

Locating Kera'a (Idu Mishmi) in Its Linguistic Neighbourhood: Evidence from Dialectology

Uta Reinöhl

1 Introduction

This paper is a first step towards locating Kera'a (otherwise known as “Idu” or “Idu Mishmi”) within its immediate linguistic environment by examining dialectological data. Kera'a, a language spoken in the very northeastern corner of Arunachal Pradesh (India)—and, by extension, in the northeastern-most corner of India—is a language for which little material is available to date. The same applies to the other two “Mishmi” languages that Kera'a is traditionally grouped with, Tawrä (aka “Digaru (Mishmi)”) and Kman (aka “Miju (Mishmi)”). As far as I am aware, next to no material has been provided that clearly distinguishes between the dialects or other sub-varieties of Kera'a so far. Disentangling the sociolinguistic landscape of Kera'a, as well as its ties to other Mishmi languages and beyond, is a first step towards understanding the linguistic as well as ethnic history of this remote and understudied area. Importantly, it will become clear that linguistic and ethnic relationships in this region often do not align in a straightforward manner. Exploring the role of language as a marker of dynamic social identity is thus key to understanding cultural dynamics past and present in this region.

This paper is based on comparative dialectal data collected in early 2017 by the author. In recent years, a small number of researchers have begun work on Mishmi languages. In this paper, I include targeted comparative data from Tawrä and Kman provided by colleagues, in addition to the Kera'a data collected by me.¹ As more material becomes available, we will hopefully soon be in a position to explore topics such as the one focused on in this paper on a much broader data foundation. Nonetheless, the relatively limited material available already suggests certain directions regarding the internal socio-lectal structure as well as the external relationships of Kera'a.

1 I thank Jonathan Evans and Johakso Manyu for kindly providing the comparative data on Tawrä, and François Jacquesson and Syndulum Ngadong for providing the same for Kman.

The suspicion already voiced by some writers (e.g. Post & Burling 2017) that Kera'a and Tawrã may be closely related while Kman does not show particularly close ties with either Kera'a or Tawrã is supported by the data reviewed in this paper. Thus, it appears at this point that Kman may link to Kera'a-Tawrã on a higher genealogical level, instead of presenting a direct sister variety. In beginning to disentangle the linguistic ties among the Mishmi varieties, dialectal evidence has proved exceedingly helpful, and constitutes crucial evidence for the argument presented here. Kera'a subdivides into two main dialects, the linguistically more progressive Midu and the more conservative Mithu. The latter is not only more conservative within Kera'a, but presents a stepping stone between Kera'a and Tawrã in several phonological domains. By taking into consideration this most conservative lect of Kera'a, i.e. Mithu, we can narrow down the linguistic gap to Tawrã significantly, and phonological correspondences emerge in much clearer light.

On a methodological level, this paper argues for an inside-out approach to the question of linguistic sub-grouping in Tibeto-Burman. The way to equip us better for the notorious challenge of disentangling Tibeto-Burman sub-grouping (see for instance the recent assessment in Post and Burling 2017) is not only to widen our knowledge of individual languages, but also to deepen our knowledge in terms of the historical layers within the individual languages.² For obvious reasons, the more conservative varieties within the various languages are the best stepping stones to the next of kin, as they have not yet diverged as much from a hypothetical common ancestral variety in comparison with more progressive lects. In view of rapid dialect levelling, language change and language loss given the spread of lingua francas such as Hindi, English, Assamese or other languages, putting a focus on the more conservative varieties—if still spoken—can provide a powerful tool in the enterprise of reaching a better understanding of the internal structure of Tibeto-Burman.

While I apply a branching model in this paper both with regard to the internal dialectal variation of Kera'a, as well as with regard to relations among, and tentatively beyond, the Mishmi languages, it must be emphasized that such an approach presents a highly simplistic model of the way in which dialects and languages develop, which cannot reflect the full complexity of actual devel-

2 Besides conservative dialects or clan-lects, another source potentially allowing insights into previous linguistic stages are the distinct linguistic varieties used in Shamanic rituals, which exist in many societies throughout the wider region. The Shamanic language of the Kera'a is currently investigated by the author of this paper.

opments. Whether a branching model may provide insights for a given group of varieties, or at least for a certain number of data points within them, is an empirical question. Normally, there will of course also be data points that can only be explained as having arisen through language contact. Even in language families that show evidence for the usefulness of the historical-comparative method such as Indo-European, not all data points and structures can be accounted for by the same (branching) model. Thus, while the evidence studied for this paper lends itself to a branching account, we cannot and should not assume that the latter will be able to accommodate all evidence unearthed in future work. In any case, it is subject to further research to determine the scope of material that may be accounted for in this way. Given what we know of the linguistic landscape within and surrounding Kera'a, I will reflect at the end of this paper on linguistic as well as ethnic parameters that may have played a role in bringing about what we know so far about the linguistic ties in and around Kera'a-speaking territory.

2 Tribes³ and Names

The Kera'a—commonly known as *Idu (Mishmi)*—are primarily concentrated in the Dibang Valley and Lower Dibang Valley districts of Arunachal Pradesh, the northeastern-most state of Northeastern India, see Map 9.1.

While the Kera'a form a single tribe, they are the equivalent of what was considered to be two tribes in the colonial British sources. These two groups correspond to what are today the two sub-groups of the Kera'a, the Midu and the Mithu. The Assamese and British perception of colonial times, believing the two groups to be separate tribes, seems to have been a misconception or at least an over-statement. It may have arisen due to the fact that the Assamese and British mainly dealt only with the Midu.⁴ The Midu are today the dominant group in the plains and lower foothills (i.e. in Lower Dibang Valley), in particular in Roing (the transport hub and economical centre of the Kera'a-speaking area). They are also the dominant group far up in the mountains around Anini,

3 The term “tribe” is the term in current use among indigenous populations of Northeastern India and is used with positive or at least neutral connotations. “Adi-vasi” (Sanskrit for ‘origin-dweller’), the official term used by the Indian government as well as by “mainland” Indian speakers for indigenous populations anywhere in India, is a term that the populations in this region do not identify with.

4 Many thanks to Razzeko Dele for clarifying this point for me, and for teaching me most of what I know of Kera'a history, as well as directing me to the pertinent historical sources.



MAP 9.1 Lower Dibang Valley and Dibang Valley (brown) in Arunachal Pradesh (yellow) and (North-east) India (beige) (Credit: Rebecca Gnüchtel)

close to the Tibetan border (i.e. in the upper part of Dibang Valley). The Mithu, by contrast, traditionally live in the middle ranges about half way up into the mountains of Kera'a territory, today clustering in the areas in and around Hunli and Desali.

The Midu were called “Chulikātā” (or “Chulikat(t)a”) by the Assamese and British, which is Assamese for ‘cropped hair’ (“kata”/“kātā” from the Sanskrit root *kṛt* ‘to cut’), a reference to the traditional haircut of the Kera'a, still worn by many members of the older generation today. “Chulikātā” is still known in the region as a name for the Kera'a, but is considered derogatory and is rejected by the community. The ethnologue code, however, remains *clk*. In the literature, “Chulikātā” has been equated by some with the Kera'a (as “Idu”) as a whole (see Sun 1993: 336, fn. 223), even though it was the Assamese (and British) term only for the Midu. One passage in Grierson (1909) suggests this erroneous

equation, as the three groups “the Chulikātā, the Digāru and the Mījū” (i.e. the Midu plus the other two Mishmi groups) are explicitly mentioned at one point as constituting the “Mishmis” (1909: 568), even though the wording suggests that these may not be all groups. Later in the book it is correctly specified that

The name Chulikātā is used by the Assamese in order to denote the tribe.
... They call themselves Midu ...

GRIERSON 1909: 613

The Mithu, by contrast, were referred to at the time as the “Bebejiya”, which is Assamese as is “Chulikātā” and means ‘outcast’:⁵

The Bebejiyas or outcast Mishmis ... Bebejiya is an Assamese name; they call themselves Mithun.

GRIERSON 1909: 613

With regard to the terrain in which the Midu and Mithu lived and live today, the contemporary situation departs slightly from the description in Grierson (1909), since the upper region around Anini is at least nowadays also regarded as Midu-speaking (with the clan-lect Mindri as an important variety within Midu in that area).

The Bebejiyas ... occupy the valleys of the Ithun River and its tributaries, between the Chulikātās and the Digārus. ... The Bebejiyas extend towards the high ranges of the Southern Tibet border on the north, and on the south they are bordered by the Chulikātās.

The linguistic and ethnic ties between the Midu and Mithu are briefly commented on in Grierson (1909):

The Bebejiya dialect is said to be almost identical with Chulikātā. The two tribes also agree in appearance and dress, and cut their hair in the same manner. They do not, however, intermarry.

