

# Phonological Koinéization in Kathmandu Tibetan

Christopher Geissler  
*Yale University*

## 1 Introduction

This paper tests the new-dialect formation model of Peter Trudgill (1986, 2000 *inter alia*) by examining the Tibetan language spoken in the diaspora community of Kathmandu, Nepal. As with previous studies of this phenomenon, this is a community where speakers of many varieties of the language under study have come together in a region where that language was not previously spoken, and by now two generations have been born and raised in this new location. Phonological and lexical data was gathered in fieldwork conducted in the summer of 2016 in an attempt to capture the process of new-dialect formation in progress.

Trudgill's model predicts that the second generation born in the new region should exhibit both simplification, the failure of marked variants to transmit to subsequent generations, and focusing, the selection of particular variants by the community. While younger Tibetans (under age 30) in Kathmandu should belong to this generation, and they do indeed exhibit simplification of marked variants, evidence of focusing is not found. Indeed, both younger and older Diaspora-born speakers exhibit a comparable degree of diversity in their use of these variants, suggesting that the process of new-dialect formation, at least among Tibetans in Kathmandu, is proceeding more slowly than predicted.

## 2 New dialect formation

**2.1 *Trudgill's model*** Perhaps the most influential voice in the study of new dialect formation is that of Peter Trudgill, who has studied the development of a dialect in a community settled by speakers of several dialects in a place where the language had not previously been spoken. Drawing especially from research on New Zealand English, Trudgill (1986) proposed that such a new dialect develops predictably over the course of three generations.

The first generation consists of the original immigrants to the new location, who bring with them the speech of their home regions. As they interact with each other, some degree of accommodation takes place on the scale of individual conversations as well as over the course of speakers' lives (e.g. Giles and Powesland 1975 and subsequent work). These first-generation speakers have children, who grow up exposed to speakers of many dialects and lack a single dialect to copy. Thus, second-generation speakers display a wide range of linguistic features as different children select different variants from the diverse input. For New Zealand English, some of these speakers were recorded later in life in an oral history project; these recordings were acquired under the leadership of Elizabeth Gordon and now maintained by the Origins of New Zealand English (ONZE) project at the Department of Linguistics at the University of Canterbury, Christchurch. Trudgill et al. (2000) and Trudgill (2004) find that these speakers exhibit *ideolects* with features from many British dialects, but in a combination that matches no individual dialect. Additionally, these speakers display high intra- and inter-speaker variability, often producing several forms of a single feature as well as combinations of forms markedly different even from other speakers of their generation from the same small town.

Finally, Trudgill finds in the third generation, the second generation of native-born speakers, the emergence of a new dialect, "a stable, crystallized variety" (Trudgill 2004:307), a process he terms "focusing." Along the way, Trudgill (1986) describes the dialect as having passed through two major

---

\* This research was made possible by the advice and support of Claire Bower, Nawang Tsering, Tsering Dorje, Tenzin Norbu, and Sonam Bhutia, as well as feedback from Erich Round and members of the Yale University Linguistics Department's Phonology Group and Historical/Contact Linguistics Reading Group. Fieldwork in Nepal was supported by a Pre-dissertation Fieldwork Grant from the South Asian Studies Council of The Macmillan Center at Yale University. All mistakes are my own.

processes: leveling and simplification<sup>1</sup>. In leveling, marked variants are reduced as speakers choose socially and/or structurally unmarked alternatives. Simplification, as defined here, involves an “increase in morphophonemic regularity” and an “increase in morphological and lexical transparency” (Trudgill 1986:103). In particular, following continued leveling processes at each generational stage, Trudgill notes that the koiné emerging in the third generation consists primarily of forms that are unmarked and common to many of the original input dialects.

In subsequent work, Trudgill (2000, 2004) has stressed the role of a variant's frequency across the first-generation speakers in its survival to the koiné stage, in an explicitly deterministic manner. Thus the resulting forms were often those of Southeast England, the largest contributing dialect region, but other forms also survived when they were common to enough of the other dialects in the original mixture. Failing to find evidence that early Anglophone New Zealand colonists selected variants that indexed particular social identities, Trudgill asserts that this kind of straightforward majority-rules framework explains the development of New Zealand English.

Trudgill's theory of new dialect formation has proven popular among linguists studying dialect contact. Kerswill and Williams (2000, *inter alia*) largely follow Trudgill in their analysis of the changing speech of the English “New Town” of Milton Keynes from its expansion in 1967 to the 1990s. While this community was not strictly formed *ex nihilo* as Anglophone New Zealand was, the original small village of Milton Keynes formed only a small portion of the population of the city. In particular, the first two “Principles of Dialect Contact” developed by Kerswill and Williams (2000) follow straightforwardly from Trudgill's theory: “Majority forms found in the mix, rather than minority forms, win out,” and “Marked regional forms are disfavored.” Importantly, they also find confirmation of Trudgill's timeline, which they distill as “From initial diffusion, focusing takes place over one or two generations.” In Milton Keynes, Kerswill and Williams find a small degree of focusing in the second generational phase--the first local-born children--but primarily observe focusing in the third generation, as Trudgill did in New Zealand. This finding qualifies Trudgill's theory by claiming that koineization can proceed somewhat more quickly, with focusing in the second generation.

Another study which draws heavily on Trudgill's model is that of Hornsby (2006), which studied changes in the regional French of the northern French town of Avion. There, with the growth of the town as a coal-mining center, a regional dialect developed that Hornsby interprets as a koiné of working-class Standard French and Picard, the formerly dominant language of the region. While often classed as different languages, Standard French and Picard are both *langue d'oïl* varieties, and so close that their interaction was one of dialect contact. Although Hornsby's interest lies primarily in the effect of changing social networks and more recent shifts toward Standard French rather than the timeframe of the development of Regional French in Avion, the approximate timeframe of 1890-1930 suggests an approximately generational scale to Trudgill's observations of New Zealand English.

