

On Becoming Mainland

Author(s): Pittayawat Pittayaporn

Source: *Sojourn: Journal of Social Issues in Southeast Asia*, March 2024, Vol. 39, No. 1 (March 2024), pp. 62-89

Published by: ISEAS - Yusof Ishak Institute

Stable URL: <https://www.jstor.org/stable/10.2307/27300975>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



ISEAS - Yusof Ishak Institute is collaborating with JSTOR to digitize, preserve and extend access to *Sojourn: Journal of Social Issues in Southeast Asia*

JSTOR

On Becoming Mainland: Unravelling Malay Influence on Moklenic Languages

Pittayawat Pittayaporn

This study examines the impact of Malay on Moklenic dialects spoken by Moklen and Moken communities to uncover the historical contact setting between Proto-Moklenic and Malay. It shows that a significant portion of Proto-Moklenic vocabulary was borrowed from Malay, while aspiration and three-way vowel height contrasts arose from Malay influence. Sociolinguistic dynamics resulting from contact between Moklenic and Malay, as reflected in the contact-induced changes, indicate a long-lasting and closely intertwined, though asymmetrical, relationship, which suggests widespread bilingualism and openness towards the influence of Malay.

Keywords: sea nomads, Moklenic languages, Moklen, Moken, Malay, language contact.

Language can provide valuable insight into the history of interethnic relations that may not be apparent from documents or archaeological materials. By analysing changes in a language resulting from contact with another language, researchers can uncover evidence of past interactions and shared history among ethnolinguistic groups, particularly for smaller communities, such as the Moklen and Moken, who are often marginalized in national narratives. Linguistic analysis of Moklenic dialects is crucial for understanding the history of the Moklenic-speaking populations because there are only limited historical traces of them.

The Moklenic people, who reside along the Andaman coast of Thailand and Myanmar, comprise various ethnolinguistic groups speaking closely related non-Malayic dialects. The Moklenic subgroup of the Austronesian language family is made up of two clusters of

mutually intelligible dialects that are referred to by similar names: “Moklen” and “Moken”. These dialect clusters are highly similar in terms of their grammar and sound systems, but they have significant differences in their vocabularies. The Moklen dialects are spoken in seaside communities within the Phuket and Phang-nga provinces of southern Thailand. Moken varieties are spoken by communities that, although increasingly settled, continue to roam the islands of Thailand’s Phang-nga and Ranong provinces, as well as Myanmar’s Mergui Archipelago. Its speakers can also be found in a village in Phuket (Hogan 1972, pp. 200–12; Larish 1999, pp. 60–77; Larish 1992, pp. 1305–307).

The literature often refers to speakers of Moklenic dialects as “Orang Laut”, which means “sea people” in Malay and includes all sea-based populations in the Strait of Melaka, Sumatra and Malay Peninsula estuaries, Riau-Lingga archipelagos and other South China Sea islands. The term includes the Malayic-speaking Urak Lawoi’, who inhabit northern Malaysia and southern Thailand, with a geographical overlap with Moklenic dialects in Phuket (Andaya 2008, p. 173; Anderbeck 2012, pp. 266–67; Hogan 1972, pp. 206–7). The use of this umbrella term in primary sources or academic writing inevitably obscures the heterogeneity among these groups. The pivotal study on the role of trade in ethnic formation within the Strait of Melaka by Andaya (2008, pp. 173–201) assumes that Moklen and Moken practices resemble those of Orang Laut groups in the south, with close political, economic and sociocultural ties to Malay lords. The Orang Laut emphasized their sea-based lifestyle as a defining feature of their ethnic identity, in contrast with the Malay.

Records of contact between the Moklenic- and Malay-speaking communities date back only to the nineteenth century. However, no documents recording historical contact between the Moklen and the Malays exist. The oldest known reference to a Moklenic population is a Moken primer published in 1846, while the earliest report of “sea nomads” dates to 1826, shortly after Britain colonized Lower Burma (Ivanoff 1997, p. 10). Oral history from the 1920s describes frequent raids by the Malays, leading to the enslavement of many

Moken (White 1922, pp. 57–58). Primary evidence of contact with Malay-speakers dates to the 1980s, when Ivanoff (2001), conducting fieldwork in the Surin Islands, recorded Moken folktales of encounters with the Malay people, some of which are adaptations of or have been influenced by Malay folktales. Notably, Ivanoff (2001, pp. 34–36) also documented the arrival of a group of Malays from Phuket to whom the Moken elders spoke in Malay.

Linguistically, the Moklen and Moken dialects are related to Malay and other Malayic languages—such as Minangkabau, Banjarese, Iban and Kerinci—and Malay dialects—including Malaysian, Indonesian, Kelantan Malay and Urak Lawoi’ (Adelaar 1992, pp. 1–2; Adelaar 2005, pp. 202–5; Anderbeck 2012, pp. 267–70). Their genealogical relationship to other languages within the Austronesian family is uncertain, but they are considered distinct from the Malayic group (Blust 1994, pp. 34–40; Blust 2010; Grey, Drummond and Greenhill 2009; Larish 1999, pp. 362–415; Smith 2017, pp. 459–60; Thurgood 1999, pp. 58–59). Research on Moklenic dialects has identified evidence of lexical and structural convergence towards Mainland Southeast Asian languages, with extensive influence from Austroasiatic and Thai languages leading to the adoption of typical Mainland Southeast Asian features in their phonological and grammatical structures (Brunelle and Pittayawat 2012; Larish 1997; Larish 1999, pp. 48, 136–42, 308–33, 362–415; Pittayawat 2005).

These observations present Moklenic languages as remnants of an ancient Mainland Austronesian population outside the linguistic ecology dominated by the Malay traditionally used as a *lingua franca* in trade (Ansaldo 2009, pp. 53–55). Larish (1999, pp. 417–50) hypothesizes that contact with Austroasiatic languages occurred between the third and sixth centuries along active land trade routes across the Malay Peninsula. He suggests that a few Moklenic etyma were borrowed from Austroasiatic languages during ancient times, implying the involvement of Moklenic groups in metal trade with Austroasiatic-speaking kingdoms. A similar view is espoused by Sidwell (2013, p. 262) and Benjamin (this issue). This perspective contrasts with Andaya (2008, pp. 173–201) and Ivanoff (2001, pp.

408–9), who view Moklenic groups as part of the Malay seascape and under greater Malay influence than Austroasiatic.

This study examines the influence of Malay on Moklenic dialects, focusing on the presence of loanwords from Malay and the Mainland Southeast Asian phonological traits in Proto-Moklenic (PMok), the hypothetical common ancestor of Moklen and Moken. It aims to uncover the historical setting for contact between PMok and Malay and determine whether the observed changes are due to non-Malay-speakers shifting to Malay or PMok-speakers borrowing Malay elements. The findings support the notion that Moklenic-speaking groups had close ties with Malay-speakers before appearing in historical records and provide valuable insight into the sociocultural and linguistic contexts of Moklenic-speaking communities in their interactions within the Southeast Asian linguistic landscape.

The primary sources of Moklenic data in this study are the Bang Sak dialect of Moklen documented by Pittayawat, Warunsiri and Loss (2022) and the Rawai dialect of Moken described by Veena (1980) and Wolff (2018, pp. 523–44),¹ with additional data from Larish (1999). The PMok forms cited are from Larish (1999, pp. 573–1801),² although modifications have been made to address concerns regarding Larish's reconstruction of the PMok vowel system. Larish's proposal includes excessively fine vowel distinctions based on a potentially inaccurate interpretation of available data and does not account for lexical tone. All PMok forms cited in this study should be considered tentative.

