First International Conference of Aceh and Indian Ocean Studies

Organized by
Asia Research Institute, National University of Singapore & Rehabilitation and Construction Executing Agency for Aceh and Nias (BRR), Banda Aceh, Indonesia

24 – 27 February 2007

The Historical Place of Acehnese: The Known and the Unknown

Graham Thurgood

California State University, Chico, USA

gthurgood@csuchico.edu

Not to be quoted without permission from the author
Abstract

The linguistic evidence has established Acehnese as one of the Chamic languages, a group including the mainland Chamic languages Phan Rang Cham, Haroi, Jarai, Rade, Chru and Roglai found in central Vietnam, Hainan Cham found near Sanya on southern Hainan Island, and Western Cham found in parts of Cambodia and Thailand. Its closest non-Chamic relatives are the Malayic languages Malay and Mingangkabau (Awak Jamee or Anuek Jamee) and the Malay dialect Tamiang (Basa Teumieng); more distantly related are the Batak languages, Karo Batak, Alas, and Kluet. Many of the frequently noticed widespread similarities in vocabulary with Malay are now recognized as the result of long-term contact.

The linguistic record also shows that Acehnese speakers migrated from central Vietnam coast to the tip of Sumatra, in anticipation of the Vietnamese conquest of the southern capital of Champa in 1471 --- an analysis that matches our non-linguistic historical material closely. In addition, it is clear that some time was spent in Kelantan, although the dating of this is unclear.

And, while it is clear that the spread of Acehnese involved absorption and replacement of other languages, major work on Acehnese dialects and their interaction with other languages in the immediate vicinity is needed to clarify our understanding of this development.

1.0 The genetic\(^{[1]}\) relationship between Acehnese and other Austronesian languages

Despite the geographical distance between Acehnese and its closest relatives, the linguistic evidence clearly and unambiguously establishes that Acehnese is one of the Chamic languages, a subgroup that includes the mainland Chamic languages Phan Rang Cham (= Eastern Cham), Haroi, Jarai, Rade, Chru and Roglai found in central Vietnam, Hainan Cham (= Tsat) found near Sanya on the southern part of Hainan Island, and Western Cham found in parts of Cambodia and Thailand. It is equally obvious from the linguistic record that Acehnese is also genetically related but more distantly to the other Austronesian languages of Sumatra (see Figure 1).

Acehnese’s closest non-Chamic relatives are the Malayic languages, a group including Malay, Mingangkabau (Awak Jamee or Anuek Jamee), and the Malay dialect Tamiang (Basa Teumieng (Adelaar (1988, 1992)). The Malayic languages and the Chamic languages including Acehnese form the Malayo-Chamic subgroup (Blust 1992:33).

Finally, the Batak languages, Karo Batak, Alas, and Kluet are more distantly related to Acehnese (again, see Figure 1). Also related is Gayo, a language treated only as Western Malayo-Polynesian, with the finer details of subgrouping apparently not yet worked out. Again, see Figure 1.

\(^{[1]}\) Genetic affiliations are done entirely on the basis of shared innovations in patterns of regular sound-meaning correspondences (correspondences, for short). Neither typological similarity nor geographical distribution constitute legitimate bases for determining genetic affiliation, nor should they as the speakers of languages move from place to place and since languages change their typological structures in contact with other languages.
Figure 1: The Acehnese and other regional languages

Acehnese

Phan Rang Cham
Western Cham
Chru
Haroi
Jarai
Hainan Cham (=Tsat)
Roglai

Malay
Tamiang
Minangkabau

Karo Batak
Alas
Kluet
Gayo (?)

Chamic

Malayo-Chamic

Western
Malayo-
Polynesian

Malayic

Batak

Notes on Figure 1:

1. Ross (1995:263) distinguishes language subgroupings established on the bases of shared innovations from collections of languages placed together for other reasons. Ross places the collections in italics, while presenting the innovation-based subgroupings in bold-faced, non-italicized type. It is the innovation-based subgroupings, which are done on the bases of shared innovations, that show a period of common development and thus indicate a common history.

2. Western Malayo-Polynesian contains countless other languages totally ignored in Figure 1.

3. It is worth bearing in mind that the historical subgrouping of languages, when properly done, as these subgroupings are, is done not on the basis of the geographical distribution of the languages, but on the basis of shared historical innovations (Blust 1995, Ross 1992). The Austronesian family tree above is based on such shared historical innovations; the fact that the family tree has such striking correlations with geography is because, to a large degree, the current linguistic distribution still reflects the older migration patterns fairly accurately.

