On the development of Urak Lawoi' Malay

HEIN STEINHAUER

ABSTRAK
Urak Lawoi' yang berarti 'orang laut' adalah nama suku nelayan pengembara laut yang tersebar di Pulau Phuket (Thailand) dan Kepulauan Adang di sebelah selatannya. Bahasa mereka yang juga dikenal dengan nama Urak Lawoi' merupakan sebuah varietas Melayu yang khas. Dari pemeriannya yang ada tampaklah bahwa bahasa itu menyimpang dari varietas Melayu di daratan Thailand Selatan dan di semenanjung Malaysia. Dalam artikel ini direkonstruksikan perubahan bunyi dalam urutannya yang diakronis, yang telah memberikan wujud khas pada kata-kata leksikal Urak Lawoi' itu. Yang menonjol dalam hal itu adalah dua perubahan bunyi yang mengingatkan bahasa-bahasa di Kalimantan Barat dan Serawak, yaitu hilangnya letupan bersuara setelah konsonan nasal di dalam morfem (*-mb- > -m-, dan seterusnya), dan penggantian nasal pada akhir kata oleh konsonan tak bersuara yang homorgan (*-m > -p, dan seterusnya).

KATA KUNCI
Malay varieties, Kedah Malay, Patani Malay, Proto Malay, Standard Malay, Malayic Dayak, Land Dayak, Urak Lawoi', Moklen-Moken, Thai Phonemes.

1 INTRODUCTION[1]

The evolutionary success of Malay has led to a great variety of isolects and language continue and the three standardized Malay-derived languages of Indonesia, Malaysia, and Brunei Darussalam. How all these varieties of Malay

[1] An embryonic version of this paper was presented at the *Simposium Internasional Ilmu-Ilmu Humaniora II Bidang Linguistik dan Sejarah*, Yogyakarta 25-27 April 1993. I am grateful to Orin Gensler for his comments on a later version.

HEIN STEINHAUER (1943) studied Slavic and Baltic languages at the Universities of Amsterdam and Zagreb. Since his PhD dissertation on Croatian dialects (University of Amsterdam 1973), he has been working on Austronesian and Non-Austronesian languages of Indonesia. From 1988 to 1994 he was stationed at Pusat Bahasa, Jakarta, for the Indonesian Linguistics Development Project (ILDEP). He is currently affiliated to Leiden University, and to Radboud University Nijmegen, where he holds the extra-ordinary chair for ethnolinguistics of Southeast Asia. He published extensively on Indonesian, varieties of Malay, and on other regional languages of Indonesia.

are interrelated, is still largely a matter of conjecture and hypothesis, if only because so many of them have hardly been reliably described.

The present paper is a comparison of a Malay "outlier" isolect, namely the Urak Lawoi' (UL) language variety with Malaysian/Indonesian (collectively referred to as Standard Malay, SM) and Proto-Malay, with the aim of establishing the sound changes which gave UL its present shape, where possible in their relative chronological order.

UL is spoken as a home language among an estimated 6000[2] "sea people" (*orang laut* in SM, *urak lawoi'* /urak lawoy/ in UL) on some of the islands along the west coast of southern Thailand[3]. For the UL data I rely on Hogan's papers on spelling and on comparative syntax (Hogan 1976, 1978), his short grammar (Hogan 1999) and especially on the description and dictionary by Hogan and Pattemore (1988)[4]. For Proto-Malay the standard study is Adelaar (1992). For SM I used the current monolingual and bilingual dictionaries (Wilkinson 1959; Iskandar 1984; Teeuw 1996; KBBI 2001).

Hogan and Pattemore (1988) and Hogan (1999) distinguish three varieties of UL: the Southern or Adang dialect, Phuket Young People's dialect and the Phuket Old People's dialect. The description concentrates on the latter, "as it has the most phonological contrasts, and many of the forms of the other dialects can be derived from it" (Hogan and Pattemore 1988: 1)[5]. According to the map on page v in this source, these other dialects also include the centrally located dialect of Lanta Island. No information, however, is given on the variety of UL spoken in this central area.

Before moving on to the sound changes which have lend the Phuket Old People's dialect of Urak Lawoi' its special phonotactic character, I shall first discuss some phonological aspects of this dialect[6]. The paper closes with a discussion about the possible relations of Urak Lawoi' with other Malay varieties.

2 Phonology

2.1 Hogan's inventory of phonemes is presented in tables 1 and 2. Where the symbols I shall use differ from those of Hogan, his are added in parentheses. Hogan 1999 uses the more common phonetic symbols.

---

[2] The actual number of speakers is probably less.

[3] Recent information on the spread of the Urak Lawoi' in the Adang Archipelago and beyond is given in Wongbusarakum (2007), which is an excellent account of their traditional culture and expert knowledge of their environment, which both are probably fatally threatened by commercial fishing, the tourist industry, and other effects of globalization and modern urban culture.

[4] Hogan 1999 is in fact a rephrasal of the descriptive introduction of Hogan and Pattemore 1988. If below I refer Hogan's analysis it is to the analysis in the latter source, of which Hogan is the main author.

[5] Below a reference to just a page number will always be a reference to this source.

[6] Below the term Urak Lawoi' and the abbreviation UL will be used to refer to this particular dialect and to its reconstructed earlier stages.

	Labial	Alveolar	Alveo-palatal	Velar	Glottal
Stops					
Voiceless Aspirated	p^h (ph)	t^h (th)	c^h (ch)	k^h (kh)	
Unaspirated	p	t	c	k	ʔ(q)
Voiced	b	d	j	g	
Fricatives		s			h
Nasals	m	n	ø (ny)	ŋ (ng)	
Lateral		l			
Semi-vowels	w	r	y		

Table 1: Consonants (see Hogan and Pattemore1988: 13)

	Unrounded		Rounded
	Front	Central	Back
High	i		u
Mid	e (ë)	ə (e)	o
Low	ɛ (ä)	a	ɔ (ö)

Table 2: vowels (see Hogan and Pattemore 1988: 21)

2.2 The series of aspirated voiceless stops are mainly found in borrowings from Thai, such as *khru* 'teacher'. Only a few occur in originally Malay words. Some of these derive from *C-h- through apocope, for example, *khana-* 'intend to' (compare SM *kehendak* 'wish'). Word-initial /ph/ especially may be the result of assimilation (compare *hadak* 'protect against evil', *phehadak* 'protection'). Some, such as *phiraw* 'tack (of sailing boat)' (SM *pirau*), cannot be explained. In contrast to SM, the glottal stop is well established as a phoneme in UL.

As Table 2 shows, UL has a richer vowel system than SM. In non-final syllables within roots, the front and back vowels other than /i/ and /u/ seem to occur only in loanwords. In closed final syllables they are frequent, notably in inherited words, but the number of minimal pairs is limited. I found the following for /o/ and /ɔ/:

(1)	*bubɔn*	'kind of tree'[7]	*bubon*	'mend (net, etc.)'
	tupɔn	'piece of wood or metal to clean the ear'	*tupon*	'blunt'
	kətəyə	'food scrape'	*kətoyə*	'pass wind'
	timɔn	'cucumber'	*timon*	'float'

[7] The examples are quoted from Hogan's dictionary. The glosses are his. I add Hogan's symbol for word-class, only if disambiguization is needed. If the example quoted is not an alphabetically retrievable dictionary entry, I add the relevant page-number(s).

lik□□	'turn, swerve'	*liko□*	'wind (rope, cloth, etc.)'[8]
m□ŋat□y□	'hit, crack open'	*m□ŋatoy□*	'sleepy'
p□t□k (*buloh* ~)	'bamboo shoots'	*p□tok*	village on Peepee Island
d□□	'dock'	*do□*	'Father! (adoptive)'

The front vowels show some complementarity in distribution. Before word-final /n, t, p, w/ they exclude each other, before others, notably /ʔ/ and /h/, both may occur. In (2) the number of occurrences before these consonants is given:

(2)	-n#	-t#	-p#	-w#	-□#	-h#	
	[e]	18	15	4	-	34	23
	[ɛ]	-	-	-	5	44	8

I found the following (near) minimal pairs:

(3)				
	b□rnek	(no meaning given), from		
	nek	'carry child on hip'	*b□rn□k*	'embrace'
	mine□	'minute'	*min□□*	(no meaning given, in:
			aye min□□	'soft drink' (p.103))
	kute□	'collect (taxes, small items)'	*ut□□*	'catfish (deep water)'
	g□rne□	'beads'	*k□rn□□*	'dots, dashes'
	maneh	'sweet'	*j□n□h*	'kind'
	sireh	'betel leaf'	*sir□□*	1) 'turn around (with hand held up as in Manohra dance)', 2) 'approach a superior'.

