally.

Devonish, H. (1978). The selection and codification of a widely understood and publicly useable variety in Guyana to be used as a vehicle of national development. Unpublished D.Phil. thesis, University of York.

Dey, Kakoli (2010). Silchar Bengali: A sociolinguistic study. Unpublished doctoral dissertation. University of Delhi, Delhi.

Drummond, Lee (1980). The cultural continuum: a theory of intersystems. *Man*, *15*(2), 352–74. https://doi.org/10.2307/2801676

Eckert, Penelope, & Rickford, John (Eds). 2001. *Style and Sociolinguistic variation*. Cambridge: University of Cambridge Press.
Eckert, Penelope, & Labov, William (2017). Phonetics, phonology and social meaning. *Journal of Sociolinguistics*, *21*(4), 467–496. https://doi.org/10.1111/josl.12244
Edwards, Walter (1975). Sociolinguistic behavior in rural and urban circumstances in Guyana. Unpublished D.Phil. dissertation, University of York.
Fishman, Joshua. A. (1964). Language maintenance and Language Shift as a field of inquiry: A Linguistics, *2* (9).
Foley, William A. (1997). Standard Languages and linguistic engineering. In William Foley, *Anthropological linguistics: An introduction* (pp. 398–416). Malden, MA: Wiley-Blackwell.
Heller, M., & McElhinny, B. (2017). *Toward a critical history*. Toronto: University of Toronto Press.
Holliday, Nicole R. (2019). Multiracial identity and complexity in sociolinguistic variation. *Language and Linguistics Compass*, *13*(8), https://doi.org/10.1111/lnc3.12345
Hutton, C. M. (2020). Linguistics and the intellectual challenge of diversity. In F. Hofmann & M. Messling (Eds.), *The epoch of universalism, 1769–1989* (pp. 105–124). De Gruyter. https://doi.org/10.1515/9783110691504-201
JMLU (2005). *The language attitude survey of Jamaica*. The Jamaican Language Unit. University of the West Indies, Mona.
King, Sharese (2021). Rethinking race and place: The role of persona in sound change reversal. *Journal of Sociolinguistics. 25*, 159–178. https://doi.org/10.1111/josl.12454
King, Jeanette, Maclagan, Margaret, Harlow, Ray, Keegan, Peter, & Watson, Catherine (2020). Prestige norms and sound change in Māori. *Language Ecology*, *4*(1), 95–114. https://doi.org/10.1075/le.00011.kin
Koopman, Nikolas (2016). Labov, vernacularity and sociolinguistic change. *Journal of Sociolinguistics*, *20*(4), 409–430. https://doi.org/10.1111/josl.12191
Kroch, Anthony, & Small, Cathy (1980). Grammatical ideology and its effect on speech. In David Sankoff (Ed.), *Linguistic variation: Models and methods* (pp. 45–55). New York: Academic Press.
Kumar, Ranjan (2021). *Linguistic Variation and Change in a Caste-Based Speech Community*. NWAV-AP6, Singapore National University (February 17–20, 2021).
Labov, William (2018). The role of the Avant Garde in linguistic diffusion. *Language Variation and Change*, *30*, 1–21. https://doi.org/10.1017/S0954394518000042
Labov, William (2001). *Principles of linguistic change. Volume II: Social Factors*. Cambridge: Blackwell.
Labov, William (1994). *Principles of linguistic change: Internal factors*. Oxford: Basil Blackwell.
Labov, William (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
Labov, William (1966). *The social stratification of English in New York City*. Washington, DC: Centre for Applied Linguistics.
Labov, Williams (1963). The social motivation of a sound change. *Word*, *19*, 273–309. https://doi.org/10.1080/00437956.1963.11659799
Labov, William, Rosenfelder, Ingrid, & Fruehwald, Josef (2013). One hundred years of sound change in Philadelphia: Linear incrementation, reversal, and reanalysis. *Language*, *89*(1), 30–65. https://doi.org/10.1353/lan.2013.0015
Labov, William, Ash Sharon, Ash, Ravindranath, Maya, Weldon, Tracey, Baranowski, Maciej, & Nagy Naomi (2011). Properties of the sociolinguistic monitor. *Journal of Sociolinguistics*, *15*(4), 431–463. https://doi.org/10.1111/j.1467-9841.2011.00504.x

Lambert, W. E., Lambert, R. C., Hodgson, R. C., Gardner, & S. Fillenbaum (1960). Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology*, *20*(1), 44–51. https://doi.org/10.1037/h0044430

LASJ (2005). *The language attitude survey of Jamaica*. The Jamaican Language Unit, Department of Language, Linguistics and Philosophy. Mona: University of the West Indies.

Le Page, Robert B. (1960). An historical introduction to Jamaican creole. In Le Page (Ed.), *Jamaican Creole* (pp. 3–121). London: Macmillan.

Lindsey, Kate (2021). Ende oration and final /n/-realisation. *APLV*, *7*(1), 30–61. https://doi.org/10.1075/aplv.19013.lin

Meyerhoff, Miriam, Abtahian, Maya, Gafter, Roey, Horesh, Uri, Kasstan, Jonathan, Keegan, Peter, & King, Jeanette (2021). Styles, standards and meaning. *Language Ecology*, *4*(1), 1–16. https://doi.org/10.1075/le.00006.mey

Milroy, James (2001). Language ideologies and the consequences of standardization. *Journal of Sociolinguistics*, *5*(4), 530–555. https://doi.org/10.1111/1467-9481.00163

Nagy, Naomi, & Meyerhoff, Miriam (2008). Introduction. In Miriam Meyerhoff and Naomi Nagy (Eds.), *Social lives in language: The Sociolinguistics of multilingual speech communities, celebrating the work of Gillian Sankoff* (pp. 1–16). Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/impact.24.02nag

Nevalainen, Terttu, & Rissanen, Matti (1985). Do you support the Do-support? Emphatic and non-emphatic Do in affirmative statements. In Sven Jacobson, Stockholm (Ed.), *Papers from the third symposium on syntactic variation* (pp. 35–50).

Preston, Dennis (1996). Where the worst English is spoken. In Edgar Schneider (Ed.), *Focus on the USA* (pp. 297–360). Amsterdam: John Benjamins. https://doi.org/10.1075/veaw.g16.16pre

Rickford, John (1987). Decreolization paths for Guyanese singular pronouns. In Glenn G. Gilbert, *99* (Ed.), *Pidgin and creole Languages: Essays in memory of John E. Reinecke* (pp. 130–38). Honolulu: University of Hawaii Press.

Rickford, John (1985). Standard and nonstandard language attitudes in a creole continuum. In Nessa Wolfson & Joan Manes (Eds.), *Language of inequality* (pp. 145–159). https://doi.org/10.1515/9783110857320.145

Rickford, John (1980). How does Doz disappear? In Day Richard (Ed.), *Issues in English creoles: Papers from the Hawaii Conference* (pp.77–96). Heildelberg: Julius Gross Verlag. https://doi.org/10.1075/veaw.g2.07ric

Rickford, John (1979). Variation in a creole continuum: Quantitative and implicational approaches. Unpublished doctoral dissertation, Philadelphia: University of Pennsylvania.

Roopnarine, J. L., & Brown, J. (1997). *Caribbean families: Diversity among ethnic groups*. Greenwich, Connecticut: Ablex Publishing Corporation.

Sankoff, David, Sali A. Tagliamonte, & Smith, Eric (2005). *Goldvarb X: A variable rule application for Macintosh and Windows*. Department of Linguistics, University of Toronto. Goldvarb 0.0b3. Available from http://individual.utoronto.ca/tagliamonte/goldvarb.html

Satyanath, Shobha (2021). Genealogies of Sociolinguistics in India. *Journal of Sociolinguistics*, *25*(5), 762–784. https://doi.org/10.1111/josl.12496 (Theme Series: Decentering the Anglosphere).

Satyanath, Shobha (2018). Kohima: Language variation and change in a small but diverse city in India. In Dick Smakman & Patrick (Eds.), *Urban Sociolinguistics: The city as a linguistic process and experience*. Heinrich (Eds.). *Globalizing Sociolinguistics* (pp. 95–112). London and New York: Routledge.

Satyanath, Shobha (2015). Language variation and change: The Indian experience. In Dick Smakman & Patrick Heinrich (Eds.), *Globalizing Sociolinguistics* (pp.107–122). New York: Routledge.

Satyanath, Shobha (1990). The use of pronouns and its implications for the implicational model of decreolozation in Guyanese. *The Penn Review of Linguistics*, *14*, 129–142.

Satyanath, Shobha, & Sharma, Richa (2016). The growth of English in Delhi: New perspectives in a multilingual setting. In Jaspal Singh, Argyro Kantara, & Dorotty Cserző (Eds.), *Downscaling culture: Revisiting intercultural communication* (pp. 192–227). Newcastle: Cambridge Scholars.

Satyanath, T.S. (2018). Mapping Multilinguality in Medieval Indian literary culture. In K. Alfons Knauth, & Subha Chakraborty Dasgupta (Eds.), *Figures of transcontinental Multilingualism* (pp.115–135). Zurich: Verlag GmbH & Co.

Schlesinger, B. (1968). Family Patterns in the English-speaking Caribbean. *Journal of Marriage and the Family*, *30*, 149–154. https://doi.org/10.2307/350235

Sharma, Devyani, Levon, Erez, & Ye, Yang (2022). 50 years of British accent bias: Stability and lifespan change in attitudes to accents. *English World-Wide*. https://doi.org/10.1075/eww.20010.sha

Sharma, K.N. (1986). Changing forms of East Indian marriage and family in the Caribbean. *The Journal of Sociological Studies*, *5*, 20–58.

Smakman, Dick (2012). The definition of the standard language: A survey of seven countries. *International Journal of Sociology of Language*, *21*, 127–140. https://doi.org/10.1515/ijsl-2012-0058

Sunny, Neethu (2013). A sociolinguistic study of Malayalam in Cherukunnam. M.Phil. Dissertation. University of Delhi, Delhi.

Suokhrie, Kelhouvinuo (2016). Clans and clanlectal contact: *Variation and change in Angami*. *Asia-Pacific Language variation*, *2*(2), 188–214. https://doi.org/10.1075/aplv.2.2.04suo

Tieken-Boon vann Ostade Ingrid (1987). *The auxiliary Do in eighteenth-century English: A sociolinguistic approach*. The Netherlands: Floris.

Trudgill, Peter (1974). *The Social differentiation of English in Norwich*. Cambridge: Cambridge University Press.

Wassink, Alicia Beckford (1999). Historic low prestige and seeds of change: Attitudes toward Jamaican Creole. *Language in Society*, *28*, 57–92. https://doi.org/10.1017/S0047404599001037

Westphal, Michael (2015). Attitudes toward Accents of Standard English in Jamaican Radio Newscasting. *Journal of English Linguistics*, *43*(4), 313–333. https://doi.org/10.1177/0075424215607327

सार

समाजिक भाषाविज्ञान में बोलचाल की भाषा की केंद्रीयता के बावजूद, इसे निरंतर मानक के विरूद्ध सखा गया है जिसके कारण यह अपनी महत्ता स्थापित करने में विफल रहा है। पश्चिमी समाजिक भाषाविज्ञान में वेरिएंट्स की मानक या गैर-मानक के रूप में पहचान करके उनमें अंतर स्थापित करने का प्रचलन आम है। लेकिन, गैर-पश्चिमी संदर्भों में अंग्रेजी पर

शोध करते समय, ऐसा भेद हमेशा मान्य नहीं होता । यह अध्ययन गयाना में भारतीय मूल के समुदाय द्वारा बोली जाने वाली अंग्रेजी में पेरिफ्रास्टिक डू से प्राप्त एक प्री-वर्बल ऑक्ज़ीलरी के प्रयोग में वेरिएशन के बारे में बताता है । मैच्ड गाइज़ टेस्ट की एक श्रृंखला से यह तथ्य सामने आया कि बोलने वाले वेरिएंट्स के बारे में स्पष्ट रूप से यह नहीं बता पाए कि वे अंग्रेजी के हैं या क्रियोल अथवा मिश्रित हैं। इसके अलावा, उन्नत चरण में चल रहे परिवर्तन के बावजूद, ऑक्ज़ीलरी वेरिएंट का मूल्यांकन उसे कम या ज़्यादा महत्व दिए बिना तटस्थ रूप से किया जाता है। चूंकि स्टेडंर्ड और स्टेडंर्ड – नेस की धारणा पश्चिमी संदर्भों में वेरिएशन के समाजिक भाषावैज्ञानिक मॉडलिंग पर हावी है इसलिए वह उसको प्रभावित करती है ।

कीवर्ड: एंग्लोफोन क्रिओल्स, स्टेडंर्ड, दृष्टिकोण, राष्ट्रीय विचारधारा, गैर-पश्चिमी, भाषा विचारधारा, स्थानीय भाषा, गयाना

Address for correspondence

Shobha Satyanath
Department of Linguistics
University of Delhi
Delhi, 10007
India
ssatyanath.du@gmail.com

Publication history

Date received: 1 May 2022
Date accepted: 7 September 2022

Author Queries

- Please provide a complete reference for the citation '(Edwards, 1978), (Lambert, Hodgson, Gardner, & Fillenbaum, 1960), (Thomas, 2001), (Labov, 2016), (Labov, 1982), (Smakman, 2015), (Labov, 2003), (Sharma, 2017), (Rodríguez-Ordóñez, 2021), (Satyanath, 1991), (Satyanath, 2006), (Satyanath, 2001), (Baker, 2013)' in this article.
- Please provide a citation for the reference id "CIT0001 (Bakker, Peter. (2003)), CIT0006 (DeCamp, David. (1960)), CIT0007 (DeCamp, David. (1971)), CIT0009 (Devonish, H. (1978)), CIT0012 (Eckert, Penelope, & Rickford, John (Eds). 2001), CIT0014 (Edwards, Walter. (1975)), CIT0024 (Kroch, Anthony, & Small, Cathy (1980)), CIT0034 (Lambert, W.E., Lambert, R.C., Hodgson, R.C., Gardner, & S. Fillenbaum (1960)), CIT0035 (LASJ (2005)), CIT0036 (Le Page, Robert B. (1960)), CIT0038 (Meyerhoff, Miriam, Abtahian, Maya, Gafter, Roey, Horesh, Uri, Kasstan, Jonathan, Keegan, Peter, & King, Jeanette (2021))" since citation is missing in the article.

The role of older men in a phonological change

(ɣ) in Raga, Vanuatu

Marie-France Duhamel
Australian National University

Contrary to classic predictions associated with the gender pattern in variationism, the results of this study on the effect of age and gender on a phonological innovation suggests that older speakers drive innovation in this small Oceanic speech community of Vanuatu. Young and old men are prone to deleting the phonemic consonant, while women and middle-aged men tend to retain it. The v-shaped distribution of the variant requires considering the interactions and social status of individuals in this community where older men occupy the highest-ranking positions. The deletion does not appear to be stylistic, and multivariate analyses reveal the effect of surrounding vowels, sex and age, on the frequency of consonant deletion. The variation is interpreted as a change in progress towards the deletion of the velar fricative and its high incidence in younger men is explained by their frequent interactions with the older men.

Keywords: Oceanic language, phonetic variation, language change, apparent study, gender variation, v-shaped pattern, velar fricative, consonant deletion, diachronic change

1. Introduction

This study concentrates on the correlation of language change and gender in language variation (Labov, 2001). The non-Western social and cultural setting of this study contributes to the discussion on this correlation and the necessity to consider the specific context of speakers and discourse in our analyses (James, 1996; Meyerhoff, 1996). The focus of this paper is on the effect of age and sex on the

https://doi.org/10.1075/aplv.21002.duh
Asia-Pacific Language Variation 8:2 (2022), pp. 206–239 | ISSN 2215-1354 | E-ISSN 2215-1362

suppression of the velar fricative in the speech of Raga speakers, in Vanuatu.[1] This variant presented a challenging distribution in the corpus investigated, whereby the groups of younger and older men were prone to deleting the consonant while the middle-aged speakers of both sexes displayed little incidence, and no steady increase, of this innovative variant. To interpret the conditioning of the variation, I needed to consider the activities, interactions, aspirations, and social status of individuals in this community where older men occupy the highest-ranking positions, and where men are experiencing significant social changes. Men at each end of the age spectrum share the frequent loss of the velar fricative and I suggest that younger men adopt this pattern from the speech of the elders, with whom they are in frequent contact. Further, young men may use this variant for symbolic purpose, to confer on themselves the customary prestige and authority that is a distinctive feature of their respected elders. This study thus contributes to the wider literature that provides a better understanding of the role of social factors in language use (e.g., clan membership (Stanford, 2009), ethnic identity (Fimone, 2020)) by applying variationist methods to non-Western societies, and it responds to the call for more diverse sources of data (Stanford, 2016).

2. Vanuatu and Pentecost

The 83 islands of the archipelago of Vanuatu spread from northwest to southeast in an area southeast of the Solomon Islands and west of Fiji. In 2016, the country counted a population of just under 266,500 (Vanuatu National Statistics Office – Bureau National des Statistiques, 2020). The Republic of Vanuatu gained independence in 1980 from the joint control of Anglo-French rulers. To this date, English and French are still the main languages of education, but the national language is Bislama, an English-lexified creole. All three languages, Bislama, English, and French are official languages of Vanuatu (Republic of Vanuatu, 1980).

Vanuatu shows extreme linguistic and cultural diversity. In 3000 years, the language spoken by the people of the Lapita cultural complex went through widespread diversification to give rise to the 138 languages identified in 2015 (François, Franjieh, Lacrampe, & Schnell, 2015). The status of these local lan-

1. My study on three variables in Raga was a key case in the ARC Laureate project 'The Wellsprings of Linguistic Diversity' (2014–2019, Chief Investigator Nicholas Evans, Australian National University) which probed the factors contributing to variation and diversification in communities of different scales. This paper is adapted from Chapter 6 in my doctoral dissertation (Duhamel, 2020).

guages, their heritage value, and the duty of the Republic to protect them, are stated in the Constitution of Vanuatu (Republic of Vanuatu, 1980).

The Raga language (ISO 639-3 lml) is one of these languages. This Oceanic language belongs to the North and Central Vanuatu linguistic group, under the Central-Eastern Oceanic group, within the Austronesian family (Clark, 2009). Raga is primarily spoken in the northern part of the island of Pentecost, in north-central Vanuatu, by one of the nine largest speech communities of Vanuatu. Many Raga speakers also live in the urban centres of Port Vila, the capital city, on Efate Island, and Luganville, on Espiritu Santo Island. In 2016, this close-knit society counted 5,778 members in North Pentecost (Vanuatu National Statistics Office – Bureau National des Statistiques, 2020) and the overall number of 6,500 Raga speakers that was proposed in 2001 (Lynch & Crowley, 2001) is likely to have grown.

In North Pentecost, Raga is spoken by all community members, and in most circumstances, whether at home, or in formal settings such as religious services and community meetings. Adult Raga speakers rely on Bislama only exceptionally, with visitors speaking another language. Children are immersed in Raga language from birth and tend to be monolingual in their first years. Since 2015, Raga children have been taught their community language and its spelling[2] for the two first years of their schooling. Bislama and the languages of education, English mainly, and more rarely French, are later added to their linguistic repertoire. Raga can be regarded as a high vitality language, on account of this healthy intergenerational transmission. Other factors contribute to the vigour of the language: the high density of speakers in North Pentecost; the use of the language in new media; the community's important political profile in the years leading to and following the country's independence;[3] the high socio-economic status of a sizeable number of highly educated Raga speakers; and more recently, the income generated by North Pentecost's high quality kava crop.[4] In addition, the frequent practice of rites, storytelling, songs, and dances, reinforces the transmission of the local culture to the younger generations.[5]

---

2. The orthography of Raga was devised in 1972 by a group of speakers and a linguist (Walsh, 2005).

3. The political party whose program moved the country towards independence, and the first prime minister of the Republic of Vanuatu, originated from North Pentecost.

4. 90.2% of households grow kava in North Pentecost, while the country average is 32.4% (Vanuatu National Statistics Office – Bureau National des Statistiques, 2016).

5. According to what I observed during my stay in the Raga community and which moreover was confirmed for the PENAMA province (PENtecost-AMbae-MAweo islands) in a recent Vanuatu report (Vanuatu National Statistics Office – Bureau National des Statistiques, 2021, p. 117).

Raga has not departed significantly from the proto language reconstructed for the languages of North Central Vanuatu (Clark, 2009). It also shows little variation in its morphosyntax (Crowley, 2002) and phonology (Vari-Bogiri, 2011), and minimal dialectal differences (Walsh, 1995). Such conservatism and regional uniformity are uncommon in Vanuatu linguistics. It is in this atypical context of linguistic conservatism that this study investigates the likelihood of a phonological innovation.