Standard Malay (SM) forms, representing dialects from the Melaka area, were obtained from works by Amon (1987; 1992) or from Winstedt (1960). Urak Lawoi' (UL) forms, which represent dialects from the north, were sourced from Amon (1987; 1992) and Hogan (1988). Proto-Malayic (PMal) forms, representing the common ancestor of all Malayic varieties and other Malayic dialects, were obtained from Adelaar (1992), whereas Proto-Malayo-Polynesian (PMP) forms were sourced from the *Austronesian Comparative Dictionary* (Blust and Trussel 2020). In addition, Austroasiatic forms are cited from *A Mon-Khmer Comparative Dictionary* (Shorto 2006).

Moklenic and Mainland Southeast Asian Languages

Researchers generally agree that the Moklenic dialects belong to the Austronesian language family, but there is ongoing debate about their position within the Malayo-Polynesian branch. Blust (2019, 1994, 2010) proposes a link to the Greater North Borneo subgroup, while Larish (1999, pp. 329–61) proposes that they belong to the Malayo-Chamic subgroup. Grey, Drummond and Greenhill (2009) also suggest a close relationship between Moken and the Chamic subgroup. While most researchers agree that the Moklenic dialects share close kinship with the Malayic and Chamic languages, Smith (2017, pp. 459–60) argues that they constitute a primary subgroup within the Malayo-Polynesian family and lack significant affinity with other subgroups.

Several hypotheses have been proposed regarding the origins of Moklenic. One hypothesis suggests, based on oral traditions of the Moken in Myanmar, that the Moklenic dialects originated in the north and migrated southward (White 1922, pp. 156–57). Another posits that Moklen originated along the Gulf of Thailand and later crossed westward to the Andaman side of the isthmus (Benjamin this issue; Larish 1999, pp. 453–82; Sidwell 2013, p. 262). Yet another proposes that Moklen originated in the south, either in the Riau-Lingga Archipelago or northern Sumatra, and later migrated northward (Blust 1994, pp. 52–54; Ivanoff 2001, pp. 408–9; Sopher 1977, pp. 345–86). Which hypothesis is most accurate remains debatable.

Nevertheless, it is evident that Moklenic dialects have undergone convergence with languages in the Mainland Southeast Asian linguistic area, as demonstrated by their structural characteristics, which exhibit greater similarity to Kra-Dai and Austroasiatic languages than their Austronesian counterparts in the insular part of the region (Larish 1997; Larish 1999, pp. 362–415; Pittayawat 2005). Specifically, comparison of the phonological characteristics in Table 1 reveals that Moklen and Moken display closer similarity to the Mainland languages Thai (Kra-Dai) and Khmu' (Austroasiatic) than to their relatives Malay and Tagalog (Austronesian). Importantly, these

TABLE 1
Mainland Southeast Asian Phonological Characteristics in Moklenic

<i>Characteristics</i>	<i>Thai</i>	<i>Khmu'</i>	<i>Moklen</i>	<i>Moken</i>	<i>Malay</i>	<i>Tagalog</i>
3-way laryngeal contrast	✓	✓	✓	✓		
Neutralization in coda position	✓	✓	✓	✓		✓
3-way vowel height contrast	✓	✓	✓	✓		
Vowel length contrast	✓	✓	✓	✓		
Monosyllabism/ sesquisyllabism		✓				
Bimoraicity of foot head	✓	✓	✓	✓		
Word-final stress	✓	✓	✓	✓		
Tone/register	✓	✓	✓	(✓)		

Source: Modified from Pittayawat 2005, p. 192.

characteristics are observed in all documented Moklenic varieties, indicating they were already present in PMok. Although Veena (1980, pp. 24–26) claims that Moken differentiates tones at “the phonetic level”, no study has reported contrastive tones in the language. However, the presence of lexical tones in Moklen (Pittayawat, Warunsiri and Loss 2022, pp. 10–11) implies that Moken may also exhibit lexical tones.

Previous research has identified Austroasiatic and Thai languages as sources of the Mainland characteristics present in Moklenic. Larish (1999, pp. 417–50), based on certain everyday vocabulary (see Table 2), considers contact with Austroasiatic languages to be the most significant and ancient.

In addition, Larish (1997, pp. 127–32; 1999, pp. 380–83, 402–3) suggests that Moken may share features—including word-final stress, diphthongs, vowel length distinction and lexical tone—with Austroasiatic languages as a result of language contact. Larish (1999, pp. 417–50) proposes that this contact occurred through trade across the Thai-Malay peninsula beginning in the sixth century. But some authors argue that these similarities are motivated by factors internal

TABLE 2
Examples of Austroasiatic-Related Vocabulary in Moklenic

<i>Glosses</i>	<i>Moklen</i>	<i>Malay</i>	<i>Cham</i>	<i>Aceh</i>	<i>Old Mon</i>
bracelet	dəlá:ŋ	gəlaŋ	klà:ŋ	glə:ŋ	—
iron	bəc ^h új	bəsi	pat ^h ej	buusəə	bərsəj
bird	ticúm	—	cim	cicem	kəncem
crab	kətá:m	kətam	—	—	kənta:m

Source: Extracted from Larish 1999, pp. 573–1081.

to the Moken language itself. For instance, Pittayawat (2005, pp. 195–96) demonstrates that vowel length distinction arose from the lengthening of vowels preceding Proto-Austronesian *-q and *-R, while the distinction between high and low vowels is due to the lowering of *i and *u in closed syllables. Brunelle and Pittayawat (2012, pp. 414–19, 422–25) suggest that Moken might have evolved from a trochaic to an iambic language through a stressless stage induced by word-final lengthening. Therefore, a comprehensive linguistic analysis is needed to determine whether these features truly resulted from direct contact with Austroasiatic languages.

Language contact with Thai has had a significant impact on Moklenic dialects. Larish (1999, p. 48) suggests that Moklenic dialects' change from Malayo-Polynesian *R to /-l-/ between vowels and /-n/ in word-final position may be due to contact with Thai starting in the thirteenth century. The most crucial claim is that Moklenic is evolving a prosodic system similar to the Southern Thai dialects (Larish 1997, pp. 136–42). Additionally, Pittayawat (2005, pp. 192–93) argues that Thai serves as the primary source of contrastive aspiration and presents numerous Thai loanwords featuring aspirated stops as onsets, such as /məc^haj/ 'to use', /p^həlu:ŋ/ 'hole' and /k^hiŋ/ 'half'. Interestingly, he highlights certain Moken terms with initial aspirated stops that can be traced back to Proto-Austronesian *s-, such as /cəc^həj/ 'milk', derived from *susu. We argue, however, that these Austronesian-related forms are loanwords from Malay, which introduced aspirated consonants into the Moklenic sound system.

Malay, a cousin of Moklenic, has often been overlooked as a source of influence, most likely because of their shared Austronesian linguistic roots. Nevertheless, the Moklenic vocabulary contains irregular sound reflexes that can help identify borrowings. Blust (1994, pp. 42, 53) asserts that many lexical items in Moklenic were borrowed from Malay dialects but provides only one clear example; namely, the Moken word /bəc^huj/ ‘iron’ from the Malay word /bəsi/, ultimately of Austroasiatic origin. Furthermore, Larish (1999, pp. 335–45, 394–414) speculates that the insertion of the glottal stop /ʔ/ word-finally, the variation between /s/ and /c^h/, and the distinction between short /a/ and long /a:/ are also a result of contact with Malay. There has not, however, been an in-depth study of the Malay-Moklenic contact.

