Malayic varieties of Kelantan and Terengganu: description and linguistic history

Wu, J.

CHAPTER 9

General discussion and conclusions

9.1 Overview

In this dissertation, I have examined the Malayic varieties spoken in Kelantan and Terengganu (NEPMs) with two goals: first, to provide sketch grammars of NEPMs (presented in Chapters 2 to 6), and second, to investigate NEPMs' historical development (discussed in Chapters 7 and 8). In this concluding chapter, I summarise and synthesise the main findings in §9.2. Furthermore, I explore the implications of the linguistic data for the Malayic migration history in §9.3 by placing the linguistic findings against the non-linguistic background summarised in §1.4.1. In §9.4, I acknowledge some limitations in the present study and point at some directions for future research.

9.2 Summary of main findings

In the phonology, all three varieties of NEPMs share a consonant inventory comprised of twenty native phonemic consonants. These include nine stops /p, b, t, d, c, j, k, g, ʔ/ (among which /t/ is dental [t̪], /c/ and /j/ are phonetic-

ally affricates), four nasals /m, n, ɲ, ŋ/, three fricatives /s, x, h/, two liquids /l, ɣ/ and two glides /w, y/. This consonant inventory is typical of Malayic languages, except for the voiceless velar fricative /x/, which originated from the reciprocal assimilation of the two segments making up earlier ⁺kɣ- clusters. As a result, /x/ only occurs word-initially, primarily in the geminated form /xx-/. Geminate consonants, which are analysed as clusters consisting of two identical segments, stand out as a remarkable feature in the consonant systems of NEPMs. Almost all consonants, except for glottals and glides, can appear in their geminated form and are exclusively attested in word-initial position. CTM also has a geminated bilabial glide /ww-/, resulting from the reciprocal assimilation within earlier ⁺bɣ- clusters (thus parallel to ⁺kɣ- > /xx-/). These geminate clusters are characterised by a longer duration of articulation, and they can often be contrasted with their singleton counterparts in (near-)minimal pairs. They occasionally display variation with non-geminate clusters, from which they arose diachronically. Moreover, geminate clusters can occur in both morphologically simple and complex words. In the latter case, they can often be analysed as having derived from the morphophonological process of initial gemination.

The vowel systems of NEPMs differ more significantly. KM and CTM make a four-way distinction with regard to vowel height, each having eight oral vowels /a, i, u, ə, ɛ, ɔ, e, o/. Additionally, KM has four nasal vowels /ã, ũ, ɛ̃, ɔ̃/, and CTM has five nasal vowels /ã, ĩ, ũ, ɛ̃, ɔ̃/. ITM, as represented by the subvariety spoken in Kampung Dusun, lacks the set of mid-high vowels, thus having only six oral vowels /a, i, u, ə, ɛ, ɔ/. It also features four nasal vowels /ã, ĩ, ũ, ɔ̃/. ITM further differs from KM and CTM in having two phonemic diphthongs /ɛi, əʊ/, and its vowel phonemes generally display more complex allophonic variation, conditioned by the presence and the nasality of the onset, as well as the presence of the coda.

NEPMs share similar syllable and word structures, yet they vary in their phonotactic constraints on phoneme distributions, particularly concerning vowel distribution. The canonical syllable template is (C)(C)V(C). Words typically comprise two syllables following a (C)(C)V(C).(C)V(C) pattern. With the exception of the glottal stop, all consonant phonemes can occupy the onset position of a syllable, among which the velar fricative /x/ only occurs as geminated /xx/ in KM and CTM. In the coda position, only nasals and glottals are allowed. The syllable template further reveals that consonant clusters only occur in the onset position, which may consist of

two identical or two dissimilar segments. Non-geminate clusters usually comply with the SSP. ITM allows the largest range of possible combinations in consonant clusters, but non-geminate clusters that violate the SSP tend to be assimilated regressively to become geminates. This tendency is also observable in some clusters that conform to the SSP.