3. The velar fricative in Raga

Raga's phonemic inventory presents 18 consonants and five vowels (Table 1).

Table 1. The phonemes of Raga

Consonants					Vowels		
Labial	**Labio-velar**	**Alveolar**	**Velar**	**Glottal**	**Front**	**Central**	**Back**
b	b^w	t d	k ^{ŋ}g		i		u
					e		o
v	v^w	s	ɣ	h		a	
m	m^w	n	ŋ				
		l					
		r					
	w						

The voiced velar fricative appears in the phonemic inventory of many North Vanuatu languages. More specifically, the languages of northeast Vanuatu, including Raga, all share this phonological innovation which reflects Proto Oceanic **k* or **ŋk* (Tryon, 1976, pp. 11–18). Clark (2009) proposes that Raga /ɣ/ reflects Proto North Central Vanuatu (PNCV) **k*, as in the example of *gamali*,[6] 'men's house', where ɣamali reflects PNCV *kamali. It is noteworthy that Raga /k/ is also a reflex of *k, but that the velar stop has a very restricted distribution and occurs only word-initially in a dozen grammatical lexemes (Walsh, 1982). The velar fricative, on the contrary, is a frequent phoneme in Raga and one of its rare phonetic innovations. It is found word initially or intervocalically, but not word-finally or next to a consonant, as exemplified in the minimal pairs:

6. In Raga orthography, <g> stands for /ɣ/, <g̃> for /ŋg/, <ñ> for /ŋ/.

golo	'wither, fade away'	***k**olo*	'small'
		***l**olo*	'inside'
geli	'dig, plant'	***w**eli*	'small centipede'
lago	'walk'	*la**g̃**o*	'be ready'
		*la**ñ**o*	'a fly'
ligo	'tie up'	*li**v**o*	'tooth'
manogo	'be cooked'	*manon**o***	'embers'
saga	'be wearing (clothes)'	*sara*	'open space; field'

This phonemic consonant shows broad variability in the speech of the 52 participants surveyed in this study, with variants displaying a large phonetic distribution, but no features that allowed for regrouping them. Realisations of the variable (the 'Other' variant category in Table 2) presented a lenited form of the velar fricative, e.g.[h], or a fortition of the consonant, e.g. [k], while for other sounds e.g. [ʁ] there was no gain or loss of sonority.

Table 2. Categories of realisations of (ɣ) observed in my overall corpus

Variant category	Description	Example	Total tokens
ɣ	velar fricative	*dogo* [doɣo] 'stay'	1,067
Other	segment other than the velar fricative	*bisigai* [bisi**h**ai] 'refuse' *geki* [**k**eki] 'this'; *sigai* [si**ʁ**ai] 'negative'; *gabwe* [**w**ambwe] 'break'	223
Zero	deleted segment	*ige* [ie] 'fish'	197
Total			**1487**

Previous work on the language has reported on the variability of the velar fricative (Codrington, 1885; Ivens, 1938; Vari-Bogiri, 2011; Walsh, 1982) with a focus on the frequent alternation between the phonemic velar fricative /ɣ/ and velar stop /k/. The authors observed that this variation occurred in the initial position of high frequency grammatical lexemes, such as the demonstratives *geki* 'this' and *gea* 'that', the adverb *gaha* 'now' and the preposition *gabe* 'if'. Ivens proposed that euphony was a possible factor in the alternation between the two velar consonants,[7] but no other structural conditioning factor was suggested in the

7. Ivens does not describe the euphonic rule that he proposes to be at play in this alternation, but the examples he mentions (Ivens, 1938, p.736) suggest that the initial velar fricative is favoured when preceding /e/, and the velar stop when preceding /a/.

descriptive material cited above. It was remarked that the alternation was based on the choice of individuals (Codrington, 1885; Walsh, 1982), used in free variation (Vari-Bogiri, 2011, p. 268), and not a dialectal variant (Codrington, 1885, p. 432). My primary intention, in the study of this phonetic variable, was to investigate the linguistic and social factors constraining the /k/ and /ɣ/ alternation in my corpus. It quickly emerged, however, that tokens of zero variants were occurring more readily in the speech of male speakers. None of the previous studies mentioned cases of deletion of the consonant. A variationist approach to the zero variant was the occasion to test the hypothesis of an ongoing weakening of the proto-velar consonant, and to shed light on the underlying cause of the alternation between the phonemic velar fricative /ɣ/ and velar stop /k/ and the overall instability of the velar fricative.

3.1 Hypothesis: PNCV *$k > \gamma > 0$

It is proposed in this study that the historical weakening of the proto velar stop, from PNCV *k to γ, is a continuing process in Raga, and that the deletion of the Raga consonant observed in the speech of groups of male participants provides evidence of the final stage of lenition.

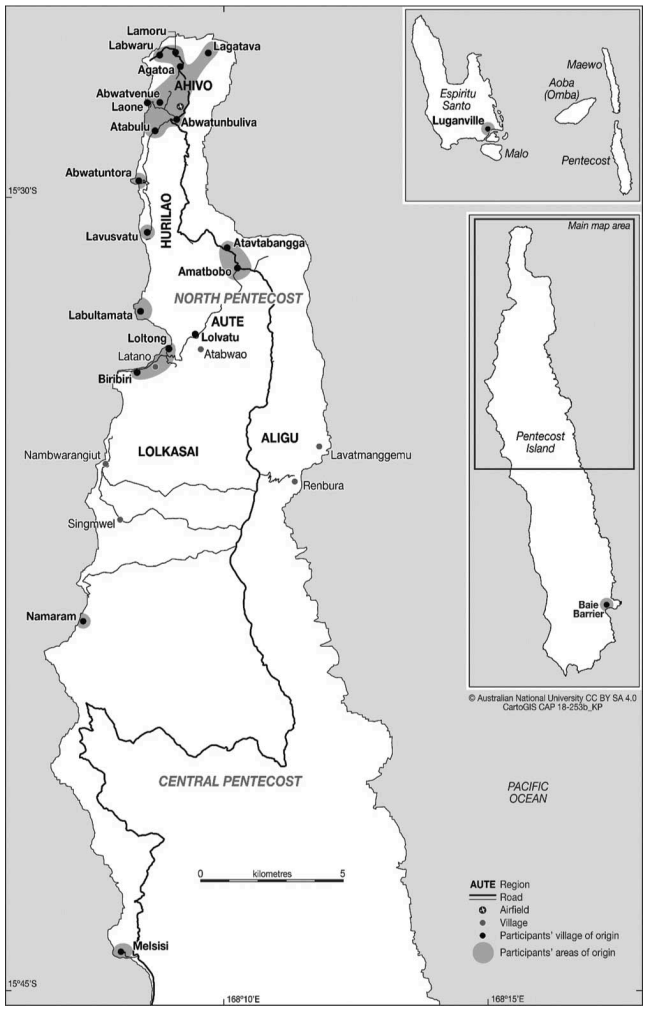
4. Dataset and methodology

Tokens of the variable were taken from 74 recordings of natural speech of different performative styles, collected from 52 Raga speakers, 25 women and 27 men, aged 15 to 89 years. These recordings are drawn from an overall corpus of 137 narratives recorded in North Pentecost, during four trips, approximately 11 months, from 2015 to 2017.[8] The participants surveyed for this phonetic variable were selected to ensure that all the potentially explanatory social factors were represented. Not all speakers could be recorded for each speaking style, and, for this study, recordings were selected from my corpus to ensure that a balanced number of variable tokens were extracted for each performative style.

The surveyed speakers came from different areas in North Pentecost and beyond (Map 1), and all but two participants were native speakers of Raga.

Tokens were extracted from recordings of Raga speakers of both sexes, distributed by age, representing all education levels and both matrimoieties, with a history of high or low mobility, and for whom Raga was their first language or a language acquired subsequently.

8. See Chapter 3 in Duhamel (2020), for maps and description of trips.



Map 1. (Duhamel, 2020, p. 33). The places of origin of the speakers surveyed in this study, in North Pentecost and elsewhere in Vanuatu

Speakers were divided into three age groups which reflected their distinct social roles and positions in the society. Community members under the age of 25 still had to establish their place in Raga society, although there was a marked gender difference within this age group: many young women under 25 already had the charge of a family and participated in the activities of older adults, whereas most young men were still single, living with their male peers, largely free of family commitments, they had specific tasks in the preparation of kava and duties towards older men. Men and women aged between 25 and 50 years, were rarely single and typically in charge of a family. The backbone of the community,

they managed and participated in most community activities, whether producing food, building structures, or organising events. The roles of women and men in this age group were mostly complementary: women's tasks included caring for children, preparing food, weaving, doing the laundry, feeding pigs, contributing food and other resources to public events; men did the heavy work in the gardens and on buildings and shared structures. In the oldest group, 50 years and over, were the leaders and highest-ranking members of the society, whose advice was sought and valued by members of the community. Some women in this age group had gained chief titles but these did not grant them the authority of male chiefs.

Phonological and lexical differences are likely to be more salient to speakers (and possibly to language investigators) than linguistic differences in syntactic and morphological domains (Walker, 2010). In the case of Raga, there was a paucity of detectable phonological variability in the corpus under study. The different realisations of the velar fricative constituted the most perceptible phonological variable, and it is noteworthy that an analysis of speakers' variation in their production of vowels revealed variability within the speech of the same speaker, but negligible inter-speaker differences (Duhamel, 2020, pp. 173–174). Dropping of the voiced velar fricative was not a noticeable feature in the corpus; its incidence in the speech of groups of speakers was only revealed when the corpus was transcribed and the phonological variable annotated for its range of variants.

The data was transcribed by language assistants, native speakers of Raga who had been trained on ELAN [Computer software] (2021). The transcription was done in my presence, or I would later review and discuss it with the transcribers. Variants of the velar fricative were then annotated auditorily by the author in ELAN and assigned to one of the three categories:

- the velar fricative variant 'ɣ'
- the zero variant 'Ø'
- the 'other realised' category, including all variants of the consonant other than the velar fricative

When auditory impressions of the acoustic signals did not allow their categorisation, the author resorted to spectral analysis with the software PRAAT (Boersma & Weenink, 2019), as shown in the spectrograms in Figure 1 and Figure 2 for the word *atagu-n* 'after-POSSESSED'.[9] Figure 1 shows the zero realisation of the word-medial variable and Figure 2 its velar fricative realisation.

9. The suffix *-n* indicates that a possessor is present but not indexed by the suffix.

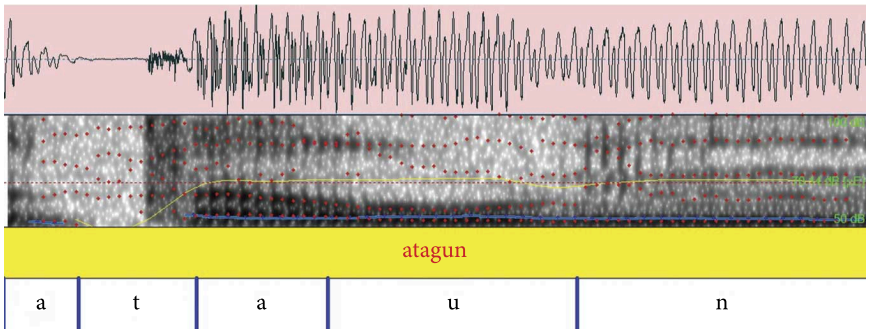


Figure 1. Spectrogram of the zero realisation of the variable (ɣ) in the word atagu-n 'after-POSSESSED'

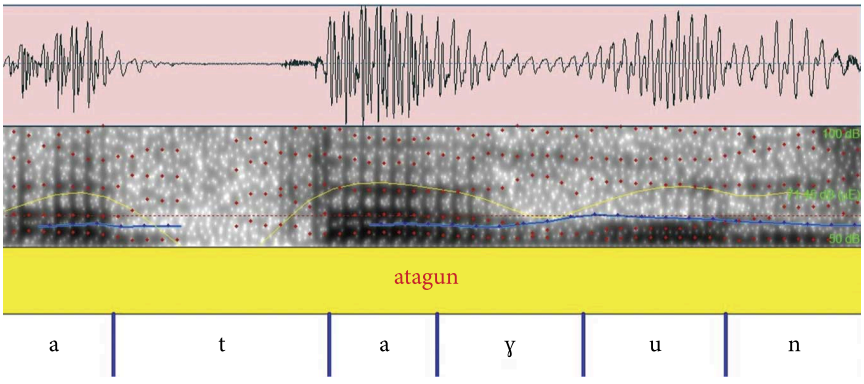


Figure 2. Spectrogram of the velar fricative realisation of the variable (ɣ) in the word atagu-n 'after-POSSESSED'

Initially, 1,487 tokens were coded. However, for the study reported here, only two categories were considered: the zero variants ($n = 197$), and renditions of the velar fricative ($n = 1{,}067$). The other variants ($n = 223$) were excluded from the statistical analysis. Lenited forms of the velar fricative which fell in this third category could have been included, but, firstly, the deletion of the consonant offered a clear contrast with the velar fricative,[10] and secondly, the zero variants fitted the hypothesis that they represent the final stage of a continuing lenition of the proto-consonant **k* for this language.

This study is therefore restricted to the 1,264 tokens coded for (a) velar fricative realisations, which account for 84.4% of these tokens, and (b) deleted vari-

10. Such a contrast can be difficult to establish with certainty for lenited variants of fricatives, as discussed in Mansfield (2015).

ants, which account for 15.6%. Table 3 below presents the distribution by age and sex of the 52 surveyed speakers and, for each variant, the number of tokens extracted from the narratives collected for these speakers.

Table 3. Distribution of the participants by age and sex. In brackets, number of tokens of the variants ɣ: Number of tokens of variant Ø, by age and sex cohort

Speakers	> 50 years	25 to 50 years	< 25 years	Total (ɣ : ø)
Women	4 (100:12)	12 (287:38)	9 (150:29)	25 (537:79)
Men	13 (234:44)	7 (195:25)	7 (101:49)	27 (530:118)
Total	**17 (334:56)**	**19 (482:63)**	**16 (251:78)**	**52 (1067:197)**

4.1 Conditioning factors of deletion

Multiple factors, structural, social, and stylistic, were considered for their potential impact on the deletion of the velar fricative. These factors are listed in Table 4:

Table 4. The factors predicted to influence renditions of (ɣ)

Factor	Nature	Levels	
Words surveyed	structural	277	
Position in the word	structural	3	morpheme boundary; word initial; word medial
Following vowel	structural	5	the five phonemic vowels
Preceding vowel	structural	6	the five phonemic vowels and 'P' for pause
Individual speakers	social	52	
Sex of the speaker	social	2	female; male
Speaker's age group	social	3	under 25 years; above 25 and under 50; above 50
Speaker's native language	social	3	Raga; Bislama; Apma
Speaker's district of residence	social	2	Loltong Bay; Ahivo
Mobility of the speakers	social	3	low; high; unknown
Speaker's matrimoiety	social	2	Bule; Tabi
Speaker's education level	social	4	Primary; secondary; tertiary/training; unknown
Narrative type	stylistic	4	traditional stories; disaster stories; personal stories; interviews

4.1.1 *Structural factors*

The lexical items ($n = 277$) from which the variants were extracted, were annotated. Care was taken to diversify the words that were coded, and the lexical class that they fell into. It was not however entirely possible to avoid that some common lexical items, especially function words, were more frequently coded than others. To account for the effect of the repeated occurrence of these words, word was included in our statistical test as a random effect, though see Section 4.2 for further details.

The phonological context of the variable was coded in ELAN for each variant. The position of the variable in the word was annotated, which could be initial or medial; a further distinction was added for variants at a morpheme boundary, e.g., *doroni-go* 'want-2SG.OBJECT'.

The segments surrounding the variable were also annotated, that is, the vowel preceding and the vowel following the variants. Word-initial variants preceded by a pause were coded as such. When no pause was observed after a word-final vowel and before a word-initial variant, word boundaries were ignored, and the vowel was considered as preceding the variant. In the case of the following vowel, since the velar fricative does not occur word-finally, the following vowel is always within the word boundary.

4.1.2 *Social and stylistic factors*

Social factors were coded based on data elicited from participants' interviews and my conversations with members of the community. Informal talks were especially valuable for the information they provided on the social and cultural workings of the Raga society, people's aspirations and motivations, and their attitudes to some social groups. Combined with my observations in the field, the remarks of participants substantiated my interpretation of the social meaning of this variable.

The 52 surveyed speakers in this study were stratified by age and sex to test for intergenerational and intergender differences that may reveal language change over time.

Native language and place of residence were also explored as possible independent variables. Fifty participants were native speakers of Raga, a proportion of native speakers that was representative of speakers in North Pentecost. The two non-native speakers of Raga were fluent speakers of the community language and therefore were included in this study. The native language of the first non-native speaker, a 43-year-old woman, was Apma, the Oceanic language spoken south of Raga community's boundary. Married to a Raga man, she had been speaking Raga for 10 years. The second non-native speaker, a 20-year-old woman, had grown in a Raga-speaking household, but on Espiritu Santo Island, where Bislama, the

national language, is used in most domains. Neither Apma nor Bislama include the velar fricative in their phonemic inventory, which suggested that the native language of these two speakers might influence null realisation of the velar fricative in their speech.

Data was collected in two main locations of North Pentecost, in Loltong Bay ($n = 40$ participants), in the southwest region of North Pentecost, and in Ahivo district ($n = 12$ participants), the northern tip of Pentecost (Map 1). Tracking the place of residence of speakers allowed to explore regional variations in the choice of variant. The region of Ahivo also has special significance in North Pentecost since creation stories and oral history give the northern tip of the island as the original location of the settlement of Pentecost (Taylor, 2008; Yoshioka, 1987), which would add a diachronic perspective to any regional variation that might be observed for this variable.

Mobility referred to the current and past frequency of movement and length of stay of participants outside the Raga linguistic community. In this study, it was assumed that a low ($n = 33$ participants) or high ($n = 12$ participants) degree of mobility of speakers would reveal the impact of contact with other languages in the choice of variant. Information on the movements of participants[11] also revealed the exposure of Raga speakers to other societies and cultures, and it helped refine their social profile for this study.

Raga people are divided into exogamous moieties, Bule and Tabi, whose membership is inherited from one's mother. These matridescent lines determine who one can marry and regulate one's inherited land rights, the gardens that one can work and the locations where one can set up a dwelling. Moiety membership influences who Raga speakers regularly associate with, and therefore their social circle, and it also represents a significant facet of participants' identity; it was therefore fitting to explore any potential relationships between participants' realisation of the variable and this social partition, given well-established principles of change with regards to social networks (e.g., Milroy, 1980). The two descent lines were represented by a near equal number of participants: Bule ($n = 27$) and Tabi ($n = 25$).

The level of Western education, ranging from no schooling to tertiary education (Table 5), was considered for its impact on Raga's linguistic features of the languages of education (English or French), and the national language (Bislama), which is the language of communication of participants schooled away from North Pentecost. The level of education is correlated with age, in part because access to the Western education system has been made easier over the years. Most young speakers reported several years of secondary education, and were too

11. Information on their mobility could not be collected for seven participants.

young to have attended tertiary education, whereas only half of individuals in the middle age group had a secondary education and many older speakers had only attended primary school or did not mention their years of schooling.

Table 5. Distribution of the surveyed speakers by education level

Education level	>50 years	25 to 50 years	<25 years	Total speakers
None	1	0	0	1
Primary	6	8	3	17
Secondary	1	7	12	20
Tertiary/Training	3	2	1	6
Unknown	6	2	0	8
Total	**17**	**19**	**16**	**52**

The stylistic factor reported here consisted in the performative style of the narratives annotated for this variable, out of the 137 narratives collected in this Raga corpus.[12]

Narratives were divided into four main types: traditional stories; natural disaster stories; interviews; personal stories. The stylistic context of the variable, i.e., the type of narrative, was thus assigned to each variant. Some of the recorded narratives presented more careful speech than others, which allowed for the quantitative distribution of the deleted consonant to be tested across a stylistic continuum from more to less careful speech. The selection of narratives was to ensure that all types of data would be represented for the different age groups and each sex (Table 6). Between 15 and 20 tokens of the variable were coded for each narrative.