2.3.1 Hogan rejects a bivocalic interpretation of [Vi] before word-final glottals because "the language has no non-suspicious vowel clusters" (p. 16). Now, suspiciousness is in the eyes of the beholder, and it is unclear why Hogan writes these sequences *phonetically* as sequences of vowels. Since [-Vi□] and [-Vih] are said to be part of the final syllable, the suspect sequences may be diphthongs. Yet, they are not treated as such, although the language does have diphthongs, also in Hogan's analysis[9]. Instead, the [-ih] part of the [-Vih] sequence is analysed as an allophone of /s/, because of the /s/ in Malaysian cognates. Historically this can be justified, synchronically it is arbitrary, since [h] < *s after a single vowel is analysed as /h/.

Similarly, when the glottal stop [□] < *t occurs after a single vowel it is analysed as /□/, but when this [□] < *t is preceded by a [Vi] sequence, the [-i□] segment is again analysed as the allophone of a single consonant. But this time it cannot be */t/, in spite of the SM cognates with /-t/, since UL *has*

[8] These are the meanings given on p. 132. On p. 23 *lik□□* is only 'swerve', and *liko□* 'cover up'

[9] An example is the second syllable [baw] of the word for 'water-buffalo', *krerbaw* (Hogan's spelling), as opposied to [bau] 'smell'. The latter word is bisyllabic.

word-final [-Vt] sequences, which Hogan analyses synchronically as /-Vt/. His way-out is to consider [-i□] after a vowel as the syllable-final allophone of the alveopalatal stop /c/. This is a) arbitrary, b) apparently contrary to the phonetic facts, c) historically unjustifiable, and d) at variance with what is generally found as a phonotactic constraint in Malay dialects. In my analysis therefore, Hogan's word-final /s/ and /c/ will be reinterpreted as /yh/ and /y□/ respectively. Compare the following SM words, followed by their UL cognates in broad phonetic notation, in Hogan's (phonemic) spelling, and in my phonemic interpretation (in that order):

(4)	SM	*tikus*	'mouse'	UL	[tikuyh]	*tikus*	/tikuyh/
		hangus	'scorched'		[haŋ□yh]	*hangös*	/haŋ□yh/
		beras	'husked rice'		[brayh]	*bras*	/brayh/
		habis	'complete'		[habih]	*habih*	/habih/
		manis	'sweet'		[maneh]	*manëh*	/maneh/
		takut	'fear'		[takoy□]	*takoc*	/takoy□/
		semut	'ant'		[s□m□y□]	*semöc*	/s□m□y□/
		pantat	'buttocks'		[pantay□]	*pantac*	/pantay□/
		pahit	'bitter'		[pahe□]	*pahëq*	/pahe□/

2.3.2 Hogan's description of /r/ is confusing. In his Table of consonants (p. 13), /r/ is qualified as an alveolar semi-vowel. Syllable-initially [r] (Hogan's phonetic symbolization) is said to vary "from a retroflex vocoid to a slight flap" (p. 18). Whether /r/ in the position C-V varies in the same way is unclear. As the coda of a syllable, whose nucleus is then always [□], /r/ is also said to be "a retroflexed vocoid" (p. 18). This /r/, however, appears to be Hogan's interpretation of "the second mora" of a "vocoid cluster [□□]", which is "phonetically … a lengthened [□]" (p. 18). When this lengthened schwa "occurs in the penultimate syllable, that syllable bears the major word stress and the vowel cluster [sic, HS] fluctuates phonetically between the allophones [r ~ re ~ er]".[10] Hogan's spelling of words with *r* is insufficiently consistent; there are quite a number of examples such as *prerlëh* (pp. 18 and 144) vs. *perlëh* (p. 143) '(the Malaysian state of) Perlis', *trerbëq* (pp. 18 and 157) vs. *terbëq* (p. 155) 'depart', *krerja* (pp. 19 and 128) vs. *kerja* (p. 127) and even *kr□ja* 'work' (Hogan 1999: 46), *k□b□lot*, *k□rb□rlot* 'starve' (Hogan 1999: 22, 40) vs. *k□rb□lot* 'starving' (Hogan and Pattemore 1988: 127). I assume that the optional (additional) *r* in these words is the reflection of vowel lengthening in stressed or secondary stressed position, and in the case of spellings such as *krerja* of the anticipatory retroflex articulation (triggered by the syllable closing *r). In the position C_V there is an opposition /r/ ~ /□r/, given Hogan's observation that "words with the /br/ cluster … must be distinguished from words commencing with the prefix /ber/ which sometimes omits the /e/" (p. 15, read /b□r/ and /□/).

[10] To be read as [r ~ r□ ~ □r].

2.3.3 Some problems are raised by Hogan's analysis of the semivowel [y] between /i/ and another vowel, and of [w] between /u/ and another vowel. Given near minimal pairs – at least in writing – such as those under (5) in Hogan's spelling, one is inclined to conclude that the glides are phonemic.[11]

(5)	*tuwah*	'luck', SM *tuah*	*tu.at*	'master', SM *tuan*
	giya	'gear wheel'	*gi.ak*	'cymbals'
	liyak	'wound, sore', SM *liang* 'hole	*'li.aŋ*	'feast' (<Thai)
	liya	'ginger', SM *halia*		
	siyak	'light, daylight', SM *siang*	*si.ap*	'Thai', SM *siam*
	siya□	'prepare', SM *siap* 'prepared'.		

Below I shall stick to Hogan's analysis without trying to find an explanation why some Malay words appear with a glide between a high and a non-high vowel, while others do not.

3 Sound changes/correspondences

3.1 An estimated one third of the mono-morphemic entries of the dictionary cannot be reconstructed as Malay words. Some of these are marked as loanwords (from standard Thai mainly, also from English). Others are geographical names, or exclamations. Some of the remaining words have a non-Malay shape: they are monosyllabic or show foreign phonotactic patterns.[12] The remainder consists of entries which could belong to the Malay core of UL, or which derive from another Malay dialect. Further research may point out Malay cognates for these words.

3.2 As in other varieties of Malay the most conspicuous sound innovations have occurred in the final syllables of roots or stems. UL does not show any trace of verbal suffixes comparable to SM *–i* and *–kan*. UL does not show a trace of the frequent SM suffix (or homophonous suffixes) *-an* either, except in the following three words: *duriat* 'durian' (SM *duri-an*), *s□latat* 'south-east' (SM *selat-an* 'south'), *b□rhaget* 'share (n)' (presumably cognate to SM *bahagi-an*). It is likely that Proto-Malayic *-an was lost in UL. Some striking examples of the absence of an expected suffix are given in (7) (left UL, right SM in standard spelling).

(7)	*kawat*	'friend, group of'	*kawan* 'friend'
			kawanan 'flock, herd, swarm'
	paso□	'group'	*pasukan* 'troops, group, formation'

[11] Hogan uses a dot to indicate the presence of a syllable boundary.

[12] The most likely source of these words is Southern Thai. According to Larish (1997: 126 footnote) UL "has been heavily influenced by Southern Thai." Larish (1997: 127, 135, 141) also refers in passing to contacts between UL speakers and speakers of Moklen-Moken. Possible UL influence on these languages is adduced as an explanation for some deviating sound patterns (Larish 1997: 135, 141).

ramoy□	'rambutan'	*rambutan* '(fruit of) Nephelium tree, *rambutan'*
tunaŋ	'fiancée'	*tunangan* 'fiancé(e)'
p□rbuway□	'acts, things done'	*perbuatan* 'deed, act, action, conduct, performance'
p□laboh	'anchorage'	*pelabuhan* 'anchorage, harbour'.

3.3 Unless otherwise indicated, I shall use "Proto-UL" to compare UL with. This Proto-UL is largely similar to what Adelaar (1992) reconstructs as Proto-Malayic, with the main difference, that there are already some forms with mid vowels, and that Proto-Malayic *a and *schwa in final syllables already had merged to *a. Where Adelaar reconstructs *A (for what may have been either *schwa or *a), I use *□ instead. For Adelaar's word-final *glottal stop I use *-k, with the understanding that its phonetic realization may indeed have been glottal. Finally, I include in Proto-UL some early loanwords of non-Malayic origin.

