Languages of South East Asia and China

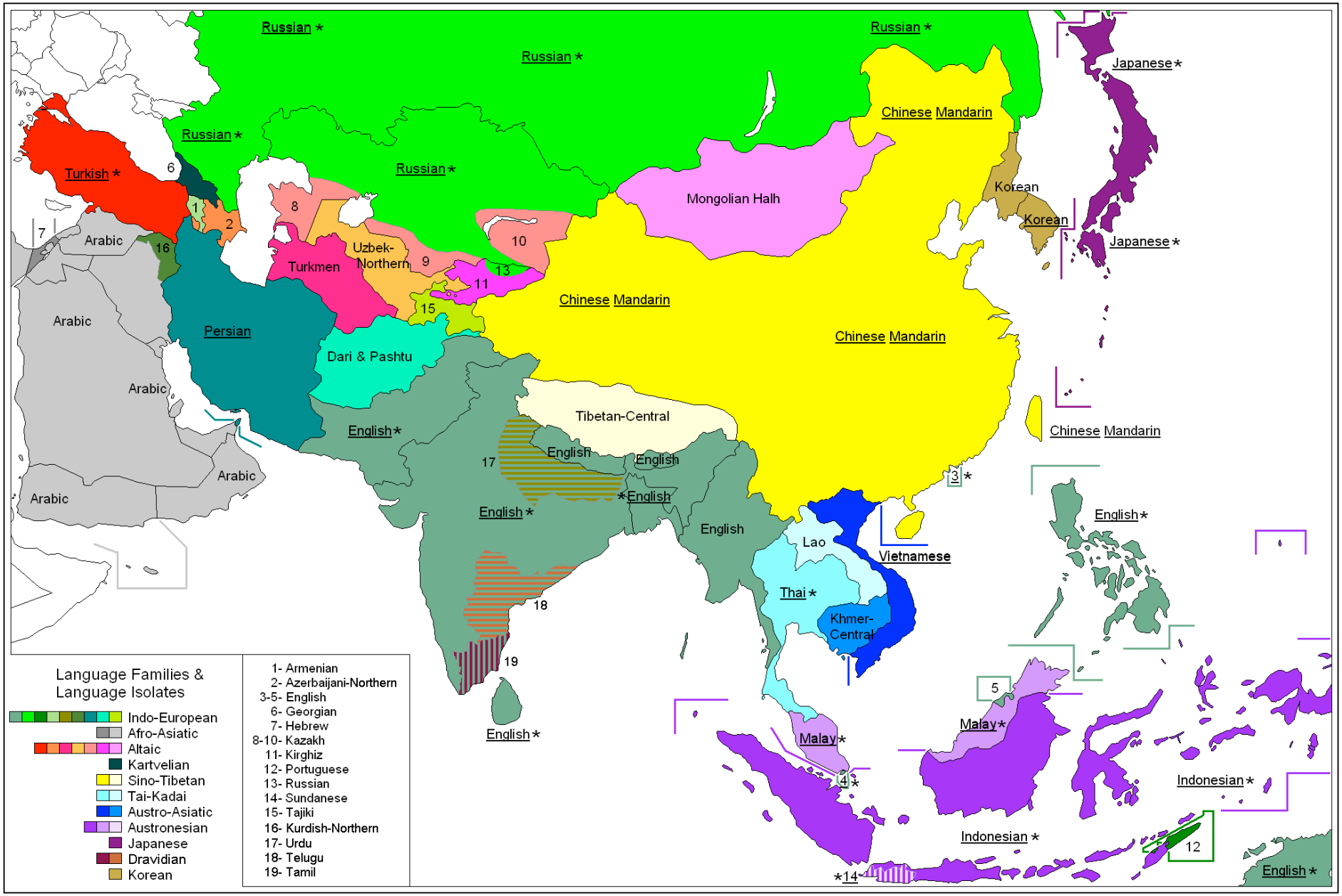
Francis Bond
based on slides by František Kratochvíl

2025

Overview

- Recognised language families of the region
- Impact of geography on human population movement, linguistic diversity and cultures of MSEA
- Indospheric and Sinospheric influences
- Typological features of MSEA languages
  - Lexical tone
  - Phoneme inventories
  - Zero anaphora
  - Classifier systems
  - Serial verb constructions
- What makes MSEA a linguistic area, and why?

Official Languages



The major language families of SEA

- **Austronesian**: mainly in coastal regions of Vietnam, Burma and western peninsular SEA.
- **Austroasiatic** (Mon-Khmer branch – Munda branch is only found in South Asia): widely scattered throughout MSEA region and extending into NE India (Khasian branch).
- **Sino-Tibetan**
  - **Tibeto-Burman** Burma, NE India, mountainous regions of Western and northern Thailand, Laos and northern Vietnam.
  - **Sinitic** predominantly China
- **Tai-Kadai**: southern China, SEA, NE India.
- **Hmong-Mien**: (a.k.a. Miao-Yao): mostly in southern China, extending into SEA.
- **Korean, Japanese and Ainu**: Korean Peninsular and Japan

MSEA forms a linguistic area in which many languages share similar features despite being geneDcally unrelated.

The countries/regions of MSEA

- Burma
- Cambodia
- Laos
- Malaysia
- Thailand
- Vietnam
- NE India (arguably eastward of the south bank of the Brahmaputra River)

Much of Mainland Southeast Asia is very mountainous. The people of the uplands have long resisted being governed by central states of the lowland basins.

