

The Malayic-speaking Orang Laut

Dialects and directions for research

KARL ANDERBECK

ABSTRACT

Southeast Asia is home to many distinct groups of sea nomads, some of which are known collectively as Orang (Suku) Laut. Those located between Sumatra and the Malay Peninsula are all Malayic-speaking. Information about their speech is paltry and scattered; while starting points are provided in publications such as Skeat and Blagden (1906), Kähler (1946a, b, 1960), Sopher (1977: 178–180), Kadir et al. (1986), Stokhof (1987), and Collins (1988, 1995), a comprehensive account and description of Malayic Sea Tribe lects has not been provided to date. This study brings together disparate sources, including a bit of original research, to sketch a unified linguistic picture and point the way for further investigation. While much is still unknown, this paper demonstrates relationships within and between individual Sea Tribe varieties and neighbouring canonical Malay lects. It is proposed that Sea Tribe lects can be assigned to four groupings: Kedah, Riau Islands, Duano, and Sekak.

KEYWORDS

Malay, Malayic, Orang Laut, Suku Laut, Sea Tribes, sea nomads, dialectology, historical linguistics, language vitality, endangerment, Skeat and Blagden, Holle.

1 INTRODUCTION

Sometime in the tenth century AD, a pair of ships follows the monsoons to the southeast coast of Sumatra. Their desire: to trade for its famed aromatic resins and gold. Threading their way through the numerous straits, the ships' path is a dangerous one, filled with rocky shoals and lurking raiders. Only one vessel reaches its destination. This ship is in the express service of the ruler of Srivijaya, and is guided through the treacherous waters by skilful people

KARL ANDERBECK has been involved with Indonesian linguistics since 2000, doing Masters research on Malay dialects in Jambi, and later, field research in Lampung, South Sumatra and Bengkulu provinces. His research interests are dialectology, historical linguistics, and sociolinguistics. His main publications include "Initial reconstruction of Proto-Lampungic: Phonology and basic vocabulary" (2007), in: *Studies in Philippines Languages and Cultures* 16: 41–165 and *Malay dialects of the Batanghari River Basin (Jambi, Sumatra)* (2008), SIL e-Books 6 (<http://www.sil.org/silepubs/abstract.asp?id=50415>). He is currently in a PhD programme at Universiti Kebangsaan Malaysia. Karl Anderbeck may be contacted at: anderbeck@gmail.com.

of the sea (*Orang Laut* in Malay). The other ship, seeking to trade without paying duties to the ruler, is boarded by raiders, the same people of the sea. Its sailors are executed and its booty confiscated.

The famous Malay kingdoms of history, from Srivijaya and Melayu to Malacca and Johore, could not have existed were it not for their loyal subjects, people of the sea. These sea nomads gathered important products for trade, enforced the use of authorised trading ports and punished transgressors. This brief and preliminary study focuses attention on a narrow aspect of these historically important groups, namely their speech. What language(s) to these sea people speak (particularly those in the Riau Archipelago), and how is their speech related (or not) to that of their traditional vassals, the Malays? With the drastic changes in the world, do these groups even exist anymore?

1.1 WHO AND WHAT ARE THE SEA NOMADS?

Given the archipelagic nature of Island Southeast Asia, it is not surprising that numerous ethnic groups have made their living primarily from the sea. Some have exploited the sea while living on land, while others have lived a more nomadic existence, even to the point of living on their boats rather than land. This nomadism allowed these latter groups to move from place to place, harvesting different products in different seasons (Chou and Wee 2002: 334). These groups have been variously called *Sea Tribes*, *Sea Nomads*, *Sea Gypsies*, *Boat People*, and, in Indonesian/Malay, *Bajau*¹ (sea gypsies), *Orang Laut* (sea people), *Orang Suku Laut* (people of the sea tribes), *Ra(k)kyat Laut* (sea subjects [to Malay rulers]), or *Orang Perahu/Sampan* (boat people).² In this paper, examining as I do the interrelationships of various groups, I use the term (*Malayic-speaking*) *Sea Tribes* to refer to them and their collective speech varieties.

B.W. Andaya and L. Andaya (2001: 14) note that evidence of “communities of able seafarers” in western Nusantara has been dated to as long as three thousand years ago. As far back as the maritime state of Srivijaya in the seventh century, *Sea Tribes* have played a key role in the history of the region, enforcing the dominance of certain ports, functioning as the navy of Malay rulers, and gathering important sea products for trade, usually in a patron-client relationship with those rulers.³ They steadfastly maintained their own identity, resisting the pressure to *masuk Melayu* (become Malay) with all its cultural trappings. Their glory days, however, waned sharply beginning in 1699 following the arrival of the Bugis into the West Nusantara trading network and the assassination of Sultan Mahmud Syah II, the ruler of Johor. Not only was their linkage with Malay dynasty (and patronage) weakened,

¹ “Bajau” usually is used as an ethnonym referring to non-Malayic-speaking seafaring groups further east (see below) but has also been used as a general term, particularly in colonial-era literature.

² We do well to keep in mind T. Barnard’s (2007: 34) caution that “[a]lthough the title *Orang Laut* suggests a certain amount of homogeneity, it is a relatively artificial designation”.

³ In Barnard’s (2007: 34) memorable phrase, “[t]he Straits were a highway in which the *Orang Laut* were the toll collectors”.

but the Riau Sea Tribes were also soon eclipsed technologically by Illanun raiders in the late eighteenth century (Barnard 2007: 45). The second half of the nineteenth century, with the rise of Industrial-Age colonialism, saw many groups abandoning nomadism (Sopher 1977: 114), a process which still continues. Maintaining a distinct ethnic identity from Malays, formerly economically profitable (L. Andaya 2008: 184), has become an economic and cultural liability (Lenhart 1997: 586). Today, “enormous changes [are] occurring at a very rapid rate” (Lenhart 2001: 67) in traditional Sea Tribe areas. Commercial fishing, resettlement programs, development and seizure of traditional fishing grounds, among other factors, have seriously marginalized the Sea Tribes. Traditional lifestyles are being abandoned, populations are shrinking or assimilating to the majority culture, and a number of historical groups have completely disappeared.

1.2 TAXONOMY OF SEA TRIBE LECTS⁴

Sopher (1977: 50) provides a macro division of the languages spoken by the Sea Tribes of Nusantara:

1. Mawken
2. Malayic
3. Bajau

Map 1 shows the location of the Mawken and Malayic groups. The Mawken (Moken/Moklen) of central and northern Thailand, and the Bajau of the southern Philippines, eastern Borneo, Sulawesi and further east, both speak Malayo-Polynesian but non-Malayic languages. As the focus of this paper is on Malayic-speaking groups, I will not speak further of Mawken or Bajau.

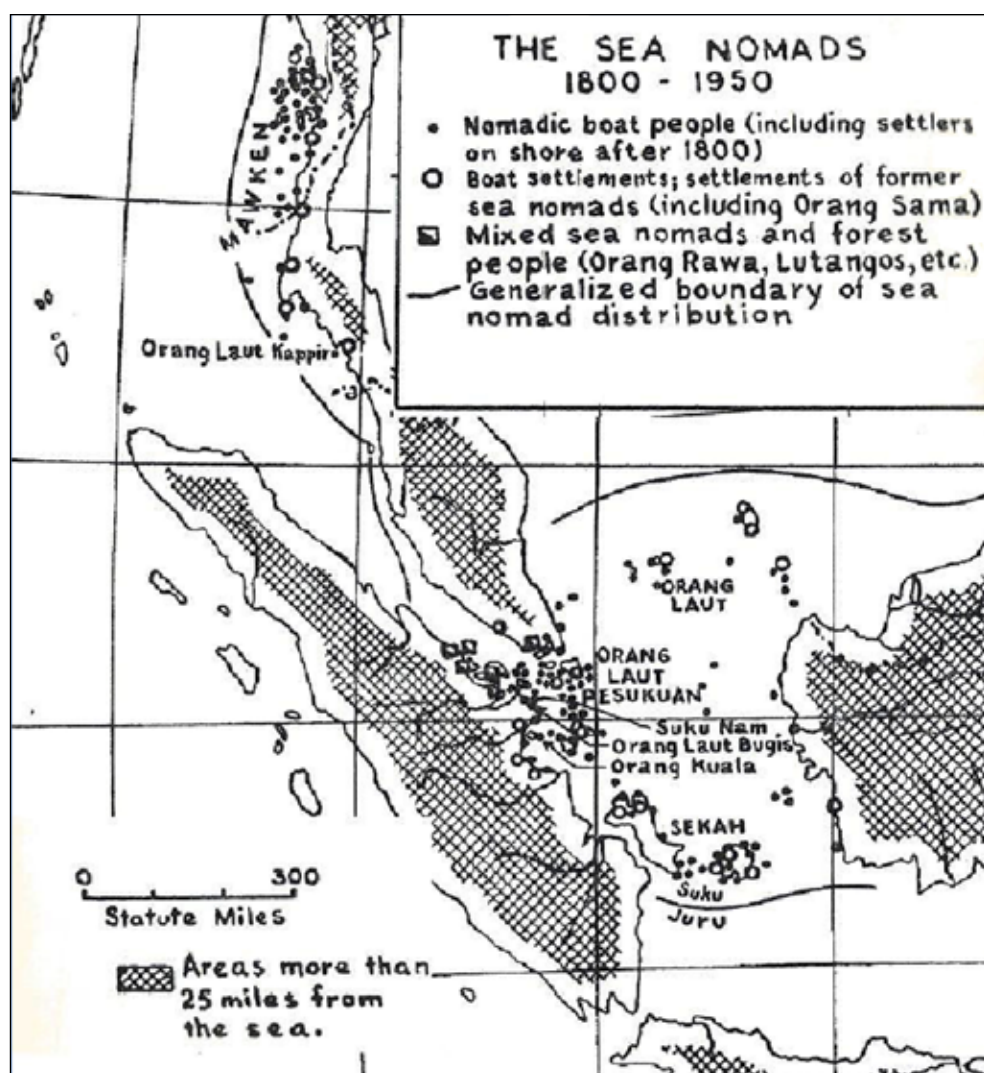
Within the Malayic grouping, my analysis (this paper) shows a four-fold division of Sea Tribe lects (see Map 2):

- A. Kedah
- B. Riau Islands
- C. Duano
- D. Sekak

Group A refers to the small subset of Malayic Sea Tribe lects found from southern Thailand to Kedah in Malaysia, individually embedded within what has been called the *Kedah* dialect (Collins 1988). The best-known of these groups is Urak Lawoi', which will be discussed in more detail below. *Riau Islands*, the focus of this paper, refers to those groups clustered in the

⁴ In this paper I use the neutral term *lect* to denote a given speech variety, agnostic as to where it may fall on the language-dialect continuum. It may be helpful to note that the Malay/Indonesian term *bahasa* (< Sanskrit), commonly translated as 'language', is actually better translated as the more noncommittal 'lect', denoting everything from idiolect (*bahasa aku* 'my way of speaking') to a plurality of languages (*bahasa daerah* 'local language(s)'). I prefer *lect* over the semantically equivalent but distractingly polysemous term *variety* (Chambers and Trudgill 1998: 5) or the somewhat redundant *isolect* (Hudson 1967: 12).

Riau-Lingga Archipelago between Sumatra, Singapore, and (to a lesser extent) the west coast of Kalimantan. I argue here that this geographical collection of lects has a collective set of features which both distinguishes it and sets it within the larger dialect network surrounding it. *Duano* is the name of a single ethnolinguistic group found off the east coast of the Sumatran mainland (Riau and Jambi provinces) and the west coast of Johor. Although technically located in the Riau-Lingga Archipelago, linguistically Duano is very different than anything else found there and is thus classified separately. *Sekak* is the lone described Sea Tribe lect in the islands of Bangka and Belitung (although reports exist of other groups) and is quite distinct from the sedentary Malay lects of Bangka and Belitung. Although I will touch on Kedah, Duano and Sekak in this paper, the focus will most strongly be on the Riau Islands lects.



Map 1. Distribution of Sea Tribes in western Nusantara (Sopher 1977: Plate III)



Map 2. Malayic Sea Tribe lect groupings.

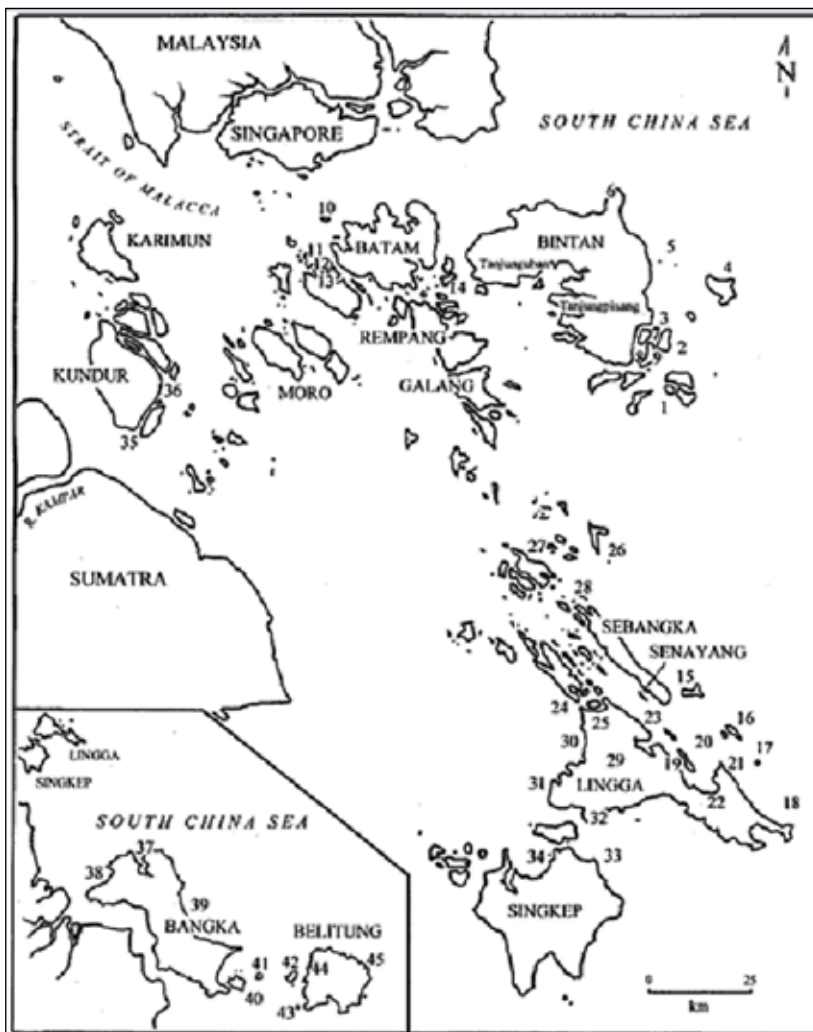
Within the Sea Tribes of the Riau Islands, Lenhart (2001: 84) makes another four-fold social division:

- a. Mantang
- b. Mapor (Mapur)⁵
- c. Barok
- d. Galang (sedentarized and mostly assimilated)

Due to a lack of data, it is currently unclear whether Lenhart's division is also reflected dialectally. Moreover, this division should be not accepted uncritically, for two reasons. First, many more Riau Sea Tribes had been reported in the past (for example Seletar, Tambus, Moro, et cetera) but many of these have since assimilated into Malay society. Second, every ethnic listing contains slightly different membership, and conceptions of ethnic affiliation are quite fluid; compare Chou (2003: 34, 45). Chou (see Map 3) lists forty-five different Riau Sea People territorial groupings which evidently still exist,

⁵ While samples of Mantang, Galang, and Barok are represented in the dataset for this study, it is unclear if any of the lists in the dataset actually correspond to Mapor. The identity of the Sea Tribe known as Mapor is somewhat of a mystery. O. Smedal's (republished) thesis (1989), per M. Kartomi, reports that the Mapur Tribe are named after the Mapur mountain in Northeast Bangka and are part of the Lom Tribe, indeed that Lom customary law (*adat*) is called *adat Mapur*. In Smedal (1987: 1-2), he states flatly, "Orang Lom and Orang Mapur are two terms for the same group - and language - in the district Belinyu" and demonstrates that the term Mapur has been used by colonial writers since the mid-nineteenth century. However, Kartomi (personal communication 2010) writes, "My informants in Belinyu [Bangka]/Suku Laut [Sekak] said some of their group contained people from an island to the north called Mapur, and that they were forced off the island by the government and sailed to Bangka where they settled among the Sekak. Local Bangka officials confirmed this". Smedal and Kartomi, after comparing notes, concluded there must be two Mapur groups.

and reports (2003: 18) that, with the collapse of the feudal system into which the Sea Peoples were socially ranked, they less often refer to themselves by a particular *suku* and more by what territory they belong to. She additionally reports (2003: 10) that “[e]ach Orang Laut clan possesses its own dialect”, although this information may be based more on hearsay than evidence.