Table 6. Distribution of the surveyed narratives, by speakers age group and sex (F:M)

Narratives	>50 years	25 to 50 years	<25 years	Total (f:m)
Traditional stories	10 (3:7)	9 (4:5)	3 (1:2)	22 (8:14)
Disaster stories	5 (2:3)	7 (5:2)	3 (2:1)	15 (9:6)
Interviews	6 (2:4)	9 (7:2)	10 (6:4)	25 (15:10)
Personal stories	1 (0:1)	2 (1:1)	8 (4:4)	11 (5:6)
Total	**22 (7:15)**	**27 (17:10)**	**24 (13:11)**	**73 (37:36)**

12. A total of 74 narratives were annotated for the variable, but one young woman's very short traditional story yielded no zero or velar fricative variants and is therefore not included here.

The traditional stories, *veveve* ($n=22$ texts), the Raga legends, creation, and historical stories that Raga speakers recite since childhood, gave the most careful speech. These foundation stories teach younger generations their culture and history, and their telling appeared to be greatly normalised. Many young participants were reluctant to tell these stories, stating that they did not know them well enough, and some young participants who offered to recite them would check for the approval of the Raga audience during their performance. The three other types of narratives provided samples of spontaneous speech. Accounts of natural disasters ($n=15$ texts) prompted the least controlled speech. Recollecting such events can be unpleasant for participants these recordings were only collected from the participants whom I had met on several occasions. Interviews of participants ($n=25$ texts) were conducted in Raga by native speakers who were instructed to create opportunities for conversation. The interview consisted of a list of questions to gather information about the speakers: age, place where they grew up, native and other languages, marital status, education, mobility, activities in the community and network of friends. Some interviewees did not need to be prompted to comment or add clarification, thus providing samples of natural speech, but other participants kept their answers very brief. Additional stories (overall, $n=11$ texts) were collected from participants for whom we had little recorded speech, and consisted of participants' personal stories, for example a fishing party or traveling to an event, narratives prompted by board stories,[13] or participants' account of the 'North wind and the Sun' story.

4.2 Analysis

After a first exploration of the distribution of extracted tokens, three of the social predictors described above, native language, mobility, and place of residence, were excluded from the statistical analysis. Sample sizes were skewed for native language and mobility factors, since most annotated variants were extracted from the speech of native Raga speakers ($n=50$), and low mobility participants ($n=33$). A first exploration with a chi-square test also found that speakers' district of residence had no significant correlation with incidence of deletion ($\chi^2(1, n=1264)=1.61$, $p=0.203$). The social factors included in the analysis were therefore reduced to speaker's sex, age, matrimoiety, and education.

The factors hypothesised to correlate with the zero variant were tested for statistical significance using a generalised linear regression model. The statistical

13. Including 'The Family Problems Picture Task' (Barth & Evans, 2017; San Roque et al., 2012) and the 'Bandel Banana Cards' (von Prince, Krajinović, Krifka, Guérin, & Franjieh, 2018).

models were run in the R software environment (R Core Team, 2013). The model used for this categorical dependent variable was the generalised linear mixed-effect model (GLMM, in the R package *lme4*) which allowed for including fixed effects and random effects in the statistical tests. The models were run with the random effects of speaker.[14] The fixed effects comprised the structural, social, and stylistic factors described above and not previously excluded from the analysis.

The study then ranked the individual speakers for their incidence of deletion of the consonant. This allowed for the exploration of social factors which were not included in the statistical models, or whose nature was more amenable to qualitative examination (i.e., status in the society, and attitudes to other languages or social groups).

5. Results

5.1 Statistical models

Following a stepwise procedure, the independent variables not selected for their influence by the mixed-effect models were eliminated.

Separate models were run for structural factors on the one hand, and social and stylistic factors on the other. This allowed for reducing issues of collinearity among factors.

The predictors were tested with the random effect of speaker for its potential explanatory effect. The model with the only random effect of speaker was used as the base model to which I added the independent variables to be tested for significance and model fit. The base model's AIC score was 1039. Subsequent models needed a lower AIC score to show that they fitted the data better.

5.1.1 *Structural model*

The vowel preceding and the vowel following the variable (see Appendix A. for tables showing distribution of the variants) were selected by a statistical model for their significant effect on the deletion of the velar fricative. This model (Table 7) also proved a better fit for the data (AIC = 989.7) than the base model.

The model shows the vowels that have a significant effect on increasing the rate of deletion of the velar fricative when preceding or following the variable, as

14. The model used did not allow for combining the two random effects of word and speaker in a single model, and models with the random effect of word ($n = 277$) issued convergence warning messages, which seems to occur when a model is run with a large set of random effects, and these models were therefore not reliable.

per the positive coefficients associated with deletion of the consonant. High vowels /i/ and /u/ were returned as significant factors when preceding or following the variable; mid-high vowels were also selected: /o/ when preceding the variable, and /e/ when following it. Place of articulation, and in particular tongue height, is therefore a significant factor in the deletion of the velar fricative.

Table 7. Result of the mixed effect linear regression model for the zero variant of the velar fricative. The model estimates the fixed effects of the vowel preceding and vowel following the variable, accounting for the individual effect of speaker. AIC = 989.7

Random effects					
Groups	**Variance**	**std.dev.**			
Speaker	0.9781	0.989			
N = 1264, Speakers = 52					
Fixed effects					
Independent variables	**Estimate**	**std. error**	***z*-value**	***p*-value**	**Significance**
(intercept)	−2.6952	0.2684	−10.043	< 0.001	***
Preceding vowel					
/i/	1.0770	0.2638	4.083	< 0.001	***
/o/	0.7376	0.2523	2.923	< 0.01	**
/u/	1.0283	0.3719	2.765	< 0.01	**
Following vowel					
/e/	1.0899	0.2571	4.240	< 0.001	***
/i/	0.8254	0.2683	3.077	< 0.01	**
/u/	0.7935	0.3410	2.327	< 0.05	*

A slightly higher rate of deletion of the consonant was reported in word-initial position (Table 8), however the position of the variable in the word was not selected for its effect on the likelihood of the consonant deletion.

Table 8. Distribution of the variants by position of the variable in the word

	[ɣ] retained		[ɣ] dropped		Total
Morpheme boundary	21	87.5%	3	12.5%	24
Word initial	525	81.6%	118	18.4%	643
Word medial	521	87.3%	76	12.7%	597
Total	**1067**		**197**		**1264**

5.1.2 *Social and stylistic model*

Statistical tests, including the random effect of speaker, revealed that speaker's age had a significant influence on the incidence of deletion in the speech of the 52 speakers. Speaker's sex also showed some tendency towards affecting the rate of deletion. An earlier model had shown that the speakers under 25 years were the only group whose speech displayed significant increase in the incidence of the deletion of the consonant, and for this model the three age groups were therefore reduced to two, dividing the speakers on each side of 25 years. This model (Table 9, AIC = 1036.8) fitted the data slightly better than the base model.

Grouping males over 25 together may suggest that the middle-aged men pattern like the older men, though the analysis by individual speakers (Figure 3, below) shows that the speech of men over 65 departs from what is observed in men in the 25 to 65 age bracket in their rate of dropping the velar fricative. This is not picked up by the model since age groups, as they were defined in this study, do not align precisely with the observed rates of deletion.

Table 9. Summary of the mixed effect linear regression model for the zero variant of the velar fricative. The model estimates the fixed effects of speaker's sex and speaker's age, accounting for the individual effect of speaker. AIC = 1036.8

Random effects					
Groups	**Variance**	**std.dev.**			
Speaker	0.5769	0.759			
N = 1264, Speakers = 52					
Fixed effects					
Independent variables	**Estimate**	**std. error**	***z*-value**	***p*-value**	**Significance**
(intercept)	−2.4241	0.2438	−9.941	< 0.001	***
Speaker's sex					
Male	0.5317	0.2867	1.855	0.063	tendency
Speaker's age group					
Young	0.7669	0.3086	2.291	< 0.05	*

The two other social predictors, education and matrimoiety, showed no effect on the rate of deletion of the velar fricative. Chi-squared tests of speaker's education (Table 10) and moiety (Table 11) had revealed correlation between each of these two social factors and the dependent variable, however the models did not select these two factors for their effect and returned worse fit than the base model.

Table 10. Number of tokens of the variant ɣ and variant Ø by speakers' education level

Education level	**[ɣ] retained**		**[ɣ] dropped**		**Total**
None	19	90.5%	2	9.5%	21
Primary	358	89.5%	42	10.5%	400
Secondary	398	79.6%	102	20.4%	500
Tertiary/Training	175	84.1%	33	15.9%	208
Unknown	117	86.7%	18	13.3%	135
Total	**1067**		**197**		**1264**

$\chi 2$ (4, $n = 1264$) = 17.79, $p = p < 0.05$.

Table 11. Number of tokens of the variant ɣ and variant Ø by speakers' matrimoiety

Moiety	**[ɣ] retained**		**[ɣ] dropped**		**Total**
Bule	600	81.7%	134	18.3%	734
Tabi	467	88.1%	63	11.9%	530
Total	**1067**		**197**		**1264**

$\chi 2$ (1, $n = 1264$) = 9.49, $p < 0.05$.

Similarly, Chi-squared tests had revealed correlation between the zero variant and the style of the narrative (Table 12), but style was not returned as a significant factor. The rate of dropped [ɣ] displayed by the least controlled narrative type, disaster stories, is barely higher than for the most controlled type, traditional stories, and the deletion of the consonant does not appear to be influenced by the attention paid to speech. Furthermore, including this parameter as fixed effect in the models increased AIC scores.

Table 12. Number of tokens of the variant ɣ and variant Ø by narrative style

Narratives	[ɣ] retained		[ɣ] dropped		Total
Traditional stories	334	88.4%	44	11.6%	378
Disaster stories	264	84.9%	47	15.1%	311
Interviews	117	78.5%	32	21.5%	149
Personal stories	352	82.6%	74	17.4%	426
Total	**1067**		**197**		**1264**

$\chi 2$ (3, $n = 1264$) = 9.48, $p < 0.05$.

5.1.3 *Combined model*

Finally, mixed-effect models were run to assess the combined influence of the structural and social predictors previously selected. Models issued convergence warning messages when the structural predictors were included as fixed effect alongside the social predictors, and consequently separate models were run for each structural factor combined with the social factors.

The resulting best fit model (Table 13) combines the fixed effects of preceding vowel, age, and sex, as, and the random effect of speaker.[15] Preceding vowel, young participants and men, were selected for their significant effect on increasing the rate of deletion of the velar fricative when preceding it.

This unifying model (AIC = 1005.3) fits the data better that the model with the random effect of speakers only, but worse than the model with the effects of preceding and following vowel (AIC 989.7), both of which could not be included in the model.[16]

In summary, this analysis reveals that the surrounding vowels have a strong positive effect on the deletion of the consonant, and that the young group of speakers, and men, also favour the dropping of the velar fricative.

15. A model with the fixed effects of following vowel, age, and sex was a worse fit for the data (AIC = 1021.2).

16. Another model, with the fixed effects of preceding vowel and interaction between age and sex, allowed establishing the effect of age-by-sex groups, but its AIC score was higher (1007.2).

Table 13. Summary of the mixed effect linear regression model for the zero variant of the velar fricative. The model estimates the fixed effects of the interaction of speakers' age and sex and the vowel preceding the variant. It also accounts for the individual effect of speaker as random effect. AIC = 1005.3

Random effects					
Groups	**Variance**	**std.dev.**			
Speaker	0.5839	0.7641			
N = 1264, Speakers = 52					
Fixed effects					
Independent variables	**Estimate**	**std. error**	***z*-value**	***p*-value**	**Significance**
(intercept)	−2.8938	0.3070	−9.427	< 0.001	***
Preceding vowel					
/ i/	1.0792	0.2607	4.140	< 0.001	***
/ o/	0.7367	0.2491	2.957	< 0.01	**
/u/	1.0064	0.3675	2.738	< 0.01	**
Speaker's sex					
Male	0.5948	0.2907	2.046	< 0.05	*
Speaker's age group					
Young	0.8599	0.3150	2.729	< 0.01	**

5.2 Individual speakers

The value used for ranking speakers on their rate of deletion of the consonant was the estimated deviation between each speaker's average incidence of deletion and the overall average. This value was calculated in R by the function RANEF of the package *plyr* and was derived from the generalized linear mixed model (function GLMER) that considered the sole random effect of speakers on the deletion of the consonant. This value (henceforth called (average) deletion rate/incidence) was plotted on the graph in Figure 3 (see Appendix B for detail of values obtained with RANEF). The graph shows the distribution of the deviation between a speaker's average incidence and the average incidence of all surveyed speakers (y-axis), male (triangular symbol) and female (round symbol), by speakers' age (x-axis), for the deletion of the velar fricative. Values plotted above the '0' horizontal line mark speakers who showed a higher average rate of deletion than the calculated average of speakers, whilst speakers with a lower than average rate of deletion of the consonant are plotted underneath the '0' horizontal line.

The distribution of the deleted velar fricative, by speakers' age and sex plotted in Figure 3 shows that most speakers are plotted underneath the '0' horizontal line, which means that most speakers across all age levels tended to not drop the velar fricative. It also shows that speakers whose speech deviated most from the other speakers for this variable and presented a positive average incidence of the deleted variant, are mostly men, under the age of 40 (left circle), or over 65 years (right circle). Figure 3 thus reveals the grouping of three generations of male speakers (circled): middle aged men, aged 40 to 65, who presented a positive average incidence of velar fricative realisations and a negative average incidence of its deletion, and men (under 40 and over 65) who showed a low incidence of the velar fricative variant while also showing a high incidence of its deletion. In their speech, women over the age of 40 showed a clear tendency to use the velar fricative variant, while younger women presented equal average incidence of both variants. The highest rate of deletion was found in the speech of a young woman (top left of the graph).[17]

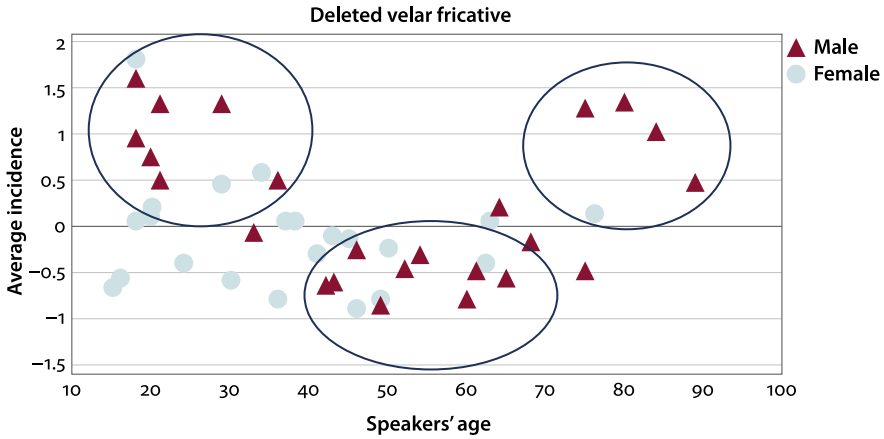


Figure 3. Distribution of the deleted velar fricative by speakers' age and average incidence of deletion

17. The speech of this 18-year-old woman exhibited variants that constantly differed from the average of speakers, in the other variables that I examined in Raga. For a discussion of her profile refer to Chapter 7 of my doctoral thesis (Duhamel, 2020).

6. Discussion

The statistical tests have selected the surrounding vowels, the age and sex categories, for their significant positive influence on the rate of deletion of the velar fricative.

The effect of the surrounding vowel had been noticed at the time of transcription, in particular for the preceding /i/, in common terms such as *ige* 'fish', *naturigi* 'child', *ratahigi* 'leader'. The articulatory work needed for the tongue to move from the fronted vowel to its raised position close to the velum may in part explain the dropping of the velar fricative in this environment. The examples provided by Ivens (1938, p.736) suggested that /e/ had a weakening effect when following the phonemic velar consonants, and this effect is confirmed in this study. A study on the realisation of the velar fricative in the Turkish language has also investigated the effect of the vowels preceding or following the consonant (Ünal-Logacev, Żygis, & Fuchs, 2017). The authors determined the constraining effect of the surrounding vowel on the velar fricative, which was deleted in all intervocalic contexts and realised as a lengthening of the preceding vowel.

It is also worth noting the lack of influence of the speech style on the deleted variant. This is of importance since the lack of variability in the deletion of the velar fricative across more to less formal styles of speech precludes the interpretation of the supressed variant as stylistic.

The focus of this paper is on the social explanatory factors, and the discussion now turns to interpreting their results, and in particular the curvilinear shape of the pattern displayed in Figure 3.

6.1 Interpreting the v-shaped pattern

The v-shaped pattern displayed in Figure 3 offers two possible interpretations of the deletion of the velar fricative: (1) the pattern indicates a stable, age-graded, variant, or (2) it indicates an innovative variant, on account of its high incidence in the group of young speakers.

Our interpretation here hinges on the notion of a linguistic norm, and variation associated with this norm. In a small language community such as Raga, whose language is largely unwritten and non-standardised, what constitutes a linguistic norm is negotiated, rather than codified, hence the necessity to examine the important social structures that speakers might orient to (Meyerhoff et al., 2020; Stanford, 2016). In this study, it was assumed that the phonemes identified for the Raga language constitute the norm; there does not appear to be consensus as to how Raga speakers should speak, and the use of 'standard' to refer to linguistic norm is best avoided. For this variable, the velar fricative is the most com-

monly occurring variant, when in variation with other sounds, or when deleted, but it is not the variant used by the people with most prestige; indeed, this study shows that these speakers, the older men, are more likely to delete the consonant. There appeared to be some awareness of the lenited forms of the velar fricative,[18] but no participants surveyed in my corpus remarked on the speakers that might use lenited forms of the phoneme. Neither did they mention the deletion of the phoneme.

The pattern of usage across the different generations of speakers surveyed in this apparent time study presents the main argument in favour of interpreting the deletion of the velar fricative as a stable age-graded variant. Higher incidence of deletion was observed in younger and older speakers but the speech of those aged 40 to 65 revealed lower than average rates of deletion. From this distribution it appears that speakers use this variant at the onset of adulthood, lose the variant when reaching middle age, and that some of them return to using it in old age.

Previous studies have found that stable variation tends to involve variables above the level of consciousness that show stylistic variation. Such phonetic variation was reported for the variable (ing) in diverse varieties of English (for example, Trudgill, 1974, in British English; Wagner, 2012, in American English). In these varieties, the stigmatised alveolar variant [ɪn] was found to correlate with lower social classes and casual speech. The variation between inflected and periphrastic future in Canadian French (Wagner & Sankoff, 2011) provides another example of stable variation, the former variant correlating with upper social classes.

The social class hierarchy used in these studies on Western languages does not apply to the context of the Raga speech communities, and one challenge presented by the present study was to work out which social criteria the Raga speakers might consider in order to evaluate variation. As mentioned previously, some Raga speakers were aware of weakened forms of the velar fricative and the variation between the velar fricative and the velar stop was mentioned in my conversation with participants, but speakers did not seem to link the variant to any specific social group.[19] As for the deletion of the velar fricative, firstly, speakers did not

18. Posted on a social media forum, a comment in Raga mentioned that speakers were expected to 'make noise' when pronouncing words that include the velar fricative (i.e., *go-m* [ɣom] '2SG.S-PROG'), which they opposed to [hom] that they heard in the speech of some Raga speakers (Duhamel, 2020, p. 158).

19. The reason may have been that participants did not wish to give a commentary on the linguistic performance of fellow speakers to an outsider.

appear to be aware of this variant[20] and, secondly, the statistical models showed that this variation was unlikely to be stylistic since it did not correlate with the topic or register of the surveyed narratives.

The possibility that Raga speakers did not associate social norms with some variants also needed to be considered. Reflecting on her study on Eastern Sutherland Gaelic, Dorian (2001) remarks that several case studies show that 'a considerable amount of unweighted linguistic variability exists in small communities that are buffered in one way or another from the development of normative judgments' (Dorian, 2001, p. 147). This absence of normative judgment towards individual variation could be expected in linguistically diverse Vanuatu, whose population celebrates the country's diverse linguistic varieties, an attitude that has been reported in several small Melanesian societies (François, 2012; Haudricourt, 1961; Wurm, 2007). In the context of this deleted consonant in Raga, however, I propose another explanation for this absence of normative judgement. Based on the reaction of transcribers, I suspect that, by and large, the deletion of the velar fricative was not detected by speakers. However, in the eventuality that some speakers may have detected the sound being dropped, the elevated status of the older individuals, whose speech exhibited a high incidence of this variant, would have inhibited a negative evaluation of the variant.

The deletion of the velar fricative, therefore, does not appear to be a stable, stylistic variant and I propose to analyse it as change in progress. To interpret the distribution of this variant across different generations of speakers and discuss the specific role of older men in this innovation, I will draw from previous studies and my observations of the social life in North Pentecost.

