

PERBANDINGAN KUANTITATIF BAHASA MIRI DENGAN BAHASA MELAYU

Noriah Mohamed

Abstrak

Bahasa Miri sejenis bahasa Austronesia yang dituturkan oleh orang Miri. Bahasa ini dikelompokkan ke dalam bahasa-bahasa jenis Sarawak di bawah subkelompok bahasa Baram Lembah yang juga mengandungi bahasa-bahasa seperti Dalik, Berawan, Kiput, Narum, Lelak dan Lementing, Belait dan Tutong. Makalah ini membincangkan bahasa ini dari segi linguistik komparatif dengan cara meninjau dan menyenaraikan kosa kata asas bahasa ini dan membandingkan kosa kata asas bahasa ini dengan bahasa Melayu. Tinjauan dan penyenaian kosa kata asas akan memanfaatkan senarai 100 dan 200 kata Morris Swadesh. Perbandingan pula akan cuba menentukan tingkat kekerabatan bahasa Miri dengan bahasa Melayu. Makalah ini adalah sumbangan kecil terhadap langkah mendokumentasikan unsur linguistik bahasa Miri, memandangkan bahasa ini semakin pupus oleh sebab dari segi penggunaan, penuturnya (terutama generasi muda), sudah beranjak ke bahasa Melayu Sarawak atas sebab-sebab sosiolinguistik (lihat Bibi Aminah Abdul Ghani dan Abang Ahmad Ridzuan 1993).

Abstract

Miri language is a type of Austronesian language which is spoken by a group of speakers known as Miri. It is grouped under Lower Baram languages, a subgroup of Sarawakan languages, which also comprises languages such as Dalik, Berawan, Kiput, Narum, Lelak and Lementing, Belait and Tutong. This article will offer a linguistic comparative discussion of this language by observing and identifying the basic vocabulary of the language, and comparing the basic vocabulary of the language with Malay language. The observation and identification of basic vocabulary will employ the 100 and 200 word lists

by Morris Swadesh. The comparison between the two languages will determine the genetic relationship between Miri and Malay. This article represents a small contribution towards the process of documentation of the linguistic aspect of Miri language in view of its verge of extinction due to the fact that its speakers (especially the younger generation) have shifted to the Sarawak Malay for sociolinguistic reasons (Bibi Aminah Abdul Ghani dan Abang Ahmad Ridzuan 1993).

PENDAHULUAN

Bahasa Miri merupakan sejenis bahasa Austronesia yang dituturkan oleh sekelompok penutur yang digelar orang Miri. Menurut Tunku Zainah (1978: 24–25), berdasarkan sebutan orang tempatan, orang Miri digelar orang Mirek (sebutannya mighek [mi(e)]¹). Di kalangan penduduk tempatan juga, istilah orang Miri boleh membawa dua pengertian. Pertama, istilah orang Miri digunakan untuk merujuk kumpulan etnik yang khusus dan kedua, untuk merujuk penduduk di pekan Miri dan kawasan sekitarnya, tidak kira sama ada mereka penduduk asal atau pendatang ke kawasan berkenaan. Dalam makalah ini istilah orang Miri digunakan untuk merujuk satu kumpulan etnik yang khusus. Dalam hal ini, kumpulan etnik Miri dianggap sebagai salah satu kumpulan yang tergolong dalam kelompok bumiputera di Sarawak, dan statusnya sama seperti kumpulan etnik bumiputera yang lain seperti Melayu, Iban, Melanau, Bidayuh dan sebagainya.

Dari segi penampilan, orang Miri sukar dibezakan daripada orang Melayu kerana kebanyakan orang Miri telah memeluk agama Islam dan menyerap identiti Melayu dan secara am diterima oleh masyarakat Sarawak sebagai orang Melayu. Orang Miri hanya dapat dibezakan daripada orang Melayu menerusi dua aspek, iaitu sejarah dan bahasa mereka yang disebut bahasa Miri.

Bahasa Miri masih dituturkan dalam komuniti orang Miri, walaupun penggunaannya semakin berkurangan. Mungkin tiga atau empat generasi akan datang bahasa Miri akan pupus kerana generasi muda komuniti ini tidak lagi menggunakannya. Hal ini berlaku kerana dari segi fungsi, bahasa Miri terhad domain penggunaannya. Di samping itu, bahasa ini tidak diketahui oleh komuniti yang lain, jadi tidak dapat digunakan sebagai sarana komunikasi sekiranya komunikasi melibatkan komuniti lain. Hal ini juga disebabkan orang Miri khususnya yang berpendidikan, lebih betah menggunakan dialek Melayu Sarawak.

Menurut Tunku Zainah (1978: 26), berdasarkan pengelompokan yang dilakukan oleh Blust (1972), bahasa Miri boleh dikelompokkan ke dalam

bahasa-bahasa jenis Sarawak di bawah subkelompok Baram Lembah yang mengandungi bahasa Dalik, bahasa Berawan, bahasa Kiput, bahasa Narum, bahasa Lelak dan bahasa Lementing, bahasa Belait dan bahasa Tutong. Dalam pengelasan bahasa-bahasa jenis Sarawak oleh Grimes (1996: 4), bahasa Miri dikelompokkan sebagai salah satu dialek bahasa Narum.

Secara am dapat dikatakan bahawa bahasa Miri dan bahasa Melayu termasuk dalam rumpun bahasa yang lebih besar, iaitu rumpun bahasa Austronesia atau dengan kata lain kedua-dua bahasa ini mempunyai hubungan kekerabatan. Walau bagaimanapun, dari segi kefahaman menyaling antara penutur, didapati hubungan kekerabatan yang dimaksudkan sudah agak jauh, kerana penutur kedua-dua bahasa ini sudah tidak saling memahami kecuali mempelajari bahasa penutur yang seorang lagi. Untuk menentukan setakat mana hubungan kekerabatan antara kedua-dua bahasa ini, makalah ini membincangkan perbandingan kosa kata asas kedua-dua bahasa berkenaan dengan menggunakan kaedah perbandingan kuantitatif.