GRIERSON 1909: 613

5 The notion that the Mithu were “outcast” once upon a time lives on in contemporary oral history, and is recounted also by Mithu speakers. It is said that the Mithu hid in the particularly steep mountain ranges of the upper region of today’s Lower Dibang Valley, facing persecution for certain actions.

Oral history does not only suggest that the Midu and Mithu always formed one tribe, but also that they did intermarry, in contrast to Grierson's assertion. This is in accordance with the far-reaching overlaps between the Midu and Mithu dialects. However, there are some notable features dividing the two varieties at least today, which I turn to in the following section.

On the topic of the names "Midu" and "Mithu", they involve, first, the element *mi-* 'person', cognates of which are wide-spread in Tibeto-Burman languages.⁶ In the case of Mithu, the second part of the name is similarly straight-forward, relating to the river Ithu, already mentioned in the quotation in Grierson above, which flows through the heartland of the Mithu area.⁷

While the etymological explanation of "Mithu" seems straightforward enough, the origins of "Midu" are less clear with regard to its second component. The latter may be *idu* (which would involve a contraction of the vowel, i.e. *mi- idu*) or perhaps more likely, *du*. (I discuss "Idu" as the wide-spread term for Kera'a language and culture below.). Two etymological connections for *idu* or, rather, *du* are conceivable at present, but there is no unequivocal support for either yet. I outline them in turn.

Firstly, several neighbouring varieties refer to the Kera'a with terms that resemble "Midu". To the west, the Adi call the Kera'a "Midi". To the northwest, the Milang refer to them as "Madə". Both these terms mean 'hill people' (Mark Post, in personal communication). At least some Tani words in *-i* have a cognate in *-u* in Kera'a, e.g. Bokar *iki* or Apatani *aki* 'dog' ~ *iku* (Midu).⁸ Thus, an equivalency between "Midi" or similar variants in certain Tani languages and "Midu" is in principle conceivable. However, note that at least some tribal names in the region that start with *mi-* are exonyms (such as e.g. "Milang" (Post and Modi 2011) or "Miju (Mishmi)"), so that a phonological correspondence along the lines just sketched would be of little help, if "Midu" wasn't in fact a Kera'a term. And indeed, at least "Idu", which is most likely derived from "Midu" (see below) is considered by the Kera'a an only recently introduced term.

6 *Mi-* is the most frequent allomorph, but the vowel shows some variation, e.g. *meto* 'chicken' for this domesticated animal. Following morphemes starting in *-d* show prenasalization (e.g. the clan lect "Mindri" (the variety spoken near the river Dri), or the name "Mindele" for members of the Dele clan).

7 In some literature, as in the Grierson quotation, the spelling is "Mithun" (or Mithun Mishmi, as in Matisoff 1996), which does not represent modern Kera'a pronunciation, however. Also, there is no connection with the bovine species "Mithun", which is of great ritual and economic significance in the region. It is conceivable that the similarity of the terms may have motivated the spelling "Mithun" as a folk etymology.

8 STEDT database. Department of Linguistics at the University of California, Berkeley. Online at <https://stedt.berkeley.edu/>.

Secondly, the Tawrä use the word “Dai” (without the *mi(n)*- prefix) for the Kera’a and the Kman call them “Mindau” (or “Mindao”). Recall from the remarks on “Mithu” and also “Mindri” (see fn. 6) that groups of people may be named after rivers, which is quite common in this region. In Kera’a, *du* means ‘clear water’. There is a river called “Dau” (Kman) or “Du” (Tawrä), which may involve the same element as is found in the Kman term “Mindau”, and which may correspond to Kera’a *du* ‘clear water’. Importantly, note that the Kman term “Mindau” suggests that we are indeed dealing with a morphological boundary between *mi(n)*- and *-du/dau*, and not a contraction between *mi* and a hypothetical *-idu*, given the presence of the homorganic nasal (now absent) separating the two morphemes (cp. fn. 6). Thus, the modern term Idu is possibly the secondary result of the productive loss of initial consonants in Midu (e.g., *iku* ‘dog’ in Midu corresponding to *miku* in Mithu)—which will be discussed in detail over the next sections. An extension of the term Idu from only designating the Midu to covering both the Midu and the Mithu does not come as a surprise given the power and prestige relations between the two groups.

While linguistically a plausible connection, the river in question flows through Anjaw district, which is Tawrä and Kman territory today. The Du/Dau is a tributary of the Lohit; the latter does figure prominently in Kera’a mythology, including in the route taken by the migrating soul of a deceased person as guided and chanted by the Shaman (*igu*) during *ya*, the death ritual (Dele 2017), so there may indeed be a connection here (see Dele 2017 for references to studies suggesting a historical migration of the Kera’a from the Lohit river into the Dibang river area).⁹ However, since there is no direct evidence that the Kera’a ever settled in the particular area around the river Du/Dau in question (or another one by that name), this etymology must also remain speculative at this point.¹⁰ If this link was further substantiated, the similarities with terms for the Midu in Tani languages might be independent exonyms for the Midu, possibly in connection with a folk etymological component given the phonological similarity.

The term “Kera’a” is considered by the members of the tribe to be the original autonym for today’s tribal group of Kera’a as a whole. This includes Midu- and Mithu-speaking clans, and it is the term used by the community for themselves

9 A close connection is also suggested by the fact that the Lohit bears the same, cognate name in Kera’a (*ilu*) and Tawrä (*telu*) (see section 4 on the correspondence between zero (Midu) and /t-/).

10 In terms of migratory routes, there is an old trail between the Mithu area around Desali and Tawrä/Kman territory, i.e. we are talking about connected, neighbouring areas.

when talking among themselves. The word is an ancestral term, being a compound of *kera* (the name of a mythological figure) and *a* 'child'.¹¹ Kera'a shares with Tawrã and Kman—also considered autonyms—the lack of the prefix *mi-*, which underscores that it is indeed probably the old autonym.

As mentioned above, "Idu" or "Idu Mishmi" is the widely used name for the Kera'a tribe and language by outsiders today. In the literature, "Idu" has implicitly been portrayed as the autonym in the absence of any mention of "Kera'a" and contrasting with the obvious exonym "Chulikātā". But the fact that Chulikātā is so obvious an exonym doesn't make Idu the autonym—at least not of the whole tribe including the Mithu. Kera'a speakers who have discussed this issue with me expressed the view that Idu is quite a recent name, whereas Kera'a is their old name. This fits a scenario where Idu (from Midu) was generalized for the whole tribe by outsiders, because the Midu as the group bordering the plains and therefore in direct contact with outsiders were taken as representative for both Midu and Mithu.

While the exact origins of "Midu" and "Idu" are not finally settled at this point—although especially the second etymology involving the Lohit region appears promising—what is clear is that "Idu" has not long been a term applying to the entire group of the Kera'a. In this paper, the whole group is referred to by the autonym Kera'a, even though the otherwise more widely used "Idu" or "Idu Mishmi" are not perceived as derogatory. I here follow the recent trend (e.g. in Post and Burling 2017) of switching to the autonyms of the neighbouring Mishmi languages despite the fact that "Digaru (Mishmi)" and "Miju (Mishmi)" are not considered derogatory either. While "Digaru", "Miju", and "Idu" may not be viewed negatively, the autonyms are viewed very positively, a crucial aspect taken into account here.

Coming now to the cover term for the Kera'a, Tawrã and Kman, the name "Mishmi" is of unclear origin. I am not aware of any convincing suggestion of a possibly etymology above and beyond the fact that we once more encounter the element *mi-*. In contrast to "Idu"/"Midu", however, the term

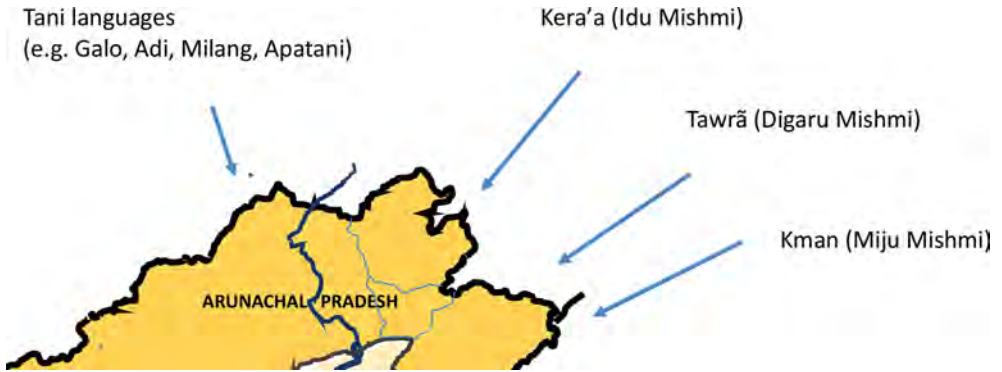
11 Alternative spellings used in the community include Kera-Ah, Kera-Aa, Kera Ah, and Kera Aa. I opt for the version Kera'a for the following reasons. There is no phonological vowel length distinction in Kera'a, and the spellings of Aa or Ah indicate a non-phonemic, lengthened realization of the vowel due to its occurrence as a mono-syllabic, mono-segmental word. Given that Kera'a is one prosodic word, I also choose to write the two words together rather than separated by a blank. A remaining alternative would be Kera-a (or Kera-A), which is not however a spelling that seems to be commonly used. Thus, I here use the same spelling as is also used by my primary language teacher, Usha Wallner, acknowledging the slight inconsistency of indicating the glottal stop onset by the apostrophe, despite it not being phonemic.

