CAMBRIDGE

Cambridge University Press
978-1-108-47633-1 — The Languages of Mainland Southeast Asia
N. J. Enfield
Excerpt
More Information

1

Context

1.1 The Mainland Southeast Asia Region

Mainland Southeast Asia can be broadly defined as the area occupied by present-day Cambodia, Laos, Peninsular Malaysia, Thailand, Myanmar, and Vietnam, along with areas of China south of the Yangtze River. Also sometimes included are the seven states of Northeast India.

There are different interpretations of what is and is not included in MSEA, but a core area is generally assumed (Comrie 2007: 45). MSEA is always taken to include former Indochina – Vietnam, Laos, and Cambodia – together with Thailand, and, usually, Peninsular Malaysia and part or all of Myanmar. In this book, our scope is Greater MSEA, thus including regions of southern China and northeast India, but we will naturally tend towards a focus on Core MSEA (see Map 1.1).

MSEA is a tropical and sub-tropical area with rugged and well-forested hills and river systems running from higher altitudes in the northwest to the plains and deltas of the south. Among the biggest rivers are the Mekong, the Brahmaputra, the Red River in north Vietnam, the Salween and Irrawaddy rivers in Myanmar, the Pearl and Yangtze rivers in China, and the Chaophraya in central Thailand. The lower reaches of these river systems are fertile alluvial plains, which have attracted people partly because of the mobility the environment affords, but also because of their suitability for paddy rice farming. Paddy farming, in which rice plants are kept continually flooded as they grow, requires management of water via systems of dykes and channels (O'Connor 1995, Hartmann 1998; note, though, that shallow-flooded plains do occur naturally, as around the Tonle Sap in Cambodia). This method is significantly more productive than upland dry-field methods, and can support larger populations (Bellwood 1992: 90). It also reduces biodiversity. Geography has played an important role in the historical demography of the area, as outlined in section 1.4 on the history of migrations, livelihoods, communications and politics over the last several millennia.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE
Cambridge University Press
978-1-108-47633-1 — The Languages of Mainland Southeast Asia
N. J. Enfield
Excerpt
More Information



Map 1.1 *Mainland Southeast Asia*
Present-day Cambodia, Laos, Peninsular Malaysia, Thailand, Myanmar, and Vietnam, along with China south of the Yangtze River, and Northeast India.
Sources: Esri, Airbus DS, USGS, NGA, NASA, CGIAR, N. Robinson, NCEAS, NLS, OS, NMA, Geodatastyrelsen, Rijkswaterstaat, GSA, Geoland, FEMA, Intermap, and the GIS user community

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-108-47633-1 — The Languages of Mainland Southeast Asia
N. J. Enfield
Excerpt
More Information

1.2 Mainland Southeast Asian Languages

The degree of linguistic diversity in MSEA (i.e., the number of languages per square km) is high (Enfield 2011b), and it is highest in upland areas. Lower language density in lowland areas is due in part to the effects of geography on the nature of social networks (see Nettle 1999) and in part to politics and the exercise of social power. Historical demographic processes of the kinds discussed in Chapter 2 caused formerly diverse lowland communities in MSEA to become homogenized through a combination of two processes. One process was ethnolinguistic shift. Some groups stayed where they were but stopped passing on their languages and identities to their children, instead adopting the languages and identities of new dominant groups. This process has been taking place for at least 2,000 years in the area and can be observed all over MSEA today. Another process was out-migration, typically to more isolated hill areas (Scott 2009). Geographical isolation is a force that still promotes language diversity in the region, where former diversity of lowland areas is decreasing. Many MSEA languages are heavily endangered (Matisoff 1991a, Enfield 2006a, Bradley 2007, Premsrirat 2007). This is exacerbated by effects of the concentration of political power of modern nation states in the lowlands. In recent decades, processes of language standardization in MSEA nations (Simpson 2007) are heavily reducing language diversity.

