EKSPLORASI HUBUNGAN GENETIK BAHASA BATAK KARO DAN BAHASA BATAK TOBA: PENDEKATAN LEKSIKOSTATISTIK

Ria Clara Ginting¹, Warisman Sinaga², Ramlan Damanik³

Universitas Sumatera Utara^{1,2,3}

pos-el: riaclaraq123@gmail.com¹, warisman@usu.ac.id², ramlan1@usu.ac.id³

ABSTRAK

Penelitian ini bertujuan untuk mendeskripsikan hubungan genetik antara bahasa Batak Karo dan Batak Toba, menentukan waktu perpisahan linguistik mereka, dan memperkirakan usia kedua bahasa tersebut. Menggunakan teori linguistik historis komparatif, khususnya pendekatan leksikostatistik dan glotokronologi yang dikemukakan oleh Gorys Keraf, penelitian ini menyediakan kerangka analisis untuk memahami evolusi linguistik dari bahasa-bahasa ini. Data untuk penelitian ini dikumpulkan melalui kombinasi metode observasi, wawancara, teknik pencatatan, dan perekaman. Data tersebut kemudian dianalisis menggunakan metode deskriptif kuantitatif. Hasil penelitian menunjukkan bahwa dari 500 kosakata yang diteliti, terdapat 320 pasangan kata kerabat yang menunjukkan hubungan genetik sebesar 65%. Perpisahan antara bahasa Batak Karo dan Batak Toba diperkirakan terjadi sekitar tahun 1028 M atau pada abad ke-11, sekitar 995 tahun yang lalu (dihitung dari tahun 2023). Hasil ini memiliki implikasi penting bagi klasifikasi dan pemahaman hubungan genetik dalam bahasa-bahasa Batak. Temuan ini menekankan pentingnya penelitian komparatif yang lebih komprehensif yang melibatkan semua varian bahasa Batak guna mendapatkan gambaran yang lebih lengkap tentang hubungan genetik dalam keluarga bahasa Batak. Penelitian ini tidak hanya berkontribusi pada studi linguistik historis komparatif bahasa-bahasa Austronesia di Indonesia, tetapi juga menyoroti pentingnya melestarikan keunikan dan kekayaan linguistik setiap varian bahasa Batak.

Kata Kunci: *Hubungan Genetik, Karo, Toba, Leksikostatistik, Pelestarian Bahasa.*

ABSTRACT

This research aims to describe the genetic relationship between the Karo Batak and Toba Batak languages, determine the timeline of their linguistic divergence, and estimate the age of these languages. Using the comparative historical linguistics theory, specifically the lexicostatistical and glottochronological approach by Gorys Keraf, this study provides an analytical framework for understanding the linguistic evolution of these languages. Data for this research was gathered through a combination of observation, interviews, note-taking, and recording techniques. The data was then analyzed using a quantitative descriptive method. The findings reveal that out of 500 vocabularies studied, 320 pairs are genetically related, indicating a 65% genetic relationship. The divergence between the Karo Batak and Toba Batak languages is estimated to have occurred around 1028 AD or in the 11th century, approximately 995 years ago (calculated from 2023). These results have significant implications for the classification and understanding of the genetic relationships within the Batak languages. They highlight the need for further comprehensive comparative studies involving all variants of the Batak languages to obtain a more complete picture of the genetic relationships within the Batak language family. This research not only contributes to the historical comparative linguistic studies of Austronesian languages in Indonesia but also underscores the importance of preserving the linguistic uniqueness and richness of each Batak language variant.

Keywords: *Genetic Relationship, Karo, Toba, Lexicostatistics, Language Preservation.*

1. PENDAHULUAN

Indonesia ialah negara yang memiliki banyak kekayaan terkhusus pada keanekaragaman suku, budaya, serta bahasa. Setiap suku memiliki budaya dan bahasanya sendiri yang berbeda dengan suku-suku lainnya. Sebagai negara yang kaya, Indonesia memiliki banyak bahasa daerah. Menurut *National Agency for the Development of Language and Books* (dalam Afria et al., 2020) Indonesia mempunyai 719 bahasa yang tersebar dari Sabang hingga Merauke bersumber pada data 2019 Badan Pengembangan Bahasa dan Perbukuan.