Human and linguistic diversity in MSEA

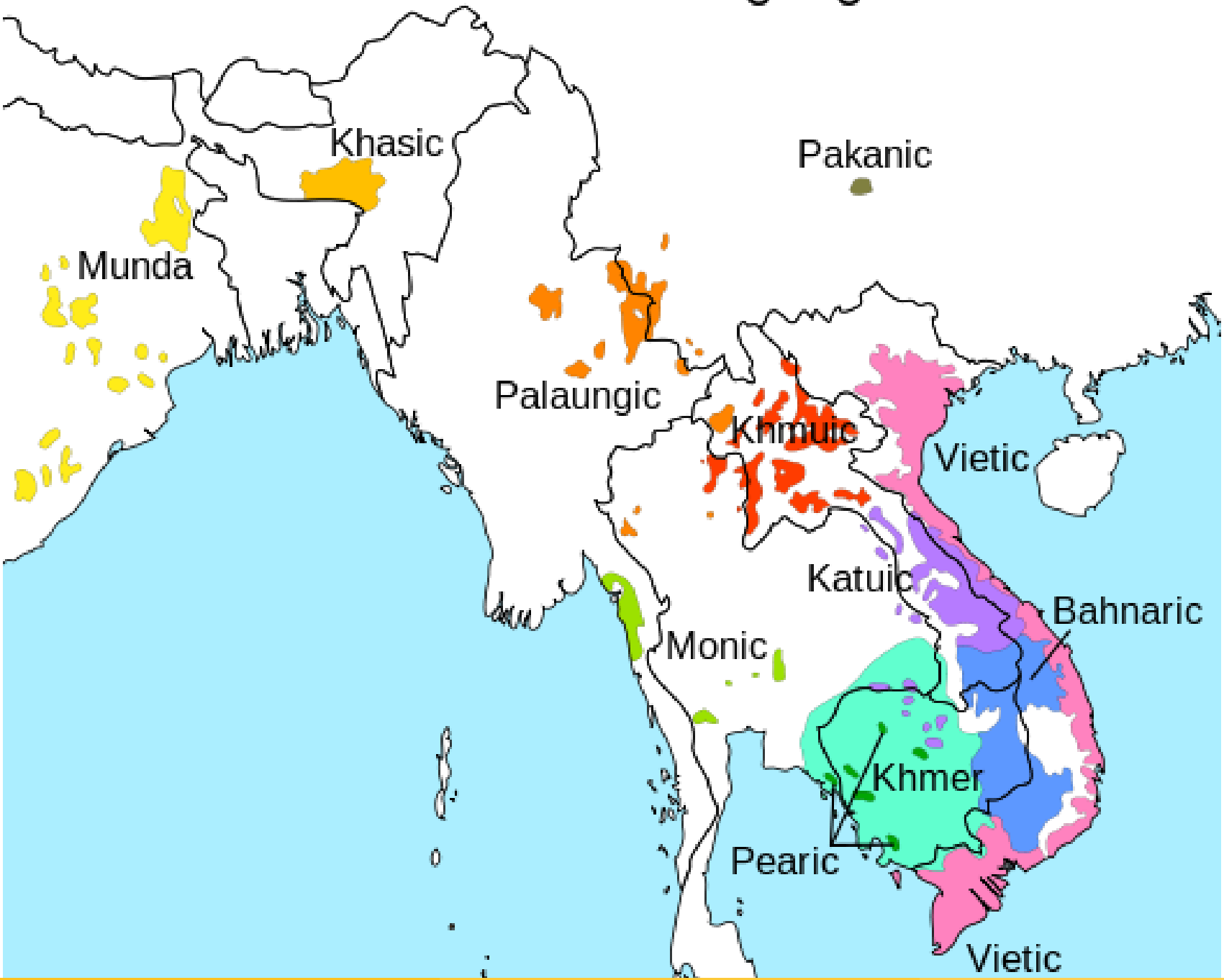
- Mainland Southeast Asia (MSEA) is recognised for its high degree of human diversity in comparison to the global average.
- This refers to diversity in all measures of distinction among human groups, including social structure, material culture, and genetic markers.
- The diversity of languages, however, has been regarded as relatively low, although they number around 2000. Why?

- **Language Families**: Many languages in MSEA belong to a small number of language families (Austroasiatic, Tai-Kadai, Sino-Tibetan, Hmong-Mien), leading to similarities that reduce the perception of diversity.
- **Contact and Convergence**: Long history of contact, trade, migration, and cultural exchange has led to linguistic convergence, creating similar features across unrelated languages.
- **Multilingualism and Language Shift**: High rates of multilingualism and the use of lingua francas (e.g., Thai, Vietnamese) often overshadow smaller languages, making linguistic diversity less visible.
- **Linguistic Classification**: Differences in how linguists classify languages and dialects can influence the perceived number of languages, sometimes reducing the apparent diversity.

Austroasiatic

The Austroasiatic family (AA) forms two major branches: **Munda**, spoken only in India, and **Mon-Khmer**, constituting more than 120 languages spoken in an area extending from Meghalaya state in NE India to Vietnam in the east, and the Malay peninsula in the south.
AA languages are also spoken in China.