The languages of MSEA are from five major language families:[1] Sino-Tibetan, Tai-Kadai, Hmong-Mien, Austroasiatic, and Austronesian.[2] More than 550 distinct languages are spoken in greater MSEA. If we exclude the China and India data, thus representing the core MSEA area, the number of languages is about half this amount; see Table 1.1.[3]

The high linguistic diversity in northeast India and southern/southwestern China adds dramatically to the number of languages included in the estimate for this area. It also reverses the relative proportion of Sino-Tibetan and Austroasiatic languages.

[1] Languages of other language families not noted in Table 1.1 include widely used languages of colonial origin – English and French – along with other languages used in the MSEA area by (descendants of) migrants and travellers. These include languages of South Asia and the Middle East, and languages from Japanese to Yoruba that may be encountered in major urban centres of MSEA such as Bangkok, Phnom Penh, or Hanoi. They also include sign languages of the Deaf (see section 2.9). In Myanmar, the Indo-Aryan language Rohingya (arguably in a dialect relationship with Bengali) is spoken by around a million people in Rakhine State.

[2] The Andamanese languages are located just outside MSEA as defined here. For work on these lesser-known languages, see the bibliography in Comrie and Zamponi (2019).

[3] These figures were assembled by Weijian Meng. Language coordinates are from: Glottolog 4.0, https://glottolog.org (accessed 3 July 2019). Administrative borders are from Natural Earth, www.naturalearthdata.com (accessed 26 Aug 2019), river centre lines are from Natural Earth www.naturalearthdata.com (accessed 5 Sep 2019) and the Harvard WorldMap, https://worldmap.harvard.edu/data/geonode:chiangyangtze_river_tnc (accessed 26 Aug 2019). Core MSEA was defined for this count as Cambodia, Laos, Myanmar, Thailand, and Vietnam; Greater MSEA includes this together with Peninsular Malaysia, areas of India east of 90 degrees (i.e., the states of Arunachal Pradesh, Nagaland, Manipur, Mizoram, Assam, Meghalaya, and Tripura) and China south of the Yangtze river (specifically, the provinces of Zhejiang, Jiangxi, Hunan, Guizhou, Yunnan, Guangxi, Guangdong, Fujian, and Hainan).

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE
Cambridge University Press
978-1-108-47633-1 — The Languages of Mainland Southeast Asia
N. J. Enfield
Excerpt
More Information

Table 1.1 *Languages of MSEA: (a) numbers of languages in Core MSEA, by family; (b) numbers of languages in Greater MSEA, by family*

	Core MSEA		Greater MSEA	
Austroasiatic	122	(44%)	138	(24%)
Sino-Tibetan	74	(26%)	288	(49%)
Tai-Kadai	51	(18%)	93	(16%)
Austronesian	25	(9%)	26	(4%)
Hmong-Mien	8	(3%)	38	(7%)
Total	280		583	

Core MSEA is defined as Cambodia, Laos, Myanmar, Thailand, and Vietnam. Greater MSEA is defined as Core MSEA plus states of India to the east of 90ºE, China south of the Yangtze river, Peninsular Malaysia.

There is good agreement among specialists of MSEA languages as to the basic language family affiliation of known languages. There are unresolved issues about lower level subgroupings and there are unresolved hypotheses about possible macro-groupings. But for every known language, scholars agree about which of the five main language families it fits into. This is unusual firstly because it means that each language's basic affiliation is apparently uncontroversial, and secondly because no language isolates have yet been identified (cf. Blench 2011: 125–6). To put this in context, there are an estimated 136 language isolates in the world, of which nearly two-thirds come from the Americas, and only 7% come from anywhere in Asia (and none from MSEA; Campbell 2010). For more on the historical linguistic background of MSEA, see Chapter 2.