At the word level, syllables with a complex onset typically occur initially. Exceptions are found in a few trisyllabic derivatives, where non-geminate clusters occur word-medially after a prefix *bə-* 'INTR; MID'. Consonant sequences formed across two syllables commonly consist of a nasal + a homorganic voiced obstruent, but sequences of /-ʔ.C-/ are also attested. In word-final position, only three consonants /ʔ, ŋ, h/ are allowed. NEPMs thus exhibit a strong left-oriented consonantal asymmetry in their syllable and word structures (Ogloblin 2018: 330). As for vowels, ultimate syllables allow for more vowels and more phonemic contrasts, which may be taken as an indicator of their greater prominence. In KM and CTM, the mid-high vowels /e, o/ are only permitted in ultimate syllables. In ITM, both the mid vowels and the diphthongs exclusively occur in this position. Nasal vowels, which carry a relatively low functional load, are also solely attested in ultimate syllables. Another noteworthy generalisation about the word structure in NEPMs is that both morphologically simple words and derivatives share a similar word shape, following the same maximal CCVC.CVC template, and the phonotactic constraints apply to both word types.

In the morphological system, NEPMs have a small inventory of affixes with an overwhelming preference for prefixing. KM and CTM have five prefixes, namely *bɣ-* 'INTR; MID', *tɣ-* 'NVOL', *pɣ-* 'CAUS; FCT', *NN$_1$-* 'IPFV' and *NN$_2$-* 'NMLS'. ITM has one prefix less as it lacks the causative/factitive marker. Overall, NEPMs display a largely isolating-analytic profile. The prefixation processes are further limited by the phonological constraints on permissible word-initial clusters. The full forms of *bɣ-* 'INTR; MID', *tɣ-* 'NVOL' and *pɣ-* 'CAUS; FCT' only emerge before vowel-initial bases. They are reduced to *b-, t-, p-* respectively before bases with a single initial consonant, with which they combine to form a consonant cluster. Bases with initial complex clusters, on the other hand, generally do not undergo prefixation. Since the clusters resulting from prefixation need to be phonologically well-formed, the reduced single-segment prefix mainly appears before bases with an equally or more sonorous initial consonant. In other cases, the prefix assimilates to the base-initial consonant, creating morphologically complex word-initial

geminates. I have demonstrated that while these complex initial geminates may arise from the assimilation of an underlying prefix to a base-initial consonant, this unified analysis based on allomorphic alternations cannot comprehensively account for all occurrences of complex geminates. Therefore, a morphophonological process of initial gemination must be acknowledged. Initial gemination can also operate as a locative marker, which is best treated as a special type of preposition cliticisation. The nasal prefixes *NN$_1$-* 'IPFV' and *NN$_2$-* 'NMLS' are underlyingly geminated, and they have a restricted distribution. They only attach to bases with an initial vowel or *ɣ-*, appearing as *ŋŋ-*, or to bases with an initial voiceless obstruent, wherein the underspecified nasal element takes the place of articulation of the base-initial segment and substitutes it.

In additional to prefixation and initial gemination, compounding and reduplication are other word-formation processes in NEPMs. The majority of compounds are attributive compounds, constituted by a noun + a noun or a noun + a stative verb. Reduplication is a productive process that can be applied to bases from various word classes, performing a variety of grammatical functions. Formally, reduplication may be considered a special type of compounding, as it is restricted to full reduplication and a few instances of echo reduplication, which produce complex words with two roots.

As far as word classes are concerned, NEPMs have two open word classes, namely nouns and verbs, which are primarily differentiated based on their syntactic properties rather than morphological features. NEPMs lack a distinct category of adjectives. Semantic adjectives are subsumed as a subclass of verbs, referred to as stative verbs, on account of their shared grammatical properties with dynamic verbs. Stative verbs may also function as manner adverbs without any formal changes; thus, adverbs are best viewed as a closed word class with a limited number of members. Other word categories that may occur in an NP include quantifiers, numerals and classifiers, which as a rule precede the head noun. On the other hand, possessive pronouns, relative clauses and demonstratives follow the head noun within an NP.