Bahasa merupakan aspek penting yang ada dalam kehidupan manusia, sebab fungsinya adalah sebagai alat interaksi sosial dalam masyarakat. Dalam Chaer (2014:32) menyatakan bahasa sebagai alat untuk berkomunikasi dan bekerja sama yang digunakan oleh kelompok sosial berupa sistem lambang bunyi yang arbitrer. Bahasa pada dasarnya merupakan identitas diri penuturnya, karena bahasa merupakan ciri khas yang dapat menandai seseorang, sekelompok orang (masyarakat tertentu), ataupun suatu bangsa.

Bahasa merupakan objek kajian bidang ilmu linguistik. Menurut Chaer (2018:4) salah satu pembedaan ilmu linguistik adalah linguistik sejarah (linguistik historis). Linguistik historis merupakan ilmu yang mempelajari bahasa serta perubahan unsur-unsurnya pada suatu bidang waktu Keraf (2018:22). Tujuannya adalah untuk mengadakan suatu perbandingan mengenai unsur-unsur yang menunjukkan kekerabatan antar bahasa serumpun. Dalam linguistik historis, untuk mengadakan suatu perbandingan diperlukan sebuah metode, metode ini disebut leksikostatistik.

Leksikostatistik merupakan teknik yang lebih memprioritaskan peneropongan kata-kata melalui statistik

dalam pengelompokan bahasa Keraf (2018:121). Dengan demikian, pendekatan ini akan mengelompokkan bahasa berlandaskan persentase kesamaan dan perbedaannya. Metode leksikostatistik ini digunakan untuk mengetahui tingkat kesamaan antar bahasa melalui pengelompokan bahasa. Sejalan dengan leksikostatistik digunakan juga metode glotokronologi. Glotokronologi merupakan metode yang melakukan pengelompokan bahasa yang memprioritaskan perhitungan usia ataupun waktu Keraf (2018:121). Dalam glotokronologi usia bahasa dihitung memakai satuan ribuan tahun.

Bahasa Batak ialah satu di antara bahasa daerah di Indonesia. Bahasa Batak terdiri dari lima bahasa berdasarkan dengan sub etnis Batak, yaitu bahasa Simalungun, Angkola Mandailing, Karo, Pakpak, serta bahasa Toba. Bahasa yang akan dikaji adalah bahasa Batak Karo dan Batak Toba. Bahasa Batak Karo dan Batak Toba dipilih sebagai objek dari penelitian ini karena meskipun daerah penggunaan bahasa Batak Karo dan bahasa Batak Toba berseberangan, dimana bahasa Batak Karo berada di sebelah Utara sedangkan bahasa Batak Toba berada di sebelah Selatan, dan dibatasi oleh daerah penggunaan bahasa Simalungun dan Pakpak, namun kedua bahasa ini tetap memiliki kesamaan atau kemiripan di antara kedua bahasanya, yaitu kedua bahasa ini merupakan dua bahasa dengan rumpun yang sama (yaitu rumpun keluarga bahasa Batak). Sehingga kedua bahasa ini dapat dikatakan sebagai bahasa yang berkerabat atau bahasa yang memiliki hubungan genetik. Oleh sebab itu peneliti merasa perlu untuk melakukan penelitian terhadap dua bahasa Batak ini, untuk mengetahui seberapa besar hubungan genetik/ hubungan kekerabatan yang dimiliki oleh kedua bahasa ini.

Bahasa Batak Karo dan bahasa Batak Toba, kedua bahasa ini memiliki

keunikan, yaitu: penggunaan kata dalam kedua bahasa ini menggunakan kata bahasa yang kasar dan kata bahasa yang halus. Penggunaan kata bahasa yang kasar dan halus ini bergantung pada siapa dan pada situasi apa digunakan. Biasanya kata bahasa yang kasar digunakan kepada orang/ lawan tutur yang lebih muda atau setara dengan penutur, kata bahasa yang kasar ini lebih sering digunakan pada pergaulan kehidupan sehari-hari, sedangkan kata bahasa yang halus digunakan kepada orang/ lawan tutur yang lebih tua dari penutur, kata bahasa yang halus ini lebih sering digunakan dalam kegiatan atau upacara adat/ acara formal suku Batak. Selain itu kedua bahasa Batak ini sama-sama memiliki beberapa ragam dialek. Menurut Tarigan (dalam Sukapiring, 2019:12) bahasa Karo memiliki tiga dialek bahasa, yaitu dialek Kabanjahe, dialek Gunung-Gunung, dan dialek Jahe-jahe. Sedangkan pada bahasa Batak Toba, bahasa Batak Toba memiliki lima variasi bahasa (dialek), yaitu dialek Toba, Silindung, Sibolga, Samosir, dan dialek Humbang (Hutabarat et al., 2013:3).

