

LEKSIKOSTATISTIK DAN GLOTOKRONOLOGI BAHASA BATAK: HUBUNGAN KEKERABATAN BAHASA BATAK DIALEK TOBA, SIMALUNGUN, MANDAILING DAN KARO

Wartono

Pengkaji Bahasa Balai Bahasa Sumatera Utara

Abstract :

Language genetic relationship is a study of the linguistics historic comparative. Bataknese has several dialects such as Toba, Mandailing, Simalungun, and Karo. Using quantative method, the four dialects language relationship are predicted using lexicostatistics technique. In lexicostatistics, language genetic is seen based on sounds similarity in lexicon in those languages. Phonetic similarity will be based whether a word in a language has correlation with other language. 200 Swadesh base vocabolaries is used as indicator to determine cognate words. Using Crowley tehniqe finds that the time separated among the four dialects is in a family level. Another findings, Mandailing is a dialect proto from the four dialects. Based on glotocronology technique that time seperated of the four dialects is 1000-2000's years ago that means those dialects has existed since 205 BC – 934 M. This dialect migrates from Mandailing to Toba then Simalungun and Karo. Language migration shows one chain of language migration.

Keywords: *language genetic, lexicostatistics, time seperated, language age, language migration*

Abstrak :

Hubungan Kekerabatan Bahasa adalah merupakan kajian linguistik historis komparatif. Bahasa Batak mempunyai beberapa dialek diantaranya Toba, Mandailing, Simalungun, dan Karo. Dengan menggunakan metode kuantitatif hubungan kekerabatan keempat dialek tersebut dihitung secara leksikostatistik. Dalam leksikostatistik, kekerabatan bahasa dilihat berdasarkan persamaan bunyi-bunyi yang ada dalam leksikon yang muncul pada bahasa-bahasa tersebut. Kemiripan secara fonetis ini akan menjadi dasar apakah sebuah kata dalam satu bahasa memiliki hubungan dengan bahasa yang lain. Indikator yang digunakan untuk menentukan kata berkerabat adalah kosa kata dasar yang disebut kosa kata dasar Swadesh yang berjumlah dua ratus kosa kata yang dianggap ada pada semua bahasa di dunia. Dengan menggunakan cara Crowley diperoleh hasil bahwa hubungan kekerabatan keempat dialek tersebut adalah dalam satu keluarga. Diketahui juga bahwa Mandailing merupakan dialek proto dari keempat dialek yang ada. Diperoleh hasil berdasarkan penghitungan glotokronologi bahwa waktu pisah keempat dialek ini terjadi antara 1000-2000an tahun yang lalu yang berarti dialek tersebut sudah ada sejak tahun 205SM sampai 934M. Dialek ini bermigrasi dari Mandailing ke Toba lalu ke Simalungun dan Karo. Arah Migrasi menunjukkan satu mata rantai migrasi bahasa.

Kata kunci: *kekerabatan, leksikostatistik, waktu pisah, usia bahasa, migrasi bahasa*

1. Pendahuluan

a. Latar Belakang

Dengan perpindahan manusia ke seluruh bumi menyebabkan bahasa terpisah-pisah dari bahasa induknya karena menyesuaikan dengan kontak sosial, alam, dan lingkungannya. Pada akhirnya memungkinkan kita manusia sekarang

mengetahui bagaimana hubungan bahasa-bahasa di dunia melalui kajian perbandingan bahasa atau yang lebih dikenal dengan linguistik bandingan historis atau linguistik diakronis, atau apapun namanya yang pasti bahwa bahasa-bahasa di seluruh dunia ini dapat dicari hubungan kekerabatannya dan diketahui kapan berpisahannya dan bagaimana

9MEDAN MAKNA	Vol. XI No. 1	Hlm. 61 - 75	2013	ISSN 1829-9237
---------------------	----------------------	---------------------	-------------	-----------------------

bisa terpisah. Kajian-kajian tentang bahasa dari sisi sejarahnya dalam kajian linguistik termasuk dalam kajian Linguistik Historis Komparatif atau Linguistik Bandingan Historis.

Linguistik Bandingan Historis atau Linguistik Historis Komparatif adalah sebuah cabang dari ilmu bahasa yang mempersoalkan bahasa dalam bidang waktu serta perubahan-perubahan unsur bahasa yang terjadi dalam bidang waktu tersebut. Dalam Linguistik Historis Komparatif dipelajari data-data dari suatu bahasa atau lebih, sekurang-kurangnya dalam dua periode. Data-data itu diperbandingkan dengan cara cermat untuk memperoleh kaidah-kaidah perubahan yang terjadi dalam bahasa tersebut (Keraf, 1991: 22).

Dalam penelitian ini, peneliti akan meneliti hubungan kekerabatan bahasa Batak Toba, Madailing, Simalungun dan Karo yang ada di provinsi Sumatera Utara. Penelitian ini berasumsi pada penelitian terdahulu bahwa bahasa Batak Dialek Toba, Madailing, Simalungun dan Karo mempunyai hubungan kekerabatan namun tingkat kekerabatannya berbeda-beda. Dengan demikian peneliti menganggap penting dilakukan pengkajian perbandingan bahasa dengan menggunakan teknik leksikostatistik dan glotokronologi sehingga kebutuhan informasi perbandingan bahasa-bahasa daerah di Sumatera Utara dapat dipenuhi.

Indikator yang akan digunakan untuk merekonstruksi dan melihat glotokronologi bahasa Batak Dialek Toba, Mandailing, Simalungun dan Karo adalah kosa-kata dasar Swadesh yang selama ini menjadi dasar untuk mempelajari kekerabatan bahasa-bahasa di dunia. Kosakata yang diambil ini karena seluruh bahasa di dunia mempunyai kosakata tersebut.

b. Rumusan Masalah

Ada beberapa masalah utama yang menjadi dasar penelitian ini, yaitu:

1. Bagaimanakah hubungan kekerabatan bahasa Batak Dialek Toba, Mandailing, Simalungun dan Karo?

2. Apakah fonem proto dan bahasa proto dari bahasa Batak Dialek Toba, Mandailing, Simalungun dan Karo?
3. Kapan waktu pisah dan usia bahasa Batak Dialek Toba, Mandailing, Simalungun dan Karo?
4. Bagaimana migrasi Bahasa Batak Dialek Toba, Mandailing, Simalungun dan Karo?

c. Batasan Masalah

Batasan masalah digunakan untuk memudahkan pengerjaan penelitian ini. Peneliti membatasi penelitian dengan hanya menggunakan 200 kosakata dasar menurut Swadesh yang diproses untuk mengetahui hubungan kekerabatan, fonem proto dan bahasa proto, waktu pisah dan usia bahasa, serta migrasi dari 4 bahasa yang menurut penutur bukan bahasa yang sama yaitu bahasa Batak Dialek Toba, Mandailing, Simalungun dan Karo.

