The “Greater Austric” Hypothesis

John D. Bengtson

Association for the Study of Language in Prehistory

“Austric” is a hypothetical language family proposed by Wilhelm Schmidt, a Lutheran clergyman from Germany who worked as a missionary in Southeast Asia. He hypothesized (1906) that the Austroasiatic language family and the Austronesian language family (see below) came from a common origin, and should be grouped together into a new phylum named Austric (from Latin auster ‘south’).

[Schmidt] presented phonological, morphological, and lexical evidence in support of this hypothesis. The consensus of the linguists who have looked into the matter has been that the phonological and morphological evidence is convincing, but not the lexical, and for that reason, Austric has never been generally accepted as a valid taxonomic unit (La Vaughn Hayes).

In the ensuing century various scholars have sought to “expand” the Austric family, to include other languages that seem to have the same grammatical and lexical features as Schmidt’s Austric. Beginning around 1996 I proposed a “Greater Austric” macro-phylum of macro-family that includes the following languages:

- **Nihali**: a tribal language of central India, thought to be the last remnant of an ancient language family.
- **Austroasiatic (Munda - Mon-Khmer)**: about 155 languages spoken in India, Nicobar Islands, and Southeast Asia: includes Santali, Mundari, Sora (India), Nancowry, Mon, Khmer (Cambodia), Vietnamese, Pear, Katu, etc.
- **Hmong-Mien (Miao-Yao)**: a small number of languages (4-7?) scattered through southern China, Hainan, Vietnam, Laos, and Thailand: includes Hmong, Ke-cheng, Kao-p’o, Kao-t’ung, Yao-lu.
- **Daic (Tai-Kadai)**: about 55 languages spoken in Southeast Asia: includes Thai, Lao, Kam, Sai, Lakkia, etc.
- **Austronesian**: a vast family of nearly 1000 languages ranging from Madagascar through Indonesia, Philippines, Taiwan (aboriginals, not Chinese), to the Pacific islands: includes Malagasy, Malay (= Bahasa Indonesia), Javanese, Paiwan, Bontok, Tagalog, Fijian, Maori, Tahitian, Samoan, Hawaiian, etc.
- **Ainu**: an almost extinct minority language of Japan. Formerly more extensive, Ainu was also spoken on the Kuril Islands, Sakhalin Island, and the Kamchatka peninsula. It is thought to have been the language of the ancient Jomon culture (about 10,000 BC to 300 BC).

**Grammatical evidence for Greater Austric**

For many linguists the existence of a common grammatical system is the best evidence of original kinship. One of the most convincing pieces of grammatical evidence for Austric is the basic first-person singular pronoun, widely of the type that can be reconstructed as *ku ‘I, me’.

---

1 See http://home.att.net/~lvhayes/Langling/langpg3.htm.
2 Asterisk * is used by historical linguists to denote reconstructed words or morphemes (grammatical elements) that are not directly attested but can be inferred from descendant forms.
• Nihali jò ‘I’
• Austroasiatic: Proto-Viet-Muong *kwa ‘we, I’, Central Nicobar chuu-ö ‘I’
• Hmong-Mien: Biao kou, Hmong ko ‘I’
• Austronesian: *aku ‘I’ > Pazeh yako, Javanese aku, Tagalog akó, Tahitian au, vau, etc.
• Daic: Proto-Tai *ku ‘I’
• Ainu *ku ‘I’

Other grammatical evidence includes prefixes or infixes of the type *pa-/ *-ap- “causative,” both in Austroasiatic [AA] and Austronesian [AN]:

• AA: Katu (Vietnam, Laos) sooq ‘to flee’: pa-sooq ‘to cause to flee’; Sora (India) jUm ‘to eat’: ab-jUm ‘to cause to eat = to feed’
• AN: Bontok (Philippines) kán ‘to eat’: pa-kán ‘to cause to eat = to feed’, téd ‘die’: pa-téd ‘to cause to die = to kill’

and an “agentive marker”: 3 *ma- / *-am- in Austroasiatic = *mu- / *-um- in Austronesian:

• AA: Nancowry (Nicobar) itkéch ‘to pluck’: m-itkéch ‘one who plucks, plucker’; Pear (Cambodia) snik ‘light’: am-snik ‘to lighten’; Khmer (Cambodia) rut ‘to run away’: r-m-ut ‘one who runs away, fugitive’
• AN: Bontok ‘ákew ‘to steal’: -um-ákew ‘one who steals = thief’; inum ‘to drink’: -um-inum ‘one who drinks = drinker’

Unfortunately for the comparative linguist, some Austric languages (e.g., Vietnamese, Hmong, Thai) have become monosyllabic and the old prefixes have dropped off. However there are still some clues about their earlier presence. For example, the Hmong-Mien languages have the contrast of initial consonants in *day ‘die’ vs. *tay ‘kill’, a contrast that seems to go back to Proto-Austric stative *ma-t(r)ay ‘to die’ vs. causative *pa-t(r)ay ‘to make die, to kill’ (see the etymology die, below).