6.2 The role of older speakers

The main argument in favour of interpreting the loss of the consonant as a progressing change is that the rates of consonant deletion were observed to be peaking in the speech of young adults. Speakers' lack of awareness of the innovation may also support this interpretation, as an incoming variant may be undetected by speakers in the early stages of its spreading to the community (Labov, 1972), when not yet associated with social meaning. These two characteristics were met by the deletion of the velar fricative. However, a change in progress would not only present a high incidence of the new variant in young speakers but also a

20. My assessment of speaker awareness was impressionistic. The non-correlation of variation with speech style was in my opinion a stronger criterion on which to base my interpretation of this pattern of variation. However, an anonymous reviewer remarked that the perception of speakers could be thoroughly examined, as explored in McGowan and Babel (2020).

steady increase of the innovative variant by age group. This was not the pattern observed for this variant in Figure 3, which showed a consistently low incidence of the consonant deletion in middle-aged speakers. Another difficulty remained in how to understand the high incidence of the deleted consonant in both young and older speakers.

Previous studies have explained this curvilinear pattern other than by a stable variable, through uncovering associations between the linguistic performance of each generation and the social setting of the speech community. Examining deletion of onset glottal stop in three generations of Fiji-based Rotuman speakers, Fimone (2020) interpreted the v-shaped pattern as a progress towards glottal stop retention. Middle-aged speakers of this endangered Oceanic language displayed high incidence of deletion of this Rotuman phoneme which denotes Rotuman identity, while young and old speakers showed a tendency towards retaining it. Fimone suggests that this intergeneration variability is explained by the fluency in Rotuman of older speakers and the positive attitude towards Rotuman identity of young speakers who benefited from a revitalization program of the language, while westernised middle-aged speakers deleted this marker of Rotuman identity in their speech. Likewise, a study on Cajun English (Dubois & Horvath, 1998) has attributed the unusual v-shaped pattern displayed by intergenerational variability in the retention of Cajun English alveolar stop variants to social characteristics specific to each generation.

A premise of apparent time studies is that older speakers mostly use linguistic features that they acquired in young age and before they reached the critical period when language acquisition becomes difficult. It is therefore not anticipated that the speech of this older group would present innovative variants but rather that it would exhibit linguistic features that prevailed in their youth. There have however been instances where older speakers have adopted a fast-diffusing incoming variant, without this invalidating the premise of apparent time (Sankoff, 2019). The interpretation of the linguistic behaviour that is proposed here for the older speakers of this small Pacific society is not that of a late adoption of an incoming change. Rather it is proposed that the respected elders are the innovators, and the young group the early adopters of the deletion of the phonemic fricative. Other studies on Melanesian societies have come to a similar interpretation, whose rationale is formulated here:

> [C]ould it be the case that in some speech communities it is older individuals who drive innovation, giving them time to accumulate more variants in their linguistic portfolio, with younger speakers taking their cues from substantially older speakers rather than those just older than them. This pattern may, in turn, reflect a delayed life point when speakers feel unconstrained enough to deviate from norms. (Evans, 2018, pp. 26–27)

In her study on the Idi speech community of southern New Guinea, Schokkin (2021) reveals the effect of age as a significant factor in a phonetic variable. Variability in dropping the verb-final alveolar nasal is a reported diachronic and synchronic feature of Idi, but the elderly speakers, respected for their linguistic knowledge and performance, show the highest increase in /n/-elision. Backing Evans' comment, Schokkin contends that their position in society may be what allows these speakers to feel less constrained by linguistic norms.

If we accept the hypothesis that the men from the older generation are the innovators in dropping the velar fricative, we need to address the fact that this innovation appears in the group of young men, while bypassing the middle-aged men who mainly produce the velar fricative. I propose that this pattern may be explained, at least in part, by the high frequency of interactions between the two age groups of male speakers showing a high rate of dropping /ɣ/.

On Pentecost, every evening the traditional drink of kava is prepared by young men. It is predominantly old men who consume kava regularly. Young men and women do not drink kava, and middle-aged men do not drink it regularly. Kava-drinking sessions are therefore occasions for the older drinkers to talk among themselves, away from the rest of the community but in presence of the young men. The drinking sessions start at sunset and may carry on until late in the night, kava being prepared throughout the session by young men who mostly listen to the conversations of the old men. Other opportunities of interactions between old and young men are presented by the frequent traditional ceremonies that the high chiefs have a duty to attend. When travelling, old men are accompanied by young men who look after their accommodation and prepare their food. I have observed other interactions between young and old men that did not extend to other members of the community: an older community leader being looked after by his grandsons; and an elderly participant conferring at length with a young male relative before accepting to record a story to which their family had rights. Frequent interactions between these two groups of male speakers probably have some bearing on the forms that these men share. It is conceivable that the variants used by prestigious old speakers can influence young speakers, while bypassing the middle-aged speakers with whom older speakers have fewer interactions.

Where, then, do middle-aged men get their source of the velar fricative variant? Possibly from women, who display a consistent positive rate of using the linguistic norm throughout their life and with whom middle-aged men might interact more frequently in this period of their life. This interpretation has merit to consider both the age-grading pattern shown in Figure 3 and the hypothesised long-term trend of the consonant lenition, now reaching its final phase.

6.3 A societal innovation: Making a living from kava

The frequent loss of the velar fricative in the speech of young men may be interpreted as indicating a new way of being a Raga community member. The surge in the kava trade appears to be a societal change that may be correlated with the incoming loss of the velar fricative in Raga and documenting the phenomenon as it is happening will allow for future analysis and interpretation of this linguistic feature. Pentecost is renowned for the quality of its kava, and a high proportion of Pentecost households grow the crop.[21] The drink is nowadays consumed by people from all walks of life in Vanuatu urban centres and the powdered roots are exported to Australia, New Zealand/Aotearoa and the USA. Thus, the demand for high quality kava has increased, and so has the price of freshly uprooted kava. The young, unmarried men have seized this new opportunity of generating a good income from their plots of land without having to leave their village for Vanuatu's urban centres, or for New Zealand/Aotearoa and Australia to work as seasonal laborers.[22] It may be that this new way of making a living for rural young men will see the rise of a new identity in Raga male villagers. It remains to be seen if the loss of the velar fricative, which peaks in the speech of these young men, will acquire social meaning in correlation with this group of young rural kava growers, and whether, ultimately, the spread of this variant will be allowed to continue.

7. Conclusion

The deletion of the velar fricative is both structurally and socially conditioned. The variable was initially tested with a range of eleven independent variables, but only three factors were selected for their significant effect on this variant. The consonant is more likely to be deleted when surrounded by a high or mid-high vowel, and the young and males show a higher incidence of deletion than the rest of the speakers.

---

21. During my stay in the coastal village of Loltong, about twice a week, heavy bags of kava roots were transported down from the plateau to the shore and left on the beach, to be shipped to Port-Vila or Luganville.

22. The two countries have been hiring temporary workers from the Pacific region, mostly for seasonal work in the agriculture sector. The Recognised Seasonal Employer (RSE) scheme has been in place in New Zealand/Aotearoa since 2007 (https://www.immigration.govt.nz/about-us/research-and-statistics/research-reports/recognised-seasonal-employer-rse-scheme); its equivalent program in Australia, the Pacific Labour Mobility (PLM) program, started in 2018 (https://www.dfat.gov.au/geo/pacific/engagement/pacific-labour-mobility).

The deletion of the velar fricative does not appear to be a stylistic variant, despite displaying an age-graded pattern, mostly for male speakers. It is instead proposed that the loss of the velar fricative is spreading in the group of young and older men, due to their frequent interactions. Furthermore, on account of their high status, the prestigious older men could be seen as the innovators, and the younger group the early adopters of this variant. In their middle years, men take on different activities and roles in the community, their interactions with some groups are modified, and their source of the prevailing variant may be women, who make consistent use of this variant throughout their life.

I propose that this pattern of deletion of the consonant can be interpreted as an innovative variant embedded in an historical trend in the weakening of this reflex of the proto-Oceanic unvoiced velar stop. A follow-up study, to examine the speech of the same speakers, or to survey the speech of a population like the one surveyed in this study, would be needed to determine whether this interpretation fits the situation in Raga, and to disentangle the age level variation from the generational change. Such a study would also reveal whether the young men will adjust their use of (ɣ), when they become middle aged and do not engage in the kava ritual, and other frequent interactions with the elderly. A perception experiment might also reveal more about the salience or attitudes towards deletion. The real-time observations of such follow-up studies would establish whether the deletion of the consonant has progressed in speakers of all ages or whether the low incidence of deletion is repeated in the middle-age group, thus ascertaining which of the two interpretations, change in progress or age-graded variation, best fits this variable. If it were revealed that, once again, older men are more deletion-prone than middle aged speakers, such a study would also confirm that older speakers may indeed be closely involved in linguistic innovation in some societies.

Acknowledgements

I would like to warmly thank Jonathan Kasstan for his constant support and constructive comments and two anonymous reviewers for their valuable suggestions. I am very grateful to the Raga people of Pentecost Island for their essential contribution to this study and to the Wellsprings of Linguistic Diversity Laureate Project (Australian National University, 2014–2019) for giving me the opportunity to carry out extensive fieldwork in Vanuatu.

References

Barth, Danielle, & Evans, Nicholas (2017). SCOPIC Design and Overview. In Danielle Barth & Nicholas Evans (Eds.), *Social Cognition Parallax Interview Corpus (SCOPIC)* (pp. 1–21). Language Documentation & Conservation. Special Publication No. 12. Honolulu: University of Hawai'i Press.

Boersma, Paul, & Weenink, David (2019). Praat: doing phonetics by computer [Computer program]. Version 6.1.05. Retrieved from http://www.praat.org/

Clark, Ross (2009). **Leo tuai : a comparative lexical study of north and central Vanuatu languages*. Canberra: Pacific Linguistics.

Codrington, Robert Henry (1885). *The Melanesian languages*. Oxford: Clarendon Press.

Crowley, Terry (2002). A grammar sketch of Raga. In John Lynch, Malcolm Ross, & Terry Crowley (Eds.), *The Oceanic languages* (pp. 626–637). Abingdon: Routledge.

Dorian, Nancy (2001). Surprises in Sutherland: Linguistic variability amidst social uniformity. In Paul Newman & Martha S. Ratliff (Eds.), *Linguistic fieldwork* (pp. 133–151). Cambridge, U.K.;: Cambridge University Press. https://doi.org/10.1017/CBO9780511810206.007

Dubois, Sylvie, & Horvath, Barbara M. (1998). Let's tink about dat: Interdental fricatives in Cajun English. *Language Variation and Change*, *10*(3), 245–261. https://doi.org/10.1017/S0954394500001320

Duhamel, Marie-France (2020). Variation in Raga – a quantitative and qualitative study of the language of North Pentecost, Vanuatu. Unpublished doctoral dissertation, Australian National University. Retrieved September 28, 2021, from http://hdl.handle.net/1885/205516

ELAN [Computer software] (2021) (Version 6.2). Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved September 28, 2021, from https://archive.mpi.nl/tla/elan

Evans, Nicholas (2018). The dynamics of language diversity. In R. Mesthrie & D. Bradley (Eds.), *The dynamics of language: Plenary and focus lectures from the 20th International Congress of Linguists*. Cape Town, South Africa: Juta and Company (Pty) Ltd.

Fimone, Wilfred (2020). Onset glottal stop deletion in Suva Rotuman. *Asia-Pacific Language Variation*, *6*(2), 196–221. https://doi.org/10.1075/aplv.19016.fim

François, Alexandre (2012). The dynamics of linguistic diversity: egalitarian multilingualism and power imbalance among northern Vanuatu languages. *Int'l. J. Soc. Lang*, *214*, 85–110. https://doi.org/10.1515/ijsl-2012-0022

François, Alexandre, Franjieh, Michael, Lacrampe, Sébastien, & Schnell, Stefan (2015). The exceptional linguistic density of Vanuatu. In Alexandre François, Michael Franjieh, Sébastien Lacrampe, & Stefan Schnell (Eds.), *The Languages of Vanuatu: Unity and diversity* (pp. 1–21). Canberra: Asia-Pacific Linguistics.

Haudricourt, André-Georges (1961). Richesse en phonèmes et richesse en locuteurs. *L'Homme*, *1*, 5:10. https://doi.org/10.3406/hom.1961.366337

Ivens, Walter George (1938). A grammar of the language of Lamalanga, North Raga, New Hebrides. *Bulletin of the School of Oriental Studies, University of London*, *9*(3), 31. Retrieved September 28, 2021, from http://www.jstor.org/stable/608236. https://doi.org/10.1017/S0041977X00078514

James, Deborah (1996). Women, men and prestige speech forms: a critical review. In Victoria L. Bergvall, Janet M. Bing, Alice F. Freed (Eds.), *Rethinking language and gender research: Theory and practice* (pp. 98–125). London: Routledge

Labov, William (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
Labov, William (2001). *Principles of linguistic change. Volume 2: Social factors*. Oxford, UK: Blackwell.
Lynch, John, & Crowley, Terry (2001). *Languages of Vanuatu : A new survey and bibliography*. Canberra: Pacific Linguistics.
Mansfield, John Basil (2015). Consonant lenition as a sociophonetic variable in Murrinh Patha (Australia). *Language Variation and Change*, *27*(2), 203–225. https://doi.org/10.1017/S0954394515000046
McGowan, Kevin, & Babel, Anna (2020). Perceiving isn't believing: divergence in levels of sociolinguistic awareness. *Language in Society*, *49*(2), 231–256. https://doi.org/10.1017/S0047404519000782
Meyerhoff, Miriam (1996). Dealing with gender identity as a sociolinguistic variable. In Victoria L. Bergvall, Janet M. Bing, Alice F. Freed (Eds.), *Rethinking language and gender research: Theory and practice* (pp. 202–227). London: Routledge
Meyerhoff, Miriam, Abtahian, Maya Ravindranath, Gafter, Roey J., Horesh, Uri, Kasstan, Jonathan R., Keegan, Peter, & King, Jeanette (2020). Styles, standards and meaning: Issues in the globalisation of sociolinguistics. *Language Ecology*, *4*(1), 1–16. https://doi.org/10.1075/le.00006.mey
Milroy, Lesley (1980). *Language and social networks*. Oxford: Blackwell.
R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved September 28, 2021, from https://www.r-project.org/
Republic of Vanuatu (1980). *Constitution of the Republic of Vanuatu*. Port-Vila. Retrieved September 28, 2021, from http://www.paclii.org/vu/legis/consol_act/cotrov406/
San Roque, Lila, Gawne, Lauren, Hoenigman, Darja, Miller, Julia C., Rumsey, Alan, Spronck, Stef, … Evans, Nicholas (2012). Getting the story straight: Language fieldwork using a narrative problem-solving task. *Language Documentation and Conservation*, *6*, 135–174.
Sankoff, Gillian (2019). Language change across the lifespan: Three trajectory types. *Language*, *95*(2), 197–229. https://doi.org/10.1353/lan.2019.0029
Schokkin, Dineke (2021). Variable realisation of verb-final /n/ in Idi. *Asia-Pacific Language Variation*, *7*(1), 1–29. https://doi.org/10.1075/aplv.19011.sch
Stanford, James N. (2009). Clan as a sociolinguistic variable. In James N. Stanford & Dennis R. Preston (Eds.), *Variation in indigenous minority languages* (pp. 463–484). Amsterdam: John Benjamins. https://doi.org/10.1075/impact.25.23sta
Stanford, James N. (2016). A call for more diverse sources of data: Variationist approaches in non-English contexts. *Journal of Sociolinguistics*, *20*(4), 525–541. https://doi.org/10.1111/josl.12190
Taylor, John P. (2008). *The other side: Ways of being and place in Vanuatu*. Manoa: University of Hawai'i Press. https://doi.org/10.21313/hawaii/9780824833022.001.0001
Trudgill, Peter (1974). *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.
Tryon, Darrell T. (1976). *New Hebrides languages : An internal classification*. Canberra: Pacific Linguistics.
Ünal-Logacev, Özlem, Żygis, Marzena, & Fuchs, Susanne (2017). Phonetics and phonology of soft 'g' in Turkish. *Journal of the International Phonetic Association*, *49*(2), 183–206. https://doi.org/10.1017/S0025100317000317

Vanuatu National Statistics Office – Bureau National des Statistiques (Producer). (2016). map: A1A. Household growing selected vegetables and cash crops – Proportion of households growing kava – 2016 Mini Census. Retrieved September 28, 2021, from http://vanuatu.popgis.spc.int

Vanuatu National Statistics Office – Bureau National des Statistiques (2020, 24 June 2020). 2016 Mini Census. Retrieved September 28, 2021, from https://vnso.gov.vu/index.php/en/census-and-surveys/census/2006-mini-census

Vanuatu National Statistics Office – Bureau National des Statistiques (2021). *Well-Being in Vanuatu. 2019–2020 NSDP Baseline Survey*. Retrieved September 28, 2021, from https://vnso.gov.vu/index.php/en/special-reports/well-being-survey

Vari-Bogiri, Hannah (2011). Phonology and morpho-syntax of Raga, Vanuatu. Unpublished doctoral dissertation, The University of the South Pacific, Port-Vila.

von Prince, Kilu, Krajinović, Ana, Krifka, Manfred, Guérin, Valérie, & Franjieh, Michael (2018). Mapping irreality: Storyboards for eliciting TAM contexts. *Proceedings of Linguistic Evidence 2018*.

Wagner, Suzanne (2012). Real-time evidence for age grad(ing) in late adolescence. *Language Variation and Change*, *24*(2), 179–202. https://doi.org/10.1017/S0954394512000099

Wagner, Suzanne, & Sankoff, Gillian (2011). Age grading in the Montréal French inflected future. *Language Variation and Change*, *23*(3), 275–313. https://doi.org/10.1017/S0954394511000111

Walker, James A. (2010). *Variation in linguistic systems*. New York: Routledge.

Walsh, David Scott (1982). The restricted distribution of Raga /k/: starting point for a subgrouping hypothesis. In Rainer Carle, Martina Heinschke, Peter W. Pink, Christel Rost, & Karen Stadltander (Eds.), *GAVA: studies in Austronesian languages and cultures. Studien zu austronesischen Sprachen und Kulturen*. Berlin: Dietrich Reimer Verlag.

Walsh, David Scott (1995). Raga. In Darrell Tryon (Ed.), *Comparative Austronesian dictionary: An introduction to Austronesian studies* (Vol. *1* Part *1*, pp. 807–818). Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110884012.2.807

Walsh, David Scott (2005). Some misconceptions remedied for Raga, an Oceanic language. *Journal of the Polynesian Society*, *114*(2), 157–163.

Wurm, Stephen A. (2007). Threatened languages in the Western Pacific area from Taiwan to, and including, Papua New Guinea. In M. Brenzinger & EBSCOhost. (Eds.), *Language diversity endangered* (pp. 374–390). Berlin ; New York: Mouton de Gruyter.