As will be shown below, the following sound changes have shaped the final syllables of UL in the same order of occurrence:

1. Insertion of a glide in vowel clusters beginning with a *high vowel.
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.
6. Simplification of homorganic nasal-stop sequences.
7. Change of final *-l into -n.
8. Lateralization of *-r

In Table 3 the UL reflexes are shown of all Proto-UL nucleus-coda combinations in final root syllables with *i, *u or *a as the nucleus, and with a nasal or another consonant as its onset. Final syllables without an onset follow the latter pattern. The number of attested occurrences of each sound change is added in parentheses. What appear to be the regular sound changes are printed in bold. Exceptions to added in a regular and smaller font. As far as the data go, exhaustive examples of the regular sound changes are given throughout this paper. The apparent exceptions, which may turn out to be indicative for connections with other Malay varieties, are dealt with in the Appendix.

	*-u-		*-i-		*-a-	
	N-	C- (≠N)	N-	C- (≠N)	N-	C- (≠N)
*-m	-	**op** (5)	-	**ip** (1), et (1)	**am** (4)	**ap** (26)
*-n	□**n** (1), un (1)	**ot** (9), ut (1)	**en** (2)	**et** (7)	**an** (7)	**at** (34), an (1)
*-ŋ	□**ŋ** (2)	**ok** (25), □k (3), □ŋ (1), uŋ (1)	**i** (3)	**ik** (14), iŋ (2), ek (1)	**aŋ** (6)	**ak** (92), aŋ (3)
*-p	-	**o**□ (3), u□ (1)	-	**e**□ (2), i□ (1), ɛ□ (1)	**a**□ (17), e□ (1), o□ (1)	
*-t	□**y**□ (4)	**oy**□ (26), uy□ (2), □y□ (1)	□□ (4)	**e**□ (16), ɛ□ (2)	**ay**□ (64), e□ (1)	
*-k	□□ (2)	**oq** (20)	-	**i**□ (10)	**a**□ (40)	
*-h	□**h** (2)	**oh** (22)	□**h** (2), **eh** (1)	**eh** (14)	**ah** (26)	
*-s	□**yh** (1)	**uyh** (8)	□**h** (2), **eh** (1)	**ih** (9), eh (2)	**ayh** (30)	
*-l	-	**on** (8)	-	**en** (6), in (2)	**an** (16)	
*-r	□**l** (3), ul (1)	**ol** (18), ul (2), un (1)	□ (2)	**el** (3), e (2)	**al** (32)	
Ø	**u** (65)		**i** (80)		**a** (112)	
*-w	-		-		**aw** (17)	
*-y	-		-		**ay** (25)	

Table 3: UL reflexes of root-final *-VC sequences, preceded by a nasal (N-) or another consonant (in parentheses the number of attestations).

3.4 Below the regular sound-changes illustrated in Table 3 will be discussed in detail and with examples. The sound changes are formulated in their supposed chronological order. Intermediate stages of development are marked with "+", the initial Proto-UL forms with "*". The numbers of the following paragraphs (3.4.1-8) correspond to the numbers of the chronological list of eight regular sound changes. The notation "*XXX $>_{1,2,4}$ YYY" means that the Proto-UL form XXX has developed into YYY via the sound changes 1, 2 and 4.

The Proto-UL forms are glossed as their SM cognates. Where the current UL meaning differs essentially from this SM meaning the former is given for all forms later than Proto-UL. Attested forms are written in italics. SM forms are spelled phonemically.

3.4.1 Insertion of a glide in vowel clusters beginning with a *high vowel

As indicated above (2.3.4) the appearance of a glide between a high vowel and a following other vowel (whether or not *high) does not seem to be automatic. Another conditioning factor than the tendency for borrowings from Thai to be pronounced without a glide cannot be formulated. The articulation of the glide corresponds to the preceding vowel: [y] after [i], [w] after [u]. Examples:

(8) *tuah > *tuwah* 'luck'
*tuil > ⁺tuwil 'lever'
*siaŋ > ⁺siyaŋ '(day)light'
*siuŋ > ⁺siyuŋ 'horn of an animal'

This glide insertion after *high vowels cannot be pinpointed in time. In fact, it may have taken place after all other changes, of which it is independent.

3.4.2 Lowering of *high vowels in closed final syllables

This is a sound change, which UL may have had in common with many Malay varieties, although the details may differ.

After a nasal onset the resulting vowel was lower than after another consonant, undoubtedly a corollary of progressive vowel nasalization (see also 3.4.5).

For *-u- there are many examples. Some cases of *-u- after non-nasal and nasal consonants are presented in (9) and (10). There are only rare exceptions, the only systematic exception being the preservation of ⁺-u- < *-u- in the position between a non-nasal onset and a closing *-s (11). After *-a- a glide -w- was inserted before the reflex of *-u- and a glide -y- before the reflex of *-i- (see (12) and (16)).

(9)	*hidup		> ⁺hidop[13]	'live'
	*mulut	'mouth'	> ⁺mulot	'mouth, voice'
	*masuk		> ⁺masok	'enter'
	*tujuh		> *tujoh*	'seven'
	*jarum		> ⁺jarom	'needle'
	*k□bun		> ⁺k□bon	'garden'
	*siuŋ	>$_1$ ⁺siyuŋ	> ⁺siyoŋ	'horn of an animal'
	*tumpul		> ⁺tumpol	'blunt'
	*ikur		> ⁺ikor	'tail'.
(10)	*lumut	'moss, algae'	> ⁺lum□t	'marine growth'
	*□amuk		> ⁺□am□k	'mosquito'
	*p□nuh		> *p□n□h*	'full'
	*haŋus	'scorched'	> ⁺haŋ□s	'scorch in fire, blister'
	*gunuŋ		> *gun□ŋ*	'mountain'

[13] See the Appendix for the deviating pattern(s).

	*timur	> +tim□r	'east'
(11)	*tikus	> +tikus	'mouse'
	*kurus	> +kurus	'lean, thin'
(12)	*laut	> +lawot	'sea'

For *-i- the pattern appears to be less clear. Not all *-iC combinations are sufficiently represented in the data to show a clear pattern. Some of them were scarce in Proto-Malay anyway. As table 3 shows, *-i- was lowered before *-t, *-n, *-r and *-h, between a nasal onset and *-s, and possibly also between a non-nasal onset and *-l. Examples of lowered *-i- after a non-nasal onset are given in (13).

*-i- remained unchanged before *-k and *-ŋ, and with more than chance frequency between a non-nasal onset and *-s (see (14)). Also before *-p and *-l, there is a tendency for lowering. In other positions no pattern emerges (see the Appendix for the data). The degree of lowering after a nasal onset tends to be maximal, for example, -ɛ-, but there are also some instances of *-i- becoming -e- (see (15)).

(13)	*t□ritip	> +t□ritep	'oyster'
	*isit 'gums'	> +iset	'female genitals'
	*kasih 'love, affection'	> *kaseh*	'love illicitly, commit adultery'
	*masin 'salty'	> +masen	'salty, brackish'
	*kudil	> +kudel	'scabies'
	*bibir	> +biber	'lip'
(14)	*itik	> +itik	'duck'
	*nipis	> +nipis	'thin'
	*t□biŋ	> +t□biŋ	'slope'
	*kuniŋ	> *kuniŋ*	'yellow'.
(15)	*tumit	> +tumɛt	'heel'
	*b□nih	> *b□n□h*[14]	'seed, offspring'
	*j□nis	> +j□nɛs[15]	'kind'
	*aŋin	> *aŋen*	'wind'
	*ha□ir	> +ha□□r	'fishy smell'.
(16)	*air	> +ayer	'water'.
	*kain	> *kayen*	'cloth'.

In open final syllables the high vowels did not change:

(17)	*t□bu	> *t□bu*	'sugarcane'
	*bahu	> *bahu*	'shoulder'
	*itu	> *itu*	'that, those'

[14] Also *j□rn□h* 'clear (water)' (< *j□rnih), but *k□meh* 'urinate' (< *k□mih).

[15] Also *m□naŋ□h* 'cry' (< +m□naŋɛs < *m□naŋis), but *maneh* 'sweet' (< +manes < *manis).

*padi		> *padi*	'unhusked rice'
*bu□i	'noise'	> *buŋi*	'noise (of engine etc.)'[16]
*p□ti	'case, chest'	> *piti*	'box'
*isi	'contents'	> *isi*	'contents, nature, inner meaning'.