1.1 Older literature

The existence of a connection between Acehnese and the mainland Chamic languages has been noticed for a long time. As early as 1822, John Crawfurd recognized that Cham was Austronesian when he described Cham as the “Malay of Champa”. In 1891, G. K. Niemann recognized the genetic connection between Acehnese and the mainland Chamic languages. However, around the turn of the twentieth century, misled by typological similarities caused by contact with Mon-Khmer languages and by material borrowed from Mon-Khmer, several scholars misanalyzed the relationship: Schmidt (1906) described the Chamic languages as “Austroasiatische Mischsprache” [A Austroasiatic mixed language]; Blagden (1929) recognized the unique similarities between Acehnese and the other
Chamic but suggested that these were due to similar influences on both languages from Austroasitic sources;\(^2\)\(^2\) and Sebeok (1942) claimed that these languages were Austroasiatic, a totally incorrect analysis.

A number of researchers were not misled by the Mon-Khmer influences, recognizing the mainland Chamic languages as closely connected to Acehnese: Cowan (1974, 1981, 1988, 1991), Marrison (1975), Shorto (1975), I. V. Collins (1975)\(^3\)\(^3\), Blust (1981), Adelaar (1985), Durie (1985:3), and Sidwell (p.c.) have all argued for a special connection between the Chamic and the Malayic languages.\(^4\)\(^4\)

The use of lexicostatistical evidence led several scholars to argue that Acehnese was closer genetically to Malay than to the mainland Chamic languages: the failure to recognize that a large number of Malay words in Acehnese were loans resulted in the impression that Malay and Acehnese are much more closely related than they actually are. For example, Dyen (1965), relying on lexicostatistical accounts, mistakenly places Acehnese closer to Malay, Madurese, and Lumpung than to the mainland Chamic language Chru (Cru). Similarly, I. V. Collins (1969), also using lexicostatistical methods, places Acehnese closer to Malay than to the mainland Chamic language Northern Roglai.

While the conclusions reached using lexicostatistics are incorrect, the numbers that Dyen and I.V. Collins\(^5\)\(^5\) came up with still require some sort of explanation: the numbers reflect some sort of special relationship between Acehnese and Malay. The explanation is, as Shorto (1975) and Durie (1996) both suggest, the lexicostatistical results reflect a long-term, close contact relationship with Malay. Shorto (1975) is quite specific in suggesting that the misanalysis of the genetic relationship between Malay and Acehnese was the result of “the very high incidence of Malay loans” in Acehnese.

\(^2\)\(^2\) The regular correspondences between the two groups in Thurgood 1999 eliminate the possibility that the similarities are due to mutual influence from some third Austroasiatic source.

\(^3\)\(^3\) I. V. Collins (1969) argued that Acehnese was closer to Malay than Chamic on the basis of lexicostatistical evidence, but by the time of his 1975 dissertation had changed his mind. It is worth noting that lexicostatistical evidence that Acehnese is closer to Malay than to Chamic may reflect, not a genetic closeness, but rather the effects of Malay contact on Acehnese. However, note that I have not personally been able to consult this work yet.

\(^4\)\(^4\) J. Collins has two qualifications, both of which need be dealt with. First, Collins adds a word of caution (1991:110), noting that without full reconstructions of Proto-Chamic, Proto-Malayic, and Proto-Acehnese, “especially, with respect to their phonological history,” “comparisons are superficial”. However, now with a detailed reconstruction of Proto-Chamic (Thurgood 1999, building on and expanding Lee 1966 to include Acehnese), with Adelaar’s proto-Malayic (1992), and Durie’s partial reconstruction of proto-Acehnese (1990), these reconstructions have been done, and, now, on the bases of these, it is evident that the uniquely Malay elements in Acehnese are borrowed, while the uniquely Chamic elements are inherited. The publication of Thurgood (1999) seems to confirm what scholars have long maintained, that is, there exists a subgrouping relationship between Acehnese and mainland Chamic languages. Earlier advocates have not changed their position, and both reviewers who dealt with that issue, Blust (2000) and Dyen (2001), confirm their endorsement of the relationship.

Second, although Collins is alone among scholars in arguing that the most immediate relatives of Acehnese are Malayic, not Chamic (p.c. in Blust 1992:74, fn. 6; Collins 1991), Collins is unquestionable right about there being a distinctly Malayic element in modern Acehnese. This will be dealt with next.