Menurut Syafi'i dan Ibrahim (2019) bahasa kerabat ialah perkumpulan bahasa yang termasuk pada sebuah kelompok bahasa yang mempunyai sejarah perkembangan yang tidak berbeda/ pernah membentuk satu bahasa proto yang serupa. Kedua bahasa Batak tersebut juga dipilih karena latar belakang penulis, dimana penulis merupakan keturunan dari dua sub-etnis Batak yang berbeda, yaitu etnis Batak Karo dan etnis Batak Toba. Selama hidup dan berbahasa di dalam keluarga, penulis menyadari bahwa kedua bahasa tersebut merupakan bahasa dengan rumpun yang sama yang kemudian banyak ditemukan kesamaan pada kedua bahasa ini karena hubungan genetik yang kedua bahasa ini miliki, karena itu penulis merasa penting untuk melakukan eksplorasi atau penelitian terhadap

hubungan genetik yang dimiliki oleh kedua bahasa ini dengan menggunakan metode leksikostatistik dan glotokronologi, karena dengan menggunakan teknik leksikostatistik kekerabatan/ hubungan genetik bahasa dapat dilihat dari persamaan bunyi dalam leksikon/kata yang ada pada bahasa Batak Karo dan Batak Toba, dengan begitu akan dapat diketahui seberapa besar hubungan genetik dari kedua bahasa tersebut, serta perubahan-perubahan yang terjadi pada bahasa tersebut, lalu dengan menggunakan teknik glotokronologi akan dapat diketahui waktu pisah antara kedua bahasa tersebut.

Penelitian ini dibangun di atas dasar penelitian-penelitian terdahulu yang telah memberikan kontribusi signifikan dalam pemahaman kekerabatan bahasa-bahasa Batak. Studi Damanik (2019) tentang leksikostatistik bahasa Batak Simalungun dan bahasa Batak Mandailing, serta penelitian Napitupulu dan Silaban (2020) mengenai telaah leksikostatistik terhadap bahasa Batak Toba dan bahasa Batak Angkola Mandailing menjadi rujukan penting, baik dalam hal objek kajian (rumpun bahasa Batak) maupun metodologi analisis yang digunakan.

Lebih lanjut, studi-studi terdahulu telah mengungkapkan temuan-temuan menarik tentang hubungan genetik antara bahasa-bahasa Batak. Wartono (2018) menemukan bahwa dialek-dialek Batak (Toba, Mandailing, Simalungun, dan Karo) termasuk dalam keluarga bahasa yang sama, dengan Mandailing diidentifikasi sebagai proto-dialek. Pemisahan dialek-dialek ini diperkirakan terjadi antara 1000-2000 tahun yang lalu, sekitar tahun 205 SM hingga 934 M. Penelitian Napitupulu dan Silaban (2020) mengungkapkan hubungan kognat sebesar 55% antara Batak Toba dan Batak Angkola Mandailing, menunjukkan koneksi tingkat keluarga dengan pemisahan terjadi sekitar tahun

640 M. Temuan serupa juga dilaporkan oleh Mualita (2015), yang menemukan hubungan kognat sebesar 57% antara Batak Toba dan Batak Angkola, dengan perkiraan pemisahan pada tahun 681 M. Sementara itu, Dardanila (2018) dalam penelitiannya tentang bahasa Karo dan Gayo menunjukkan hubungan kognat sebesar 43,5%, mengindikasikan pemisahan sekitar 1.926 tahun yang lalu.

Dengan demikian, penelitian ini tidak hanya bertujuan untuk mengungkap hubungan genetik antara bahasa Batak Karo dan Batak Toba, tetapi juga untuk memberikan kontribusi pada pemahaman yang lebih luas tentang keanekaragaman dan evolusi bahasa di Indonesia. Melalui analisis leksikostatistik dan glotokronologi, penelitian ini diharapkan dapat memberikan wawasan baru tentang sejarah, perkembangan, dan hubungan antara dua bahasa Batak yang signifikan ini, serta memperkaya kajian linguistik historis di Indonesia. Dengan mempertimbangkan temuan-temuan dari penelitian terdahulu, studi ini akan dapat memosisikan hubungan antara Batak Karo dan Batak Toba dalam konteks yang lebih luas dari evolusi bahasa-bahasa Batak, serta memberikan kontribusi pada pemahaman yang lebih komprehensif tentang sejarah linguistik di wilayah Sumatera Utara.