“Mishmi” has a long history in oral literature, where it is said to date back as far as the time of the Chutiya rule (around 1000AD). The first written attestation appears to be the inscription on the “Snake Pillar”, nowadays on display at the Guwahati State Museum, which dates to the early 16th century (cp. Phukan 1985).¹²

While the term Mishmi thus appears to be quite old, it is unclear what exactly defines a Mishmi culture or language as opposed to its neighbours. On the one hand, while identifying as separate tribes, the Tawrä and Kman partially form one ethnic group today in the sense that they intermarry and have a shared Shamanic culture, i.e. where the same language and chants are used in ritual. By contrast, the Tawrä and Kman neither traditionally intermarry with the Kera'a, nor do they seem to share the Shamanic language with the latter.¹³ There are several other ways in which there is a clear cultural difference between the Tawrä and Kman on the one hand, and the Kera'a on the other hand. The Kera'a are notorious for their extremely strict marriage taboos, where kinship relations not only on the paternal, but also on the maternal side need to be extremely distant in order for a marriage to be socially accepted. It is often asserted by members of the community that one may not be related up to ten generations in order to be able to marry. While this restriction may be weakened somewhat in reality, it is a fact that every person is related to a very significant number of other people in the community, and so eligible partners are few. At the same time, it is strongly preferred to marry inside the tribe even today. It has been speculated that these marriage taboos are a factor in the very high suicide rates in recent times among the younger Kera'a (cf. Tarun 2011). The Kera'a constitute a separate ethnic group from the Tawrä and Kman also in the domain of other taboos, such as regarding the consumption of meat. Traditionally, only Kera'a men are allowed to eat red meat; when they do, they must abstain from sexual intercourse for a week. Women are only permitted to consume fish and selected birds.

12 Note that it is not clear whether the early attestations of “Mishmi” refer to today’s three “Mishmi” tribes or to only one of them. Colonial as well as more recent writings discussing the history and prehistory of the Kera'a and of the other Mishmi tribes—and particularly where they may have migrated from—are summarized in Dele (2017).

13 Razzeko Dele (in personal communication) informs me there used to be a somewhat closer connection between Kera'a and Tawrä, e.g. in the form of an exemption from intermarriage taboos. There are also similarities between some clan names and regarding elements of Shamanic ritual. These points fit the linguistic evidence of a close historical relationship between Kera'a and Tawrä.



MAP 9.2 Kera'a and its neighbours

Not only the three societies in the northeastern corner of what is today Arunachal Pradesh were classified together as “Mishmi”, but this classification has long been carried over to their languages (see Map 9.2).

In earlier literature, the grouping of the Mishmi languages as a sub-branch was assumed uncritically. Still in 1985 does Sastry write in his “Mishmi grammar” that “Mishmi has three principle dialects ...” (Sastry 1985: 2). But even if this misrepresentation of the three languages as “dialects” is left aside, it is doubted in recent classifications that the three languages even present a linguistic sub-group. Instead, it has been suggested that Idu and Tawra may be closely related, but that Kman does not form part of the same sub-group (e.g. Post and Burling 2017, van Driem 2014). While the evidence available to date is still in great need of expansion and substantiation, I argue in this paper that this doubt can be supported. In particular, Idu and Tawra share considerable portions of lexicon which can be related by few sound changes. This does not seem to be the case for how these two languages relate to Kman, where lexemes often do not even appear to be cognate, at least given what we know at this stage. At the same time, despite the linguistic diversity more generally found in the wider region, I see no particular reason at this point to consider the three “Mishmi” languages to be isolates, as recently claimed by Blench (2017). At the very least, there is evidence that Kera'a and Tawra are closely related. Otherwise, the “Mishmi” languages do not seem any less Tibeto-Burman (TB) than other languages of the area, showing core TB vocabulary not only among content words, but also among function words (e.g., pronouns).

To sum up what we know about the cover term “Mishmi”: there are clear ethnic ties between Kman and Tawra, and clear linguistic ties between Tawra and Kera'a, but, at this point, there is no obvious way in which the three together

form a sub-group, whether ethnic or linguistic. At best, the three groups and languages can be defined negatively, in particular as clearly not belonging to the Tani languages to the west. “Mishmi” can otherwise simply be understood as a cover term for the tribes and languages located in the very far northeastern corner of Arunachal. In fact, a geographic understanding has for a long time been the other most common use of the term “Mishmi”, designating the section of the Eastern Himalayas where the three tribes are concentrated as the “Mishmi hills”, temporarily an official district after 1948 (e.g. Swain 2008: 66).

3 The Dialectal and Clanlectal Profile of Kera'a

I now discuss in some more detail the dialectal and clan-lectal profile of Kera'a. There is a clearly perceived dialect division between Midu and Mithu. While there is a lot of variation on all linguistic levels, this division manifests itself primarily in the phonetics of tone besides some salient lexical differences.¹⁴ Tone realization also varies significantly on the lower sociolinguistic level of clanlects, with a particular (patrilineal) clan belonging predominantly to either Midu and Mithu.¹⁵ In fact, the perception of a dialect division appears to be a generalization, bundling together Midu and Mithu clan-lects respectively, and stereotypically taking certain clan-lectal realizations as representative of the dialect as a whole. It is not unlikely that references to the dialects has to do with an entrenchment of Midu and Mithu as two separate groups and linguistic varieties through external perception in colonial times, as briefly suggested above. I return to the importance of the clan-lectal level below.

14 Kera'a has a four-way (perhaps even five-way in some clanlects), syllable-based tone system. The comparatively large toneme inventory in Kera'a fits Evans' (ms) analysis of four tonal distinctions for Tawrä. It contrasts however with most other languages of the region, which tend to have a two-tone system or lack tone. The tone system of Kera'a is also noteworthy because of its very high functional load, which is unusual for Northeast Indian languages. This is in part to do with the significant number of mono-syllabic words. Thus, every or virtually every syllable in the language tends to form three or more words as distinguished by tone. I do not mark tone in this paper as the clan-lectal and other sociolectal differences in phonetic realization—as well as whether we are consistently dealing with a 4-way system or not—still awaits further study.

15 While most clans are considered either Midu or Mithu, some include both members identifying as Midu and ones identifying as Mithu (e.g. the Mipi, Meto and Mega clans). Moreover, today's clans mythologically descend from certain original or 'root' clans (*asi*). In some cases, one root clan gave rise both to what are Midu and Mithu clans today, or ones associating with both groups. Thanks to Razzeko Dele (personal communication) for discussing these points with me.

The phonetic differences in tone realization function as salient markers of social identity. There is a perceived comprehensibility asymmetry between Midu and Mithu, which is often attributed to tone realization. Midu speakers sometimes claim that Mithu is difficult to understand (or virtually incomprehensible if they are not fully competent in Kera'a, but instead dominantly speak Arunachali Hindi, see below). By contrast, Mithu speakers (i.e. members of a Mithu clan) are generally not only able to understand Midu without difficulty, but will also switch to Midu fully or partially in many situations, or might even speak Midu as their primary variety if living, for instance, in Roing.

The outlined perception asymmetry is unsurprising given the sociolinguistic and geographic circumstances of the Midu and Mithu speaking communities. While traditionally inhabiting the mountains, many members of the community have settled in recent decades in Roing in the plains, just at the foot of the Mishmi hills in Lower Dibang Valley.¹⁶ While the population of Roing is mostly Kera'a, there are also speakers of several other tribal groups of the area, in particular Adi and also Galo, both Tani groups. Moreover, as in other towns of Northeast India, there is a significant number of "mainland" Indian officials working in public service (administration of the town and district, police, military etc.), and there are further minorities of e.g. Nepalese, Tibetan or Bihari origin. Hindi, or more precisely the contact variety Arunachali Hindi, is the primary lingua franca. Thus, a significant portion of the Midu-speaking population, which centres on Roing, speaks Hindi, or even English in some cases, every day. Arunachali Hindi is the primary language for many members of the younger generations (roughly, under 30yr olds). In terms of lifestyle, the younger generations do lead a life still somewhat in touch with their heritage identity—e.g. they will regularly experience Shamanic rituals or performances of traditional dance—but with strong influences of urban Indian lifestyle. Of the two dialects, it is Midu which is associated with these and other aspects of life in multi-ethnic and multi-lingual Roing.¹⁷

16 This migration into newly formed towns at the foot of the mountains has taken place all around the Brahmaputra valley in the last decades, with many families at the same time still retaining their traditional homes in the hills. Thus, while Roing is the main town for the Kera'a, Tezu to its south-east is the main hub for the Tawrä and Kman. To the west, Pasighat serves the same role for the Adi, and so on.