Map 2.1 The spread of Orang Laut in the Riau-Lingga Archipelago.

Key: (1) Orang Pulau Toi, (2) Orang Tanjung Sekuang, (3) Orang Pulau Buton, (4) Orang Mapur, (5) Orang Berakit, (6) Orang Panglung, (7) Orang Tanung Senkuang, (8) Orang Air Kelobi, (9) Orang Pulau Malim, (10) Orang Dapur Arang, (11) Orang Pulau Bertam, (12) Orang Pulau Padi, (13) Orang Pulau Boyan, (14) Orang Nginang, (15) Orang Kentar, (16) Orang Kojong, (17) Orang Pulau Buluh, (18) Orang Mensemut, (19) Orang Sungai Liang, (20) Orang Pulau Hantu, (21) Orang Air Kelat, (22) Orang Pongok, (23) Orang Kungki, (24) Orang Linau, (25) Orang Air Batu, (26) Orang Mamut, (27) Orang Pulau Medang, (28) Orang Limas, (29) Orang Pancur, (30) Orang Tembuk, (31) Orang Lelumu, (32) Orang Mentuda, (33) Orang Penuba, (34) Orang Sungai Buluh, (35) Orang Tanung Batu, (36) Orang Sebele, (37) Orang Mantung, (38) Orang Teluk Kampa, (39) Orang Baturusa, (40) Orang Pulau Lepar, (41) Orang Pulau Liat, (42) Orang Pulau Mendanau, (43) Orang Pulau Seliu, (44) Orang Tanung Pandan, (45) Orang Teluk Pring.

Map 3. Sea Tribes in the Riau Archipelago (Chou 2010: 26).

1.3 DIALECTOLOGY IN THE MALAY WORLD

The vernacular Malayic lects of Sumatra, the Malay Peninsula and (at least) western Kalimantan form what could be considered a geographical dialect continuum. A dialect continuum is where adjacent areas can understand each other but have small linguistic differences. As the geographic distances increase, the linguistic differences increase and hence, intelligibility decreases, to the point where the extreme end points may be considered separate languages. These continua are rife in our world, for example the massive Romance dialect continuum in Europe. In fact, they may be the rule rather than the exception (Chambers and Trudgill 1998: 6). Just as the Romance dialect continuum ultimately derives from (Vulgar) Latin, the Malayic dialect continuum derives from Proto-Malayic (PM; reconstructed in K.A. Adelaar 1992).

Dialect continua mean that languages are not like marbles, but neither are they a perfectly graded scale. M. Paul Lewis, editor of the *Ethnologue*, puts it like this, "Language is a lot more like oatmeal, where there are some clearly defined units [lumps] but it's very fuzzy around the edges" (Erard 2005). The "lumps" are frequently areas which have achieved sociopolitical, demographic or economic dominance. That the Malay World forms a dialect continuum is elucidated well in Collins (1989), which describes complex, layered relationships and often gradual gradation from one area to the next. Collins cautions us not to view the sea (or current political borders) as boundaries. In Nusantara, "the sea functions as a major communication route guaranteeing a high density of communication" (1989: 255).

The next section describes some "lumps" pertinent to this study.

1.4 PROFILES OF URAK LAWOI', DUANO, JAKUN, AND SEKAK

Since the focus of the paper is the Riau Sea Tribe lects, I will try to dispose of the non-Riau lects all at once. Accordingly, here I give brief portraits of a few of the more distinctive Sea Tribes lects plus one inland "tribal Malay" lect.

1.4.1 URAK LAWOI'

Urak Lawoi' [urk],⁶ which translates as 'sea people' and is cognate with Malay *orang laut*, is spoken by approximately three thousand people "located in villages on the islands off the west coast of Thailand from Phuket Island to the Adang Island group" (Hogan 1988: 1). In contrast with the nomadic Moken (Mawken), as of the 1960's the Urak Lawoi' were strand-dwellers, making long journeys in their boats but returning to fixed settlements (Hogan 1972: 215). Robert (2010) is a dissertation documenting how traditional Urak Lawoi' lifeways and language are rapidly losing ground to outside influences, both Western and Thai.

Previous linguistic research on the group, besides the two works by Hogan

⁶ When introducing a lect which has been separately identified in the *Ethnologue* (Lewis 2009), for clarity's sake I will include the language's three-letter ISO 639-3 identifier.

(1972, 1988), includes A. Saengmani's (1979) MA thesis about Urak Lawoi' phonology, while H. Steinhauer (2008) used Hogan's data for further analysis of Urak Lawoi' sound changes, focusing on the Phuket Old People's dialect.

Noting that the most salient sound changes occur in the final syllables, Steinhauer (2008: 125) details a number of sound changes from PM, including:

1. Insertion of a glide in vowel clusters beginning with a *high vowel. The example **laut* 'sea' > *lawoi?* illustrates innovations 1-4.
2. Lowering of *high vowels to their mid and lower-mid pendants in closed final syllables.
3. Diphthongization of non-front vowels before *-s and *-t.
4. Glottalization of final *stops and *fricatives.
5. Change of final nasals into their corresponding voiceless stop, unless the onset of the final syllable was also a nasal (for example *urak* 'person' < **uraj*).
6. Simplification of homorganic nasal-stop sequences. Sequences with voiced stops reduce to the nasal component (for example *taja* 'ladder' < **taja*), while those with voiceless stops reduce to the stop component (for example ***tupol* 'blunt' < **tumpul*).
7. Change of final *-l into -n (for example *tupon* 'blunt' < ***tupol*).
8. Lateralization of *-r (for example *lapal* 'hungry' < **lapar*)
9. In terms of its lexicon, Urak Lawoi' is fairly mainstream, containing some seemingly unique words but probably no more than the typical Malay dialect.

1.4.2 DUANO

Duano [dup] is spoken by approximately 17,500 people, the majority in the coastal region of Riau and Jambi Provinces, with a minority on the facing coast of Johor, Malaysia. They also go by the names *Orang Kuala* (people of the estuaries) and *Desin Dola'*, a phrase meaning 'people of the sea' in what presumably is the non-Malayic but still Austronesian substratum of their language. As with the Urak Lawoi', they are not (currently) nomadic.

Previous linguistic documentation of Duano includes:

- J.G. Schot (1884) provided about 150 Duano lexical items in a geographical survey of the Kateman river basin in Riau.
- W. Skeat and H.N. Blagden (1906; henceforth SnB) published a short list collected near Malacca (Tanjung Seginting). It was listed as "unidentified", but comparing the items to other Duano data makes its provenance clear.
- H. Kähler (1946a, b) furnished a lexicon of nearly 450 Duano items spoken on Rangsang Island.
- E. Seidlitz (2007) wrote a brief phonology of Duano based on sites in both Malaysia and Indonesia. His paper also contains a more complete history and bibliography for this group. He has kindly shared with me two Malaysian and two Indonesian 500-item wordlists.
- M.S. Yusof's (2006) dissertation focuses on language obsolescence among the Orang Seletar (Johor and Singapore) and (Malaysian) Duano. Since I

have not seen it, I am unaware whether it contains linguistic data, but it is one of the few sociolinguistically oriented works on Sea Tribes.

Although Duano includes significant internal dialect variation, it is nevertheless clear what qualifies as Duano and what does not. Duano is likely the most aberrant Malayic lect in terms of its lexicon, with only 59% of its basic vocabulary derived from PM, and 56% similarity to both Standard Indonesian and Standard Malay (McDowell and Anderbeck 2008). Its sound changes make it one of the more phonologically divergent Malayic lects as well. Innovations include: strengthening of PM **h* to *ʔ* (or retention of PMP **q* as *ʔ*), PM **k* > *y* (or *y* as intervocalic allophone of **k*; compare Seidlitz (2007)), loss of final **r*, final open **a* > *u*, and raising of **a* in closed syllables after a voiced stop, particularly in final syllables (described in more detail in Section 3.2). Describing Duano sound changes is difficult, however, for two reasons. The first is due to the internal dialect variation. For example, two dialects regularly occlude final **nasals* (for example *kəyeəʔ* 'dry' < **kəriŋ*) while the others do not or do so only occasionally. The second difficulty is the likely existence of a non-Malayic substratum, which requires ferreting out which sound changes belong to which layer.

I will say a little more here about possible substratum, although a whole paper could easily be written on the issue. C. Pelras (2002: 6) listed a number of non-Malay words, being mostly unsuccessful in his attempt to determine their origin. I have a very broad database of Malayic lects yet most are mysteries to me as well. A few likely Mon-Khmer terms can be found. Duano *coʔ* 'fall' seems to correspond with the Proto-Aslian-derived Jahai *co* (SnB)/*cərah* (T. Phillips In progress), *cooy* 'wood' is also likely Aslian (Mon-Khmer languages of Peninsular Malaysia). However, T. Phillips (personal communication) assures me that, on the whole, Aslian is not the primary source of the unknown words.

In terms of sound changes, I will make a brief comment on Duano reflexes of a few Proto-Malayo-Polynesian (PMP) phonemes. First, PMP **q*, which came to Proto-Malayic as **h*, is reflected in Duano in a way that is basically unknown in Malayic (see Table 1).

PMP (Proto-Malayo-Polynesian)	Duano	Proto-Malayic
*qulu 'head'	kulu	*hulu(?)
*qijuhun 'nose' (Adelaar 1992)	kəloŋo	*hidun
*taqu 'know'	taɣu	*tahu(?)
*qatay 'liver'	ɣati	*hati
*bunuq 'kill'	buna?	*bunuh
*buaq 'fruit'	bua?	*buah

Table 1. Reflexes of PMP **q* in Duano.

One other deviation from Malayic, which merged PMP *-uy with *-i, is the Duano lexeme *maloŋoy* 'swim' < PMP **laŋuy*. To add to the mystery, Kähler (1946a, b; as discussed in G. Benjamin 2009: 317) documents the use of Austronesian verb infixes *-(u)m- 'actor focus' and *-in- 'perfective aspect' in Duano, which are otherwise nearly non-existent as productive forms in Malayic. Is this a retention from very early Malayic, or from a substratum language?

Clearly, something is different in Duano from any other Malayic lect yet described. Since Duano otherwise shares many Malayic (even Malay) innovations, my preliminary and scantily supported conclusion is that Duano has an Austronesian non-Malayic substratum.⁷

1.4.3 JAKUN

The Jakuns are a Malay-speaking inland *Orang Asli* (original people) group in southern peninsular Malaysia (Pahang and Johor) with an ethnic population of nearly 28,000 (Seidlitz 2005). Although they are not a Sea Tribe, their inclusion here is primarily due to the fact that Jakun shows some linguistic resemblance to various Sea Tribes discussed in the paper, and has been linked by earlier writers to Sea Tribe lects.

As Seidlitz's (2005) MA thesis on Jakun phonology discusses previous research on the Jakun language and society, I will only mention the data sources for this paper, which are fourteen 200-item wordlists from Seidlitz, and a handful of century-old lists (mostly under 50 items) from *Orang Asli* groups of southern Malaya compiled by Blagden (1906).

These British colonial writers (Skeat and Ridley 1900; Skeat and Blagden 1906) judged the various tribal groups they encountered, including Sea Tribes, as being Aslian ("Sakai") in origin. At least for the Jakun, there would seem to be some basis for this claim. The Jakun vocabularies⁸ provided by SnB were a mix of Malayic and Aslian (Mon-Khmer) words. My examination of the modern-day Jakun language reveals continuity with that of a century ago, but the process of Malayization of their language has continued to the point where Aslian words form only a small minority.⁹ This trajectory gives credence to SnB's assertion that the Jakun people originally spoke an Aslian language (perhaps akin to the extinct Kenaboi) but fell into the Malay sphere, versus the

⁷ The discussion in W. Mahdi (2009) regarding the possible trade role of pre-Moken (Mawken) speakers in the Malacca Straits, as well as the shared and relatively rare change PMP *q > k (G. Thurgood 1999: 58–59) would indicate Moken as a possible substrate language of Duano. However, M. Larrish, an expert in Moken, was not able to identify any lexical innovations shared between Moken and Duano (personal communication), nor have I identified any unusual *aba* sequences in Duano as are seen in Moken (Mahdi 2009: 80).

⁸ Until fairly recently the term Jakun was used broadly to denote any Austronesian-speaking *Orang Asli* of Malaysia. Nevertheless, the lists labeled Jakun in SnB do show broad similarity to each other.

⁹ One example: *ayih/ajih* 'you (singular)', is found in both nineteenth century and modern Jakun, with a probable cognate in extinct Kenaboi *yei* identical, and reflexes in many Aslian languages as something like *ajih* 'that; there'.

opposite possibility (as promoted by P. Schebesta (1926)) that the Jakun were originally of Malayan stock but were (linguistically) Aslianized by virtue of their location.¹⁰ This Malayization accords with the general pattern of *masuk Melayu* (become Melayu) known from time immemorial (see Benjamin 2002).

1.4.4 SEKAK

In the Malay zone of Bangka and Belitung is a Sea Tribe variously identified as *Loncong* (*Lontjong* in earlier Dutch spelling), *Sekah*, *Sekak*, and *Sawang* [Ice]. A rough estimate of their current population is under 600, down from 1700 or more when the Dutch counted them in the late nineteenth century. According to an article in *Kompas* magazine (Mama 2010), they are still nomadic. It is unknown how much of the Sekak language is still spoken.

A few works have been published containing linguistic data on Sekak:

- In 1881 J.G.F. Riedel published two folk stories in the Sekak language (Riedel 1881).
- Exactly 100 years after Riedel's publication of Sekak folk stories, the Indonesian government's *Pusat Bahasa* published *Struktur bahasa Sekak* (Napsin et al. 1981) which included some grammatical aspects and an appended wordlist.
- In addition, K. Anderbeck and U. Tadmor in a work-in-progress explore some historical linguistic aspects of Sekak based on the data cited above.

The existence of samples separated by a century provides the opportunity to examine the linguistic changes that have occurred over that period. The vocabulary extracted from the 1881 folk stories shows some interesting things. First, a large number of words are of Javanese origin, including but not limited to: *ucul* 'become loose', *sədulur* 'sibling', *lacut* 'break off', and probably *milu*, *melu* 'go, depart' < JV. *milu*, *melu* 'follow'. A significant percentage of the lexicon (probably below 25%) cannot be attributed to Malay or Javanese, such as *bəŋkur* 'breast', *ayau* 'person', *mənam* 'female', *umar* 'blood', *marus* 'white', *sawan* 'sea', *ŋaŋui* 'answer', *tuyu* 'rice', *di rapak* 'where' and the relativiser *mo(h)*. Additionally, many of these items break Malay phonotactics as outlined in Adelaar (1992). Of the unknown items, only five have identifiable links to PMP. For example, *alum* 'inside' is certainly Austronesian but probably non-Malayic given the reflex of PMP *e.