Socio-historical Contexts of Contact-Induced Changes

The sociolinguistic environments in which PMok and Malay potentially interacted offers insight into the socio-historical circumstances surrounding the encounter of speakers of these closely related languages. Despite differences in opinions and terminology among linguists (e.g., Matras 2020; Sankoff 2002; Thomason and Kaufman 1988; van Coetsem 1988; Winford 2005), there is a general consensus that contact-induced changes can arise from two types of transfer, each of which emerges in different socio-historical settings involving varying degrees of bilingualism, language attitudes, language policies, types of population movement, and population sizes, among other factors.

The first type is borrowing, which occurs when speakers of the recipient language adopt lexical or structural elements from the source language in which they are less proficient (Thomason and Kaufman 1988, pp. 35–64; van Coetsem 1988, pp. 7–24). English-speakers incorporating French words while speaking English (van Coetsem 1988, p. 3) is a clear example. From a population perspective, borrowing is closely associated with situations in which the native language is maintained despite pressure from an outside language.

TABLE 3
Borrowing Scale Proposed by Thomason and Kaufman

Casual contact	Category 1: content words
↓	Category 2: function words, minor phonological features, lexical semantic features
	Category 3: adpositions, derivational suffixes, phonemes
	Category 4: word order, distinctive features in phonology, inflectional morphology
Intense contact	Category 5: significant typological disruption, phonetic changes

Source: Adapted from Matras 2020, p. 168.

This type of transfer primarily involves vocabulary, although structural borrowing can occur with intense contact (Sankoff 2002; Thomason and Kaufman 1988, pp. 37–45; Winford 2005). Thomason and Kaufman (1988, pp. 74–76) propose a borrowing scale of the types of materials borrowed into the recipient language as a function of the intensity of contact. As summarized in Table 3, the scale ranges from casual contact, in which only a few non-basic words are borrowed, to intense contact under strong cultural pressure, in which a significant portion of the basic vocabulary and major structural features are adopted from the source language.

The second type of language transfer is imposition, where speakers of the source language retain characteristics of their native language when speaking a less proficient language (Thomason and Kaufman 1988, pp. 35–64; van Coetsem 1988, pp. 7–24); for example, a French speaker using French articulatory features while speaking English (van Coetsem 1988, p. 3). This transfer type primarily affects phonological and grammatical features and is associated with scenarios in which a group of speakers abandon their native language and shift to a dominant language (Sankoff 2002; Thomason and Kaufman 1988, pp. 37–45; Winford 2005). According to Thomason and Kaufman (1988, pp. 110–46), small shifting groups or perfect target language acquisition are not expected to result in transfer-induced changes. But rapid shifts or large populations with imperfect acquisition can significantly influence the target language, especially in terms of phonological and grammatical structures.

Focusing on phonological structure, contact-induced changes involve the replication of phonological forms from a donor language by a recipient language. These processes can affect original sound patterns differently, with the first three associated with borrowing and the last with imposition. When speakers of the recipient language are monolingual or semi-bilingual with a conservative attitude to foreign linguistic influence, the replicated forms are adjusted to fit the constraints of the recipient language. For example, *current*, when borrowed into English from French, underwent changes such as stress shifting, aspiration of /k/ and reduction of unstressed vowels. The overall English sound system was not affected (Matras 2020, pp. 242, 245; Thomason and Kaufman 1988, pp. 77–78).

Second, in cases in which bilingualism is more widespread among speakers of the recipient language and they are flexible in their language use, the borrowed forms may be maintained as they are or undergo partial modifications. For example, German *Dschungel*, borrowed from English *jungle*, is pronounced with an initial /dʒ-/ sound non-native to German (Matras 2020, pp. 242–43; Thomason and Kaufman 1988, pp. 78–73). Third, when native speakers maintain their community language despite intensive bilingualism with a prestigious target language, the sound patterns of the native language may converge towards those of the target language. For instance, in Domari, /p/ is becoming /b/ under the influence of Arabic, which lacks a voiceless labial stop (Matras 2020, pp. 243–45; Thomason and Kaufman 1988, pp. 115–46).

Finally, in situations involving limited bilingualism or a shifting large population, there can be two influencing factors on language choice: a strong group identity and the reason for acquiring the target language. These factors can impose constraints of the native language sound patterns on the phonological forms of the target language. For example, the pronunciation of /t/ as an unaspirated retroflex in South Asian English is influenced by native Indian languages (Matras 2020, pp. 243–45). This situation is commonly observed among immigrant groups that sustain long-term bilingualism or local communities that are bilingual in an imported language (Sankoff 2002, pp. 644–49).

Examining the impact of Malay on Moklenic can reveal the type of language transfer and historical dynamics of contact. With borrowing, we would expect substantial presence of Malay borrowings, especially cultural vocabulary and high-register words, with fewer phonological and grammatical interferences. With imposition, we would expect limited presence of substrate vocabulary but significant phonological and grammatical interference from Malay.

Malay Loanwords in PMok

Many PMok words are clearly Malay loanwords. These words are of PMP origin, display irregular reflexes of PMP phonemes and resemble corresponding Malay forms. For instance, the PMok word *plóh for ‘ten’ probably did not come directly from PMP, as the reconstructed PMP form *puluq would be *plú:k in PMok. The presence of /-h/ in PMok is surprising because it differs from the expected sound change from PMP *q to Moklenic /-k/ observed in words like PMok *pənú:k ‘full’ and *butú:k ‘penis’ (Blust 1994, pp. 34–40; Wolff 2018, pp. 538–39). Notably, the attested PMok forms closely resemble the SM and UL word /səpuluh/, suggesting that PMok *plóh is a Malay loanword. Examples are provided in Table 4.

In the first subset, we encounter another unexpected feature: word-final *-h. Final *-s in PMP should be deleted; e.g., PMok *bəlà: ‘unhusked rice’ and *tilí: ‘to leak’, derived from PMP *beRas and *tiRis, respectively. Instead of the expected forms *latú: and *nipí:, we find *latóh and *nipíh. Importantly, final *-h is found only in words not directly inherited from PMP (Wolff 2018, pp. 539–40), and the retention of *-s as *-h aligns with characteristics of northern Malay dialects, such as UL, and the dialects spoken in Satun and Nakhon Si Thammarat (Steinhauer 2008, pp. 127–29; Amon 1987, pp. 19–22). This suggests that Moklenic had contact with the northern dialects rather than those spoken in Melaka and the Riau-Lingga archipelagos.