17 While the Roing area is predominantly Midu-speaking, there are some villages outside of Roing which are inhabited (almost) exclusively by members of Mithu clans, and where Mithu thus remains the predominant dialect. This includes Bhishmagnagar, Kurunu and Injonu.

By contrast, Mithu is predominantly spoken in the steep ranges of the upper region of Lower Dibang Valley in the comparatively much smaller communities of Hunli and Desali, and in numerous very small villages (sometimes consisting of one or two houses) scattered through the area, quite a few of which remain reachable only by foot. External influences can be felt here too, but the lifestyle by and large remains much closer to the traditional customs as regards housing construction, food supply, Shamanic culture etc. Midu speakers and small numbers of speakers of other languages can be found in this area, too, but the sociolinguistic diversity is much lower than in Roing.

Given these geographic and socioeconomic circumstances, it comes as no surprise that Midu is the linguistically more progressive variety, whereas Mithu retains certain conservative features, as will be outlined. In terms of language endangerment, Kera'a is endangered on the whole with fully competent speakers often no younger than around 30 at least in the larger communities. As for the dialects, the situations of Midu and Mithu are quite different. Whereas Midu is endangered predominantly by Arunachali Hindi, Mithu is in danger of being replaced by the prestige variety Midu as well as by Arunachali Hindi.¹⁸

Apart from the phonetic differences in tone realization, there is one other major phonological difference between the dialects, which is that certain consonantal onsets are retained in Mithu, while dropped in Midu in minimally two-syllable words. Some further segmental differences exist in that Mithu forms retain certain segments lost in Midu, however these patterns show a complexity that is not as well understood yet as that of the initial consonant loss.¹⁹ For the remainder of this paper, I will primarily focus on the presence vs. absence of consonantal onsets, both within Kera'a as well as in teasing out relations between Kera'a and its neighbouring varieties.

Kera'a has three stop series (bilabial, alveolar²⁰, velar) including voiceless unaspirated, voiceless aspirated, and voiced phonemes. Of these, the voice-

18 This is painting the picture in broad strokes. For a more detailed understanding of the viability of Midu and Mithu, several sociolinguistic variables play important roles, in particular area, size of the community, gender, and education level. On the whole, it is especially elderly women who are least exposed to varieties other than their home language, many of whom are monolingual.

19 Examples include the loss of nasalization on word-internal vowels, vowel assimilation, consonant-glide reduction to glides, and consonant deletion inter-vocally. Other variation is more symmetrical, e.g. involving different segments in the two varieties. There are also various lexical differences, the most-often cited one by speakers probably being the difference between Mithu *kadji ma?* 'where?' (literally, 'what in/at?'), where Midu speakers ask *hanu ma?*

20 Some minimal pairs suggest a further distinction between a dental and alveolar series in the lects of some speakers, a point awaiting further scrutiny.

less unaspirated stops are lost in Midu onsets, but retained in Mithu (with the evidence being substantial for /k/ and /p/, with some indications for it affecting also /t/). Moreover, not only /p/, but all bilabial onsets (/p/, /b/, /m/) are lost in Midu while retained in Mithu, again to the exclusion of the aspirated stop /ph/. Furthermore, the same loss affects the glottal fricative /h/ and the palatal glide /y/ (and there is some evidence that /r/ is also affected). All the bilabials are lost before /a/, /e/, and /i/, while the other consonants (including the glides) are lost before /a/. The following four tables illustrate the presence vs. absence of these consonantal onsets in Midu and Mithu. These correspondences affect at least many dozen if not hundreds of lexemes for onsets in /k/, bilabials, and /y/. The onset in /h/ is rarer in the data collected so far. (The Midu forms start with a glottal stop, which is not however marked here, it not being phonemic.)

As can be gleaned from Table 9.1 through Table 9.4, the lexemes affected belong to various semantic fields, including flora and fauna, body parts, and kinship terms. While the majority of lexemes seem to be nouns, there are also some verbs partaking in the correspondence sets. The seemingly lesser role of verbs appears connected to the fact that verbs are often mono-syllabic, and the phenomenon seems to be restricted to words that have minimally two syllables. For example, while there is a correspondence between *ata* (Midu) and *hata* (Mithu) for 'food', the corresponding verb *ha* 'to eat' is found across both dialects.²¹

In contrast to what we have just seen, a vocalic onset in Midu does not allow for the conclusion that the cognate in Mithu has a consonantal onset. For instance, the word for 'elephant' is *ata* in both varieties. Similarly, while *ili* 'pig' in Midu compares to *bili* in Mithu, both dialects share the form *iliku* 'down'. A plausible explanation would be that lexemes such as *ata* never had a consonantal onset, but this is speculation at this point. Comparative data from Tawrã will be especially useful for shedding light on this question. Note also that consonantal onsets are not lost in all cases in Midu, which has for instance *matshi* 'water' or *matsu* 'cow'. Also recent loans do not seem to partake in the correspondence (e.g. *mekari* 'cat' from Assamese in both Midu and Mithu, in contrast to the correspondence set of the indigenous word (*m*)*arjari*, also 'cat').

The lack of regular correspondences throughout the lexicon raises the question of whether we are indeed dealing with a phonological change or not. In

21 A morphological analysis is unlikely, as e.g. /m-/ is lost (e.g. in (*m*)*iku* 'dog'), even though speakers are aware that the whole syllable /mi/ is the meaning-bearing unit, 'person'.

TABLE 9.1 k- before -a

English	Kera'a (Midu)	Kera'a (Mithu)
'mosquito'	alondrō	kalondrō
'finger, toe'	atosu	katosu
'house mouse'	atʃɪŋgu	katʃɪŋgu
'house fly'	am ^w ebra	kam ^w ebra
'small cockroach'	alepĩ	kalepĩ
'bat'	aphe	kaphe

TABLE 9.2 h- before -a

English	Kera'a (Midu)	Kera'a (Mithu)
'food'	ata	hata
'cheek'	aphu	haphu

TABLE 9.3 Bilabials before -a, -e, -i

English	Kera'a (Midu)	Kera'a (Mithu)
'brother's sister'	athi	pathi
'ant (small, black)'	aroka	paroka
'pig'	ili	bili
'Russell's viper'	etambō	betambō
'tree'	asimbō	masimbō
'leaf'	asina	masina
'dog'	iku	miku
'chicken'	eto	meto

TABLE 9.4 y- before -a

English	Kera'a (Midu)	Kera'a (Mithu)
'monkey'	ami	yami
'bear'	ahũ	yahũ

TABLE 9.4 y- before -a (*cont.*)

English	Kera'a (Midu)	Kera'a (Mithu)
'Mishmi takin'	akrũ	yakrũ
'wild cat'	amrã	yamrã
'wild boar'	am ^w e	yam ^w e
'(long) squirrel'	adaŋgu	yadaŋgu

view of all factors, I lean at this point towards a phonological change despite apparent exceptions as just cited. Firstly, as stated above, a vowel-initial form in both varieties may have to do with the fact that there was never a consonantal onset to lose. Secondly, it is conspicuous that lexemes not partaking in the correspondence sets tend to be function words or high-frequency content words. Thus, I would tentatively consider a frequency effect, with lexical diffusion not having been completed. Thirdly, it is important to again consider the wide-spread shift of Mithu speakers towards Midu. Middle-aged and younger Mithu speakers can be uncertain of traditional Mithu pronunciations, as they are often primarily conversing in Midu themselves. Thus, vowel-initial forms in both varieties may in fact have a consonantal onset in the Mithu of elderly speakers. While I have worked with elderly speakers, the data collected predominantly come from speakers in their 30s and 40s. For similar reasons, it is also conceivable that high-frequency lexemes have been borrowed from Midu into Mithu, which would again be relevant for vowel-initial forms in both varieties. On balance, the significant scope of the correspondences at this point suggest a phonological phenomenon despite its limits. The following remarks on the importance of clan-lects suggest further important directions towards a better understanding of the variation.