In terms of structure, the language seems Malay in some sense and non-Malay in another. The word order, particularly that of nouns and deictics, seems reverse from the typical Malay order, for example *iti jukut* 'that fish' versus SM (Standard Malay) *ikan itu*. Possession similarly is odd: *aku əmpun gaŋur* 'my weapon'. This lect also did not seem to employ the ML nominalizing suffix *-an*, for example *pərbuat* versus SM *perbuatan*, *pərmula* versus SM *permulaan*, *ucap* 'utterance' versus SM *ucapan*.¹¹

¹⁰ Of course both possibilities could be true to some extent; likely some of ancient Malay stock never took up the *Melayu* identity.

¹¹ These features are not unknown in the Malay World, particularly among contact

The 1981 Sekak sample, while clearly a descendant of the earlier Sekak, falls much more in line with other neighbouring Malayic lects (spoken by both Bangka/Belitung Malays and Lom “tribals”) than its older self. Of the Sekak words which rarely occur in Malayic lects or have different semantics, only a quarter were retained. Even more striking, of the unknown lexemes from the nineteenth century, basically all had disappeared by 1981.

In terms of phonological innovations, both 1881 and 1981 Sekak innovations are documented rather comprehensively in Table 3.

1.5 ORIGINS

The origins of the West Nusantara Sea Tribes and their Malayic speech are obscure; linguistic evidence of a non-Malayic substratum is slim to none. Skeat and H.N. Ridley, in their brief report on river nomads of Singapore, the Orang Kallang, cite the words *koyok* ‘dog’, *kiyan* ‘come, come here’, *kiyun* ‘go away’ and *kiyoh* ‘far off’ as the only non-Malay terms their informants could produce for them. They correctly noted that cognates of the latter three terms are very common in the tribal lects (both Malayic and Aslian) of Johor, Pahang, Negeri Sembilan and Selangor including Jakun. It is unlikely, however, that these terms should be classified as Aslian loans into (mostly) tribal Malayic lects: more likely as Malayic into southern Aslian languages which are already rife with Malayic loans.¹² Beyond a few exotic terms, however, Riau Sea Tribe lects are extremely lexically mainstream (see Section 3.3).

Sound changes are unlikely to provide much evidence for substratum in a case like this. While it is conceivable that unique features to, say, Aslian languages could be evident in the target language, like vowel length distinctions, or an expanded phoneme inventory matching Aslian phonemes, most sound changes could just as easily be explained by other causes. Sound changes, however, can more easily shed light on the relationship between the Malay spoken by the Sea Tribes and other Malayic lects. Such questions will be explored in detail below.

Benjamin (2002: 18) criticises the *kuih lapis* (layer cake) view of ethnology that has characterized much scholarship of the past century, and which assumes that “tribals” are unchanged products of earlier migrations than “civilized” groups. He argues that “tribal”, while a valid ethnological category, is better understood as a synchronic strategy of maintaining separateness from centres of power than as a retention of ancient lifeways. For this and other reasons, he states that the “search for the remoter ‘origins’ of any of the constituent populations will therefore be misconceived – and with it the search for a supposedly single ‘origin’ for the Malays themselves” (Benjamin 2002: 23).

Malay varieties. Sekah’s relationship with other Malay varieties is explored more in Anderbeck and Tadmor (In progress).

¹² The troika *kian*, *kiun*, and *kiuh* are all likely bi-morphemic, with Malay *ke* ‘to, toward’ as the first element. *Kian* ‘this, thus, this way’ (as in *sekian*) is found in standard Malay dictionaries, and Blagden notes Belitung Malay *siun* ‘there’, likely using the *-un* morpheme in a typically Malay manner. *Koyok* ‘dog’ is actually a fairly common term in Kalimantan and not unknown elsewhere, so should not be considered Aslian.

1.6 SOME NOTES ON TERMINOLOGY

The previous section leads well into a needed side-trip regarding terminology. First, identifying ethnic groups in the Malay World is a task most confusing. Ethnonyms come in a bewildering thicket; terms like *orang asli* (original or aboriginal people), *orang darat* (people of the inland), *orang hutan* (people of the forest), and *orang laut* have been applied at will to various groups, with some groups (for example Orang Seletar) being labeled both *orang laut* and *orang hutan*! In addition there are place names (which can overlap with other place names in hierarchical fashion), names given to reflect social or political organisation (like *orang batin*, referring to a special type of tribal leader, and *orang pesukuan* 'people of the tribal divisions'), religion or lack thereof (*orang kappir* 'infidel', *orang (be)lom* 'people yet without religion'). Care is thus needed. In any attempt at ethnic labelling or categorisation, we do well to heed the caution of those who study ethnicity and who tell us that ethnic conceptions are variable and layered.

Second, I must define some terms used in this paper. By "Malayic" I follow Adelaar's definition as presented in (1992) as any lect which descends from Proto-Malayic, whether or not its speakers make any claim to being culturally Malay.

The definition of "Malay", both as a linguistic as well as an ethnic term, is much more complicated, in fact entire books have been written on the subject (for a linguistic perspective, see for example Collins (1998b); regarding Malay ethnic identity, see for example Benjamin and Chou (2002), Barnard (2004) and L. Andaya (2008)). Benjamin (2002: 50) argues that the modern conception of Malay or *Melayu* – since the late nineteenth century – is not a state but an achievement: "one must *act Melayu*". This acting *Melayu* involves three components: "language, Islam and an acceptance of social hierarchy". Benjamin and others demonstrate the fluidity of this category, where various tribals in some contexts consider themselves or are even considered by others as Malay. Yet it is the modern sense of *Melayu* which I employ in this study. Thus the cultural-linguistic term canonical Malay used in this paper, as I understand introduced by Collins (1997), refers to the Malayic lects spoken by people who define themselves by the narrower *Melayu* label. I also follow Benjamin's (2002: 7) restricted definition of Malay World, "the areas currently or formerly falling under *kerajaan Melayu*, the rule of a Malay king".

I mention above my use of the term tribe and tribal. Benjamin (2002: 8) argues that tribal status is a response to the "classical civilizing process" and thus does not involve autonomy but rather heteronomy, and is a fluid conception that can involve movement both out of as well as into this category. Hence, he argues that many Sea People today may be more tribal than when they were in the direct service of Malay rulers. However, it seems that the strongest pressure over the past half century has been in the direction of de-tribalization and assimilation.

1.7 RESEARCH QUESTIONS

To summarize, this study employs existing research to examine the relationships between Sea Tribe speech varieties and with other Malay dialects. Such historical linguistic/ dialectological analysis can round out the emerging picture of Malay(ic) dialectal variation throughout Nusantara, shed light into historical relationships, and may also prove theoretically interesting. We also face some urgency here: many of these speech varieties are endangered and subject to strong pressures, therefore unique matters are being lost, both to researchers and to the Sea People themselves.

Questions

- Do Sea Tribe dialects cohere or adhere? What relationships seem to be closest? In other words, how do these speech varieties fit within the larger Malayic dialect continuum? Can we find any lumps, areas which cluster linguistically?
- Do we see significant unknown (non-Malayic) linguistic characteristics in these lects?
- Do all seem to be Malay (versus Malayic)?
- What do we know, what do we not know? What field research is needed/ where?

2 DATA SOURCES AND METHODOLOGY

2.1 PREVIOUS RESEARCH

The state of the art of research into Nusantara Sea Nomads remains Sopher (1977, originally defended in 1954). That dissertation, however, was mostly composed of secondary research rather than fieldwork. Chou (2003: 7-8) writes,

[R]ecent literature based on field research on the Orang Laut remains dismal [...] most of the published literature on the Orang Laut dates from the mid-nineteenth century. It comprises a heterogeneous collection of travel accounts, geographical monographs, local histories and administrative reports.

This dismal status has been remedied a great deal by the fieldwork of Chou (2003, 2010), Lenhart (1997, 2001, 2002) and some others, but little work has been linguistically- or sociolinguistically-focused.

The great majority of existing linguistic descriptions, such as they are, focus on one dialect or another. This paper attempts to stitch together these resources into a more coherent picture. Beyond the studies mentioned in Section 1.4, here is a summary of resources which contain linguistic information on Sea Tribes Malayic. I will highlight the aggregators first:

- Skeat and (primarily) Blagden (1906) collected published and unpublished tidbits of information on the speech of six Sea Tribes (among others).¹³ In

¹³ My thanks go to T. Phillips for digitizing and sharing these lists with me.

the analysis below the date of the list is given as 1906 unless more specific information on the time of fieldwork was provided in SnB. The ethnonym and number of wordlist items of the six lists are:

- Barok (74)
- Galang (32)
- Temiang (6)
- Kallang (Singapore; SnB ‘Orang Laut’ but actually river dwellers) (17)
- Trang/Kappir (southern Thailand) (12)
- Tanjung Segenting (Duano, Malaysia) (15)
- Similarly, W.A.L. Stokhof (1987) published a number of colonial-era wordlists labeled “Holle lists”, including of three non-canonical Malay-speaking groups of Riau. All three were elicited in 1905:
 - Orang Hutan/Darat of Batam (list #93, 66 items),¹⁴ referred to in this study as Orang Darat Batam;
 - Tambus, Sanglar Island (between Moro island and the mainland), Riau, 1905 (list #94, 72 items);
 - Mantang (list #92, 215 items).
- Kähler (1960, based on fieldwork in the 1930’s) contains ethnographic and linguistic description of *Orang Darat* of Batam Island with contrastive examples from other tribal groups of Riau: *Orang Akit* (‘Raft People’; Rupal Island off the Sumatran coast), *Orang Hutan* (Tebing Tinggi Island, off the Sumatran coast further south), and *Orang Laut* whose location is given only as the Riau Archipelago.¹⁵ The book includes a fairly extensive comparative vocabulary of the four groups, which was re-purposed for this study.
- M.D. Kadir et al. (1986) studied three dialek Orang Laut of Riau, containing 230-item wordlists (Barok, Mantang, Galang). These lists are referred to by their location, appended with “Kadir”.

Here are more specific studies, from north to south:

- S. Smith et al. (1814: 182) documented 21 words of the “Language of the Orang laut” of “certain of the islands lying off the western coast of Queda [Kedah], particularly Pulau Lontar”. In this paper I refer to this list as “Kedah OLaut”. Although the Urak Lawoi’ identifies Lanta Island (presumably the same island) as the place from which they spread (Hogan 1988: 1), it is not immediately clear how close the two varieties are; this

¹⁴ All “Holle” wordlists published by Stokhof were numbered. The total number of lexical items contained in each list varied, so the Mantang list, given number 92 by Stokhof, contained 215 words, while the Orang Hutan/Darat list, given number 93, contained merely 66 words.

¹⁵ However, see also the caveat of Adelaar (2008: 15), “Kähler’s sketchy data are difficult to read because he basically tries to describe four dialects at once, because of the nature of his analysis which is more outdated than is suggested by his terminology and, last but not least, because of the inferior layout of his book which was typical for Dietrich Reimer publications at the time”.

question is addressed in Section 4.2.

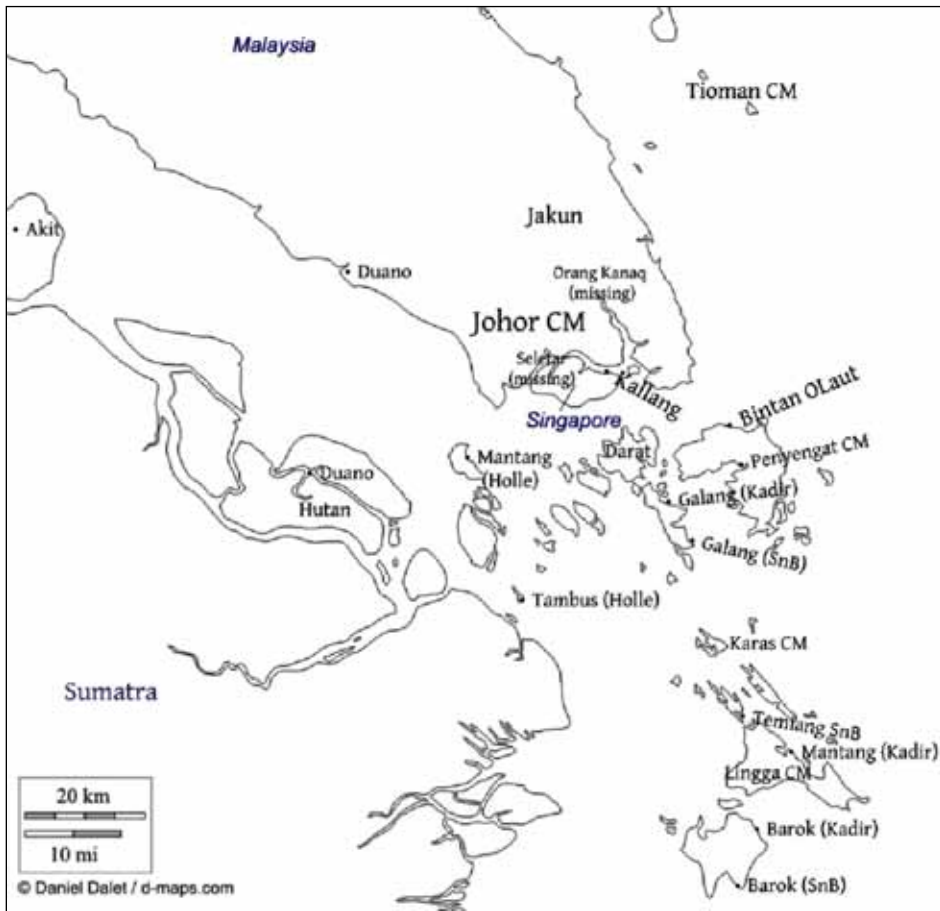
- B. Usman's et al. (1983) *Struktur Bahasa Orang Laut* provides a basic documentation of the speech of "Orang Laut" off the North Sumatran (Sumut) coast. These data are discussed further in Collins (1988). I refer to this documented lect as Sumut OLaut.
- S. Umar et al.'s (1991) *Struktur Bahasa Akit* includes a 214-item wordlist, texts, and sentences of the abovementioned Akit (Raft) Tribe. These are nomadic foragers but not currently sea nomads. The authors report 3,436 Akit people in two villages on Rupid Island,¹⁶ as well as vigorous usage of the vernacular.
- Collins (1995), in his bibliography of Malay dialects on Sumatra, provides basic information regarding previous studies of Sea Tribe lects there. One study mentioned therein (Collins 1995: 190) is S. Syamsiar et al.'s (1986) 79-pages *Struktur Bahasa Sokop*, documenting the speech of 700 Orang Laut in three villages on the island of Rangsang. Although Collins¹⁷ equates their dialect with that of the Duano researched by Kähler (1946a, b), the very limited data I have seen do not support this conclusion.¹⁸
- Orang Seletar (Singapore) and Duano wordlists were published in V. Arnaud et al. (1997). The wordlist database has 900 lexical slots but many of them are probably empty for any given list. My guess is the data are from C. Pelras (see Pelras 2002). I do not yet have access to this resource.
- I conducted extremely brief fieldwork of the partly-sedentarized Sea Tribe on Bintan Island in 2010, gathering a 200-item wordlist and some sociolinguistic information. The wordlist can be found in the appendix, and the group is referred to here as Bintan Orang Laut.

Map 4 shows the Sea Tribe data points in the Riau Archipelago.

¹⁶ Per G. Benjamin (personal communication), Orang Akit also live on Karimun Island in the Riau Archipelago.

¹⁷ Collins (1995: 189) also cites Suwardi's (1993) article as "an analysis of aspects of the phonology and morphology of the dialects of Riau Malay, Talang Mamak and Orang Laut of the Riau-Lingga Archipelago, as well as the dialects spoken by Orang Bonai and Sakai in the Rokan River basin" (translation mine). However, Suwardi's Orang Laut data are from Kadir et al. (1986), and his analysis contributes little if anything to the present study.

¹⁸ Duano (including Kähler's Rangsang dialect) contains highly unique body part terms not reflected in this 1986 book, for example Sokop *kaki* 'leg' versus Rangsang *kəyəmpan*, *perut* 'belly' versus Rangsang *bətəŋ*, *kepale* 'head' versus *kulu*, *idong* 'nose' versus *kələŋə*, and *gigi* 'tooth' versus *ləpə*. Judging by this tiny sample, 'bahasa Sokop' looks like nearly any other Riau Malay lect.