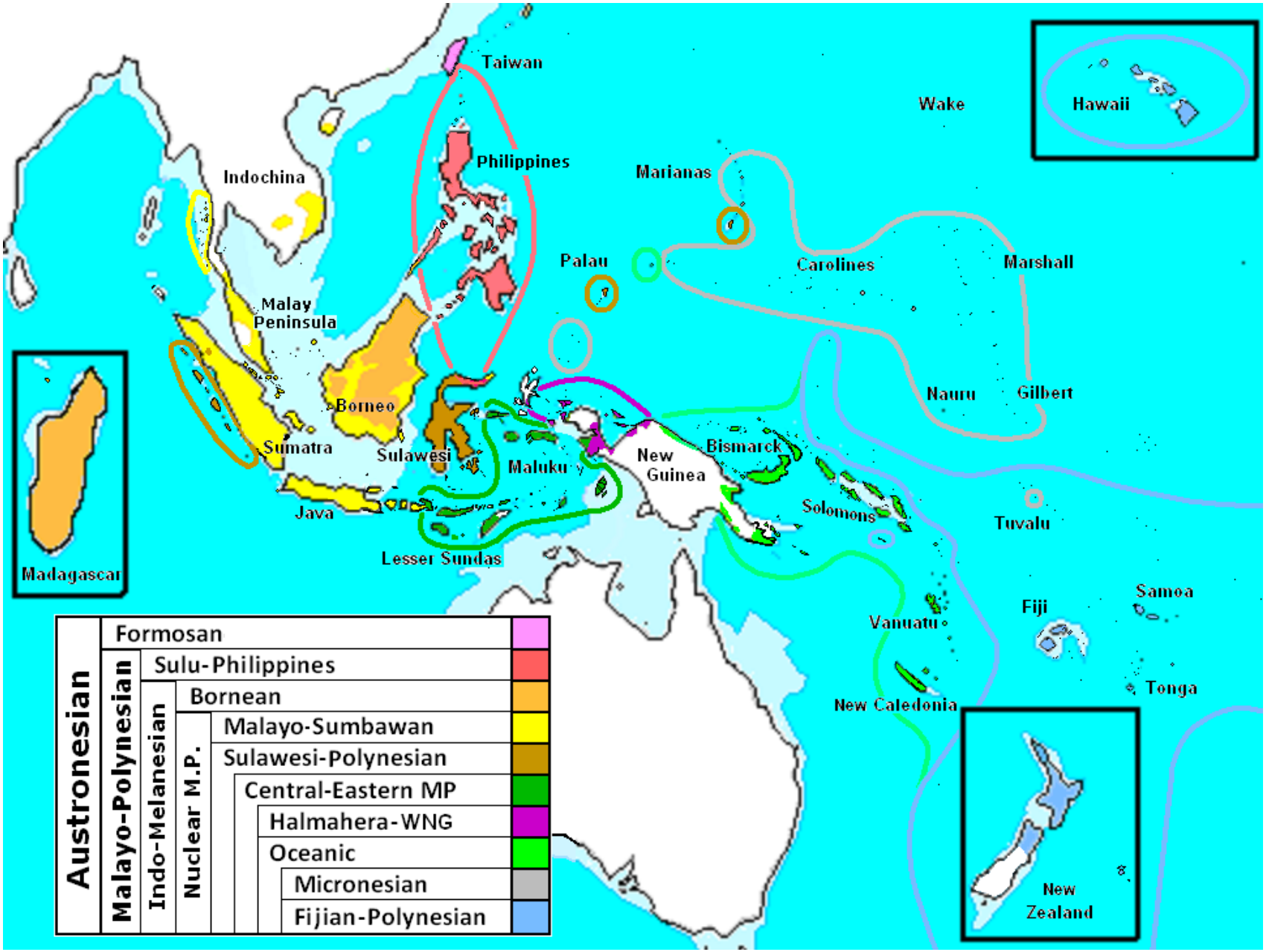
Austro-Asiatic Languages



Austronesian

Austronesian is the largest and most widespread family in the world, with somewhere around 700 (maybe as many as 1,200) languages altogether and 300 million native speakers. Aside from Southeast Asia, Austronesian languages are found on numerous islands in the eastern and central Pacific Ocean all the way to Easter Island. There is also a western outpost language (Malagasy), spoken on the island of Madagascar.
Major languages include:

- Malay (200 million speakers, about 40 million as a first language) Indonesian (Bahasa Indonesia) and Malaysian (Bahasa Melayu)
- Javanese (75 million)
- Sundanese (30 million)
- Pilipino (Tagalog) (50 million, 17 million as a first language).


Taiwan
Philippines
Indochina
Malay Peninsula
Borneo
Sumatra
Sulawesi
Java
Lesser Sundas
Maluku
New Guinea
Bismarck
Solomons
Wake
Hawaii
Marianas
Palau
Carolines
Marshall
Nauru
Gilbert
Tuvalu
Fiji
Samoa
Tonga
Vanuatu
New Caledonia
New Zealand
Madagascar
Austronesian
Formosan
Malayo-Polynesian
Sulu-Philippines
Indo-Melanesian
Bornean
Nuclear M.P.
Malayo-Sumbawan
Sulawesi-Polynesian
Central-Eastern MP
Halmahera-WNG
Oceanic
Micronesian
Fijian-Polynesian

Mon-Khmer branch of Austroasiatic

- Aslian (mostly endangered), spoken in southern Thailand and Western Malaysia.
- Katuic-Bahnaric, spoken in Vietnam, Laos, NE Thailand and Cambodia.
- Northern Mon-Khmer, which divides into Palaungic and Khumic, spoken in Yunnan and adjacent areas of Burma, Northern Thailand, and NE Laos.
- Vietic languages, spoken in Vietnam (including Vietnamese)
- Monic, spoken by increasingly diminishing numbers of speakers in Burma and Central Thailand.
- Pearic, spoken by small groups in Western Cambodia and adjacent areas of NE Thailand.
- A Pakanic branch has recently been proposed for a group of languages spoken in northern Vietnam, and in Guanxi and Yunnan, China.

There are many competing classifications of the Mon-Khmer branch in the literature, none of which has received universal acceptance.

Major Mon-Khmer languages

- Khmer (Cambodian, about 13 million speakers)
- Vietnamese (nearly 70 million)

Tibeto-Burman

Most of the roughly 300 Tibeto-Burman languages are found in the South Asian region, but many of them straddle the buffer zone between South Asia and Southeast Asia.
Overall, the Tibeto-Burman languages can be quite different from Sinitic languages, with agglutinative morphology, verb final word order, and postpositions.
Major languages:

- Burmese (30 million speakers)
- Tibetan (5 million)
- Karen (4 million combined: S'gaw is the largest, at about 2 million)
- Lolo (Yi) (3 million, spoken in China)
- Bai (1 million, spoken in China).


MONGOLIA
Northern Mandarin
Northwestern Mandarin
CHINA
Northern Mandarin
Ladakhi
TIBET
Amdo
Central Tibetan
Khams
Dzonghkha
River Mandarin
Wu
Wu
Southwestern Mandarin
NEPAL
BHUTAN
Bai
Gan
Min
Southwestern Mandarin
Xiang
INDIA
BANGLA DESH
Kejia
Min
Burmese
Cantonese
TAIWAN
MYANMAR
LAOS
Min
Hainan Island
Bay of Bengal
THAILAND
South China Sea
Andaman Sea
VIETNAM
Gulf of Thailand
Chinese
Tibetan
Rungic
Tani
Konyak
Naga
Mikir-Meithei
Kuki-Chin
Bodo-Garo
Lolo-Burmese
Kachinic
Karenic
Bai
Qiangic
© 2013 A. Gutman & B. Avanzati

Sinitic Languages

Sinitic languages are spoken by a huge number of people (over 1,000 million) in mainland China, and Taiwan, and in Southeast Asia.