**2.2 Critiques of Trudgill** The koineization model of Trudgill (1986, *inter alia*) outlined in the previous section has also come under criticism, in particular for its claim that the characteristics of a new dialect are purely the result of the ratios of variants present in the input dialects. The only exception to this is forms that are especially marked, which usually do not survive; otherwise, no other social or linguistic factors factor into the development of a new dialect.

Baxter et al. (2009) attack the claim of determinism with reference to Trudgill's own observation that New Zealand English had largely crystallized in three generations, as manifest in the speech of the second generation of settlers born in New Zealand. To do this, they created a mathematical model of speaker networks to estimate how long it would take for inter-speaker interaction to converge on a single, consistent form of a linguistic variable. This model of neutral interactor selection with a population the size of New Zealand requires “hundreds of generations” to come to completion, far longer than was observed in history. Indeed, follow-up work on mathematical models of linguistic change by Blythe and Croft (2012) indicates supports the notion that in order to achieve the S-shaped curve in variable frequency over time observed across sociolinguistic studies, “any theory of language change must put some kind of differential weighting of linguistic variants in a central role” (Blythe and Croft 2012:294). In other words, speakers necessarily exhibit asymmetric preference for one form of a linguistic variable over another, and change proceeds too quickly to rely on unweighted interactions across a social network.

This finding should come as no surprise, given the enormous amount of work that sociolinguists have put into learning how and why linguistic variables propagate through a population. To take only one

<sup>1</sup> A third, reallocation, involves retaining a feature used as a social marker and reanalyzing it to mark a different social category

example, Labov (2006) investigated the way speakers chose forms for a single variable on the basis of a complex web of interacting social pressures, identities, and aspirations. Indeed, Eckert (2000) has found these dynamics play out in quite sophisticated ways in even small communities. Although Mufwene (2008) agrees with Trudgill that a simple drive to create a group identity through similar speech characteristics does not explain the formation of a new dialect, he does still maintain that the choice of variants in a contact situation is affected by weighting factors. Nevertheless, in the case of dialect contact, he stresses inter-speaker accommodation and more deterministic factors, such as the Founder Principle (section 2.3, below). Still, the place of mechanistic vs. social factors in contact-induced change remains an important question.

**2.3 The founder principle** The selection by speakers of particular variants from the set of many possibilities present in the preceding generation, a crucial aspect of Trudgill's (1986) theory, closely resembles the idea of a "feature pool" proposed by Mufwene (2001, *inter alia*), a fact that Mufwene (2008) himself has noted. Originally developed to model the formation of creole languages, this idea holds that speakers select the linguistic variants of their speech from the full range of input forms they are exposed to; this process is said to apply in all cases of linguistic evolution across generations, from the small-scale changes within a language to the more dramatic cases of creole formation.

Mufwene (2001) discusses a corollary to this idea, the Founder Principle, which is adapted from biological evolutionary theory<sup>2</sup>. For a population with an initial founding population as well as subsequent immigration, the Founder Principle claims that the variants present among the founders exert a stronger effect on the later stages of the community than would be predicted from their numbers alone, as they set the initial conditions of the feature pool to which each subsequent wave of immigration contributes only a small portion at a time, spread out over a long duration. Thus, even if later immigrants bring substantially different features, the next generation will likely select forms present among the founders because these predominate in the feature pool to which they are exposed. The input of later immigrants is thus weakened through their arrival over time. Mufwene invokes this idea in the context of creole languages, which developed in colonies whose founding populations consisted largely of lower-class Europeans. Thus creoles exhibit many of the features of regional, working-class variants of their lexifier languages, rather than the prestigious and politically dominant standard varieties.

### 3 Tibetan in Diaspora

Of approximately six million total speakers of Tibetan, around 150,000 currently live outside traditional Tibetan regions in a diaspora which began in 1959, following the Chinese conquest of Tibet. While several thousand Tibetans now live outside South Asia, most Tibetans in diaspora live in India, Nepal, and Bhutan. Their communities form a web of interconnected enclaves: formally-designated settlements, monastic institutions, boarding schools, and urban neighborhoods. While they often use the Tibetan language among themselves, Tibetans are often multilingual, regularly using the dominant languages of their host countries such as Nepali, Hindi, or English, and many also speak Chinese (Denwood 1999, Roemer 2008). Although spread over a geographically wide area, Tibetans in diaspora frequently move between these areas and continue vibrant use of the Tibetan language, thus defining a speech community separate from those inside China.

Over the past six decades, several waves of migrants have come from all Tibetan regions of China into diaspora. However, Roemer (2008) reports that 70% of the first wave of 85,000 Tibetans to arrive in India and Nepal in 1959-1962 came primarily from the central U-Tsang region. Of the remainder, 25% came from the eastern region of Kham, and only 5% from the northeastern region of Amdo. Subsequent rates of immigration varied over time, but generally more balanced proportions of Tibetans from the three regions. U-Tsang, however, remains prominent in diaspora, particularly due to the presence of this historic capital city, Lhasa, and since it was the region where the Dalai Lamas and many other elites traditionally lived.

<sup>2</sup> Note that this differs from the "Founder Effect," also from evolutionary biology, whereby a daughter population can differ from its source population because of the random effects of small groups of founders. The application of the Founder Effect to historical linguistics has been critiqued by Bowerman (2011), and indeed the populations discussed in this paper have relatively large founding populations, making this sort of random drift less likely.