So far, I have presented the variation internal to Kera'a as portraying a dialectal dichotomy between Midu and Mithu. This division has a foundation in the self-portrayal of the speakers, who regularly make reference to it. However, at closer inspection, it becomes clear that the primary socio-lectal level in Kera'a is that of the clan-lect, rather than the higher level of the dialect, the importance of which was already briefly mentioned in the context of tone realization. The correspondence pairs between Midu and Mithu shown above are based on an idealized categorical distinction. However, instead of a bundle of neat dialectal isoglosses, we are dealing with a continuum of clan-lects, which differ regarding how close they are to an idealized Midu standard (i.e. where consonants are dropped in all the environments outlined) or closer to an idealized,

i.e. very conservative Mithu, which retains consonants in all environments. For instance, the lect of the Dele clan, which belongs to Mithu, overlaps in several respects with standard Midu. As an approximation, Mithu clan-lects spoken further to the east (i.e. in the area around Desali) are more conservative (i.e. retain consonants in more types of environments and instances of use) than the varieties spoken in and around Hunli. This comes as no surprise considering that Hunli is the main hub on the main mountain road connecting Roing with the Kera'a hill regions, from which the Desali area is considerably removed.

While certain features are clearly recognized by the speakers I worked with as being associated with certain clan-lects, it must be noted that there is considerable variation in naturally produced data. Clan-lectal differences are often veiled by a shift from Mithu to Midu among members of the younger generations, among speakers who live in larger communities, among speakers that have close social ties with Midu speakers, and among speakers primarily identifying with life in the plains rather than traditional life in the hills. It is conceivable that frequency effects compound the issue, an investigation of which will however have to await the collection of larger data samples. This variation notwithstanding, the idealized distinction of forms associated with Midu and Mithu are identified uncontroversially as such by speakers. For this reason and to simplify the presentation, I rely on the higher-level dialect division here, too.

It is an open question at this point whether the dialectal distinction between Midu and Mithu will in fact be substantiated empirically or whether there is no relevant level between language and clan-lect, other than one serving speaker attitudes. The clear division between Midu and Mithu in terms of social identity—which has at least some historical depth given evidence from the British documents of colonial times—may have led to at least certain actual linguistic divisions between Midu and Mithu. However, it is not clear how prominent these are in comparison with the clearly important level of the clan-lect. At this point, it appears that we are dealing with a case not unlike that of the Dutch-German border area where standard varieties are clearly distinct, whereas spoken, non-standard varieties (used to) form a continuum.

For the remainder of this paper, I take the following approach: Kera'a can be sub-divided into two dialects, which have reality at least on the level of conscious socio-lectal differentiation, but which at the same time represent idealizations that are not in fact representative for the respective dialect as a whole. “Midu” and “Mithu” are shorthands in the remainder of this paper for an idealized Midu norm as spoken in Roing on the one hand, and for a conserva-

tive variety of Mithu as found in the eastern area of the Mithu-speaking region and as spoken by members of the older generations, on the other hand.

4 Kera'a and Tawrã (Digaru Mishmi)

Even though there is not much information on Tawrã available in published form, it was possible to collect comparative data tailored to the present study. The following tables include the above data from Midu and Mithu as well as the Tawrã equivalents, as found in Sastry (1985, marked as Tawrã (S)) and as provided by Jonathan Evans and Johakso Manyu ((Tawrã (E/M)).²² While most forms quoted above re-appear below, I excluded a couple of lexemes where the forms may not be cognate. In some cases, variation is likely to be orthographic (e.g. tʃ vs. ch). Otherwise, we find variation in particular where Tawrã preserves codas lost in Kera'a, some differences in vowels in the coda syllable, and variation in the choice of classifier, if present as a word-final, typically mono-syllabic morpheme (e.g. Kera'a (*m*)*asina* 'leaf' vs. (*m*)*asimbõ* 'tree' involving the classifiers *-na* for a flat, small-ish shape, and *-bõ* indicating a trunk-like form). Since the focus is here on consonantal onsets, variation at the right edge of the word forms is here not regarded as critical, so long as the initial part of the word shows the same or phonologically corresponding form and likely appears to be a cognate.

I begin in Table 9.5 with lexemes that in Mithu have onsets in /k/ or in a bilabial, but which lack this onset in Midu. We can see that this onset is retained in several lexemes that are probable or possible cognates in Tawrã.²³

22 Sastry (1985) is viewed critically by many linguists working in the region, one reason for which is the unclear title 'Mishmi Grammar'. However, at closer inspection it becomes clear that the book focuses on Tawrã: "The present work is based on the data collected in two field trips ... from the Digaru dialect" (1985: 5). Importantly, there are overlaps in the material in Sastry (1985) and more recently collected material by Jonathan Evans, and some by myself. For these reasons and since new material on Tawrã is only slowly becoming available, I include here forms from Sastry's study for comparative purposes. It is important to note, however, that it is not clear in either of the studies whether there is relevant variation within Tawrã, and if yes, what varieties are represented by the word forms collected. In a recent presentation by Evans, Manyu & Post (2019), a sub-division into three dialects is proposed.

23 The spelling suggests differences in aspiration in some forms, which calls into question whether we are indeed dealing with cognates. Since the documentation of Tawrã is in its infancy (at least in terms of published material), we do not yet understand variation within Tawrã, or between transcribers, which is why I chose to include these forms as potentially cognate.

TABLE 9.5 k- and bilabial onsets

English	Kera'a (Midu)	Kera'a (Mithu)	Tawrã (S)	Tawrã (E/M)
'house mouse'	atʃɪŋgu	katʃɪŋgu	–	kachey 'type of rat'
'small cockroach'	alepĩ	kalepĩ	–	khalũ
'bat'	aphe	kaphe	kapìn	kapũng
'pig'	ili	bili	bĩlig	baley
'tree'	asimbõ	masimbõ	masán	masang

TABLE 9.6 'Dog' and 'chicken'

English	Kera'a (Midu)	Kera'a (Mithu)	Tawrã (S)	Tawrã (E/M)
'dog'	iku	miku	kwág	kwak
'chicken'	eto	meto	tjú ('cock')	tju

TABLE 9.7 Three-way distinction

English	Kera'a (Midu)	Kera'a (Mithu)	Tawrã (S)	Tawrã (E/M)
'monkey'	ami	yami	tàmym	tamyum
'bear'	ahũ	yahũ	–	tahum
'wild boar'	am ^w e	yam ^w e	támè	tamè

With regard to bilabial onsets, and specifically onsets in *m-*, there are two data points which seem to contradict the evidence we just saw (see Table 9.5). In these examples, the Tawrã forms have not only no consonantal onset, but do not even have the vocalic onset typical of Midu. These examples, however, must be analysed differently. In Tawrã, we see direct reflexes of TB forms, while the prefix *mi-/me-* is added in Kera'a. One may speculate that the addition of this prefix in Kera'a is connected to the domestication of these types of animals, marking them as belonging to 'a person'.

Apart from the two data points shown in Table 9.6—which require a morphological rather than phonological analysis—we have so far seen the same onsets in Mithu and Tawrã. In Table 9.7, however, we see a three-way distinction of /t/ (Tawrã) corresponding to /y/ (Mithu) and zero (Midu).

TABLE 9.8 Overlaps between Midu and Tawrã

English	Kera'a (Midu)	Kera'a (Mithu)	Tawrã (S)	Tawrã (E/M)
'snake'	tabu	yabu	tabáb	tabaw
'meat'	tambrẽ	yambrẽ	tàbré	taabrẽR ¹

1 R indicates a rising tone.

In some cases where we find a correspondence between Tawrã /t/ and Mithu /y/, however, the Midu data point does not show the expected dropping of the consonant, but shows the same form as in Tawrã, i.e. an onset in /t/. Compare Table 9.8.

One may entertain various types of speculations to explain this pattern. For instance, the phonological differentiation mentioned in fn. 20 may play a role, i.e. we may not necessarily be dealing with one and the same segment /t/. Possibly, the phonological environments are not in fact equivalent, given the bilabial continuation of the forms in Table 9.8. Another possibility is that the lexemes in question were borrowed at some point from Tawrã into Midu. Again, it may be relevant that we are dealing with semantically general, highly frequent terms. While the two data points shown clearly require an explanation, which is not yet available at this point, it is worth noting that Midu *tabu* and *tambrẽ* are not only an exception to the specific correspondence set at issue, but also present an exception more generally to the pattern that consonants are lost in Midu. For this reason, I tentatively consider the three-way distinction shown in Table 9.7 to be evidence for a regular phonological correspondence in the same way as the cases of exact overlaps between Mithu and Tawrã in Table 9.5.

In contrast to the data points just cited, cases where Midu and Tawrã overlap lexically pose less of a challenge, as we are dealing with variation or overlap in a single data point, rather than with exceptions to an otherwise more general rule. For instance, the already quoted *hanu ma?* 'where' in Midu overlaps with Tawrã *hanu?*, contrasting with Mithu *kadjí ma?*.