Clauses in NEPMs can be classified into verbal and non-verbal ones. Verbal clauses can be further categorised as dynamic and stative, or as intransitive, transitive and ditransitive clauses based on the verb's valency. Notably, NEPM verbs do not have active/passive voice-marking morphology. In transitive verbal clauses, the differentiation between active and

passive constructions is primarily indicated by word order. An active transitive clause has the basic word order of AVP, whereas the a passive construction is marked by the order of PV(A), often combined with an adversative auxiliary. The agent is introduced by a preposition as an adjunct, and when the adversative auxiliary is absent, the expression of an agent is compulsory.

Overall, NEPMs exhibit more similarities in their morphosyntactic structures, but there are also some distinct features that set ITM apart from KM and CTM. For example, while ITM distinguishes exclusivity in the first-person plural pronouns as many other Malayic varieties do, such as distinction is neutralised in KM and CTM. ITM also stands out in having a reduced set of demonstratives and the prevalent usage of an anaphoric pronoun *ŋə*. Non-verbal clauses with prepositional predicates are typically attested in ITM, but not in KM and CTM.

The synchronic descriptions form the foundation for investigating the historical development of NEPMs. By comparing the phonological and morphological systems of NEPMs with existing PM reconstructions, the diachronic changes that have taken place in the history of NEPMs are established. The evolution from PM to NEPMs can primarily be characterised as showing a tendency towards reduction, both in phonology and morphology. Word-final consonants underwent various mergers and losses: final stops merged to /ʔ/, final nasals merged to /ŋ/, *-s underwent lenition and merged with *-h, and all final approximants were eliminated. Additionally, word-medial consonant sequences consisting of a nasal + a homorganic voiceless obstruent were reduced to the obstruent component.

The most prominent change in NEPMs is the process of syllable reduction, which affected both disyllables and trisyllables. As a consequence of losing initial *(h)ə, some PM disyllables have become monosyllabic. More significantly, trisyllables, encompassing morphologically simple words and prefixed derivatives, have been reduced to disyllables through vowel contraction or antepenultimate vowel syncope. These shortened disyllables often have initial clusters, some of which have undergone cluster assimilation, leading to the formation of geminate clusters.

These changes in the consonant system are shared by all three NEPM varieties (with some minor differences in the processes involving syllable reduction), which gives the impression that NEPMs resemble each other to a great extent. However, the development of the vowel systems reveals signi-

ficant divergence, which again distinguishes ITM from KM and CTM. In certain aspects, there is a trend of change towards complexification rather than reduction, as NEPMs have acquired more vowel phonemes, including phonemic nasal vowels. KM and CTM underwent the lowering of high vowels in both penultimate and ultimate syllables, whereas ITM has retained high vowels in the penultimate syllables and diphthongised high vowels in ultimate syllables. Some subvarieties of ITM have undergone a two-stage diphthongisation process with particularly complex conditions. While there are some common vowel changes across NEPM varieties, such as the raising and rounding of ultimate *a, the environments in which such changes occurred are not identical.

The establishment of a relative chronology of sound changes demonstrates that the seemingly shared development in the consonant systems must have been preceded by distinct changes in the vowel system of the individual languages. That is to say, the common changes in the consonant systems cannot be attributed to shared innovations in a single ancestral language, but more likely have emerged as the result of diffusion.