Rumusan masalah berdasarkan latar belakang penelitian ini adalah:

- 1) Berapa persentase hubungan genetik bahasa Batak Karo dan bahasa Batak Toba?
- 2) Kapan perkiraan waktu pisah antara kedua bahasa tersebut ?
- 3) Berapa usia antara kedua bahasa tersebut ?

2. METODE PENELITIAN

Metode penelitian kuantitatif dengan pendekatan leksikostatistik dan glotokronologi adalah metode yang digunakan pada penelitian ini. Sugiyono (2018:13) mendefinisikan metode

kuantitatif sebagai metode yang data penelitiannya berlandaskan data konkrit. Penelitian ini menggunakan data berupa angka yang diukur, sebagai alat untuk menguji perhitungan yang berhubungan terhadap objek penelitian guna memperoleh suatu kesimpulan.

Leksikostatistik adalah teknik yang memprioritaskan peneropongan kata-kata secara statistik dalam pengelompokan bahasa (Keraf, 2018:121). Metode ini mengelompokkan bahasa berdasarkan persentase kesamaan dan perbedaannya. Dalam menetapkan pasangan kata kerabat, teknik leksikostatistik mengikuti prosedur tertentu. Langkah-langkah prosedur ini meliputi: pertama, mengidentifikasi gloss yang tidak diperhitungkan, yaitu kata yang tidak memiliki pasangan atau merupakan kata jadian. Kedua, melakukan pengisolasian morfem terikat untuk mempermudah penetapan pasangan kata kerabat. Ketiga, menetapkan kata kerabat berdasarkan kriteria tertentu. Kriteria ini mencakup: (a) pasangan kata yang identik, yaitu yang memiliki fonem dan maksud yang sama persis; (b) pasangan kata dengan korespondensi fonemis, di mana terjadi perubahan fonemis secara teratur dan timbal balik; (c) pasangan kata dengan kemiripan fonetis dalam artikulatoris yang sama; dan (d) pasangan kata dengan satu fonem berbeda namun tetap memiliki kemiripan fonetis dalam artikulatoris yang sama. Dengan mengikuti prosedur dan kriteria ini, penelitian dapat secara sistematis mengidentifikasi dan menganalisis hubungan kekerabatan antar bahasa yang diteliti.

Setelah melakukan perhitungan jumlah kata yang memiliki hubungan genetik/ kata berkerabat, selanjutnya jumlah kata kerabat ini dibagi dengan jumlah kata yang dibandingkan (tidak termasuk gloss yang tidak diperhitungkan) lalu dikali dengan 100%, kemudian hasil dari inilah yang

menjadi persentase kata kerabat. Kosakata yang digunakan pada penelitian ini adalah kosakata oleh Mahsun yang berupa 500 kosa kata, semua kosakata ini dapat ditemukan di semua bahasa.

Teknik glotokronologi menurut Keraf (2018:121) merupakan teknik yang melakukan pengelompokan bahasa dengan mengutamakan perhitungan usia. Dalam glotokronologi usia bahasa dihitung secara umum, memakai satuan ribuan tahun (milenium). Waktu pisah bahasa dalam teknik glotokronologi hanya dapat dihitung setelah persentase kata kerabat kedua bahasa diketahui. Setelah itu carilah logaritma dari persentase kata kerabat dan logaritma dari retensi pada tabel logaritma, kemudian logaritma retensi dikali 2, lalu hasil logaritma persentase kata kerabat dibagi dengan hasil dari logaritma retensi yang dikali 2, kemudian hasil dari pembagian inilah yang menunjukkan waktu pisah bahasa dalam satuan ribuan tahun.

Lokasi penelitian pada Desa Sibandang, Kecamatan Muara, Kabupaten Tapanuli Utara dan Desa Lingga, Kecamatan Simpang Empat, Kabupaten Karo. Sumber data penelitian bersumber dari penelitian lapangan, dimana data didapatkan dari kegiatan penelitian dengan pengamatan dan wawancara langsung kepada informan tentang daftar 500 kosa kata oleh Mahsun (Mahsun, 1995), nantinya jawaban yang diperoleh berupa bahasa Batak Karo dan bahasa Batak Toba dari daftar 500 kata tersebut. Informan yang digunakan pada penelitian ini sebanyak 2 orang dengan kriteria sebagai berikut: 1) pria atau wanita usia 25-60 tahun, 2) informan merupakan masyarakat yang lahir dan besar di desa tersebut, 3) menguasai bahasa Batak Karo/ bahasa Batak Toba, 4) sehat jasmani maupun rohani. Instrumen pada penelitian ini adalah Tape recorder (alat rekam), dan alat tulis (kertas dan pulpen), untuk

merekam dan mencatat hasil wawancara dari informan berupa 500 kata bahasa Batak Karo dan Batak Toba. Instrumen pada penelitian ini adalah Tape recorder (alat rekam), dan alat tulis (kertas dan pulpen), untuk merekam dan mencatat hasil wawancara dari informan berupa 500 kata bahasa Batak Karo dan Batak Toba.