This book provides information on the history and classification of MSEA languages, and the main linguistic properties of the languages. The scope is necessarily limited, given the number and diversity of languages under discussion. For more detailed coverage of specific sub-areas or subsets of languages in the area, see Jenny and Sidwell (2015) on Austroasiatic languages, Diller *et al.* (2008) on Tai-Kadai languages, Thurgood and LaPolla (2017) on Sino-Tibetan languages, Grant and Sidwell (2005) on Chamic and other mainland Austronesian Languages, and Ratliff (2010) on Hmong-Mien languages, along with Simpson (2007), Goddard (2005), Comrie (1990, 2007), and Vittrant and Watkins (2019) on MSEA more broadly, and the many references in all of those sources, as well as in this book.

1.3 Nomenclature, System Ontology, and Language Data Selection

In any discussion of languages, language families, and ethnic or cultural groups, we inevitably encounter problems of nomenclature. Countless terms in hundreds of

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE
Cambridge University Press
978-1-108-47633-1 — The Languages of Mainland Southeast Asia
N. J. Enfield
Excerpt
More Information

languages have been used to label human groups in present-day and historical MSEA. Not one of them refers to a homogenous entity. Like any other linguistic category, a label for a human group is a useful but imperfect tool that glosses over differences between members of the category.

This glossing over of difference can be used for good or ill. The utility of labels for human groups is not only practical but political. This is obvious in the case of terms like Thai and Vietnamese, which can refer to identities associated with nation states, or terms like Hmong and Karen whose members are politically organized. But it is no less true in the case of names for small ethnolinguistic groups such as Kri and Chong. And the referents of labels for historical groups, such as Pyū and Hoabinhian, can be similarly strategic – and potentially misleading – in their implication of category-internal uniformity (see below for discussion). As we shall see, there is little evidence that these terms refer to people who shared a unitary linguistic, cultural, or ethnic profile.

All ethnolinguistic terms are fluid and contestable. They all involve some form of essentialization and/or naturalization by insiders and outsiders, including academics and other researchers. Such essentialization or naturalization can be problematic if it is based on false or inaccurate understanding, or if it is put to certain prejudicial uses. But it is ultimately necessary for any discourse about language, for two reasons. First, we cannot talk about any natural phenomena without generalizing in some way. The issue is not whether we generalize. It is whether we generalize in relevant, accurate, fair, and productive ways. Second, while there is always variation within any language ecology, languages exist. They have system properties. They enable communication, just as they create real barriers to communication. They serve conventional and collective functions within definable human groups, and as such they bind groups of people together, defined as those who are socialized within a shared linguistic system, who have an understanding of that system, and who can use it.

Of course, the names of language systems do not equal the systems themselves. There are important issues concerning ethnolinguistic nomenclature and its uses. One problem is that most languages or human groups can be referred to by more than one label. An example in MSEA is the Northern Tai language spoken in certain villages of the middle Nam Noi valley in upland Central Laos. This language is referred to as Saek [sæ:k] in the linguistic literature (e.g., Hudak 2008: 53–8, 2010), [tɕʰɛ:k] by the villagers themselves, [zɻəː?] by Kri speakers in neighbouring villages a few kilometres upstream, and [jɔ:j] by Brou speakers in neighbouring villages a few kilometres downstream. This sort of situation is globally the norm. English, for example, is also called *Engels*, *Anglais*, *ʔaŋkit*, and *Yīng Wén*, among many other names, depending on the language spoken. It is important to keep track of these different labels.

A second issue is that language names can change over time, as can the connotations of those names, and in turn people's preferences and practices in language naming.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE
Cambridge University Press
978-1-108-47633-1 — The Languages of Mainland Southeast Asia
N. J. Enfield
Excerpt
More Information

A language name that is established and accepted now may become marked or pejorative later, and speakers of the language may come to prefer that the name not be used. For example, Lao speakers today widely refer to the Hmong language as *mong4*. Another word for the language, and the ethnic group more broadly, is *mèèw4* (cf. Chinese *Miao*). A few decades ago, the word *mèèw4* was an unmarked term in Lao, but today it is considered by many Hmong, and some Lao, to be pejorative and offensive.