KAEDAH PERBANDINGAN

Kaedah perbandingan ialah cara yang digunakan untuk membandingkan dua atau lebih daripada dua dialek atau bahasa. Kaedah ini secara am digunakan untuk menyingkap perkaitan kekerabatan antara bahasa pada zaman prasejarah, yakni sebelum wujud catatan bertulis. Disebabkan tidak wujud catatan bertulis atau rakaman pertuturan pada zaman prasejarah, maka untuk menelusuri perkaitan kekerabatan yang dimaksudkan, digunakan dialek atau bahasa semasa, yang wujud pada hari ini. Dialek atau bahasa semasa digunakan kerana diandaikan wujud hasil perkembangan bahasa terdahulu. Dengan kata lain, terdapat bahasa yang lebih tua² yang dianggap memencarkan dialek atau bahasa yang lebih muda.³ Walau bagaimanapun, disebabkan perkembangan masa, maka dialek atau bahasa yang lebih muda mungkin berubah atau tidak sama lagi dengan bahasa induk yang menjadi pemencar itu. Oleh sebab itu, selain untuk mendapatkan gambaran tentang bahasa induk itu, perbandingan secara khusus digunakan untuk meneliti perubahan terhadap dialek dan bahasa yang dipencarkan daripada bahasa induk. Oleh hal yang demikian bidang perbandingan ini dianggap bidang yang mempelajari data daripada suatu bahasa atau lebih, sekurang-kurangnya dalam dua periode. Data daripada suatu bahasa dari dua periode atau lebih itu diperbandingkan secara cermat untuk memperoleh kaedah-kaedah perubahan yang terjadi dalam bahasa itu (diubahsuai daripada Gorys Keraf 1984: 22).

Perbandingan bahasa dalam makalah ini tidak bertujuan untuk memperoleh gambaran bahasa induk, tetapi lebih kepada usaha meneliti perubahan dua bahasa; bahasa Melayu dan bahasa Miri, dan dengan demikian dapat ditentukan lama waktu,⁴ abad atau tahun pisah dan tingkat kekerabatan⁵ kedua-dua bahasa ini.

Dalam usaha menentukan kedua-dua perkara di atas, makalah ini akan memanfaatkan perbandingan bersifat kuantitatif, atau secara khusus dikenali sebagai leksikostatistik dan glotokronologi. Walau bagaimanapun, sebagai latar belakang akan dihuraikan secara am kaedah kualitatif atau kaedah yang memeriksa unsur bahasa yang diretensi (terawet) dan berinovasi.

KAEDAH KUALITATIF

Kaedah kualitatif atau pemeriksaan unsur retensi dan unsur inovasi meninjau, mengenal pasti dan membandingkan kosa kata berkognat dalam bahasa-bahasa yang dianggap sekerabat. Kosa kata berkognat bermaksud kosa kata yang sama atau mirip dalam bahasa-bahasa sekerabat yang wujud disebabkan kosa-kosa kata berkenaan berasal usul daripada bahasa induk yang sama. Kosa kata berkognat mungkin sama atau mirip. Misalnya kata **daging* dalam bahasa Austronesia purba yang disebut sebagai *daging* dalam bahasa Melayu dan bahasa Miri merupakan kosa kata berkognat yang sama bentuknya. Kata **abu* dalam bahasa Austronesia purba yang disebut sebagai *abu* dalam bahasa Melayu dan *aboh* dalam bahasa Miri merupakan kosa kata berkognat yang mirip sahaja.

Kosa kata berkognat dalam bahasa-bahasa turunan yang sama dengan kosa kata bahasa induk dianggap kosa kata yang dikekalkan. Kosa kata berkognat yang mirip wujud oleh sebab berlaku perubahan atau inovasi bahasa.

Terdapat dua jenis inovasi bahasa, yang tak tulen dan yang tulen. Salah satu sebab inovasi tak tulen berlaku mungkin kerana kesilapan dalam pengucapan atau penulisan sesuatu kata. Misalnya dalam *Hikayat Hang Tuah*, *Hang Lekir* dan *Hang Lekiu* sebenarnya orang yang sama tetapi disebabkan kesilapan penulisan dalam hikayat Melayu lama, maka wujud dua tokoh yang mewakili orang yang sama. Walau bagaimanapun, inovasi akibat kesilapan jarang-jarang berlaku. Inovasi tak tulen juga mungkin berlaku kerana pencemaran terhadap kosa kata akibat proses peminjaman kata daripada bahasa-bahasa lain. Inovasi akibat peminjaman kosa kata amat sukar dibezakan daripada inovasi tulen kerana inovasi jenis ini kerap berlaku, konsisten, dan mungkin berlaku dalam kalangan bahasa sekerabat. Walau bagaimanapun, inovasi tak tulen yang disebutkan ini tidak diambil kira sewaktu meneliti inovasi kosa kata

kerana kedua-duanya dianggap bukan inovasi tulen. Yang dimaksudkan dengan inovasi tulen kosa kata ialah inovasi yang berlaku disebabkan peredaran atau perubahan masa. Hal ini berlaku kerana terdapat andaian bahawa bahasa-bahasa sekerabat akan memperbaharui satu atau lebih daripada satu kosa katanya pada suatu waktu dengan alasan tertentu. Berbeza daripada dua bentuk inovasi sebelum ini, inovasi tulen sepatutnya berlaku secara sistematik atau dengan kata lain berlaku secara teratur. Perubahan yang teratur bermaksud perubahan yang bukan sahaja berlaku dalam satu kosa kata sahaja tetapi berlaku juga dalam kosa kata lain yang memiliki susunan fonem yang sama.

Melalui kaedah memeriksa unsur, kosa kata berkognat daripada bahasa-bahasa yang dibandingkan dikumpulkan. Kata-kata boleh merupakan kata dasar atau kata budaya asalkan kata-kata berkenaan memperlihatkan kesepadanan. Selepas ditentukan bentuk yang terawet dan yang berinovasi, bahasa-bahasa dikelompokkan berdasarkan data yang memperlihatkan kesepadanan yang lebih mirip bagi setiap kelompok. Dengan kata lain, jika dua atau lebih daripada dua bahasa memperlihatkan ciri-ciri retensi dan inovasi yang sama, maka bahasa-bahasa berkenaan tergolong dalam kelompok yang sama. Sila lihat contoh dalam data Jadual 1.

Berdasarkan kaedah pemeriksaan unsur retensi dan inovasi keempat-empat bahasa yang dibandingkan, maka secara am dikatakan bahasa Melayu mempunyai hubungan yang lebih dekat dengan bahasa Iban berbanding

Jadual 1 Perbandingan Kosa Kata Empat Bahasa Austronesia

Bahasa Austronesia Purba ⁶	Bahasa Melayu	Bahasa Iban	Bahasa Kadazan	Bahasa Dusun
*ka/*ka'en/ka'i/pangan*	Makan	makan	akan	akan
*(t)elur	Telur	telu	tontohu	tontolu
*bulu	Bulu	bulu	wulu	vuhu
*apuj	Api	api	tapui	tapui
*kaki	Kaki	kaki	gakod	takod
*andang/pandang	Panjang	panjai	anaru	ana'u
*b/ah/aru/*beru	Baru	baru	wagu	agu

Sumber: Dikutip secara rawak daripada Noriah (1999).

dengan bahasa Kadazan dan bahasa Dusun. Oleh sebab itu bahasa Melayu akan dikelompokkan dengan bahasa Iban sementara bahasa Kadazan dan bahasa Dusun yang memperlihatkan persamaan yang lebih dekat membentuk satu lagi kelompok bahasa sekerabat.