Map 4 Sea Tribe data points, Riau Archipelago.

2.2 ADDITIONAL DATA SOURCES

Additional sources of comparative linguistic data and/or analysis are given in Table 2.

Lect	Source
Proto-Malayic; Standard Indonesian [ind]	Adelaar (1992)
Standard Malay [zsm]	Collins (2008)
Penyengat (Bintan, Riau) Malay [zlm]	S. Dahlan et al. (1990)
Karas (Riau) Malay [zlm]	Dahlan (1989)
Lingga (Riau) Malay [zlm]	Stokhof (1987)
Tioman Island Malay (southeast coast of Peninsular Malaysia) [zlm]	Collins (1985; fieldwork 1982)
Natuna Islands Malay [zlm]	Collins (1998a)

Lect	Source
Johor Malay [zlm]	A. Moain (1999)
Kedah Malay [zlm]	Asmah Haji Omar (2008)
Nineteenth Century Jakun [jak?] ¹⁹	Skeat and Blagden (1906)
Modern Jakun, Northern Pahang [jak]	Seidlitz (2005)
Temuan [tmw]	A. Baer (1999), A. Lee (2004)
Deli Malay [zlm]	E. Rafferty (1983)
Sakai [zlm]	H. Kalipke and M.A. Kalipke (2001)
Minangkabau [min]	Y. Kasim et al. (1987)
Jambi Malay [jax]	Anderbeck (2008)
Kubu (Jambi) [kvb]	A. Maryono et al. (1997)
South Sumatran Malay (for example Ogan, Enim, Besemah [collectively pse], Palembang, Musi, Rawas [collectively mui], Lembak [liw], others)	J. McDowell and Anderbeck (2008)
Belitung Malay [zlm]	S. Napsin et al. (1986)
Haji [hji]	Anderbeck (2007)
Bangka Malay [mfb]	B. Nothofer (1997)
Lom (Bangka) [mfb]	O. Smedal (1987)
West Kalimantan Malayic [zlm, xdy etcetera] and Bidayuhic [scg etcetera]	H. Astar (2002), W. Kurniawati (2002), N. Martis et al. (2002)

Table 2 External data sources (ISO 639-3 code in brackets if available).

2.3 METHODOLOGY

This study employs techniques from historical linguistics and dialectology, utilizing lexical data from a broad variety of Malay dialects to reach conclusions of shared unique lexical items and shared consistent sound changes from Proto-Malayic (PM). Sound changes analysed include reflexes of PM *s, *r and *h, *high vowels, ultimate *open syllables and others; see Section 3.2.

2.3.1 DATA LIMITATIONS

A number of data limitations were encountered in this study:

- Lack of data. The shortest list is 6 items (Temiang – SnB)!
- No data from some known groups like Orang Seletar and Orang Kanaq (former Sea Tribe from around Batam, resettled in Johor; compare M.

¹⁹ It is unclear (as is explained in Section 1.4.3) whether one can draw a straight linguistic line from the Jakun data of Skeat and Blagden (1906) to the Jakun language of today. Are they the same language? That remains to be determined.

Musa (2011: 51)).

- No or very limited data from some geographical areas, for example Natuna, Anambas, many Riau islands, and peninsular Malay varieties.
- Lack of phonetic detail, and varying orthographic standards:
 - most of the thirty-plus samples were from different elicitors;
 - many were produced before standardized transcription systems like the International Phonetic Alphabet (IPA);
 - the vowels are hardest to interpret; there is a frequent lack of clarity between *e* and *a*, for example;
 - different flavours of *r* are also difficult to disambiguate, as are final *k* and *ʔ*.
- Lack of morphological information.
- Lack of comparable lexeme set - some have “tooth”, others don’t, but have “nose”, etcetera.
- Lack of sociolinguistic background.

2.3.2 METHODOLOGICAL LIMITATIONS

In addition to the limitations imposed by the lack of data, some methodological limitations exist also. For example, the summary of innovations discussed below sometimes glosses over important factors. For example, even in varieties where PM **a* is raised after voiced obstruents, different lexemes might be affected in one variety than another. One lect may show raising in *babi* ‘pig’ but not in *balik* ‘return’, while the other lect may show the reverse pattern. Nevertheless, both varieties are classed here as frequently raising PM **a*, obscuring the individual differences. Also, complex conditioning environments or broader phenomena, like Temuan additive initial and final *h* in a variety of *environments (Lee 2004: 4), are classed as similar to simpler phenomena, such as Mantang’s occasional addition of *h* after open vowels.

3 ANALYSIS

In this section I present the detailed results of this study, including a discussion of the classification of Sea Tribes lects, their primary sound changes, and distinctives in the lexicon of the Bintan Sea Tribe.

3.1 DO THE SEA TRIBES SPEAK MALAYIC?

Adelaar (2005: 360), in an update to his published reconstruction of Proto-Malayic (Adelaar 1992), set forth 14 phonological developments from Proto-Malayo-Polynesian which, taken together, define the Malayic subgroup. He adds (1992: 38) a nearly-universal post-PM development, the merger of PMP **e* (schwa) and **a* in final closed syllables. I will list only a few innovations here, enough to demonstrate that all Sea Tribe lects should be classified as Malayic.

- PMP **j* > *d*, for example PMP **qijunʔ* ‘nose’ > PM **hidunʔ*. All Sea Tribes lects share in this development, with only Duano somewhat inconsistent in this regard (see Section 1.4.2).

- PMP **w* > Ø, for example PMP **wahiR* ‘water’ > PM **air*. Same as above.
- PMP **q* > *h*, for example PMP **tuqa* ‘old’ > PM **tuha*. Same as above.
- PMP **e* > *a* / *_C#*, for example PMP **enem* ‘six’ > most Malayic *enam*. Even Duano is consistent in this area.

Unquestionably, the Sea Tribe lects, from north to south, are Malayic.

3.1.1 SEA TRIBE LECTS – MALAY OR MALAYIC?

Tadmor (2002) makes the distinction between “Malay (a language) and Malayic (a group of related languages)”. Since the sixteenth edition, the *Ethnologue* (Lewis 2009) has attempted to reflect this distinction in its classification system. For example, Minangkabau is considered to have derived from Malay (Adelaar 1992) so is classified as “Malayic, Malay”, while Iban is not, so is classified simply as “Malayic”.

This distinction is simple to make in theory and difficult to make in practice. What is called for is a historical argument – if a particular variety is not spoken by people who consider themselves *Melayu*, is it at least historically derived from Malay? The problem is, Malay has been around since at least the seventh century. We might consider a geographical argument – if the homeland of Malayic is Kalimantan, while the homeland of Malay is southeast Sumatra, then perhaps a speech variety outside Kalimantan is most likely to be Malay and not merely Malayic. But there are a couple problems with this: first, that we do not *know* that our ‘homeland’ suppositions are correct; and second, speakers of non-Malay Malayic languages are theoretically as capable of migrating as Malay speakers are. Also, if speakers of a non-Malayic language shift over time to Malay in a process of gradual assimilation (as seems likely to have happened with Duano and Jakun, at least), at what point do we consider their speech Malay, especially if we still see some substratum effects?

For the upcoming seventeenth edition, I recommended to the *Ethnologue* editors that all listed Malayic languages be classified as “Malay” with the following exceptions:²⁰ the six Ibanic lects (Iban [iba], Balau [blg], Mualang [mtd], Remun [lkj], Seberuang [sbx] and Sebuyau [snb]), Kendayan/Selako [knx], Malayic Dayak [xdy], Keninjal [knl], Urak Lawoi’ [urk] and Duano [dup]. Iban and Kendayan seem to be the clearest examples of Malayic-not-Malay, while Malayic Dayak (better termed “Dayak Malayic” and including Keninjal) I regard as only Malayic until proven Malay. Perhaps the same agnostic stand can be taken for Urak Lawoi’ and Duano, two fairly divergent varieties. Benjamin (2001: 101) says as much for Duano; Steinhauer (2008) does not differentiate between Malay and Malayic in his discussion of Urak Lawoi’. I close this section with the dampening words of Adelaar (1992: vi), which seem as valid today as twenty years ago: “The question of the internal classification of the Malayic subgroup, and hence the question of the difference between ‘Malay’ and ‘Malayic’ [...] remains unanswered, and it is doubtful whether

²⁰ The codes in brackets are the ISO 639-3 language codes.

sound solutions will be obtained from the comparative method alone".²¹

3.2 MALAYIC SEA TRIBE SOUND CHANGES

The chief post-Proto-Malayic innovations discovered and compared on the charts below are interpreted here.

- Mutation of PM **a* in closed penultimate (listed on the chart as **a > e* PU) and ultimate (**a > e* ULT) syllables. This innovation is nearly always connected in some way with voiced onset (whether syllable or word), specifically the voiced obstruent series *b, d, j*, and *g*, for example Akit *jalan* 'walk' < **jalan* but no raising in *kanan* 'right'. This phenomenon is discussed in greater detail in, among others, Collins (1998a: 548), McDowell and Anderbeck (2008), and most recently and in greatest detail, in T. Mckinnon (2012).
- Lowering of PM **high* vowels in penultimate position (HV↓ PU) and in ultimate position (HV↓ ULT). An example of this is Barok 1986 *esol* 'boil (noun)' < PM **bisul*. This innovation, which invariably only affects some lexemes and not others, is discussed in detail in Adelaar (1992: 10, 42) and elsewhere.
- Mutation of final open **a*, for example Tambus *matə* 'eye' < PM **mata*. This innovation, its manifestations and possible origin, is discussed in Tadmor (2003). In Malayic lects outside Kalimantan this mutation at present is nearly universal, but which vowel to which it changes is quite variable between lects.
- Closing of final open syllables (**-V* closed) with glottal stop (presumably represented as *k* in some lists) or with *h*. For example Sekak *duaʔ* 'two' < PM **dua*.
- One lexical innovation is included in the charts below, namely whether the lect has the triplet *kian* '(come) here', *kiun* '(go) there' and *kiuh* 'far off'.
- Reflexes of final diphthongs **ay* and **aw*, like *sunje* 'river' < **sunjay*, or *piso* 'knife' < **pisaw*.
- Reflexes of PM **h* in initial, medial and final position (usually *h* or Ø) are presented by position.
- Reflexes of PM **r* are presented similarly to **h*, though I additionally looked at reflexes in medial clusters like in *bersih* 'clean'. Given that many wordlists probably did not distinguish orthographically between a voiceless uvular fricative *χ* and an *h*, I marked all devoiced reflexes as special.
- Medial nasal-voiced-obstruent consonant clusters may show a reduction or elimination of the obstruent component (NC > N), like *tiŋ^si* or *tiŋi* 'high' < *tiŋgi*.
- Medial nasal-voiceless-obstruent consonant clusters may show a reduction or elimination of the nasal component (NC > C), like *gutur* 'thunder' < *guntur*.
- "Preplosion" on the chart is a cover term for basically any sort of variable

²¹ Perhaps the most conservative classificatory approach, therefore, would simply be to dissolve the *Malay* node and keep everything in the *Malayic* group.

plosion or nasalisation word-finally such as discussed in R.A. Blust (1997), Seidlitz (2005), Anderbeck (2008), and Steinhauer (2008). Examples of this can range from final *nasals to stops, as in *urak* 'person' < PM **uraŋ*, to pre- or post-plosion (*urakŋ*, *uraŋk*), to pre- or post-nasalisation of stops (for example *urant*, *uratn* 'vein' < PM **urat*).

- A few lects, particularly Barok 1986, exhibit frequent loss of the initial consonant (*C-loss) or even initial syllable in a word, for example *tes* 'calf (of leg)' < PM **bətis*.
- The final two innovations tracked, both in final consonants, are *-*iŋ* > *in*, for example Sumut OLaut *kəsin* 'dry' < **kəriŋ*, and *-*s* > *ih*, for example *tikuuh* 'rat' < **tikus*.

Here are a few more notes on the chart:

- The chart is broken up into four pieces to fit on the page, and moves roughly south-to-north.
- Canonical Malay lects are suffixed with CM.
- Asterisked cells are those which require more explanation. These are noted below the chart.
- I did not end up finding all the tracked innovations ultimately interesting for subgrouping and left them all uncoloured on the chart. Interesting (to me) innovations are shaded. Since presence of initial **h* is very rare in modern Malayic lects, I also shaded *retention* of initial **h*.

features	Belitung CM 1986	Sekak 1881	Sekak 1981*	Barok 1906	Barok 1986*	Temiang 1906	Tambus 1905	Galang 1906	Galang 1986
*a > e PU	no	some	strong	strong	strong	?	no	some	strong
*a > e ULT	no	?	strong	strong	strong	?	no	some	strong
HV → PU	no	no	rare	some	some	?	common	?	some
HV → ULT	common	rare*	rare	most	most	?	common	common	common
final *a	ə	a	a	e	ə	a	ə	a?	ə
*-V closed	some	20%+h	strong	20%	some	?	no	strong	strong
diphthongs	aw, ay	ai>e	aw>o(?)	ay>e	aw, ay	aw	ay, ey	?	aw, ow
kian/kiun?	no	?	yes?	yes	yes	?	yes	yes	?
*h INIT	Ø	Ø	Ø	h	Ø	?	h	?	Ø
*h MED	Ø	?	Ø	Ø, V1hV1	Ø, V1hV1	?	h, Ø	?	Ø
*h FIN	Ø	h+Ø	Ø	h	h	h	h	?	h
*r INIT	r	r	r	χ	h	?	χ	r	Ø
*r -V-	r	r	r	χ	h	?	χ	r	Ø
*r CLUST	r	Ø?	r	χ	Ø	?	χ	?	?
*r FIN	r	r	r/?	χ	h	Ø	Ø	r/Ø	Ø
NC > N	no	yes	yes	no	no	no	NC	some	yes
NC > C	no	no	no	no	no	?	no	no	no
prepllosion	no	no	no	no	no	?	no	no	no
*C-loss	no	no	no	no	strong	no?	no	no	no
*-itj > in	no	no	no	no	no	?	no	?	no
*-s > ih	no	no	no	no	no	?	no	?	no

Table 3 Comparison of innovations, part 1 (Belitung and Sea Tribes).

features	Mantang 1905	Mantang 1986	Kallang 1900	Bintan OLaut 2010	Karas CM 1989	Lingga CM 1895	Penyengat CM 1990	OLaut 1930's
*a > e PU	some	common	?	common	no	no	no	no
*a > e ULT	rare	some	?	some	common	no	no	no
HV↓ PU	common	common	some	common	?	some	common	common
HV↓ ULT	common	most	some	common	?	some	some	some*
final *a	a, ə	ə	a	u	ək	a	ə	a(h)
-V closed	no	rare k or h	no	pro/dem	strong	no	no	no
diphthongs	aw, ay	aw	aw>o	aw	aw, ay	aw, ay	aw, ay	aw, ay
kian/kiun?	?	?	yes	yes	?	no	?	?
*h INIT	h	Ø	?	Ø	Ø	Ø, h	Ø	Ø
*h MED	h?	h, some Ø	h	Ø, V1hV1	?	h	Ø, V1hV1	Ø, V1hV1
*h FIN	h, some Ø	h	h	h	ək?	h	h	h
*r INIT	h, some r	ɤ	?	Ø	ɤ/Ø	ɤ	ɤ/Ø	r
*r -V-	h, some r	ɤ	h	Ø	?	ɤ	ɤ	r
*r CLUST	?	Ø	?	Ø	?	ɤ	ɤ	r
*r FIN	h, some r	Ø	?	h	?	Ø/ɤ	Ø	h
NC > N	some	no	?	no	no	no	no	some
NC > C	no	no	no	no	no	no	no	no
preplosion	rare	no	no	no	yes	no	no	no
*C- loss	no	some	no	no	no	no	no	no
*-ij > in	no	no	?	no	no	no	no	no