Summarizing, when it comes to consonantal onsets, either Mithu and Tawrã share the onset which is lost in Midu, or there is a three-way correspondence set, with some exceptions where Midu rather than Mithu aligns with Tawrã. The phonological correspondences outlined may be taken as one piece of evidence for a close relationship between Kera'a and Tawrã, and will help reconstruct Proto-Kera'a-Tawrã. Above and beyond the suggestion of a close relation between Kera'a and Tawrã, the data may be interpreted as Mithu occupying an

intermediate position between Midu and Tawrä. While Mithu clearly forms a language together with Midu, in terms of a far-reaching overlap on all linguistic levels and in terms of mutual intelligibility, it is clearly the more conservative variety within Kera'a. Moreover, it shares certain features with Tawrä or shows what may be analysed as intermediate forms (namely the glide /y/ as the lenited form corresponding to Tawrä /t/). One may speculate that the geographical proximity of the Mithu-speaking regions with the Tawrä-speaking ones played a role in Mithu remaining more similar to Tawrä. The geographical proximity becomes especially clear when we take into account ancient trails (see fn. 10), with directly link the eastern areas of the Mithu area to Tawrä territory, in contrast to today's route through the plains.

5 Initial Consonant Loss in Typological Perspective

The loss of word-initial consonants in CV syllables is typologically rare and rather unexpected, in accordance with the notion that CV syllables are the least "marked" and most wide-spread syllable type. However, there is one area where this change is frequently found, which is in a significant number of indigenous languages of Australia, both Pama-Nyungan and Non-Pama-Nyungan. Blevins (2001) quoting Hale (1964) mentions that the phenomenon is found in at least 50 languages. Traditionally, initial consonant loss in Australian languages has been linked to a stress shift away from the first syllable, with lenition or loss of the initial C or the entire first syllable as consequence (Hale 1964: 256, Blevins 2001). There are, however, some languages for which a simple stress-shift scenario does not yield satisfactory answers including Ogunyjan and Oy kangand-Olgol (Alpher 1976: 86–7). In these languages, the vowel in the second syllable has undergone reduction and deletion, which would not be expected at least when applying a simple stress-shift account (Blevins 2001: 485). Thus, Blevins (2001: 485) instead proposes a two-step stress shift, where the stress ends up on the initial syllable again after having shifted to the second, which could explain the reduction of the second syllable.

Similarly to Ogunyjan and Oy kangand-Olgol, Kera'a shows no obvious indication of a stress shift. Midu and Mithu varieties of the same word show no rhythmic differences. In contrast to Ogunyjan and Oy kangand-Olgol, there is also no evidence for reduction or lenition of the vowel in the second syllable, so that there is also no support for two stress shifts having occurred. Similarly, no evidence suggesting a change in prosodic parsing direction from left to right as proposed for Umpithamu by Verstraete (2019) can be determined.

A stress-shift scenario would not be necessary if we were dealing with the reduced word forms having been borrowed from a neighbouring language, as opposed to resulting from an internal change. A borrowing scenario is advanced for some Cape York peninsula languages in Verstraete (2019). However, there is no language in the vicinity of Kera'a which closely resembles Midu and which could have been the donor of the phonologically reduced lexemes. The closest relative, Tawra, tends to show the long forms, not the reduced ones. It seems safe to say that a borrowing scenario can be excluded. **Midu** is thus a particularly intriguing case of initial consonant loss because there is no evidence whatsoever for a stress shift, nor for a change in prosodic parsing direction, nor for language contact having played a role (at least not in the form of borrowing). While a stress shift is plausibly connected to initial consonant loss in Australian languages, it thus remains an open question how to account for the phenomenon in Kera'a. It is also worth adding here that the consonant loss increases rather than decreases the number of homonyms—especially if tonal differences are regarded—which will make it hard to argue for any kind of functional explanation.

There are also interesting differences regarding the segments affected when comparing Kera'a with Australian languages. For Umpithamu, Verstraete (2019) mentions that peripheral consonants except *p are retained, whereas coronal consonants and *p are lost or lenited. In **Midu**, by contrast, some peripherals are lost including /k/ and the bilabials. Others are retained, such as the aspirated stops. Coronals do not appear to form a relevant natural class either. Thus, the phenomenon targets neither peripherals nor coronals in any systematic way in **Midu**, in contrast to Umpithamu (if disregarding the exceptional behaviour of *p in that language). As for peripherals, it does not come as a surprise that they form a relevant natural class in an Australian language. The feature +/- peripheral is known to form a natural class targeted by phonological processes in many Australian languages, but not normally in other languages. It is not expected to select a relevant natural class in other languages, and **Midu** indeed yields no evidence for it.

In conclusion, there is neither overlap in possible triggering factors for the initial consonant loss between **Midu** and Australian languages, nor in the segments affected at least when comparing **Midu** with Umpithamu. Future work will hopefully shed more light on the phenomenon of initial consonant loss, which does not easily fit with the assumption of CV as the basic, unmarked syllable type, especially in the absence of any obvious triggering factor as in **Midu**.

TABLE 9.9 Consonantal codas in Tawrä (Evans, ms)

/-p/	'elbow'	/la ^H kɪap ~ la ^H kɪaw/	lakrap ~ lakraw
/-k/	'son's child'	/a ^H jik ~ a ^H jej/	ayik ~ ayej
/-m/	'cloud'	/am ^H /	m
/-ŋ/	'house'	/aŋ ^L /	ng
/-j/	'buffalo'	/madʒaj ^H /	majay
/-w/	'young woman'	/ʔjã ^H ɪaw ~ ʔjã ^H ɪap /	qyāraw
/-u/	'soak'	/tsau ^H /	tsaü

6 Tawrä as the Most Conservative Variety within Kera'a-Tawrä

The preservation of consonantal onsets in Tawrä already points to Tawrä being more conservative than Kera'a in this particular phonological domain, especially in comparison to Midu. Tawrä also displays more conservative features in other respects. Thus, it does not only retain onsets where Midu has lost them, but it also retains consonantal codas, where there are none in Kera'a, neither in Midu nor in Mithu. Table 9.9 shows consonantal codas in Tawrä (from Evans, ms).

In the previous section, we already saw several cases where Tawrä retains a consonantal coda (e.g. *kwak* 'dog' in Table 9.6 or the retained nasals in Table 9.7). The next table (Table 9.10) shows some more examples of cognates where a consonantal coda (here: velar nasals) appears in Tawrä, but is lost in Kera'a.

When comparing the sets of personal pronouns, kin terms and numbers in Kera'a and Tawrä, further phonological reductions in the former in comparison with the latter can be observed, see Table 9.11. Besides the loss of word-final velar nasals, we also see the reduction of a consonant-glide combination (-*ty*), the loss of vowel nasalization (*nyũ*), reduction of a vowel-glide combination (-*ay*), and the absence of a vowel onset (*ane* vs. *ne*). At this point, it is not yet possible to ascertain the scope of these differences, i.e. whether they are restricted to the realm of high-frequency elements or are of more general scope, and how much they may be influenced by orthographic preferences of the person collecting the data. These data points, however, also support the impression that Tawrä is more conservative in its phonology in comparison with Kera'a.

While we see phonological conservatism in Tawrä, we see the occasional morphologically progressive change in Kera'a on the other hand. The example of *meto* 'chicken' (Mithu) vs. *tju* (Tawrä) discussed above shows, on the one hand, the maintenance of a consonant-glide combination in Tawrä, and on the other hand the morphologically progressive strategy in Mithu of adding the

TABLE 9.10 Examples for preservation and loss of codas

English	Kera'a	Tawrã
2sg	ni	niŋ
'three'	kasõ	kasəŋ
'yellow'	mi	miŋ

TABLE 9.11 Some core lexemes in Kera'a and Tawrã

English	Kera'a	Tawrã
2SG	nu	nyũ
1PL	ni	niŋ
2PL	ne	ane
father's father (FF)	nata	natya
'three'	kasõ	kasəŋ
'four'	kapri	kapray

morpheme *me-* (*-mi-*). Similarly, *kwak* 'dog' (Tawrã) shows phonological conservatism in its consonantal coda²⁴ in comparison with *mi-ku* (Mithu), while the later shows the morphologically extended form.

The same applies here as elsewhere, however, namely that future research will be necessary to determine whether these phenomena are sporadic or attest to more general patterns. Importantly, it is of course not necessarily the case that one variety is more conservative than another in every respect. With regard to the question of reconstruction, the evidence so far only allows for the determination of the hypothetical proto-forms in some cases. Thus, where we see an overlap between Mithu and Tawrã, we may safely reconstruct the proto-form, as in the case of onsets in /k/ and certain bilabials. However, in the case of the contrast of /t/ (Tawrã) and /y/ (Mithu), it may be speculated that the original

24 The labialization of the velar onset suggested by the spelling is not necessarily indicative of a more conservative form in comparison with Kera'a, as there is much synchronic variation with regard to labialization in the latter. The labialized variant does appear to be the more conservative pronunciation in Kera'a, but as it is not clear whether the Tawrã form is labialized in all instances, we cannot at this point draw strong conclusions.