A third problem concerning the labelling of linguistic entities is the theoretical question – alluded to already – of whether there is in fact any real entity to be labelled. A classical puzzle in linguistics is to determine, for two language varieties, whether the varieties are two separate languages or two dialects of the same language. One way to answer the question is with reference to the criterion of mutual intelligibility. If the two varieties are mutually intelligible, despite certain readily defined differences, then they are dialects of a single language. If they are not mutually intelligible, then they are separate languages. Another way to answer the question is with reference to sociopolitical criteria. If the labeller's goal is to demarcate two communities as having separate identities, then their varieties may be labelled as different languages, even if they are mutually intelligible. This is the case, for example, for the two language varieties of MSEA known as Thai and Lao. On linguistic grounds, their high degree of mutual intelligibility would lead us to conclude that they are dialects or varieties of a single language. Conversely, if our goal of labelling were to convey that the two groups are part of one sociocultural entity, then we might want to label these significantly different language varieties in the same way. This explains why Mandarin, Cantonese, and Shanghainese are often referred to as 'dialects of Chinese'. These three language varieties are mutually unintelligible and would on technical linguistic grounds be regarded as separate languages. But on sociopolitical grounds, they may be grouped together on the basis that they are spoken by communities who are regarded as part of a single Sinitic cultural sphere.

When we consult secondary sources, it is important to be mindful of the possibly quite distinct reasons a certain language name may have been used. There are different kinds of categorization, and it is easy to mistake one kind of categorization for another. Many first-time travellers to Southeast Asia will assume that Thai and Lao are not mutually intelligible, simply because they are referred to as different languages and are written using different scripts. But their categorization as different languages is made on primarily political grounds. Particular caution is needed with linguistic categorization and nomenclature offered by non-specialists, such as in the writings of early colonial surveyors and explorers, or modern-day tourists and amateur observers. On the problem of defining language and culture groups in Myanmar, Watkins (2007: 277) remarks: 'The taxonomic fervour of colonial ethnographic surveyors and their lack of accuracy in

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE
Cambridge University Press
978-1-108-47633-1 — The Languages of Mainland Southeast Asia
N. J. Enfield
Excerpt
More Information

discerning genuine divisions among continua of cultures and languages has been frequently remarked on with regard to South Asia, and a warning about the over-simplification of ethno-linguistic categorization is equally well warranted in the context of Burma/Myanmar'. The same is true for MSEA more broadly.

This point has been appreciated in postcolonial anthropology since at least Leach (1954). Leach's classic study of the social interrelations among highland Burmese groups established that the classical assumption of ethnography – 'treating culture groups as social isolates' – does not hold up:

> In the Kachin Hills Area as a whole we find a considerable number of named groups culturally distinct or partly distinct. In places these groups are segregated into fairly well-defined areas, in other places they are all jumbled up. A study of Kachin social organization cannot therefore proceed in the classical manner which treated culture groups as social isolates.
>
> (Leach 1954: 60)

Leach showed that 'the social system is not uniform'. Instead, people of different culture groups – people who speak different languages and identify as ethnically distinct – are 'part of a single social system', where the possibility of shifting cultural identity is natural and commonplace (Leach 1954: 60). There may be periods of equilibrium, but 'any such equilibrium as may appear to exist may in fact be of a very transient and unstable kind' (Leach 1954: 61).

We can draw an analogy between the stability of ethnic groups and the stability of landscape features around small-scale alluvial river systems. When I first travelled to the upland riverine area of Laos in which Kri speakers live, I naïvely assumed that the streams, pools, crossings, and riverbanks I encountered were permanent features. I was puzzled that stairways and bridges were only crudely constructed with a minimum of effort, and thus with no longevity. In subsequent trips I learned that river morphology can develop rapidly. Changing flows quickly transform streams into stagnant ponds, washing old walking trails away, opening up new lines of access, and so on. The same fluid nature is found in social systems, ethnic identities, and inter-ethnic relationships. What upon first impression may seem like an established structure can be quickly reorganized. People will adapt to the changes with equal flexibility.