Table 4 Comparison of innovations, part 2 (Sea Tribes and canonical Malay).

features	ODarat Batam 1905	ODarat Batam 1930's	Akit 1930's	Akit 1991*	Duano*	OHutan 1930's	Tioman CM 1982	Natuna CM 1996
a > e PU	?	strong	strong	common	no	no	strong	strong
*a > e ULT	some	common	common	some	rare	rare	before *r	strong*
HV → PU	common	common	common	common	no	some	strong	rare
HV → ULT	common	common	some	common	rare	rare	common	some
final *a	ə/a	aʔ, ɔʔ	aʔ	e	u	a, ɔ	i	ə
*-V closed	no	strong	strong	pro/dem	no	no	some	strong
diphthongs	?	ow, ae	ow, ay/ae	aw, ay	aw, i	aw, ay	aw, ay	ay
kian/kiun?	?	yes	yes	?	no*	?	?	?
*h INIT	Ø, h	Ø	Ø	Ø	Ø	Ø	Ø	Ø
*h MED	?	Ø	Ø	Ø	Ø, VlhV1	Ø	VlhV1	h/Ø
*h FIN	H	h	h	h	h/?	h	h	h
*r INIT	R	ɣ	x	h/Ø	ɣ	r	ɣ	ɣ
*r -V-	R	ɣ	x	h/x	ɣ	r	ɣ	ɣ
*r CLUST	?	ɣ	Ø	Ø	V:	r	Ø	?
*r FIN	?	ə	x	χ	Ø	r	əØ	V:
NC > N	no	some	yes	yes	no	some	yes*	no
NC > C	no	rare	no	no	no	no	no	no
preplosion	no	yes	no	no	yes	common	no	no
*C-loss	no	no	no	no	no	no	no	no
*-ij > in	no	no	no	no	no	common	no	no
*-s > ih	?	no	no	no	no	no	no	no

Table 5 Comparison of innovations, part 3 (Riau tribal, Duano and eastern Riau Islands).

features	Jakun 2003	Jakun 1906	Temuan*	Deli CM 1983	Kedah CM 2008	Sumut OLaut 1980*	Kedah OLaut 1814	Kappir 1906	Urak Lawoi 1988*
*a > e PU	no	rare	no	no	no	no	no	?	no
*a > e ULT	no	no	no	no	no	no	no	?	no
HV↓ PU	some	rare	rare	common	some	some	some	?	no
HV↓ ULT	some	low	some	common	some	some	low rare	?	common
final *a	a	a	a	e	a	a	a	a	a
-V closed	no; 2S-D yes	no	h/?	rare	no?	no	no	h?	no
diphthongs	aw	aw, ay	aw, ay	aw, ay	aw, ay	aw, ay	lower	aw	aw
kian/kiun?	yes*	yes	no	?	?	?	?	?	no?
*h INIT	h	h	h	h	∅	∅	h	?	h
*h MED	h	h	h	h	∅	∅, V1hV1	h	h	h
*h FIN	h	h	h	h	∅	h	h	h	h
*r INIT	r	r	∅	r	ɛ	ɛ	?	?	r
*r -V-	r	r	uɭ	r	ɛ	ɛ	?	r	r
*r CLUST	∅	r	∅	r	r	ɛ	?	r	r
*r FIN	r	r	uɭ	r	ɛ	k	wi/ɣu	∅	l
NC > N	no	no	yes	no	yes*	some	no	?	yes
NC > C	no	no	no	no	no	no	yes	no	yes
preplosion	yes	yes	yes	no	no	no	no	?	yes
*C-loss	no	no	no	no	no	no	no	no	no
*-itj > in	no	no	yes	no	yes	yes	no	?	no
*-s > ih	no	no	no	no	yes	yes	?	?	yes

Table 6 Comparison of innovations, part 4 (Peninsular and northern).

Notes

- Sekak 1881: Final high vowels before **h* are frequently lowered with subsequent elision of *h*, for example *taro* ‘place (v.)’ < *taruh*, *ole* ‘by, by means of’ < *ulih*. This innovation seems to have been lost/reversed in 1981 Sekak.
- Another Sekak 1981 innovation not listed here: Subsequent to undergoing the raising of ***a*, CVVC lexemes experienced apocope, with the apparent path being *jahit* > *jait* > *jit*, *laut* > *laut* > *lut*, etcetera.
- Barok 1906 exhibits epenthesis of palatal glide in *muxtendianj* ‘kick’ (< *təndanj*), *haməŋgianj* ‘roast’ (< *paŋganj* I assume) but not *sirampanj* ‘fish with a trident’ – perhaps this occurs only after nasal-voiced stop (NvC) clusters? No further examples are in the list.
- Barok 1986 also differs from Barok 1906 in reference to the bullet above. Specifically, the two extant examples of final *-*anj* following nasal-voiced-stop clusters (*umbanj* ‘beetle’ < *kumbanj*, and *injanj* ‘waist’ < *pinjanj*) do not exhibit any epenthesis of palatal glide in the vowel.
- Mantang 1905 frequently exhibits ‘stretching’ of the penult for example *tahali* ‘rope’ < **tali*, *kəhuniŋ* ‘yellow’ < **kuniŋ*, *pəhutih* ‘white’ < **putih*. (Not seen in Mantang 1986.)
- Bintan Orang Laut innovations not shown on the chart include merger of final **k* and **ʔ* to *ʔ* (a common Malay innovation). Glottal stop is appended to most pronouns, demonstratives, and interrogatives.
- Karas Malay, spoken on Karas Besar island in Galang sub-district 1989: many stops and nasals reportedly go to glottal stop, *k* or zero, like *pəʔut* ‘stomach’ < **pərut*, *tiʔəʔ* ‘three’ < ***tiga*.
- Johor Malay (not on chart) is well-known to exhibit loss of final **r* and frequent lowering of *high vowels (Moain 1999: 648).
- Penyengat 1990: Elision of final **r* is accompanied by high vowel lowering. Per. Dahlan et al. (1990), Senayang Malay, spoken on Senayang island in Riau, is similar to Penyengat Malay but does not exhibit the loss of final **r*, for example *bebək* ‘lip’ < **bibir*, *aək* ‘water’ < **air*.
- All Kähler (1960) lists are transcribed as having voiced final consonants. This is suspicious given that no other linguist has noted the same, and his lists span territory from offshore mainland Riau to Batam and Bintan islands.
- Kähler OLaut frequently lowers final *open high vowels (*gige* ‘tooth’ < **gigi*), a rarely-seen innovation.
- Kähler ODarat (also to some extent in Akit and Hutan) maintains the distinction between final **k* and **ʔ*.
- Akit 1991 also exhibits frequent loss of final **l* (not seen in Kähler Akit).
- Kähler’s notation is ambiguous, but the Hutan data seem to optionally debuccalize final stops (for example, *kuli(d)* ‘skin’ < *kulit*).
- Tioman Malay 1982: NC > N occurs only when the consonant cluster in question is in the penult, for example *təmagə* ‘copper’ < ***təmbaga*.
- Natuna Malay 1998: Conclusions are from the Sedanau dialect. The Serasan

- dialect seems more conservative. The raising of **a*, mostly after voiced stops, usually results in [o] rather than [e] (Collins 1998a: 548).
- Temuan innovations not shown in the chart: Not only is *h* frequently appended to the ends of words, but 1) it is added to the beginning of words reconstructed with an initial vowel (for example, *hayam* < PM **ajam* ‘weave’); and 2) final oral consonants are frequently but not universally debuccalized to *h* (for example, *awah* ‘early’ < *awal*, *isah* ‘gills’ < *isanj*).
 - Jakun 2003: Of the fourteen Jakun 2003 wordlists, eleven exhibit open *a*, one exhibits *a*, and two exhibit consistent closure with glottal stop. Because the wordlists are short it is unclear whether Jakun 2003 has anything like *kian*, *kiun* and *kioh* but note *can* ‘near’ and *cun* ‘far’.
 - Duano also has something like ***kia*/*kəna*/*kəsut* in place of the *kian*/*kiun*/*kiuh* troika.
 - Sumut OLaut 1980 additionally innovates final **l* > *i*, for example *tumpui* ‘dull’ < *tumpul*.
 - Kedah Malay final **l* > *i* in the same pattern as Sumut OLaut. Some Kedah Malay subdialects elide voiced stops after nasals, while others do not.
 - Urak Lawoi’: See Section 1.4.1 for additional Urak Lawoi’ innovations.

3.3 LEXICAL ANALYSIS

Due to the exigencies of time and space, I will centre my lexical analysis on the 2010 Bintan Sea Tribe wordlist (see Appendix).

One of the most noticeable observations from looking at the approximately 230 words in the list is how standard is the vocabulary, and favouring mainstream Peninsular Malay over Indonesian. Even though various sound changes have acted upon the words, the lexemes themselves are usually the same as mainstream Peninsular (like Johor) Malay. For example, *kii* ‘left’ is obviously cognate to standard Malay (SM) *kiri*, albeit with the loss of *r*. Likewise, *bexʔ* ‘when’ is cognate to standard Malay *bila* versus Indonesian *kapan* and *pokoʔ* ‘tree’ versus Indonesian *pohon*. But not always is this the case; compare *bəgʁymʁnu* ‘how’ (Indonesian *bagaimana*) with Johor Malay *macam manə*, and *etəŋ* ‘count’ (Indonesian *hitung*) vis-à-vis Johor Malay *bilang*. Only 10% of the list shows up as different from SM, and a good half of that 10% consists of items which also appear in SM, albeit as a variant form (for example *pərah* ‘squeeze’) or with slightly different semantics, for example *ŋet* ‘mosquito’ (mainstream Malay *rengit* ‘gnat’). Compare this with Jambi Ilir, with approximately 80% shared SM cognates, Jambi Ulu at approximately 74% (Anderbeck 2008: 77) or the Dayak Malayic of the Melawi River basin at approximately 70% (Anderbeck and Sellato In progress).

Another way of looking at this is how few characteristically Sumatran Malay lexical items appear in this list. Some examples: *dii* ‘stand’ instead of *təgəʔ*, *bə-kəju* ‘work’ instead of *bə-gawe*, *sayap* ‘wing’ versus *kəpəʔ*, *panas* ‘hot’ versus (*h*)*aŋat*, *meah* ‘red’ versus *abaŋ*, *pokoʔ* ‘tree’ versus *bataŋ*. But common Sumatran Malay words sometimes do appear, like Bintan *kəuʔat* ‘cut, hack’ (Jambi Malay *kərat*), *pəpaʔ* ‘chew’ (South Sumatran Malay same), *ekʔ* ‘this’

(Musi, Jambi Malay, also found in Jakun).

The aberrant words in the list are so few as to allow mentioning all of them. First, two forms of which I have no record appearing elsewhere: *pəŋa?* ‘cloud’, *usa?* ‘don’t’ (although for the latter compare Indonesian *tidak usah* ‘not necessary’). The list also has some rare but not completely unknown items:

- *koyo?*, the nearly universal word for ‘dog’ among the Riau Sea Tribes, can also be found in other tribal lects of the area, the Holle Lingga (canonical Malay) list, in southwestern Sumatra (Musi, Enim, Lembak, Haji), in Sekak, in West Kalimantan Bidayuhic, and probably in northwestern Kalimantan Malayic *ukuy*. Although Seidlitz (2005) did not elicit ‘dog’ in modern Jakun, the same lexeme can be found in four of SnB’s nineteenth century ‘Jakun’ lists.
- *ua?*, the most common term for ‘father’ among the Riau Sea Tribes, is also seen in a dialect of Pontianak Malay, Duano, Jakun, and Bangka/Lom, and possibly cognate with the dialectal Ogan form *uba?*. It is likely cognate with SM *uak* ‘elder (in family relationships); term of address for father or mother’s grandfather’.
- *tikam* ‘throw’ exhibits an evident semantic shift from ‘stab’ shared with Duano, Bangka/Lom and a few West Kalimantan Malayic isolects. It is possible this semantic shift occurred among maritime populations as the main thing they would be ‘stabbing’ would be sea creatures using a thrown trident.
- *mə-leleh* ‘flow’ also appears in Duano, and as a rare form in Besemah, Jakun, Kalimantan Malayic and Kalimantan Bidayuhic. Also compare Urak Lawoi’ *nileh* ‘to flow out’.
- *kote?* ‘tail’: Besides Sekak and Riau Sea Tribe lects, I have only otherwise found this in R.J. Wilkinson’s (1959) Malay-English dictionary.
- *kiun* ‘go over there’ appears in a dialect of Jambi Kubu as ‘that’ and in a dialect of Minangkabau as ‘there’. See also the discussion in Section 1.5.
- Bintan *miku* ‘you (singular)’ seems rare, shared with Perak Malay (Kin 1999: 46), Orang Darat Batam (Holle) and Kubu (and one data point in Jambi Ulu close to Kubu as ‘you (plural)’).
- *saŋap* ‘yawn’ is shared with Kähler Orang Hutan, Darat, and Akit (latter as *səlaŋab*), and Duano and Jakun.²²

Although four rare items are shared with Duano, as I was eliciting the Bintan list I asked the informant if he recognized a number of other unique Duano forms and in all cases except the above the answer was negative. Since Duano has a number of unique sound changes not shared by Bintan, I consider the possibility of a significant relationship between the two lects as unlikely, and the shared vocabulary as due to contact.

A few rare lexical connections can be seen variously with Bangka/Lom, Jakun, Kubu and Besemah, but it is hard to know what if anything to conclude

²² It is unclear whether Barok *uŋap* is closely related to *saŋap* or is a metathesis of (*mə*) *ŋuap*.

from this.

Here are a number of brief notes on Urak Lawoi' (UL) lexical items:

- UL *apok* 'father' is probably cognate to Kedah Orang Laut *apung* identical. Another shared item is *nanak* 'child'.
- UL also has a number of commonalities with the SnB Kappir list: *nibini* 'woman' (< *bini-bini*); Kappir *ma* 'mother' and *pa* 'father' resemble UL *ma?* and *pa?*; and Kappir *calaki* 'man' resembles UL *kilaki* (< laki-laki).
- UL *lulu* and modern Jakun *lolok* 'meat'. These are possibly reduplications of *lauk* 'solid food to be eaten with rice'.
- UL *brulak* and modern Jakun *bəlolan* 'skin' (Malay 'pelt; hide').
- UL *sudə?* and modern Jakun *cedə?* 'dig'.
- UL *gabə?* 'rotten', Duano *yəpə?*, *lapə?*, Sekadau *əpəpə* all seem related, if distantly.
- Interesting set 'throw': Urak Lawoi' *məlatik*, Rawas *luti*, Musi *letok*, Muko-Muko *latian*, Bangka *məlidə?*, Jakun *lətə*.

Here are a few notes on Duano lexical items:

- See above for mention of 'father', 'flow', 'throw', 'yawn' and 'rotten'.
- *bisa* 'sick, painful'; SM 'venom' is also found in Jakun, Orang Darat Batam, Akit (Kähler) and Tambus.
- Duano *kukut* 'hand' is shared with SnB 'Jakun' *kokot*.

4 CONCLUSIONS

4.1 DIACHRONIC FINDINGS

4.1.1 ORIGINS

Is there evidence of a non-Malayic substratum in Sea Tribes lects? Besides the doubtful lexical examples discussed in Section 1.5, nothing of note has been uncovered for Riau Sea Tribe lects.²³ One contrary note, a seeming archaism, is found in SnB's Barok list: *murtendian* 'kick', composed of Malay *təndan* 'kick' plus a **mar* prefix found in Old Malay but rarely elsewhere (Adelaar 2008: 13; Benjamin 2009: 307). If this (lone) prefix is truly a retention from ancient forms of Malayic, the case would be strengthened that at least some Sea Tribe lects represent a continuation of Proto-Malayic.

I do not have time or space here to explore the same for the more aberrant Urak Lawoi' or Sekak except to say that an exploration of possible substratum in Sekak is undertaken in the manuscript mentioned above, while the question of Duano origins is touched upon in Section 1.4.2.