Yoshioka, Masanori (1987). The story of Raga, a man's ethnography on his own society. Retrieved September 28, 2021, from http://web.cla.kobe-u.ac.jp/staff/yoshioka/english-ragastory.htm

Appendix A. Distribution of tokens and % of the variants by surrounding segments

Vowel, or pause, preceding the variable

	[ɣ] retained		[ɣ] dropped		Total
Pause	261	87.6%	37	12.4%	298
a	338	91.4%	32	8.6%	370
o	214	80.1%	53	19.9%	267
i	152	75.6%	49	24.4%	201
u	48	75.0%	16	25.0%	64
e	54	84.4%	10	15.6%	64
Total	**1067**		**197**		**1264**

Vowel following the variable

	[ɣ] retained		[ɣ] dropped		Total
a	577	87.2%	85	12.8%	662
o	195	87.4%	28	12.6%	223
i	123	79.4%	32	20.6%	155
u	65	81.3%	15	18.8%	80
e	107	74.3%	37	25.7%	144
Total	**1067**		**197**		**1264**

Appendix B. Speakers' average incidence of deletion, sorted from high to low incidence

	Deletion rate	Sex	Age	L1	Moiety	Mobility	Education	Village
1	1.834	F	18	Raga	Bule	high	secondary	Loltong
2	1.589	M	18	Raga	Bule	low	secondary	Loltong
3	1.337	M	29	Raga	Tabi	low	primary	Loltong
4	1.331	M	80	Raga	Bule	low	secondary	Abwatvenue
5	1.309	M	21	Raga	Bule	low	secondary	Loltong
6	1.282	M	77	Raga	Bule	unknown	unknown	Atavtabanga
7	1.019	M	84	Raga	Tabi	high	tertiary/ training	Laone
8	0.940	M	18	Raga	Tabi	low	secondary	Agatoa
9	0.744	M	20	Raga	Bule	low	secondary	Loltong
10	0.596	F	34	Raga	Bule	low	secondary	Loltong

	Deletion rate	Sex	Age	L1	Moiety	Mobility	Education	Village
11	0.497	M	36	Raga	Tabi	low	secondary	Lagatava
12	0.497	M	21	Raga	Bule	high	secondary	Loltong
13	0.472	M	89	Raga	Bule	low	primary	Lamoru
14	0.465	F	29	Raga	Bule	high	tertiary/ training	Loltong
15	0.445	F	30	Raga	Tabi	low	primary	Namaram
16	0.233	F	20	Raga	Tabi	low	primary	Loltong
17	0.192	M	64	Raga	Bule	unknown	unknown	Biribiri
18	0.157	F	76	Raga	Bule	high	tertiary/ training	Amatbobo
19	0.105	F	20	Bislama	Tabi	high	secondary	Luganville (Santo)
20	0.080	F	63	Raga	Tabi	unknown	unknown	unknown
21	0.069	F	37	Raga	Tabi	low	secondary	Abwatuntora
22	0.069	F	18	Raga	Bule	low	secondary	Atabulu
23	0.069	M	15	Raga	Tabi	low	primary	Loltong
24	0.061	F	38	Raga	Bule	low	primary	Atavtabanga
25	−0.047	M	82	Raga	Tabi	unknown	unknown	Biribiri
26	−0.047	F	18	Raga	Bule	low	secondary	Loltong
27	−0.089	M	33	Raga	Bule	low	primary	Abwatuntora
28	−0.099	F	43	Apma	Tabi	high	tertiary/ training	Melsisi
29	−0.138	F	45	Raga	Bule	unknown	unknown	Biribiri
30	−0.171	M	68	Raga	Tabi	low	none	Abwatunbuliva
31	−0.233	F	50	Raga	Tabi	low	primary	Loltong
32	−0.236	M	11	Raga	Bule	low	primary	Loltong
33	−0.279	M	46	Raga	Tabi	low	secondary	Santo
34	−0.289	F	41	Raga	Bule	high	secondary	Loltong
35	−0.318	M	52	Raga	Tabi	unknown	primary	Loltong
36	−0.383	F	24	Raga	Tabi	high	tertiary/ training	Loltong
37	−0.419	F	62	Raga	Bule	low	primary	Labultamata
38	−0.475	M	80	Raga	Bule	high	unknown	Amatbobo
39	−0.475	M	61	Raga	Bule	high	primary	Laone
40	−0.476	M	52	Raga	Bule	low	primary	Lolkasai
41	−0.496	F	25	Raga	Bule	low	secondary	Abwatuntora
42	−0.537	F	16	Raga	Bule	low	secondary	Lolkasai
43	−0.568	F	30	Raga	Tabi	low	secondary	Loltong
44	−0.576	M	65	Raga	Tabi	unknown	unknown	Lagatava

	Deletion rate	Sex	Age	L1	Moiety	Mobility	Education	Village
45	−0.608	M	43	Raga	Tabi	low	primary	Loltong
46	−0.647	F	15	Raga	Tabi	low	secondary	Loltong
47	−0.648	M	42	Raga	Bule	low	secondary	Loltong
48	−0.769	F	49	Raga	Bule	low	unknown	Loltong
49	−0.796	F	36	Raga	Tabi	low	primary	Loltong
50	−0.796	M	60	Raga	Tabi	high	tertiary/ training	Lolvatu
51	−0.872	F	46	Raga	Tabi	low	primary	Loltong
52	−0.872	M	49	Raga	Tabi	low	primary	Loltong

Résumé

Cette étude variationniste d'une variable phonétique dans une communauté linguistique au Vanuatu révèle que le discours des hommes jeunes présente une forte tendance à supprimer la fricative vélaire. Cette tendance est également présente chez les hommes les plus âgés. À l'inverse, la suppression de la fricative vélaire est rare dans le discours des femmes et celui des hommes entre 40 et 65 ans. Ces résultats ne correspondent pas aux prédictions classiques associées au modèle sociolinguistique du variationnisme et la distribution de la suppression de la fricative vélaire, par âge et sexe des locuteurs, nécessite donc de considérer les interactions et le statut social des membres de cette communauté où les hommes plus âgés occupent les postes d'autorité. Le style du discours ne semble pas influencer le taux de suppression de ce phonème, alors que des analyses multivariées révèlent l'impact de la voyelle précédente et suivante, ainsi que celui de l'âge et du sexe du locuteur, sur la non-réalisation de la consonne. La suppression de ce phonème est interprétée comme la phase finale d'un processus diachronique de perte de matériel phonétique. La forte incidence dans le discours des hommes jeunes de la suppression de ce phonème s'explique par leurs fréquentes interactions avec les hommes les plus âgés et les plus hauts placés dans la hiérarchie de cette communauté.

Address for correspondence

Marie-France Duhamel
Australian National University
49 Church Bay Road
Oneroa, Auckland 1081
New Zealand
mfduhamel@gmail.com

Publication history

Date received: 9 November 2021
Date accepted: 12 August 2022

Authenticity in language ideology

Social variation in Chanka Quechua

Natalie Povilonis and Gregory Guy
Indiana University Bloomington | New York University

Like many marginalized languages, Chanka Quechua (Peru) lacks community-wide prestige norms associated with standard-language ideology. Formal situations require Spanish, and few speakers are literate in Quechua, so normative speech styles are absent. Speakers' evaluative judgments do not reference notions of correctness; rather, they value *puro* 'pure' speech and authenticity.

This paper explores alternative approaches to accessing sociolinguistic judgments with a study of the variably present uvular phoneme in the past tense /–rqa/ morpheme, as exemplified in the following alternation:

(1) *ri-rqa-ni* ~ *ri-ra-ni*
go-PST-1SG go-PST-1SG
'I went' 'I went'

To contrast speech from sociolinguistic interviews, careful, self-monitored speech is elicited through oral retelling of material presented aurally, rather than in writing. Of 38 participants, rural speakers tend to have higher rates of /q/ than urbanites and reflects idealized *puro* Quechua. We argue that authenticity guides variation, in place of standard-language ideology.

Keywords: sociolinguistic authenticity, variation, Spanish, Quechua

1. Introduction

Sociolinguistic research has long identified shared social norms guiding language use (Labov, 1972). In majority languages with a broadly disseminated writing system and language authority, these top-down norms often come from orthography, linguistic elites, or other community-wide prestige norms in standard language ideology (SLI). However, in languages like Quechua, which do not have these top-down impetuses, the source of the norms that guide variation is less clear. This

https://doi.org/10.1075/aplv.22004.pov
Asia-Pacific Language Variation 8:2 (2022), pp. 240–273 | ISSN 2215-1354 | E-ISSN 2215-1362

paper takes steps toward identifying common ground for the speech forms preferred by Quechua speakers, and finds that discourses around authenticity are a major contributing force.

Quechua is a diverse language family spoken natively by around eight million people originally from countries around the Andes mountain range, including Peru, Ecuador, Bolivia, Colombia, and Chile (Chirinos Rivera, 1998; Chirinos Rivera, 2001). In Peru, the total population of Quechua speakers is estimated at about 3.7 million according to the 2017 census (Instituto Nacional de Estadística e Informática, 2017, p. 48), which is approximately 14% of the total population of the country (32 million). These numbers are likely low estimates, as the census allowed for a single language response (Spanish OR Quechua). Many bilingual interview participants reported having answered Spanish due to widespread social acceptance, including by the census recorder – who asked questions in Spanish and sometimes marked Spanish without consulting the interviewee.

Quechua dialects are generally considered by linguists to be as typologically distinct as Romance or Slavic languages.[1] In Peru, the Ministry of Education has developed a standardized orthography and written norms for the country's six major Quechua dialects, and in recent years it has invested significant effort in distributing printed Quechua materials around the country. Chanka Quechua, the dialect of focus for this study, is spoken by just under one million people in the departments of Ayacucho, Huancavelica, and western Apurímac in southern Peru (Lewis, Simons, & Fennig, 2016). Data for this study was gathered in Andahuaylas, on the Eastern side of the Chanka dialect region, approaching the Cuzco-Collao region.

Despite the government's literacy efforts, Quechua is still widely an oral language. Only since the early 2000s has education in Quechua (and thus, in Quechua reading and writing) become more common – and even then, discrepancies in access between urban and rural residents, and between men and women, mean that a large majority of the adult Quechua-speaking population still has low literacy and little knowledge of written norms, especially in Quechua (INEI, 2017). In fact, many of the rural women interviewed during this study did not know how to write: they gave oral consent and "signed" payment receipts with a fingerprint.

Moreover, Quechua has limited formal social domains of use. In Peru, formal environments historically require Spanish, such that Quechua is only spoken in informal domains such as the home. Speaking Quechua in public, even if both interlocutors are bilingual, risks implying that the speakers do not know

1. Some effort has been made to unite Ayacucho-Chanka and Cuzco-Collao dialects under a single Southern Quechua orthography, but they are phonologically distinct.

Spanish – a marker of inferior class or status (Zavala, 2014). Thus, the comparison between formal and informal language settings, often used to describe different registers in sociolinguistics, has no clear application for Quechua.

Sociolinguistic norms are often specific to a speech community (Morgan, 2004). This may be even more relevant for Quechua, as speakers within the same dialect region are separated from other communities by many hours on foot or long bus rides. In fact, even if Quechua is used publicly in the local community, little engagement in Quechua occurs outside of the immediate geographic area. When speakers visit another town or even another village, they use Spanish due to its prestige as the public language. Thus, they have little reference for how Quechua is spoken elsewhere, and a lower chance of following norms from outside of their own community. Certainly, they do not orient themselves toward the regional standard proposed by the Ministry of Education.

This paper reports results from a study of a sociolinguistic variable in Chanka Quechua: uvular alternation in verbal affixes. The data in focus was elicited through a novel oral sentence correction task, intended to increase participants' attention paid to speech. It will be briefly compared with data from more traditional sociolinguistic interview-type conversations. Given the social situation of Quechua speakers, they have not developed a standard language ideology as traditionally imagined (such as from a widespread written standard, the media, or high-status speakers). Instead, authenticity was found to be a potential alternative ideology that guides language variation. This ideology and the speech forms associated with it seem rooted in national political and language maintenance discourses that have developed over the past century. The following section of the paper will provide background on the uvular alternation of interest (2.1) and discourses around authentic Quechua language (2.2). Section 3 describes the participant makeup and methodology of the study. The results in Section 4 show variability in uvular presence for both production and perception, though ultimately the uvular-full variant is still preferred. In Section 5, we discuss the limitations of the attention-to-speech model for accessing stylistic variation in Quechua and other endangered languages, as well as evaluate our attempts to adapt it. Much of the discussion revolves around the apparent lack of a widespread standard language ideology for Quechua. However, we argue that a search for authenticity may be a growing force guiding variation, in place of written normative pressures.

2. Background

2.1 Uvular alternation in Chanka Quechua

The uvular phoneme /q/ is variably present in the experienced past morpheme /-rqa/. Chanka has a single uvular phoneme,[2] which has been described as being in free variation with zero in this morpheme and two others (exhortative /–rqu/ and 1st-to-2nd person future /–sqa/). Besides these morpheme-specific alternations and some other cases of phonological deletion, uvular presence does not vary (Cerrón-Palomino, 1987; Parker, 1969; Soto Ruiz, 1976).

The /-rqa/ morpheme is used for relating past facts or events that the speaker directly witnessed or performed (Quintero Bendezú, 1996; Zariquiey & Córdova, 2008). In Quechua's agglutinative verbal template, /–rqa/ appears near the end, after aspectual morphemes and before person and tense morphemes. Examples (1) and (2) show a verb in experienced past tense (Quintero Bendezú, 1996, p. 57).

(1) *puklla-**rqa**-ni*
play-PST-1S
'I played'

(2) *puklla-**rqa**-nki*
play-PST-2S
'you played'

This morpheme can appear word-finally in 3rd person singular verbs, as the 3rd person singular marker *–n* in both experienced and reportative past tenses is not obligatory in Chanka. In fact, it is generally not used, and a verb without it "is more natural" (Zariquiey & Córdova, 2008, p. 164). This contrasts with the Cuzco variety, which does usually add the *-n* person morpheme (*muna-**rqa*** vs. *muna-**rqa**-n* 'he/she wanted') (Parker, 1969; Zariquiey & Córdova, 2008). In Example (3), no person suffix appears (Quintero Bendezú, 1996, p. 57).

(3) *puklla-**rqa**-∅* (*puklla-**rqa**-n* in Cuzco)
play-PST-3S
'he/she played'

The variable presence of the uvular in the experienced past is evident in Chanka and Cuzco Quechua (Adelaar with Muysken, 2004; Cerrón-Palomino, 1987;

2. The Chanka uvular phoneme is generally realized as a fricative /χ/. This paper utilizes the <q> grapheme from the standardized orthography to allow for easier comparison with other Quechua dialects like Cuzco-Collao, where the uvular phoneme is a stop.

Parker, 1969). This differs from many central and northern Quechua dialects, where the past marker /–rqa/ has completely lost the uvular. While the latter dialects are typologically quite distinct from Chanka/Cuzco, the uvular has also been deleted in Argentinian dialects southeast of, but more closely related to, Chanka[3] (Cerrón-Palomino, 1987). To illustrate this variation in the past tense morpheme, alternate pronunciations of the previous Examples (1)–(3) are as follows in (4)–(6):

(4) *puklla-**ra**-ni*
play-PST-1S
'I played'

(5) puklla-***ra**-nki*
play-PST-2S
'you played'

(6) *puklla-**ra**-∅* (*puklla-**ra**-n* in Cuzco)
play-PST-3S
'he/she played'

Besides a brief mention in some descriptions (Cerrón-Palomino, 1987; Parker, 1969; Soto Ruiz, 1976), only one known quantitative study on uvular alternation has been done for Quechua. Povilonis (2016) found near categorical absence of the uvular in the exhortative morpheme (1.4% presence, $N = 142$), variable presence of the uvular in the experienced past morpheme (87.3% presence, $N = 167$), and categorical presence of the uvular in all other morphemes ($N = 425$). The data extracted for that study were taken from 15 hours of radio speech from two announcers on the Chanka news program *Llactamanta, llactapaq* broadcasted via *Radio Titanka* in Andahuaylas, following a methodology similar to Kroch and Small (1978) to identify formal speech through the announcers and casual speech through guests and callers. The results showed higher rates of uvular presence in the speech of the announcers hosting the show, than in the usage of guest speakers. The quantitative analysis also identified a linguistic factor that significantly affected uvular presence: the number of following morphemes, with higher rates of uvular presence when there were no following morphemes.[4]

While this study offered some initial insight into uvular variation patterns in Chanka Quechua, token counts were quite small, and questions arose as to

3. In these dialects, evidence of the uvular remains: adjacent high vowels still show lowering (i.e., *rerani* appears instead of *rirani* or *rerqani* 'I went'), a process that usually only occurs in the phonetic context of a uvular consonant (Cerrón-Palomino, 1987).

4. This study was unable to resolve potential interactions between number of following morphemes and other correlated factors: stress position and number of syllables.

whether the announcers were using a special radio style different from their casual style, or whether speakers with other social characteristics would show the same patterns. As will be described in the following sections, the quantitative study of uvular alternation described below expands on Povilonis (2016) to include more speakers of a range of social characteristics, casual and conscious speech styles, and higher token counts.

2.2 Authenticity in discourse

In the 1920s, an intellectual discourse called *indigenismo* reaffirmed the Quechua language as intrinsically tied to authentic Indigeneity. Rather than the "problem" that it was historically conceptualized to be, Quechua was reframed as a celebrated representative component of a renewed national Indigenous identity. Peruvian identity became rooted in a diverse history, and Indigenous people were the link to it. Since this diversity was conceived as categorical rather than on a continuum, Quechua became fixed as part of an authentic Indigenous identity. Conversely, Spanish became part of the *mestizo* 'mixed' identity. While *indigenismo* gave Quechua intrinsic value, it did not provide a place for Quechua or Quechua speakers to participate in modern society (Zavala, Mujica, Córdova, & Ardito, 2014).

Quechua has many overtly positive associations, which were frequently mentioned in the interviews for this study. In addition to being revered as an ancestral language, Quechua is also associated with "an idealized communal life in a rural context" (Zavala et al., 2014, p. 39), essentially a "lost paradise" (Zavala et al., 2014, p. 39) marked by "rural purity"[5] (Zavala et al., 2014, p. 38). In interviews, Pedro (42-year-old urban male) located broadly valued reciprocity relationships and feelings in the countryside:

> *Porque kanraqmi chay sentimiento, corazón, prójimo nisqanchikta, runa masinchikta yanapaykunaypaq. Pero huklawpiqa manañam kanñachu. Kunan másta kaypi kanraq aslla. Riruyraq campun lawman; campun lawman másta tarinki.*
> Because there is still that feeling, heart, to help our neighbor, our neighbor. But elsewhere, that does not exist anymore. Now here [in the city], it is just a little. You ought to go to the countryside; in the countryside, you'll find more.
> (Pedro, 42 years, urban male)

The "living culture" is believed to still exist "higher up" or "farther inside" (Babb, 2020, p. 13) and must be preserved (Zavala et al., 2014, p. 40). The geographic references reflect the fact that urban areas are generally at lower altitudes, with easier

5. All translations are my own.

access via roads. In some ways, reciprocity and solidarity are part of this culture that has been transferred to the urban realm via the Quechua language. Leocadía (58-year-old urban female) described how Quechua is hidden until two Quechua speakers come together, whether or not they know each other:

> *A veces riqsisqaykuwanyá musyarquniku quechua simi rimasqanta, hinaspaqa kacharinikuyá.*
> Sometimes we recognize those who speak Quechua, and we let it out.
> (Leocadía, 58 years, urban female)

To fit into the Indigenous identity mold, Quechua, like other indigenous languages around the world, is surrounded by "notions of authenticity – of who counts as an 'authentic' Indigenous person or representative." Even though "authenticity in itself does not naturally exist" (Zavala et al., 2014, p. 38), "traditional clothing, language use and other culture practices" from the past "become the only 'legitimate' forms of Indigeneity" (Patrick, 2007, p. 119). Importantly, "speakers can be locked into fixed or essentialized notions of identity, 'authenticity', and place" (Patrick, 2007, p. 127). Essentialization is a process in which cultural or biological attributes of social groups are assumed to be inherently associated. This supposes that each group can be clearly defined, and assumes homogenization within the group (Bucholtz, 2003). The essentialized representation of Quechua delegitimizes anyone who does not fit into the mold, and it erases urban Quechua use from the realm of possible ways of speaking (Irvine & Gal, 2000). This disempowers many Quechua speakers (Zavala et al., 2014).

As the Quechua language is a key component of the Indigenous identity, speakers also seek to define authentic Quechua. An authentic speaker is idealized as a monolingual who speaks *quechua neto* 'pure Quechua' (Zavala et al., 2014, p. 13) that is not "contaminated" by Spanish, as one of acquaintance in Andahuaylas mentioned. Despite the fact that "new forms of language are constantly developing" (Patrick, 2007, p. 125), especially in in indigenous languages in contact, "mixed language forms … are often considered illegitimate" and "lack authenticity" (Patrick, 2007, p. 125). This widely documented goal of linguistic purism in language revitalization (see Dorian, 1994) also exists for Quechua. Laurimar (36-year-old urban female) reflected the sentiment of several other urban speakers when she described how their Quechua is less well-spoken because it includes aspects of Spanish:

> *Manañam kunanqa quechua quechua, sapachallan quechuataqa rimanchikñachu. Aswanqa rimanchikña castellano; chapunchikñam castellano simiwan. Campopi allinta rimanmanku quechuataqa,¿riki? Kunan kaypi ñuqaykuqa manañan chay quechua puro nisqantaqa rimanchikñachu, castellanuwan chapuspañan.*

> We no longer speak just Quechua. We often speak Spanish; we mix with Spanish. In the countryside they would speak Quechua well, no? Now here [in the city] we no longer speak that pure Quechua anymore, "mixing" with Spanish.
>
> (Laurimar, 36 years, urban female)

Geographically, as Laurimar suggested, authentic Quechua is located in the *campo* 'countryside', as declared by almost all urban interview participants. There, the Quechua is *sumaq* 'beautiful', *chuya* 'clean', *puro* 'pure (Spanish)', *allin* 'good', *miski* 'sweet'. Conversely, urban Quechua is *contaminado* 'contaminated (Spanish)' or *qayma* 'tasteless'. Temporally, authentic Quechua existed more in the past, by older generations. Several urban participants claimed that their parents spoke it, and that it used to be spoken more *ñawpaq* 'long ago'. In these conversations, discussions around *ñawpaq* evoke a nostalgia of more distant past events, especially family and community history.