\(^5\)\(^5\) By the time of his 1975 dissertation, I. V. Collins no longer argued that the lexicostatistical evidence showed a especially-close genetic relationship between Acehnese and Malay.
2.0 Dating the migrations

Two migrations of Chamic speakers can be dated with confidence. The first migration involves what the literature on the Champa Kingdom calls the Northern Cham, Chamic speakers originally located around Hue (see Figure 2). The linguistic record makes it clear that the modern Hainan Cham, found near Sanya City, and the Northern Roglai, now found in southern Vietnam, were once the same group.\footnote{It is important to note that the Acehnese were not part of this Northern Cham group, as Acehnese shows no evidence of the two shared innovations that characterize the Northern Cham group.} The subgrouping of Northern Roglai and Hainan Cham is evident from the fact that these two groups share two rather distinctive sound changes, the loss of the word-final Proto-Chamic *-s after the vowel *-a and the change of word-final nasals to word-final stops.\footnote{See Thurgood (1999:224ff) for a more detailed discussion.} However, one part of the group is now on Hainan Island and the other part is now inland in southern Vietnam; the modern distribution only makes sense if they both descend from the Northern Cham, that is, the northern part of the Champa kingdom, with the merchants fleeing to Hainan and the non-merchant class moving south with the fall of the northern capital. This scenario also explains why the Northern Roglai are reputed to have had in their position some of the royal regalia from the kingdom of Champa---when the northern capital fell, it was the royal capital. In addition, the Chinese dynastic records date the arrival of the Hainan Cham. The northern capital of Champa, Indrapura, fell in 982, splitting the Northern Cham into two, with those engaged in trading migrating to the southern tip of Hainan Island while the remaining Northern Cham, most of those engaged in farming, crafts, and the like, migrated south. The Chinese records show these refugees arriving in Hainan and paying tribute to the Chinese just four years after the fall of Indrapura.\footnote{A second and apparently smaller group of Chamic speakers are recorded as arriving in Hainan in 1486.}

The second migration that can be dependably dated correlates with the fall of the southern capital in 1471 (see Figure 2).\footnote{The most phonologically distinctive of the Chamic languages, Hainan Cham, migrated to Hainan in two waves, the first in 986 after the fall of the northern capital and the second in 1486 after the fall of the southern capital. Its differentiation from the rest of the Chamic languages} The date 1471 fits remarkably well with the...
earliest date recorded for the Aceh dynasty, a date which is found on a Chinese bell (Reid 2006:10). This bell, which Reid describes as “part of the regalia of the Aceh kingdom,” has inscriptions in Arabic/Persian and in Chinese. The Chinese date (Reid 2006:10, and fn. 33) is equivalent to late 1469 or January 1470. Not only does the date on the bell fit with the 1471 date of the fall of the southern capital Vijiaya, but the linguistic fit is just as close: the linguistic evidence shows that at the time Acehnese split from the other Chamic languages, it was largely undifferentiated from the other Chamic dialects. As Durie (1990:111) notes, Acehnese does not share features consistently with any single group of Chamic languages; my own examination also concludes that the separation of Chamic and Acehnese predates the breakup of the Chamic group.

Thurgood (1999), expanding and documenting Durie’s position, concluded that, aside from the Hainan Cham, the break-up of Chamic into the distinctive modern descendents did not occur until after the 1471 fall of the southern capital at Vijaya. The fall interrupted the communication networks between the various Chamic communities along the coast and led to the formation of new communication networks with largely Bahnaric speaking rather than Chamic speaking groups. Here the linguistic evidence is compelling, but the evidence provided by the date inscribed on the Chinese bell rings even louder.

The sources of Malay contact

Without question, Acehnese displays significant Malay influence, enough for some analysts to have mistaken the Malay-related aspects as evidence of a genetic rather than a contact relationship.

One source of the Malay influence on Acehnese is the use of Malay as a language of wider communication between Acehnese and non-Acehnese speakers and, within Aceh, as a medium for writing, schooling, and so on. Malay and Acehnese had a diaglossic

---

10[10] Reid (2006:20) notes in an endnote that there is a widespread missing reading of the date as 1409 which was subsequently corrected to 1469-70 by Wolfgang Franke, Chinese Epigraphic Materials in Indonesia, Vol. I (Singapore: South Seas Society, 1988), p. 44.

11[11] It has been speculated that Acehnese and the rest of Chamic are sisters to each other, but no linguistic evidence has been presented to support this position. Methodologically, there is no reason to split it off; Acehnese is more conservative than the mainland dialects but one would expect that mainland dialects to be less conservative (far longer and more intense contact with Mon-Khmer) and there is no evidence of shared innovations that indicate a split.

12[12] It is also important to note that the Chamic languages remaining on the mainland do not have a group of shared innovations not also shared by Acehnese. The conclusion here is thus the same: Acehnese was one of the Chamic languages at the time it split off.