Proses pengumpulan data penelitian ini dikumpulkan dengan: 1) teknik observasi (pengamatan secara langsung terhadap objek penelitian), 2) wawancara langsung kepada informan, 3) teknik catat dan teknik rekam (peneliti mencatat serta merekam hasil wawancara agar nantinya dapat memperhatikan pelafalan informan dengan baik). Setelah memperoleh data dari penelitian, data kemudian diolah dengan cara:

- 1) Mentranskripsikan data rekaman wawancara
- 2) Menghitung jumlah kata kerabat yang memiliki hubungan genetik berdasarkan prosedur leksikostatistik
- 3) Membuat persentase kata kerabat hubungan genetik menggunakan rumus:

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

(Keraf 2018:124)

- 4) Menghitung waktu pisah bahasa Batak Karo dan Toba dengan rumus:

$$W = \frac{\text{Log. C}}{2\text{Log. r}}$$

(Keraf 2018:130)

- 5) Menarik kesimpulan serta saran.

3. HASIL DAN PEMBAHASAN

Berdasarkan temuan penelitian bahasa Batak Karo dan Batak Toba yang diperoleh di lapangan, dari 500 kosa kata atau gloss oleh Mahsun (Mahsun, 1995) yang telah diklasifikasikan, terdapat 490 pasangan kata, 320 pasang kata kerabat, dan 170 pasang kata yang tidak

berkerabat, selebihnya yaitu 10 kata adalah kata yang tidak diperhitungkan.

Menghitung Kata Kerabat

Kata kerabat bahasa Batak Toba dan Batak Karo dihitung dengan menggunakan prosedur-prosedur leksikostatistik, yaitu:

a. Gloss yang tidak diperhitungkan
Pasangan kata ini adalah kata (gloss) yang tidak mempunyai pasangan, dan gloss yang merupakan kata jadian (bukan kata dasar) karena dalam data gloss terdapat gloss yang sama yang merupakan kata dasar.

Tabel 1. Gloss yang tidak diperhitungkan

No	Gloss	Bahasa Batak Karo	Bahasa Batak Toba
93	dukun bayi	--	<i>simargala gala</i>
99	melaHIRkan	<i>mupus</i>	<i>manubuh on</i>
251	di atas	<i>idatas</i>	<i>diginjang</i>
309	berjalan	<i>ardalan</i>	<i>mardalan</i>
317	berobat	<i>artambar</i>	<i>martambar</i>

Gloss nomor 93 merupakan gloss yang tidak mempunyai pasangan, sedangkan gloss nomor 99, 251, 309, dan 317 merupakan gloss yang merupakan kata jadian.

b. Pengisolasian morfem terikat
Isolir morfem terikat harus dilakukan agar lebih mudah dalam menetapkan pasangan kata kerabat.

Tabel 2. Pengisolasian morfem terikat

No	Gloss	Bahasa Batak Karo	Bahasa Batak Toba	Morfem Terikat
305	bicara (ber)	<i>arkata</i>	<i>marhata</i>	/ər/~/mar/
307	renang (ber)	<i>arlangi</i>	<i>marlange</i>	/ər/~/mar/
318	siul (ber)	<i>arsiwel</i>	<i>marsiul</i>	/ər/~/mar/

341	(mem) bawa di bahu	<i>ngusung</i>	<i>mangusung</i>	/ng/~/mang/
345	(mem) buru hewan	<i>arburu</i>	<i>marburu</i>	/ər/~/mar/
457	kejut (ter)	<i>tarsənggət</i>	<i>tarsonggot</i>	/tər/~/tar/

Dalam hal ini isolasi morfem terikat perlu dilakukan agar dapat mengetahui bentuk dasar atau kata dasar dari gloss yang memiliki morfem terikat, hal ini dilakukan agar dapat lebih mudah untuk mengetahui apakah pasangan kata yang dibandingkan berkerabat atau tidak.

c. Penetapan Kata Kerabat

Sebuah pasangan kata dianggap berkerabat jika memenuhi ketentuan leksikostatistik berikut:

a) Pasangan Kata Identik

Apabila pasangan kata ini mengandung fonem serta maksud yang sama betul maka pasangan kata ini disebut pasangan identik.