Jumlah bahasa dan kosa kata yang dipilih untuk perbandingan perlu diperluas atau ditambah sehingga sampai ke tahap yang benar-benar dapat dibezakan hubungan dekat, jauh atau di atas, di bawah tatatingkat dalam rajah pohon salasilah kekeluargaan bahasa.

KAEDAH KUANTITATIF

Kaedah kedua yang digunakan dalam aktiviti perbandingan bahasa ialah kaedah kuantitatif yang lebih dikenali sebagai kaedah leksikostatistik. Kaedah leksikostatistik sering dikaitkan dengan kaedah glotokronologi. Leksikostatistik merujuk kaedah yang digunakan untuk menentukan peratus kata sekerabat dengan cara menghitung jumlah kata berkognat. Glotokronologi pula merupakan kaedah yang menggunakan formula tertentu untuk menghitung usia bahasa sekerabat (dalam ribuan tahun), dan menentukan bilakah berlakunya perpisahan bahasa-bahasa sekerabat berkenaan daripada bahasa induk. Lazimnya, kaedah glotokronologi mengikuti kaedah leksikostatistik kerana glotokronologi memanfaatkan data hasil perbandingan leksikostatistik.

Leksikostatistik dan glotokronologi menggunakan kaedah penelitian unsur retensi dan unsur inovasi. Perbezaan kaedah leksikostatistik dan glotokronologi berbanding dengan kaedah kualitatif dapat dilihat dari segi jenis kosa kata yang digunakan untuk perbandingan. Kaedah leksikostatistik dan glotokronologi menggunakan kosa kata dasar sebagai asas perhitungan. Berdasarkan pengasasnya, khususnya Dyen (1975) dan Gudschinsky (1980), kosa kata dasar yang dimaksudkan dalam kaedah ini diandaikan memiliki dua sifat. Pertama, kosa kata dasar dianggap sukar berubah dibandingkan dengan kosa kata budaya dan lazimnya bersifat sejagat, yakni sepatutnya terdapat dalam semua bahasa di dunia. Kosa kata ini meliputi kata ganti, bilangan, anggota badan, alam sekitar, alat kelengkapan seharian, dan sebagainya. Kedua, kosa kata dasar dianggap memiliki sifat awet atau retensi yang malar setiap seribu tahun. Dengan kata lain, kosa kata dasar dianggap dapat bertahan secara malar setiap seribu tahun dalam peratusan tertentu. Seribu tahun berikutnya kosa kata dasar akan bertahan dengan peratusan yang sama. Berdasarkan kajian terhadap tiga belas bahasa Indo-Eropah yang memiliki naskah bertulis, didapati kosa kata dasar dapat bertahan

sekitar 86.4% sehingga 74.4% setiap seribu tahun.

Kosa kata dasar yang digunakan dalam kaedah ini terbatas jumlahnya. Senarai kosa kata dasar yang paling lazim digunakan dicadangkan oleh Morris Swadesh,⁷ yang mengandungi seratus atau dua ratus patah perkataan. Senarai seratus kosa kata menggunakan harkat retensi 86% (0.86) dan senarai dua ratus kosa kata pula menggunakan harkat retensi 81% (0.81).

Kaedah ini dimulakan dengan meneliti kosa kata dasar kedua-dua bahasa yang dibandingkan dan menentukan kosa kata yang dapat digunakan atau yang perlu ditolak. Kosa kata yang harus ditolak ialah kata pinjaman, kata kosong (ada dalam satu bahasa tetapi tidak ada dalam bahasa yang satu lagi) dan kata majmuk. Jika terdapat kata terbitan, imbuhan perlu disingkirkan dan digunakan hanya kata dasar sahaja. Jika kata dasar berkenaan sudah wujud dalam kata lain sebelumnya, maka kata dasar tersebut perlu disingkirkan. Contohnya apabila membandingkan bahasa Melayu dengan bahasa Miri, terdapat kata *berjalan* sebagai terjemahan kepada *to walk*. Apabila imbuhan *ber* disingkirkan, maka kata dasar yang tinggal ialah *jalan*. Disebabkan perkataan *jalan* sudah muncul sebagai salah satu kosa kata yang dibandingkan sebelum ini, iaitu untuk *road*, maka kata *jalan* yang merujuk *to walk* tidak perlu dibandingkan lagi.

Setelah selesai ditolak semua kata yang tidak diperlukan, baharulah ditetapkan jumlah kata yang boleh dibandingkan. Langkah seterusnya ialah membandingkan kosa kata dan menentukan taraf kognat antara kosa kata. Untuk menentukan sama ada pasangan kosa kata itu kata berkognat atau tidak, kriteria berikut diambil kira.

- a. Pasangan kata itu identikal atau sama seratus peratus.
- b. Pasangan kata itu memiliki bentuk dan makna yang sepadan / mirip, yakni memiliki fonem / bunyi yang sama pada posisi yang sama atau memiliki fonem / bunyi yang mirip tetapi dapat diterangkan bahawa kemiripan itu akibat pengaruh lingkungan / berlaku secara konsisten.

Hasil perbandingan akan memperlihatkan jumlah kata berkognat di antara kedua-dua bahasa. Untuk mendapatkan peratus kata kerabat (c), jumlah kata berkognat harus dibahagikan dengan jumlah kata yang dibandingkan (n), yakni jumlah kata selepas ditolak kata-kata yang tidak diperlukan.

Setelah ditentukan peratus kata sekerabat menerusi kaedah leksikostatistik, maka langkah seterusnya adalah menerapkan kaedah glotokronologi untuk menentukan waktu pisah kedua-dua bahasa yang dibandingkan. Rumus yang digunakan untuk menghitung waktu pisah antara dua bahasa

berkerabat yang dikemukakan oleh Robert B. Lees⁸ adalah seperti yang berikut.

$$T = \frac{\log c}{2 \log r}$$

dengan

T = Waktu pisah dalam ribuan tahun / abad

C = Peratus kata sekerabat

R = Kadar bertahan kosa kata dalam 1000 tahun dan angka yang digunakan ialah 86% atau 0.86 (untuk daftar 100 kata) atau 81% atau 0.81 (untuk daftar 200 kata)

Log = Logaritma asli

Logaritma yang digunakan dalam kaedah ini ialah logaritma asli seperti yang dipaparkan dalam Jadual 2.