The morphological history of NEPMs reveals an even clearer pattern of reduction. Out of more than sixteen affixes that could be reconstructed to PM, only five are preserved in KM and CTM, and merely four in ITM. PM prefixes are typically retained, whereas all suffixes and circumfixes have either been lost or have become fossilised in a few instances. Since original prefixes mostly fell on the antepenultimate syllables which were the target of syllable reduction, the retained prefixes are phonologically reduced. Some prefixes have transformed into geminated segments, and such contrasts between a base form with an initial singleton and a corresponding complex form with an initial geminate have served as the basis for generalising initial gemination as a morphophonological operation, presumably through analogy. Other prefixes have only been retained in a restricted set of phonological environments.

Furthermore, I have showed that the fossilisation of certain suffixes was not random, and the loss of other suffixes and circumfixes was not a wholesale process. Both processes can be understood as being influenced by the structural pressure of maintaining a canonical disyllabic structure. If a disyllabic target had been achieved through other means of reduction, or if the original suffixed forms were semantically disassociated from their bases, suffixes tended to become fossilised. The observation that the reduction of

morphology followed a specific pattern suggests that internal phonological changes were the impetus behind the morphological reduction in NEPMs.

Lastly, I discussed the potential role of language contact and substratal influences in morphological reduction, and examined probable indications of external interference in various aspects of NEPM grammars. Clear traces of contact-induced change are not found in the lexicon, phonology or morphology of NEPMs. Therefore, there is little evidence supporting a contact scenario, and it is not immediately obvious that external influences played a significant role in motivating the morphological reduction.

To sum up, the perception of NEPMs as some of the most aberrant Malayic varieties, whether expressed by laypeople or earlier general studies on Malayic varieties, is not unfounded. This study has demonstrated that NEPMs exhibit many distinct features, both from Malayic and cross-linguistic perspectives. While many general characteristics are common to all three NEPM varieties, KM and CTM appear to be more closely related to each other, while ITM stands out as more distinctive.

9.3 Implications for the Malayic migration history

It is generally presumed that the dispersal of Malayic languages followed a route via Sumatra from the homeland in West Borneo before reaching the Malay Peninsula (see a summary in §1.2). As the ancestors of Malayic-speaking populations on the peninsula arrived from overseas, they likely settled initially along the coasts and river mouths before spreading inland along the riverine systems (§1.4.1.1). Importantly, previous studies also suggest that the Malayic varieties spoken on the east coast of the peninsula form a distinct subgroup within Malayic (§1.4.2). Assuming that the pattern of language dispersal reflects the prehistoric movements of the ancestors of those who speak them (cf. linguistic migration theory, Sapir 1968 [1916]; Dyen 1956), it can be inferred from the existing literature that the ancestors of NEPM speakers migrated into the region in a single group, presumably from Sumatra.

However, in light of the findings from the current study, a revision and adjustment of the aforementioned scenario is required. It appears improbable that the ancestors of NEPM speakers reached the Malay Peninsula as a singular group. Instead, I propose a two-wave migration pattern based on

the present-day distribution and differentiation of NEPMs.

Several important points should be highlighted here. First, despite the superficial similarities across NEPMs, there is a fair amount of variation among the three varieties. In §9.2, I drew attention to the similarities between KM and CTM on the one hand, and the distinctiveness of ITM on the other hand. The closer relationship between KM and CTM is evident in various aspects of the grammar, including phonology, morphology and certain syntactic structures. ITM stands out with both conservative and innovative features that are not attested in KM or CTM (e.g., the retention of penultimate high vowels and the diphthongisation of ultimate high vowels). Second, ITM displays an exceptionally high degree of internal variation (see Collins 1983a), which suggests more significant differentiation *in situ* for a longer period. Third, it is also noteworthy that ITM shares some of its peculiarities with non-Peninsular Malayic varieties such as Kerinci, the most prominent one being the diphthongisation of earlier ultimate high vowels (see early remarks in van Reijn 1974). All these observations imply that ITM is the older variety among NEPMs, and it may have closer relationships with other non-Peninsular Malayic varieties. Moreover, as I argued in §7.6, NEPMs do not form a discrete subgroup within Malayic, and they do not have an immediate common ancestor. This indicates that the ancestors of NEPM-speakers did not migrate to the region as a unified group.