3.4.3 Diphthongization of non-front vowels before *-s and *-t

The non-front vowels which at this stage reflected *-u- and *-a- developed a front offglide [y] in anticipation of, and adjusting to a following coronal obstruent (*-s and *-t). Examples:

(18) *tikus	$>_2$	⁺tikus	> ⁺tikuys	'mouse'
*haŋus	$>_2$	⁺haŋ□s	> ⁺haŋ□ys	'scorch in fire, blister'
*mulut	$>_2$	⁺mulot	> ⁺muloyt	'mouth, voice'
*laut	$>_2$	⁺lawot	> ⁺lawoyt	'sea'
*lumut	$>_2$	⁺lum□t	> ⁺lum□yt	'marine growth'
*b□ras			> ⁺brays	'husked rice'
*tunas			> ⁺tunays	'sprout, shoot'
*hampas 'waste'			> ⁺hampays	'husk'[17]
*h□mpas			> ⁺h□mpays	'throw down violently'
*ubat			> ⁺ubayt	'medicine'
*baŋat			> ⁺baŋayt	'quick'

This change can also be observed in (recent?) loanwords, such as *bay□* 'baht (Thai currency)', *joy□* 'jute' and *mutuboy□* 'motorboat'. Apparently these words have also undergone the next sound change.

Theoretically it is possible that the diphthongization took place before the lowering of the *high vowels in closed final syllables. But that would require unlikely conditions for the lowering of *-u-: lowering before ⁺-yt, no lowering before ⁺-ys, but extra lowering in both cases if the onset was a nasal. It makes phonetically more sense to condition both vowel lowering and diphthongization by the immediately surrounding sounds, for example, by assuming that diphthongization came after vowel lowering.

3.4.4 "Glottalization" of final *stops and *fricatives

After the diphthongization of non-front vowels before *-t and *-s, these latter sounds merged with the other word-final stops and with *-h respectively and became glottal:

---

[16] Although it is not clear what "*etc.*" means here, the explanation suggests a more specific sound than what is indicated by the SM etymon. It is possible that the change of the nasal has an onomatopoeic background.

[17] Hogan gives this word with this meaning only in the collocation *hapas kelamël* 'coconut husk' (< *hampas k□lambir). I assume that that meaning has to be understood as 'waste of a coconut'.

(19) *-t, *-k, *-p > -□
*-s, *-h > -h.

In (20) examples are given of the word-final stops, in (21) of the fricatives.

(20)	*t□ritip	>$_2$	+t□ritep	>	*t□rite□*	'oyster'
	*isit	>$_2$	+iset	>	*ise□*	'female genitals'
	*tumit	>$_2$	+tumɛt	>	*tum□□*	'heel'
	*itik			>	*iti□*	'duck'
	*hidup	>$_2$	+hidop	>	*hido□*	'live'
	*mulut	>$_{2,3}$	+muloyt	>	*muloy□q*	'mouth, voice'
	*laut	>$_{2,3}$	+lawoyt	>	*lawoy□*	'sea'
	*lumut	>$_{2,3}$	+lum□yt	>	*lum□y□*	'marine growth'
	*s□mut	>$_{2,3}$	+s□m□yt	>	s□*m□y□*	'ant'
	*masuk	>$_2$	+masok	>	*maso□*	'enter'
	*□amuk	>$_2$	+□am□k	>	□*am□□*	'mosquito'
	*asap			>	*asa□*	'smoke'
	*g□nap			>	*g□na□*	'enough, complete'
	*ubat	>$_3$	+ubayt	>	*ubay□*	'medicine'
	*baŋat	>$_3$	+baŋayt	>	*baŋay□*	'quick'
	*badak			>	*bada□*	'rhinoceros'
	*ba□ak			>	*ba□h□*	'much, many'
	*h□ntak	'stamp, pound'		>	*h□nta□*	'begin'.[18]
(21)	*kasih	>$_2$			*kaseh*	'love illicitly, commit adultery'
	*b□nih	>$_2$			*b□n□h*	'seed, offspring'
	*nipis			>	*nipih*[19]	'thin'
	*j□nis	>$_2$	+j□nɛs	>	*j□n□h*	'kind'
	*tujuh	>$_2$			*tujoh*	'seven'
	*p□nuh	>$_2$			*p□n□h*	'full'
	*tikus	>$_{2,3}$	+tikuys	>	*tikuyh*	'mouse'
	*haŋus	>$_{2,3}$	+haŋ□ys	>	*haŋ□yh*	'scorch in fire, blister'
	*tuah	>$_1$			*tuwah*	'luck'
	*l□mah			>	*l□mah*	'soft, weak'
	*b□ras	>$_3$	+brays	>	*brayh*	'husked rice'
	*tunas	>$_3$	+tunays	>	*tunayh*	'sprout, shoot'

[18] The reconstruction is dubious, given the difference between the SM and UL meanings.

[19] The following roots have an unexpected mid vowel in the final syllable: *tuleh* 'write' (< *tulis), *p□rleh* 'the Malaysian state of Perlis' (< *p□rlis), *gar□h* 'write, draw a line, strike a match' (< *garis).

The above mergers took place after the diphthongization of non-front vowels before *-s and *-t, for otherwise one would have expected forms like **asay and **l mayh instead of observed *asa* for 'smoke' and *l mah* for 'soft, weak'.

3.4.5 Change of final nasals into their corresponding voiceless stops

The crucial change which conditioned the later split of the root-final nasals had been the nasalizing effect of a nasal consonant on an immediately *following* vowel. As indicated above (see 3.4.2), this effect was already noticeable at the time of the lowering of *high vowels in closed final syllables. For current UL Hogan describes this phenomenon as an ideosyncratic feature of speech, which has no phonemic function. The articulatory effect of a nasal may even spread to the next syllable if the nasalised vowel is separated from the next vowel by a glide (see Hogan and Pattemore 1988: 27-28 for some examples). This progressive nasalization is the origin of the intervocalic nasal in words such as *mu aŋ* 'great-great-grandparent' (SM *moyaŋ* 'forefather'), and *ma ay* 'corpse' (SM *mayat*), and *p ŋa oh* 'paddle' (p. 135) alongside *p ŋayoh* (p. 125) from the root *kayoh* (< *kayuh).[20]

From Hogan's data it cannot be inferred whether this kind of nasalization was also operative in the past. In any case it seems that present day nasalization cannot be equated with the historical process (see the discussion around (30) in the next section). The closed final syllables, which existed at the time this process was operative, were –CVC, –NVC, –CVN, and –NVN (in which N symbolizes a nasal, C another consonant, and V a vowel which before –h or – could be followed by a palatal glide). If one assumes that regressive nasalization was not completely absent, the following prosodic picture emerges (||| = nasal articulation, --- = non-nasal articulation):

(22)	|||||||||||	||||||||--	---||||||	------
	-NVN	-NVC	-CVN	-CVC

In order to check the effect of this regressive nasalization and achieve a maximal contrast between final syllables with a nasal onset and those without one, the non-nasal articulation was protracted, resulting first in a preploded nasal coda:

(23)	---||||||		-----|||
	-CVN	>	-CV^{c}N

The preplosive element was homorganic with the final nasal. It may have been voiced, but that was phonemically irrelevant.

---

[20] The higher mid /o/ suggests that this nasalization is a recent phenomenon: after the nasal () one would have expected a lower mid **. Another example of such an innovative nasal is *naŋgri* 'country, city, area' (< *nagri). The nasal in *i aŋ* 'great-grandparent' (SM *eyaŋ*) is probably a matter of analogy with *mu aŋ* 'forefather'.

The next stage in the development was the complete reduction of the nasal part of the preploded final nasal, leaving the preplosive element as the new coda. At the end of the word they were voiceless now.

(24) -CV^{c}N > -CVC

Examples of unchanged final nasals are given in (25) and of stops from *nasals in (26).