- Mandarin
  - Northern
  - Jiang-Huai/Southeastern (Jiangsu and Anhui)
  - Southwest (Hubei, Sichuan, Yunnan, Guizhou, eastern Hunan, western Guangxi)
- Wu (Southern Jiangsu, Zhejiang)
- Xiang (Hunan)
- Yue (Guangdong, Guangxi, parts of Hainan) Cantonese
- Min (Fujian, Taiwan, parts of Hainan) Hokkien
- Gan (Jiangxi, north-eastern Guangdong)
- Hakka (southern Jiangxi, north-eastern Guangdong, western Fujian, part of Sichuan, Guangxi, and Taiwan)

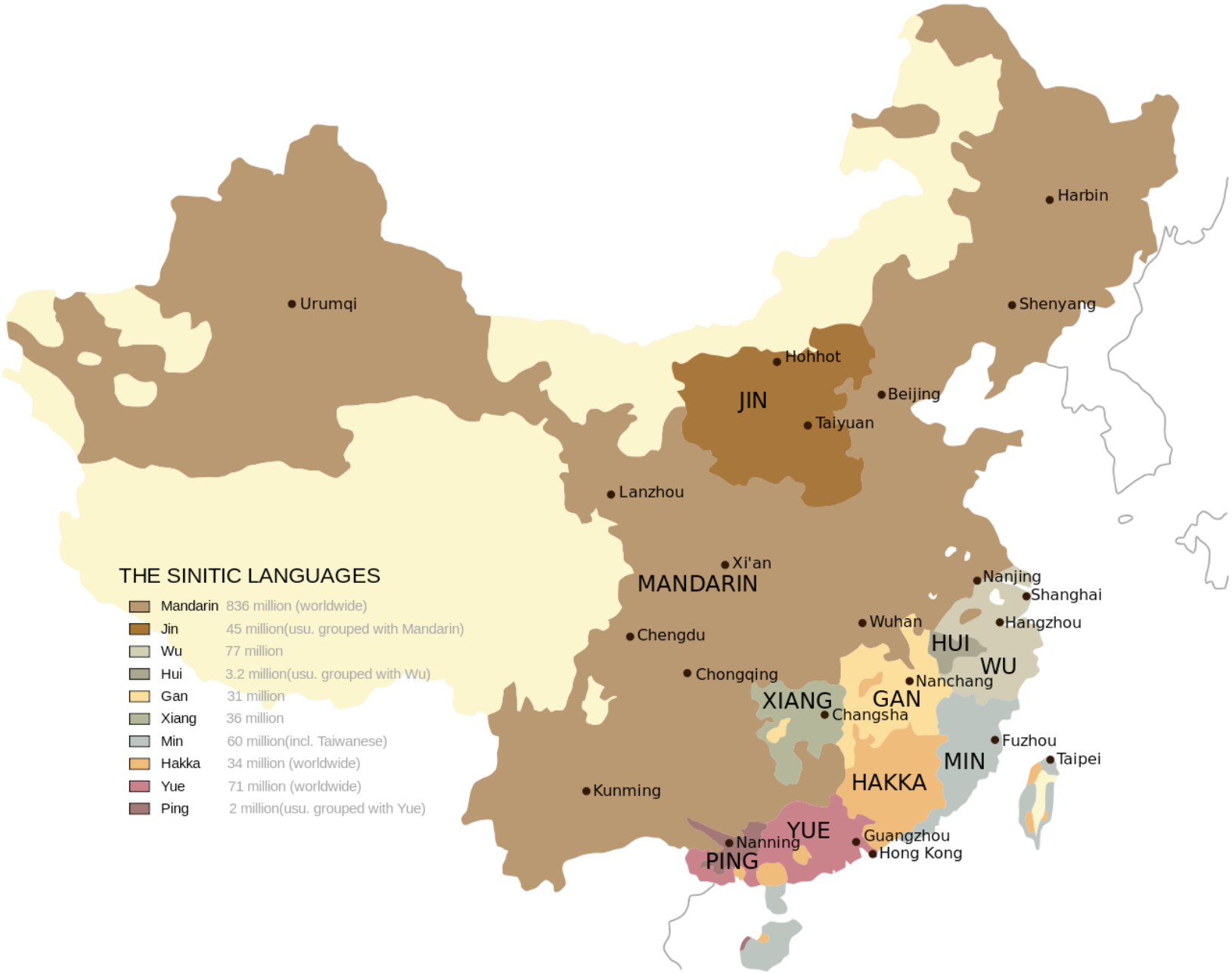
Major Chinese Dialect Groups and Number of Speakers

Dialect Group	Number of Speakers (millions)
Mandarin	900+
Wu (e.g., Shanghainese)	80+
Xiang (Hunanese)	36+
Yue (Cantonese)	85+
Min (e.g., Hokkien, Teochew)	70+
Gan	22+
Hakka	45+

The numbers are approximate and can vary due to factors such as migration and classification.

According to Goddard (p55)

- In the Chinese tradition any variety other than Mandarin is referred to as a "dialect".
- This reflects the cultural and historical unity of China and the use of a common script
  although not all words can be written in Hanzi [FCB]
- In normal usage, dialects are mutually intelligible
- Linguists usually refer to the seven language groupings as "Sinitic languages"
- But some of these "languages" also have many mutually unintelligible varieties
- E.g., Min has as many as nine mutually unintelligible varieties in the Fujian province alone (Li 1992, cited in Goddard 2005).
- There are probably hundreds of mutually unintelligible Sinitic languages in China


THE SINITIC LANGUAGES
Mandarin 836 million (worldwide)
Jin 45 million(usu. grouped with Mandarin)
Wu 77 million
Hui 3.2 million(usu. grouped with Wu)
Gan 31 million
Xiang 36 million
Min 60 million(incl. Taiwanese)
Hakka 34 million (worldwide)
Yue 71 million (worldwide)
Ping 2 million(usu. grouped with Yue)
Harbin
Urumqi
Shenyang
Hohhot
JIN
Beijing
Taiyuan
Lanzhou
Xi'an
MANDARIN
Nanjing
Shanghai
Wuhan
Hangzhou
Chengdu
HUI
WU
Chongqing
Nanchang
XIANG
GAN
Changsha
Fuzhou
MIN
Taipei
HAKKA
Kunming
YUE
Guangzhou
Nanning
Hong Kong
PING

Return the money to me

(1) 把 钱 还 给 我
pa^{324} tɕ^{h}jæn^{25} xwan^{25} kej^{324} wɔ^{324}
take money return give 1SG

Mandarin (Putonghua)

(2) 畀 返 D 钱 我
pei^{35} fæ:n^{55} ti^{55} tɕ^{h}in^{35} ŋɔ^{13}
give back PL.CL money 1SG

Yue (Hong Kong)

(3) 铜钱 还 给 我
doŋ^{22}-dʑi^{55} wæj^{35} paʔ^{5} niʔ^{5}
money return give 1SG

Wu dialect (Suzhou)

(4) 钱 倒 还 与 我
tɕĩ^{35} tɔ^{55} huan^{35} ho^{22} gua^{53}
money invert return give 1SG

Southern Min (Quanzhou)