**Lexical evidence for Greater Austric**

The most convincing word comparisons are those that involve the most basic (non-cultural) meanings, such as parts of the body (blood, head), basic natural phenomena (root, fire, sky), simple social terms (man), and basic verbs (die). Studies have shown that basic words of these types are far more likely to remain in a given language over hundreds and even thousands of years than words with non-basic meanings (such as ‘chair, table, tea, card’, items that are likely to be traded between cultures along with their names). 4 Only a few of the many Austric word comparisons are cited below. Some phonetic transcriptions have been simplified.

**man:** prototype *qulo

• Nihali Kol-ta, Kal-to ‘men, Nihals’ (self-name), kol ‘wife, woman’
• Austroasiatic: Korku koro ‘man’; Khmer kur ‘Bahnar or Sr. ’

---

3 Agentive marker: like -er, -or in English bak-er, sail-or, etc.
4 While most basic words in English are still of Anglo-Saxon (Germanic) origin, the four words cited in the parentheses are of non-Anglo-Saxon origin: chair < French < Latin < Greek; table < French < Latin; tea < Dutch < Malay < Chinese, card < French < Latin < Greek < Egyptian (originally ‘a sheet of papyrus’).
• Austronesian: Proto-Malayo-Polynesian *qulun ‘outsiders, alien people’ > Bintulu, Katingan ulun ‘person’, Merina olona, Punan Kelai lun, etc.
• Ainu *kur ‘man, person’

blood: prototype *(m)-kyem
• Austroasiatic:Mundari mayam, Khmer jha:m, Mon chim
• Hmong-Mien: Yao dzhyaam, Hmong ntsheng
• Austronesian: Formosan5 *dzamu(‘) ‘blood’ > Paiwan djamu, etc.
• Ainu kem

head: prototype *Pengu
• Nihali peng, pyeng ‘head’
• Austroasiatic: Khmer tpu:ng ‘above’
• Austronesian: *bungh ‘head’ > Formosan: Kanakanabu nabung, Siraya bung, Central Amis fungoh, Tsou fngiu ‘head’, etc.
• Ainu *pa ‘head’ (if from earlier *pang)

root: prototype *riat(s)
• Austroasiatic: Santali rehe’d, Mundari re:’d ‘root’, Khmer ris; Mon ruih; Sre rias, Vietnamese re ‘root’
• Austronesian *uKat ‘blood vessel, vein, sinew, tendon’ > Formosan: Kanakanabu uratsə ‘blood vessel’, etc.; Iban, Maloh urat ‘root’, Dayak Ngaju umat, Tagalog ugát, etc.
• Ainu *rit ‘root, blood vessel, tendon’

fire: prototype *-apoy
• Nihali a:po ‘fire’
• Austroasiatic: Brao pacy, Tampuon pae, Katu mpoih ‘fire’
• Hmong-Mien: Ke-cheng fvi ‘ash’, Yao-lu fui, Thailand Yao whi ‘ash’
• Austronesian *Sapuy > Paiwan sapuy, Pazeh hapuy, Malay api, Tonga afu, Maori, Hawaiian ahi, etc. ‘fire’
• Daic: Thai fay, Kam-Sui *pwai ‘fire’
• Ainu *apOy ‘fire, hearth’ > ape, abe, ambe, aboi (in various dialects)

sky: prototype *langit(s)
• Nihali lēgē ‘up’
• Austroasiatic: Sora leng.leng ‘very high, inaccessible’, bo leng ‘roof’, Juang aling ‘top’; Khmer ling ‘to climb, ascend’, Bahnaric le:ng ‘sky’
• Austronesian: *langiT ‘sky’ > Saaroa langica, Puyuma ranget, Merina lānitra, Tagalog langit, Malay langit, Maori rangi, Hawaiian lani ‘day, heaven, sky’, etc.
• Ainu *nis ‘sky, cloud’ > nish ‘clouds, heavens, air, sky’ (if from *langiT[s])

die: prototype *(ma)-t(r)ay / *(pa)-t(r)ay
• Nihali pāDa: ‘to kill’ (if pa- corresponds to the Proto-Austric causative prefix *pa-)
• Hmong-Mien: *day ‘die’ vs. *tay ‘kill’ > Hmong dua ‘to die’