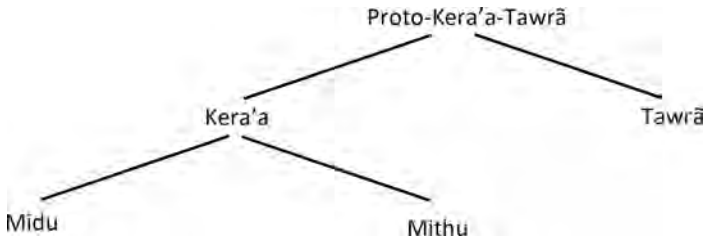


FIGURE 9.1 Kera'a-Tawrã (aka “Digarish”) based on selected phonological evidence

phoneme was /t/ (as attested in Tawrã) which lenited to /y/ in Kera'a (as attested in Mithu), and which was then lost in Mīdu. Alternatively, however, the original phoneme may have been neither /t/ nor /y/, but for instance /ty/ or yet another phonologically related form.

What does seem safe to say at this point is that Kera'a and Tawrã are most probably closely related, with many cognates as well as regular phonological correspondences. Cognates are not only found in the realm of content words, but also among function words, as illustrated with the set of pronominals above. The latter are not only clearly Tibeto-Burman, but also strongly resemble each other. Loosely speaking, Kera'a forms can be characterized as often somewhat phonologically reduced variants of the Tawrã forms. The branching of Kera'a-Tawrã as it appears at this point relating to the phonological data reviewed here is shown in Figure 9.1—bearing in mind that the scope of this tree structure, i.e. how much material it will be able to account for, remains to be determined.

Until we can say for sure that Kera'a and Tawrã form a sub-branch, at least with regard to the type of phonological evidence reviewed this paper, external evidence is of course necessary. Such evidence would need to show that Kera'a and Tawrã share a change, which is however not shared by other, more distantly related languages. Thus, I now turn to the wider linguistic landscape.

7 The Wider Linguistic Landscape

Despite the long-standing grouping of Kera'a, Tawrã and Kman as the “Mishmi” tribes and languages, more recent approaches tentatively group together only Kera'a and Tawrã (typically under the label “Digarish”), and consider Kman as part of a separate branch (e.g. Thurgood and LaPolla 2003; Post and Burling 2017). This new classification is convincing. On the face of it, there is little lexical overlap between Kera'a and Tawrã on the one hand and Kman

TABLE 9.12 Comparison with Kman

English	Kera'a (Midu)	Kera'a (Mithu)	Tawrã (S)	Tawrã (E/M)	Kman (J/Ng)
'monkey'	ami	yami	tàmyim	tamyum	amo'ŋ
'bear'	ahũ	yahũ	–	tahum	ku'm2
'Mishmi takin'	akrũ	yakrũ	–	takang	khyam
'wild boar'	am ^w e	yam ^w e	támè	tamě	tə-sə'm
'squirrel'	adaŋgu	yadaŋgu	tadyá	–	də'əu d'əu

on the other hand. Consider for instance some of the Kera'a-Tawrã cognates discussed in this paper with their corresponding Kman forms in Table 9.12. So far, no regular phonological correspondences that include Kman could be detected.

Some cases where Tawrã seems to be closer to Kman than to Kera'a need to be regarded with caution. I already discussed the forms for 'dog' above. Somewhat resemblant to Tawrã *kwak* 'dog', Kman has *kwi*, as opposed to Midu *iku* and Mithu *miku*. Since *mi-* in Kera'a is an addition—compare proposed Proto-Tibeto-Burman reconstructions such as **kwi*y or **kwe*y (Benedict 1972, Matisoff 1983)—it becomes clear that the Tawrã and Kman forms seem similar because the forms are Tibeto-Burman and lack the added morpheme of the Kera'a word, not necessarily because the two languages are particularly closely related. Certainly, there are certain resemblances in a number of lexemes between Kman on the one hand and Kera'a and Tawrã on the other hand, but there is so far no compelling evidence that would include Kman in a sub-branch with the other two languages. Instead, Kera'a-Tawrã and Kman are likely to be connected on a higher level—if again casting the relationship in the simplistic terms of a family tree—attesting to a certain amount of shared inherited material.

As for other languages of the region, it has been suggested that there might be a closer connection of Kera'a and Tawrã with the Tani languages (including/and/or Milang) e.g. Burling (2003: 181) or Post (2007: 95). Post writes:

There is very little evidence available to shed light on the nature of any post-Proto-Tibeto-Burman, pre-Proto-Tani ancestor language(s). **The nearest languages likely to be relatable are the “Mishmi” languages Idu, Digaru and Miju (in that order) ...**

POST 2007: 95, emphasis mine

The order of “*Idu*, *Digaru* and *Miju*” roughly mirrors geographical proximity. However, more recently, Post and Burling (2017) highlight a particularly strong resemblance with *Tawrã*, citing work by Sun (1993) and Modi (2013). This resemblance emerges in particular when *Milang*, a language associated and sometimes grouped with *Tani*, is included in the picture (Modi 2013).

It is clear that *Idu* and, especially, *Tawrã*, share a number of similarities both with *Tani* languages and with *Koro-Milang* (Sun 1993, Modi 2013); this is striking when one considers that it is *Idu*, not *Tawrã*, that is spoken in areas contiguous to the *Tani* languages. Nevertheless, it is not yet possible to determine whether this is indicative of a special genealogical relationship, or whether it may reflect earlier language contact.

POST and BURLING 2017, emphasis mine

Given what I have outlined in this paper, the greater similarity of the *Tani* languages with *Tawrã* over *Kera'a* does not come as a surprise. With *Tawrã* being more conservative than *Kera'a* in several phonological respects, *Tawrã* is expected to more closely resemble other languages which *Kera'a-Tawrã* connect to through a higher node. Thus, I tentatively confirm Post and Burling's (2017) speculation that the similarities they observe are indeed “indicative of a special genealogical relationship”. Note that “special” here does not so much relate to distance or proximity in terms of branching—as *Kera'a* and *Tawrã* may be equally close or distant from the *Tani* languages on this count—but to number of phonological (and possibly other) changes since a hypothetical shared ancestral language. Also note that the scenario just outlined does of course not preclude the possibility of “earlier language contact” also having played a role.

8 Conclusion

This paper sheds some light on the linguistic landscape of the Eastern Himalayas by focussing on the “*Mishmi*” languages and in particular on *Kera'a*. I argue that the recommended starting point for gaining insights into relationships between languages is their most conservative sub-varieties. While this may be a somewhat trivial point in the context of better-studied languages, it is often a challenge to cover more than the standard variety in research on lesser-studied languages due to constraints of wo/manpower, time, and finances. Nonetheless, families which defy straightforward genealogical classification to such an extent as *Tibeto-Burman* make such an approach rather desirable, if it can at all be realized. The more or most conservative variety naturally serves as the best

stepping stone to related languages by being closest to a hypothetical shared ancestral language. With regard to several phonological features, the most conservative clan-lectal varieties within Kera'a are spoken in the eastern area of the Mithu dialect group. These varieties retain consonantal onsets in environments where these onsets are lost in Midu, the more progressive dialect of Kera'a. The preservation of these onsets in Mithu narrows down the linguistic gap to the neighbouring language Tawrã, Kera'a's immediate geographical neighbour to the east and south-east. In several cases, Mithu overlaps exactly with Tawrã, or it shows what appears to be an intermediate form between Tawrã and Midu in terms of phonological features. On the whole, Tawrã appears most conservative within Kera'a-Tawrã, also with regard to other phonological features, such as in its coda phonology. I conclude that Kera'a and Tawrã show clear signs of a close relationship, with Tawrã as the variety that is most conservative. The disentangling of the internal dialectal (and clan-lectal) landscape of Kera'a considerably helped in reaching this conclusion.

The recent views that the Mishmi languages may not in fact constitute a sub-branch and that, instead, Kera'a and Tawrã are closely related, but Kman is a more distant relative, are supported by the data presented here. However, it must be stressed that much research on the three Mishmi languages is still necessary to place these claims on firmer empirical grounds. As for other external relationships, the observations by some researchers of a particular close resemblance of the Tani languages with Tawrã, rather than with their geographical neighbour Kera'a, can be explained by Tawrã having undergone fewer changes since its departure from the hypothetical ancestral variety Proto-Kera'a-Tawrã.

In the remainder of this conclusion, I would like to return to the topic raised at the beginning of this paper, namely the question of what the data reviewed here can tell us with regard to the ethnolinguistic history of the area. The first observation I would like to make relates specifically to Kera'a, while the second observation concerns an ethnolinguistic parameter that is more generally relevant for the wider region.