Mandarin Romanization Comparison

Characters	Wade-Giles	Hanyu Pinyin	Notes
中国/中國	Chung^{1}-kuo^{2}	Zhōngguó	China
北京	Pei^{3}-ching^{1}	Běijīng	PRC Capital
台北	T'ai^{2}-pei^{3}	Táiběi	RoC Capital
毛泽东/毛澤東	Mao^{2} Tse^{2}-tung^{1}	Máo Zédōng	Mao
蒋介石/蔣介石	Chiang^{3} Chieh -shih^{2}	Jiǎng Jièshí	Chiang Kai-shek
孔子	K'ung^{3} Tsu^{3}	Kǒng Zǐ	Confucius

There is also a native transliteration system for Standard Chinese and other Sinitic languages: **Bopomofo**. It is commonly used in Taiwan. It consists of 37 characters and five tone marks, which together can transcribe all possible sounds in Mandarin Chinese: ㄓㄨㄥㄍㄨㄛ , ㄅㄟ ˇ ㄐㄧㄥ, ㄊ ㄞ ㄅㄟ ˇ, ㄇㄠ ㄗㄜ ㄉㄨㄥ, ㄐㄧㄡ ˇ ㄐㄧㄝ 尸 , ㄎㄨㄥ ˇ ㄗ ˇ

Sinospheric features

Sino-Tibetan languages found in SEA are more likely to have features in common with Sinitic languages – e.g.

- complex tone
- classifiers
- limited morphology
- isolating word formation
- pragmatically determined syntax
- reliance on zero anaphora

Indospheric features

Sino-Tibetan languages found in South Asia are more likely to have features in common with Indic languages, e.g.

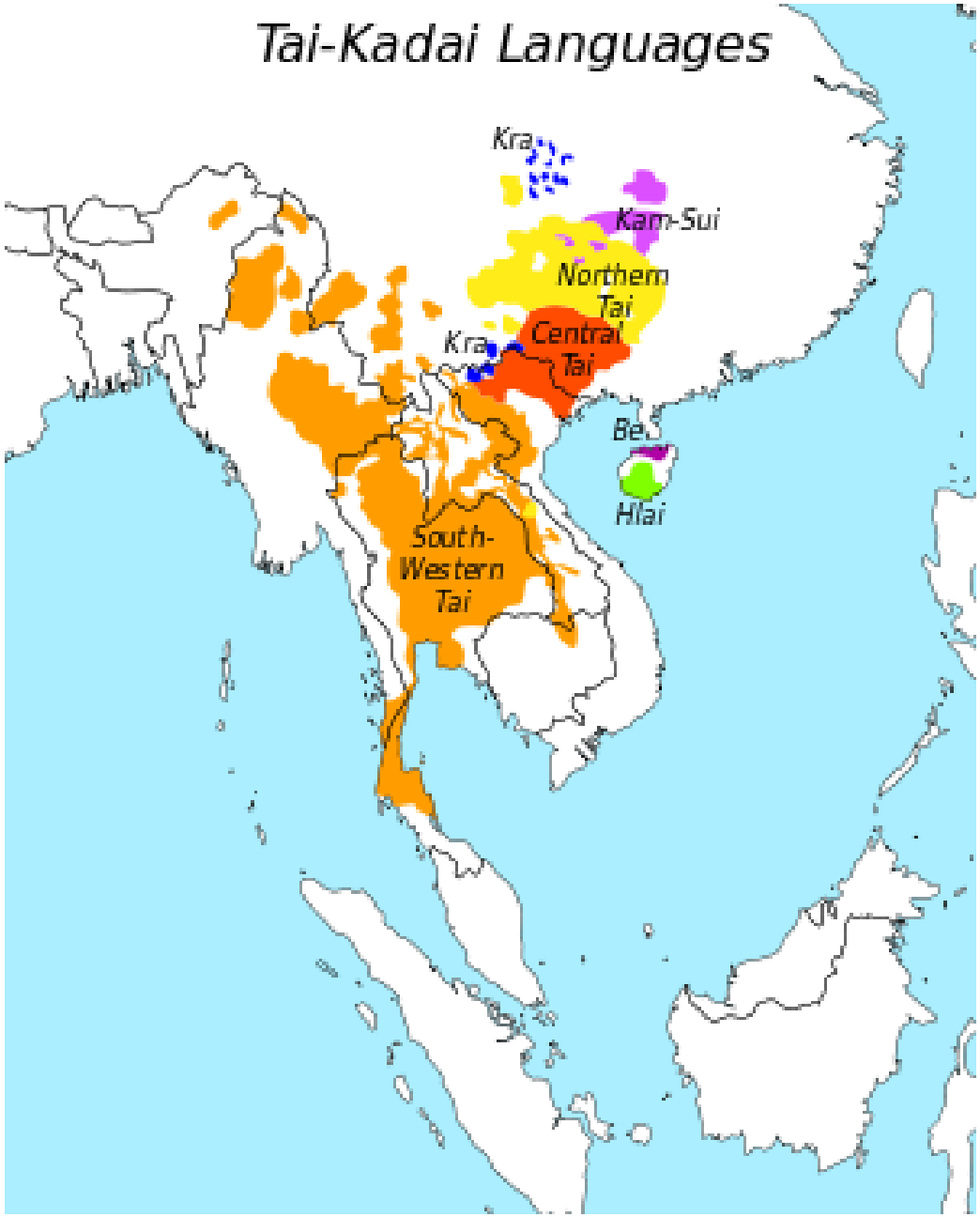
- simple tone systems or no lexical tone
- morphologically complex stems
- fusional or agglutinative word formation
- breathy and retroflex consonants
- pronominal cross-referencing on the verb
- well-developed case-marking paradigms

Tai-Kadai

The Tai-Kadai (onen simply Tai) family splits into three branches:

- Northern
- Central
- South-western

Thai is a member of the South-western branch, members of which can be found as far west as Assam in NE India.
The established spelling convention is that 'Tai' (pronounced with an unaspirated dental stop *t*) is used for the family name, and 'Thai' (pronounced with an aspirated dental stop *th*) is normally used to refer to the national language of Thailand.