d. Tujuan Penelitian

1. Mengetahui hubungan kekerabatan bahasa Batak Dialek Toba, Madailing, Simalungun dan Karo
2. Mengetahui fonem proto dan bahasa proto dari bahasa Batak Dialek Toba, Madailing, Simalungun dan Karo
3. Menghitung dan mengetahui waktu pisah dan usia bahasa Batak Dialek Toba, Madailing, Simalungun dan Karo
4. Mengetahui migrasi bahasa Batak Dialek Toba, Madailing, Simalungun dan Karo

e. Manfaat Penelitian

Secara teoritis manfaat dari penelitian ini adalah untuk menambahkan informasi kebahasaan tentang bahasa Batak, memberikan informasi tambahan tentang penelitian linguistik historis di Sumatera Utara, memberikan gambaran tentang hubungan kebudayaan dan bahasa dari pemilik bahasa-bahasa yang diteliti di masa lampau. Sedangkan manfaat praktis adalah memberikan informasi tentang bahasa Batak

9MEDAN MAKNA	Vol. XI No. 1	Hlm. 61 - 75	2013	ISSN 1829-9237
---------------------	----------------------	---------------------	-------------	-----------------------

kepada pengguna, pemilik dan peneliti bahasa tersebut, menjadi bahan rujukan untuk penelitian linguistik historis, memberikan informasi kepada pemegang kekuasaan untuk mengambil kebijakan politik terkait otonomi daerah dengan sumbangsih penelitian kekerabatan bahasa.

2. Konsep dan Landasan Teori

a. Kekerabatan Bahasa

Kekerabatan dalam istilah linguistik diartikan sebagai hubungan antara dua bahasa atau lebih yang diturunkan dari sumber yang sama (KBBI, 2008). Sedangkan, bahasa berkerabat diartikan sebagai bahasa yang mempunyai hubungan genealogis dengan bahasa lain. Bahasa-bahasa yang berada dalam satu rumpun yang sama tentulah memiliki kekerabatan. Akan tetapi, tingkat kekerabatan bahasa-bahasa yang berada dalam satu rumpun ini kemungkinan tidaklah sama. Sejauh mana tingkat keeratatan hubungan bahasa yang satu dengan yang lainnya dapat dilihat dari kemiripan atau perbedaan dari bahasa-bahasa yang dibandingkan. Semakin mirip kedua bahasa, semakin eratlah hubungan kekerabatannya. Semakin berbeda kedua bahasa, semakin rengganglah hubungan kekerabatannya.

Dalam Kamus Linguistiknya, Kridalaksana menyebut bahasa Batak dengan sebutan “Dialek-Dialek Batak”, yang menyiratkan bahwa bahasa Batak memiliki banyak variasi yang masih tergolong dalam dialek. Terdiri atas bahasa Batak (dialek) Karo/Alas/pakpak, bahasa Batak (dialek) Toba, bahasa Batak (dialek) Simalungun, bahasa Batak (dialek) Angkola Mandailing. Hal ini kemudian didukung oleh temuan Balai Bahasa Medan dalam kegiatan pemetaan bahasa-bahasa di seluruh Nusantara dengan perhitungan dialektologi yang masih mengategorikan bahasa batak dan varian-variannya dalam bentuk dialek-dialek, bukan bahasa-bahasa yang berdiri sendiri. Perlu dicatat bahwa data yang digunakan dalam kegiatan itu adalah data bahasa yang diperoleh pada tahun 1990.

Dengan kemungkinan adanya perubahan dalam kurun waktu setiap 10 tahun, kenyataan yang ada saat ini bisa jadi tidak lagi sama.

Sementara itu, Panggabean dalam tesisnya pada tahun 1994 menyebut “Bahasa-Bahasa Batak” yang dimaknai bahwa variasi-variasi yang ada dalam kelompok bahasa Batak sudah berdiri sebagai bahasa. Ini ditopang dengan kenyataan bahwa masing-masing masyarakat Batak akan menganggap sukunya sebagai suku yang berbeda dan berdiri sendiri. Masyarakat Batak Mandailing/Angkola bertahan bahwa mereka adalah masyarakat yang berbeda, bukan bagian dari Batak. Saat ini, hal itu berkembang lagi dengan kenyataan bahwa Mandailing dengan Angkola pun tidak mau disamakan lagi. Begitu juga yang terjadi dengan Simalungun, Karo, Pakpak/Dairi. Karena itu perlu ada penelitian lanjutan untuk membuktikan apakah dialek-dialek ini benar bagian dari bahasa Batak atau sudah menjadi bahasa sendiri-sendiri.

b. Leksikostatistik

Leksikostatistik adalah suatu teknik yang memungkinkan kita untuk menentukan tingkat hubungan di antara dua buah bahasa, dengan menggunakan cara yang paling mudah, yaitu dengan membandingkan kosa kata pada bahasa-bahasa tersebut yang kemudian dapat dilihat dan ditentukan tingkat kesamaan di antara kosa kata kedua bahasa (Crowley: 1992:168). Dengan demikian, sejauh mana hubungan kekerabatan satu bahasa dengan bahasa lainnya dapat diketahui.

Menurut Crowley (1987: 191—192), metode leksikostatistik beroperasi di bawah dua asumsi dasar. Asumsi pertama ialah bahwa beberapa bagian kosakata dari sebuah bahasa sukar berubah daripada bagian lainnya. Apa yang dimaksud dengan kosakata yang sukar berubah adalah kosakata dasar, yakni kata-kata yang sangat intim dalam kehidupan bahasa, dan merupakan unsur-unsur yang menentukan mati hidupnya suatu bahasa (lihat juga Keraf, 1991: 123).

9MEDAN MAKNA	Vol. XI No. 1	Hlm. 61 - 75	2013	ISSN 1829-9237
--------------	---------------	--------------	------	----------------

Kemudian, istilah ‘perubahan’ mengacu pada penggantian sebuah kata dengan sebuah kata nonkerabat karena bentuk asli berubah maknanya sehingga kemunculannya merujuk kepada sesuatu yang lain, atau karena sebuah kata dipinjam dari bahasa lain untuk mengekspresikan makna tertentu.

Asumsi kedua ialah bahwa perubahan kosakata dasar pada semua bahasa adalah sama. Asumsi ini telah diuji pada 13 bahasa, di antaranya bahasa yang memiliki naskah-naskah tertulis. Hasilnya menunjukkan bahwa dalam tiap 1.000 tahun, kosakata dasar suatu bahasa bertahan antara 86,4—74,4 %, atau dengan angka rata-rata 80,5%. Tentu saja hal itu tidak dapat diartikan bahwa semua bahasa akan bertahan dengan persentase rata-rata tersebut, karena semua bahasa yang digunakan dalam eksperimen itu (kecuali dua bahasa) adalah bahasa-bahasa Indo-Eropa.

Bila asumsi kedua diterima, retensi rata-rata kosakata dasar suatu bahasa dalam tiap 1.000 tahun dapat dinyatakan dalam rumus: $80,5\% \times N$. Simbol N adalah jumlah kosakata dasar yang ada pada awal kelipatan 1.000 tahun yang bersangkutan. Dari 200 kosakata dasar (N) suatu bahasa sesudah 1.000 tahun pertama akan tinggal $80,5\% \times 200$ kata = 161 kata. Sesudah 1.000 tahun kedua akan tinggal $80,5\% \times 161$ kata = 139,6 kata atau dibulatkan menjadi 140 kata. Sesudah 1.000 tahun ketiga kosakata dasarnya tinggal $80,5\% \times 140$ kata = 112,7 atau dibulatkan menjadi 113 kata, dan seterusnya.