13[13] Reid (2006:19, fn. 25) notes that Paul Sidwell, in an unpublished reaction to my dating of the Acehnese migration “argues that the original connection of Aceh with Chamic languages may be much older, linked with similar patterns in Mon-Khmer languages of the Malayan peninsula”. I am unhappy reacting to an unpublished paper that I have no access to. However, the existence of Mon-Khmer material in Acehnese that does not occur in mainland Chamic has an obvious source: There is an abundance of Chamic place names there attesting to a Chamic and, thus, probably, an Acehnese presence. The existence of contact between these people and the Orang Asli is highly likely and the existence of some Orang Asli borrowings is certainly a possibility, but thus far no evidence for this has been put forth. In any event, neither the presence of Acehnese speakers on the east coast of Malaysia nor the existence of Mon-Khmer borrowings unique to Acehnese argues for an earlier dating of the primary Acehnese migration to Aceh.

It will be interesting to try to date the Acehnese presence on the east coast of the Malayan peninsula and to investigate whether some of the Mon-Khmer material in Acehnese is Orang Asli.
relationship,\textsuperscript{14} a relationship in which Acehnese was used in some contexts, while Malay was used in others. It is clear that Malay was the language of wider communication—the \textit{lingua franca} of the Malay world, as Durie (1996) and others have pointed out. Linguistically Malay dominated this world, playing the role not just of the language of wider communication but the role of the written language. Durie (1996:114) describes Malay as “the dominant \textit{lingua franca}, the language of the royal courts, letters, legal documents, scholarship, education, and cross-group communication”; complementing the use of Malay, Acehnese was the “language of the village and family, and of poetic tradition.” Poetry was ultimately for oral recitation, but when it was written down, poetry was always ‘framed’ in Malay, as the appropriate language for introductory prose material” (Durie 1996; Reid 2006:8). This older pattern of usage accounts in part for Durie’s (1996:1155) observation that written Acehnese first appeared only in 1658 and for the fact that it is only in modern times, according to Reid (2006:8), that Acehnese was used for prose, first by “the Dutch language purists in the 1930s, and more recently by some Aceh nationalist purists in the 1990s.”\textsuperscript{15}

There are other sources of Malayic, if not always Malay, influence. Significant pockets of Malay (or Malayic speakers) still exist within Aceh and others may have existed earlier. Tamiang is a Malay dialect, one spoken in part of Aceh, and one affecting the local Acehnese dialect. Tamiang doubtless has some effects on at least the local variant of Acehnese. In addition, Minangkabau is a Malayic language, although this seems to be a relatively recent presence. In any case, although much remains to be learned about the linguistic history of Aceh, it already seems apparent that insofar as Malay dialects and Malayic languages have been absorbed into Acehnese there is every expectation that these languages, both those languages that are contemporary and languages that have already disappeared have left their imprint on the local variants of Acehnese.

The spread of Acehnese

The Acehnese language appears to have spread quite rapidly in the last 500 years or so. Some of these speakers may or may not have come from the east coast of Malaysia. Evidence in the place names of the east coast makes it clear that there was an Acehnese population there at one point. As Thurgood (1999:23-24) writes,

\begin{quote}
“With respect to the path the Acehnese took on their travels, there is evidence that there was an important, Chamic presence in Kelantan, on the east coast of the Malaysian peninsula but this influence looks to have been quite late. As Abdul Rahman al-Ahmadi (1994 [1987]: 105-106) notes, local Kelantanese traditions have the King of Kelantan coming from Kembayat, an area some authors believe to be Champa, although others maintain it is Cambodia. Less controversial evidence of a Chamic presence is found in the numerous place names related to Champa: Pengkalan Cepa, Kampong Cepa, and Gong Cepa, to cite but a few. The Cepa of these names is obviously Champa, with the expected sound changes. These place names and other influences were the result of an Acehnese presence in Kelantan, not just a Chamic presence.”
\end{quote}

\textsuperscript{14} The relationship was broader than pictured here. As Durie (1996:117) shows, traditional literacy involved Arabic (> Malay (> Acehnese)), that is, literacy in Arabic (at least to some degree) came first, followed by literacy in Malay, and only then followed by literacy in Acehnese.

\textsuperscript{15} Durie (1985:1) notes that “With the exception of language mixing on the West and East coasts, Acehnese village people traditionally live day to day life in a purely Acehnese linguistic environment.” On the west coast, the Acehnese are mixed with Minangkabau speakers, with the Acehnese having come from the north and the Minangkabau from the south, both replacing earlier groups of speakers. On the east coast are speakers of Tamiang, a Malay variant which has influenced the Acehnese spoken there.