(25)	*timun	$>_2$			*timun*	'cucumber'
	*gunuŋ	$>_2$			*gunuŋ*	'mountain'
	*aŋin	$>_2$			*aŋen*	'wind, air'
	*kuniŋ	$>_2$			*kuniŋ*	'yellow'
	*dmam		>	*dmam*		'fever'
	*snaŋ		>	*snaŋ*		'happy, contented'
(26)	*daun	$>_{1,2}$	+dawon	> +dawo^{t}n	> *dawot*	'leaf'
	*kbun	$>_2$	+kbon	> +kbo^{t}n	> *kbot*	'garden'
	*mbun	$>_2$	+mbon	> +mbo^{t}n	> +mbot	'dew'
	*tuntun	$>_2$	+tunton	> +tunto^{t}n	> +tuntot	'clasp, lead by holding upper arm'
	*jarum	$>_2$	+jarom	> +jaro^{p}m	> *jarop*	'needle'
	*untuŋ	$>_2$	+untoŋ	> +unto^{k}ŋ	> +untok	'profit'
	*kmbuŋ	$>_2$	+kmbo	> +kmbo^{k}ŋ	> +kmbokŋ	'swollen, distended'
	*kirim			> +kiri^{p}m	> *kirip*	'send, entrust'
	*kain	$>_2$	+kayen	> +kaye^{t}n	> *kayet*	'cloth'
	*tariŋ			> +tari^{k}ŋ	> *tarik*[21]	'tusk of wild pig'
	*guntiŋ			> +gunti^{k}ŋ	> +guntik	'scissors'
	*tajam			> +taja^{p}m	> *tajap*	'sharp'
	*rndam			> +rnda^{p}m	> +rndap	'soak, dip in, sit inwater'
	*uban			> +uba^{t}n	> *ubat*	'grey hair'
	*intan			> +inta^{t}n	> +intat	'diamond'
	*jamban	'latrine'		> +jamba^{t}n	> +jambat	'bridge, wharf'[22]
	*siaŋ	$>_1$	+siyaŋ	> +siya^{k}ŋ	> *siyak*	'(day)light'
	*tlajaŋ			> +tlaja^{k}ŋ	> +tlajak	'naked'
	*bintaŋ			> +binta^{k}ŋ	> +bintak	'star'

The change of *-n into –t (and consequently the parallel changes for the other nasals) came after the diphthongization of *-a- before *-t. Otherwise one would have expected forms like **ubayt instead of observed *ubat* for 'grey hair'.

[21] The word for 'sheep', *kamek* (< *kambiŋ) has an unexpected –*e*- in the final syllable.

[22] Traditional latrines were wooden constructions above a river.

3.4.6 Simplification of a medial *nasal followed by a homorganic *stop

An intervocalic sequence of a *nasal followed by a homorganic voiced stop lost the stop, an intervocalic sequence of a nasal followed by a homorganic voiceless stop lost the nasal. In other words:

(27) *-V_2NBV$_1$- > $^+$-V_2NV$_1$- (in which B symbolizes a homorganic voiced stop)
*-V_2NPV$_1$- > $^+$-V_2PV$_1$- (in which P symbolizes a homorganic voiceless stop)

Examples of the former are given in (28)[23], of the latter in (29).

(28)	*s□□ja	'sunset'	> *s□□a*	'late afternoon'
	*k□nduri	'ritual meal'	> *k□nuri*	'spirit feast'
	*l□mbu		> *l□mu*	'cow'
	*taŋga	'ladder, stairs'	> *taŋa*	'ladder, stair, boom of a dredge'.
(29)	*ku□ci		> *kuci*	'key'
	*inti	'kernel'	> *iti*	'filling, topping (in cake etc.)'
	*limpa	'spleen'	> *lipa*	'internal organ'
	*tumpul	>$_2$ $^+$tumpol	> $^+$tupol	'blunt'
	*maŋkuk	>$_{2,4}$ $^+$maŋkoq	> *makoq*	'bowl'

I assume that the process started with a change in articulation of the voiced nasal stop sequence. If it ever was a sequence of two phonemes, it became one phoneme now: a nasal with an oral non-nasal release. Such sounds have been described for several western Austronesian languages (for example, Rejang and Acehnese in Sumatra, Narum in Sarawak, and Mualang in West Kalimantan; see Coady and McGinn 1982: 443, Ladefoged and Maddieson 1996: 106; Blust 1997: 170, and Tjia 2007: 24-25). They have been reconstructed for Kerinci (Central Sumatra, Steinhauer 2002), Tunjung (Central Borneo), Lom (Bangka Island), Proto-Chamic (mainland Southeast Asia) and indeed UL (see Blust 1997: 170-171).

The result of this change in articulation was that many non-final syllables became open, which may have been the mould for other non-final syllables to become open as well. For the *-VNBV- sequences it was possible to shift the articulatory syllable boundary to the left and reduce the voiced stop to a short - voiced - oral release. For the *-VNPV- sequences with its articulatory break between the voiced nasal and the voiceless stop this was a less probable

[23] One effect of this voiced stop deletion is a difference in verbal morphology as compared to SM. A voiced stop at the beginning of a root is preceded by a nasal in SM, but replaced by one in UL if the root is prefixed by the "active" prefix {*m□N-*} or the actor prefix {*p□N-*}. Hogan's information (pp. 40-41) is minimal and confusing, and his very few examples do not suggest high productivity nor semantic transparency. Compare *dideh* 'to boil' (p. 41), *(m□)nideh* 'boiling' (p. 165); *gar□y□* 'scrape', *m□ŋar□y□* 'coconut scraper', *j□l□□* 'pound (curry ingredients, etc.)', *p□□□l□□* 'stone mortar'.

device. It seems likely that the articulatory contrast between the voiceless stop and the following vowel was too much part and parcel of the identity of the word for the stop to be dropped. Consequently it was the nasal which was dropped, either or not through a stage in which the preceding vowel was nasalized.

Above it was demonstrated that onset-driven nasalization was responsible for the preservation of nasal codas in final syllables. The forms such as those in (30) show that such nasalization did not have the same effect in penultimate syllables, either because it had never been there, or because it had stopped to be there. In spite of the nasal onset the nasal coda before a voiceless stop onset of the next syllable was not preserved:

(30)	*(m□)nanti	> (*m□*)nati*	'wait'
	*m□nantu	> *m□natu*[24]	'son/daughter-in-law'

The reduction of *-V_2NPV_1- > $^+$-V_2PV_1- also occurred when V_2 was a schwa, unless the consonant preceding it was *h. If there was no preceding consonant schwa was also dropped:

(31)	*t□mpat	>$_{3,4}$	$^+$t□mpayq	> *t□pay□*	'place'
	*t□ntu			> *t□tu*	'sure'
	*b□□ci			> *b□c^{h}i*[25]	'hate'
	*j□ŋkal			> $^+$j□kal	'span'
	*□mpat	>$_{3,4}$	$^+$□mpay□	> *pay□*	'four'
	*□ŋaw			> *kaw*	'you (sg.)'
	*hampas	>$_3$	$^+$hampayh	> *hapayh*	'husk'

But:

*h□mpas	>$_{3,4}$	*h□mpayh*	'throw down violently'
*h□ntak	>$_4$	*h□nta□*	'begin'.

These are the only instances of h□NP-, and in the last example the etymology is suspect. Yet, I believe that the persisting nasal is not accidental. Thanks to the [h-] onset the first syllable was not reduced to zero as in the case of $^+$pat 'four' **(<** *□mpat). Between this [h-] onset and a nasal coda, however, schwa could be reduced to such an extent that the nasal became syllabic: [$^+$h□mpas/hmpas, $^+$h□ntak/hntak]. And syllabic nasals stood a better chance of being preserved than nasals which merely functioned as a coda.

Some recent loanwords, such as *h□nda* 'Honda engine', apparently entered the language too late to be subject to the sound change.

---

[24] The verb *cuken*, *m□□uken* '<u>scrape out</u>' (<*cuŋkil*) is probably another example; but *m□□uken* instead of **m□□uŋken* may be a matter of analogy with the prefixless form.

[25] I have no explanation for the aspirated stop. Maybe it is an effect of the emotion.

The simplification of the voiced nasal-stop sequences post-dated the extra lowering of *-u- after a nasal onset. Otherwise *-u- would have become **-□- also after -N- < *-NB-. Compare:

(32) *s□m□□*	'ant'	<$_{2,4}$	*s□mut
samo□	'receive, welcome'	<$_{2,4,6}$	*sambut
j□ŋ□□	'peer/look at, visit'	<$_{2,4}$	*j□ŋuk
aŋo□	'nod, great'	<$_{2,4,6}$	*aŋguk
p□n□h	'full'	<$_{2}$	*p□nuh
suŋoh	'true'	<$_{2,6}$	*suŋguh
haŋ□yh	'scorch in fire, blister'	<$_{2,3,4}$	*haŋus 'scorched'
t□muyh	'pierce'	<$_{3,4,6}$	*t□mbus
tim□n	'cucumber'	<$_{2}$	*(m□n)timun
mot	'dew'	<$_{2,5,6}$	*□mbun
gun□ŋ	'mountain'	<$_{2,5}$	*gunuŋ
k□mok	'swollen, distended'	<$_{2,5,6}$	*k□mbuŋ.

As some of the examples show, the simplification of medial consonant clusters must also have post-dated the changes of the root-final nasals. Otherwise one would have expected final nasals to have been preserved, not only after -VNV- < *-VNV-, but also after –VNV- < *-VNBV-, and consequently forms like **mon* 'dew' and **k□moŋ 'swollen, distended' instead of observed *mot* and *k□mok*.