Tabel 3. Pasangan kata identik

No	Gloss	Bahasa Batak Karo	Bahasa Batak Toba
13	hidung	<i>igung</i>	<i>igung</i>
68	siapa	<i>ise</i>	<i>ise</i>
116	rumah	<i>ruma</i>	<i>ruma</i>
142	tikar	<i>amak</i>	<i>amak</i>
164	akar	<i>urat</i>	<i>urat</i>
175	daun	<i>bulung</i>	<i>bulung</i>
191	rebung	<i>tubis</i>	<i>tubis</i>
266	hujan	<i>udan</i>	<i>udan</i>
338	bakar	<i>tutung</i>	<i>tutung</i>
375	mimpi	<i>nipi</i>	<i>nipi</i>

Gloss pasangan kata identik merupakan gloss yang memiliki makna dan fonem-fonem yang sama pada kedua bahasa yang dibandingkan.

b) Pasangan Kata Yang berkorespondensi fonemis

Apabila pasangan kata ini mengandung perubahan fonemis yang terjadi secara teratur dan timbal balik, maka bentuk yang berimbang dianggap berkerabat.

Tabel 4. Pasangan kata yang berkorespondensi fonemis

No	Gloss	Bahasa Batak Karo	Bahasa Batak Toba	Korespondensi Fonemis
1	alis	<i>səlibən</i>	<i>salibon</i>	/ə~/a/ /ə~/o/
19	jari tengah	<i>jari tənəgəh</i>	<i>jari tonga</i>	/ə~/o/ /h~/ə/
29	lema k	<i>tabəht abəh</i>	<i>tabotabo</i>	/ə~/o/ /h~/ə/
135	periu k	<i>kudin</i>	<i>hudon</i>	/k~/h/ /i~/o/
159	nenas	<i>kənas</i>	<i>honas</i>	/k~/h/ /ə~/o/
192	rotan	<i>kətang</i>	<i>hotang</i>	/k~/h/ /ə~/o/
229	tikus	<i>mənci</i>	<i>monsi</i>	/ə~/o/ /c~/s/
306	bisik (ber)	<i>kusik</i>	<i>husip</i>	/k~/h/ /k~/p/
411	dingin	<i>bərgəh</i>	<i>borgo</i>	/ə~/o/ /h~/ə/
423	kuat	<i>gəgəh</i>	<i>gogo</i>	/ə~/o/ /h~/ə/

Gloss-gloss diatas memiliki bentuk yang berimbang walaupun mengalami perubahan fonemis yang timbal balik namun terjadi secara teratur, sehingga gloss-gloss di atas masih dianggap berkerabat.

c) Kemiripan secara fonetis

Pasangan kata ini dianggap berkerabat apabila dalam artikulatoris yang sama mengandung kemiripan secara fonetis.

Tabel 5. Kemiripan secara fonetis

No	Gloss	Bahasa Batak Karo	Bahasa Batak Toba	Kemiripan secara fonetis
2	bahu	<i>bara</i>	<i>abara</i>	/ə~/a/
36	mulut	<i>babah</i>	<i>baba</i>	/h~/ə/
114	pintu	<i>pintun</i>	<i>pintu</i>	/n~/ə/
227	taring	<i>sawit</i>	<i>sait</i>	/w~/ə/
289	sawah	<i>sabah</i>	<i>saba</i>	/h~/ə/
334	lihat	<i>idah</i>	<i>ida</i>	/h~/ə/
382	tambah	<i>tambah</i>	<i>tamba</i>	/h~/ə/
416	hijau	<i>ratah</i>	<i>rata</i>	/h~/ə/
478	delapan	<i>ualuh</i>	<i>ualu</i>	/h~/ə/
490	lima puluh	<i>lima puluh</i>	<i>lima pulu</i>	/h~/ə/

Berdasarkan tabel di atas, gloss bahasa pada bahasa Batak Karo dan bahasa Batak Toba cenderung memiliki kemiripan secara fonetis, perbedaannya ialah pada gloss bahasa Batak Karo yang lebih cenderung menggunakan fonem /h/ pada akhir kata.

d) Satu fonem berbeda

Asal segmennya cukup panjang, pasangan kata yang mengandung satu fonem berbeda dianggap berkerabat.