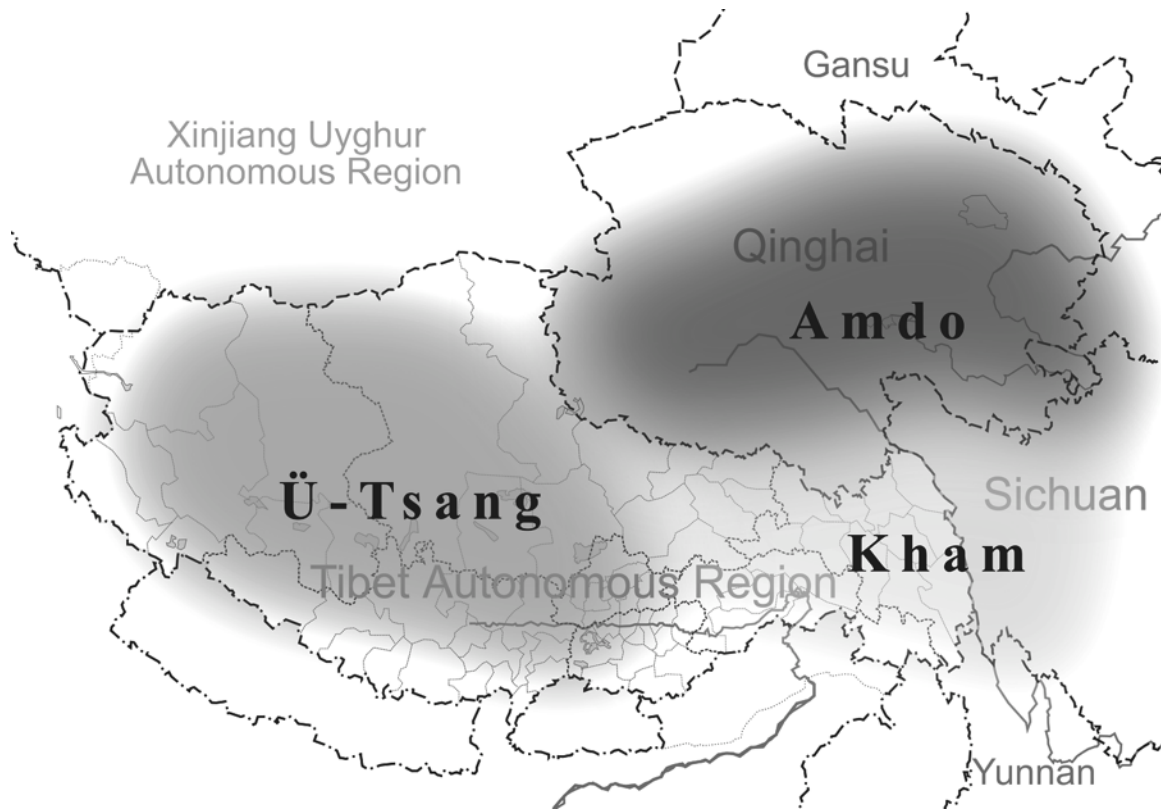


Figure 1: Traditional regions of Tibet, corresponding to broad macro-dialects. (Musser 2011).

The three large regions of U-Tsang, Kham, and Amdo are broad cultural regions, and all are composed of several sub-dialects. Nevertheless, that these three do correspond to macro-dialectal regions has been upheld by modern linguists such as Denwood (1999), Tournadre and Dorje (2003), and Dondrup Tsering (2011). Following such consensus, this paper treats the three as source dialects from which the Tibetan variety spoken in diaspora developed.

**3.1 Predictions for Diaspora Tibetan** According to the model of Trudgill (1986), the younger Tibetans raised in diaspora today--those born after about 1985--would correspond to the third generation of settlers, whose grandparents were the age of the founding diaspora population. Middle-aged Tibetans raised in diaspora would correspond to the second generation of settlers, the first raised in diaspora. These older speakers would be expected to exhibit a mix of features not corresponding to any particular Tibetan dialect, while the younger speakers should be well on their way to converging toward a new dialect unique to diaspora.

What would this new dialect look like? A deterministic theory that takes into account the Founder Principle would predict it to group with the U-Tsang dialects and to share features broadly common to many Tibetan dialects. In line with the notion of leveling, features closely associated with any particular regiolect should be absent; likewise, simplification predicts particularly complex forms or rules not to survive into the new dialect. Given the high social prestige of the Lhasa dialect, however, some features particularly characteristic of the Lhasa dialect may be apparent. This may be supported by the intuition common to many speakers interviewed that Tibetans in diaspora speak *dbus skad* 'U-Tsang dialect' or *lha sa skad* 'Lhasa dialect' (anonymous interviews, 2016).

## 4 Methods

Interviews for this study were conducted in the Kathmandu Valley, Nepal (henceforth, "Kathmandu") in the summer of 2016, one of several population centers of the Tibetan diaspora. Most interviews were conducted in the major Tibetan population center of Boudhanath, the neighborhood surrounding a major Buddhist pilgrimage site and home a substantial portion of Nepal's Tibetan community. This includes a number of monasteries and nunneries and two Tibetan settlements, as well as schools, shops, restaurants,

and small factories. Many Tibetans live, work, shop, attend school, worship, and engage in recreation in Boudhanath, which is near the edge of the urban agglomeration of Kathmandu; however, some Tibetans also live in other areas of the city and regularly travel around the city and valley.

As the author is obviously not a native speaker of Tibetan, all interviews were primarily conducted by one of two native speakers hired by the author, with the author present and offering occasional input or questions. As relative insiders, the interviewers stood a greater chance of eliciting naturalistic speech than the author alone, and helped mitigate against speaker accommodation to a clearly non-native speaker. Both interviewers were ethnically Tibetan, were born and raised in Nepal, and attended Tibetan schools through the “+2” (grade 12) level. One speaker was aged 23 and male, the other aged 19 and female; each conducted approximately half of the interviews.