This is the essence of Leach's point about the dynamic social system that Shan people and Kachin people shared in the Kachin Hills. Their interdependent social subsystems allowed such flexibility that it was possible to speak of 'Kachins becoming Shans or of Shans becoming Kachins' (Leach 1954: 61). The system 'has *no* stability through time', Leach suggested. 'What can be observed now is just a momentary configuration of a totality existing in a state of flux' (Leach 1954: 63).

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE
Cambridge University Press
978-1-108-47633-1 — The Languages of Mainland Southeast Asia
N. J. Enfield
Excerpt
More Information

This leads us to a more fundamental question about labelling languages. If a language is an imagined entity and/or merely a momentary configuration of something that is in flux, do we have the right to label languages and thereby treat them as things at all? There are real issues of language ontology, genuine questions about what kind of a thing a language is (Harris 1980). But these issues should not prevent us from describing such systems as if they were stable structures. For one thing, the fact that the systems are distributed, dynamic, and reconfigurable does not mean that the systems do not exist. They do exist, as evidenced by the fact certain groups of people collectively understand, and can productively use, certain sets of words, phrases, idioms, and not others. Another reason is that the dynamic, distributed processes at hand are not always obvious or even detectable within our lifetimes, so from the point of view of the individuals who live within these systems, as members of social groups and as agents of cultural transmission, the systems in question may be effectively fixed and permanent. A third reason is that without labelling them and treating them as real, we would be unable to talk about, analyse, or understand the linguistic phenomena we wish to focus on in this work.

Ontological questions about the reality of social systems like languages and cultures are dealt with by taking the nuanced conceptual approach that social reality demands.[4] As Searle (2010) argues, a piece of social reality can be both subjective and objective in nature, because it is *ontologically subjective* (its existence depends on human perspective and experience) while being *epistemically objective* (the truth of claims about it can be settled as a matter of fact). In any case, for the purposes of this book we need to leave these questions aside and get on with the job. This may be an act of essentialism. If so, it is a strategic one. It is necessary for our purposes.

The naming of language *families* runs up against two further problems in addition to those just discussed for the labelling of languages and sociocultural groups.

The first problem with language-family labels is that individual language names are often used to index entire groups of languages. In MSEA linguistics, language-family labels including Mon-Khmer, Hmong-Mien, and Tibeto-Burman do this, following a common recipe. They pair two individual language names to denote a group of many languages. The languages that are foregrounded in the family labels happen to be spoken by the largest populations and happen to be culturally and sociopolitically

[4] Discussing anthropological research in our area, O'Connor (1995: 968–9) notes: 'Once ethnologists studied Southeast Asia as a region (Heine-Geldern 1956; Sharp 1962; Burling 1965; Provencher 1975; Keyes 1977), but their empirical generalizations have given way to an anthropology of discrete cases. While archaeologists still piece together regional types (Bellwood 1985; Higham 1989), ethnographers dote on the people or village they study … [M]any ethnographies naively treat their subjects as if they existed and were knowable apart from larger conditions and logical models. In this fieldwork-driven empiricism where anthropology collapses into ethnography (Stocking 1992: 357, 362–72), local description, synchronic constructs, and inductive reasoning have come to displace the regional comparison, historical perspective, and deductive logic that rigor and balance require.'

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE
Cambridge University Press
978-1-108-47633-1 — The Languages of Mainland Southeast Asia
N. J. Enfield
Excerpt
More Information

dominant. For this reason (in part), Blench and Post (2014: 92–3) regard Sino-Tibetan as a 'highly inappropriate name for the phylum'. They say the same for Tibeto-Burman. They argue that the named languages have 'no special classificatory significance' and that the political or cultural prominence of these languages is not a good reason to foreground or privilege them. Furthermore, privileging them in this way comes at the expense of sociopolitically less powerful groups of people, or of languages that do have special classificatory significance. Blench and Post (2014: 93) endorse an alternative term Trans-Himalayan, which they say 'would capture the geographical locus of much of the phylum without emphasizing individual subgroups'.