Jadual 2 Logaritma Asli

N	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.1	-2.203	-2.207	-2.120	-2.040	-1.966	-1.897	-1.833	-1.772	-1.715	-1.661
.2	-1.609	-1.561	-1.514	-1.470	-1.427	-1.386	-1.347	-1.309	-1.273	-1.238
.3	-1.204	-1.171	-1.139	-1.109	-1.079	-1.050	-1.022	-.994	-.968	-.942
.4	-.916	-.892	-.868	-.844	-.821	-.799	-.777	-.755	-.734	-.713
.5	-.693	-.673	-.654	-.635	-.616	-.598	-.580	-.562	-.545	-.528
.6	-.511	-.494	-.478	-.462	-.446	-.431	-.416	-.400	-.386	-.371
.7	-.357	-.342	-.329	-.315	-.301	-.288	.274	-.261	-.248	-.236
.8	-.223	-.211	-.198	-.186	-.174	-.163	-.151	-.139	-.128	-.117
.9	-.105	-.094	-.083	-.073	-.062	-.051	-.041	-.030	-.020	-.010

Sumber: Gudschinsky (1964), dan Swadesh (1971).

Walau bagaimanapun, nilai t yang diperoleh menerusi formula di atas menimbulkan keraguan kerana tidak mungkin perpisahan antara dua bahasa berlaku tepat pada tahun-tahun tertentu. Hal ini demikian kerana bahasa berpisah secara beransur-ansur dalam julat waktu. Oleh hal yang demikian harus dihitung pula jangka kesalahan (ralat) menggunakan rumus kesalahan standard. Hitungan jangka kesalahan dalam makalah ini didasarkan pada jangka kesalahan standard⁹ $7/10$. Untuk memperoleh jangka kesalahan ini digunakan rumus berikut.¹⁰

$$S = \sqrt{\frac{c(1-c)}{n}}$$

dengan

- S = Kesalahan standard
- C = Peratus kata sekerabat dua bahasa yang dibandingkan
- N = Jumlah kata yang dibandingkan

Selanjutnya dihitung pula nilai c yang baru, iaitu $c1$ menggunakan rumus yang berikut:

$$c1 = c + s$$

Setelah diperoleh c yang baharu, iaitu $c1$ maka dapat ditentukan waktu pisah yang baharu, iaitu $t1$ menggunakan rumus berikut:

$$t1 = \frac{\log c1}{2 \log r}$$

Untuk memperoleh jangka kesalahan maka formula $t - t1$ digunakan. Penentuan jangka waktu pisah antara dua bahasa berkerabat dilakukan dengan cara menambah atau menolak waktu pisah lama (t) dengan jangka kesalahan, iaitu $t + t1$ dan $t - t1$.

Berdasarkan peratus kata sekerabat dan pengiraan abad pisah, dapatlah ditentukan tingkat kekerabatan antara bahasa seperti yang ditunjukkan yang ditunjukkan dalam Jadual 3.

Makalah ini selanjutnya akan menerapkan kaedah leksikostatistik dan glotokronologi, terhadap perbandingan bahasa Melayu dan bahasa Miri.¹¹

Jadual 3 Jadual Tingkat Kekerabatan Bahasa

Tingkat Kekerabatan ¹²	Abad Pisah	Peratus Kata Berkerabat
Bahasa	0 – 5 abad	100 – 81
Keluarga	5 – 25 abad	81 – 36
Rumpun	25 – 50 abad	36 – 12
Mikrofilum	50 – 75 abad	12 – 4
Mesofilum	75 – 100 abad	4 – 1
Makrofilum	100 abad ke atas	kurang daripada 1

Sumber: Gudschinsky (1964: 621).

PENENTUAN PERATUS KATA SEKERABAT ANTARA BAHASA MELAYU DENGAN BAHASA MIRI

Daftar 100 Kosa Kata Dasar Bahasa yang Dikaji

Bahasa Melayu

1.semua 2.abu 3.kulit pokok 4.perut 5.besar 6.burung 7.gigit 8.hitam
 9.darah 10.tulang 11.tetek 12.terbakar 13.kuku 14.awan 15.sejuk/dingin
 16.datang 17.mati 18.anjing 19.minum 20.kering 21.telinga 22.tanah
 23.makan 24.telur 25.mata 26.lemak 27.bulu 28.api 29.ikan 30.terbang
 31.penuh 32.kaki 33.beri 34.baik 35.hijau 36.rambut 37.tangan 38.kepala
 39.dengar 40.jantung 41.tanduk 42.aku 43.bunuh 44.lutut 45.tahu 46.daun
 47.berbaring 48.hati 49.panjang 50.kutu 51.lelaki/laki-laki 52.banyak
 53.daging 54.bulan 55.gunung 56.mulut 57.nama 58.leher 59.baru
 60.malam 61.hidung 62.tidak 63.satu 64.orang 65.hujan 66.merah 67.jalan
 68.akar 69.bulat 70.pasir 71.ungkap/sebut 72.lihat 73.benih 74.duduk
 75.kulit 76.tidur 77.kecil 78.asap 79.berdiri 80.bintang 81.batu 82.matahari
 83.berenang 84.ekor 85.itu 86.ini 87.kau 88.lidah 89.gigi 90.pokok
 91.dua 92.berjalan 93.panas 94.air 95.kami 96.apa 97.putih 98.siapa
 99.perempuan 100.kuning

Bahasa Miri

1.semua 2.aboh 3.ulait ajuh 4.jen 5.ayak 6.manok 7.gutap 8.mitam
 9.daghek 10.tulang 11.titait 12.mekuh 13.ilaoh 14.awan 15.sajuk 16.dusuk

- 17.mataih 18.au 19.sap 20.magha 21.telingah 22.tanak 23.umin 24.tejuk
 25.matah 26.lemaoh 27.buluh 28.igun 29.hutak 30.tulut 31.fenaok/henaoh
 32.ha'ah 33.tuke 34.jek 35.hijau 36.fuk 37.tujuk 38.uhaok 39.nyait
 40.jantung 41.ugang 42.koi 43.binaoh 44.tut 45.ta'an 46.da'un 47.lupaek
 48.atai 49.bet 50.gutaoh 51.lakaih 52.akik 53.dagin 54.teka 55.lat
 56.munong 57.aden 58.bataok 59.baghuk 60.lem 61.singot 62.oumlok
 63.seh 64.jemulon 65.ujen 66.ma'ang 67.lahau 68.ghamot/gamot
 69.belileng 70.ghaet 71.matang 72.ni'an 73.tulaek 74.utaok 75.ulait
 76.makel 77.diik 78.sap 79.masang 80.hatak 81.bataoh 82.matah daloh
 83.meta 84.ihaok 85.jih 86.itoi 87.aok 88.jilek 89.nyipan 90.ajuh 91.debeh
 92.melahau 93.la'oh 94.fek 95.amaeh 96.upaeh 97.hutaek 98.sai 99.tughai
 100.(k)uneng