Tai-Kadai Languages
Kra
Kam-Sui
Northern Tai
Central Tai
Kra
Be
Hlai
South-Western Tai

Historical origin of Tai-Kadai languages

Tai-Kadai languages are descended from a parent language estimated to have been spoken by a single group of people approximately fineen hundred to two thousand years ago.
Most linguists think that the Tai homeland was somewhere near the present-day Vietnam/China border. Tai speakers are believed to have migrated from this region to northern Vietnam

- Laos
- Thailand
- southern provinces of China
- northern Burma
- Assam in north-eastern India

Major Tai-Kadai Languages

- Thai (60 million speakers)
  could be several languages
- Lao (4 million)

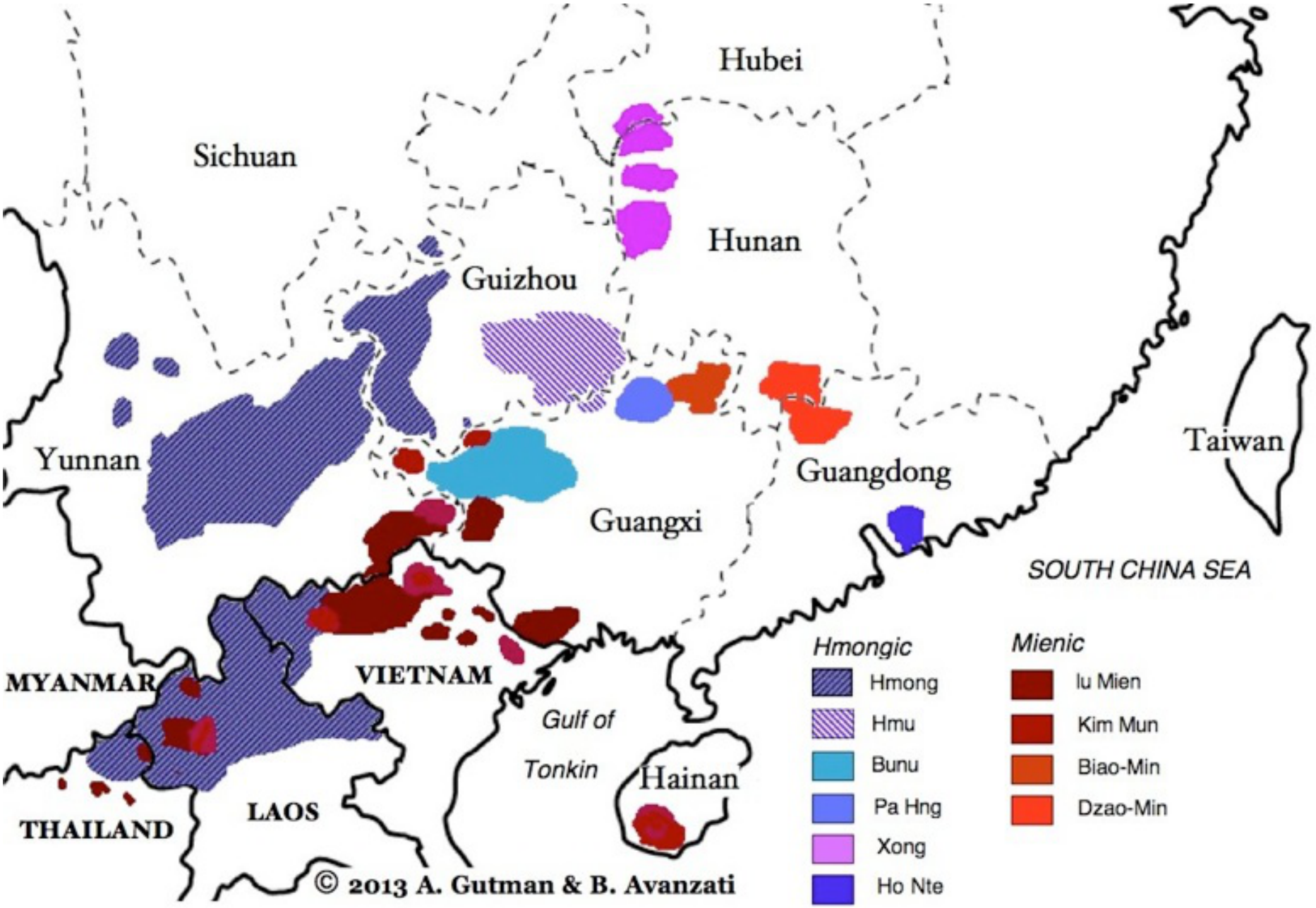
Over 50 other Tai-Kadai languages.

Hmong-Mien

Thirty-five languages spoken mainly in southwestern China, with several spoken in adjacent parts of Southeast Asia.

- Hmong (also known as Miao: 5 million speakers)
- Mien (also known as Man and as Yao: 2 million)

Most work has been done on this family by Chinese scholars, and it is relatively little known in the West. It is mainly spoken in the hills.


Hubei
Sichuan
Hunan
Guizhou
Yunnan
Guangdong
Taiwan
Guangxi
SOUTH CHINA SEA
MYANMAR
VIETNAM
Gulf of
Tonkin
Hainan
LAOS
THAILAND
© 2013 A. Gutman & B. Avanzati
Hmongic
Hmong
Hmu
Bunu
Pa Hng
Xong
Ho Nte
Mienic
Iu Mien
Kim Mun
Biao-Min
Dzao-Min

Japanese, Korean, and Ainu

- Japanese (125 million speakers)
- Korean (70 million)
- Ainu (isolate, fewer than 300 speakers)

Japanese and Korean share very similar syntax (agglutinative, case marking, verb final, pre-modification) and both have borrowed much vocabulary from Chinese.

Mainland Southeast Asia as a Linguistic Area

Mainland Southeast Asia as a Linguistic Area

- Languages of different families often converge when in close contact.
- A **linguistic area** (*Sprachbund*) is where unrelated languages share features.
- Mainland Southeast Asia is a linguistic area due to long mutual influence.
- Shared features span phonology, lexicon, and grammar.
- Example: Limited syllable structures and presence of lexical tone.
- Many features shared with non-mainland languages

Shared Phonological Features

- Many languages have a limited range of syllable structures.
- Lexical tone is common across the region.
- Example: Tones in Vietnamese, Thai, and many Chinese dialects.
- Mon-Khmer languages often lack tones, with some exceptions.
- Phonological convergence results from prolonged contact.