The next subset presents an additional irregularity in PMok, with the presence of *c^h-. Regular sound changes from PMP suggest that

TABLE 4
Examples of Malay Loanwords in PMok with Phonological Irregularities

<i>Glosses</i>	<i>PMok</i>	<i>Moklen</i>	<i>Moken</i>	<i>SM</i>	<i>UL</i>	<i>PMal</i>	<i>PMP</i>
ten	*plóh	plóh	cəpəh	(sə)puluh	(sə)puluh	*(sA-)puluh	*puluq
hundred	*latóh	latóh	latə:	(sə)ratus	(sə)ratusih	*ratus	*ratus
thin	*nipih	nipih	nipih	nipih	nipis	*nipis	*nipis
breast	*cuc ^h ɟj	tuc ^h ɟj	cəc ^h əj	susu	susu	*susu(?)	*susu
sour	*mac ^h ám	mac ^h ám	mac ^h am	masam	masap	*m-asəm	*ma-qalesam
one	*c ^h á:ʔ	c ^h á:ʔ	c ^h aʔ	sa(tu)	sa	*əsaʔ	*sa
scale	*cəcik	təcik	cəcik	sisik	sisiʔ	*sisik	*siksik
navel	*pəcát	pəcát	pəcat	pusat	pusaiʔ	*pusət	*pusej

the fricative *s- should have dropped out, as seen in *ʔɟj ‘dog’ and *ʔikɟj ‘elbow’ derived from PMP *asu and *siku, respectively (Wolff 2018, pp. 539–40). In PMok, the forms *tuc^hɟj ‘breast, milk’ and *mac^hám ‘sour’ are found instead of the expected *cuʔɟj and *maʔám. The default word for ‘one’ is *c^há:ʔ, a Malay loanword. The original Moklenic form survives in Moklen as the clitic /ʔá-/ used in compounds and with classifiers; e.g., /ʔá-latóh/ (one hundred) and /ʔá-lɛj/ ‘one day’. Moklenic dialects do not have /s/ as an independent phoneme but as a variant of /ch/. The presence of the aspirated palatal stop can be regarded as an approximation of the Malay /s/.

In the last subset, PMP *s- irregularly appears as *c- in PMok, instead of *ʔ, which would yield *ʔəʔik and *pəʔát for ‘scale’ and ‘navel’, respectively. The reason for the unaspirated *c- instead of complete loss or aspirated *c^h- commonly found in Malay loanwords remains unclear.³ Possible explanations include borrowing at an early stage when aspirated stops were not allowed, or origins in a non-Malayic language where PMP *s- became /c-/. Excluding them from the list of Malay loanwords does not affect the overall argument because there are very few cases of PMok *c- for PMP *s-.

Other Malay loanwords cannot be reconstructed for PMP. Table 5 contains a set of forms identified by Adelaar (1992, pp. 200–204) as distinctively Malayic. For instance, Moklen /ŋán/ and Moken

/ŋan/ closely resemble Malay /taŋan/, which is derived from Proto-Malayic *taŋan that replaced PMP *(qa-)lima. As Moklenic does not belong to the Malayic subgroup, morphemes originating in this subgroup and found in Moklenic cannot have been directly inherited from PMP. Intriguingly, all of these items belong to Swadesh's basic vocabulary list of two hundred words.

Additionally, many words not in Adelaar's list are common among Malay dialects, and thus are specifically of Malay origin rather than of other Malayic languages. Many of these words, illustrated in Table 6, have phonological irregularities, such as final *-h, found only in loanwords. Among these, PMok *pədə́h 'spicy' has the front vowel *ɛ, suggesting that the word was imported from UL or other northern Malay dialects. The preposition *datá: 'on', of Malay origin, suggests relatively intense contact between the two languages.

Remarkably, Sanskrit and Austroasiatic loanwords in PMok reveal an intriguing linguistic connection with Malay, as shown in Table 7. The first subset consists of Sanskrit loanwords also present in Malay, which were probably introduced into PMok through contact. The second subset includes words previously proposed as Austroasiatic borrowings, many of which are also found in Malay, indicating they may have entered PMok dialects through contact with Malay and were not directly acquired from Austroasiatic sources. The exceptions, 'bird' and 'duck', were not attested in Malay but may have been borrowed from Acehnese or directly from an Austroasiatic source.

Malay loanwords in Moklenic extend beyond non-basic vocabulary associated with distant entities to basic words such as 'hand', 'rat' and 'fat', as well as non-basic 'spicy', 'shoulder', 'grandchild', 'crab', 'easy' and 'lean'. This indicates very close contact between Malay and Moklenic. Furthermore, not all Malay loanwords in both Moklen and Moken can be traced back to PMok. For instance, Moklen /labóh/ and Moken /labɔh/, meaning 'anchor', cannot be reconstructed at the PMok level because of the lack of regular correspondence. The irregular /o/ in the modern Moklen form suggests that these words were independently borrowed from Malay by each Moklenic dialect. A similar explanation applies to the case of Moklen /buc^hóʔ/ and Moken /bac^hoʔ/, which mean 'rotten'.

TABLE 5
Examples of Malay Loanwords in Moklenic That Are Distinctively Malayic

<i>Glosses</i>	<i>PMok</i>	<i>Moklen</i>	<i>Moken</i>	<i>SM</i>	<i>UL</i>	<i>PMal</i>	<i>PMP</i>
hand	*ŋán	ŋán	ŋan	taŋan	taŋan	*taŋan	*(qa-)lima
seven	*dujú:k	dujú:k	luju:k	tujoh	tujoh	*tujuh	*pitu
grandchild	*cəcə:ʔ	təcə:ʔ	cəcə:ʔ	cucu	cucu	*cucuʔ	*(ma)kəmpu
he, she, it	*ŋá:	ŋá:	ŋa:	ia	ŋa	*ña	*ia
we (inclusive)	*kaməj		kaməj	kami	kami	*kami	*k-ami
to sit	*(mə-)dó:k	dó:k	mədó:k	duduk	duduʔ	*duduk	*untud
person	*ʔolá:ŋ	ʔolá:ŋ	ʔola:ŋ	orang	ʔurak	*uraŋ	*tau
rat	*tikú:	tikú:	tiku:	tikus	tikuih	*tikus	*labaw
fat	*ləmək	ləmək	ləmak	ləmak	ləmaʔ	*ləmək	*miñak
grass	*ləpót	ləpót	ləpət	rumput	rupuiʔ	*rumput	*baliji

TABLE 6
Examples of Malay Loanwords in Moklenic Not on Adelaar's (1992, pp. 200–204) List

<i>Glosses</i>	<i>PMok</i>	<i>Moklen</i>	<i>Moken</i>	<i>SM</i>	<i>UL</i>	<i>PMal</i>	<i>PMP</i>
deer	*ləcʰá:ʔ	ləcʰá:ʔ	ləcʰa:ʔ	rusa	rusa	*rusaʔ	
cow	*ləmú:	ləmú:	ləmu:	ləmbu	ləmu		
spicy	*pədéh	pədéh	pədəh	pədas	pədaih		
bamboo shoot	*ləbó:ŋ	ləbó:ŋ	ləbó:ŋ	rəbuŋ	rəbuk		*rebuŋ
to boil	*ləbúh	ləbúh	ləbuh	rəbus	rəbuih		*luab
medicine	*ʔəbát	ʔəbát	ʔəbat	ubat	ʔubaiʔ		
horn, tusk	*gadiŋ	dadiŋ	gadiŋ	gadiŋ			*uReŋ
on	*datá:	datá:	data:	di-atas	ʔataih	*atas	*qi taqas

TABLE 7
Examples of Malay Loanwords in Moklenic of Austroasiatic or Sanskrit Origin