4.1.2 BETWEEN THE CENTURIES

We have the benefit of a number of putative pairs of wordlists which can be examined for sound changes and other trends:

²³ One word in the Orang Darat Batam (Holle) list may have an Aslian origin: *kot/kat* 'no', compare Besis *ηot* identical. See also the discussion of Duano in Section 1.4.2.

- Barok 1906 and Barok 1986: In spite of the identical names given to these two lists, they seem to represent separate subvarieties. I make this judgement on the basis of significant differences in their lexicon and sound changes. The same is the case for the two Galang lists (1906 and 1986), and the two Mantang lists (1905 and 1986).
- Sekak 1881 and 1981: As discussed in Section 1.4.4, the Sekak of 1981 is much more mainstream lexically than its 1881 counterpart, although the two varieties are clearly akin.
- The Duano lists from 1884, 1906, 1939, and 2005 show clear affinity. It is unclear whether their inevitable lexical and phonological differences should be attributed to change over time or to dialectal variation which is relatively significant in modern-day Duano.
- In spite of the small and seemingly concentrated population of the (Rupat) Akit group, the 1991 data show significant discontinuities (as well as some continuities) with the data from the 1930's. Both show complete devoicing of PM **r* and reduction of medial nasal-voiced stop clusters to the nasal component. Both show raising of **a* after voiced plosives, although the proportion of affected lexemes is much smaller in the later sample, for example 1930's Akit *dōtaŋ* 'come' versus 1991 *dataŋ*. Initial **r* is often lost in 1991 but not in the earlier sample. 1930's Akit evinces the closure of basically all *open final syllables with glottal stop, which is much rarer in 1991 Akit. Instead, we see mutation of final **a* in the newer sample, with an occasional excrescent *h* (for example, *bətina^h* 'female' < earlier *bətina*), neither of which is seen in 1930's Akit. Neither occasional metathesis of **r* in 1991 Akit (*rusa* 'deer' > *usəχ*, *səribu* 'thousand' > *sibuχ*, *bərsih* 'clean' > *bəsīχ*) nor loss of final **l* (*jəŋkal* 'hand span' > *jeŋge*, *tun^hgal* 'one and only' > *tun^he*) are seen in older Akit. Lexically, very few non-mainstream items in 1930's Akit appear in the 1991 list (*naŋuy* 'pig' is the sole exception) or the reverse. If not for the likely historical continuity between the two samples, I would conclude that these are two distinct but geographically proximate dialects.
- The Batam Orang Darat samples are separated by about three decades; Holle Orang Darat in 1905 and the Kähler list in the 1930's. The two lists show a perplexing diversity similar to the Akit lists. Because of the shortness of the Holle list, some questions about potential innovations remain unanswered. The commonalities: some **a* raising, lowering of *high vowels and sporadic mutation of final open **a* in approximately equal measure. The differences: *open final syllables remain open in Holle, but are regularly closed with glottal stop in Kähler. Significantly, Kähler regularly transcribed the final nasals as preploded (for example, *o^hyaŋən* 'person' < **uraŋ*) while no preplosion is transcribed in the Holle list. Equally significantly, Holle reflexes of penultimate **a*, as far as can be discerned, usually remained ə (transcribed as *e*), while in Kähler most go to *o*: compare Holle *kera* 'monkey' (SM *kəra*) to Kähler *koyag*. In terms of non-mainstream lexical items, a couple of shared rarities (*kot* 'no' and

bələŋkok ‘crooked’) seem to be outweighed by differences (for example *oroi* ‘walk’ versus *jalan*, *molek* ‘good’ versus *böig*, *mika* and *ape* ‘you (singular)’ versus *dikou*, *matei apa apa* ‘why’ versus *apa kona?*).²⁴ My conclusion (or confusion) with these two lists is the same as for the Akit lists.

We can see a very interesting progression of certain “cosmopolitan” innovations over the century, namely mutation of final open **a*, and elision of initial **h*. Final open **a* mutation, for example *matə* ‘eye’ < **mata*, is nowadays nearly universal in canonical Malay lects. This is not necessarily true of non-canonical Malay. For example, one of the salient differences between Jakun and its *Melayu* neighbours is whether final **a* is raised. Similarly in Jambi, *Melayu* **a* has universally rounded to *o*, while in Kubu, rounding is variable. Among the Riau lists, unmutated final open **a* is preserved in Temiang 1906, Galang 1906, Kallang 1906 and even the canonical Lingga 1895, and variably preserved in Mantang 1905 and Orang Darat 1905. The new Riau lists, both canonical and non-canonical, consistently reflect *a*.

The picture is rather the same with initial **h*. Barok 1906, Tambus 1905 and Mantang 1905 retain **h*, and in Orang Darat 1905 and canonical Lingga 1895 retention of initial **h* is variable, while in modern Riau lects (or even those documented by Kähler in the 1930’s) the phoneme has completely disappeared from this position. The spread of these two innovations in Riau, the seeming epicentre for many Malayic innovations, is thus given a valuable chronology. Both of these innovations are quite new, in terms of the 1500-plus-year history of Malay in the region.

Another cosmopolitan innovation is the lowering of **high* vowels (Anderbeck 2008: 31–33). It is noteworthy that, unlike the innovations affecting **-a* and **h*, no trend toward greater lowering can be seen in this time period. We can perhaps conclude that this innovation, at least in Riau, is the result of an older process that had already completed by the previous century.

The raising of **a* after voiced obstruents also shows a noticeable increase over the century in Sekak, Galang and Mantang. This is puzzling, for if one looks at the distribution of this innovation in the Malay World, it is most frequently found in peripheral and tribal areas. If this is indeed a relic innovation, why does it show an increase over time in the Riau Islands?

4.2 SYNCHRONIC FINDINGS

Karas Malay seems to be a real outlier whether compared to its fellow canonical Riau Malay lects, or to Riau Sea Tribe lects, but data are lacking.

The non-canonical Malay groups in Riau (and Johor?), Duano excepted, are extremely mainstream lexically, more mainstream than canonical Malay lects in Bangka, Jambi, Kalimantan, et cetera. However, they are phonologically much more divergent than canonical Malay varieties in Riau (Karas Malay excepted). Are they more phonologically divergent than Malayic lects

²⁴ Of course, any or all of these differences might be solely attributable to synonymy rather than complete lack of the non-mainstream lexeme in question.

elsewhere? It is difficult to tell given the lack of phonetic detail in the data, and difficult to define 'divergence' as well. Certainly they are more conservative than the most divergent lects like Kerinci, Duano and Urak Lawoi'.

The most significant finding of this study is that Riau Sea Tribe lects seem to basically cohere as a dialect group (Tambus and Kähler's Orang Laut excepted), with the following innovations:

- **a* raising after voiced obstruents
- frequent *-*V* closing
- (lexically mainstream)
- *kian*/*kiun*/*kiuh*
- (less universal) NC > N
- **r* devoicing (Mantang 1986 evidently an exception)
- loss of initial and final **r* (shared with Riau canonical Malay)
- final open **a* raised to ə/ə̃/ĩ (shared with Riau canonical Malay)

The fact that this bundle of innovations exists, largely independent from canonical Riau Malay, seems to provide evidence of substantial social separation or at least diglossic differentiation between canonical Malays and tribal groups. The latter (differentiation) may be more likely, given the fact that Galang, historically the group most intimately involved with Riau Malay leadership (Chou 2003: 18), evinces some of the strongest adherence to these innovations.

What about other groups in the area? Akit, a tribal but not (currently) a Sea Tribe, fits right in to this group linguistically. The presumably now-extinct Batam Orang Darat (1905 and 1930's) shows significant similarity to the Sea Tribe cluster, while Kähler's Orang Hutan does not fit that well, nor do other tribal groups, notably Duano, Jakun, and Temuan. Another outlier, problematic for my thesis, is Kähler's Orang Laut dataset. It is nearly as phonologically conservative as Lingga canonical Malay (1895), with only some reduction of medial voiced consonant clusters and devoicing of final **r* to show that it belongs in the area. I do not have a good explanation for this pattern.

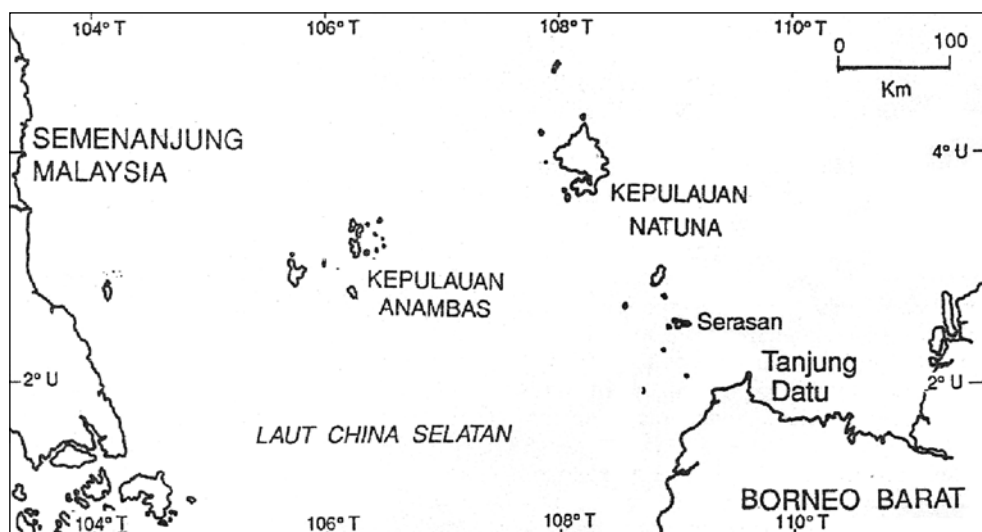
Sekak (old and new) linguistically mirrors its geographical position in relation to Riau: obviously related but not intimately. Its lexicon has many disjunctures with Riau, as does its phonology, with minimal vowel lowering, no devoicing of **r* and reduction of final *diphthongs.

How does Bangka Malay/Lom relate to these? The short answer is that Bangka dialects are very different from the Riau Sea Tribe cluster as well as from Sekak, lexically and phonologically. There are some similarities; three Bangka subdialects show NC > N: Lom, Gunung Muda, and Gadung. Penultimate raising of **a* after voiced obstruents can also be found frequently in Pelaik, Kacung, Kacung 2, Gunung Muda, Perlang, and Pakuk, and less frequently in Arung Dalem and Lom Tuatunu. Note that the locations with **a* raising overlap only slightly with those displaying the NC > N innovation; Gunung Muda alone shows both innovations.²⁵

²⁵ Intriguingly, the Upstream Kutai Malay (Kota Bangun) dialect evinces both NC > N and raising of **a* after voiced obstruents (Collins 1991: 9). However, because of other significant

What about Jambi Suku Anak Dalam/Kubu (another tribal group)? On the affirmative side, **a* raising is frequent. We see *kiun* 'there' in one dialect (and also in a couple Minangkabau dialects). On the negative side, lexically Kubu is not very mainstream, no NC > N is seen, it is conservative about **h*, even initially, and it does not seem to have loss of initial **r* while final **r* is mixed. I conclude that connections are not that tight between Kubu and the Riau Sea Tribe cluster.

The second major finding is the similarity between Riau Sea Tribes and the canonical Malay lects spoken on the Natunas and Tioman (Map 5); both share **a* raising, **-V* closing and loss of final **r* with Riau Sea Tribes. We can go a few different directions with these observations:



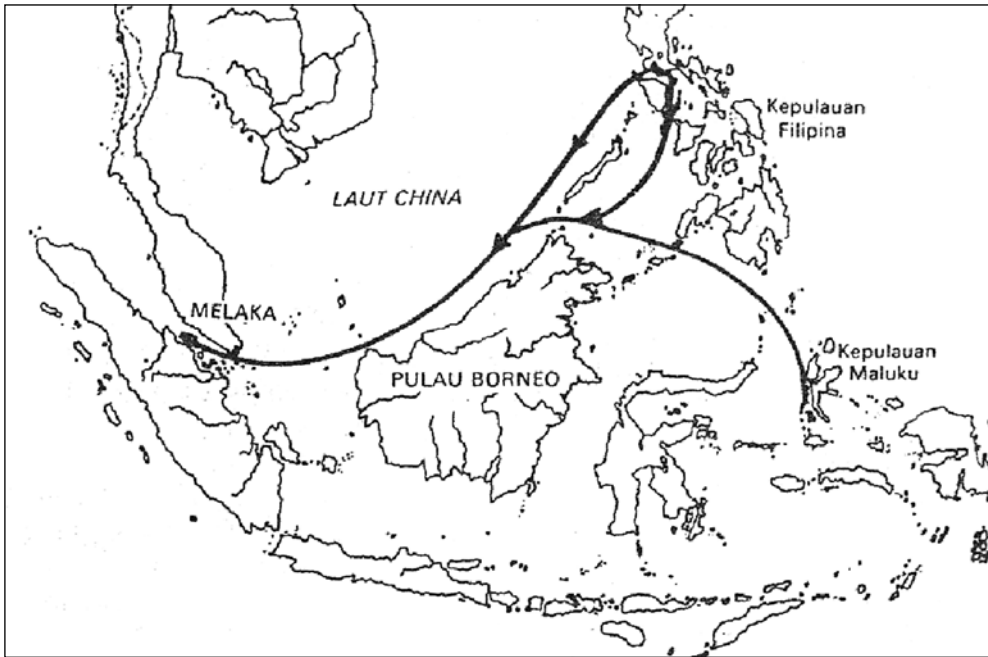
Map 5 Pulau Tujuh including Anambas and Natuna Islands (Collins 1998a: 540).

- Sopher (1977: 114) documents the extensive Sea Tribe population on the Natuna islands which by the end of the nineteenth century was nearly fully assimilated with the Malays. We could therefore take the sociolinguistically-doubtful position that Riau Sea Tribes influenced the canonical Malay there.
- We could posit the origin of the Riau Sea Tribes from the Natuna islands and environs. This seems unlikely from a geographical perspective – migration from the periphery to the centre.
- The most likely scenario would be to posit an earlier Riau Islands Malay including both sedentary Malays and Sea Tribes, a form of speech which was later levelled out via contact to a more unmarked form in most canonical Malay groups except those on the periphery (Natuna, Tioman).²⁶

innovations, most saliently the merger of penultimate PM **ə* and **a*, Collins classifies Kutai Malay with other eastern Borneo Malay lects.

²⁶ Given that initial **r* has not been lost in Natuna or Tioman, while it has been lost virtually everywhere in the core Riau Archipelago, signifies that loss of initial **r* may be a more recent innovation.

Map 6 shows ancient trade routes which would have encouraged close relationships within the span of the archipelago.



Map 6 Ancient trade routes from Maluku and Philippines to Malacca (Collins 1986: 896).

Moving north, Sumut OLaut, Kedah Malay, Kedah OLaut (1814), Kappir, and Urak Lawoi' seem to be part of the same loose dialect network (while maintaining their own distinctives):

- palatalization of *-s (Kedah OLaut and Kappir unknown);
- *-ij > in (only Kedah Malay and Sumut OLaut; also Kähler OHutan);
- NC > N (Urak Lawoi', some subdialects of Kedah Malay, very limited in Sumut OLaut but not Kedah OLaut, and unknown in Kappir);
- NC > C (Urak Lawoi' and Kedah OLaut);
- loss of final *r (Kedah OLaut and Kappir);
- lexical connections between UL and Trang/Kappir.

Despite the scanty evidence, it seems Kedah OLaut was not the 200 year-old progenitor of Urak Lawoi'. Although they seem to have shared the same island, if early Urak Lawoi' had lost final *r, this phoneme could not have later reappeared then lateralized into *l*.

As Collins (1988) argues, it would *not* be appropriate to make the leap from noting these broad similarities between Sumut OLaut, Kedah OLaut, Kappir, and Urak Lawoi' to proclaiming the existence of "Northern Sea Tribes Malay". Instead, these lects (past and present) are embedded within, not separate from, the larger network of Malayic lects.