3.4.7 Change of final *-l into -n

With the word-final nasals practically removed, there was room for a new nasal to take their place. Word-final *-l, being phonetically closest to [n], filled the gap.

(33) *kudil	>$_{2}$	+kudel	> *kuden*[26]	'scabies'
*tumpul	>$_{2,6}$	+tupo	> *tupon*	'blunt'
*j□ŋkal	>$_{2,6}$	+j□kal	> *j□kan*	'span'.

3.4.8 Lateralization of *-r

Once word-final *-l had disappeared, *r , which in current UL no longer is a trill (if it ever was), but a flap or a retroflex vocoid, became lateral in word-final position. After *-i-, however, *-r tends to be dropped.

(34) *lapar			> *lapal*	'hungry'
*d□ŋar			> *d□ŋal*	'hear'
*bibir	>$_{2}$	+biber	> *bibel*[27]	'lip'

[26] Also *baten* 'bowl' (< *batil), *cuken* 'scrape out' (< *cuŋkil), *siken (gigi)* 'have severe toothache' (< *siŋkil 'pain'), *kayen* 'fishing line' (< *kail), but *paŋin* 'summon' (< *paŋgil), and *tuwin* 'lever' (< *tuil).

[27] See the Appendix.

*ikur	$>_2$	+ikor	> *ikol*	'tail'
*timur	$>_2$	+tim□r	> *tim□l*	'east'
*air	$>_2$	+ayer	> *aye*	'water'
*ha□ir	$>_2$	+ha□ɛr	> *ha□□*	'fishy smell'.

3.4.9 Other changes, differences with SM

3.4.9.1 Mid vowels

Only in some (recent?) loanwords does UL have mid vowels in non-final syllables, for example, in *roti* 'bread', *nori* 'lorry' (with change of initial consonant), *kopi* 'coffee' (alongside *kupi*), *ceti* 'Indian money lender', *h□nda* 'Honda engine', *b□r□ŋ* 'boring', *m□ken* 'Moken tribe', *l□t□ŋ* 'upper storey'. The syllable-final nasals in the last four words also suggest that these words are recent additions to the lexicon. *L□moŋ* 'large pond (as tin mine pool)' is possibly related to SM *lomboŋ* 'mine shaft, pit', but □ and *ŋ* are anomalous; it must be a loanword as well.

Many words which have a mid vowel in a non-final syllable in SM have a corresponding high vowel in UL. Compare the following SM and UL cognates (both in phonemic notation).

(35) SM	*b□ndera*	'flag, banner'	UL	*k□nira*	(with change of initial consonant)
	setan	'devil'		*sitat*	
	dosa	'sin'		*dusa*	
	enjin	'engine'		*ijen*	(with -n instead of **-t and -j- instead of **-n-)
	stokiŋ	'stocking'		*s□tukin*	(with -in instead of **-ik)
	motobot	'motorboat'		*mutuboy□*	

The first three loanwords (from Portuguese, Arabic and Sanskrit) are probably old, but the last three examples (all from English) suggest that mid vowels also in (at least some) recent foreign borrowings were assimilated to UL sound patterns.

In final syllables mid vowels in Malay and in foreign words remained mid, usually lower mid with some exceptions. Compare the following SM and UL forms.

(36) SM	*contoh*	'example'	UL	*cut□*[28]	
	bomor	'shaman'		*bum□l*	'doctor, shaman'
	g□romboŋ	'group'		*grum□ŋ*	'act as a mob'
	bom	'bomb'		*b□m*	'dynamite (fish)'

[28] After *-o-, *-h disappears. Compare *bud□* 'stupid', SM *bodoh,* and *jut□* 'destined (marriage partner)', SM *jodoh* 'marriage partner' (with a difference in the medial consonants of the latter pair).

SM	*c□kek*	'strangle'	UL	*c□k□□*	
	tiket	'ticket'		*tik□□*	
	seret	'drag'		*sir□□*	'turn around (with hand held up as in Manohra dance)'.[29]
	pohon	'tree'		*puhot*	
	leher	'neck'		*lihel*	
	roŋgeŋ	'paid dancing girl'[30]		*ruŋek*	'Malay type dance'

3.4.9.2 Retrograde nasalization

Above it was indicated that "retrograde nasalization avoidance" triggered preplosion of final nasals. This also happened with a few words which had a non-final syllable with a nasal coda followed by *-s-, if the preceding vowel was *-a-. Also here the nasal was finally lost completely. If the preceding vowel was *-u- or *-i-, there is no stop either.

(37)	*baŋsa	$>_5$ $^{+}$ba^{k}ŋsa	> *baksa*	'nationality, race'
	*baŋsat 'scoundrel, pauper'	$>_{3,4,5}$ $^{+}$ba^{k}ŋsay□	> *baksay□*	'wander'[31]
	*buŋsu	($>_5$? $^{+}$bu^{k}ŋsu)	> *busu*	'youngest (of children)'
	*insaŋ	($>_5$? $^{+}$i^{t}nsa^{k}ŋ)	> *isak*	'gills'

In a number of polysyllabic words there seems to be a retrograde, and apparently optional, nasalizing effect on an initial labial stop, if the intervening vowel was schwa. The data contain the following examples (UL compared with SM):

(38) SM	*binataŋ*	UL	*m□natak*, *b□natak*	'animal'
	b□nua		*m□nuwa*, *b□nuwa*	'world'
	p□ŋa□ciŋ		*m□ŋacik*, *p□ŋacik*	'button'
	p□ŋayuh		*m□ŋayoh*, *p□ŋayoh*, *p□ŋa□oh*	'oar'
	garut 'scrape'[32]		*m□ŋar□y□*	'coconut scraper'

[29] Semantically a questionable correspondence, although it is possible that 'drag' is the dancing term for the movement in question.

[30] According to the dictionaries which give such etymological information (Iskandar 1984; Teeuw 1996) the word is of Javanese origin. Traditionally a *ronggeng* had ritualised prostitute functions (see the meaning given in KBBI 1996). Iskandar 1984 just gives the meaning *penari perempuan* 'female dancer'. In UL the meaning seems to have shifted. According to Wongbusarakum (2007: 45) the dance was introduced by Malays from Penang.

[31] Given the apparent difference in word class and meaning one may doubt that SM *bangsat* is cognate with UL *baksay□*.

[32] The actor noun form of this verbal base would be *p□ŋgarut*, which, however, is not found in the dictionaries.

3.4.9.3 NON-FINAL SYLLABLE REDUCTION

Three types of initial syllable reduction can be observed: loss of initial *schwa (39), reduction (basically vowel reduction) of the antepenultimate syllable of polysyllabic words (40-41), complete loss of such a syllable (42-44).
The first type seems to be without exception.

(39)	*□laŋ	$>_{5}$	*lak*	'eagle, hawk, kite'[33]
	*□mas	$>_{3,5}$	*mayh*	'gold'
	*□nam	>	*nam*	'six'
	*(□)sa	>	*sa*	'one'
	*□mbun	$>_{2,5,6}$	*mot*	'dew'
	*□mpat	$>_{3,4,6}$	*pay□*	'four'

The second type is less systematic. Compare again UL with SM:

(40) SM	*bahasa*	UL	*basa*	'manners, language'
	binasa 'ruined'		*b□nasa*	'broken, spoilt'
	binataŋ		*b□natak*, *m□natak*	'animal'
	dunia		*d□nia*	'world'
	paŋlima 'military commander'		*p□lima*	'captain (of boat)'
	pusaka		*p□saka*	'inheritance'
	b□ri tahu		*bitahu*, *b□rtahu*	'tell, inform'
	b□rk□lahi		*b□rkahi*	'quarrel'
	p□luru 'bullet'		*prulu*	'bullet, arrow' (with metathesis < +pluru)
	s□b□raŋ		*s□rbak*	'other side of river' (with metathesis < +s□braŋ)
	b□laŋa		*b(□)laŋa*	'frying pan'
	p□landuk		*p(□)lano□*	'mouse-deer'

Parallel to the last two examples is the phenomenon that the prefix *b□r-* before stems which begin with *r-*, *l-* or a vowel, often appears as *b-*:

(41) SM	*b□rlari*	'run'	UL	*b(□r)lari*	'running'
	b□rlayar	'sail'		*b(□r)layal*	'sailing'
	ramay	'crowded, lively'[34]		*b(□)ramay*	'keeping festival, having fun'
	b□lajar			*b(□)lajal*	'study, learn'

[33] UL *h□lak* (< *h□laŋ) also occurs with the same meaning.

[34] There is no prefixed form in the SM dictionaries parallel to UL *b(□)ramay*.