Firstly, it is notable that there is so far not much evidence for convergent change among the Mishmi languages, or at least not regarding Kera'a. There seems to be no or hardly any lateral borrowing of elements or structures from the other two Mishmi languages into Kera'a. This is also true for the apparent absence of borrowings from other, non-Mishmi tribal languages. Mark Post (personal communication) informs me that, similarly, there seems to be no or next to no borrowing from Kera'a into neighbouring Tani languages, while at least some of the Tani languages do borrow among themselves. Folklore in the region has it that Kera'a is particularly hard to learn in comparison to other tribal languages, primarily due to its complex tone system and the sig-

nificant number of segmental homophones, many of which monomorphemic, that only differ in tone. Any lexemes that have been borrowed into Kera'a seem recent and come predominantly from Hindi, Assamese or English. Structurally, however similar, Kera'a and Tawrã are clearly developing apart, at least phonologically. In section 3 I mentioned the multi-ethnic and multi-linguistic environment which exists today in particular in Roing. While speakers may now switch mid-sentence from Kera'a to Arunachali Hindi or English, there are signs that multi-lingualism is recent. At least some members of the older generations, particularly those who have remained in the mountain villages, seem to be mostly monolingual, with very limited or no knowledge even of Arunachali Hindi, Assamese, Tawrã, Kman or Adi. This applies in particular to elderly women, who often remain monolingual even if now living in Roing. Traditionally, the steep Mishmi mountain ranges may have discouraged local migration apart from individual trading or hunting expeditions, preventing widespread bi- or multi-lingualism. There is also no traditional hegemonial power in the region which would have imposed its language as a lingua franca. A culture of predominantly monolingual communities fits well with the impression of a lack of convergent change. This suggests that Kera'a qualifies as an esoteric society with little contact-induced change (DeLancey 2014); such a categorization however will require more research in view of the fact that it does not show (or retain) particularly complex morphology.

Coming to the second point, a noteworthy feature of Northeast Indian societies is that tribal identity is negotiable, at least since British colonial times. In section 2 I outlined that the Mithu and Midu were considered two distinct groups, if not tribes, at least for a certain period in previous times, but today unequivocally identify as a single tribe. Similar mergers of tribes, even where languages are not quite so similar, have occurred more recently or are currently ongoing in other corners of the region (see e.g. Post & Burling 2017). Importantly and strikingly, it seems that language is not an impeding factor in these mergers. While still identifying as separate tribes, the fact that Tawrã and Kman intermarry and share a Shamanic culture is illustration for such a (partial) ethnic merger, even while the languages are quite different and not mutually intelligible. The fact that tribal identity can be re-aligned independent of language will crucially shape how the languages develop, as it impacts on language contact scenarios. Besides the possibility of a role of substrate influence or other types of language contact (e.g. Blench & Post, ms), this relationship between language and tribe may help explain some of the significant linguistic diversity that we find in the region, where language may change or may have changed in a way which is less impeded by social identity, nor held in check by convergent forces of bi- or multi-lingualism.

Acknowledgements

I want to express my deepest gratitude to Usha Wallner, who has been the most amazing friend and support in all matters linguistic, cultural and personal. I am also most grateful to Anoko Mega, Razzeko Dele, Rajkumar Keche and Pinda Pulu for teaching me their language and culture, and for their kind hospitality. Many thanks are also due to many others including Jyoti Linggi, Lomi Linggi, Mengo Dele and Disiya Dele. Special thanks are due to Rebecca and Kaling Tamut who have been most generous and have become true friends. I would like to thank Mark Post, Yankee Modi and Scott DeLancey for their help and support in setting up a research project on one of the Mishmi languages. Particular thanks are due to Mark Post, Razzeko Dele and Usha Wallner for their valuable comments on this paper. I thank the audiences at the ICEHEP 3 workshop, Feb 2017, and at the Centre of Excellence for the Dynamics of Language (CoEDL, ANU node), April 2017, for stimulating comments. Many thanks to the Alexander von Humboldt Foundation, to CoEDL, and to the German Research Foundation (Emmy Noether project on “Non-hierarchicality in Grammar”, project no. 406074683) for sponsoring my research on Kera'a.

References

- Alpher, Barry. 1976. Some linguistic innovations in Cape York and their structural correlates. In Peter Sutton (ed.), *Languages of Cape York. Australian Aboriginal Studies, Research and Regional Studies No. 6*. Canberra, Australian Institute for Aboriginal Studies: 84–101.
- Benedict, Paul K. 1972. *Sino-Tibetan: A conspectus*. (Princeton-Cambridge Series in Chinese Linguistics, #2). Cambridge, Cambridge University Press.
- Blench, Roger. 2017. The “Mishmi” languages, Idu, Tawra and Kman: A mismatch between cultural and linguistic relations, Presentation at the International Consortium for Eastern Himalayan Ethnolinguistic Prehistory, La Trobe University, Melbourne, Feb 8th–10th 2017.
- Blench, Roger & Mark Post. ms. (De)classifying Arunachal languages: reconsidering the evidence.
- Blevins, J. 2001. Where have all the onsets gone? Initial Consonant Loss in Australian Aboriginal Languages. In Jane Simpson, David Nash, Mary Laughren, Peter Austin, & Barry Alpher (eds.) *Forty years on: Ken Hale and Australian languages*. Canberra, The Australian National University: 481–492.
- Burling, Robbins. 2003. The Tibeto-Burman languages of Northeastern India. In Gra-

- ham Thurgood & Randy LaPolla (eds.), *The Sino-Tibetan Languages*, 169–192. London, Routledge. <https://doi.org/10.4324/9781315399508>
- DeLancey. 2014. Sociolinguistic typology in North East India: A tale of two branches. *Journal of South Asian Languages and Linguistics* 1 (1): 59–82. <https://doi.org/10.1515/jsall-2014-0004>
- Dele, Razzeko. 2017. *Ya. Idu Mishmi Shamanic funeral ritual*. Darya Ganj, IN, Bookwell Publications.
- Driem, George van. 2014. Trans-Himalayan. In Nathan Hill & Thomas Owen-Smith (eds.), *Trans-Himalayan Linguistics*, 11–40. Berlin, de Gruyter. <https://doi.org/10.1515/9783110310832.11>
- Evans, Jonathan. Tawrā (Digaru Mishmi). Ms.
- Evans, Jonathan, Johakso Manyu & Mark W. Post. 2019. Ta-wrā phonetics and phonology in areal and historical contexts. 52nd International Conference on Sino-Tibetan Languages and Linguistics. University of Sydney, June 24–26.
- Grierson, G.A. (ed.). 1909. *Linguistic Survey of India. Volume III, 1, Tibeto-Burman family*. Calcutta, Superintendent of Government Printing.
- Hale, Kenneth. 1964. Classification of Northern Paman Languages, Cape York Peninsula, Australia: A research report, *Oceanic Linguistics* 3: 248–265.
- Matisoff, James A. 1983. Translucent insights: A look at Proto-Sino-Tibetan through Gordon H. Luce's comparative word-list. *Bulletin of the School of Oriental and African Studies, University of London* 46(3): 462–476. <https://doi.org/10.1017/S0041977X00043172>
- Matisoff, James A. 1996. *Languages and Dialects of Tibeto-Burman*. Berkeley, CA, University of California Center for Southeast Asia Studies
- Modi, Yankee. 2013. The nearest relatives of the Tani group. Paper presented at the 19th Himalayan Languages Symposium. Canberra, Australian National University.
- Phukan, J.N. 1985. The date of Sadiya pillar inscription. *NEIHA proceedings, 6th session*. Shillong: 85–89.
- Post, Mark W. 2007. *A Grammar of Galo*. Melbourne, La Trobe University dissertation.
- Post, Mark W. & Robbins Burling. 2017. The Tibeto-Burman languages of Northeastern India. In Graham Thurgood & Randy J. LaPolla (eds.), *The Sino-Tibetan Languages*, 2nd edn. London, Routledge: 213–242.
- Post, Mark W. & Yankee Modi. 2011. Language contact and the genetic position of Milang (Eastern Himalaya). *Anthropological Linguistics* 53(3): 215–258. <https://doi.org/10.1353/anl.2011.0014>
- Sastry, G. Devi Prasada. 1985. *Mishmi Grammar*. Mysore, Central Institute of Indian Languages.
- Sun, T.-S.J. 1993. *A Historical-Comparative Study of the Tani Branch of Tibeto-Burman*. Berkeley, Ca, University of California dissertation.

- Swain, Pratap Chandra. 2008. *Panchajati Raj: The grassroots dynamics of Arunachal Pradesh*. New Delhi, A.P.H. Publishing Corporation.
- Tarun, Mene. 2011. *Suicide among the Idu Mishmi tribe of Arunachal Pradesh*. Itanagar, IN, Rajiv Gandhi University dissertation.
- Thurgood, Graham & Randy J. LaPolla (eds.). 2003. *The Sino-Tibetan Languages*. London, Routledge.
- Verstraete, Jean-Christophe. 2019. Initial consonant loss and its consequences: Evidence from Cape York Peninsula (Australia). Presentation at Ninth European Australianist Workshop, KU Leuven, 18–19 April 2019.