Perbandingan kosa kata dasar dan peratusnya

Kosa kata yang ditolak:

- 3. kulit pokok - ulait ajuh (kata majmuk)
- 82. matahari - mata daloh (kata majmuk)
- 92. berjalan - lahau (sebab kata dasar jalan sudah wujud dalam 67. road - jalan)

Kosa kata yang dibandingkan: $n = 100 - 3 = 97$

Bil.	Bahasa Melayu	Bahasa Miriek	Taraf kognat
1.	scmua [səmu ^w ā]	scmuak [səmu ^w āʔ]	+
2.	abu [abu]	aboh [aboh]	+
3.	kulit pokok [kulit ^ʔ pokoʔ]	ulait ajuh [ulait ^ʔ ajuh]	tolak
4.	perut [perut ^ʔ]	jen [jen]	-
5.	besar [besar]	ayak [ayaʔ]	-
6.	burung [burug]	manok [mānōʔ]	-
7.	gigit [gigit ^ʔ]	gutap [gutap ^ʔ]	-
8.	hitam [hitam]	mitam [mitam]	+
9.	darah [darah]	daghcek [dareʔ]	+
10.	tulang [tulag]	tulang [tulag]	+
11.	tetek [teteʔ]	titait [titait ^ʔ]	+
12.	terbakar [terbakar]	mekuh [møkuh]	-
13.	kuku [kuku]	ilaoh [ilaoh]	-
14.	awan [awan]	awan [awan]	+
15.	sejuk [səjuʔ]	sajuk [səjuʔ]	+
16.	datang [datan]	dusuk [dusuʔ]	-

PERBANDINGAN KUANTITATIF BAHASA MIRI DENGAN BAHASA MELAYU

17.	mati [māti]	matach [mātach]	+
18.	anjing [anjɪŋ]	au [a ^u]	-
19.	minum [minum]	mesap [mesap ^ʼ]	-
20.	kering [kəriŋ]	magha [māra]	-
21.	telinga [təliŋā]	telingah [təliŋāh]	+
22.	tanah [tanāh]	tanak [tanāʔ]	+
23.	makan [mākan]	umin [umin]	-
24.	telur [təluɾ]	tejuk [təjuʔ]	-
25.	mata [māla]	malah [mālah]	+
26.	lemak [ləmāʔ]	lemaoh [ləmāōh]	+
27.	bulu [bulu]	buluh [buluh]	+
28.	api [api]	igun ^ʳ [igun]	+
29.	ikan [ikan]	hutak [hutaʔ]	-
30.	terbang [tərbəŋ]	tulut [tulut ^ʼ]	-
31.	penuh [pənuh]	fenaok/henaoh [fənāōʔ]/[hənāōh]	+
32.	kaki [kaki]	ha'ah [haʔah]	-
33.	beri [bəri]	tuke [tuke]	-
34.	baik [baiʔ]	jek [jeʔ]	-
35.	hijau [hijau]	hijau [hijau]	+
36.	rambut [rambut ^ʼ]	fuk [fuʔ]	+
37.	tangan [taŋān]	tujuk [tujuʔ]	-
38.	kepala [kəpala]	uhaok [uhaoʔ]	-
39.	dengar [dəŋār]	nyait [nāit ^ʼ]	-
40.	jantung [jantun]	jantung [jantun]	+
41.	tanduk [tanduʔ]	ugaag [ugag]	-
42.	aku [aku]	koi [koi]	+
43.	bunuh [bunuh]	binaoh [bināōh]	+
44.	lutut [lutut ^ʼ]	tut [tut ^ʼ]	+
45.	tahu [tahu]	ta'an [taʔan]	-
46.	daun [daun]	da'un [daʔun]	+
47.	berbaring [bariŋ]	lupaek [lupaeʔ]	-
48.	hati [hati]	atai [atai]	+
49.	panjang [panjaŋ]	bet [bet ^ʼ]	-
50.	kutu [kutu]	gutaoh [gutaoh]	+
51.	lelaki / laki-laki [ləlaki]/[laki laki]	lakaih [lakaih]	+
52.	banyak [bajāʔ]	akik [akiʔ]	-
53.	daging [dagiŋ]	dagin [dagin]	+
54.	bulan [bulan]	teka [təka]	-
55.	gunung [gumun]	lat [lat ^ʼ]	-
56.	mulut [mulut ^ʼ]	munong [munōŋ]	-
57.	nama [nāmā]	aden [aden]	-
58.	leher [leher]	bataok [bataoʔ]	-
59.	baru [baru]	baghuk [baruʔ]	+
60.	malam [mālam]	lem [lem]	+
61.	hidung [hidun]	singot [siŋōt ^ʼ]	-
62.	tidak [tidaʔ]	oumluk [oumlōʔ]	-

63.	satu [satu]	seh [seh]	+
64.	orang [oraŋ]	jemulon [jəmulon]	-
65.	hujan [hujan]	ujen [ujen]	+
66.	merah [merah]	ma'ang [mā'āŋ]	-
67.	jalan [jalan]	lahau [lahau]	-
68.	akar [akar]	ghamot/gamot [ʔamōtʔ]/[gamōtʔ]	-
69.	bulat [bulatʔ]	belileng [bəliləŋ]	-
70.	pasir [pasir]	ghaet [ʔaetʔ]	-
71.	sebut [səbutʔ]	matang [mātaŋ]	-
72.	lihat [lihatʔ]	ni'an [niʔān]	-
73.	benih [bənih]	tulaek [tulacʔ]	-
74.	duduk [duduʔ]	utaok [utaoʔ]	+
75.	kulit [kulitʔ]	ulait [ulaitʔ]	+
76.	tidur [tidur]	makel [mākel]	-
77.	kecil [kecil]	diik [diiʔ]	-
78.	asap [asapʔ]	sap [sapʔ]	+
79.	berdiri [bərdiri]	masang [māsang]	-
80.	bintang [bintaŋ]	hatak [hataʔ]	-
81.	batu [batu]	bataoh [bataoh]	+
82.	matahari [mātahari]	matah daloh [mātah daloh]	tolak
83.	berenang [bərənāŋ]	meta [məta]	-
84.	ekor [ekor]	ihaok [ihaoʔ]	-
85.	itu [itu]	jih [jih]	-
86.	ini [ini]	itoi [itoi]	-
87.	kau [kau]	əok [aoʔ]	+
88.	lidah [lidah]	jilek [jileʔ]	-
89.	gigi [gigi]	nyipan [ɲipan]	-
90.	pokok [pokoʔ]	ajuh [ajuh]	-
91.	dua [du ^w a]	debeh [dəbeh]	-
92.	berjalan [bərjalan]	melahau [məlahau]	tolak
93.	panas [panās]	la'oh [laʔoh]	-
94.	air [air]	fek [feʔ]	-
95.	kami [kami]	amach [amāch]	+
96.	apa [apa]	upaeh [upaeh]	+
97.	putih [putih]	hutack [hutacʔ]	+
98.	siapa [si'apa]	sai [sai]	+
99.	perempuan [pərəmpu ^w an]	tughai [tuyai]	-
100.	kuning [kuniŋ]	(k)uneng [(k)unəŋ]	+