There are not many mono-morphemic examples in the data of complete syllable loss (see 42), but there are several verbal forms in which the prefix *m□N-* is reduced to *N-* (those for which I found a Malay cognate are given in (43)).

(42) SM	*halia*		UL	*liya*	'ginger'
	har(□)ga			*r□ga*	'price, value'
	k□pala			*pala*	'head'
	utara			*tara*	'North'
	manusia	'human being'		*s□miya*	'man' (with metathesis < ⁺m□siya < ⁺m□n□siya)
	b□si paku			*(b□)sipaku*	'iron nail'

(43)[35] SM	*m□mbawa*	'bring'		UL *mawa*	'callawitchdoctor'
			m□mawa		'summon (a witch doctor)'
	m□ŋantuk	'sleepy'	*ŋato□*		'be sleepy'
			m□ŋato□		'sleepy'
	m□□□lam'	dive'	*□□lap*		'dive'
			m□□□lap		'dives'

Most conspicuous, however, is the complete reduction of initial syllables in reduplication:

(44) SM	*bajik*	'good, virtuous'	UL	*ji□-baji□*	'well' (p.105)
	buli-buli	'small bottle/jug'		*libuli*	'bottle'
	g□laŋ-g□laŋ	'tapeworm'		*lak-g□lak*	'centipede'
	kadaŋ-kadaŋ			*dak-kadak*	'sometimes'
	mata-mata	'spy'		*tamata*	'policeman'
	moyaŋ-moyaŋ			*□aŋ-mu□aŋ*	'ancestors'
	pagi-pagi	'early in the morning'		*gipagi*	'in the morning'.

3.4.9.4 PRESERVATION, ADDITION, AND LOSS OF *H

As compared to SM and most other varieties of Malay, UL shows archaic preservation of *h, both in initial and medial position. UL speakers must have been aware of this, since there are a few cases of apparent hypercorrection. On the other hand there are also three roots where *h- was lost. In (45) examples are given of roots where SM lost *h-, or preserved it optionally. Banjarese Malay, which is diagnostic in this respect has preserved *h- in these cases (see Hapip 1977).

[35] For many of these words both the forms with the long and the reduced prefix are entered in the dictionary, sometimes with slightly different semantic descriptions. It must be doubted that for instance *m□naŋ□h* glossed as 'cries' (p. 135) and *naŋ□h* 'cry' (p. 137) represent different morphological categories. The more so as similar formal differences are not accompanied by any semantic difference, while not infrequently the same forms in different places in the book have different semantic descriptions. Hogan's morphological description is insufficiently elaborate to get a clear picture of what is going on.

(45) SM	*abu*		UL	*habu*	'ashes'
	alu			*halu*	'pestle'
	(h)a□ir			*ha□□*	'fishy smell'
	apit			*hape□*	'press, squeeze, hold, clip'
	araŋ			*harak*	'charcoal'
	atap			*hataq*	'leaf roof'
	ayam			*hayap*	'chicken'
	□laŋ			*h□lak, lak*	'eagle, hawk, kite'
	(h)□mpas			*h□payh*	'throw down violently'
	(h)□ndap			*h□na□*	'secretly'
	(h)ujuŋ	'tip, point, end'		*hujok*	'cape'
	ulat	'caterpillar, maggot'		*hulay□*	'(small) caterpillar, worm'
	urup	'change money, barter'		*huro□*	'change, exchange (money)'.

Word-medially *-h- was preserved in the following cases:

(46) SM	*tiaŋ*	UL	*tihaŋ*	'mast'
	gua		*guha*	'cave'
	tua		*tuha*	'old'
	s□mua		*s□muha*	'all'.[36]

In (47) the roots are listed in which I assume that UL *h-* is a hypercorrection, and in (48) those in which *h- appears to have been lost:

(47) SM	*adat* 'custom, tradition'	UL	*haday□*	'custom, tradition, age'
	adam		*hadap* (also: *adap*) 'Adam'	
	adaŋ		*hadak*	'Adang Island'.

(48) SM	*haŋgaw*	UL	*aŋaw*	'reach out'
	hampar		*apal*	'spread out (planks, slates, etc.)'
	hari		*ari*	'day'.

3.4.9.5 Medial consonant clusters

The few cases of a syllable-final stop in a penultimate syllable appear to have been subject to the same process of glottalization as word-final stops:

(49) SM	*s□laksa*	UL	*s□la□sa*	'ten thousand'
	saksi		*sa□si*	'witness'
	napsu		*na□su*	'sexual desire'.

4 Discussion

In the dictionary some twenty items are marked as "M", for example, of Malaysian origin. At least seven of them have an Islamic flavour (such as *nabi* 'prophet', *lahtala* 'God (Muslim)' and *s□d□kah* 'Muslim gifts to the poor'), or they belong to the most frequently used SM words, such as *aku* 'I', *satu* 'one',

[36] Adelaar (1985) does not reconstruct any form for 'all'. Banjarese (Hapip 1977) has *samua*, however. So it is possible that here UL *-h-* is an innovation.

saja 'only', and *p□rgi* 'go'.

In matters pertaining to religion the UL community must therefore have been in contact with SM. Consequently, there must be sufficient awareness among UL speakers of regular sound correspondences between SM and their own language, which enables them to adapt recent loanwords according to the established patterns. The fact that some presumably recent loanwords such as *mutuboy□* 'motorboat' appear to have followed changes reconstructed as less recent is therefore not necessarily counter-evidence to the order of sound changes as proposed above. But further research in these matters is necessary.

The information available on UL morphology and syntax is concise, but potentially useful for comparative historical purposes. These are other fields, which deserve further study.

UL appears to have some archaic features, such as the preservation of *h- in words where most other Malay varieties lost it. UL retained the Austronesian root for 'dog', for example, *asu,* instead of a cognate of the widespread innovation *a□jiŋ*. For a proper evaluation, however, of the lexico-semantic peculiarities of UL one needs extended lexical data-bases for as many Malay(ic) varieties as possible.

In spite of its name UL has very little in common with the Malay varieties of other (semi-)nomadic "sea-people" (such as the *orang laut* described by Kähler (1960)). Neither does it show immediate correspondences with neighbouring Malay varieties, such as Kedah Malay and Patani Malay, as far as can be judged from Collins 1986. For most, if not all regular UL sound changes discussed above, parallel developments can be pointed out in other Malay varieties. The most striking change is the replacement of root-final nasals by their homorganic stops, in combination with the simplification of homorganic nasal-stop sequences (3.4.5 and 3.4.6 respectively). It is possible that this combination of changes was a mere local affair, but the similarities with patterns and reconstructed changes in languages of West Kalimantan and adjacent areas of Sarawak among both Land Dayak and Malayic Dayak languages (see Blust 1997: 157 and Tjia 2007) are too conspicuous not to look for further similarities. I hope to do so on another occasion.

References

Adelaar, K.Alexander. 1992. *Proto Malayic: reconstruction of its phonology and of part of its lexicon and morphology.* Canberra: Pacific Linguistics, Research School Of Pacific and Asian Studies, Australian National University. (Pacific Linguistics C-119).

Blust, R.A. 1997. "Nasals and nasalization in Borneo", *Oceanic Linguistics* 36: 149-179.

Coady, James and Richard McGinn. 1982. "On the so-called omplosive nasals of Rejang", in: Rainer Carle et al., *GAVÁ; Studies in Austronesian languages and cultures dedicated to Hans Kähler*, pp. 437-449. Berlin: Reimer. (Veröffentlichungen des Seminars für Indonesische und Südseesprachen