But this strategy of using a geographical designation to label a language family raises a second potential problem, as LaPolla writes:

> If we look at the various proposals for subgrouping in Tibeto-Burman, almost all include geographic designations such as 'Western Himalayan'. These designations assume that either the languages involved have always been at that location or that all the languages developed from a single ancestor which migrated to that location at some time in the past. Yet we know that there have been waves of migration, particularly into and/or through certain regions (see LaPolla 2001), and so there is a suspicion that the languages were not originally closely related, but have come to seem similar because of long-term contact.
>
> (LaPolla 2012: 120)

Issues of nomenclature are important, particularly for our practices in future research in which we have a responsibility for referring to language varieties in clear, informed, and appropriate ways. Labels for ethnic groups are not straightforwardly discovered or created. Linguistics has to make do with secondary sources, and often the materials we consult are produced in very different contexts from our own, by people with different motivations, standards, and knowledge from our own.

However, with all of this in mind, it would be impossible to revise or update all of the terminology in the literature, or to otherwise adjudicate on the implications, connotations, or appropriateness of every language or culture label cited in this book. I therefore make no attempt to change the nomenclature derived from secondary sources. I will note if or when clarification of reference to a language or a language family is needed.

The language selection drawn on in this book is, by necessity, a selection of convenience. Of the many hundreds of languages in the MSEA area, only a small fraction have been described in detail, and so there is a bias towards languages that happen to be well described in published sources. A problem with this is that the best-described languages may not necessarily be the most important for research purposes, depending on the question (though they may be of broader public interest).

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE
Cambridge University Press
978-1-108-47633-1 — The Languages of Mainland Southeast Asia
N. J. Enfield
Excerpt
More Information

If we are going to understand human language, then we need representative data. Our null hypothesis must be that each language is of equal value as a data source (with the caveat that many languages do not represent independent data points because of their historical relationship with each other; see Galton's remarks in Tyler 1889). The starting assumption has to be that a language like, say, Kri, spoken by 500 people in an isolated upland pocket of central Laos, will be no less important as a source of evidence for a science of language than Vietnamese, which is spoken by up to a hundred million people. That said, there are reasons why we might want to concentrate our efforts in research on a big language like Vietnamese. Such work might be of interest to larger numbers of people (including non-specialists). And the available information on Vietnamese is likely to be more reliable, detailed, and extensive. But to the extent possible, in this book I have endeavoured to shift the balance to lesser-known languages.

My chosen emphasis on lesser-known languages raises two points I would like to emphasize. The first is that even with a significant recent increase in the amount and quality of research reports on lesser-known languages, we are still in need of good quality information about the many languages of MSEA: field research with, and by, speakers of the languages is needed. The second point is that when you think of MSEA languages, you should think first of minority, non-official languages like Semelai, Lahu, Saek, Moken, Kri, or Mien, and not those that are usually thought of first or cited most often: Thai, Lao, Khmer, Vietnamese, or Burmese. As we shall see through the rest of this book, the minority languages are more representative of the area, in terms of history, sociolinguistics, and typology.

1.4 History of MSEA

MSEA has seen a long and complex history of human movement, contact, and diversification, leading to the linguistic diversity that we see today. This section provides an outline of the historical processes that – often indirectly – help to answer questions of what MSEA languages are like and why.

1.4.1 Prehistory

The land mass that we now identify as MSEA is a recent formation in geological time. The mainland area as we know it was formed by processes of separation, rotation, and collision of tectonic plates over many millions of years. By around 40 million years ago, the MSEA area had begun to take recognizable shape, most notably due to the collision with Asia of the Indian subcontinent, moving up from the southwest, and the coming together of the Indochina, Shan-Tai and south China cratonic areas (Bunopas and Vella

© in this web service Cambridge University Press www.cambridge.org