Speakers to interview were identified through the author's and interviewer's social networks, largely friends-of-friends and extended family members. In several cases, the author and an interviewer visited a monastic institution, retirement home, or school following previous contact by the interviewer; once there, several residents, students, and/or staff were referred to be interviewed by members of that institution. One school, one retirement home, and seven monastic institutions in Kathmandu proper and the nearby town of Pharping were visited in this manner. The remainder of the interviews were conducted at the home of the speaker, at an interviewer's home, or at a neighboring monastery. Familiar locations and acquaintance relationships were used to elicit trust and comfort despite the presence of recording equipment and an unfamiliar foreign researcher.

Each interview began with speakers reading from a wordlist or, for the small number of speakers who were non-literate or had vision difficulties, equivalent lexical items were elicited by pictures in a slideshow shown on a laptop computer. This was followed by a choice task between light verbs (see section 4.1), a short reading passage, elicitation from storyboards, and finally a period of free conversation or narration according to the interests of the speaker and interviewer, but often involving discussion of the speaker's language attitudes. The speaker was provided with a paper copy of the wordlist, word choice task, and reading passage; the interviewer also asked for the speaker's preference for each item in the word choice. Interviews took between 20-60 minutes. Recordings were made on a Zoom H4N recorder with an Audio-Technica ATM73a headworn microphone and an Audio-Technica AT202 microphone on a small tripod.

Sound files were edited and reviewed in Praat (Boersma and Weemick 2011); while acoustic measurements were made, the categorization of speaker variables was made by the judgement of the author. An attempt was made to identify admixture using STRUCTURE (Pritchard et al 2000), while NeighborNets were generated using SplitsTree (Huson and Bryant 2006).

In total, 73 speakers were interviewed. Of these, 43 were male (of whom 14 were monastic) and 29 were female (of whom 6 were monastic). Basic demographic questions asked of speakers included what regions they were born in and had lived in, their educational history, and what languages they spoke. Before each interview, speakers were told they would remain anonymous, and in most cases the author did not know the speaker's name. For statistical purposes, speakers were classified by gender, and the region where they had spent the substantial majority of childhood and teenage years, and for speakers raised in diaspora, aged 18-30 or over 40. An “other” category used for certain tests consists of five speakers from ethnically-Tibetan regions outside Tibet or the Diaspora communities, four speakers in their thirties, and seven who moved between regions such that they could not be said to have been raised in either diaspora or one of the major Tibetan regions. The resulting demographic breakdown is represented below:

Category	Women	Men	Total	Phonological data	Verb-choice data
Diaspora, younger	9	8	17	17	14
Diaspora, older	2	6	8	8	7
U-Tsang	11	9	20	17	16
Kham	3	5	8	7	7
Amdo	1	2	3	3	3
Other	3	13	16	15	13

Note that, because several interviews were not complete, full data was not able to be elicited or analyzed for each speaker. The two rightmost columns in the above table show how many speakers' data of each category was analyzed; phonological data from 67 speakers was analyzed, as was verb-choice data from 60 speakers.

**4.1 Features** A range of phonological and lexical features were chosen in order to capture a general portrait of the state of each speaker's idelect. Of these, eight were phonological variables coded based on each speaker's pronunciation of the wordlist, and eleven were light verbs coded according to speakers' conscious judgements of their own speech. The phonological variables were selected based on the author's experience with the Tibetan language and reflect known variables in the literature such as reported in Tsering (2011); also following Tsering (2011), orthography was used as a proxy for etymological correspondences. The phonological features and their coding are as follows:

Phonological variable	Variants
Contrastive tone	present; absent
Voicing of onset plosives	no voicing; voice only orthographic <i>ga</i> -series clusters (e.g. <i>rga</i> ); voice all orthographic <i>ga</i> -series; other
Voicing of orthographic <i>z/zh</i> fricatives	present; absent
Realization of orthographic velar stop-palatal fricative clusters (i.e. <i>kj</i> -series)	cluster; affricate; fricative
Realization of orthographic labial stop-palatal fricative clusters (i.e. <i>pj</i> -series)	affricate, fricative
Vowel Harmony	present, absent
Nasal codas	prominent; absent/highly reduced
Realization of post-alveolar fricatives	0-3 scale of alveolo-palatal to palatal

In addition to the phonological features, the verbal choice task was based on a list compiled by a native-speaker scholar *g.yu sgang 'od zer*, titled *bya tshig sbyong thabs gsar ma grangs nges kun grok* (Dirk Schmidt, personal communication, July 10, 2016). Speakers were asked which sounds more natural to use in ordinary speech. Each item consisted of an object-verb collocation, where the verb could either be a more specific lexical verb or one of several frequent light verbs: *brgyab* 'strike', *bzos* 'make', *btang* 'send', and in one case, *shor* 'slip'. The hypothesis was that younger diaspora speakers would prefer the light verbs, while older and non-diaspora speakers would prefer the lexical verbs:

Translation	More-specific light verb	Less-specific light verb
'build a house'	<i>khang pa rtseg</i> 'stack a house'	<i>khang pa rgyab</i> 'strike a house'
'roll up paper'	<i>shog bu sgril</i> 'roll paper'	<i>shog bu ril ril bzos</i> 'make paper round'
'mail a letter'	<i>yi ke skur</i> 'mail a letter'	<i>yi ke btang</i> 'send a letter'
'play the lute'	<i>sgran snyan dkrol</i> 'strum a lute'	<i>sgra snyan btang</i> 'send a lute'
'dance'	<i>zhabs bro khrab</i> 'dance a dance'	<i>shabs bro rgyab</i> 'strike a dance'

Translation	More-specific light verb	Less-specific light verb
`release from prison'	<i>btson khang nas glod</i> `release from prison'	<i>btson khang nas btang</i> `send from prison'
`separate'	<i>khag khag dbye</i> `separate a difference'	<i>khag khag bzos</i> `make a difference'
`do circumambulations'	<i>skor ba skor</i> `circumambulate circumambulations'	<i>skor ba rgyab</i> `strike circumambulations'
`laugh'	<i>gad mo gad</i> `laugh a laugh'	<i>gad mo shor</i> `let-slip a laugh'
`have fun'	<i>skyid po byung</i> `receive enjoyment'	<i>skyid po btang</i> `send enjoyment'