Tabel 6. Satu fonem berbeda

No	Gloss	Bahasa Batak Karo	Bahasa Batak Toba	Fonem Berbeda
3	bibir	<i>bibər</i>	<i>bibir</i>	/ə~/i/
11	gigi	<i>ipən</i>	<i>ipon</i>	/ə~/o/
49	telinga	<i>pinggə l</i>	<i>pinggo l</i>	/ə~/o/
96	kepala desa	<i>pəngul u</i>	<i>pangul u</i>	/ə~/a/

131	lesung	<i>lasung</i>	<i>losung</i>	/ə/~/o/
163	ubi	<i>gadun</i> <i>g</i>	<i>gadon</i> <i>g</i>	/u/~/o/
195	tebu	<i>tabu</i>	<i>to bu</i>	/ə/~/o/
215	kupu- kupu	<i>kaba-</i> <i>kaba-</i>	<i>raba-</i> <i>raba-</i>	/k/~/r/
331	ingat	<i>ingət</i>	<i>ingot</i>	/ə/~/o/
376	minum	<i>inəm</i>	<i>inum</i>	/ə/~/u/

Gloss bahasa Batak di atas mengandung satu fonem yang berbeda, namun tetap dianggap berkerabat/ memiliki hubungan genetik, karena gloss bahasa Batak tersebut tetap mengandung makna yang sama walaupun memiliki satu buah fonem yang berbeda.

Setelah menetapkan kata kerabat sesuai prosedur leksikostatistik, maka dapat dihitung persentase hubungan genetik berdasarkan kata kerabatnya dengan menggunakan rumus:

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

Keterangan:

C: cognates (kata kerabat)

K: jumlah kosa kata kerabat

N: jumlah kata yang diperbandingkan

Penyelesaian:

Dari 500 kosa kata (gloss) yang dicatat di bahasa Batak Karo dan Batak Toba, terdapat 490 pasangan kata, 320 pasang kata kerabat, dan 10 gloss yang tidak diperhitungkan, maka:

$$C = \frac{K}{N} \times 100\% = \frac{320}{490} \times 100\% = 65,3\%$$

C = 65% (pembulatan)

Berdasarkan perhitungan di atas, persentase hubungan genetik berdasarkan kata yang berkerabat dari bahasa Batak Karo dan bahasa Batak Toba adalah sebesar 65%.

Menghitung Waktu Pisah

Waktu pisah bahasa Batak Karo dan Batak Toba dihitung melalui rumus:

$$W = \frac{\text{Log. C}}{2\text{Log. r}}$$

Keterangan:

W: waktu pisah

r: retensi

C: persentase kerabat

Log: logaritma dari

Penyelesaian:

$$W = \frac{\text{Log. C}}{2\text{Log. r}} = \frac{\text{Log } 65}{2 \times \text{Log } 0,805} = \frac{-0,432}{2 \times -0,217} = \frac{-0,432}{-0,434} = 0,995$$

W= 0,995 dikali 1.000 tahun

W= 995 tahun

Berlandaskan rumus di atas waktu pisah bahasa Batak Karo dan bahasa Batak Toba dijelaskan di bawah ini:

1. Bahasa Batak Karo dan bahasa Batak Toba adalah bahasa tunggal sekitar 995 tahun yang lalu.
2. Dihitung dari tahun 2023, bahasa Batak Karo dan bahasa Batak Toba terpisah dari bahasa proto tahun 1.028 Masehi atau abad ke XI Masehi.

Untuk mengadakan pengelompokan bahasa-bahasa kerabat, Swadesh (dalam Keraf (2018:135) mengusulkan suatu klasifikasi/ tingkatan bahasa menggunakan dasar-dasar leksikostatistik. Berdasarkan hasil dan pembahasan di atas, persentase hubungan genetik berdasarkan kata kerabat bahasa Batak Karo dan Batak Toba, adalah sebesar 65% dan waktu pisahnya adalah pada tahun 1.028 atau pada abad ke XI Masehi, dengan menggunakan klasifikasi bahasa yang diusulkan oleh Swadesh (dalam Keraf (2018:135) dapat disimpulkan bahwa klasifikasi bahasa Batak Karo dan Batak Toba masuk ke tingkatan bahasa bagian keluarga (*family*).