<i>Glosses</i>	<i>PMok</i>	<i>Moklen</i>	<i>Moken</i>	<i>SM</i>	<i>UL</i>	<i>PMal</i>	<i>Possible source</i>
iron, nail	*bəcʰúj	bəcʰúj	bəcʰuj	bəsi	bəsi	*bəsi	Old Mon <i>birsey</i>
duck	*ʔadá:	ʔadá:	ʔada:	kətam	kətap		PMonic *(ʔa)da:
bird	*cicúm	ticúm	cicum				PMonic *kŋciəm
crab	*kətá:m	kətá:m	kəta:m				PMonic *kntam
grandchild	*cəcə:ʔ	təcə:ʔ	cəcəʔ	cucu	cucu	*cucuʔ	Semelai /cuʔ/
elephant	*gajáh	dajáh	gajah	gajah	gajah		Sanskrit <i>gaja</i>
shoulder	*bahó:j	bəhó:j	bahəj	bahu	bahu	*bahu	Sanskrit <i>bāhu</i>
lean	*kəlóh	kəlóh	kələh	kurus	kuruih	*kurus	Sanskrit <i>kṛśa</i>

Phonological Changes in Moklenic Induced by Contact with Malay

Following Pittayawat's (2005) examination of Moken, we investigated two phonological features in PMok that are a result of contact with Malay. These shared characteristics between the Moken and Moklen dialects indicate their presence in PMok, the common ancestral language of the Moklenic varieties. This section focuses on exploring how PMok acquired these traits through sound changes induced by contact with Malay.

The Moklenic dialects' distinctive aspirated consonants differentiate them from Malay and other Austronesian languages, as well as from PMP with its two-way voiced and voiceless series. However, the complete series of aspirated stops was not present in PMok. Most words with aspirated onsets in contemporary Moklen and Moken dialects are recent loanwords from Southern Thai (Pittayawat 2005, pp. 192–93), as shown in Table 8. These words are likely to have been imported after the fifteenth century, given the change in onset voicing from voiced stops to aspirated voiceless stops that occurred in Thai (Shinnakrit 2020). If they had been borrowed before the fifteenth century, they would have retained the voiced onsets of Old Thai.

Building on the investigation of Malay loanwords in the preceding section, it is evident that Malay introduced aspirated consonants into the PMok sound system. Although most Moklenic dialects exhibit a robust aspirated series, as shown in Table 9 and Table 10, only *c^h is reconstructed as an aspirated stop in PMP, as shown in Table 11. It is worth noting that in both Moklen and Moken, the phoneme /c^h/ can be phonetically realized as either a stop [c^h] or a fricative [s], although the stop [c^h] appears to be more commonly observed. The presence of the other aspirated stops in modern Moklenic dialects can be attributed to loanwords from Thai or sporadic changes. While Pittayawat (2005, pp. 192–93) acknowledges that many Moken forms with initial aspirated stops have /c^h-/, he overlooks their Malay loanword origins.

Intriguingly, a significant number of Austronesian words in Moklenic dialects with /c^h/ discussed in the preceding section can be traced back to Malay words that originally had /s/, as shown

TABLE 8
Some Thai Loanwords with Aspirated Stops in Moklenic Dialects

<i>Glosses</i>	<i>Moklen</i>	<i>Moken</i>	<i>Southern Thai</i>
to use		məc ^h aj	c ^h a:j ⁶
to carry on head	t ^h ú:n	mət ^h u:n	t ^h u:n ³
to strike		mək ^h ɔʔ	k ^h ɔʔ ⁶
to arrive at		kat ^h aŋ	t ^h uŋ ¹
herd	p ^h ú:ŋ	p ^h u:ŋ	fu:ŋ ¹
to cross		mək ^h a:m	k ^h a:m ⁵
to weigh		məc ^h aŋ	c ^h aŋ ⁴
half		k ^h iŋ	k ^h ruŋ ⁴
hole	p ^h lɔ:ŋ	p ^h əlu:ŋ	p ^h lɔ:ŋ ³

Source: Modified from Pittayawat 2005, p. 193.

TABLE 9
Moklen (Bang Sak Dialect) Consonant Inventory

	<i>Labial</i>	<i>Alveolar</i>	<i>Palatal</i>	<i>Velar</i>	<i>Glottal</i>
Unaspirated stop	p	t	c	k	ʔ
Aspirated stop	p ^h	t ^h	c ^h	k ^h	
Voiced stop	b	d		g	
Fricative					h
Nasal	m	n	ɲ	ŋ	
Liquid		l			
Glide	w		j		

Source: Adapted from Pittayawat, Warunsiri and Loss 2022, p. 9.

TABLE 10
Moken (Rawai Dialect) Consonant Inventory

	<i>Labial</i>	<i>Alveolar</i>	<i>Palatal</i>	<i>Velar</i>	<i>Glottal</i>
Unaspirated stop	p	t	c	k	ʔ
Aspirated stop	p ^h	t ^h	c ^h	k ^h	
Voiced stop	b	d	ɟ	g	
Fricative			(s)		h
Nasal	m	n	ɲ	ŋ	
Liquid		l			
Glide	w		j		

Source: Modified from Chantanakomes 1980, p. 53.

TABLE 11
Proto-Moklenic Consonant Inventory

	<i>Labial</i>	<i>Alveolar</i>	<i>Palatal</i>	<i>Velar</i>	<i>Glottal</i>
Unaspirated stop	*p	*t	*c	*k	*ʔ
Aspirated stop			*c ^h		
Voiced stop	*b	*d	*j	*g	
Fricative					*h
Nasal	*m	*n	*ɲ	*ŋ	
Liquid		*l			
Glide	*w		*j		

TABLE 12
Malay Loanwords with *c^h- in Proto-Moklenic

<i>Glosses</i>	<i>PMok</i>	<i>Moklen</i>	<i>Moken</i>	<i>SM</i>	<i>UL</i>	<i>PMP</i>
banana	*pɛc ^h áŋ	pɛc ^h áŋ		pisarŋ	pisak	
civet cat	*mɔc ^h áŋ	mɔc ^h áŋ		musarŋ		*musarŋ
one	*c ^h á:ʔ	c ^h á:ʔ	c ^h aʔ	sa(tu)	sa	*sa
earing	*c ^h ɔbáŋ	c ^h ɔbáŋ	c ^h ɔbaŋ	subarŋ	subak	
iron, nail	*bɔc ^h új	bɔc ^h új	bɔc ^h uj	bəsi	bəsi	
sour	*mac ^h am	mac ^h ám	mac ^h am	masam	masap	*ma-qaləsəm

in Table 12. The aspirated palatal stop /c^h/ played a pioneering role in the development of the aspirated consonant series. Once one consonant of this type became established in the phonological system, it facilitated the incorporation of other consonants with similar features into the system.

The absence of the alveolar fricative /s/ in PMok, which can be traced back to the loss of PMP *s, presented a challenge when incorporating Malay words with /s/ into PMok. Higher-proficiency speakers of Malay would have been able to produce the fricative [s], but those with limited proficiency would have approximated it with an aspirated palatal stop [c^h]. The phoneme was probably acquired

by imitating bilingual early adopters who used [c^h]. This phonetic adaptation would have resulted in the coexistence of the [s] and [c^h] variants, which continues to this day. The prevalence of the [c^h] variant confirms that the majority of PMok-speakers had some proficiency in Malay but limited speaking skills in the language. The adaptation of Malay /s/ to PMok *c^h was limited to loanwords and did not impact the pronunciation of inherited vocabulary. This suggests a widespread but not highly intensive bilingualism among speakers.