Phonological Features: Examples I

- **Limited syllable structures**:
  - **Mandarin Chinese** primarily uses simple syllables like CV (Consonant-Vowel) and CVN (Consonant-Vowel-Nasal), e.g.,
    - *mā* (mother)
    - *màn* (slow)
  - Complex consonant clusters are rare in these languages.
- **Lexical tone examples**:
  - **Mandarin Chinese** has four tones:
    - *mā* (mother) –high-level tone
    - *má* (hemp) –rising tone
    - *mǎ* (horse) –dipping tone
    - *mà* (scold) –falling tone

Phonological Features: Examples II

- **Thai** has five tones:
  - *maa* (come) –mid tone
  - *máa* (horse) –rising tone
  - *màa* (dog) –low tone
  - *mâa* (mother) –falling tone
  - *mǎa* (new) –high tone
- **Vietnamese** has six tones, e.g.,
  - *ma* (ghost) –level tone
  - *má* (cheek) –high rising tone
  - *mà* (but) –low falling tone
  - *mả* (tomb) –falling-rising tone
  - *mã* (horse) –glottalized rising tone
  - *mạ* (young rice plant) –glottalized falling tone

Shared Morphological Features

- Lack of inflectional morphology is widespread.
- Use of classifier constructions is a common feature.
- Example: Classifiers used when counting nouns in Thai and Mandarin.
- Morphological similarities facilitate language learning across groups.
- Reflects a mutual adaptation over time.

Classifiers in Thai and Mandarin

- **Mandarin Chinese** classifiers:
  - 一只猫 (yī zhī māo) –"one (zhī) cat"
  - 三本书 (sān běn shū) –"three (běn) books"
  - 五辆车 (wǔ liàng chē) –"five (liàng) cars"
- **Thai** classifiers:
  - เด็กหนึ่งคน (dèk nùeng khon) –"one (khon) child"
  - บ้านสองหลัง (bâan săawng lăng) –"two (lăng) houses"
  - เสื้อผ้าสามตัว (sûea phâa săam tua) –"three (tua) shirts"
- Classifiers are essential when quantifying nouns.

Areal Lexicon

- Beyond shared words, languages share conceptual frameworks.
- Term **areal lexicon** coined by James Matisoff.
- Shared worldview and consensus on topics worth discussing.
- Example: Similar cultural vocabulary across unrelated languages.
- Reflects deep cultural and linguistic exchange.

Areal Lexicon: Shared Cultural Vocabulary I

- **Shared terms for "rice" at different stages**:
  - **Thai**:
    - *khao* (ข้าว) –rice in general
    - *khao niao* (ข้าวเหนียว) –sticky rice
  - **Vietnamese**:
    - *lúa* –rice plant
    - *gạo* –uncooked rice grains
    - *cơm* –cooked rice
  - **Japanese**:
    - 稲 *ine* –rice plant
    - 米 *kome* –uncooked rice grains
    - 御飯 *gohan* –cooked rice
  - **Indonesian**:
    - *padi* –rice plant (still growing in the field)
    - *gabah* –harvested but unhusked rice
    - *beras* –husked, uncooked rice grains
    - *nasi* –cooked rice

Areal Lexicon: Shared Cultural Vocabulary II

- **Kinship terms emphasizing relative age**:
  - **Mandarin Chinese**:
    - 哥哥 (gēge) –older brother
    - 弟弟 (dìdi) –younger brother
  - **Thai**:
    - พี่ (phîi) –older sibling
    - น้อง (nóng) –younger sibling
- Reflects shared cultural concepts due to prolonged contact.

Shared Syntactic Features

- Prevalence of serial verb constructions.
- Topic prominence influences sentence structure.
- Example: Word order changes based on the topic of interest.
- Ellipsis is common, relying on context for interpretation.
- Sentence-final particles express speaker's attitude.

Topic Prominence: Word Order Examples

- **Mandarin Chinese**:

  (5) Topic-comment structure:

  这 本 书 ，我 看 过 了。
  zhè běn shū , wǒ kàn guò le .

  "This book, I have read."

- **Japanese**:

  (6) Topic marker は (wa):

  象 は 鼻 が 長い 。
  zou wa hana ga nagai .

  "Elephants have long noses."

  (As for elephant, nose-NOM long)

Topic Prominence and Ellipsis

- Sentence structure is guided by discourse considerations.
- Less rigid grammatical rules for word and phrase order.
- Ellipsis involves omitting understood participants.
- Example: Dropping the subject when it's contextually clear.
- Emphasizes the importance of context in communication.

Ellipsis: Omitting Understood Participants

(7) **Japanese**:

ϕ 行きます 。
ϕ iki masu .

"(I) am going."

(8) **Mandarin Chinese**:

ϕ 吃 了吗?
ϕ chī le ma?

"Have (you) eaten?"

(9) **Thai**:

ϕ ไปไ หน?
ϕ pai năi?

"(You) go where?"

Subjects and objects are often omitted when contextually clear.

Sentence-Final Particles

- Small expressive words at the end of sentences
- Indicate speaker's feelings or attitudes
- Also used to mark questions
- Example: Thai particle นะ (na) to soften statements
- Common in many mainland Southeast Asian languages
- Enhance nuance and expressiveness in speech

Sentence-Final Particles: Examples

- **Thai**:
  - คุณสบายดีไหมคะ?
    (Khun sabai di mai kha?)
    "How are you?" (*kha* adds politeness from a female speaker)
  - ไปนะ
    (Pai na)
    "I'm leaving now, okay?" (*na* softens the statement)
- **Mandarin Chinese**:

(10) 我们 走 吧。
wǒmen zǒu ba .
we go SUGG .
'Let's go.' (*ba* indicates a suggestion)

(11) 好 吃 吗 ?
hǎo chī ma ?
good eat Q ?
'Is it delicious?' (*ma* turns a statement into a question)

Summary of Areal Features

- Languages share features like lexical tone and classifiers.
- Variation exists within language families.
- Mon-Khmer: Some languages lack tones.
- Tibeto-Burman: Generally verb-final and use postpositions.
- Sinitic languages have both prepositions and postpositions.

Geographical Influence on Language Distribution

- Rivers and mountains shape language geography.
- Major rivers: Irrawaddy, Chao Phraya, Mekong, Red River.
- Fertile deltas support large populations and language spread.
- Mountain ranges create linguistic boundaries.
- Southern China included in the linguistic area.

Uplands vs. Lowlands Sociolinguistics

- Primary dichotomy between upland and lowland regions.
- Lowlands: Wet rice cultivation supports large populations (e.g., Thai, Vietnamese).
- Uplands: Shifting agriculture supports smaller groups (e.g., Hmong).
- Elevation influences agricultural methods and settlement patterns.
- Example: Hmong villages at elevations of 1,000–1,500 meters.

Historical Language Dispersal

- Mon-Khmer speakers inhabited inland areas 4,000 years ago.
- Tai-speaking peoples migrated from southern China 2,000 years ago.
- Tai expansion followed river valleys into lowlands.
- Han Chinese expanded southward, influencing language shifts.
- Austronesian languages present in southern Vietnam (Cham empire).