5 “Formosan” denotes aboriginal languages of Taiwan a.k.a. Formosa (Portuguese for ‘beautiful’).
6 *C denotes “hypothetical consonant of unknown quality.”
Historical Implications of the Greater Austric Hypothesis

If the Greater Austric hypothesis is correct, people speaking a “Proto-Austric” language spread out over a large part of Asia, an area probably stretching from India in the west, continuing through Southeast Asia (the old “Indo-China”) and southern China, and to the northeast possibly as far as Korea and Manchuria. There is much evidence that Austric-speaking peoples were formerly more numerous and widespread in all of these areas, and that the subsequent expansions of other language families (mainly Indo-European, Dravidian, and Sino-Tibetan) have pushed Austric languages to the south. Austric languages have left traces (loanwords) in non-Austric languages ranging from Pakistan in the northwest (Burushaski, Indo-Aryan and Dravidian languages) through the Himalayan regions (Tibeto-Burman languages), in Tibetan and Chinese, and in the northeast, in languages such as Korean and Japanese.

The Austric proto-language has been identified by some with the Hoabinhian archaeological industry dating from the late Pleistocene to mid-Holocene (roughly 6,000 to 12,000 years ago). Primary Hoabinhian sites have been identified in Sumatra, Thailand, Laos, Myanmar and Cambodia, while isolated inventories of stone artefacts displaying Hoabinhian elements have been found in Nepal, South China, Taiwan and Australia. Except for Nepal and Australia all of these areas are home to Austric languages, and, as mentioned in the preceding paragraph, there is evidence that Austric may formerly have been spoken in the Himalayan foothills.

The time-depth of the Hoabinhian industry is consistent with the diversification of Proto-Austric into about 1000 languages, and the lexical diversity among them. In this light Austric seems comparable in age to the Afro-Asiatic language family (consisting of the Semitic, Egyptian, Berber, Chadic, Cushitic, Omotic families). There is tremendous lexical diversity among the Afro-Asiatic languages, yet some grammatical features are widely found among them.

Secondary to the original dispersal of Austric, one Austric sub-family, Austronesian, has made a tremendous expansion to the south, southwest, and east. Austronesian speakers are thought to have migrated from southern China about 8,000 years ago, first to Taiwan (where aborigines speaking Austronesian languages still live, though the languages are now endangered or extinct), then (about 6,000 years ago) to other islands ranging from Madagascar through Indonesia, Philippines, and throughout the Pacific islands as far as Hawai‘i and Easter Island. During these 8,000 years Austronesian has diversified into a huge number of languages (959 according to Merritt Ruhlen, 1268 according to Ethnologue), making up about one fifth of the world’s languages. Of these a few are official (national) languages, especially Malay (Indonesia and Malaysia), Tagalog (Philippines), Malagasy (Madagascar), and several Pacific languages (Samoa, Tahitian, Maori, etc.).

Of the remaining Austric languages only a few are official or national languages, namely Vietnamese, Khmer (Cambodia), Thai, and Lao. Apart from these four the Mon language (Myanmar, Thailand) has a long literary tradition dating back to the 7th Century A.D. The vast

---

1 Historical linguists use the term “loanword” to denote words that are not part of a language’s native lexicon, but were “borrowed” from other languages. See the discussion of English chair, table, tea, card, above, and footnote 4.
2 Thanks to Wikipedia article on Hoabinhian.
3 No Austric languages are native to Australia, but Austronesian languages are spoken close to Australia, in parts of New Guinea and Indonesian islands.
majority of Austric languages are minority or “tribal” languages that have only recently acquired a written form (if at all). Nevertheless, the Austric language macro-family is of tremendous importance in understanding the history and prehistory of Asia and Oceania.

References

Cooper, Doug. *Thai Lexicography Resources.*
[http://seasrc.th.net/index.html?main=http%3A//seasrc.th.net/proto/]
Evolution of Human Languages. (An international project on the linguistic prehistory of humanity coordinated by the Santa Fe Institute.) [http://ehl.santafe.edu/]
Mother Tongue (Journal) 2: 67-74.
-------- 1996. Comments on entries in “Lexical Parallels between Ainu and Austric”. Ms.