Summing up the linguistic evidence, I suggest that ITM represents an older variety, presumably an earlier offshoot of PM, whose speakers settled inland in Terengganu at an early stage. KM and CTM, on the other hand, reflect a migration of more recent origin, and their speakers spread primarily along the coastline before advancing inland into Kelantan.

Similar remarks were previously made by Mohd Tarmizi (2018a), who contends that the inland Malayic varieties of the east coast of the peninsula (Kelantan, Terengganu and Pahang) have a longer history compared to the coastal varieties. He further suggests that the east coast varieties originated in the hinterland and then dispersed downstream towards the coast. However, this downstream migration pattern is unconvincing for two reasons. First, the direction of migrating from the inland to the coasts contradicts the commonly accepted pattern of Malayic settlements, which typically began at river mouths. Second, in the case of Terengganu, there is no evidence indicating that CTM represents an offshoot of ITM as a result of the

speakers' migration. The two varieties spoken in Terengganu do not form a dialect continuum, but are separated by a number of clear isoglosses (see Collins 1983a). The geographic and political proximity between inland Terengganu and coastal Terengganu does not suggest a closer linguistic relationship between CTM and ITM (against KM).

The migration scenario proposed here is also supported by non-linguistic data, as historical records demonstrate early settlements in the inland areas of the east coast. Chinese records documented the existence of several inland kingdoms, such as Chi'tu around the sixth century, which was likely situated in inland Kelantan, and Fo-lo-an around the twelfth century, possibly located in inland Terengganu (§1.4.1.2). According to *Sui-shu*, it took a month's journey to reach the capital of Chi'tu from the coast (Wheatley 1973: 36), suggesting that the kingdom was situated in the hinterland. Moreover, the establishment of human settlements in the inland must have begun thousands of years ago before the arrival of the Malays (and possibly pre-Malayic Austronesian speaking people), likely driven by the abundant mineral resources such as gold and tin on the peninsula's interior (Benjamin 1987; Bellwood 1993). The discovery of the Terengganu inscription stone further underscores the historical significance of the Hulu Terengganu region. Sheppard (1949: 3–4) postulated that a (probably Buddhist) Malay kingdom might have already existed in inland Terengganu in the eighth century, predating the foundation of Malacca by several centuries. The founding legend of the Patani Kingdom, on the other hand, suggests that its ancestors arrived from overseas, probably in the early fifteen century (§1.4.1.2).

This timeline is consistent with the two-wave migration pattern. It is plausible that ITM descends from a group of Malayic-speaking people who had already settled in the inland area during the first millennium, though the exact date cannot be determined. Subsequently, the ruler of the inland kingdom converted to Islam and erected the Terengganu inscription stone in the early fourteenth century. The ancestors of KM and CTM speakers (as well as Patani Malay speakers) arrived in the region during a later period, either from other parts of the Malay Peninsula or possibly from Sumatra. These people primarily inhabited the coasts and established kingdoms in the nearby regions before eventually moving inland.

Based on the historical scenario proposed above, it remains unclear how and why the languages spoken by these groups of people underwent sub-

stantial phonological changes, which further drove morphological reduction. It is plausible that the ancestors of NEPM speakers came into contact with pre-Malayic speaking people who were already present on the Malay Peninsula before the arrival of the Malays. However, as discussed in §8.4.3, there are no clear indications in present-day NEPMs that speak in favour of such a contact scenario. Further comprehensive research is required to thoroughly investigate this aspect.

9.4 Limitations and directions for future research

Given the limited time and data available, compounded by the negative impact of the Covid-19 pandemic on field trips, it is important to acknowledge that the present study has several limitations. There are also a number of unresolved issues and questions that require further research in the future.