Draft Copy – Not to be Quoted Without Permission from the Author
In personal conversation, Gérard Diffloth has mentioned that Acehnese contains loanwords restricted to Aslian.\textsuperscript{16}\textsuperscript{16}

However, there seem to be too many Acehnese speakers to account for them simply by the migration from the coast of Vietnam, even if a group from the east coast of Malaysia were added in. What is more likely as a source is speakers of other languages shifting to Acehnese, a solution that at once accounts for the large number of modern speakers and for the considerable dialect diversity.

There is certainly evidence of speakers of other languages with the area dominated by Acehnese. Durie (1985:2) suggests that the Gayo speakers around Bireuen could have become Acehnese speakers; certainly, some have within recent history but an even larger group may have shifted earlier. Durie further notes (1985:1-2) that the “extent of the Acehnese language today owes much to the expansion of the Acehnese coastal kingdom based in Banda Aceh,” noting that the “expansion began during the time of Sultan Ali Mughayat Syah, when in the space of five years Daya, Pidie, and Pase were subjugated” (here Durie cites Iskandar 1958:38). As Figure 1 shows, other languages are still found within the area in which Acehnese is dominant: Tamiang, Mingangkabau, and Gayo, which have already been mentioned; Karo Batak, Alas, and Kluet are also spoken. Further, Reid (2006:5; cf. also pp. 6-7) notes the existence of a south Indian population, which would have spoken Dravidian e.g. Tamil or Indic languages. Certainly these speakers were present; the question of whether these speakers contributed to the diversity among Acehnese dialects remains open.

The apparent diversity that preceded the current dominance of Acehnese is certainly suggested in descriptions in The Travels of Marco Polo. Marco Polo writes of the whole of Sumatra that there are eight kingdoms on the island, each with its own king and its own language (Latham 1958:252). Durie (1985:2) tentatively places six of the eight kingdoms in North Sumatra, stressing that each had its own king and its own language.\textsuperscript{17}\textsuperscript{17} Durie, frequently citing Cowan (1950), notes that five of them were on the coast: \textsuperscript{18}\textsuperscript{18}

1. Ferlec (Perlak; In Pase: referred to by the Acehnese today as Syamtalira
2. Basman (Peusangan), see Cowan 1950),
3. Samara (Samalanga, see Cowan 1950),
4. Dagroian (somewhere in present day Pidie, and probably the Nakur of Chinese sources (Groeneveldt 1880-96-97 [reprinted 1960]) and Nagô of Acehnese oral tradition. And,
5. Lambri (probably near present day Krueng Raya in Greater Aceh)

Durie suggests that several of the groups now living inland once inhabited the coast; when these areas were incorporated into Aceh, some speakers were absorbed and some moved inland (only to be absorbed later). The general pattern is clear: initially Acehnese was found along the coastlines, but with the expansion of commercial agriculture in the interior, Acehnese also moved into the interior.\textsuperscript{19}\textsuperscript{19}

\textsuperscript{16}\textsuperscript{16} Certainly Diffloth is fully qualified to make such an observation. 
\textsuperscript{17}\textsuperscript{17} Reid (2006:8) similarly writes that the earlier diversity of the eight tiny port-states of Sumatra was noted by Marco Polo in 1292, who in reference to Sumatra as a whole, wrote that each of the eight kingdoms has its own language (The Travels of Marco Polo, trans. Ronald Latham (Harmondsworth: Penguin Books, 1958: p. 252 (not page 224)). Ferlec (Latham 1958:253); Basman (Latham 1958:253); Samara, called in Latham Sumatra (Latham 1958:254); Dagroian (Latham 1958:255); Lambri (Latham 1958:256). Oth; and, finally, Fansur (Latham 1958:253), not commented on by Durie.
\textsuperscript{18}\textsuperscript{18} This section is taken straight from Durie, only slightly paraphrased.
\textsuperscript{19}\textsuperscript{19} Durie (1985:3) suggests that “Greater Aceh and Daya on the west coast form the oldest Acehnese speaking area, for these are where the greatest dialect variation is to be found.”
Much work needs to be done on Acehnese dialects, however, before we can do much other than speculate. Such study of the Acehnese dialects should be a tremendous source of wider interest both for historical insights and, from a broader linguistic perspective, for what these dialects would contribute to our knowledge of linguistic typology, as the grammatical structure of Acehnese is remarkably interesting in and of itself.

Acknowledgements

My debt to Mark Durie is obvious throughout this paper, just as it is obvious throughout my book on Chamic (Thurgood 1999). Others who have contributed significantly are Bob Blust, Sander Adelaar, and Gerárd Diffloth. Finally, I wish to thank Anthony Reid, who brought me back to this fascinating topic.