“Leksikostatistik adalah metode pengelompokan bahasa yang dilakukan dengan menghitung prosentase perangkat kognat/kerabat (Mahsun,1995:115)”. Dalam penghitungan leksikostatistik, kata-kata yang memiliki kemiripan dari segi fonetis atau morfologi akan dianggap sebagai kata yang berkerabat atau dikenal dengan istilah kognat (*cognate*). Melalui kata-kata berkerabat inilah dilakukan penghitungan waktu pisah dari bahasa protonya atau usia bahasa.

Menurut Keraf,(1984: 121) Leksikostatistik itu suatu teknik dalam

pengelompokan bahasa yang lebih cenderung mengutamakan peneropongan kata-kata (leksikon) secara statistik, untuk kemudian berusaha menetapkan pengelompokan itu berdasarkan persentase kesamaan dan perbedaan suatu bahasa dengan bahasa lain. Dari konsep di atas, Keraf kemudian menjabarkan metode kerja dalam leksikostatistik yang dapat dijadikan acuan dalam penelitian kekerabatan bahasa.

Dalam leksikostatistik, tataran yang berbeda dari subkelompok dinamai sebagai berikut:

Penamaan
Subkelompok Bahasa

Level subkelompok	persentase kerabat pada kosakata inti
Bahasa (<i>language</i>)	81—100%
Keluarga (<i>family</i>)	36—81%
Rumpun (<i>stock</i>)	12—36%
Mikrofilum	4—12%
Mesofilum	1—4%
Makrofilum	0—1%

Dalam klasifikasi leksikostatistik, kesamaan pada tingkat 81-100% disebut bahasa, kesamaan pada tingkat 36—81% disebut keluarga, kesamaan pada tingkat 12-36% disebut rumpun, kesamaan pada tingkat 4-12% disebut mikrofilum, kesamaan pada tingkat 1-4% disebut mesofilum, dan kesamaan pada tingkat 0-1% disebut makrofilum. Namun, perlu dicatat bahwa ahli bahasa yang berbeda adakalanya menggunakan hitungan yang berbeda.

Keraf (1991: 127—130) mengatakan bahwa dalam membandingkan kata-kata untuk menetapkan kata-kata kerabat dan kata-kata nonkerabat terdapat asumsi bahwa fonem bahasa proto yang berkembang secara berlainan dalam bahasa-bahasa kerabat akan berkembang secara konsisten dalam lingkungan linguistik bahasa kerabat masing-masing. Dalam perbandingan itu, fonem-fonem dalam posisi relatif sama dibandingkan satu sama lain. Bila terdapat hubungan genetis, pasangan fonem tersebut akan timbul kembali dalam banyak pasangan

lain. Tiap pasangan yang sama yang timbul dalam hubungan itu merupakan pantulan suatu fonem atau alofon dalam bahasa protonya (lihat juga Crowley).

c. Glotokronologi

Glotokronologi adalah suatu teknik dalam dialektologi yang berusaha mengadakan pengelompokan dengan mengutamakan penghitungan waktu atau penghitungan usia bahasa-bahasa kerabat. Dalam hal ini usia bahasa tidak dihitung secara mutlak dari suatu tahun tertentu, tetapi dihitung secara umum dengan menggunakan satuan ribuan tahun. (Keraf 1984:121).

Pendapat senada dikemukakan oleh Crowley (1992:79) yang menyatakan metode kedua yang biasanya digunakan untuk menentukan waktu tepatnya kapan bahasa yang berkerabat berpisah disebut glotokronologi. Metode ini memungkinkan ahli bahasa mengetahui sudah berapa lama bahasa-bahasa yang berkerabat yang dalam hal ini termasuk pada level sub-grouping telah berpisah.

Hasil perhitungan tersebut dapat dikelompokkan Crowley (1992:179) menjadi:

Tingkat Pengelompokan	Tahun Pisah
Dialek dari satu bahasa	Kurang dari 500
Bahasa dari satu keluarga	500 sampai dengan 2500
Keluarga dari satu rumpun	2500 sampai dengan 5000
Rumpun dari satu mikrofilum	5000 sampai dengan 7500
Mikrofilum dari satu mesofilum	7500 sampai dengan 10.000
Mesofilum dari satu makrofilum	Lebih dari 10.000 tahun

Jika jumlah kata berkerabat antara dua bahasa yang ditelaah antara 200-162, dengan persentase 100-81, maka waktu pisah

diperkirakan 0-500 tahun yang lalu. Jika jumlah kata yang berkerabat antara 162-132 dengan persentase 81-66, maka waktu pisah kedua bahasa diperkirakan antara 500-1000 tahun yang lalu. Jika jumlah kata berkerabatnya 132-106, dengan persentase 66-53, maka waktu pisah kedua bahasa itu diperkirakan 1000-1500 tahun yang lalu, dan seterusnya. Setelah menghitung waktu pisah bahasa-bahasa yang dijadikan objek penelitian, selanjutnya dapat pula dilakukan penghitungan usia bahasa seperti yang dijabarkan di atas dikaji dalam dialektologi dengan metode yang disebut *glotokronologi*.

Berdasarkan prinsip itu, waktu pisah kedua bahasa kerabat dengan prosentase kata kerabat yang diketahui adalah seperti tertera dalam tabel berikut ini (Keraf: 1984: 125):

Perkiraan Waktu Pisah dan Usia Bahasa

Jumlah kata kerabat antara A—B	Persentase kata kerabat	Usia (waktu pisah) antara bahasa A—B sekian tahun yang lalu (sudah dibagi 2)
200-162	100-81	0-500
162-132	81-66	500-1000
132-106	66-53	1000-1500
106-86	53-43	1500-2000
86-70	43-35	2000-2500
70-56 dan seterusnya	35-28	2500-3000

d. Migrasi Bahasa

Migrasi bahasa adalah sebuah usaha pengumpulan asumsi, batasan-batasan, dan hipotesa-hipotesa yang membicarakan gerak dan arah migrasi bangsa-bangsa pada jaman prasejarah. Adapun tujuan mengetahui migrasi bahasa adalah untuk mengetahui negeri asal bahasa dan daerah penyebaran bahasa-bahasa yang berkerabat sekarang ini.

Keraf (1984:172) menyatakan wilayah suatu bahasa diidentifikasi dengan daerah penutur-penuturnya. Ada dua istilah dalam teori migrasi yaitu wilayah dan daerah. Wilayah bahasa adalah tempat-tempat

di mana terdapat penutur-penutur suatu bahasa. Dalam kenyataan suatu bahasa dapat terdiri dari suatu tempat yang secara geografis bersinambungan, atau dapat pula terdiri dari sejumlah tempat yang secara geografis terpisah satu dari yang lain. Tiap satuan tempat yang secara geografis terpisah dari yang lain tetapi dihuni oleh penutur-penutur bahasa yang sama disebut daerah bahasa. Terjadinya daerah bahasa bisa karena penuturnya berpindah.