Peratus kata berkerabat atau $c = 41/97 = 0.423 = 42\%$

PENENTUAN JANGKA WAKTU PISAH DAN TAHUN/ABAD PISAH BAHASA MELAYU DENGAN BAHASA MIRI

(1) Cari nilai t :

$$t = \frac{\log c}{2 \log r} = \frac{\log 0.42}{2 \log 0.86} = \frac{-0.868}{2 \times -0.151} = \frac{-0.868}{-0.302} = \underline{\underline{2.874}}$$

(2) Cari nilai s :

$$s = \sqrt{\frac{c(1-c)}{n}} = \sqrt{\frac{0.42(1-0.42)}{97}} = \sqrt{\frac{0.42 \times 0.58}{97}} = \sqrt{\frac{0.2436}{97}}$$

$s = \underline{\underline{0.0501}}$

3) Cari nilai c yang baru atau c_1

$c_1 = c + s = 0.42 + 0.0501 = 0.4701$

4) Cari nilai $t_1 = \log c_1 = \log 0.4701 = \frac{-0.777}{2 \log r} = \frac{-0.755}{2 \times -0.151} = \underline{\underline{2.5}}$

(5) Cari jangka kesalahan = $t - t_1 = 2.874 - 2.5 = \underline{\underline{0.374}}$

(6) Cari jangka waktu pisah: t dicampur dan ditolak dengan $t - t_1$

$t + (t - t_1) = 2.874 + 0.374 = \underline{\underline{3.248}}$

$t - (t - t_1) = 2.874 - 0.374 = \underline{\underline{2.5}}$

Jangka waktu pisah ialah **antara 2.5 ribu tahun hingga 3.248 ribu tahun (25 abad hingga 32 abad yang lalu).**

(7) Tentukan tahun / abad pisah (dari tahun 2000).

$3.248 - 2.000 = \underline{\underline{1.248}}$ (abad ke-13 S.M.)

$$2.5 - 2.000 = \underline{0.5} \text{ (abad ke-6 S.M.)}$$

Tahun Pisah ialah **antara abad ke-13 S.M. hingga abad ke-6 S.M.**

Daripada jangka waktu pisah dan tahun pisah yang diperoleh hasil perbandingan bahasa Melayu dengan bahasa Miri maka berdasarkan Jadual 3 sebelum ini maka dapatlah ditentukan tingkat kekerabatan antara kedua-dua bahasa seperti yang ditunjukkan dalam Jadual 4. Jadual 4 menunjukkan bahawa hubungan kekerabatan antara bahasa Melayu dengan bahasa Miri ialah pada tahap sempadan antara keluarga dengan rumpun bahasa. Hal ini bermakna walaupun dipencarkan daripada bahasa induk yang sama, kedua-dua bahasa tersebut telah mencapah dan berkembang secara bersendirian semenjak kira-kira 25 hingga 32 abad yang lalu.

Jadual 4 Tingkat Kekerabatan Bahasa Melayu dengan Bahasa Miri

Bahasa yang Dibandingkan	Peratus Kata Berkerabat	Jangka Waktu Pisah	Tahun/ Abad Pisah	Tingkat Kekerabatan
Bahasa Melayu - bahasa Miri	42%	25 – 32 abad yang lalu	Abad ke-13 S.M.– ke-6 S.M.	Keluarga / abad rumpun

KESIMPULAN

Makalah ini telah membincangkan dan menerapkan kaedah perbandingan ke atas perbandingan dua bahasa Austronesia, iaitu bahasa Melayu dan bahasa Miri. Sebagaimana yang dijelaskan pada awal perbincangan, makalah ini merupakan sumbangan kecil, khususnya terhadap kajian bahasa Miri yang dapat dikatakan hampir pupus. Sebagai kajian permulaan, masih banyak kekurangan dalam perbincangan ini, khususnya dari dua sudut. Pertama, kajian yang dibuat agak dangkal kerana hanya melibatkan 100 kosa kata sahaja. Kedua, kaedah yang digunakan juga merupakan kaedah yang masih diperbahaskan dalam kalangan ahli linguistik sejarah dan perbandingan. Sebaik-baiknya kedua-dua kekurangan ini dapat diatasi dalam kajian lanjutan dengan cara mengemukakan data kosa kata yang jumlahnya berpatutan dan juga penerapan kaedah kualitatif, yakni kaedah pemeriksaan unsur diretensi dan berinovasi, untuk menghuraikan bahasa ini secara lebih terperinci.

LAMPIRAN 1

Daftar 100 Kosa Kata Morris Swadesh

1. All 2. Ashes 3. Bark 4. Belly 5. Big 6. Bird 7. Bite 8. Black 9. Blood 10. Bone
 11. Breast 12. Burn 13. Claw 14. Cloud 15. Cold 16. Come 17. Die 18. Dog 19. Drink
 20. Dry 21. Ear 22. Earth 23. Eat 24. Egg 25. Eye 26. Fat-grease 27. Feather
 28. Fire 29. Fish 30. Fly 31. Full 32. Foot 33. Give 34. Good 35. Green 36. Hair
 37. Hand 38. Head 39. Hear 40. Heart 41. Horn 42. I 43. Kill 44. Knee
 45. Know 46. Leaf 47. Lie 48. Liver 49. Long 50. Louse 51. Man-male 52. Many
 53. Meat-flesh 54. Moon 55. Mountain 56. Mouth 57. Name 58. Neck 59. New
 60. Night 61. Nose 62. Not 63. One 64. Person 65. Rain 66. Red 67. Road
 68. Root 69. Round 70. Sand 71. Say 72. See 73. Seed 74. Sit 75. Skin 76. Sleep
 77. Small 78. Smoke 79. Stand 80. Star 81. Stone 82. Sun 83. Swim 84. Tail
 85. That 86. This 87. Thou 88. Tongue 89. Tooth 90. Tree 91. Two 92. Walk
 93. Warm 94. Water 95. We 96. What 97. White 98. Who 99. Woman 100.
 Yellow

Sumber: Dipetik dan diubah suai daripada Lehmann (1973: 108).