**4.2 Analytic tools** The first analytic goal in this project was to determine which speakers group together according to the several variables. In order to do this, NeighborNets were generated using the SplitsTree. As discussed by Bryant and Moulton (2004) and Bryant et al. (2005), NeighborNets, which are primarily used in evolutionary biology, are also used in historical linguistics to develop networks of language families. The NeighborNet algorithm takes a matrix of data coded by taxa (e.g. species, languages) and computes the distances between the taxa. Each taxon is assigned a node, and each node is paired with two other nodes (rather than one, as in other related algorithms). These three nodes are then reduced to two, a process that is repeated to generate the splits that represent groupings of taxa. The distances involved in these splits are represented by branch lengths. Splits and branch lengths are depicted in graphical form, and while these are not evolutionary trees, the similarities constructed can be used to infer historical relationships.

In this study, the Neighbor-Net approach is still valuable as a tool to identify and depict groupings between taxa, in this case, individual speakers. Two types of information can be conveyed in these visualizations: branch lengths depict distance, the degree of difference between speakers, and the pattern of splits creates clusters of speakers with more similar traits. Following the predictions in section 3.1, this tool is predicted to find loose groupings of speakers for each of the three traditional regions as well as a cluster of younger Diaspora speakers, the latter perhaps within the U-Tsang cluster; older Diaspora speakers should appear scattered around the tree, in particular mixed among the U-Tsang speakers.

Another statistical tool borrowed from evolutionary biology, STRUCTURE (Pritchard et al 2000), was also used in an attempt to identify admixture between groups. This algorithm uses a Bayesian Markov Chain Monte Carlo estimation, randomly assigning taxa to a specified number of clusters (K), and continually re-assigns the taxa according to frequency estimates. The result is an estimate of the admixture of traits between populations, in this case, linguistic features. For these STRUCTURE runs, all 19 features were used, with a burning period of 10,000 and 50,000 repetitions, and K values from 2 to 10, with five runs per value of K.

## 5 Results

**5.1 Overall trends** Among the data, several phonological variables immediately stood out as patterning closely with particular regions. Firstly, vowel harmony was only found in four speakers. Of these, three of the four were born and raised in Lhasa, and the fourth, though coded as “other” for having spent time in both Kham and Lhasa, also spent a significant portion of her childhood in the city. Since none of the diaspora speakers exhibited vowel harmony, it can be said that this feature has not been transmitted past the first generation of migrants from Lhasa.

In contrast, a group of Kham and Amdo speakers clearly emerged with reference to other features. As coded, all 7 Kham and 3 Amdo speakers lacked tonal contrasts, while all 7 Kham speakers exhibited voicing in their orthographic <z> and <zh> fricatives; in these cases, no member of other groups had these feature values. Additionally, six of seven Kham speakers and two of three Amdo speakers were the only speakers recorded to produce orthographic bilabial stop + glide clusters as fricatives rather than affricates, as all other speakers did. These contribute to the divergence of Kham and Amdo speakers from the other groups, as shown below.

Neighbor-Net analysis of phonological and verb-choice features together failed to produce the expected clusters. Kham and Amdo speakers did indeed cluster together, separate from the U-Tsang and diaspora speakers, but other groups, including the expected younger diaspora group, were not found. This is shown in Figure 1.