der Universität Hamburg 17).
Collins, James T. 1986. *Antologi kajian Dialek Melayu*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
Hapip, Abdul Djebar. 1977. *Kamus Banjar-Indonesia.* Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.
Hogan, David W. 1976. "Urak Lawoi' (Orang Laut)", in: W.A. Smalley, *Phonemes and orthography: language planning in ten minority languages of Thailand*, pp. 283-302. Canberra: Pacific Linguistics, Research School Of Pacific and Asian Studies, Australian National University. (Pacific Linguistics C-43).
Hogan, David W. 1978. "Urak Lawoi', Malay and Thai: some syntactic comparisons", *Te Reo Journal of the Linguistic Society of New Zealand* 21: 15-33.
Hogan, David W.1999. *Urak Lawoi'*. Prepared for publication by Stephen Pattemore. Muenchen: Lincom Europa. (Languages of the World/ Materials 268).
Hogan, David W., in collaboration with Stephen W. Pattemore. 1988. *Urak Lawoi'; basic structures and a dictionary*. Canberra: Pacific Linguistics, Research School Of Pacific and Asian Studies, Australian National University. (Pacific Linguistics C–109).
Iskandar, Teuku. 1984. *Kamus dewan*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
Kähler, Hans. 1960. *Ethnographische und linguistische Studien über die Orang darat, Orang akit, Orang laut und Orang utan an der Ostküste von Sumatra.* Berlin: Dietrich Reimer. (Veröffentlichungen des Seminars für Indonesische und Südseesprachen der Universität Hamburg, Band 2).
KBBI. 2001. *KBBI, Kamus besar bahasa Indonesia.* Edisi ketiga. Jakarta: Balai Pustaka.
Ladefoged, Peter and Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell.
Larish, Michael D. 1997. "Moklen-Moken Phonology: Mainland or Insular Southeast Asian Typology?", in: Cecilia Odé and Wim Stokhof (eds), *Proceedings of the Seventh International Conference on Austronesian Linguistics; Leiden 22-27 August 1994*, pp.125-150. Leiden University: Department of Languages and Cultures of South East Asia and Oceania, Projects Division, Leiden University.
Steinhauer, Hein. 2002. "More (on) Kerinci sound changes", in: K. Alexander Adelaar, and Robert Blust (eds), *Between worlds: linguistic papers in memory of David John Prentice,* pp 149-176. Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University. (Pacific Linguistics 529).
Teeuw, A. 1996. *Indonesisch-Nederlands woordenboek*. Vierde, herziene en uitgebreide druk, met medewerking van I. Supriyanto. Leiden: KITLV uitgeverij. (First edition 1990).

Tjia, Johnny. 2007. *A grammar of Mualang: an Ibanic language of Western Kalimantan, Indonesia*. Utrecht: LOT.

"Urak Lawoi", (http:// www2.unescobkk.org/elib/publications/2007Urak_Lawoi/Urak_Lawoi.pdf).

Wilkinson, R.J. 1959. *A Malay-English dictionary*. London: MacMillan.

Wongbusarakum, Supin. 2007. *The Urak Lawoi' of the Adang Archipelago, Thailand*. Bangkok: Themma Group.

Appendix

There are a few cases of *ia and *ua which seem to have undergone contraction rather than glide insertion:

(A1) SM	*k□tiak*	'armpit'	UL	*k□t□□*	
	buaya	'crocodile; log to put the mast in'		*b□ya*	'stem piece of boat'
	bahagian	'part, division		*'b□hag□t*	(< ⁺b□hagɛn)
	p□luaŋ	'opportunity' (base: *luang* 'hole, ….')		*p□rl□k*	'hole, crevice; opportunity' (< ⁺p□rloŋ)

In a number of words a final *nasal was preserved in spite of the fact that the syllable concerned had no original nasal onset:

(A2) SM	*bimbiŋ*	'lead by the arm'	UL	*bimiŋ*	'carry at side (arm extended)'
	dindiŋ	'wall'		*diniŋ*	
	tuŋgiŋ	'bend head down, back up'		*tuŋiŋ*	'lie on face with knees drawn up'

These reflexes suggest that for this particular phonotactic pattern (*-VMBiŋ) the simplification of the medial consonant cluster preceded the change of the final nasal into the homorganic voiceless stop, with the effect that the new nasal onset prevented that latter change from happening. If this is true, the reflex *kamek* 'sheep' <*kambiŋ presents a problem. But the reflex *-ek* instead of expected **-ik < *-iŋ poses a problem anyway. Preserved *-ŋ is incidentally found with other preceding vowels, again in spite of the fact that onset of the final syllable is or was not a nasal:

(A3) SM	*lindung*	'shelter, hide'	UL	*lin□ŋ*	(also with unexpected low vowel)
	timbaŋ	'weigh'		*timaŋ*	'weigh, ballast, balance'
	lambuŋ	'rise/jump up'		*lamuŋ*	'throw upwards'[37]
	lomboŋ	'mine shat, pit'		*l□moŋ*	'large pond (as tin mine pool)'
	loteŋ	'upper strorey'		*l□t□ŋ*	

[37] With unexpected high vowel. The expected form *lamok* has the meaning 'soar, go far (of a sound)'.

In some loanword too, final nasals are maintained, for example, in *s□l□ŋ* 'Ceylon tea', *s□priŋ* 'spring', *b□r□ŋ* 'boring for tin samples', *b□m* 'dynamite (fish)', *pam* 'pump', *m□ken* 'Moken tribe', *panton* 'pontoon'. A loanword which did denasalize the final nasal is *kaptat* 'captain'. Also in loanwords the nasal may be maintained in intervocalic consonant clusters. Most examples are from Thai, for example, *amph□r* 'district', *yiŋkali* 'prostitute', *baŋk□□* 'Bangkok'. Non-nasal word-final consonants may be retained as well, for example, *tep* 'tape recorder', *pukol* 'hour, time' (SM *pukul* … '… o'clock') and *s□tat* 'start'. Apparently the latter loanword is more recent than *mutuboy□* 'motorboat', *bay□* 'baht (Thai currency), *way□* 'wat (Thai temple)'.

In passing some roots have been mentioned which seem to be cognates with SM forms but differ in a minor phonemic aspect. Compare the following SM and UL forms (A4) show differences in voice, (A5) other consonantal differences, (A6) vowel differences):

(A4) SM	*pucuk*	'sprout'	UL	*buco□*	'summit, sprout'
	k□niŋ	'forehead'		*g□niŋ*	
	kutu	'louse'		*gutu*	
	jaguŋ	'maize'		*jakok*	
	b□ŋkok	'bent'		*piko□*	'curved, bent'
(A5) SM	*b□ndera*	'flag'	UL	*k□nira*	
	c□mburu	'jealous'		*k□mburu*	
	m□rpati	'dove'		*b□rpati*	
	s□mbu□i	'hide'		*s□luni*	'hide (intr.)'
				buni	'hide (trans.)'
	udaŋ	'prawn, lobster'		*hurak*	
	p□rtanda	'indications'		*p□rnana*	'indications, bearings'
	s□□ja	'sunset'		*s□□a*	'later afternoon'.
				h□□n	'early evening'[38]
	s□□jata	'weapon'		*s□nata*	
	p□□jara	'prison'		*p□jara*	
	geleŋ (*k□pala*)	'shake (head)'		*ilik* (*pala*)	
(A6) SM	*m□ntah*	'half-cooked, unripe'		*matah*	
	s□gi	'aspect'		*sagi*	'angle, corner (of table)'
	s□maŋat	'spirit'		*s□m□ŋay□*	
	guliŋ	'roll (an object)'		*gerlik*	
	d□ŋan	'with, and'		*jaŋan*	(with additional change of consonant)

[38] In the dictionary both forms refer to each other. The glosses differ only because the forms are far apart.

l□mbar	'sheet'	*lamal*	classifier for leaves and pieces of paper or cloth
□nlaŋ	'open (eyes)'	*□□lak*	'open eyes'.

Finally I present a list of roots which show an unexpected or too unpredictable reflex of the vowel in the final syllable. Other such cases, mentioned in footnotes above, complement this list. The leftmost column represents the regular reflex, in the central one the SM cognates are given of the deviating UL forms which are presented with their meaning in the column on the right.

*-ir /-el	*sisir* 'comb'	(*p□hape□*) *sise*	'hair-slide'
	air	*aye*	'water'
*-it/-e□	*cicit*	*cic□□*	'great-grandchild'
	s□mpit	*s□p□□*	'narrow'
*-ap/-a□	*s□rgap*	*s□go□*	'leap on'
	(problematic also because of the absence of *–r–* in UL)		
	s□bap	*s□be□*	'because'
*-at/-ay□	*aŋkat*	*ake□*	'lift, erect, carry (with one's hand)'
*-up/-o□	*c□lup*	*c□lu□*	'dip, dye'
*-ut/-oy□	*garut*	*gar□y□*	'scrape'
	t□rk□jut	*t□rk□juy□*	'startled'
	j□mput 'take between one's fingers'	*j□puy□*	'snatch away'
*-un/-ot	*ayun*	*ayut*	'rock, swing'
*-ur/-ol	*ubur-ubur*	*bul-ubul*	'white jelly-fish'
	s□mbur	*s□bul*	'spray'
	(problematic also because of UL *–b–* instead of expected *–m–*)		
	aŋgur	*aŋun*	'grape'
	(also UL *–n* instead of expected *–l* is problematic)		
*-uŋ/-ok	*arungan* 'sea-route'	*ar□k*	'open sea, ocean'
	t□ruŋ	*tr□k*	'eggplant'
	kutuŋ 'cut off'	*kut□k*	'cut into sections'.