Another Mainland phonological trait in Moklenic is the three-way vowel contrast, which sets it apart from the more concise vowel system of PMP. Larish (1999, pp. 318, 394–403) and Pittayawat (2005, pp. 194–95) attribute the vowel distinctions to the influence of Austroasiatic languages or internally driven changes, respectively, but fail to consider the impact of Malay contact on the Moklenic vowel system.

According to Pittayawat (2005, pp. 194–95), Moken has a limited number of mid vowels, which are either derived from the contraction of *iw, such as in PMP *kasiw (reflected as /kaʔe:/ ‘wood, tree’ in Moken), or found in non-Austronesian words such as Moken /p^he:/ from Southern Thai /p^heː/ and /jiʔo:/ from the English “radio”. A brief examination of the Moklen dictionary (Pittayawat, Warunsiri and Loss 2022) reveals that most words with mid vowels in Moklen are loanwords, such as /tənó:t/ ‘sugar palm’ from Southern Thai /tano:t⁴/. But there are a few words with mid vowels in both Moklen and Moken that show a regular correspondence between them, such as /dajó:ŋ/ ‘tall’ and /ʔuʔé:n/ ‘water’. It is possible, therefore, that mid vowels existed as marginal phonemes in PMok. Tables 13, 14 and 15 show monophthongs in stressed final syllables. In unstressed initial syllables, however, the vowels are neutralized with respect to vowel length and vowel qualities. Note that /ə/ represents unstressed neutral vowels in the initial syllable.

Low vowels in PMok can be found in both PMP words and Malay loanwords, which shows they were not solely from loanwords but also from the lowering of PMP high vowels *i and *u to PMok *ɛ

TABLE 13
Proto-Moklenic Inventory of Monophthongs

	<i>Front</i>	<i>Back</i>	
		<i>Unrounded</i>	<i>Rounded</i>
High	*i, *i:		*u, *u:
Mid	*e:	(*ə)	*o:
Low	*ɛ, *ɛ:	*a, *a:	*ɔ, *ɔ:

TABLE 14
Moklen (Bang Sak Dialect) Inventory of Monophthongs

	<i>Front</i>	<i>Back</i>	
		<i>Unrounded</i>	<i>Rounded</i>
High	i, i:	ɯ, ɯ:	u, u:
Mid	e, e:	(ə) ɣ, ɣ:	o, o:
Low	ɛ, ɛ:	a, a:	ɔ, ɔ:

Source: Modified from Pittayawat, Pornpottanamas and Loss 2022, p. 11.

TABLE 15
Moklen (Rawai Dialect) Inventory of Monophthongs

	<i>Front</i>	<i>Back</i>	
		<i>Unrounded</i>	<i>Rounded</i>
High	i, i:		u, u:
Mid	e, e:	(ə)	o, o:
Low	ɛ, ɛ:	a, a:	ɔ, ɔ:

Source: Adapted from Chantanakomes 1980, p. 53.

and *ɔ. In the development from PMP to PMok, two vowel changes occurred, one in stressed final syllables and the other in unstressed initial syllables. The connection between these two changes provides evidence that the lowering process in PMok was influenced by contact with Malay.

With stressed final syllables, lowering primarily applied to closed syllables, except for those ending in PMP final *-q, *-R and *-j, which caused vowel lengthening and blocked the lowering process

(Pittayawat 2005, pp. 194–98). As shown in Table 16, PMP forms with these final consonants are reflected with high vowels in PMok, such as ‘to kill’, ‘lip’ and ‘fire’, while most forms ending in other consonants exhibit low vowels, like ‘bark of tree’, ‘bitter’ and ‘hundred’. Exceptions, like ‘alive’ and ‘wind’, have no clear conditioning factor. But this semi-regular lowering pattern is also observed in Malayic, including Malay dialects spoken in Malaysia and Thailand, such as SM and UL (Adelaar 1992, pp. 45–46; Steinhauer 2008, pp. 125–27).

With the unstressed initial syllable, vowel lowering operates as vowel harmony with the final syllable. Larish (1999, p. 321) notes that vowel lowering in unstressed initial syllables is conditioned by the final syllable through vowel harmony. This is observed in PMok disyllabic words, where the initial syllable’s vowel height is determined by the height of the vowel in the final syllable.

TABLE 16
Examples of Proto-Moklenic Words
That Show the Lowering of Proto-Malayo-Polynesian *i and *u

<i>Glosses</i>	<i>PMok</i>	<i>Moklen</i>	<i>Moken</i>	<i>SM</i>	<i>UL</i>	<i>PMP</i>
bark of tree	*kəlét	kəlét	kələt	kulit	kuleʔ	*kulit
wax/candle	*ləlén	dəlén	lelen	lilin	lilet	*lilin
bitter	*pakét	pakét	paket	pahit	paheʔ	*paqit
behind	*ləkót	ləkót	ləkət			*likud
hundred	*latóh	latóh	latəh	(sə)ratus	(sə)ratuih	*Ratus
chicken	*manók	manók	manək	manok		*manuk
lip	*bibí:n	díbi:n	bibi:n	bibir	bibe	*bibiR
to kill	*munù:k	mənù:k	munu:k	bunuh	bunəh	*bunuq
tail	*ʔikú:n	ʔikú:n	ʔiku:n	ekor	ʔikoj	*ikuR
pig	*babú:j	babú:j	babuj	babi	babi	*babuy
fire	*ʔapú:j	ʔapú:j	ʔapuj	api	ʔapi	*hapuy
alive	*kudíp	kudíp	kudip	hidup	hiduʔ	*qudip
wind	*ʔaŋin	ʔaŋin	ʔaŋin	ʔaŋin	aŋen	*haŋin
thin	*nípíh	nípíh	nípíh	nipis	nipih	*nipsis
(animal) scale	*cəcík	kəcík	cəcik	sisik	sisiʔ	*siksik
pin/needle	*jalúm	dalúm	jalum	jarum	jarup	*zaRum

TABLE 17
Examples Showing Vowel Height Harmony in Proto-Moklenic

<i>Glosses</i>	<i>PMok</i>	<i>Moklen</i>	<i>Moken</i>	<i>SM</i>	<i>UL</i>	<i>PMP</i>
alive	*kudíp	kudíp	kudip	hidup	hidu?	*qudip
to sleep	*mə-tidú:n	didú:n	midu:n	tidor		*ma-tiduR
bark of tree	*kólét	kólét	kølet	kulit	kule?	*kulit
mother	*ʔenónj	ʔenónj	ʔenonj	indonj	ʔinuk	
needle	*jalúm	dalúm	jalum	jarum	jarup	*zaRum
full	*pənú:k	pənú:k	pənuh	pənuh	pənoh	*penuq
moon	*bulá:n	bulá:n	bula:n	bulan	bulat	*bulan
hair	*bulɔ̃j	bulɔ̃j	bulɔj	bulu	bulu	*bulu
louse	*dutɔ̃j	dutɔ̃j	gutəy	kutu	gutu	*kutu

Exceptions are /a/ and /ə/, which do not have high counterparts and are not subject to raising or lowering. This pattern aligns with Pittayawat's (2005, pp. 197–98) findings for Moken and is also evident in Moklen data.