4.3 DIRECTIONS FOR FUTURE RESEARCH

This is only a preliminary armchair survey of Sea Tribe lects. Much additional field research is needed to obtain a more complete picture of the dialects of this area, whether canonical Malay or tribal. Specifically, research is needed to fill in the data limitations mentioned earlier: learning more about Anambas Malay, Karas Malay and sampling the forty-five Sea Tribes territories mentioned in Chou (2010). The Riau Archipelago through to Bangka and Belitung should be the highest priority. Benjamin (2009: 318) calls researchers to seek for more archaic features (such as the infixes *-um-* and *-in-*, or the prefix *mar-/mər-*) in hopes of finding “a distant echo of the kind of linguistic matrix out of which Malay proper began to emerge around two millennia ago”. The time has also come to go deeper than just collecting a wordlist here and there, to focusing on grammar,²⁷ text collection,²⁸ sociolinguistics, and anthropological linguistics (Sea Tribe directionals for example). Duano in particular would benefit from more substantial historical linguistic attention.

Research could probe for additional connections with other Malay lects, particularly with Natuna and Sarawak Malay. In general, the question could be asked whether the older Riau lects more resemble the Malay of Sumatra, Peninsular Malaysia or Borneo. In that vein, the type of work Collins has done (1988, 1996, 1998a), seeking to tease apart various layers of contact and influence, has only here been attempted in the most rudimentary fashion.

Any linguistic research into historical connections faces limitations as more recent influences obscure older patterns. If “language history is a kind of palimpsest” (Blust 2010: 64), our experience here with the lects identified as Barok, Sekak, Akit, and Orang Darat seems to indicate that many of these groups are particularly susceptible to the fading and disappearance of older linguistic layers. Caveat emptor, therefore, to future researchers!

In the arena of sociolinguistics, it would be helpful to examine the language ecology as the physical and cultural environment is rapidly changing. Certainly, some groups have already assimilated and others will soon follow, so there is some urgency to research in this area. When looking at intergroup relations, it would be good to keep in mind Chou’s (2003: 45) picture of significant rivalry between groups. Reports of unintelligible languages (“rivalling Orang Laut may pretend not to comprehend each other” (Chou 2010: 6)), or of people with exotic/evil practices or completely distinct origins, should be taken with a grain of salt and triangulated with other evidence whenever possible.

²⁷ I have exploited only a fraction of the information in Kähler (1960). In general, we would do well to heed Steinhauer’s (1988: 151) cautionary conclusion: “Macassarese Malay (or should it be Malay Macassarese?) lends strong support to the theory that large scale lexical replacement is a possible cause of language birth [...] Comparative historical research with regard to languages such as Malay should take such possibilities into account. Undue emphasis on lexical aspects with disregard of grammar may lead to unrealistic pictures of language history”.

²⁸ Chou reports having collected a substantial number of Galang texts. These texts could be analysed as a starting point.

4.4 FINAL THOUGHTS

We set out on this journey, not seeking aromatic resins or gold, but a clearer understanding, driven not by the monsoons but by curiosity and scientific interest. I hope we can say we reached our initial destination, guided by the experts in the matter, the people of the sea. Here is some cargo for our ship's return voyage:

The Malayic languages spoken by the Sea Tribes are not some exotic creatures; in fact they fit quite nicely (with the exception of Duano) within the larger network of Malayic lects of the region. We have seen that the fourfold division of Sea Tribe lects (Kedah, Duano, Riau Islands, and Sekak) seems generally to hold while, given similarities to canonical Malay lects like Kedah, Johor, and Natunas, we should not gratuitously assume that the Orang Laut ethnic identity confers some special linguistic status. That having been said, certain lects seem to have a quite salient identity, particularly Duano, Urak Lawoi', Sekak, and the Riau Sea Tribes cluster.

No longer at the centre of the Malay kingdoms, the *Orang Laut* way of life continues to fade away, and language loss is a casualty in the cultural tumult the Sea Tribes are facing. May we learn all we can about these unique people and their speech before it is too late.

REFERENCES

- Adelaar, K. Alexander. 1992. *Proto-Malayic; The reconstruction of its phonology and parts of its lexicon and morphology*. Canberra: Australian National University. [Pacific Linguistics C-119.]
- Adelaar, K. Alexander. 2005. "Malayo-Sumbawan", *Oceanic Linguistics* 44 (2): 357-388.
- Adelaar, K. Alexander. 2008. "On the classifiability of Malayic", in: Yuri Lander and Alexander Oglöblin (eds), *Language and text in the Austronesian world; Studies in honour of Ülo Sirk*, pp. 1-22. Munich: LINCOM. [LINCOM Studies in Austronesian Linguistics 06.]
- Andaya, Barbara Watson and Leonard Andaya. 2001. *A history of Malaysia*. Second edition. London: Palgrave. [First edition 1982.]
- Andaya, Leonard. 2008. *Leaves of the same tree; Trade and ethnicity in the Straits of Melaka*. Honolulu: University of Hawaii Press.
- Anderbeck, Karl. 2007. "Haji; One language from twelve", in: Chong Shin, Karim Harun, and Yabit Alas (eds), *Reflections in Southeast Asian seas; Essays in honour of professor James T. Collins*, Vol. II, pp. 51-91. Pontianak: STAIN Pontianak Press.
- Anderbeck, Karl. 2008. *Malay dialects of the Batanghari River Basin (Jambi, Sumatra)*. Dallas: SIL International. [SIL e-Books 6: <http://www.sil.org/silepubs/abstract.asp?id=50415>, accessed on 15-8-2008.]
- Anderbeck, Karl and Bernard Sellato. In progress. *Pattern in the patchwork? Language, ethnicity and classification in inland Borneo*.
- Anderbeck, Karl and Uri Tadmor. In progress. *Language assimilation among the Sea People of Belitung*.

- Arnaud, Véronique, Henri Campagnolo, and Christian Pelras. 1997. *Lexique thématique plurilingue de trente-six langues et dialectes d'Asie du Sud-Est insulaire [Thematic multilingual lexicon of thirty-six languages]*. Paris: L'Harmattan.
- Asmah Haji Omar. 2008. *Susur galur bahasa Melayu*. Second edition Kuala Lumpur: Dewan Bahasa dan Pustaka. [First edition 1985.]
- Astar, Hidayatul. 2002. *Kosakata dasar Swadesh di kotamadya Pontianak dan kabupaten Pontianak*. Jakarta: Pusat Bahasa.
- Baer, Adela. 1999. *A Temuan-English-Malay lexicon*. Keene, NH. [Adela S. Baer Papers 2: Manuscripts, 1968-1991, 2001, Digital File.]
- Barnard, Timothy (ed.). 2004. *Contesting Malayness; Malay identity across boundaries*. Singapore: Singapore University Press.
- Barnard, Timothy. 2007. "Celates, Rayat-Laut, pirates; The Orang Laut and their decline in history", *Journal of the Malaysian Branch of the Royal Asiatic Society* 80(2): 33-49.
- Benjamin, Geoffrey. 2001. "Orang Asli languages; From heritage to death?", in: Razha Rashid and Wazir Jahan Karim (eds), *Minority cultures of peninsular Malaysia; Survivals of indigenous heritage*, pp. 101-122. Penang: Academy of Social Science (AKASS).
- Benjamin, Geoffrey. 2002. "On being tribal in the Malay world", in: Geoffrey Benjamin and Cynthia Chou (eds), *Tribal communities in the Malay world*, pp. 7-76. Singapore: Institute of Southeast Asian Studies.
- Benjamin, Geoffrey. 2009. "Affixes, Austronesian and iconicity in Malay", *Bijdragen tot de Taal-, Land- en Volkenkunde* 165(2-3): 291-323.
- Benjamin, Geoffrey and Cynthia Chou (eds). 2002. *Tribal communities in the Malay World; Historical, cultural and social perspectives*. Singapore: Institute of Southeast Asian Studies.
- Blagden, Charles Otto. 1906. "Language", in: Walter William Skeat and Charles Otto Blagden (eds), *Pagan races of the Malay peninsula*, Vol. 2, pp. 379-775. London: Macmillan.
- Blust, Robert. 1981. "Variation in retention rate among Austronesian languages". Paper, The Third International Conference on Austronesian Linguistics, Bali, 19-24 January.
- Blust, Robert. 1997. "Nasals and nasalization in Borneo", *Oceanic Linguistics* 36(1): 149-179.
- Blust, Robert. 2010. "The Greater North Borneo hypothesis", *Oceanic Linguistics* 49(1): 44-118.
- Chambers, J.K. and Peter Trudgill. 1998. *Dialectology*. Second edition. Cambridge: Cambridge University Press. [First edition 1981.]
- Chou, Cynthia. 2003. *Indonesian sea nomads; Money, magic, and fear of the Orang Suku Laut*. New York: Routledge.
- Chou, Cynthia. 2010. *The Orang Suku Laut of Riau, Indonesia; The inalienable gift of territory*. London: Routledge.
- Chou, Cynthia and Vivienne Wee. 2002. "Tribality and globalization; The Orang Suku Laut and the 'Growth Triangle' in a contested environment",

- in: Geoffrey Benjamin and Cynthia Chou (eds), *Tribal communities in the Malay world*, pp. 318–363. Singapore: Institute of Southeast Asian Studies.
- Collins, James T. 1985. "The phonology of Tioman Malay and the reconstruction of Proto-Malay", in: Suriya Ratanakul, David Thomas, and Suwilai Premasirat (eds), *Southeast Asian Linguistic Studies Presented to André-G. Haudricourt*, pp. 541–566. Bangkok: Institute of Language and Culture for Rural Development, Mahidol University. [[Http://purl.org/sealang/collins1985phonology.pdf](http://purl.org/sealang/collins1985phonology.pdf).]
- Collins, James T. 1986. "Dunia Melayu dan dialek Melayu; Wilayah Laut China barat daya", *Dewan Bahasa* 30: 895–907.
- Collins, James T. 1988. "Wilayah dialek Kedah; Bukti dari Sumatera Utara", *Dewan Bahasa* 32: 494–525.
- Collins, James T. 1989. "Malay dialect research in Malaysia; The issue of perspective", *Bijdragen tot de Taal-, Land- en Volkenkunde* 145: 235–264.
- Collins, James T. 1991. "Bahasa Melayu di Sungai Mahakam; Siasatan fonologi dialek Kutai", *Sari* 10: 3–24.
- Collins, James T. 1995. *Bibliografi dialek Melayu di pulau Sumatera*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Collins, James T. 1996. "Dialek Melayu Selangor; Tinjauan di Jugra", *Khazanah dialek Melayu*: 89–108.
- Collins, James T. 1997. "The Malays and non-Malays of Kalimantan Barat; Evidence from the study of language". Paper, The International Conference on Tribal Communities in the Malay World, Singapore, 24–27 March.
- Collins, James T. 1998a. "Bahasa Melayu di Kepulauan Natuna; Tafsiran dan klasifikasi awal", *Dewan Bahasa* 42: 539–555.
- Collins, James T. 1998b. *Malay, world language; A short history*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Collins, James T. 2008. "Lampiran 1: Enam varian Bidayuhik di lembah Sekadau, Kalimantan Barat", in: Chong Shin (ed.), *Bahasa Bidayuhik di Borneo barat*, pp. 103–118. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Dahlan, Saidat. 1989. *Geografi dialek bahasa Melayu Riau Kepulauan*. Jakarta: Pusat Bahasa.
- Dahlan, Saidat, M. R. Yunus, Abdullah Manan, and Abdul Razak. 1990. *Struktur bahasa Melayu Riau dialek Pasir Pangaraian*. Jakarta: Pusat Bahasa.
- Erard, Michael. 2005. "How linguists and missionaries share a Bible of 6,912 languages", *New York Times* 19 July.
- Hogan, David. 1972. "Men of the sea; Coastal tribes of South Thailand's west coast", *Journal of the Siam Society* 60: 205–235.
- Hogan, David. 1988. *Urak Lawoi'; Basic structures and a dictionary*. Canberra: Australian National University. [Pacific Linguistics C-109.]
- Hudson, A. B. 1967. *The Barito isolects of Borneo; A classification based on comparative reconstruction and lexicostatistics*. Ithaca, NY: Southeast Asia Program, Department of Asian Studies, Cornell University.
- Kadir, Mohd. Daud, M. Yunus R., and Sitti Syamsiar. 1986. *Dialek bahasa Orang Laut*. Jakarta: Pusat Bahasa.

- Kähler, Hans. 1946a. "Ethnographische und linguistische studien von den Orang Laut auf der Insel Rangsang an der Ostküste von Sumatra", *Anthropos* 41/44(1/3): 1-31.
- Kähler, Hans. 1946b. "Ethnographische und linguistische studien von den Orang Laut auf der Insel Rangsang an der Ostküste von Sumatra", *Anthropos* 41/44(4/6): 757-785.
- Kähler, Hans. 1960. *Ethnographische und linguistische Studien über die Orang Darat, Orang Akit, Orang Laut und Orang Utan im Riau-Archipel und auf den Inseln an der Ostküste von Sumatra*. Berlin: Dietrich Reimer. [Veröffentlichungen des Seminars für Indonesische und Südseesprachen der Universität Hamburg Band 2.]
- Kalipke, Hans and Mohamad Agar Kalipke. 2001. *Wörterbuch Sakai-Indonesisch-Deutsch*. Hamburg: Buske.
- Kasim, Yuslina, Marjusman Maksan, Syamsir Arifin, and Muhamad Yamin Zailoet. 1987. *Pemetaan bahasa daerah di Sumatra Barat dan Bengkulu*. Jakarta: Pusat Bahasa.
- Kin, Quah Chiew. 1999. "A relocation of Malay language varieties based on the diglossic description", *Sari* 17: 37-55.
- Kurniawati, Wati. 2002. *Kosakata dasar Swadesh di kabupaten Sanggau dan Sintang*. Jakarta: Pusat Bahasa.
- Lee, Alan. 2004. "The Temuan language of Ulu Selangor". Paper, The Eighth International Symposium on Malay/Indonesian linguistics (ISMIL 8), Penang, 31 July-2 August.
- Lenhart, Lioba. 1997. "Orang Suku Laut ethnicity and acculturation", *Bijdragen tot de Taal-, Land- en Volkenkunde* 153(4): 577-604.
- Lenhart, Lioba. 2001. "Orang Suku Laut communities at risk; Effects of modernisation on the resource base, livelihood and culture of the 'Sea Tribe people' of the Riau Islands", *Nomadic Peoples* 5(2): 67-86.
- Lenhart, Lioba. 2002. "Orang Suku Laut identity; The construction of ethnic realities", in: Geoffrey Benjamin and Cynthia Chou (eds), *Tribal communities in the Malay world*, pp. 293-317. Singapore: Institute of Southeast Asian Studies.
- Lewis, M. Paul (ed.). 2009. *Ethnologue, languages of the world*. Sixteenth edition Dallas: SIL International. [First edition 1951.]
- Mahdi, Waruno. 2009. "In search of an historical Sea-People Malay dialect with -aba-", in: K. Alexander Adelaar and Andrew Pawley (eds), *Austronesian historical linguistics and culture history; A festschrift for Robert Blust*, pp. 73-89. Canberra: Pacific Linguistics. [Pacific Linguistics 601.]
- Mama, Joe. 2010. "Panggil saja kami Orang Sawang", *Kompas*. [Http://cetak.kompas.com/read/xml/2010/01/22/05540726/panggil.saja.kami.orang.sawang, accessed on 22-1-2010.]
- Martis, Non, Buha Aritonang, Hidayatul Astar, and Wati Kurniawati. 2002. *Monografi kosakata dasar Swadesh di kabupaten Ketapang, Kapuas Hilir, dan Sambas*. Jakarta: Pusat Bahasa, Departemen Pendidikan Nasional..
- Maryono, Albertus Sinaga, Agus Setyonegoro, and Ade Kusmana. 1997.