Teori migrasi bahasa didasarkan dua dalil, yaitu wilayah asal bahasa-bahasa kerabat merupakan suatu daerah yang bersinambungan. Hal ini bisa menjadi dasar untuk menentukan satu daerah asal yang merupakan daerah kesatuan bagi bahasa-bahasa yang terpisah letaknya dewasa ini dari pada harus mengambil semua daerah secara bersama-sama sebagai wilayah asal. Dalil yang kedua adalah jumlah migrasi yang mungkin direkonstruksi akan berbanding terbalik dengan jumlah gerak perpindahan dari tiap bahasa. Maksud dari dalil ini adalah sebagai kaidah gerak yang paling minimal, yaitu bila jumlah gerak dalam dua buah peluang migrasi yang direkonstruksi itu berbeda, maka migrasi dengan jumlah gerak yang paling kecil mempunyai peluang yang paling besar sebagai migrasi yang sesungguhnya pernah terjadi.

e. Penelitian Terdahulu

Sudah banyak penelitian kekerabatan bahasa yang dilakukan para peneliti bahasa di linguistik historis komparatif, diantaranya adalah Ika Indriani H. dengan judul *Leksikostatistik Bahasa Batak Toba dengan Bahasa Pakpak Dairi* (2007). Dalam penelitian ini digunakan daftar kosa kata yang disusun oleh Mahsun sebanyak 809 kosa kata. Akan tetapi, dalam penelitian ini tidak digambarkan secara jelas perubahan-perubahan yang terjadi pada bahasa-bahasa yang diteliti. Hasil penelitian tersebut menyatakan bahwa bahasa Batak Toba dan bahasa Pakpak Dairi merupakan bahasa tunggal pada 2.320-2200 tahun yang lalu. Bahasa Batak Toba dan bahasa Pakpak Dairi

mulai berpisah dari suatu bahasa proto antara 320-200 sebelum Masehi (dihitung dari tahun 2000).

Suyata Pujiati *Dari Leksikostatistik ke Glotokronologi: Analisis Sembilan Bahasa di Indonesia* (1998). Dalam penelitian ini dikaji kekerabatan Sembilan bahasa di Indonesia, yaitu bahasa Batak, bahasa Minang, bahasa Melayu, bahasa Banjar, bahasa Sunda, bahasa Jawa, bahasa Madura, bahasa Bali, dan bahasa Bugis dengan alat bantu 100 kosakata dasar Swadesh. Penelitian ini dilakukan di Yogyakarta dengan menggunakan mahasiswa sebagai informan. Pujiati menggunakan teknik yang dikemukakan oleh Dyen dan hasil yang didapatkannya adalah bahasa Batak dan bahasa Bugis merupakan bahasa proto atau bahasa tertua, yang menurunkan bahasa yang lain, dari kesembilan bahasa yang diteliti sehingga Pujiati menyebutnya sebagai bahasa proto Batak-Bugis. Bahasa proto Batak-Bugis awalnya adalah bahasa satu bahasa. Dalam perjalanannya, induk bahasa tersebut berpisah menjadi 3 subgrup: 1. Subgrup Batak, Melayu, Minang, dan Banjar; 2. Subgrup Sunda, Jawa, dan Madura; 3. Subgrup Bali dan Bugis. Dalam penelitian itu disimpulkan bahwa bahasa yang memiliki kekerabatan paling renggang atau jauh adalah Batak dan Bugis, sedangkan bahasa Melayu dan bahasa Minang memiliki tingkat kekerabatan yang paling erat atau dekat.

Himpun Panggabean *Telaah Bahasa-Bahasa Batak dari Segi Leksikostatistik* (1994). Dalam penelitian ini Panggabean menggunakan 300 kosa kata dasar yang merupakan kombinasi dari Swadesh, Gudschinsky, Travis, Rea, dan Keraf. Panggabean menyimpulkan bahwa bahasa-bahasa Batak mempunyai tingkat kekerabatan dan waktu pisah yang tidak sama antara satu dengan yang lainnya. Bahasa Batak Toba dan Angkola berada dalam satu bahasa yang sama, dengan kata lain kekerabatan keduanya masih sangat erat sehingga salah satunya berstatus dialek dari yang lain. Sedangkan bahasa-bahasa Batak yang lain berada dalam lingkup keluarga. Dengan perincian, bahasa

9MEDAN MAKNA	Vol. XI No. 1	Hlm. 61 - 75	2013	ISSN 1829-9237
--------------	---------------	--------------	------	----------------

Karo, bahasa Alas, dan bahasa Dairi berada dalam satu kemompok, dan bahasa Simalungun tidak berada dalam kelompok kedua kelompok bahasa tersebut. Ini berarti, bahasa Simalungun berdiri sendiri.

Selain itu, telaah kekerabatan bahasa-bahasa yang ada di Nusantara juga banyak dilakukan oleh para pakar, seperti Kridalaksana, Blust, Dyen, Fernandes, Mbete, Mahsun, dan lain-lain yang dalam penelitian ini dijadikan rujukan.

Berdasarkan hasil penelitian-penelitian diatas, maka dilakukan penelitian ini untuk menambahkan informasi yang ditemukan dalam penelitian-penelitian diatas.

3. Metodologi Penelitian

a. Jenis Penelitian

Penghitungan leksikostatistik diaplikasikan untuk menentukan tingkat kekerabatan bahasa-bahasa yang diteliti, menghitung waktu pisah, bahkan usia bahasa serta migrasi bahasa. Dalam penelitian ini nantinya akan banyak menggunakan angka sebagai acuan penentuan hasil penelitian. Penganalisisan dengan menggunakan rumus-rumus yang ada pada teknik leksikostatistik dan glotokronologi sehingga semakin kuatlah metode kuantitatif dalam penelitian ini.

b. Lokasi dan Waktu Penelitian

Penelitian ini dilakukan di kota Medan dengan mengambil data bahasa Batak dialek Toba, Mandailing, Simalungun dan Karo yang diambil dari data penelitian pemetaan bahasa Pusat Bahasa tahun 2007. Data tahun 2007 diverifikasi ulang dengan mengecek beberapa narasumber dari dialek-dialek dalam kajian ini. Sedangkan waktu penelitian dilaksanakan selama lima bulan dari bulan Mei 2013 sampai bulan Oktober 2013.

c. Data

Data yang digunakan untuk instrumen penelitian adalah kosa kata dasar (*basic core vocabulary*) yang berisi daftar 200 kosa kata yang dianggap ada di seluruh bahasa yang ada di dunia dan tidak akan berubah dalam

kurun waktu selama 1000 tahun. Seperti telah dijelaskan di atas, dalam penelitian kekerabatan bahasa, kosa kata yang paling sering digunakan sebagai data adalah kosa kata milik Morris Swadesh yang kemudian direvisi dan disesuaikan dengan lingkungan alam, budaya, dan kebiasaan masyarakat yang akan diteliti.

d. Teknik Analisis Data

Teknik yang digunakan adalah teknik leksikostatistik. Dalam teknik ini, untuk menganalisis data dilakukan dengan langkah-langkah sebagai berikut:

- a. Mendaftarkan kata-kata yang didapatkan dari lapangan.