Subsets 1, 2 and 3 in Table 17 show vowel harmony, while subset 4 contains exceptions. Closer examination, however, reveals blocking environments, such as voiced initial consonants discouraging vowel lowering in words like ‘moon’ and ‘two’. These blocking environments also appear in Moken, with gaps in low vowel distribution in the initial syllable. Specifically, /b/, /d/ and /g/ in the initial syllable never precede low vowels (Veena 1980, p. 23). Importantly, the vowel harmony rule is also found in SM and some other Malayic dialects (Adelaar 1992, p. 10; McDonnell 2008, p. 414).

The interconnectedness of lowering and vowel harmony in Moklenic and Malay suggests contact between the two languages. One possible explanation for this shared phonological development is that pre-PMok-speakers, influenced by Malay or internal factors, pronounced high vowels *u and *i with varying heights. For example, the pre-PMok term *tulúnj ‘eggplant’ could have varied as [túlónj], [tolónj] or [təlónj]. Yet, when incorporating Malay words with lowered vowels, pre-PMok-speakers consistently pronounced them with low vowels. Malay /təronj/, likely pronounced as [təronj] in

the Malay source, could have been adopted by pre-PMok as [təlón̄], thereby maintaining close similarity to the Malay pronunciation. The introduction of [ɔ] via Malay loanwords thus established [ɔ] as a distinct phoneme from pre-PMok *u. In response, pre-PMok-speakers reclassified their previously variable pronunciations. Consequently, [təlón̄], reflecting the low vowel realization of pre-PMok *tulún̄, became fixed, yielding the consistent PMok term *təlón̄.

Similar to contrastive aspiration, the differentiation between high and low vowels, and the effect of vowel lowering on native vocabulary, indicates a significant level of bilingualism. PMok-speakers demonstrated their proficiency by accurately reproducing authentic Malay pronunciations in borrowed words.

Moklenic as Part of the Malay Linguistic Ecology

The transfer of language features in the Moklenic-Malay contact situation can provide insight into the dynamics of language contact. If PMok evolved from Malay-speakers shifting to Moklenic, we would expect limited Malay loanwords but significant influence on PMok's phonological characteristics. On the other hand, if PMok originated from Moklenic-speakers borrowing from Malay, barring intense contact, we would anticipate numerous Malay loanwords and limited presence of Malay phonological and grammatical features.

Our findings align with the borrowing scenario under significant cultural pressure described in the subsection titled “Socio-Historical Contexts of Contact-Induced Changes”. In cases of less intense contact, phonological changes induced by contact are not typically observed, and any new phonemes are limited to the borrowed vocabulary. As contact intensifies, however, new phonemes can emerge within the native vocabulary accompanied by phonological rules and allophonic alternations from the donor language (Thomason and Kaufman 1988, pp. 74–76). The high number of Malay loanwords and the introduction of *c^h indicate a significant level of contact between the two languages. Malay also influenced the PMok sound system through the phonemicization of *ε and *ɔ and the adoption of

vowel harmony. The vowel lowering process affects both native and loanwords, emphasizing the impact of Malay on PMok's phonology.

Within Matras's (2020, pp. 241–51) framework, the introduction of *c^h and vowel lowering in PMok represent approximation of Malay pronunciation and reflect widespread bilingualism and an openness towards Malay, considered a prestigious language of wider communication. Some Moklenic-speakers must have spoken Malay and maintained the pronunciation of borrowed words with a certain degree of authenticity. Monolingual Moklenic-speakers, in turn, may have also adopted the innovation from these bilingual early adopters. These results support Andaya's (2008, pp. 173–201) socio-historical context that highlights strong political, economic and sociocultural ties between Moklenic and Malay communities. The presence of Malay elements in PMok reflects a long-standing and intertwined, albeit asymmetrical, relationship between the two groups.

The presence of loanwords in PMok dialects that show characteristics of northern dialects of Malay, such as UL, indicates that the community may have interacted with individuals from the northern regions of the Malay Peninsula rather than power centres along the Strait of Melaka. The exact timing of the contact between Moklenic and Malay prior to the nineteenth century is unclear. But the high degree of homogeneity among Moklenic dialects suggests that diversification occurred relatively recently and that PMok, the common ancestor of Moklenic dialects, was most likely spoken only a few hundred years ago. The absence of Portuguese loanwords in PMok implies that the peak of contact predated the Portuguese conquest of Melaka in the sixteenth century.

It is plausible that ties between the two ethnolinguistic groups began to weaken in the sixteenth century, when the ancestors of Moklenic-speakers migrated northward to the Mergui Archipelago, probably reaching it before the arrival of the British in the nineteenth century. The differentiation of PMok into Moklen and Moken is unlikely to have occurred before the Mergui era. Furthermore, the absence of archaic Thai and recognizable Burmese loanwords suggests that substantial interaction between Moklenic-speakers and

their two expanding neighbours from the north began only recently, casting doubt on the speculation that Moklenic-speaking populations have inhabited the isthmus for over a thousand years (Larish 1999, pp. 416–88).

Moklenic communities have maintained continuous contact with Malay-speakers up to the present day, as evidenced by the inclusion of later Malay loanwords in both Moklen and Moken. Moklenic words like those for ‘anchor’ and ‘rotten’ demonstrate the continued influence of Malay. Moklen people at the northern tip of Phuket Island have intermarried with UL-speakers and replaced the original terms for “father” and “mother” with UL borrowings. Similarly, the Moken community on Rawai beach coexists with UL-speakers in the same village. To understand the dynamics of Malay-Moklenic contact, it is crucial to study the impact of contemporary interactions with UL on different Moklenic dialects alongside the Moken varieties in Myanmar, which are believed to have had minimal contact with UL-speakers.

Conclusion

We found that Malay loanwords significantly influenced the vocabulary and the sound system of the Moklenic languages. In this scenario, close political, economic and sociocultural connections with speakers of Moklenic’s linguistic kin, Malay, resulted in areal convergence. As such, our study aligns with Andaya’s (2008, pp. 173–201) socio-historical context. It also provides insights into the Malay influence on Moklenic languages and contributes to an understanding of the historical development and sociolinguistic dynamics of Moklenic communities. The implications extend beyond linguistics, offering a deeper understanding of the interplay between language, culture and social relationships in the region.

Acknowledgements

This research was part of the project “Research and Documentation of the Moklen Language and Culture in the Southeast Asian Context”

supported by the Institute of Suvarnabhumi Studies, Thailand Academy of Sciences, Humanities, and Arts, and the Ministry of Higher Education, Science, Research, and Innovation. I would like to express my sincere gratitude to the Moklen communities for their unwavering kindness and their invaluable data, and extend a special thanks to Narumon Arunotai and her team, who provided insights and practical advice and facilitated community connections. I appreciate the insightful comments I received at the “Placing Sea Nomads” webinar and “Historical Relationships of East and Southeast Asian Languages 2022” conference. I acknowledge the role of technology in refining this manuscript; ChatGPT3.5 and Paperpal improved language quality. The manuscript was also edited by a native-English speaker, but the content and structure are entirely my own.

Pittayawat Pittayaporn is Associate Professor in the Department of Linguistics and Head of the Center of Excellence in Southeast Asian Linguistics at the Faculty of Arts, Chulalongkorn University, 254 Phayathai Rd., Wangmai, Patumwan, Bangkok, 10330, Thailand; email: Pittayawat.P@chula.ac.th.