Linguistically, thus, authentic Quechua is an idealized way of speaking rooted in the past. Monolithic and static, it is believed to no longer exist in modern urban society; it must be recovered and preserved. Much of this search for authentic Quechua occurred during development of the standardized orthography and grammar in the final decades of the twentieth century (Fishman, 1988; Hornberger, 1993), and now the norms from the Ministry of Education influence what educated speakers consider authentic speech. Occasionally, these speakers also reference notions of correctness related to these norms, but this is most often used for teacher certification exams or in a classroom, rather than for policing colloquial or spoken language.

Positive ideologies idealizing Quechua were most commonly expressed by urban speakers. All of the urbanites in the study (except for two over 70 years old) were fluent (native or almost native) Spanish speakers, educated professionals, and economically secure – and importantly, they and their families conformed in daily life to the *mestizo* social norms of the city. In local terms, they had sufficiently *superado* 'overcome' the Indigenous identity, even if they could also speak Quechua.

In these urban *mestizo* families, Quechua has become an accessory that does not threaten their social status. These are the people who were, according to Zavala et al. (2014), less embarrassed to speak Quechua and show that identity. They participate in Quechua workshops that are carried out in Quechua, and they demand that politicians speak Quechua in order to win elections (Zavala et al., 2014). At the same time, language shift has occurred in their own families: many participants described their children as not wanting to learn Quechua when young, indicating pervasive negative ideologies in the recent past. Now that more money is involved, and universities are requiring Quechua, many of those young

adults are asking their parents to teach them, as Mario (50-year-old urban male) explained.

> *Ñuqaykutaqa niwanku, "Papá, actualízate," niwanku… Chay runasimitapas, manam warmaykuqa rimayta munanñachu… Kunan, chayllaraq chay ñakakuypi rikurichkanku*,
> They [our children] used to tell us, "Dad, update yourself," they told me… Quechua, our children did not want to speak it anymore… Now, just recently, they are waking up to some regret. (Mario, 50 years, urban male)

Similar to languages like Welsh and Gaelic in recent years (O'Hanlon, 2015), the increasing economic value of Quechua is motivating young people to learn their family language. At the same time, its economic value still depends on recovering an essentialized or romanticized form of Quechua, which is only one element of an authentic Indigenous identity. In many cases, speakers who no longer fit into the authentic identity on a daily basis, usually those who have migrated to the city, can return to their rural communities for festivals and "perform" the identity (Zavala et al., 2014, p. 40). This seems to be the type of diversity that is celebrated in the country. Once one can follow social norms, especially speaking Spanish in mainstream society, then they benefit from adopting the Indigenous identity in appropriate cultural spaces. Full adoption of the Indigenous identity comes through the use of authentic Quechua language.

Authenticity thus appears as an alternative type of standard language ideology (SLI). Urban speakers, also bilingual in Spanish and educated in Quechua grammar (as described by the Ministry of Education), cite authenticity (via purity, cleanliness, sweetness) as a main characterization of rural speech. Rural speech is also described as having more morphemes and phonological segments, while the contrary, urban speech is *qayma* 'tasteless'. A speaker who wants to sound more authentic may then incorporate more rural features. This inverts the usual status of rural people, who, aside from language variety, are typically poorer, less-educated and lower in social status than their urban counterparts.

3. Methodology

To determine the extent to which an ideology of authenticity may influence spoken Chanka Quechua, we analyzed recordings of a sentence repetition task from 38 Chanka Quechua-speaking participants in rural and urban Andahuaylas, Peru, in 2019. The dependent variable was the uvular phoneme in the experienced past morpheme. Results from the controlled task were compared with rates of uvu-

lar presence in naturalistic speech from a subset of 32 participants. Section 3.1 describes the participant sample, and Section 3.2 describes the tasks performed.

3.1 Participants

The participant sample was gathered to account for two locations of residence (18 rural, 20 urban) and both sexes (20 women, 18 men), as well as a broad age distribution where possible. The social distribution of the participant sample is summarized in Table 1.

Table 1. Final participant sample

Spanish influence	Community	Age	Male	Female	Total
less	Rural Andahuaylas	Elder (54–75)	2	4	18
		Middle (30–40)	2	5	
		Young (18–29)	3	2	
more	Urban Andahuaylas	Elder (54–74)	6	3	20
		Middle (36–53)	5	6	
		Young (18–29)	0	0	

All rural participants were born in the rural community or one nearby and had spent most of their lives there, except for two of the young males (18 and 20 years old), who had been spending weeknights in a rented room in the urban valley for six years in order to attend school. Rural females except for the youngest one (18 years old) had low Spanish proficiency and less than primary education. For rural males, Spanish proficiency and education levels correlated inversely with age, such that the oldest men had low proficiency and less than primary education, while the youngest men had high L2 proficiency and at least secondary education.

All but four urban participants were born and raised in the urban valley. Two males had come from a rural community before the end of the critical period (at ages 8 and 10), and three females and one male had arrived after secondary school. Another urban female had also spent several years of primary school living with grandparents in a rural community. Urban speakers ranged from 36 to 74 years of age. All except the two oldest speakers and the four late arrivals were simultaneous bilinguals, and everyone except the two oldest had secondary education or higher. Unfortunately, younger speakers of Quechua who had spent most or all of their lives in the city were difficult to find, and the few possible qualifiers claimed that they did not speak Quechua well enough. This diffi-

culty in gathering an age-balanced sample of speakers reflects a logistical problem commonly encountered by researchers in endangered language speech communities in contact with a majority language. Stanford and Preston (2009) and authors within the same volume observed greater influence from the majority language for younger speakers. Long-term, this also results in language shift among younger speakers first, leading to fewer young endangered language speakers in the community (Dorian, 1994, p. 481). Certainly, future research in Andahuaylas will seek to instill confidence in younger urban speakers and encourage their participation. Much as Nagy (2009, p. 401) found for Faetar, younger speakers choose not to speak Quechua in the presence of their elders, given claims that they do not speak pure Quechua.

Spanish proficiency was defined on five levels, summarized in Table 2 below, and coded based on a combination of the participant's self-description of their proficiency plus the research assistant and the first author's assessments following interactions with the participant (we initiated conversation in Spanish following each interview). A Spanish proficiency of 1 was coded for participants with almost no productive or receptive knowledge. Participants with a Spanish proficiency of 2 have minimal productive knowledge, but can understand more. Participants with a proficiency of 3 can communicate ideas as necessary and understand most everyday conversation but still do not use Spanish in their day-to-day tasks. Participants with a Spanish level of 4 have non-native fluency, in that they can understand all conversation as well as communicate in Spanish with relative ease despite clear Quechua interference; they also have some Spanish literacy. Finally, participants with a Spanish proficiency of 5 have native proficiency (having acquired Spanish either from birth or early childhood), and can fully communicate and express ideas; they all also have Spanish literacy.

Table 2. Spanish proficiency levels of participants

Level	Description
1	almost none
2	no production, some comprehension
3	some production, good comprehension
4	non-native fluency
5	native fluency

Chanka speakers' social participation is generally local, within their own physical rural or urban community. Land travel over winding Andean dirt roads prohibits regular interaction between rural and urban communities: a bus ride

between Andahuaylas and Lima lasts over 16 hours, and the rural community in this study is over two hours by foot from Andahuaylas city. In urban Andahuaylas, bilinguals have proficiency in Spanish and do not use Quechua in most domains. At the same time, they have high local status and value Chanka for its in-group associations and historical prestige of authenticity. Rural speakers, with low Spanish proficiency, are the most likely to use Chanka in daily life, but they may not attribute it much prestige since it hinders their social progress. These differing ideologies about Chanka between speech communities likely lead to variation in use of linguistic forms, at least in conscious speech.

3.2 Data elicitation task

Recordings were made of an oral sentence correction task intended to approximate the effects of reading and text correction tasks (see Bleaman, 2018) in a society with low literacy rates. Traditional tasks generally involve reading a passage and/or word lists, with the goal of accessing phonological variation. Bleaman (2018) adapted this to a text correction task to access morphosyntactic variation.

The purpose of the oral sentence correction task was to gather targeted data with the variable of interest, as participants heard and repeated identical phrases for speech samples that could be easily compared. Due to the nature of the tasks, participants had to pay close attention to their way of speaking. Reading tasks are often used in sociolinguistic research to boost conscious attention to speech, and the increased attention paid to speech often leads to more instances of a preferred variant. In the oral task with unfamiliar phrases, participants also had to pay explicit attention to their speech.

During this task, each participant heard, repeated, and corrected the same 19 recorded sentences, of which seven had stimulus verbs containing the study variable, some with and some without the uvular. (Eight phrases had verbs containing the exhortative morpheme, which also exhibits uvular alternation at different rates, briefly described below. The four remaining phrases were distractors and did not have any morpheme of interest.) Example (7) a uvular-less stimulus.

(7) *chita-kuna-ta apa-**ra**-ni-ku-m*
goat-PL-ACC take-PST-1SG-PL-EVID
'We took the goats.'

Stimuli were recorded by a native Chanka-speaking research assistant unknown to the participants. The research assistant was from an urban area and bilingual in Spanish. She assisted the first author in designing native-sounding sentences before recording. The distribution of each morpheme can be seen in Table 3.

For urban participants, the task was carried out in the following way:

Table 3. Distribution of realizations of stimuli for experienced past

With uvular [rqa]	Without uvular [ra]	Total
4 items	3 items	7

1. The participant heard a short recorded phrase
2. He/she repeated exactly what was heard (whether good or bad)
3. He/she declared whether it was good – *Allinchu kachkan?*
4. If it was bad, he/she corrected the phrase – *Imaynatam kanman?*

However, the activity was carried out differently for rural participants, who had difficulty producing unnatural or ungrammatical speech forms and thus did not repeat ungrammatical phrases. In the alternative cases, the participants corrected the phrase immediately with their first repetition. For analysis this was understood to be a representation of a better, more natural way of speaking for that participant, who preferred the phrase that he or she produced over the original phrase. When the participants did not offer an exact repetition the first time, their first phrase was considered to be their correction. For the few participants who repeated the recorded phrase exactly, with grammatical errors, they were asked to correct it. Then the second repetition, the corrected one, was considered to be their preferred way of pronouncing that morpheme.

Following recording, each stimulus was coded for whether the uvular was present or absent in the target morpheme. The final repetition by each participant was then coded for *same, inverse*, or *other*, depending on the pronunciation of the uvular compared to the stimulus. *Same* responses repeated the morpheme exactly as pronounced (whether uvular-full or -less). *Inverse* responses used the alternate morpheme variant (uvular-less instead of uvular-full, and vice versa). Finally, *other* responses did not include either variant of the morpheme (whether the morpheme was deleted or replaced by a different one).

3.3 Interview speech

Prior to the sentence repetition task, participants were recorded in a conversational setting that was structured like a classic sociolinguistic interview. These conversations were conducted with two participants instead of one, with the goal of mitigating the observer's effect. This provided a corpus of naturalistic speech with less of the self-monitoring promoted by the sentence correction task.

The interviewer was either a local Chanka-speaking research assistant or the first author, who is an intermediate speaker of Chanka and fluent speaker of

Spanish; and the two participants were Chanka speakers who were close friends or family members. The interviewer (third person) asked questions in Quechua, but since participants knew each other well, the setting evoked dialogue and narratives with minimal intervention from the interviewer. This created a situation where participants were paying less attention to their speech, and potentially using a more casual style. Specific questions and topics were designed to elicit tokens of the variable of study (the experienced past morpheme).

Rates of uvular alternation in this corpus were analyzed using a subset of 32 participants. The data from this part of the study provides a baseline characterization of speakers' production of the uvular variant, while the sentence correction task is intended to access their preferred realization (or perhaps underlying representations) of the /-rqa/ morpheme.

4. Results

The analysis included a total of 269 tokens for the sentence correction task (38 participants), and 2353 tokens for the interviews (32 participants). Data from 32 participants were included in both interview and sentence correction speech data. Two urbanites who had arrived to the city after adolescence and four rural women were excluded from the analysis of interview speech to better balance the sample.

Both rural and urban speakers show variation across the tasks, but at different rates. In this section we present results from the sentence correction task (§ 4.1) and compare them to rates in sociolinguistic interviews (§ 4.2) to illustrate the rural tendency for uvular presence in the experienced past /-rqa/. The characterization of rural speech as having more phonological segments is instantiated in their greater preference for uvular presence in the sentence correction task and have higher rates of uvular presence in production.

While the sociolinguistic interview is designed to reduce the subject's attention paid to speech by creating an informal conversational setting, the sentence repetition task calls the participant's attention directly to the utterance. In this way, this task is intended to be more sensitive to participants' ideologies of authenticity, and is the focus of our analysis here. We will further illustrate the efficacy of the sentence correction task for determining the preferred variant by comparing results for experienced past /-rqa/ with those for the exhortative /-rqu/ morpheme (§ 4.3).

4.1 Results from oral sentence correction task

The results of the sentence correction task provide evidence for the preferred realizations of the experienced past morpheme and ultimately demonstrate variability. Table 4 shows *same, inverse*, and *other*[6] answers for each stimulus variant. In this table, the -rqa- responses (the *same* responses to a -q- stimulus, and the *inverse* responses to a -Ø- stimulus) are highlighted in yellow. A chi-square test of independence showed a significant relationship between the stimuli type and repeated variants, X^2 (2, N = 269) = 26.89, p < .001.

For uvular-full stimuli, overall high rates (75%) of *same* repetition and low rates (13%) of *inverse* responses indicate a preference for processing and producing the uvular-full variant. For uvular-less stimuli, lower rates (51%) of *same* repetition and increased rates (39%) of *inverse* responses indicate an ability to process but preference for the uvular-full variant (*inverse*). These response patterns indicate that, for the experienced past, the [rqa] variant is the preferred form, perhaps in some sense the underlying representation for many of these participants. The [ra] variant is recognized and can be reproduced, but is dispreferred.

Table 4. Summary of same, inverse, and other repetitions for past morpheme

		Answer		
		Same	**Inverse**	**Other**
Stimulus	q	116	19	18
	Ø	59	46	11

Figure 1 displays response results for each stimuli word. Full words are shown under each corresponding bar. Bars to the left of the vertical black line show stimuli words that were pronounced with a uvular, and bars to the right show uvular-less words. In the bars, red indicates *same* responses, grey indicates *inverse* responses, and teal indicates *other* types of responses.

When the stimulus verb contained the uvular-full [rqa] variant (first four bars), it was repeated identically (red fill) almost 100% of the time for two words and over 50% of the time for the other two words. The higher rate of uvular absence for *miku-rqa-nki-chik* (eat-PST-2-PL) 'you all ate' may be related to the larger number of syllables in the word. The high rate of uvular presence for *upya-rqa* (drink-PST) 'he/she drank' is likely due to the location of stress preceding the uvular. The greater than expected number of *other* responses for *tuku-rqa-nki* (finish-PST-2SG) 'you finished' may be because the stimulus phrase was quite long, and participants seemed to have trouble remembering it.

6. "Other" responses included substitution of a different morpheme, deletion of the relevant morpheme, or deletion of the entire verb.

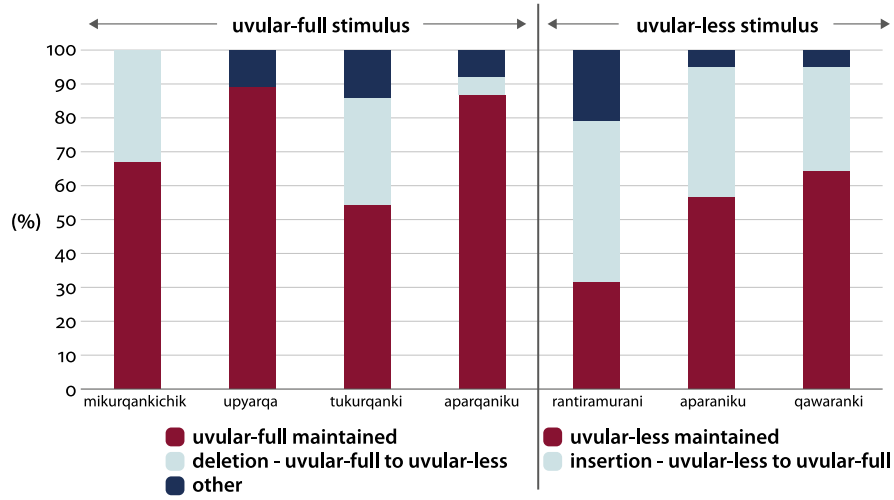


Figure 1. Results from the correction task: Variants of the experienced past /–rqa/

For the uvular-less past tense stimuli, participants were less successful at *same* identical repetition (red fill) of a verb with the [ra] variant, reaching only 30–65%. In these cases, almost all of the remaining repetitions were *inverse* and included an inserted uvular that had not been in the recording (demonstrated by the grey parts of the bars). This suggests that participants were able to understand the uvular-less [ra] variant, but the preference is to pronounce the uvular. This is especially clear for the verb *apa-rqa-niku* (take-PST-1PL.EXCL) 'we took', which appeared twice in the stimuli, once as each type of stimulus. The uvular-full stimulus was almost always repeated exactly, *same*, whereas the uvular-less stimulus was only repeated *same* (uvular-less) about half of the time, and almost all the rest of the time *inverse*, where participants inserted a uvular. It seems clear that participants prefer the uvular-full variant of *aparqaniku*.

Ultimately, these results confirm variability in the realization of the uvular in the experienced past. The uvular-full variant is easily reproducible, and the uvular-less variant is understandable. But since the uvular is recovered in production even when not heard in the recording, and rarely omitted in production when present in the stimulus, it appears that speakers expect or prefer the uvular-full variant, at least for most of the stimuli words. As for metalinguistic evaluation, speakers did not offer any conscious evaluation or overt preference for either of the two forms ([rqa] vs. [ra]), even when explicitly asked about them.

To estimate effects of social factors, responses were then broken down and analyzed according to location of residence, age, and sex. Some differences were found, but most trends were not statistically significant. Since token counts in

these breakdowns were quite small, these patterns may be reconsidered in future investigations.

As for location of residence, urbanites showed a growing acceptance and use of the uvular-less variants compared to rural residents. While urban and rural participants had similar rates of *same* responses (75%) for the uvular-full stimuli, they differed in their production and acceptance of uvular-less stimuli. With uvular-full stimuli, urbanites were also more likely to delete the uvular, doing so 15% of the time (13/84 *inverse*), while rural speakers deleted it only 10% of the time (7/69 *inverse*). For the uvular-less stimuli, urban responses were 56% *same* (35/63) and 33% *inverse* (21/63), whereas rural responses were 45% (24/53) for each. This hints at broader production of the uvular-less variant among urbanites. Rural speakers recognized the uvular-less variant, but were likely to re-insert the uvular. However, these conclusions require additional data, as location-based differences were not found to be significant, X^2 (2, $N=269$) = 26.89, $p<.001$ for the uvular-full stimuli, and X^2 (2, $N=269$) = 26.89, $p<.001$ for the uvular-less stimuli. A Cochran-Mantel-Haenszel test comparing the chi-square tables separated by location confirmed that they were not significantly different, X^2 (1, $N=269$) = 0.01, $p=.95$.

Regarding age, younger and middle-aged responses reflected the overall rates, whereas older speakers diverged. For the uvular-full stimuli, the older speakers had 63% *same* (17/27) and 22% *inverse* (6/27) responses. For the uvular-less stimuli, older speakers had 37% *same* and 37% *inverse* (7/19) responses. The lower rate of uvular-full *same* responses for older speakers is unexpected, but given the small sample, is also not significant at the .05 level. The lower rate of *same* responses for the uvular-less stimuli is consistent with a dispreference for this variant in this age group.