Data yang didapatkan dari lapangan dimasukkan dalam tabel setiap bahasa. Misalnya, data Toba yang didapatkan dari informan1 semua dimasukkan dalam tabel untuk dipilih mana yang lebih baik dijadikan sebagai data.

- b. Memilih kata-kata yang akan dijadikan data penelitian dari setiap bahasa.

Kata-kata yang telah didaftarkan kemudian diseleksi untuk mencari kata-kata yang akan dijadikan data penelitian. Pemilihan ini dilakukan dengan pertimbangan bahwa kata yang dipilih adalah yang memiliki kemiripan di antara ketiga kata yang dibandingkan dari setiap bahasa.

- c. Menentukan kata-kata yang berkerabat yang akan dianalisis.

Langkah selanjutnya adalah penentuan kata-kata yang berkerabat berdasarkan kemiripan/persamaan bunyi-bunyi yang dimiliki keempat dialek yang dibandingkan.

- d. Menghitung tingkat kekerabatan ketiga bahasa.

Selanjutnya, untuk menghitung persentase kata kerabat digunakan rumus (Keraf:1984: 127)

$$C = \frac{K}{G} \times 100 \%$$

9MEDAN MAKNA	Vol. XI No. 1	Hlm. 61 - 75	2013	ISSN 1829-9237
--------------	---------------	--------------	------	----------------

C= cognates atau kata yang berkerabat

K= jumlah kosa kata kerabat

G= jumlah glos

- e. Menentukan fonem proto untuk menentukan dialek proto.

Untuk menentukan dialek proto ditentukan dulu fonem proto untuk setiap glos. Dari 200 kosakata swadesh, setiap katanya ditemukan dulu fonem protonya. Lalu dihitunglah dari setiap dialek, dialek mana yang paling banyak menjadi fonem proto maka akan menjadi dialek proto dari bahasa Batak.

- f. Menghitung waktu pisah ketiga bahasa.

Waktu pisah antara dua bahasa kerabat yang telah diketahui prosentase kata kerabatnya, dapat dihitung dengan mempergunakan rumus berikut (Crowley, 1992:178;

$$\text{Keraf, 1984: 130): } t = \frac{\log c}{2 \log r}$$

t = waktu perpisahan dalam ribuan (melenium) tahun yang lalu

r = retensi atau prosentase konstan dalam 1000, atau disebut juga indeks

C = prosentase kerabat

Log = logaritma dari

- g. Menghitung usia ketiga bahasa.

Untuk menghitung jangka kesalahan biasanya digunakan kesalahan standar, yaitu 70% dari kebenaran yang diperkirakan. Kesalahan standar diperhitungkan dengan rumus berikut ini (Keraf: 1984:132):

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

S = kesalahan standar dalam prosentase kata kerabat

c = prosentase kata kerabat

n = jumlah kata yang diperbandingkan (baik kerabat maupun nonkerabat)

Perhitungannya dapat dilakukan dengan aturan sebagai berikut:

(1) 1 dikurangi C

(2) C dikalikan dengan hasil dari (1)

(3) hasil dari (2) dibagi dengan n

(4) menarik akar atas hasil dari 3

(5) hasil dari (4) merupakan jangka kesalahan dan prosentase kata kerabat atas dasar 0,7 perkiraan mengenai kebenaran yang sesungguhnya.

Dengan memperoleh hasil pada no (4) di atas, harus dilakukan perhitungan kesalahan standar dalam tahun, dengan mengikuti langkah-langkah berikut:

(1) jangka kesalahan dari prosentase kerabat no (4) ditambahkan kepada C;

(2) jumlah dalam (1) diperlakukan sebagai C baru, yang akan dimasukkan dalam rumus perhitungan waktu;

(3) perhitungan waktu yang baru sebagaimana diperoleh dalam (2) dikurangi dengan jumlah waktu yang pertama. Angka yang baru ini ditambah dan dikurangi dengan angka yang pertama untuk memperoleh jangka kesalahan atas dasar 0,7 dari keadaan sebenarnya.

Hasil dari kesalahan ini jumlahkan dengan prosentase kerabat untuk mendapatkan C baru, dengan C yang baru ini sekali lagi dihitung waktu pisah dengan mempergunakan rumus waktu pisah. Berdasarkan hasil penghitungan tersebut dapat diketahui perkiraan waktu pisah dan usia bahasa yang dibandingkan.

- h. Menentukan migrasi bahasa

Untuk menghitung migrasi bahasa menggunakan kesatuan distribusi. Bila bahasa-bahasa kerabat menyebar sedemikian rupa sehingga orang-orang bisa berpindah secara langsung dari suatu daerah bahasa kerabat ke daerah bahasa kerabat yang

jauh hubungan kekerabatannya dan arah migrasinya

Penentuan fonem proto dan dialek proto dari bahasa Batak dialek Toba, Mandailing, Simalungun dan Karo dengan menggunakan cara menyimpulkan suatu fonem menjadi proto dilihat dari empat

4. Analisis dan Hasil Penelitian
a. Kekerabatan Bahasa Batak

NO	KODE DP	NAMA DIALEK	KATA KERABAT	% KEKERABATAN	HUBUNGAN KEKERABATAN
1	1/2	Toba – Mandailing	120	60,0	KELUARGA
2	1/3	Toba – Simalungun	103	51,5	KELUARGA
3	1/4	Toba – Karo	85	42,5	KELUARGA
4	2/3	Mandailing – Simalungun	101	51,0	KELUARGA
5	2/4	Mandailing – Karo	89	44,5	KELUARGA
6	3/4	Simalungun – Karo	110	55,0	KELUARGA

dialek fonem apa yang paling banyak digunakan. Setelah selesai 200 glos baru

Tabel 1 Tabel Persentase Kekerabatan Dialek Bahasa Batak

Dari tabel 1 diperoleh fakta bahwa hubungan kekerabatan bahasa Batak Toba, Mandailing, Simalungun dan Karo adalah keluarga, yang bermakna hubungannya sama dan kosakata yang digunakan hampir sama disebagian daerah. Jadi benarlah secara dialektologi hubungan bahasa keempat daerah pengamatan adalah pada level dialek karena masih berada dalam satu keluarga bahasa yaitu bahasa Batak.