- Dialektologi bahasa Kubu; Laporan penelitian*. Jakarta: Pusat Bahasa, Departemen Pendidikan Nasional.
- McDowell, Jonathan and Karl Anderbeck. 2008. *Bhinnêka Tunggal Ika: unity in diversity; Malayic varieties of southern Sumatra*. [Unpublished draft.]
- Mckinnon, Timothy. 2012. "A survey of phonation driven vowel shifts in traditional Malay(ic)". Paper, The Sixteenth International Symposium of Malay/Indonesian Linguistics (ISMIL 16), Kelaniya-Sri Lanka, 22-24 June.
- Moain, Amat Juhari. 1999. "Dialek Melayu Johor", in: Hussain Safian (ed.), *Ensiklopedia sejarah dan kebudayaan Melayu*, pp. 648-649. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Musa, Mahani. 2011. "The socioeconomic history of the Orang Kanaq of Johor", *Kajian Malaysia* 29(1): 47-74.
- Napsin, Syahrul, Zainal Abidin Naning, Selamat Abdullah, and Zulkarnain Mustafa. 1981. *Struktur bahasa Sekak*. Jakarta: Pusat Bahasa.
- Napsin, Syahrul, M. Yusuf Usman, R.M. Syafruddin Zein, Tarmizi Mairu, Sudarmo, and Sofyan Silahiddin. 1986. *Morfologi dan sintaksis bahasa Melayu Belitung*. Jakarta: Pusat Bahasa.
- Nothofer, Bernd. 1997. *Dialek Melayu Bangka*. Bangi: Penerbit Universiti Kebangsaan Malaysia.
- Pelras, Christian. 2002. "A few jottings on some aquatic populations of the Nusantara Archipelago". [Http://www.keene.edu/library/OrangAsli/pelras.pdf, accessed on 6-1-2005; Translated by S. Pernia and J. Durrell-Khalife, "Notes sur quelques populations aquatiques de l'Archipel nusantarien", *Archipel* (1972) 3: 133-168.]
- Phillips, Timothy C. In progress. *Proto-Aslian; Towards an understanding of its historical linguistic systems, principles and processes*. PhD thesis, Universiti Kebangsaan Malaysia.
- Rafferty, Ellen. 1983. "Contrasts in the syntax of a Malay and an Indonesian oral narrative", in: James T. Collins (ed.), *Studies in Malay dialects*, part I, pp. 1-36. Jakarta: Badan Penyelenggara Seri NUSA, Universitas Katolik Indonesia Atma Jaya. [NUSA: Linguistic Studies of Indonesian and Other Languages in Indonesia 16.]
- Riedel, J. G. F. 1881. "De baning djoekoet singak en baning seran njeroepak amang; Twee volksverhalen in het dialekt der Orang Lawoet of Orang Sekah van Bêlitoeng", *Tijdschrift voor Indische Taal-, Land- en Volkenkunde* 26: 264-273.
- Robert, Chongkit Sripun. 2010. *Changes in patterns of communication of the Urak Lawoi*. PhD thesis, University of Wisconsin.
- Saengmani, Amon. 1979. "Phonology of the Urak Lawoi' language; Adang Island". MA thesis, Mahidol University.
- Schebesta, Paul. 1926. "Kubu und Jakudn (Jakun) als Protomalayen" (Translated by M. Kummerow and Adela Baer), *Mitteilungen der Anthropologischen Gesellschaft in Wien* 56: 192-201.
- Schot, J.G. 1884. "Het stroomgebied der Kateman; Bijdrage tot de kennis von Oost Sumatra", *Tijdschrift voor Indische Taal-, Land- en Volkenkunde*, Vol. 29: 555-581.

- Seidlitz, Eric. 2005. "A study of Jakun Malay; Some aspects of its phonology". MA thesis, Universiti Kebangsaan Malaysia.
- Seidlitz, Eric. 2007. "Duano; a first look at its phonology", in: Chong Shin, Karim Harun, and Yabit Alas (eds), *Reflections in Southeast Asian seas; Essays in honour of professor James T. Collins*, Vol. II, pp. 23-49. Pontianak: STAIN Pontianak Press.
- Skeat, Walter William and Charles Otto Blagden. 1906. *Pagan races of the Malay peninsula*. London: Macmillan.
- Skeat, Walter William and H.N. Ridley. 1900. "The Orang Laut of Singapore", *Journal of the Straits Branch of the Royal Asiatic Society* 33-34: 247-250.
- Smedal, Olaf. 1987. *Lom-Indonesian-English and English-Lom wordlists accompanied by four Lom texts*. Jakarta: Badan Penyelenggara Seri NUSA, Universitas Katolik Indonesia Atma Jaya. [NUSA: Linguistic Studies of Indonesian and Other Languages in Indonesia 28-29.]
- Smedal, Olaf. 1989. *Order and difference; An ethnographic study of Orang Lom of Bangka, West Indonesia*. Oslo: Department of Social Anthropology. [Oslo Occasional Papers in Social Anthropology 19.]
- Smith, Sydney, Macvey Napier, and William Empson. 1814. "History and languages of the Indian islands (review of Marsden's grammar and dictionary of the Malayan language)", *The Edinburgh Review; Or Critical Journal* 23(45): 151-189.
- Sopher, David E. 1977. *The sea nomads; A study based on the literature of the maritime boat people of Southeast Asia*. Singapore: National Museum. [Reprint of first edition with postscript.]
- Steinhauer, Hein. 1988. "Malay in East Indonesia; The case of Macassarese Malay", in: Mohd. Thani Ahmad and Zaini Mohamed Zain (eds), *Rekonstruksi dan cabang-cabang bahasa Melayu induk*, pp. 105-151. Kuala Lumpur: Dewan Bahasa dan Pustaka. [Monografi Sejarah Bahasa Melayu.]
- Steinhauer, Hein. 2008. "On the development of Urak Lawoi' Malay", *Wacana, Jurnal Ilmu Pengetahuan Budaya* 10(1): 117-143.
- Stokhof, W. A. L. (ed.). 1987. *Holle Lists; Vocabularies in languages of Indonesia, Vol. 10/1: Minangkabau and languages of Central Sumatra*. Canberra: Australian National University. [Pacific Linguistics D-74.]
- Suardi. 1993. "Bahasa Melayu suku asli di Riau; Satu tinjauan linguistik historis komparatif", *Dewan Bahasa* 37(2): 128-147.
- Syamsiar, Sitti, M. R. Yunus, Mohd. Daud Kadir, and Rukmini. 1986. *Struktur bahasa Sokop*. Jakarta: Pusat Bahasa.
- Tadmor, Uri. 2002. "Language contact and the homeland of Malay". Paper, The Sixth International Symposium of Malay/Indonesian Linguistics (ISMIL 6), Bintan Island, 3-5 August.
- Tadmor, Uri. 2003. "Final /a/ mutation; A borrowed areal feature in Western Austronesia", in: John Lynch (ed.), *Issues in Austronesian historical phonology*, pp. 15-35. Canberra: Australian National University. [Pacific Linguistics 550.]
- Thurgood, Graham. 1999. *From ancient Cham to modern dialects: two thousand years of language contact and change; With an appendix of Chamic reconstructions*

and loanwords. Honolulu: University of Hawaii Press. [Oceanic Linguistics Special Publications 28.]

Umar, Said Mahmud, T. Nontel, Pang Cik, and Burhan Yunus. 1991. *Struktur bahasa Akit*. Jakarta: Pusat Bahasa.

Usman, Burhanuddin Ch., Yusmaniar Noor, Apul Simbolon, and Setiana Simorangkir. 1983. *Struktur bahasa Orang Laut*. Jakarta: Pusat Bahasa.

Wilkinson, R. J. 1959. *A Malay-English dictionary (Romanised)*. London: Macmillan.

Yusof, Mohd. Sharifudin. 2006. *Obsolesensi dan kenazakan bahasa Orang Melayu Proto; Kajian sosiologi bahasa terhadap dialek Duano dan dialek Kanaq di Johor*. PhD thesis, Universiti Kebangsaan Malaysia.

APPENDIX: BINTAN ORANG LAUT WORDLIST

This wordlist follows the format of the Basic Austronesian (BAN) wordlist (Blust 1981) with a few additional items at the bottom. Speaker metadata are also provided at the bottom of the list.

BAN	English	Bintan OLaut
1	hand	taŋan
2	left	kii
3	right	kanan
4	leg / foot	kaki
5	walk (v.)	bə-jʌlʌn
6	road/path	jʌlʌn
7	come	dʌtaŋ
8	turn (left or right)	belo?
9	swim (v.)	bənaŋ
10	dirty	kətəh
11	dust	abu?
12	skin	kulit
13	back (body part)	pəŋgəŋ
14	belly	pəwət
15	bone	tulaŋ
16	intestines	uso?
17	liver	ati
18	breast	susu
19	shoulder	bəu
20	know (things)	tau
21	think	bə-pike
22	fear (v.)	takot
23	blood	dʌrʌh
24	head	kəpalu
25	neck	lee
26	hair (of the head)	ambot
27	nose	idoŋ

BAN	English	Bintan OLaut
28	breathe	bə-napas
29	sniff, smell	ciom
30	mouth	mulot
31	tooth	gigi
32	tongue	lidʌh
33	laugh (v.)	tawu
34	cry (v.)	naŋes
35	vomit (v.)	mutah
36	spit (v.)	ludʌh
37	eat (v.)	makan
38	chew (v.)	pəpa?
39	cook (v.)	masa?
40	drink (v.)	minom
41	bite (v.)	gigit
42	suck (v.)	isap
43	ear	təliŋu
44	hear (v.)	dəŋah
45	eye	matu
46	see (v.)	teŋo?
47	yawn (v.)	saŋap
48	sleep (v.)	tedoh
49	lie down (to sleep)	bʌeŋ
50	dream (v.)	mimpi
51	sit (v.)	dudo?
52	stand (v.)	bədii
53	person/ human being	uaŋ
54	man/male	jʌntʌn

BAN	English	Bintan OLaut
55	woman/ female	bətinuu
56	child	ana?
57	husband	laki
58	wife	bini
59	mother	ṃa?
60	father	ua?
61	house	umah
62	roof/thatch	atap
63	name	namuu
64	say (v.)	bə-katu
65	rope	tali
66	tie up, fasten	ikat
67	sew (clothing)	jʌet/ṃʌet
68	needle	jʌom
69	hunt (for game)	buu
70	shoot (as with a gun)	t/nemba?
71	stab, pierce (overhand)	tikam
72	hit (with stick, club)	pokol
73	steal	cui
74	kill	bunoh
75	die, be dead	mati
76	live, be alive	idup
77	scratch (an itch)	ṃgʀu/gʀu
78	cut, hack (wood)	kəuʌat
79	stick (wood)	kayu
80	split (v. tr.)	bəlʌh/məlʌh
81	sharp	tajʀm
82	dull, blunt	tompol

BAN	English	Bintan OLaut
83	work (in garden, field)	bə-kəjuu
84	plant (v.)	nanam
85	choose (v.)	pileh/mileh
86	grow (v. intr.)	tumbuh
87	swell (as an abscess)	bəṃka?
88	squeeze (as juice from a fruit)	pəuʌah
89	hold (s.t., on to s.t.)	pəgan
90	dig (v.)	gʀli/ṃʀli
91	buy (v.)	bəli/məli
92	open, uncover (v.)	buka?
93	pound, beat (as rice or prepared food)	tumbu?
94	throw (as a stone)	tikam
95	fall (as a fruit)	jʌtoh
96	dog	koyo?
97	bird	buoṃ
98	egg	təloh
99	feather	bulu
100	wing	sayap
101	fly (v.)	təban
102	rat	tikus
103	meat/flesh	dʌgij
104	fat/grease	ləma?
105	tail	kote?, eko
106	snake	ulah
107	worm (earthworm)	caciṃ

BAN	English	Bintan OLaut
108	louse (head)	kutu
109	mosquito	ŋet
110	spider	lʔbalʔbah
111	fish	ikan
112	rotten (of food, corpse)	busuʔ
113	branch	dʔʔn
114	leaf	
115	root	akah
116	flower	buŋu
117	fruit	buʔh
118	grass	umput
119	earth/soil	tanah
120	stone	bʔtu
121	sand	paseh
122	water (fresh water)	aeh
123	flow (v.)	mə-leleh
124	sea	laot
125	salt	gʔʔm
126	lake	kolam payu
127	woods/ forest	utan
128	sky	lanjet
129	moon	bulʔdn
130	star	bintan
131	cloud (white cloud; not a raincloud)	pəŋaʔ
132	fog	kabut
133	rain	ujʔn
134	thunder	guoh
135	lightning	kilat
136	wind	aŋen
137	blow (v.)	tiup

BAN	English	Bintan OLaut
138	warm/hot (tea)	panas
139	cold	səjuʔ
140	dry	kəuŋiŋ
141	wet (cloth)	bʔsʔh
142	heavy	bəuʔt
143	fire	api
144	burn (v. tr.)	bʔkah
145	smoke (of a fire)	asap
146	ash	abu
147	black	itam
148	white	puteh
149	red	meah
150	yellow	kunin
151	green	ijow
152	small	kəciʔ
153	big	bəsah
154	short	pendeʔ
155	long (of objects)	pancan
156	thin (of objects)	nipis
157	thick (of objects)	təbʔl
158	narrow	səmpit
159	wide	lebah
160	painful, sick	sakit
161	shy, ashamed	malu
162	old (of people)	tuu
163	new	bʔu
164	good	bʔeʔ
165	bad, evil	jʔhʔt
166	correct, true	bənah

BAN	English	Bintan OLaut
167	night	malam
168	day	aũi
169	year	taun
170	when (question)	belɣ?
171	hide (v. intr.)	səmuɲi
172	climb (ladder)	nae?
173	at	dəkɣt
174	in, inside	dɣlɣm
175	above	atas
176	below	bɣwɣh
177	this	ekɣ?
178	that	eto?
179	near	dəkɣt
180	far	jɣuh
181	where (question)	mɣnuu
182	I	aku
183	you (sing.)	dikəw, mikuu
184	he/she	diɣ?

BAN	English	Bintan OLaut
185	we (excl.)	kame?
185	we (incl.)	kitɣ?
186	you (plural)	dikəw səmuuu
187	they	mikɣ?, uaŋ diɣ?
188	what (question)	apɣ?
189	who (question)	sapɣ?
190	other	lain
191	all	səmuuu
192	and	dəŋan
193	if	kalaw
194	how (question)	bɣgɣymɣnuu
195	no, not	tidɣ?
196	count (v.)	eton
197	one	satu
198	two	duu
199	three	tigu
200	four	mɣpat

ADDITIONAL WORDLIST

	English	Bintan OLaut
1	five	limuu
2	six	ɲam
3	seven	tujoh
4	eight	lɣpɣn
5	nine	səmilɣn
6	ten	səpuloŋ
7	eleven	səbəlɣs
8	hundred	sa:tos
9	thousand	sibu

	English	Bintan OLaut
10	body	bədɣn
11	finger	jɣũi
12	friend	kawan
13	chicken	ayam
14	tree	poko?
15	coconut (ripe)	nioŋ
16	cassava	ubi
17	machete	paũaŋ
18	pillow	kope?, bɣntɣl

	English	Bintan OLaut
19	strong	kuat
20	many	baja?
21	here	kian 'to here'
22	there	kiun 'to there'
23	angry	mauqah
24	run	lauji

	English	Bintan OLaut
25	let's	kian
26	don't	usa?
27	there isn't	musx?
28	fish line	kodik
29	loincloth	kancut 'underwear'

Metadata

Village	Berakit Panglung Ujung
Sub-District	Teluk Sebung
District	Bintan
Province/State	Kepulauan Riau
Country	Indonesia
Informant Age	58
Sex	male
Elicited By	Karl Anderbeck
Date	22 January 2010
Notes: Informant grew up on the sea (both parents Orang Laut), and started living on land at age 17. No significant schooling. Also speaks Malay and Indonesian.	