NOTES

1. The lexical data were systematically re-transcribed based on analysis in Pittayawat (2005) and Pittayawat, Pornpottanamas and Loss (2022) to facilitate comparison.
2. The label Proto-Moklenic (PMok) in this study is equivalent to Proto-Moken-Moklen (PMM) in Larish (1999).
3. Wolff (2018, pp. 539–40) considers *s-, *c- and *∅- split reflexes of PMP *s-, which he writes as *c-.

REFERENCES

- Adelaar, K. Alexander. 1992. *Proto-Malayic: The Reconstruction of its Phonology and Parts of its Lexicon and Morphology*. Canberra: Pacific Linguistics.
- . 2005. “Structural Diversity in the Malayic Subgroup”. In *The Austronesian Languages of Asia and Madagascar*, edited by K. Alexander Adelaar and Nikolaus Himmelmann, pp. 202–26. London: Routledge.
- Amon Thavisak. 1987. *Malay Dialects in Thailand*. Nakhon Pathom: Institute of Language and Culture for Rural Development, Mahidol University.

- . 1992. *Urak Lawoi'*. Nakhon Pathom: Institute of Language and Culture for Rural Development, Mahidol University.
- Andaya, Leonard. 2008. *Leaves of the Same Tree: Trade and Ethnicity in the Straits of Melaka*. Honolulu: University of Hawai'i Press.
- Anderbeck, Karl. 2012. "The Malayic-Speaking Orang Laut Dialects and Directions for Research". *Wacana: Journal of the Humanities of Indonesia* 14, no. 2: 265–312.
- Ansaldo, Umberto. 2009. *Contact Languages: Ecology and Evolution in Asia*. Cambridge: Cambridge University Press.
- Blust, Robert. 1994. "The Austronesian Settlement of Mainland Southeast Asia". In *Papers from the Second Annual Meeting of the Southeast Asian Linguistics Society*, edited by Karen L. Adams and Thomas J. Hudak, pp. 25–83. Tempe: Arizona State University Program for Southeast Asian Studies.
- . 2010. "The Greater North Borneo Hypothesis". *Oceanic Linguistics* 49, no. 1: 44–118.
- . 2019. "The Austronesian Homeland and Dispersal". *Annual Review of Linguistics* 5, no. 1: 417–34.
- Blust, Robert, and Stephen Trussel. "The Austronesian Comparative Dictionary. Web Edition". 4 July 2023. <https://www.trussel2.com/ACD/>.
- Brunelle, Marc, and Pittayawat Pittayaporn. 2012. "Phonologically-Constrained Change: The Role of the Foot in Monosyllabization and Rhythmic Shifts in Mainland Southeast Asia". *Diachronica* 29, no. 4: 411–33.
- Grey, Russel D., Alexei J. Drummond, and Simon J. Greenhill. 2009. "Language Phylogenies Reveal Expansion Pulses and Pauses in Pacific Settlement [Supplementary material]". *Science* 323, no. 5913: 479–83.
- Hogan, David W. 1972. "Men of the Sea: Coastal Tribes of South Thailand's West Coast". *Journal of the Siam Society* 60: 205–35.
- . 1988. *Urak Lawoi': Basic Structures and a Dictionary*. Canberra: Pacific Linguistics.
- Ivanoff, Jacques. 1997. *Moken: Sea-Gypsies of the Andaman Sea Post-war Chronical*. Bangkok: White Lotus.
- . 2001. *Rings of Coral: Moken Folktales*. Bangkok: White Lotus.
- Larish, Michael D. 1992. "Who Are the Moken and Moklen on the Islands and Coasts of the Andaman Sea?" In *Pan-Asiatic Linguistics: Proceedings of the Third International Symposium on Language and Linguistics, Chulalongkorn University, January 8–10, 1992*, pp. 1305–19. Bangkok: Chulalongkorn University Printing House.
- . 1997. "Moklen-Moken Phonology: Mainland or Insular Southeast Asian Typology?" In *Proceedings of the Seventh International Conference on Austronesian Linguistics*, edited by Cecilia Odé and W.A.L. Stokhof, pp. 125–50. Amsterdam: Rodopi.

- . 1999. “The Position of Moken and Moklen within the Austronesian Language Family”. Doctoral dissertation, University of Hawai‘i.
- Matras, Yaron. 2020. *Language Contact*, 2nd ed. Cambridge: Cambridge University Press.
- McDonnell, Bradley. 2008. “A Conservative Vowel Phoneme Inventory of Sumatra: The Case of Besemah”. *Oceanic Linguistics* 47, no. 2: 409–32.
- Pittayawat Pittayaporn. 2005. “Moken as a Mainland Southeast Asian Language”. In *Chamic and Beyond*, edited by Anthony Grant and Paul Sidwell, pp. 189–210. Canberra: Pacific Linguistics.
- Pittayawat Pittayaporn, Warunsiri Pornpottanamas, and Daniel Loss. 2022. *Moklen-Thai-English Dictionary: A Pilot Version*. Bangkok: Project for the Promotion of Academic Works, Faculty of Arts, Chulalongkorn University.
- Sankoff, Gillian. 2002. “Linguistic Outcomes of Language Contact”. In *The Handbook of Language Change and Variation*, edited by J.K. Chambers, Peter Trudgill, and Natalie Schilling-Estes, pp. 638–68. Malden: Blackwell.
- Shinnakrit Tangsiriwattanakul. 2020. “Ming Shilu as Evidence of Devoicing of Voiced Obstruents in Siamese”. *Journal of the Southeast Asian Linguistics Society* 13, no. 2: 67–95.
- Shorto, Harry. 2006. *A Mon-Khmer Comparative Dictionary*, edited by Paul Sidwell, Doug Cooper, and Christian Bauer. Canberra: Pacific Linguistics.
- Sidwell, Paul. 2013. “Southeast Asian Mainland: Linguistic History”. In *The Encyclopedia of Global Human Migration*, edited by Immanuel Ness, pp. 259–69. Hoboken: Wiley-Blackwell.
- Smith, Alexander D. 2017. “The Western Malayo-Polynesian Problem”. *Oceanic Linguistics* 56, no. 2: 435–90.
- Sopher, David E. 1977. *The Sea Nomads: A Study Based on Literature of the Maritime Boat People of Southeast Asia*. Singapore: National Museum Publication.
- Steinhauer, Hein. 2008. “On the Development of Urak Lawoi’ Malay”. *Wacana: Journal of the Humanities of Indonesia* 10, no. 1: 117–43.
- Thomason, Sarah G., and Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
- Thurgood, Graham. 1999. *From Ancient Cham to Modern Dialects: Two Thousand Years of Language Contact and Change*. Honolulu: University of Hawai‘i Press.
- van Coetsem, Frans. 1988. *Loan Phonology and the Two Transfer Types in Language Contact*. Dordrecht: Foris.

- Veena Chantanakomes. 1980. "A Description of Moken: A Malayo-Polynesian Language". Master's thesis, Mahidol University.
- White, Walter G. 1922. *The Sea Gypsies of Malaya: An Account of the Nomadic Mawken People of the Mergui Archipelago with a Description of Their Ways of Living, Customs, Habits, Boats, Occupations*. London: Seeley.
- Winford, Donald. 2005. "Contact-Induced Changes: Classification and Processes". *Diachronica* 22, no. 2: 373–427.
- Winstedt, Richard. 1960. "An Unabridged English-Malay Dictionary". Singapore: Marican.
- Wolff, John U. 2018. *Proto-Austronesian Phonology and Glossary*, vol. 2. Ithaca: Cornell University Press.