Hanya saja perlu dipahami hubungan bahasa pada level keluarga antara Dialek Toba dan Karo sudah jauh karena persentasenya hampir mendekati rumput yaitu 42%. Hal ini yang memungkinkan antara penutur karu dan Toba tidak saling memahami dalam komunikasi. Begitu juga dengan Mandailing dan Karo yang hampir menyamai Karo dan Toba yaitu 44,5% yang mengindikasikan bahwa penutur Mandailing dan Karo juga sulit dalam berkomunikasi. Data diatas jelaslah menunjukkan ada perbedaan yang cukup besar dalam komunikasi.

b. Fonem proto dan Dialek Proto dari bahasa Batak Dialek Toba, Mandailing, Simalungun dan Karo

dihitung dialek mana yang paling banyak menjadi fonem protonya sehingga dialek tersebut dapat dikatakan sebagai dialek proto bahasa Batak

Diperoleh fonem proto bahasa Batak dari empat dialek untuk setiap glos. Setelah diperoleh fonem proto selanjutnya meletakkan fonem proto tersebut kedalam berian dari empat dialek tersebut. Jika fonem berian dialek mengandung fonem proto pada kolom dialek tersebut diberi tanda bintang, dan jika tidak mengandung fonem proto tidak diberi tanda bintang. Pemberian tanda bintang dimaksudkan untuk memudahkan menghitung berapa glos dari suatu dialek tersebut yang mengandung fonem proto. Semakin banyak tanda bintang maka semakin mungkin dialek tersebut menjadi dialek proto. Penghitungan penentuan dialek proto bahasa Batak dari empat dialek dalam penelitian ini adalah keluarga bahasa Batak yang menjadi dialek protonya adalah dialek Mandailing. Hal ini dibuktikan dari 200 kosakata dasar Swadesh terdapat 116 yang merupakan fonem proto dari empat dialek yang dibandingkan. Temuan ini juga bisa membuktikan bahwa kemungkinannya usia dialek Mandailing lebih tua dari usia dialek yang lain dan merupakan titik awal migrasi.

c. Waktu pisah dan usia bahasa Batak Dialek Toba, Mandailing, Simalungun dan Karo

Tabel 2 Tabel Waktu Pisah dan Usia Dialek Toba, Mandailing, Simalungun dan Karo.

N O.	KOD E DP	DAERAH PENGAMATAN	KAT A KER ABAT	PERSE NTASE KEKE RABAT AN	HUBUN GAN KEKER ABATAN	WAKTU PISAH LAMA RETENSI WAKTU PISAH BARU (dalam ribuan)			USIA SUBDIALEK (dalam ribuan)			
						t lama	$s = \sqrt{c(1-c)/n}$	t baru	Usia bawah	Usia atas	Pisah atas	Pisah bawah
1	½	Toba - Mandailing	120	60	KELUAR GA	1212	0,035	1079	1079	1345	934	668
2	1/3	Toba - Simalungun	103	51,5	KELUAR GA	1575	0,035	1417	1417	1732	596	281
3	¼	Toba - Karo	85	42,5	KELUAR GA	2030	0,035	1843	1843	2218	170	-205
4	2/3	Mandailing - Simalungun	101	50,5	KELUAR GA	1621	0,035	1461	1461	1782	552	231
5	2/4	Mandailing - Karo	89	44,5	KELUAR GA	1921	0,035	1741	1741	2102	272	-89
6	¾	Simalungun - Karo	110	55	KELUAR GA	1419	0,035	1271	1271	1566	742	447

Dari tabel diatas diketahui hubungan kekerabatan antar dialek, waktu pisah antar dialek dan usia dialek. Untuk mempermudah penamaan dialek diberikan nama dialek sesuai daerah pengamatannya, yaitu: DP1 : dialek Toba (selanjutnya disebut BT), DP2: dialek Mandailing (BM), DP3: dialek Simalungun (BS), dan DP4: dialek Karo (BK). Berikut ini adalah uraian dari tabel diatas:

1. Penghitungan waktu pisah dialek BT dan BM adalah 1212 ribuan tahun atau 1079 tahun yang lalu. Dengan kata lain, perhitungan waktu pisah BT dan BM dapat dinyatakan dengan:
 - a. BT dan BM merupakan Dialek Tunggal pada 1212 ± 133 tahun yang lalu.
 - b. BT dan BM merupakan Dialek Tunggal pada 1079 sampai 1345 tahun yang lalu dan termasuk dalam kategori keluarga dari satu bahasa(usia dialek BT dan BM antara 1079—1345tahun)

c. BT dan BM mulai berpisah dari suatu dialek proto antara 668 M sampai 934 M (dihitung dari tahun 2013).

2. Penghitungan waktu pisah dialek BT dan BS adalah 1,575 ribuan tahun

atau 1575 tahun yang lalu. Dengan kata lain, perhitungan waktu pisah BT dan BS dapat dinyatakan dengan:

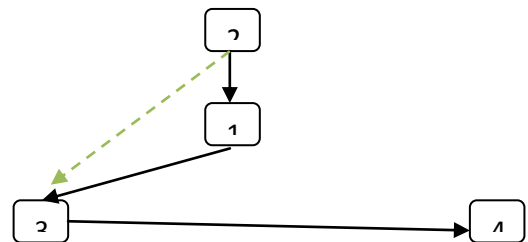
- a. BT dan BS merupakan Dialek Tunggal pada 1575 ± 157 tahun yang lalu.
 - b. BT dan BS merupakan Dialek Tunggal pada 1417 sampai 1732 tahun yang lalu dan termasuk dalam kategori keluarga dari satu bahasa(usia dialek BT dan BS antara 1417—1732 tahun)
 - c. BT dan BS mulai berpisah dari suatu dialek proto antara 281 M sampai 596 M (dihitung dari tahun 2013).
3. Penghitungan waktu pisah dialek BT dan BK adalah 2,030 ribuan tahun atau 2030 tahun yang lalu. Dengan kata lain, perhitungan waktu pisah BT dan BK dapat dinyatakan dengan:
 - a. BT dan BK merupakan Dialek Tunggal pada 2030 ± 188 tahun yang lalu.

- b. BT dan BK merupakan Dialek Tunggal pada 1843 sampai 2218 tahun yang lalu dan termasuk dalam kategori keluarga dari satu bahasa(usia dialek BT dan BK antara 1843—2218 tahun)
 - c. BT dan BK mulai berpisah dari suatu dialek proto antara 205 SM sampai 170 M (dihitung dari tahun 2013).
4. Penghitungan waktu pisah dialek BM dan BS adalah 1,621 ribuan tahun atau 1621 tahun yang lalu. Dengan kata lain, perhitungan waktu pisah BM dan BS dapat dinyatakan dengan:
- a. BM dan BS merupakan Dialek Tunggal pada 1621 ± 161 tahun yang lalu.
 - b. BM dan BS merupakan Dialek Tunggal pada 1461 sampai 1782 tahun yang lalu dan termasuk dalam kategori keluarga dari satu bahasa(usia dialek BM dan BS antara 1461—1782 tahun)
 - c. BM dan BS mulai berpisah dari suatu dialek proto antara 231 M sampai 552 M (dihitung dari tahun 2013).
5. Penghitungan waktu pisah dialek BM dan BK adalah 1,921 ribuan tahun atau 1921 tahun yang lalu. Dengan kata lain, perhitungan waktu pisah BM dan BK dapat dinyatakan dengan:
- a. BM dan BK merupakan Dialek Tunggal pada 1921 ± 180 tahun yang lalu.
 - b. BM dan BK merupakan Dialek Tunggal pada 1741 sampai 2102 tahun yang lalu dan termasuk dalam kategori keluarga dari satu bahasa(usia dialek BM dan BK antara 1741—2102 tahun)
 - c. BM dan BK mulai berpisah dari suatu dialek proto antara 89 SM sampai 272 M (dihitung dari tahun 2013)
6. Penghitungan waktu pisah dialek BS dan BK adalah 1,419 ribuan tahun atau 1419 tahun yang lalu. Dengan kata lain, perhitungan waktu pisah BS dan BK dapat dinyatakan dengan:
- a. BS dan BK merupakan Dialek Tunggal pada 1419 ± 147 tahun yang lalu.
 - b. BS dan BK merupakan Dialek Tunggal pada 1271 sampai 1566 tahun yang lalu dan termasuk dalam kategori keluarga dari satu bahasa(usia dialek BS dan BK antara 1271—1566 tahun)
 - c. BS dan BK mulai berpisah dari suatu dialek proto antara 447 M sampai 742 M (dihitung dari tahun 2013)