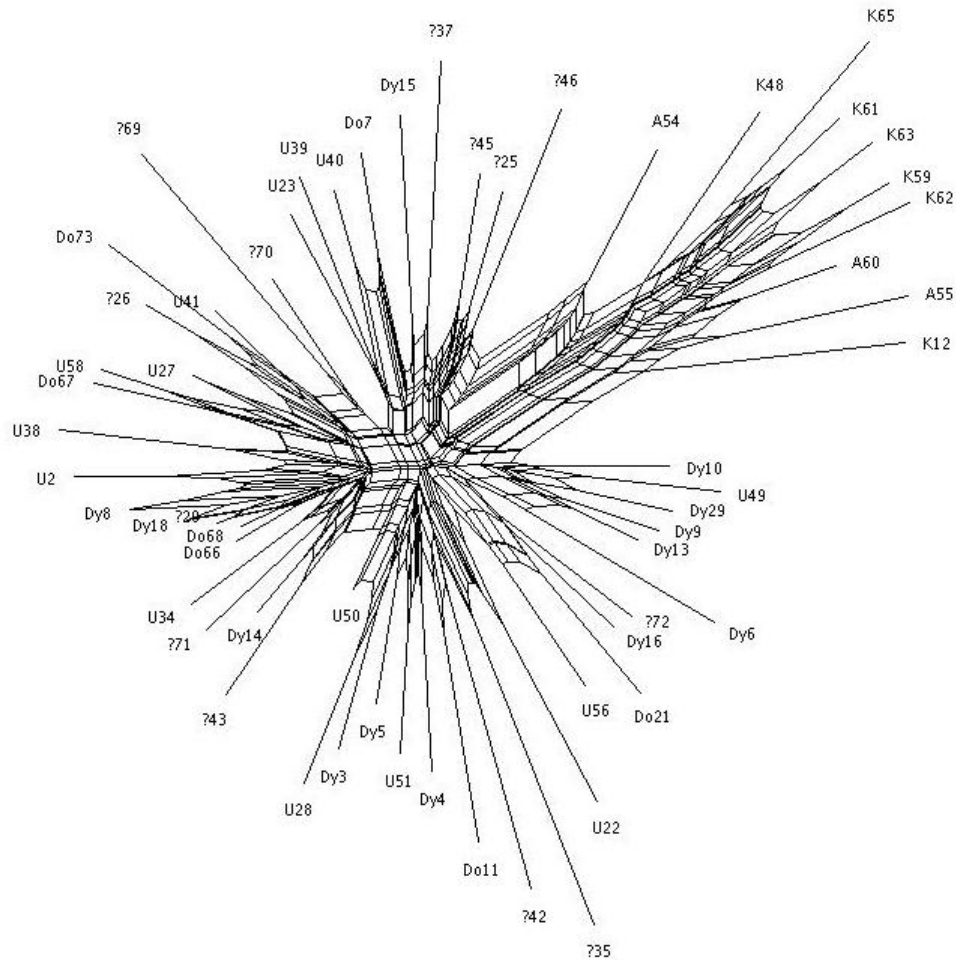


Figure 2: Neighbor Net of speakers by phonological and verb-choice data. Numbers reflect speaker codes; letter labels are as follows: U = U-Tsang, K = Kham, A = Amdo, Dy = younger diaspora, Do = older diaspora, ? = other.

However, when Neighbor Nets are generated by phonological and verb-choice data separately, it becomes clear that the separation of Kham and Amdo speakers from the rest is primarily on the basis of phonological data (Figure 2), not verb choice (Figure 3).



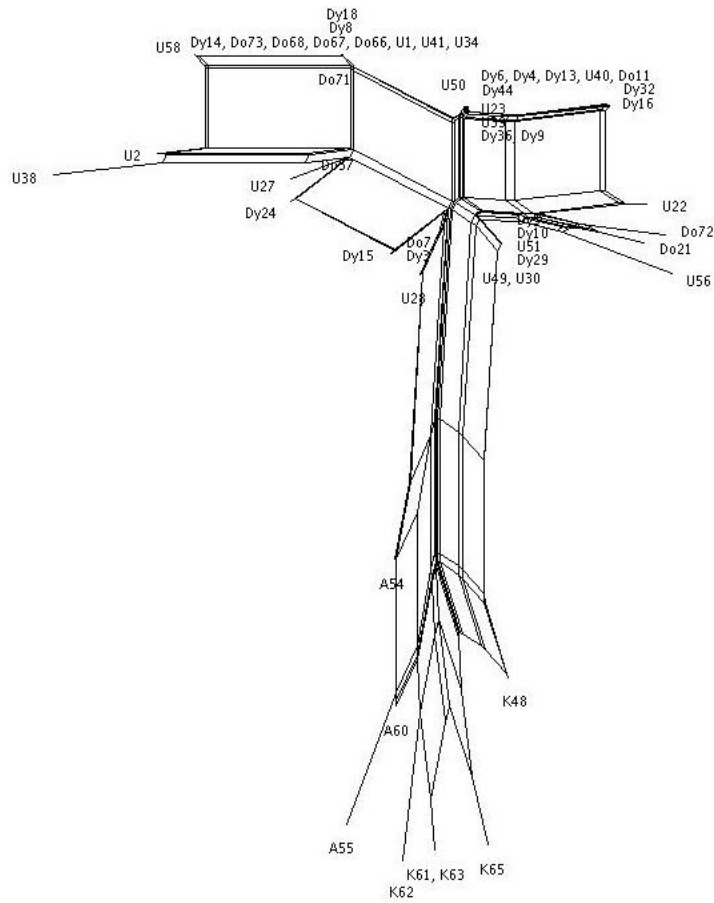


Figure 3: Neighbor Net of speakers by phonological data only. Numbers reflect speaker codes; letter labels are as follows: U = U-Tsang, K = Kham, A = Amdo, Dy = younger diaspora, Do = older diaspora.

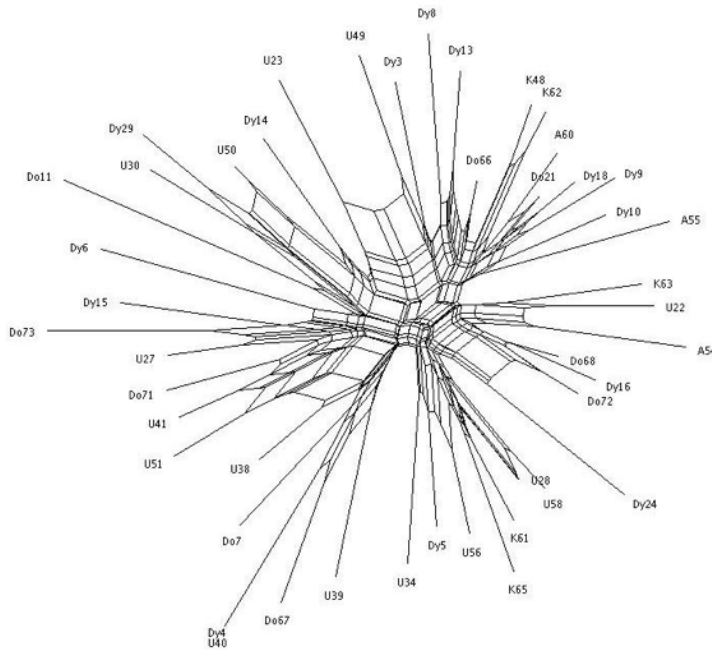


Figure 4: Neighbor Net of speakers by word-choice data only. Numbers reflect speaker codes; letter labels are as follows: U = U-Tsang, K = Kham, A = Amdo, Dy = younger diaspora, Do = older diaspora.

The highly un-tree-like image of Figure 4 suggests that this word-choice data is not a useful predictor of region. Indeed, the predicted simplification as lexical verbs are replaced by light verbs does not appear to be taking place, as speakers of all categories use a mix of both. The phonological data, however, does find a diffuse grouping of U-Tsang and diaspora speakers on the one hand and Kham and Amdo speakers on the other. This shows that the phonology of the diaspora speakers is based on that of the features of U-Tsang varieties, not those of Kham and Amdo. However, the failure to isolate a single cluster of younger diaspora speakers indicates that, with respect to these features, those speakers are not more similar to each other than are U-Tsang and diaspora speakers in general.