Rise and Fall of Empires

- 1,000 years ago: Mon-Khmer empires dominated (e.g., Khmer empire of Angkor).
- Tai kingdoms grew in Laos and Thailand, influenced by Khmer culture.
- Language shift occurred as Mon-Khmer groups adopted Tai languages.
- Tibeto-Burman peoples moved into Myanmar, ending Mon dominance.
- Mon language now survives in small pockets in eastern Burma.

Impact of Warfare and Migration

- Wars between Thai, Burmese, and Vietnamese reshaped demographics.
- Large armies and elephant corps used in conflicts.
- Kingdom of Lan Xang ("million elephants") signified military power.
- Migration led to shifts in ethnic and linguistic composition.
- Cultural exchange occurred through conquest and alliances.

Hmong-Mien Peoples

- Historically located in southwestern China until 150 years ago
- Lived under Han Chinese domination
- Migration southward due to pressure and conflicts
- Historically lived in high mountainous areas (1,000-1,500 meters)
- Minority status led to cultural and linguistic preservation
- Distinctive features reflect adaptation to mountain life

Colonial Influence in the 19th Century

- Western powers colonized Southeast Asia.
- British controlled Myanmar and Malaya.
- French ruled over Indochina (Vietnam, Laos, Cambodia).
- Thailand remained independent but lost some territories.
- Colonial borders influenced modern nation-states.

Conclusion: Linguistic Convergence

- Mainland Southeast Asia exemplifies a linguistic area.
- Languages share features due to long-term contact and mutual influence.
- Geography and history played key roles in language distribution.
- Understanding shared features aids in studying regional linguistics.
- The region showcases language evolution through cultural interaction.

A comparison of features!

Morpho-Syntactic Features I

Tone: Indicates whether the language has a tonal system (+ for tonal, – for non-tonal, +/– for mixed).

Order: Refers to the predominant word order (SVO, SOV, etc.).

Cl: Whether the language uses classifiers (+ for yes, – for no).

SFP: Presence of sentence-final particles (+ for yes, – for no).

SV: Indicates whether the language uses serial verbs (+ for yes, – for no).

Zero: Indicates if the language has zero anaphora or ellipsis (+ for yes, +/– for partial).

Inflect: Whether the language has inflectional morphology (+ for yes, – for no).

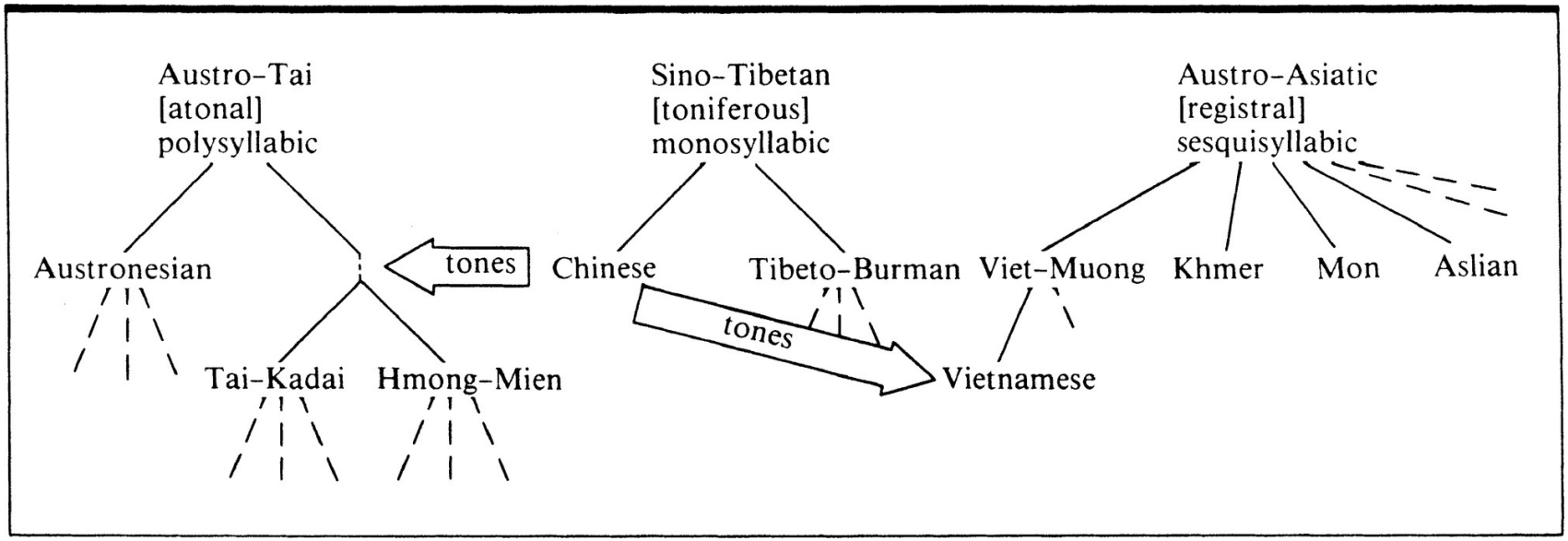
Morpho-Syntactic Features II

Language	Tone	Order	Cl	SFP	SV	Zero	Inflect
Austronesian	−	VSO/SVO	+	+/-	+	−	−
Austroasiatic	+/-	SVO	+	−	+	+/-	−
Tibeto-Burman	+	SOV	+/-	−	+	+/-	−
Sinitic	+	SVO	+	+	+	+/-	−
Tai-Kadai	+	SVO	+	+	+	+/-	−
Hmong-Mien	+	SVO	+	+	+	+/-	−
Korean/Japanese	−	{SO}V	+	+	+	+	+

We will talk about these in more detail later!

The spread of tone in MSEA languages

Vietnamese developed its 6 tones under the influence of Chinese, as demonstrated by the French linguist Andre Haudricourt. Some Tibeto-Burman languages are currently in the process of developing tonal contrasts, particularly those of the Bodic branch. Others are losing them, e.g. the Bodo languages of Assam.



Features (and vocabulary) can move across language families!

Summary I

There are five major linguistic families found in Mainland SEA.
While there are a great number of languages spoken in the region, the linguistic diversity is relatively low, probably due to centuries of language contact and human population movement.

Summary II

MSEA languages are characterized by:

- isolating word formation
- tone
- absence of inflectional morphology
- classifier systems
- aspectual and mood systems rather than tense systems
- sentence final particles
- rampant zero anaphora
- serial verb constructions
- re-duplication

Although there are many exceptions