LAMPIRAN 2

Daftar 100 Kosa Kata Morris Swadesh dalam Bahasa Austronesia Purba

1. *abih 2. *abu 3. *u(m)pak / *kulit 4. *(t)ijan 5. *raja / *laba / *bet'a(l)
 6. *bulun / *manuk 7. *karat 8. *i(n)tem 9. *(dD)arah 10. *(t)ulang 11. *t'ut'u
 12. — 13. *kuku 14. — 15. *dingin 16. *datang 17. *pataj 18. *at'u 19. *inum
 20. *kangan / *keran / *kering 21. *talinga 22. *tanah / *tanoh 23. *ka / *ka'en /
 *ka'i / *pangan 24. *(t)elur 25. *mata 26. *menak 27. *bulu 28. *apuj 29. *ikan
 / *ivak 30. — 31. *penuh 32. *kaki 33. *baraj / *bereg 34. — 35. *hid'au 36. *bu'uk
 / *d'(ae)(m)but 37. *tangan / *lima / *linga 38. *ulu 39. *danger 40. *d'antung
 41. *ta(n)duk / *tungu 42. *aku 43. *tedet 44. *(t)u'ut / *(t)Uhu(dj) 45. *tahu
 46. *daun / *(dd)aven / *dahen 47. *inep 48. *ataj / *hataj 49. *andang / *pandang
 50. *tuma / *kutu 51. *laki 52. — 53. *daging 54. *bulan 55. *gunung 56. *mulut
 57. *agan 58. *lihir 59. *baru / *b/ah/aru / *beru 60. *alem / *malem / *bengi
 61. *igung / *ugung 62. *vada 63. *et'a / *it'a 64. *diri / *uran / *ulun 65. *udan
 / *hud'an 66. *irah 67. *dalam / *d'alan 68. *akal / *vakal / *pankal 69. *bulat
 70. *pat'ir / *ket'ik 71. *uk'ap 72. *kita 73. *benih / *binih 74. *dukduk 75. *kulit
 76. *tidur / *tudur 77. *dikih / *e(n)tik / *itik 78. *a(t')y / *hat'ap 79. *d'eng-
 d'eng / *diri 80. *intang / *bintang 81. *batu 82. *a(n)dav / *ha(ng)gav / *vari

83. *languj 84. *buntut / *iku 85. *ijan 86. *ini / *itu 87. *kamu / *kav 88. *dilah
89. *gigi / *ipen 90. *puhun / *puqun 91. *duva 92. — 93. *ga(l)an 94. *(dd)anum
/ *vajer 95. *kita / *kami 96. *apa 97. *putih / *bulai / *t'akt'ak 98. *(t)a(j)i
99. *binaj / *babi / *bei 100. *kunin

Sumber: Utjen Djusen Ranabrata, 1984 (daripada Dempwolff, 1938 *Austronesisches Worterverzeichnis*).

NOTA

¹ Menurut Bibi Aminah dan Abang Ahmad Ridzuan (1993: 703) terdapat ejaan yang berbeza bagi perkataan Miri, iaitu Miri yang digunakan oleh Julaihi Bujang (1976), Tunku Zainah (1978) dan *Miri Regional Planning* (1981), Mirek yang digunakan oleh Sather (1979) dan Tunku Zainah (1982) dan Meirek yang digunakan sebagai nama terdaftar persatuan, iaitu Persatuan Jatti Meirek. Bibi Aminah dan Abang Ahmad Ridzuan (1993: 703) pula menggunakan istilah Miriek untuk merujuk perkara yang sama. Menurut hemat penulis, istilah Mirek, Meirek dan Miriek mungkin istilah mengikut sebutan penduduk tempatan (terutama orang Melayu Sarawak) untuk merujuk kepada istilah berdasarkan ejaan standard, iaitu Miri. Jadi dalam makalah ini istilah orang Miri akan dikekalkan.

² Asmah (1988: 348) menyebut bahasa tua ini sebagai bahasa induk. Istilah lain yang digunakan untuk merujuk kepada bahasa tua ini ialah istilah sumber bersama (*common source*), yakni istilah yang digunakan oleh Sir William Jones di dalam kenyataannya yang terkenal tentang kekerabatan bahasa Indo-Eropah (Arlotto, 1972: 39). Makalah ini akan menggunakan istilah bahasa induk untuk merujuk hal yang sama.

³ Hubungan antara bahasa yang lebih tua dengan bahasa yang lebih muda disebut hubungan kekerabatan atau kekeluargaan (*genetic relationship*) bahasa. Makalah ini akan menggunakan kedua-dua istilah ini untuk merujuk kepada maksud yang sama. Untuk huraian lanjut konsep ini sila rujuk Noriah (1996: 129).

⁴ Lama waktu disebut juga usia bahasa (dalam ribuan tahun). Lama waktu dapat dihitung selepas ditentukan bilakah berlakunya perpisahan bahasa-bahasa sekerabat daripada bahasa induk. Kaedah kuantitatif khusus yang digunakan untuk menentukan lama waktu ialah glotokronologi.

⁵ Jadual tingkat kekerabatan bahasa yang dimaksudkan terdapat di halaman 13 makalah ini.

⁶ Sila rujuk Lampiran 2 daftar 100 kosa kata Morris Swadesh dalam bahasa Austronesia purba.

⁷ Senarai asal kosa kata dasar yang berjumlah 200 kata yang digunakan oleh Mor-

ris Swadesh dan senarai 100 kata yang digunakan oleh John, A.Rea (1956) terdapat di dalam Winfred Lehmann (1973: 108). Makalah ini hanya akan menggunakan senarai kosa kata dasar 100 sahaja. Daftar 100 kosa kata dasar Morris Swadesh dilampirkan dalam Lampiran 1.

⁸ Lihat Lehmann (1973: 105) dan Gudschinsky (1964: 618).

⁹ Kaedah statistik menyediakan pengiraan jangka kesalahan pada mana-mana tahap keyakinan (*confidence*) atau kebarangkalian (*probability*). Walau bagaimanapun, penghitungan lazimnya menggunakan salah satu daripada tiga tahap keyakinan yang berikut, iaitu a) pengiraan yang mengandungi 90% atau 9/10 ketepatan, b) pengiraan yang mengandungi 68% atau 7/10 ketepatan (kesalahan standard [*standard error*]), dan c) pengiraan yang mengandungi 50% atau 5/10 ketepatan (kesalahan mungkin [*probable error*]). Makalah ini menggunakan kesalahan standard (7/10) mengikut Gudschinsky (1964).

¹⁰ Gudschinsky (1964) menggunakan simbol *s* sebagai lambang mewakili kesalahan standard. Makalah ini akan menggunakan simbol 's' sebagai ganti simbol berkenaan.