Dari penghitungan di atas dapat diketahui waktu pisah terlama dari empat dialek ini adalah dialek Karo dengan dialek yang lain yaitu Karo dengan Toba berpisah sekitar 1843—2218 tahun yang lalu, Karo dengan mandailing sekitar 1741—2102 tahun yang lalu, serta Karo dengan Simalungun berpisah sekitar 1271 – 1566 tahun yang lalu. Dengan demikian dialek Karo merupakan dialek yang lebih dahulu memisah dari Mandailing sebagai dialek proto. Hal ini membuktikan bahwa dialek Karo mempunyai hubungan kekerabatan yang paling jauh dengan dialek Mandailing, Toba dan Simalungun.

d. Migrasi Bahasa Batak Dialek Toba, Mandailing, Simalungun dan Karo

Dari data diperoleh gambaran arah migrasi dialek Batak dengan daerah asal adalah DP 2 sebagai berikut:



Bagan 1. Gambaran Arah Migrasi Dialek Bahasa Batak berdasarkan Persentase Hubungan Kekerabatan(Leksikostatistik)

Dari bagan di atas diketahui bahwa *homeland* (daerah asal) migrasi bahasa Batak berasal dari DP 2 yaitu daerah pengamatan Mandailing. Migrasi bergerak ke wilayah Toba (DP 1). Dari wilayah Toba (DP 1) migrasi dialek Batak bergerak ke wilayah Simalungun (DP 3), namun bisa juga dari DP 1 migrasi bergerak ke DP 3 secara langsung karena persentase kekerabatannya cukup dekat dari Mandailing maupun dari Toba. Sedangkan migrasi ke Karo bergerak dari Simalungun, dari DP 3 ke DP 4.

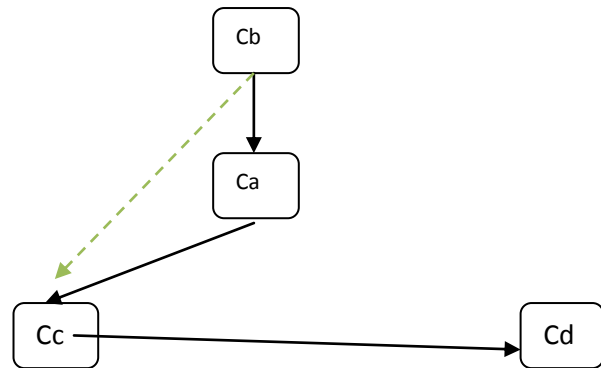
Dengan melihat hipotesa migrasi bahasa menurut Keraf yang bisa digunakan dalam penelitian migrasi dialek Batak ini adalah hipotesa: H = Ca. Dengan H adalah *homeland*, daerah asal migrasi dan Ca adalah mata rantai dialek Batak. Penggunaan arah migrasi ini dengan alasan geografis bahwa masyarakat pengguna bahasa Batak yang ada pada setiap daerah pengamatan dapat berpindah secara langsung maka dialek Bahasa Batak merupakan satu mata rantai migrasi. Yaitu:

H = Cb
 Cb → Ca
 Ca → Cc
 Cc → Cd

Dengan ketentuan:

- Cb adalah *homeland* migrasi dialek Bahasa Batak yaitu daerah pengamatan Mandailing
- Ca adalah mata rantai dialek Bahasa Batak daerah Toba
- Cc adalah mata rantai dialek Bahasa Batak daerah Simalungun
- Cd adalah mata rantai dialek Bahasa Batak daerah Karo
- → adalah penunjuk arah migrasi dialek Bahasa Batak

Dengan demikian bagan migrasi Dialek Bahasa Batak menjadi dialeknnya adalah satu mata rantai yang diawali dari Cb:



Bagan 2. Migrasi dialek Bahasa Batak

Jadi dapat disimpulkan bahwa migrasi dialek Bahasa Batak menjadi subdialek-dialeknnya adalah dari Mandailing(Cb) sebagai daerah asal dialek Batak bergerak menuju atau kearah daerah Toba(Ca) lalu bergerak ke daerah Simalungun (Cc), Dari Daerah Simalungun migrasi dialek Batak bergerak ke daerah Karo (Cd).

5. Simpulan dan Saran

a. Simpulan

Dengan menggunakan penghitungan leksikostatistik yang merupakan kajian linguisitik historis komparatif dan penghitungan glotokronologi, penelitian yang menggunakan metode kuantitatif ini menghasilkan empat kesimpulan sebagai hasil analisis dari empat masalah yang dikaji. Berikut adalah simpulan hasil penelitian ini:

1. Berdasarkan kajian 200 kosakata Swadesh ditemukan hubungan kekerabatan antara Bahasa Batak yang terdapat dalam empat daerah pengamatan yaitu Toba, Mandailing, Simalungun dan Karo menghasilkan hubungan keluarga bahasa. Dengan kata lain hubungan Toba dengan Mandailing sebesar 60%, Toba dengan Simalungun sebesar 51,5% dan Toba dengan Karo sebesar 42,5% adalah satu keluarga bahasa yaitu keluarga bahasa Batak. Begitu juga Mandailing dengan Simalungun sebesar 51% dan Mandailing dengan Karo sebesar 44,5% serta hubungan Simalungun dengan Karo sebesar 55,0% adalah satu keluarga.

Meskipun dalam satu keluarga namun ditemukan bahwa antara bahasa Karo dengan Toba sangat kecil persentase kekerabatannya yaitu 42,5% yang mendekati persentase hubungan rumpun yang berarti hampir berbeda bahasa. Begitu juga Karo dengan Mandailing dan Simalungun. Perbedaan yang ada diperkirakan akan terus semakin jauh seiring perubahan zaman.