First, more data, especially additional naturalistic conversational data, would be beneficial to supplement the corpora and synchronic description. The data for the present study were collected during two field trips. However, due to the requirement of visiting at least three field sites for collecting data from three varieties, the limited duration of stay at each site has resulted in a reduced amount of data collected for each variety. A fair amount of data used in this study was obtained through elicitation, which is arguably less representative of the natural language use. However, efforts were made to minimise the influence of the intermediate language (SM) during data collection by employing non-linguistic stimuli and prompting informants to generate sentences rather than translate them from SM. Additionally, a more optimal approach would also involve recording data from a larger number of speakers, so that inter-speaker variations can be taken into account.

Second, a few issues in the synchronic description have been glossed over or left untreated for future studies. The stress patterns of NEPMs, for instance, have only been described in a mostly impressionistic manner, and the acoustic correlates of stress have not been examined in detail. This aspect of research requires future phonetic study. Another noteworthy phenomenon that has not been explored in the present study is the tendency to shorten many disyllabic words to monosyllables in connected speech. This type of syllable reduction generally deletes the first syllable or only the

first consonantal segment, which is quite different from the historical processes of reducing disyllables to monosyllables discussed previously. This phenomenon appears to be particularly prevalent in CTM. Unfortunately, there is a limited amount of naturalistic data available for this variety. In addition to common reductions such as *buwɔh* → *wɔh* 'fruit', *buwi* → *wi* 'to give' and *buwaʔ* → *waʔ* 'to do; CAUS' (see Table 7.17 in Chapter 7), other examples are provided below in (1). Both full forms and shortened forms are attested, but typically only the full forms were given by the speakers in word list elicitation, which may be considered citation forms. It is important to note that syllable reduction as such is not limited to function words like prepositions and conjunctions, but also applies to content words, as seen in *pitih* → *itih* 'money', *ɣumɔh* → *umɔh~mɔh* 'house' and *buboh* → *uboh~boh* 'to put'. Further investigation is needed to uncover the patterns of this process and the potential factors that motivate it.

(1) CTM

buleh	→	*uleh~leh*	'can; be able to'
dəŋaŋ	→	*ŋaŋ*	'with; and'
ləpah	→	*pah*	'after; then'
tapi	→	*pi*	'but'
səbaʔ	→	*baʔ*	'because'
taʔdɔʔ	→	*aʔdɔʔ*	'NEG.EXIST'
taʔdi	→	*aʔdi*	'just now'
saŋaʔ	→	*ŋaʔ*	'very'
dulu	→	*lu*	'first'
lagi	→	*agi*	'again; more; still; yet'
pitih	→	*itih*	'money'
buboh	→	*uboh~boh*	'to put'
makaŋ	→	*akaŋ*	'to eat'
katə	→	*atə*	'to say'
ɣumɔh	→	*umɔh~mɔh*	'house'

Third, in examining of the history of NEPMs, this dissertation has focused on establishing the changes that have taken place from PM to NEPMs and comparing the three varieties against each other. In order to gain a fuller understanding of the dispersal of NEPMs and the migration history of their speakers, it is necessary to examine the relationship between NEPMs and other

Peninsular Malayic varieties, as well as non-Peninsular Malayic varieties. Before achieving this goal, however, more descriptive data must be collected. This dissertation has taken one step forward in advancing Malayic descriptive and historical linguistics, but further research is required to achieve a detailed internal subgrouping of the Malayic languages.

Lastly, in order to comprehensively study the prehistory of the Malay Peninsula, it is essential to consider all language groups present in the region and their relationships. This includes investigating language contact between NEPMs and neighbouring Aslian languages in both present and past contexts, as well as potential contact between Malayic and Mon-Khmer languages. Towards this end, further research with a more specific focus on uncovering traces of language contact is warranted. Moreover, it is crucial to combine linguistic data with insights from history, archaeology, ethnography and genetic studies. An interdisciplinary approach will provide a more comprehensive understanding of the languages in the region and the migration history of their speakers.