¹¹ Asmah (1988: 34–35) menghuraikan tingkat kekerabatan ini dengan cara yang lebih kurang sama. Menurutnya yang dimaksudkan dengan filum ialah kelompok bahasa yang lama waktu atau usianya 5000 tahun atau lebih, rumpun pula kelompok bahasa yang lama waktunya antara 2500 sehingga 5000 tahun, kelompok bahasa yang kurang daripada 2500 tahun disebut keluarga bahasa sementara keluarga bahasa pula boleh dibahagi-bahagikan kepada subkeluarga yang lebih kecil.

¹² Penulis ingin mengucapkan jutaan terima kasih kepada lima pemaklum orang Miri yang sudi mengisikan senarai dan dirakamkan sebutan mereka untuk senarai kosa kata dasar 100 dan 200 perkataan dalam bahasa Miri dan bermurah hati mencatatkan kosa kata yang sama dalam dialek Melayu Sarawak. Pemaklum yang dimaksudkan ialah Abd. Latip bin Mastan (Bakam), Zainal bin Seman (Kampung Pengkalan Lutong), Amid bin Yusuf (Pujut), Mohamad Suhaimi bin Abeng (Pujut) dan Darmie bt. Tabran (Kampung Pengkalan Lutong). Disebabkan pemaklum datang dari kawasan yang berlainan, maka terdapat variasi bagi kosa kata dan sebutan kosa kata tertentu. Untuk makalah ini akan dipilih sebutan majoriti di kalangan pemaklum. Analisis lain dalam makalah ini telah dijalankan oleh penulis sendiri. Jika terdapat kesilapan dari segi tatacara penulisan hasil analisis, maka penulis bertanggungjawab sepenuhnya terhadap kesilapan tersebut.

¹³ Perkataan igun berkognat dengan perkataan unggun dalam bahasa Melayu yang bermaksud longgokan kayu yang berapi, menyala atau terbakar.

RUJUKAN

Abd. Syukur Ibrahim, 1981. *Tujuh Bahasa Nusantara* (Sebuah Kajian Leksikostatistik). Penataran Linguistik Kontrasif dan Historis Komparatif Tahap III, Pusat Pembinaan dan Pengembangan Bahasa, Departemen

Pendidikan dan Kebudayaan FKSS IKIP, Malang.

- Arlotto, A. 1972. *Introduction to Historical Linguistics*. Boston. Houghton Mifflin Company.
- Asmah Hj. Omar, 1988. *Susur Galur Bahasa Melayu*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Bibi Aminah Abdul Ghani dan Abang Ahmad Ridzuan. "Pengekalan dan Peralihan Bahasa Miriek: Satu Pemerhatian" dalam *Jurnal Dewan Bahasa* 37: 8 Ogos 1993, hlm. 703–716.
- Bynon, T. 1994 (terjemahan). *Linguistik Sejarah*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Daripada Editor "Kuantitatif Lawan Kualitatif" dalam *Pelita Bahasa* 10: 7 Julai 1998 hlm. 2.
- Dyen, I. (ed.) 1973. "Lexicostatistics in Genetic Linguistics". *Proceedings of the Yale Conference*, Yale University, April 3–4 1971. Mouton: The Hague.
- Dyen 1975. *Linguistic Subgrouping and Lexicostatistics*. Paris: The Hague.
- Gudschinsky, S. "The ABC's of Lexicostatistics (Glottochronology)" dalam Hymes, D. 1964, *Language in Culture and Society, A Reader in Linguistic and Anthropology*. New York: Harper and Row, hlm. 612–623.
- Gorys Keraf, 1984. *Linguistik Bandingan Historis*. Jakarta: Penerbit PT Gramedia.
- Gorys Keraf, 1990. *Linguistik Bandingan Tipologis*. Jakarta: Penerbit PT Gramedia.
- Grimes, B.F. (ed.) *Ethnologue: Languages of the World*, 13th Edition 1996–1999 Internet Version. Bahan Internet di alamat <http://www.sil.org/ethnologue/ethnologue.html> dan secara lebih khusus untuk Malaysia rujuk <http://www.sil.org/ethnologue/countries/Mall.html>
- Harimurti Kridalaksana. "Perhitungan Leksikostatistik Atas Delapan Bahasa Nusantara Barat Serta penentuan Pusat Penyebaran Bahasa-Bahasa itu berdasarkan Teori Migrasi" dalam *Madjalah Ilmu-Ilmu Sastra Indonesia* Oktober 1964, hlm. 319–351.
- Jos Daniel Pareira, 1987. *Pengantar Studi Linguistik Umum dan Linguistik Bandingan (Historis) Sebuah Kisah Zaman*. Jakarta: Penerbit Erlangga.
- Lehmann, W. 1973. *Historical Linguistics An Introduction*. New York: Holt, Rinehart and Winston Inc.
- Mees, C.A. 1967. *Ilmu Perbandingan Bahasa-Bahasa Austronesia*. Kuala Lumpur: Oxford University Press.
- Noriah Mohamed. "Waktu Pisah dan Tingkat Kekeperabatan Bahasa Melayu dengan Bahasa Iban" dalam *Jurnal Dewan Bahasa* 38: 2 Februari 1994, hlm. 129–143.
- Noriah Mohamed. "Konsep Kekeluargaan Bahasa dan Kaedah Perbandingan dalam Linguistik Sejarah dan Perbandingan" dalam *Jurnal Dewan Bahasa* 40: 2 Februari 1996, hlm. 128–138.
- Noriah Mohamed. "Pengelompokan Bahasa" dalam *Jurnal Dewan Bahasa* 43: 12 Disember 1999, hlm. 1077–1107.

PERBANDINGAN KUANTITATIF BAHASA MIRI DENGAN BAHASA MELAYU

- Swadesh, M. (edited by Sherzer, J. forward by Hymes, D.) 1971. *The Origin and Diversification of Language*. London: Routledge and Kegan Paul.
- Uhlenbeck, E.M. "The Comparative Study of the Austronesian Languages" dalam *Linguistic Comparison in South East Asia and the Pacific*, Collected Papers in Oriental and African Studies 1963. School of Oriental and African Studies, University of London, hlm. 24–27.
- Utjen Djusen Ranabrata, 1984. "Refleks Fonem Proto Melayu Polinesia dalam Bahasa Indonesia". *Penataran Linguistik Historis Komparatif IV*, Pusat Pembinaan dan Pengembangan Bahasa, Departemen Pendidikan dan Kebudayaan, Jakarta.
- Tunku Zainah Tunku Ibrahim, 1978. "Malay Ethnicity in Sarawak: The Case of the Orang Miri". Tesis Sarjana Sains Kemasyarakatan yang tidak diterbitkan, Universiti Sains Malaysia.