As for sex, patterns are less clear. While both sexes had similar rates of *same* responses (males 74%, females 76%) for the uvular-full stimuli, they differed in their production and acceptance of uvular-less stimuli. In repetition of uvular-full stimuli, males were more likely to delete the uvular, doing so 17% of the time (12/69 *inverse*), while females deleted it only 10% of the time (8/84 *inverse*). In repetition of uvular-less stimuli, females were more likely to maintain uvular deletion, repeating the uvular-less variant 55% of the time (36/65 *same*) and inserting a uvular 35% of the time (23/65 *inverse*). Male responses were more mixed, repeating the uvular-less variant 45% of the time (23/51 *same*) and inserting a uvular 43% of the time (22/51 *inverse*). While males seem slightly more likely to delete the uvular in their reproduction of a uvular-full stimulus, they were also more likely to reinsert it in their reproduction of a uvular-less stimulus. Though females did not delete the uvular as often in their reproductions, they seem more likely to repeat the uvular-less variant exactly, indicating more acceptance of this

variant. While these trends are interesting, there was no significant difference between males and females for either the uvular-full, X^2 (1, $N = 269$) = 0.19, $p = .66$, or uvular-less stimuli, X^2 (1, $N = 269$) = 4.99, $p = .08$. A Cochran-Mantel-Haenszel test comparing the chi-square tables divided for sex also found that differences were not significant, X^2 (1, $N = 269$) = 1.21, $p = .54$. Additional data may clarify the exact nature of these sex-based differences.

The sentence repetition task was designed to increase attention paid to speech. To evaluate the stylistic effects of this task, experienced past task response patterns from the ten participants with under 80% uvular presence in the interviews were compared with each speaker's spontaneous speech patterns. Justina (79% presence in interview), Laurimar (75%), and Máximo (77%) repeated all seven tokens with the variant that was in the recording – they followed directions despite their own patterns, and did not show a preference for one variant over the other. Nelly (17%), Roberto (7%), and Basilia (69%) were similarly good with 6 accurate repetitions. The first two seem to have a slight preference for uvular-full speech in this task, as both inserted a uvular in one token with a uvular-less stimulus; Basilia deleted instead for one token with a uvular-full stimulus. Carmelo (37%), Elías (28%), and Ricardo (33%) had 4–5 accurate repetitions, and at least one deletion and one insertion each. Finally, Cristina (50%) pronounced uvular-full variants for all seven words, regardless of the stimulus from.

Thus, there is no clear correlation between speakers' spontaneous speech usage of the uvular in experienced past morphemes and their repetitions in this task. Several speakers, urban teachers especially, repeated the exact variant that they heard. Others did this mostly, with a few flips for both types of stimuli – which does not show a clear preference for either variant. Cristina seems to be the only participant to prefer the uvular-full pronunciation for experienced past in this task, which may be a sign that she was paying more attention to her speech if this variable is stylistically conditioned. She is a teacher with Quechua certification and thus has exposure to the Ministry of Education orthography – though she does not teach in Quechua. Further analysis of individual patterns will be treated in future research.

4.2 Results in interview speech

The preference for uvular presence that is evident in the response patterns in the sentence correction task is confirmed by the treatment of the experienced past morpheme in the participants' usage in speech. In the semi-structured sociolinguistic interview conversations, 2353 uvular tokens were analyzed, with an average of 77 tokens per speaker (ranging from 5–164 tokens).

The overall rate of uvular presence in the experienced past was 81.4%, though this varied substantially based on several social factors. Geographically, urban speakers have lower rates of uvular presence than rural speakers, as seen by the breakdown of uvular presence rates by individual participants in Figure 2. The graphs are divided by location and sex, with speaker pseudonyms and age under each bar. The orange portion of the bar represents the percentage of uvular presence; blue shows percentage of uvular absence. The percentage of uvular presence for each participant is listed on top of each bar.

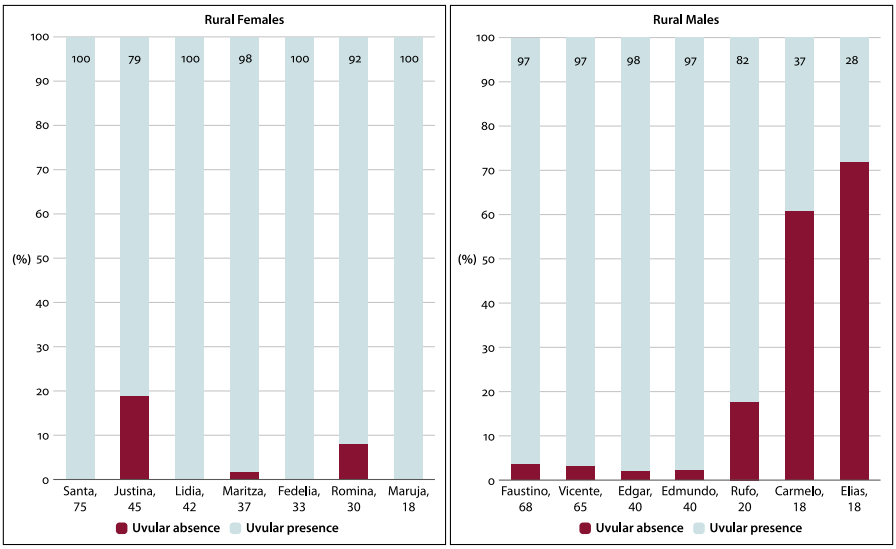


Figure 2. Individual speaker rates of uvular presence, divided by location and sex

The median presence rate for the rural speakers is 97%, and 9 of 14 people are near categorical – they have 97% or more uvular presence. For urban speakers, the median presence rate is 90%, and only 4 of 18 people have near categorical usage as defined above. Some urban speakers have much lower rates of uvular presence, including one male who has 7%, and one female who has 17% presence. The unweighted average, or mean of all the speaker percentages, shows similar effects. By this measure, the average rural speaker percentage is 86.1%, and the average for urban speakers is 75.7%.

Location of residence – rural vs. urban – is part of a cluster of highly correlated social characteristics, including Spanish proficiency and level of education. The participant selection criteria for urban speakers required them to be born in the city or have arrived in early childhood, which meant that participants were generally simultaneous bilinguals; they also had higher educational levels than

rural speakers. All of these characteristics were associated with lower rates of uvular presence, but a multivariate analysis showed the location of residence to be the most relevant predictor. Another important social correlate of uvular occurrence was speaker age: older speakers use more uvulars, among both rural and urban residents.

To measure the significance of these social correlates, uvular alternation in the past morpheme was modeled using logistic mixed-effects regression using the *lmer()* function in the *lme4* package (Bates, Machler, Bolker, & Walker, 2015) in the statistical program R (R Core Team, 2013), and using the step() function to find a best fit model by backwards elimination. The mixed effects models allowed for inclusion of fixed and random effects, which allow for inter-speaker particularities and speech preferences that diverge from what their social characteristics might predict, as well as for word-specific realizations (Johnson, 2009).

The best fit model is summarized in Model 1 below. The social predictors in the best fit model were sex, age, and location. In addition to the social predictors, relevant linguistic factors were also included: stress, preceding morphemes, and preceding syllables. The factors that were originally included but ultimately removed from the final model were: Following morphemes, Following syllables, Style, Education, Spanish proficiency.

Model 1. Best fit model for uvular alternation, with only significant factors

```
q ~ PrecMorphemes + PrecSyllables + Stress +
      sex + Age + location +
      (1|speaker) + (1|word)
```

For social factors, the results of this analysis indicated that age was significant at the $p < .05$ level, with older speakers using more uvulars. Location of residence approached significance ($p < .051$), with rural speakers using more uvulars than urban speakers. Linguistically, uvulars are favored in post-tonic syllables preceded by a monosyllabic root morpheme. Table 5 shows results from the best fit model.

Table 5. Estimates for effects from best-fit model of uvular presence

	Estimate	Std. error	*t*-value	*p*-value		*N*	% Presence
(Intercept)	0.2940	0.1510	1.94	0.582	.		
Preceding morphemes (vs. two or more)						700	72.3%
one	0.0432	0.0199	2.17	0.0300	*	1653	85.2%
Preceding syllables (vs. two or more)						1428	74.4%
one	0.1200	0.0308	3.90	0.0001	***	925	92.1%

Table 5. *(continued)*

	Estimate	Std. error	*t*-value	*p*-value		*N*	% Presence
Stress (vs. following)						528	70.8%
preceding	0.2200	0.0272	8.11	0.0000	***	952	93.7%
same	0.0266	0.0236	1.12	0.2612		873	74.3%
Sex (vs. male)						1419	81.1%
female	0.0723	0.0829	0.87	0.3903		934	81.8%
Location (vs. urban)						1626	80.3%
rural	0.1790	0.0883	2.03	0.0511	.	727	83.8%
Age (continuous)	0.0066	0.0026	2.53	0.0146	*		

significance codes: *** = < 0.001, ** = < 0.01, * = < 0.05, . = < 0.1

4.3 Comparison with exhortative morpheme results

The efficacy of the sentence correction task was confirmed when opposite results regarding uvular presence were discovered in the phonologically similar exhortative morpheme /-rqu/ and its allomorph /-rqa/. These results also highlighted the morpheme-specific nature of the uvular variation. In interview speech, the uvular exhibits almost categorical absence (99.13% absence or 0.87% presence) in the exhortative morpheme. Then, in the correction task, participants repeated almost all of the uvular-less variants exactly as heard, /-ru/ and /-ra/ – further indicating that these are the dominant forms. In fact, the uvular-less forms are so preferred that many participants had difficulty processing uvular-full forms on the correction task (they replaced the exhortative morpheme with a different one, usually the experienced past tense).

Figure 3 presents results of the correction task for the exhortative morpheme, which was found to appear uvular-less as [ru] almost all of the time in natural speech. Bars to the left of the vertical black line show stimuli words that were pronounced with a uvular, and bars to the right show uvular-less words. The full stimuli words are shown under each corresponding bar. Red indicates *same* responses, grey indicates *inverse* responses, teal indicates where response contained a past tense marker instead, and orange indicates *other* types of responses.

When the stimulus verb contained the uvular-less [ru] variant (final five bars in the figure), repetition of it was exact almost 100% of the time, as seen by the

almost completely red bars.[7] In the opposite case, participants exactly repeated a stimulus verb with the uvular-full [rqu] variant only a little more the 50% of the time. In these cases, some errors in stimuli processing occurred: sometimes the /–rqu/ morpheme was replaced with experienced past [rqa] (teal), or changed to other morphemes or completely deleted (orange) – usually after the participant had asked to hear numerous repetitions of the same phrase, seemingly with the intention of better understanding it.

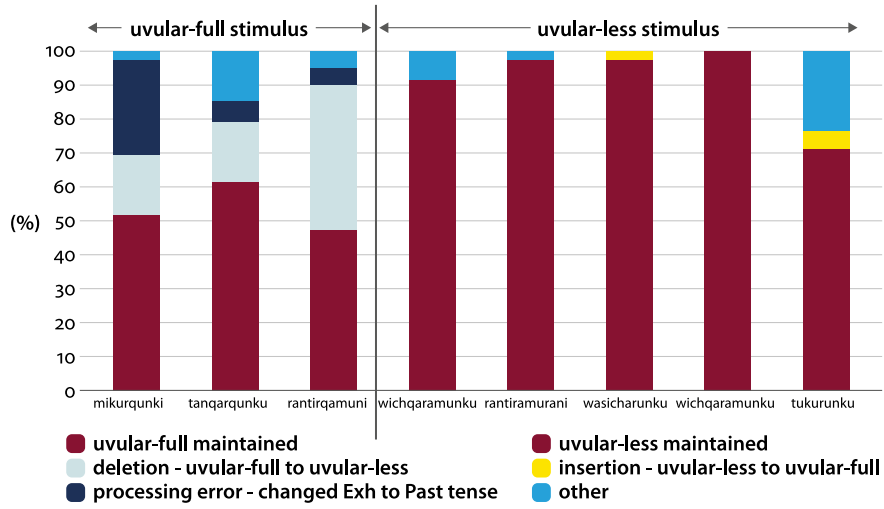


Figure 3. Results from the correction task: Variants of the exhortative morpheme /–rqu/

These results confirm that the uvular is almost never pronounced in the exhortative.[8] In fact, for some speakers this morpheme has been lexicalized as [ru], given that they seem not to understand or process the verbs with a uvular-full [rqu] morpheme very well. At the same time, limited other speakers were able to process and repeat the uvular-full exhortative. Urban speakers who have studied Chanka orthography and some rural speakers repeated some uvular-full morphemes correctly. One even recovered it when not pronounced: Rosa (25-year-old rural female) asked for repetition of the stimulus containing uvular-full *miku-rqu-nki* (eat-EXH-2)'you have eaten'. The RA interviewer gave a uvular-less pronunciation; regardless, Rosa recovered the uvular with *puchullantaña*

7. The verb with some variation in its repetition was the one in the final phrase (number 19), and it occurred after many phrases with a uvular, which could have conditioned participants to pronounce a uvular-full repetition.

8. The only exception is in conscious speech by the few speakers trained in Quechua orthography – and only in a language education environment.

miku-rqu-n (leftovers eat-EXH-3) 'he/she has eaten the leftovers'. However, very few of these insertions occurred.

As was shown previously for the past morpheme, the table below (Table 6) shows *correct, inverse*, and *other* answers for each exhortative stimulus variant. In this table, the preferred variant for the morpheme (the majority of *same* responses, with a confirming high number of *inverse* responses for the other variant) is highlighted in yellow. For the exhortative, [ru] is preferred because it is repeated correctly or the *same* 173/189 times. Moreover, this morpheme, when uvular-full, is sometimes misunderstood, which is seen by the 24 *other* answers, highlighted in red. The significance of these results was confirmed with a chi-square test of importance, X^2 (1, $N = 301$) = 62.64, $p < .001$.

Table 6. Summary of same, inverse, and other repetitions for exhortative morpheme

		Answer		
		Same	**Inverse**	**Other**
Stimulus	q	60	28	24
	∅	173	3	13

5. Discussion

Speech style and the status of speakers have long been treated in sociolinguistic research as providing two complementary windows on the social evaluation of linguistic variation. But in a marginalized minority language like Quechua, it is not clear that this model is relevant to how speakers produce and perceive the social significance of variables for at least two reasons. First, the very conception of style in Western English-speaking communities, proxied by attention paid to speech and formal/informal dimensions, depends on the relevance of these contexts and normative pressures to stylistic variation in the local community. Second, the traditional tasks used to access speech styles, usually written, depend on widespread literacy in the language of study.

The attention-to-speech model developed by Labov (1972), commonly used in sociolinguistics, postulates that speakers use more formal language when paying greater attention to speech. In majority languages such as English, formal language often reflects a prestigious standard that is reinforced through written norms in the education system. Furthermore, public venues and careful speech become associated with a standard or prestige variety. Thus, attention paid to speech seems intrinsically tied to standard language ideology, where speakers have a conception of "correct" forms, and will try to use them more when

they are aware. However, as many scholars including Labov himself (2016) have noted, "style" must be defined on local terms, especially for endangered languages.[9] This is because style is not necessarily intertwined with "standard," "conservative," and/or "prestige" in communities with fewer normative pressures (Meyerhoff et al., 2020).

Still, many variationist studies on endangered languages have identified a reduction of styles in use by speakers. This "stylistic shrinkage" (see Kasstan 2019 for an in-depth overview) is often attributed to a reduction in formal domains of use for the language, leaving only intimate domains and thus casual speech. However, while it is true that many bilingual speakers now use Quechua in fewer domains in Peru, it seems premature to assume that stylistic atrophy automatically follows. In fact, several completed studies on other endangered languages contest such an assumption. For instance, Clarke (2009) found stylistic variation for four variables in Sheshatshiu-Innu-aimun (Canada), even though the other fourteen variables in the study did not vary along the stylistic dimensions considered. Kasstan (2019) found that speakers of Francoprovençal also use stylistic variation, absent a standard or prestige norm. Furthermore, in Máíhɨ̄ki, another endangered language spoken in Peru, Skilton (2017, p. 108) found stylistic variation between 'correct' and 'incorrect' uses, despite the lack of a written standard.

Given the evidence of stylistic variation in similar contexts of language endangerment, the lack of task-dependent variation among Quechua speakers for the uvular variable does not necessarily evidence a complete lack of stylistic variation. Rather, alternative repertoires of style likely exist. In Andean communities, the diglossic situation, with Spanish being expected in formal situations (including in school until recently), limits and potentially eliminates the number of Western-like formal environments available for Quechua use. Thus, stylistic variation is likely orientated around different social norms besides the casual and formal settings explored in this study. Notably, Kasstan (2019, p. 699) found "substantial variation within the most careful speech style" in Francoprovençal, which suggests that speakers were orientated toward an alternative stylistic dimension.

Rather than assume lack of stylistic variation, we must seek evidence of notions of style along non-traditional, non-Western lines. In particular, researchers from outside the community of study can take advantage of the expertise of local speakers regarding social nuances. As we have begun to do for social categories (i.e., class and sex), gathering local insight leads to a more accurate

9. Here our focus is on the external social norms guiding stylistic variation related to the attention-to-speech model. Once the Quechua repertoire of style is identified, future research can employ agentive approaches, which argue for an individualized, event-based analysis of style for all languages.

description of the social meaning of linguistic variables in these communities. The search for an accurate conception of Quechua speech styles continues beyond this paper, as several questions remain: Do speakers have casual and formal styles in the Western sense? If so, where is each one used? Do these styles correlate with attention paid to speech? Or, does Quechua have alternative ideologies guiding variation?

As for task design, the present study developed alternative non-written methods for investigating the traditional formal-informal stylistic range in minoritized languages. As previously mentioned, the sentence correction task, compared with the interview conversation, was intended to address style as attention paid to speech, in the Labovian model (Labov, 1972). Traditionally, relatively un-self-monitored or 'casual' style is elicited through an informal oral interview, and self-monitored 'careful' speech through a formal reading task. In the present study, however, the limited literacy in the population necessitated using oral tasks for both, with the phrase correction task drawing the speakers' attention specifically to the variable in question. But although the two tasks provided complementary insight as to the most common forms, and possibly the forms preferred by each group, they did not appear to directly approximate formal and informal styles in the traditional sense. Participants exercised similar levels of carefulness throughout the entire recording session, thus prompting further review of these methods for future research in this population.

Furthermore, using attention paid to speech as a proxy for formal/informal did not lead to expected differences in use of the uvular. A higher rate of uvular presence was expected in the repetition task compared to the interview because we assumed that uvulars would vary based on formality. Again, no clear difference between the tasks was found; nor was there an inter-task correlation between rates for individual speakers or certain social groups. However, as previously mentioned, this lack of difference does not necessarily indicate a lack of stylistic variation in Quechua. Instead, it underscores the need to expand our repertoire of styles available, especially in non-Western societies that are less oriented around written language and norms.

In any case, the responses in the phrase correction task clearly indicate an overall implicit preference for the uvular-full pronunciation of /-rqa/. This also happens to be the variant used more in conversation by speakers who are older, rural, less-educated and less Spanish-dominant. In the sociolinguistic research tradition focused on majority languages with established 'standards', this would appear contradictory: in this community lower status speakers are using more of the preferred variant and also following the orthographical norm, which is the opposite of the findings in classic studies beginning with Labov's work in New York City (Labov, 1966). But in the case of Quechua, this is not surprising, given

that it lacks "'superposed' norms about 'correctness'" (Meyerhoff et al., 2020, p. 7), as the educational standard is not widely "reinforced" through instruction or social practice. In fact, those who developed Quechua orthography intended to reflect this rural speech. Furthermore, the lack of difference between conditions with supposedly distinct levels of formality highlights the fact that Quechua is not utilized in formal domains in Peru, and thus no formal register, as traditionally conceived, exists. Rather, these results suggest that the social evaluation of varieties of Quechua turns on an ideology of authenticity, in place of an association between ways of speaking and the status of speakers.

Authenticity is thus the alternative language ideology that is most evident in the current research. While most participants had no metalinguistic or social awareness of the uvular variable, teachers specifically trained in Quechua orthography attributed it social value. They positively described uvular-full speech as more *dulce* 'sweet', *puro* 'pure', and authentically Quechua. In part due to the way Quechua orthography was developed, older, rural speech more closely approximates an idealized authentic variety. The official Chanka orthography was finalized in 1989 by academics who prioritized the forms used by monolingual speakers (Hornberger, 1993; Jung & López, 1987; Zúñiga, 1987). However, it is still unclear exactly what social norms guide rural speech, despite its proximity to authenticity, given speakers' low literacy levels and unfamiliarity with authenticity discourse. As for socially established urban speakers, their uvular rates were much more varied. They attributed positive value to rural speech in general, considering it more authentic, even as their own usage diverged from it.

Thus, a desire for authenticity drives much of this idealization of the 'pure' form of Chanka Quechua. Features of authentic speech are based partially on knowledge of the official orthography and norms, and partially on local perceptions of purity, mainly meaning minimal influence from Spanish. Other ways of speaking are discredited as inauthentic, especially in the classroom – and so are the speakers. Unfortunately, the current construction of Quechua language "delegitimizes" and "can disempower" many Quechua speakers in urban areas, as Zavala et al. (2014, p. 46) describe in detail.