2. Dari keempat daerah pengamatan keluarga bahasa ini diperoleh hasil dialek proto untuk keempatnya yaitu Mandailing. Penentuan dialek proto berdasarkan banyak fonem proto yang dikandung daerah pengamatan Mandailing dibandingkan Toba, Simalungun dan Karo. Dengan demikian proto dialek dari bahasa Batak dalam penelitian ini adalah Mandailing.
3. Penghitungan waktu pisah dan usia dialek Bahasa Batak dari empat pengamatan yaitu Toba, Mandailing, Simalungun dan Karo menghasilkan data berikut:
 - a. Waktu pisah Toba dengan Mandailing adalah 1079 sampai 1345 tahun yang lalu dan usianya ada sejak sekitar 668M sampai 934M
 - b. Waktu pisah Toba dengan Simalungun adalah 1417 sampai 1732 tahun yang lalu dan usianya ada sejak sekitar 281M sampai 596M
 - c. Waktu pisah Toba dengan Karo adalah 1843 sampai 2218 tahun yang lalu dan usianya ada sejak sekitar 205 SM sampai 170M
 - d. Waktu pisah Mandailing dengan Simalungun adalah 1461 sampai 1782 tahun yang lalu dan usianya ada sejak sekitar 231M sampai 552M
 - e. Waktu pisah Mandailing dengan Karo adalah 1741 sampai 2102

tahun yang lalu dan usianya ada sejak sekitar 89SM sampai 272M

- f. Waktu pisah Simalungun dengan Karo adalah 1271 sampai 1566 tahun yang lalu dan usianya ada sejak sekitar 447M sampai 742M.
- g. Dengan demikian Karo mempunyai waktu pisah dan ada sudah sejak lama memisahkan diri dari protonya
4. Migrasi bahasa yang terjadi adalah satu rantai yang melibatkan Mandailing sebagai daerah asal migrasi. Pergerakan migrasi menuju ke Toba lalu bergerak ke Simalungun dan akhirnya ke Karo. Dengan usia dan waktu pisah Karo yang sudah begitu lama, Karo sudah memisahkan diri dan migrasi terjadi sudah lama sekali dari Mandailing ke Karo.

b. Saran

Peneliti menyarankan beberapa hal kepada pembaca dan peneliti lain yang akan melakukan penelitian sejenis yaitu:

- a. Perlu kajian mendalam dengan menggunakan data kosa kata budaya sebagai pelengkap penelitian
- b. Melaksanakan kajian sejarah terhadap migrasi penduduk untuk menentukan migrasi bahasa pada masa lalu

6. Daftar Pustaka

- Asfar, Dedy Ari. 2009. *Refleks Fonem Proto Melayu Polinesia dalam varian Bidayuhik Ribun*. Pontianak: Balai Bahasa Kalimantan Barat
- Azhar, Ridwan. 1989. *Sejarah Studi Bahasa di Indonesia*. Medan : Fakultas Sastra USU
- Blust dan Collins. 1985. *Telaah Komparatif Bahasa-Bahasa Nusantara Barat*. ILDEP. Jakarta: Djambatan.
- Crowley, T. 1992. *An Introduction to Historical Linguistics*. Oxford: Oxford University Press.
- Fernandes, Inyo Yos. 1994. *Linguistik Historis Komparatif: Pengantar di Bidang*

9MEDAN MAKNA	Vol. XI No. 1	Hlm. 61 - 75	2013	ISSN 1829-9237
--------------	---------------	--------------	------	----------------

- Teori*. Yogyakarta: Buku Tidak Diterbitkan.
- Fernandes, Inyo Yos. 2005. *Linguistik Historis Komparatif I: Ke Arah Pemahaman Teori dan Praktik Perbandingan Bahasa-Bahasa Austronesia* (Handout). Yogyakarta.
- Halim, Amran. 1988. *Politik Bahasa Nasional I*. Jakarta: Balai Pustaka.
- Indriani, Ika. 2007. *Leksikostatistik Bahasa Batak Toba dengan Bahasa Pakpak Dairi*. Tesis.
- Keraf, Gorys. 1991. *Linguistik Bandingan Tipologis*. Jakarta: Gramedia Pustaka Utama.
- Keraf, Gorys. 1984. *Linguistik Bandingan Historis*. Jakarta: Gramedia Pustaka Utama.
- Kridalaksana, Harimurti. 1963. "Perhitungan Leksikostatistik atas Delapan Bahasa Nusantara Barat serta Penentuan Pusat Penyebaran Bahasa-Bahasa itu Berdasarkan Teori Migrasi". dalam *Majalah Ilmu-Ilmu Sastra 1963-1973*. Jakarta.
- Kridalaksana, Harimurti. 2008. *Kamus Linguistik*. Jakarta: PT Gramedia Pustaka Utama.
- Mahsun. 1995. *Dialektologi Diakronis: Sebuah Pengantar*. Yogyakarta: Gadjah Mada University Press.
- Mahsun. 2005. *Metode Penelitian Bahasa*. Jakarta: Raja Grafindo Persada.
- Masyuri dan Zainuddin. 2008. *Metodologi Penelitian: Pendekatan Praktis dan Aplikatif*. Malang: Refika Aditama.
- Mbete, Aron Meko. 2002. *Metode Linguistik Diakronis*. Bali: Universita Udayana.
- Mukhamdanah, dkk. 2009. *Dialektologi Bahasa Daerah di Sumatera Utara*. Medan: Balai Bahasa Medan
- Panggabean, Himpun. 1994. *Telaah Bahasa-Bahasa Batak dari Segi Leksikostatistik*. Bandung: Universitas Padjajaran. (Tesis)
- Parera, Jos Daniel. 1991. *Kajian Linguistik Umum Historis Komparatif*. Jakarta: Erlangga. (Digitalkan tanggal 25 April 2008.)
- Ridwan, T.Amin. 2005. *Mendaulatkan Bahasa Melayu Sebagai Bahasa Utama Dunia*. Medan: USU Press.
- Robins, R. H. 1992. *Linguistik Umum: Sebuah Pengantar*. Yogyakarta: Kanisius.
- SIL International. 2006. *Bahasa-Bahasa di Indonesia*. Jakarta: SIL Internasional Cabang Indonesia.
- Siregar, Bahren Umar, dkk. 2000. *Reduplikasi dan Komposisi dalam Bahasa Batak Toba*. Dekdikbud. (Laporan Penelitian)
- Sudaryanto. 1986. *Metode Linguistik 1: Ke Arah Memahami Metode Linguistik*. Yogyakarta: Gadjah Mada University Press.
- Sudaryanto. 1988. *Metode Linguistik 2: Metode dan Aneka Teknik Pengumpulan Data*. Yogyakarta: Gadjah Mada University Press.
- Sugiyono. 2008. *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Sugono, Dendi (ed). 2008. *Kamus Besar Bahasa Indonesia*. Jakarta: Pusat Bahasa, Departemen Pendidikan.
- Suyata, Pujiati. 1998. *Dari Leksiostatistik ke Glotokronologi: Analisis Sembilan Bahasa di Indonesia*.
- Syarfina, Tengku. 2009. *Ciri Akustik Bahasa Melayu Deli*. Medan: USU Pres.
- Wurm, S.A. 1978. *English Findex of Reconstructions in Austronesian Languages*. Canberra: The Australian National University
- Badan Pusat Statistik Sumatera Utara. 2012
- Wikipedia.
http://id.wikipedia.org/wiki/Bahasa_Batak_Toba. Diunduh 11/5/2013
- Wikipedia.
http://id.wikipedia.org/wiki/Bahasa_Mandailing. Diunduh 11/5/2013
- Wikipedia.
http://id.wikipedia.org/wiki/Bahasa_Batak_Simalungun. Diunduh 11/5/2013
- Wikipedia.
http://id.wikipedia.org/wiki/Bahasa_Karo. Diunduh 11/5/2013

9MEDAN MAKNA	Vol. XI No. 1	Hlm. 61 - 75	2013	ISSN 1829-9237
--------------	---------------	--------------	------	----------------