Results of STRUCTURE did not provide any additional information. Both admixture and no-admixture models consistently identified one group of Kham and Amdo speakers and another group of U-Tsang and Diaspora speakers, thus replicating the findings of the Neighbor-Net. For K-values above 2, no additional patterns emerged. This program thus failed to find evidence of admixture between the populations.

**5.2 Phonological variables** Of the phonological variables, several were associated with speakers from Amdo and Kham, and were not present among Diaspora speakers. Contrastive tone was present for all speakers except for those of Amdo and Kham, all of whom lacked it, as did one speaker of uncertain origin. Likewise, orthographic *pj*-series clusters were realized as affricates except for two of three Amdo speakers and six of seven Kham speakers, who realized them as fricatives. Thirdly, the voicing of *z/zh*-series fricatives was only found among six of seven Kham speakers and no others. Finally, the *kj*-series clusters were realized as fricatives only by three Kham speakers, and as affricates by the remaining four Kham speakers, all three Amdo speakers, and one U-Tsang speaker; for all others, including all Diaspora speakers, they were clusters.

Vowel Harmony, a process described as unique to the city of Lhasa by Denwood (1999), Dawson (1980), and others, was predictably only observed among three U-Tsang speakers, all from the Lhasa area, as well as a fourth speaker who had spent time there in childhood (that speakers' data was not counted on account of having grown up in several regions). Like *z/zh*-voicing, *kj*-lenition, and *pj*-fricativization, this appears to be a regional feature that was not adopted by any Diaspora speaker interviewed.

Onset Stop Voicing, while not as clearly regionally-marked, was not particularly instructive. 55 out of 68 speakers voiced only <rg>-series orthographic clusters. This included two out of three Amdo speakers, all seven Kham speakers, and nearly all of the Diaspora and U-Tsang speakers as well. One Amdo speaker exhibited voicing for all velar stops, while two younger Diaspora speakers and one U-Tsang speaker did not appear to use voicing contrasts in stops. Of the remaining categorized speakers, one older Diaspora and one U-Tsang speaker exhibited voicing inconsistent with the orthographic series discussed. The differences between older and younger Diaspora speakers was not statistically significant ( $p = .547$ ).

More variable across groups was Nasal Coda Reduction. While no Amdo and only one Kham speaker were coded as exhibiting reduced nasal codas, the numbers for U-Tsang and Diaspora populations were more mixed. eight out of seventeen U-Tsang speakers were coded with reduction, as were five of eight older Diaspora speakers and four of seventeen younger Diaspora speakers. The ratio of Nasal Coda Reduction among younger and older Diaspora speakers was found to be just above statistical significance ( $p = .087$ ), though a larger sample size could produce a more robust effect.

Palatalization of post-alveolar fricatives served in part to pick out Amdo and Kham speakers, which produced notably more palatalized forms. All three Amdo and six of seven Kham speakers produced forms that were rated “2” or “3” on a 0-3 scale. In contrast, this was the case for only two of eight older Diaspora speakers and three of seventeen younger Diaspora speakers (this difference was not significant). However, nine of seventeen U-Tsang speakers were also found to have a higher degree of palatalization, indicating the presence of diversity in this feature among some U-Tsang varieties as well as those of Kham and Amdo.

The phonological variables thus pattern differently across the demographic categories. Several, namely Tone, *pj*-fricativization, *kj*-lenition, and *z/zh*-voicing, were present among speakers of particular regions but not among any speakers from the Diaspora or U-Tsang. Similarly, Vowel Harmony was only found among speakers from Lhasa itself. Onset Stop Voicing did not pattern clearly by region, while Nasal Coda Reduction was found among a mix of all U-Tsang and Diaspora speakers. Finally, Palatalization of post-alveolar fricatives was rare among Diaspora speakers, but reliably observed among Kham and Amdo speakers as well as a portion of U-Tsang speakers. Taken together, these variables suggest that Diaspora speakers generally pattern with U-Tsang speakers rather than those from Kham and Amdo, but that there is little distinction drawn between younger and older Diaspora speakers.

**5.3 Light Verbs** Responses in this task varied; speakers sometimes indicated a simple preference for one or the other, or expressed no preference. Unsurprisingly, speakers often described a difference in register, with the light verb as more informal and the lexical verb as more formal or of written style. In other cases, speakers described a difference in meaning; for instance, eleven speakers said *yi ke skur* 'mail a letter' involves sending by post, while *yi ke btang* 'send a letter' has the letter delivered by a friend or messenger. Six options were coded for most: clear preference for one form or the other, preference expressed with equivocation for one form or the other, register difference, meaning difference, or no difference. For 'build a house' and 'have fun,' speakers expressed clearer judgements, so the "preference with equivocation" categories were omitted, and for 'have fun' another category was added to reflect different judgements with respect to tense.

Due to this diversity of judgements and the generally small sample sizes, however, comparisons between older and younger Diaspora speakers consistently failed to reach statistical significance, even when clear-preference and preference-with-equivocation judgments were grouped together. As with speakers of the other demographic categories, these speakers expressed a variety of judgements about each form, suggesting that the demographic categories chosen are not strong predictors of speakers' judgements.

Further complicating the situation, many speakers offered differences in meaning, register, or style to the two options. This was especially pronounced for 'have fun', where *skyid po byung* might be translated as 'enjoy oneself', while *skyid po btang* involved seeking out an enjoyable activity. Some speakers described a difference in terms of tense, likely building on the additional role of *byung* (the lexical verb for 'to receive') as a "receptive egophoric past tense auxiliary," according to Tournadre and Dorje (2005).