This paper addresses the ideology that considers authentic Quechua to be static. The results presented here confirmed the existence of uvular variation for all social groups – including older, rural speakers. Whether or not the uvular is stylistically stratified or carries social value, the fact that it varies at all is currently ignored in the standardized orthography and its implementation. In educational materials, the uvular is always present, which leads to its being seen as the more correct, prestigious variant. Initial indications of this ideology are found in the opinions of teachers, who have knowledge of the writing system and judge uvular-full speech as more correct. These educated bilingual speakers have

begun to transfer the school-learned concept of standard language ideology from Spanish to Quechua. However, as noted above, few of our participants have any familiarity with this orthography or exposure to Quechua-language instruction. Even so, as this type of ideology spreads, it may be detrimental to revitalization efforts.

The ideological path to a search for Quechua authenticity has come alongside standardization. Back in 1950, Rowe mused, "It would be interesting to see what could be done if Quechua and Aymara were legalized for teaching and official business in the provinces where they are spoken by a significant fraction of the population" (215). Despite Quechua officialization and many language education initiatives since that time, language shift to Spanish is still occurring. In fact, discourses or methods of implementation of some of these initiatives may be contributing to it. As Zavala (2020) noted, "policies in favor of diversity can hide" the true social situation, which is not as supportive of variation in cultures and languages.

The Peruvian Ministry of Education frames Quechua as a "'discrete', bounded object" (Patrick, 2007, p. 124), indicating that only the linguistic forms in written materials are authentic and valid. This viewpoint is common in endangered language communities. As governments seek "homogeneity in nationhood," speakers find that "[language] rights will always be circumscribed by an ideology of linguistic homogeneity, and linguistic diversity will thereby be compromised" (Patrick, 2007, p. 129). While the Peruvian government recognizes 48 official languages, only the standardized form is valued.

The "correct" orthographical Quechua forms are justified as being the most authentic forms. In fact, most orthographic rules in Quechua were determined for historical, unification, preservation, or compatibility reasons (Hornberger, 1993; Jung & López, 1987; Zúñiga, 1987). Authenticity and faithfulness to (sometimes reconstructed) historical forms is generally a driving factor in determining which forms gain official status. Standardizing authorities refer to reconstructions by linguists or other professionals of what the forms of a variety or dialect used to be, or maybe to how older community members (more "authentic" ones) currently speak (Fishman, 1988; Hornberger, 1993). This may be the source of the authenticity ideology for Quechua teachers.

Even where successful implementation of the official standard Quechua orthography has occurred, it has not stopped or reversed language shift. While it gives Quechua more visibility and legitimacy as a language, it also reinforces ideologies that frame the language as static and monolithic. Government documents affirm that Quechua "is not a dead language or of the past" (Zavala et al., 2014, p. 11, quoting Gobierno Regional de Apurímac, 2009). However, "standardizing processes … lead to linguistic forms that are artificial" (Patrick, 2007,

p. 129). In fact, the Peruvian Ministry of Education standardized Chanka orthography ignores variation that is inherently present in the language. The discrepancy between a growing acceptance of uvular-less variants, shown in this paper, and the static uvular presence in the orthography, evidences this point.

Ideologies around authenticity have been found to affect variation in other minority languages, though each process is community-specific. For instance, new Basque speakers (who learn the standardized *Batua* variety in school) learn invariable Object Marking as the correct form (Rodríguez-Ordóñez, 2021), similar to how lack of variation is taught in school Quechua. However, through interaction with native speakers of local varieties, the new speakers of Basque also learn to apply variation (Differential Object Marking) to increase their own authenticity (Rodríguez-Ordóñez, 2021). In this sense, Quechua speakers and learners believe that this static variety represents authentic Quechua, rather than the variable language spoken by real, often bilingual, speakers. Currently, only monolingual Quechua language is viewed as authentic, and this aligns with the standardized variety taught in schools. While the population propagates this idealized variety, most admit that few speakers of it actually exist.

In Peru language mixing is also a clear social marker of inauthenticity. Despite being extremely common, Quechua with Spanish mixing is stigmatized as "contaminated," as Laurimar described above. While most urban participants claimed that only urban speech was contaminated, Laurimar later retracted her sentiment that rural speech is pure based on her own experiences as a teacher in a rural community:

> *En el campo te dicen igual, mezclando. Siempre se le mete algo del castellano.*
> In the countryside they say it the same way, mixing. Something from Spanish always slips in. (Laurimar, 36 years, urban female)

She recognizes that the authentic ideal does not exist, and Máximo (53-year-old urban male) agreed:

> *Mana puruchataqa rimankuchu. No hay puro, entonces una mezcla hay.*
> They [in the countryside] do not speak purely. There is no pure [Quechua]; there is a mixture. (Máximo, 53 years, urban male)

The stigmatization of Quechua varieties 'contaminated' by Spanish is similar to the antipathy for language mixing that Kroskrity (2018) describes for Tewa. He sees this "Indigenous purism" as a factor facilitating language maintenance, framing the language as an essential element of identity. But in the Quechua case, purism appears to militate against language maintenance, by devaluing the usage of so many speakers. For language maintenance purposes, reframing variable Quechua as authentic may legitimize it like local varieties of Basque have been. As

it stands, the Quechua speakers seen as most authentic have the least potential for social mobility, and thus are motivated to switch to Spanish and leave Quechua. The potential gains in authenticity of using Quechua in daily life – as speakers of Basque do – are outweighed by the continued negative stigma. Conversely, the urban, bilingual, educated, professional Quechua speakers could maintain the language without losing social or economic status – but they are not valued as speakers due to not speaking pure Quechua. Bilingual speakers are thus the most invalidated by the idealization of authentic expression in Quechua. This differs from Basque speakers, in that both traditional and new speakers are bilingual, and they can approximate authenticity by using nonstandard, contact-related linguistic features (Differential Object Marking). Using these features, Basque speakers can fall on a continuum of authenticity (Rodríguez-Ordóñez, 2021). This may be the case for Quechua speakers as well, though the popular ideology maintains a categorical definition of authentic language. Future work on Quechua should examine the possibility of describing variation on a continuum, possibly with a more salient feature.

Bilingual Quechua-Spanish speakers are the best positioned economically to maintain the language, as they have overcome most negative associations with it. The urban bilinguals in this study are often from the low middle class by regional standards, which means that their employment allows them to cover their family's basic needs in addition to a small amount of disposable income. Currently, these bilinguals can use Chanka in socially acceptable spaces such as historical cultural demonstrations – as long as they otherwise speak standard Spanish. However, they also usually have some instruction in Quechua writing and grammar, and thus strive for correct and authentic speech by those standards. Were their everyday speech also validated as authentic, they may be encouraged to use it more often. Unfortunately, these speakers tend toward less authentic, uvular-less speech, as seen in the results presented here. Similar results are likely for other variables, which heightens the need to raise awareness of the legitimacy of variation in the community.

Language shift to Spanish in Quechua-speaking communities, as shown by the absence of young Quechua speakers in the urban setting, evidences the pervasiveness of underlying negative ideologies about the value of Quechua in Andahuaylas and Peru, which certainly must continue to be addressed. At the same time, overt positive ideologies expressed by speakers, rooted in political discourse and new economic incentives, conceal the negative perceptions – though they also contribute to language shift in their own way, by only legitimizing language that is deemed authentic.

6. Conclusion

This paper found uvular alternation in Chanka Quechua (Andahuaylas, Peru) using data from an oral sentence correction task (explicitly monitored) and sociolinguistic interviews (less self-monitored). The uvular-full variant is the most common form of the experienced past morpheme, though most speakers use and understand the uvular-less variant as well. The oral task developed for this project effectively confirmed these hypotheses; however, it also highlighted the limitations of assuming a formal/informal style continuum that correlates with attention paid to speech. In endangered language communities where the dominant language is also the formal language, we can work to identify alternative style repertoires and methods for accessing them.

This paper also highlights the need to broaden our understanding of standard language ideology, seeking alternative community norms and priorities. Given the social situation of Quechua speakers, they do not align to a traditional standard language ideology based on a widespread written standard, the media, or a respected group. Instead, speakers – urbanites especially – orient their language attitudes toward authenticity, often tied to rurality and monolingualism. While authenticity can be a guiding force for language norms, we must be careful that it does not delegitimize Quechua speakers who do not fit neatly into an authentic identity. In this way, studies of variation in minority language communities can legitimize different ways of speaking in the community, which ultimately encourages all types of speakers to continue using the language.

References

Adelaar, Willem F. H., with Muysken, Pieter C. (2004). *The Languages of the Andes*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486852

Babb, Florence E. (2020). 'The real indigenous are higher up': Locating race and gender in Andean Peru. *Latin American and Caribbean Ethnic Studies*. https://doi.org/10.1080/17442222.2020.1809080

Bates, Douglas, Mächler, Martin, Bolker, Ben, & Walker, Steve (2015). "Fitting linear mixed-effects models using lme4." *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bleaman, Isaac (2018). Outcomes of minority language maintenance: Variation and change in New York Yiddish. Unpublished doctoral dissertation, New York University.

Bucholtz, Mary (2003). Sociolinguistic nostalgia and the authentication of identity. *Journal of Sociolinguistics*, *7*(3), 398–416. https://doi.org/10.1111/1467-9481.00232

Cerrón-Palomino, Rodolfo (1987). *Lingüística quechua*. Cuzco: Centro de Estudios Rurales Andinos "Bartolomé de Las Casas."

Chirinos Rivera, Andrés (1998). Las lenguas indígenas peruanas más allá del 2000 [Peruvian indigenous languages beyond 2000]. *Revista Andina*, *116*, 453–476.

Cisneros, Abilio, & Gálvez, César (2010). *Abriendo caminos. Guía de Estrategias Metodológicas para la Educación Intercultural Bilingüe en contextos urbanos* [Opening pathways. Guide of methodological strategies for intercultural bilingual educaiton in urban contexts]. Lima: Tarea.

Clarke, Sandra (2009). Sociolinguistic stratification and new dialect formation in a Canadian aboriginal community: Not so different after all? In James N. Stanford & Dennis R. Preston (Eds.), *Variation in indigenous minority languages* (pp. 109–128). Amsterdam: Benjamins. https://doi.org/10.1075/impact.25.06cla

Dorian, Nancy C. (1994). Purism vs. compromise in language revitalization and language revival. *Language in Society*, *23*(4), 479–494. https://doi.org/10.1017/S0047404500018169

ELAN (2021). [Computer program]. Version 6.1. Nijmegen: Max Planck Institute for Psycholinguistics. Available from https://tla.mpi.nl/tools/tla-tools/elan

Fishman, Joshua A. (1988). Ethnocultural issues in the creation, substitution, and revision of writing systems. In Bennett Rafoth, & Donald Donald (Eds.), *The social construction of written communication* (pp. 273–286). Norwood, NJ: Ablex.

Fishman, Joshua A. (1993). Introduction: Exploring an overlooked sociolinguistic phenomenon (The First Congress for Language X). In Joshua A. Fishman (Ed.), *The earliest stage of language planning: The "first congress" phenomenon* (pp. 1–9). Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110848984.1

Gobierno Regional de Apurímac (2016). *Plan de Desarrollo Regional Concertado Apurímac 2017–2021* [Coordinated regional development plan Apurímac 2017–2021]. 29 December 2016, retrieved on 8 July, from https://www.regionapurimac.gob.pe/2013/transparencia/wp-content/uploads/2017/01/Plan-de-Desarrollo-Regional-Concertado-Apurimac2017-2021.pdf

Hornberger, Nancy (1993). "The first workshop on Quechua and Aymara writing." In Fishman, Joshua A. (Ed.), *The earliest stage of language planning: The "first congress" phenomenon* (pp. 233–256). Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110848984.233

Instituto Nacional de Estadística e Informática (2017). *Censos nacionales XI de Población y VI de vivienda* [National censes of population XI and homes VI]. Retrieved on 23 May 2021, from https://censos2017.inei.gob.pe/redatam/

Irvine, Judith, & Gal, Susan (2000). Language ideology and linguistic differentiation. In Kroskrity, Paul V. (Ed.), *Regimes of language: Ideologies, polities, and identities* (pp. 35–84). Santa Fe: School of American Research Press.

Johnson, Daniel E. (2009). Getting off the GoldVarb Standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass*, *3*(1), 359–383. https://doi.org/10.1111/j.1749-818X.2008.00108.x

Jung, Ingrid & López, Luis E. (1987). Las dimensiones políticas de una escritura: el caso del quechua en el Perú [The political dimensions of an orthography: The case of Quechua in Peru]. *Allpanchis*, *29/30*, 483–509. Cuzco: Instituto de Pastoral Andina. https://doi.org/10.36901/allpanchis.v19i29/30.985

Kasstan, Jonathan R. (2019). Emergent sociolinguistic variation in severe language endangerment. *Language in Society*, *48*, 685–720. https://doi.org/10.1017/S0047404519000472

Kroch, Anthony, & Small, Cathy (1978). Grammatical ideology and its effect on speech. In David Sankoff (Ed.), *Linguistic variation: Models and methods* (pp. 44–55). New York: Academic Press.

Kroskrity, Paul V. (2018). On recognizing persistence in the Indigenous language ideologies of multilingualism in two Native American Communities. *Language & Communication 62*, 133–144. https://doi.org/10.1016/j.langcom.2018.04.012

Labov, William (1972). *Sociolinguistic patterns (No. 4)*. University of Pennsylvania Press.

Lefebvre, Claire (1976). Linguistic survey of Cuzco Quechua: Sampling procedures and data collection. *Anthropological Linguistics*, *18*, 328–339.

Lewis, M. Paul, Simons, Gary F., & Fennig, Charles D. (Eds.) (2016). *Ethnologue: Languages of the World* (19th ed.). Dallas, Texas: SIL International. Available from http://www.ethnologue.com

Meneses Tutaya, Norma, Alonzo Sutta, Alicia, Gonzales Rodríguez, María, & Huanca Quispe, Jaime (2020). Lineamientos para un método de enseñanza de quechua para bilingües receptivos en quechua [Guidelines for a Quechua teaching method for receptive bilinguals in Quechua]. *Lengua & Sociedad. Revista de Lingüística Teórica y Aplicada*, *19*(2), 143–168. https://doi.org/10.15381/lengsoc.v19i2.22310

Meyerhoff, Miriam, Abtahian, Maya R., Gafter, Roey J., Horesh, Uri, Kasstan, Jonathan R., Keegan, Peter, & King, Jeanette (2020). Styles, standards and meaning: Issues in the globalisation of sociolinguistics. *Language Ecology*, *4*(1), 1–16. https://doi.org/10.1075/le.00006.mey

Morgan, Marcyliena (2004). Speech community. In Alessandro Duranti (Ed.), *A companion to linguistic anthropology* (pp. 3–22). Malden, MA: Blackwell.

Nagy, Naomi (2009). The challenges of less commonly studied languages: Writing a sociogrammar of Faetar. In James N. Stanford & Dennis R. Preston (Eds.), *Variation in indigenous minority languages* (pp. 397–417). Amsterdam/Philadelphia: Benjamins. https://doi.org/10.1075/impact.25.20nag

O'Hanlon, Fiona (2015). Choice of Scottish Gaelic-medium and Welsh-medium education at the primary and secondary school stages: Parent and pupil perspectives. *International Journal of Bilingual Education and Bilingualism 18*(2), 242–259. https://doi.org/10.1080/13670050.2014.923374

Parker, Gary J. (1969). *Ayacucho Quechua grammar and dictionary*. The Hague: Mouton. https://doi.org/10.1515/9783112415641

Patrick, Donna (2007). Language endangerment, language rights and indigeneity. In Heller, Monica (Ed.), *Bilingualism: A social approach* (pp. 111–134). London: Palgrave Macmillan. https://doi.org/10.1057/9780230596047_6

Povilonis, Natalie (2016). "*Linguistic and ideological constraints on minority language variation: The case of deletion of the uvular /q/ phoneme in Chanka Quechua*." Unpublished qualifying paper, New York University.

Quintero Bendezú, Genaro R. (1996). *Gramática Básica del Quechua* [Basic Quechua grammar]. Huancavelica, Peru: Anmara Ediciones.

R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved 11 May 2018, from http://www.R-project.org/

Rodríguez-Ordóñez, Itxaso (2021). The role of social meaning in contact-induced variation among new speakers of Basque. *Journal of Sociolinguistics*, *25*, 533–556. https://doi.org/10.1111/josl.12477

Rowe, John H. (1950). Sound patterns in three Inca dialects. *International Journal of American Linguistics*, *16*(3), 137–148. https://doi.org/10.1086/464079

Skilton, Amalia (2017). Three speakers, four dialects: Documenting variation in an endangered Amazonian language. *Language Documentation & Conservation SP13* (pp. 94–115).

Soto Ruiz, Clodoaldo (1976). *Gramática quechua, Ayacucho-Chanca* [Quechua grammar, Ayacucho-Chanca]. Lima: Instituto de Estudios Andinos.

Stanford, James N., & Preston, Dennis R. (2009). The lure of a distant horizon: Variation in indigenous minority languages. In James N. Stanford & Dennis R. Preston (Eds.), *Variation in indigenous minority languages* (pp. 397–417). Amsterdam: Benjamins. https://doi.org/10.1075/impact.25.01sta

Zavala, Virginia (2014). An ancestral language to speak with the Other: Closing down ideological spaces of a language policy in the Peruvian Andes. *Language Policy*, *12*(3), 1–20. https://doi.org/10.1007/s10993-013-9297-4

Zavala, Virginia (2020). Una reflexión crítica sobre como el Estado enseña (y habla) sobre las lenguas originarias [A critical reflection about how the State teaches (and speaks) about indigenous languages]. Plenary presentation on 30 October, *Rimarisun Perú: Congreso internacional en lenguas originarias*.

Zavala, Virginia, Mujica, Luis F., Córdova, Gavina F., & Ardito, Wilfredo J. (2014). *Qichwasimirayku: Batallas por el quechua* [Because of the Quechua language: Battles for Quechua]. Lima: Fondo Editoral de la PUCP.

Zúñiga, Madeleine (1987). Sobre los alfabetos oficiales del quechua y el aimara [About the oficial Quechua and Aymara alphabets]. *Allpanchis*, *29*(*30*), 469–481. Cuzco: Instituto de Pastoral Andina.

Abstracta

Al igual que en otros idiomas marginalizados, en el quechua chanka (Perú) faltan normas de prestigio a nivel de la comunidad que se asocian con la ideología del lenguaje estándar. Las situaciones formales requieren el castellano y pocos hablantes tienen alfabetismo en quechua, así que los estilos normativos del habla están ausentes. Los juicios evaluativos de los hablantes no tienen referencia a las nociones de lo que es correcto, al contrario, valoran el habla pura y la autenticidad.

Este artículo explora los acercamientos alternativos para acceder los juicios sociolingüísticos con una investigación sobre el fonema uvular que es presente de manera variable en el morfema /-rqa/ del tiempo pasado, tal como se ejemplifica en la siguiente alternación:

(1) *ri-rɋa-ni* ~ *ri-ra-ni*
go-PST-1SG go-PST-1SG
'I went' 'I went'

En oposición al habla de las entrevistas sociolingüísticas, el habla cuidadosa, auto-monitoreada, se obtiene a través del recuento oral del material que se presenta auditoriamente, en lugar de en la escritura. De los 38 participantes, los hablantes rurales tienden a tener tasas más altas de /q/ que los urbanos, así que refleja el quechua puro idealizado. Argumentamos que la autenticidad guía la variación en lugar de la ideología del lenguaje estándar.

Address for correspondence

Natalie Povilonis
Department of Linguistics
Indiana University Bloomington
10 Washington Place #211
New York, NY 10003
USA
nsp235@nyu.edu

Co-author information

Gregory Guy
Department of Linguistics
New York University
gregory.guy@nyu.edu

Publication history

Date received: 14 January 2022
Date accepted: 16 August 2022

Author Queries

- Please provide a complete reference for the citation '(Gobierno Regional de Apurímac, 2009), (Chirinos Rivera, 2001), (Zariquiey & Córdova, 2008), (Labov, 1966)' in this article.
- Please provide a citation for the reference id "CIT0008 (Cisneros, Abilio, & Gálvez, César (2010)), CIT0011 (ELAN (2021)), CIT0013 (Fishman, Joshua A. (1993)), CIT0014 (Gobierno Regional de Apurímac (2016)), CIT0024 (Lefebvre, Claire. (1976)), CIT0026 (Meneses Tutaya, Norma, Alonzo Sutta, Alicia, Gonzales Rodríguez, María, & Huanca Quispe, Jaime (2020))" since citation is missing in the article.