Among the two options for 'play the lute', the more-specific light verb *dkrol* 'strum' and the more-general light verb *btang* 'send', younger Diaspora speakers were the only category for whom no members chose the more-specific *dkrol* 'strum'. However, this choice was not particularly popular among the other groups either, with only two of seven older Diaspora speakers choosing it; the difference in the choice of *dkrol* 'strum' was nearly but not significant ( $p = .100$ ).

## 6 Conclusion

Taken together, these results do show a degree of support for the classical model of Trudgill (1986). The failure of Lhasa vowel harmony to survive transmission in diaspora follows from the simplification and loss of socially and linguistically marked variants. Likewise, the phonological features of diaspora speakers pattern with those of U-Tsang rather than Kham and Amdo, which can also be seen as a reduction of regionally-marked variants such as palatal fricatives and lenited affricates and clusters.

The dominance of U-Tsang features in their influence of the speech of diaspora Tibetans can also be seen as an example of the Founder Principle. As noted in section 3, 70% of the first wave of Tibetan migrants came from U-Tsang, and while more Kham and Amdo Tibetans have arrived since then, the Founder Principle predicts just this kind of outsized effect of U-Tsang Tibetan.

Still, the persistent diversity of the younger diaspora speakers remains a problem for Trudgill's theory. These young people, as the third generation in diaspora, do not appear nearly as focused in their speech, and a new dialect has yet to emerge. How can this be explained? Indeed, while Trudgill and Mufwene (2008) dismiss the role of group identity in new dialect formation, the Tibetan diaspora community, with strong ideologies favoring solidarity promoted by authority figures as well as ordinary people (Roemer 2008, anonymous interviews), should offer nearly as strong a case as any for which such social factors would encourage dialect focusing. Interestingly, while many Tibetans I spoke with stressed the importance of the preservation of their language, the emphasis seemed to be on the Tibetan language as opposed to other languages, not the interactions between dialects. This suggests that, perhaps, the diaspora Tibetan conversations on language ideology have not significantly addressed dialect diversity, thus allowing changes within the language to proceed with less social interference than might be expected.

The absence of a new diaspora Tibetan dialect after nearly sixty years offers a counterexample to Trudgill and his supporters, who propose that new-dialect emergence should be well underway in the third generation, if not before. Instead, these young Tibetans continue to speak with as much variation as their parents' generation, with each speaker using a novel combination of linguistic variants. The ideas of leveling, and, to a lesser extent, simplification proposed by Trudgill (1986) remain useful tools with which to examine linguistic changes in this community, but only time will tell how long focusing will take.

## References

- Baxter, Gareth J., Blythe, Richard A., Croft, William, & McKane, Alan J. (2009). Modeling language change: an evaluation of Trudgill's theory of the emergence of New Zealand English. *Language Variation and Change*, 21(02) 257-296.
- Blythe, Richard A., & Croft, William. (2012). S-curves and the mechanisms of propagation in language change. *Language*, 88(2), 269-304.
- Boersma, Paul & Weenink, David (2017). Praat: doing phonetics by computer [Computer program]. Retrieved from <http://www.praat.org/>
- Bowern, Claire. (2011). Out of Africa? The logic of phoneme inventories and founder effects. *Linguistic Typology*, 15(2), 207-216.
- Bryant, David, Filimon, Flavia, & Gray, Russell D. (2005). Untangling our past: languages, trees, splits and networks. *The evolution of cultural diversity: A phylogenetic approach*, 67-83.
- Bryant, David, & Moulton, Vincent. (2004). Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Molecular biology and evolution*, 21(2), 255-265.
- Dawson, Willa (1980). *Tibetan Phonology*. (Doctoral dissertation). UMI– Dissertations Publishing
- Denwood, P. (1999). *Tibetan* (Vol. 3). John Benjamins Publishing.
- Eckert, P. (2000). *Language variation as social practice: The linguistic construction of identity in Belten High*. Wiley-Blackwell.
- Geissler, Christopher (2016). *Vowel Harmony in Lhasa Tibetan*. Yale University Manuscript
- Giles, Howard, and Peter F. Powesland. (1975). *Speech style and social evaluation*. Academic Press.
- Hornsby, David. (2006). Redefining regional French: koinéization and dialect levelling in Northern France. *Legenda*.
- Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular biology and evolution*, 23(2), 254-267.
- Kerswill, Paul, & Williams, Ann. (2000). Creating a new town koine: children and language change in Milton Keynes. *Language in society*, 29(01), 65-115.
- Labov, W. (2006). *The social stratification of English in New York city*. Cambridge University Press.
- Musser, Karl. (2011). *Tibet\_provinces.png*. [image file]. Retrieved from <https://commons.wikimedia.org/>
- Mufwene, Salikoko. S. (2008). *Language evolution: Contact, competition and change*. Bloomsbury Publishing.
- Pritchard, Jonathan K., Matthew Stephens, and Peter Donnelly. "Inference of population structure using multilocus genotype data." *Genetics* 155.2 (2000): 945-959.
- Roemer, Stephanie. (2008). *The Tibetan government-in-exile: politics at large*. Routledge.
- Tournadre, Nicolas & Dorje, Sangda. (2003). *Manual of standard Tibetan*. Snow Lion Publications.
- Trudgill, Peter. (1986). *Dialects in contact*. Blackwell.
- Trudgill, Peter, Gordon, Elizabeth, Lewis, Gillian, & Maclagan, Margaret. (2000). Determinism in new-dialect formation and the genesis of New Zealand English. *Journal of Linguistics*, 36(02), 299-318.
- Trudgill, Peter. (2004). *New-dialect formation: The inevitability of colonial Englishes*. Oxford University Press, USA.
- Tsering, Dondrup (*don grub tshe ring*). (2011). *bod kyi yul skad rnam bshad*. Beijing: *krung go'i bod rig pa dpe skrun khang* [Tsering, Dondrup. 2011. *An examination of Tibetan regional dialects*. Chinese Tibetan Publishing Company]