Persentase kata kerabat dan waktu pisah bahasa Batak Karo dan

bahasa Batak Toba dapat dipengaruhi oleh beberapa faktor, yaitu:

1) Faktor kontak bahasa

Kontak bahasa dapat terjadi apabila adanya hubungan yang terjalin dalam suatu kegiatan sosial suatu masyarakat yang menerima kedatangan masyarakat lain yang berbeda bahasa dari masyarakat itu sendiri. Apabila kontak bahasa ini terjalin cukup lama, maka terjadilah kontak bahasa dari bahasa lain yang dapat mempengaruhi elektabilitas bahasa itu sendiri yang nantinya terjadilah pergeseran bahasa. Akibat lain dari terjadinya kontak bahasa ini dalam suatu masyarakat ialah bilingualisme, alih kode, diglosia, campur kode, konvergensi, integrasi, dan interferensi. Faktor kontak bahasa ini akan mengurangi persentase kata kerabat suatu bahasa karena faktor ini akan menyebabkan sulitnya untuk menemukan kosakata bahasa yang sama/mirip.

2) Faktor migrasi penutur

Migrasi penutur dapat terjadi apabila suatu wilayah mengalami bencana alam, berpindah tempat tinggal, sehingga masyarakat dari wilayah tertentu bermigrasi ke wilayah lain, yang mungkin berbeda bahasa dengan masyarakat tersebut. Migrasi penutur ini juga akan menyebabkan terjadinya kontak bahasa yang nantinya akan mengakibatkan terjadinya bilingualisme, alih kode, diglosia, campur kode, konvergensi, integrasi, dan interferensi. Selain itu migrasi penutur dalam jumlah yang besar dapat menyebabkan pergeseran bahasa. Faktor migrasi penutur ini akan mengurangi persentase kata kerabat bahasa karena faktor ini akan menyebabkan sulitnya untuk menemukan kosakata bahasa yang sama/mirip.

3) Faktor sub-rumpun bahasa yang sama Seperti yang sudah dijelaskan sebelumnya bahwa meskipun daerah geografis penggunaan bahasa Batak Karo dan Bahasa Batak Toba

berseberangan, namun kedua bahasa ini berasal dari rumpun bahasa yang sama, yaitu rumpun bahasa Batak. Oleh sebab itu, dari antara kedua bahasa ini banyak ditemukan kosakata bahasa yang sama/mirip. Faktor ini akan menambah persentase kata kerabat karena bahasa yang dibandingkan merupakan bahasa dengan rumpun yang sama.

Berdasarkan analisis leksikostatistik yang telah dilakukan, dapat ditemukan jawaban untuk pertanyaan-pertanyaan penelitian ini. Persentase hubungan genetik antara bahasa Batak Karo dan Batak Toba adalah sebesar 65%. Hal ini diperoleh dari perhitungan 320 pasangan kata kerabat dari total 490 pasangan kata yang diperbandingkan.

Angka ini menunjukkan tingkat kekerabatan yang cukup tinggi antara kedua bahasa, meskipun keduanya telah berkembang secara terpisah selama periode yang cukup lama. Perkiraan waktu pisah antara bahasa Batak Karo dan Batak Toba adalah 995 tahun yang lalu. Jika dihitung dari tahun 2023, kedua bahasa ini diperkirakan terpisah dari bahasa proto pada tahun 1028 Masehi atau abad XI Masehi. Temuan ini memberikan wawasan berharga tentang sejarah perkembangan kedua bahasa tersebut dan dapat membantu dalam memahami perubahan linguistik yang terjadi selama hampir satu milenium. Usia kedua bahasa tersebut, berdasarkan perhitungan glotokronologi, adalah sekitar 995 tahun. Ini menunjukkan bahwa bahasa Batak Karo dan Batak Toba telah mengalami perkembangan yang signifikan sejak berpisah dari bahasa proto, namun tetap mempertahankan banyak kesamaan leksikal.

4. KESIMPULAN

Temuan penelitian ini memiliki implikasi penting terhadap klasifikasi dan pemahaman hubungan genetik bahasa-bahasa Batak. Dengan tingkat kekerabatan 65% antara bahasa Batak

Karo dan Batak Toba, hasil ini mengkonfirmasi bahwa kedua bahasa tersebut masih tergolong dalam satu keluarga bahasa (*family*) menurut klasifikasi Swadesh. Hal ini memperkuat pemahaman kita tentang hubungan genetik di antara bahasa-bahasa Batak. Waktu pisah yang diperkirakan pada abad ke-11 Masehi (tahun 1028) memberikan wawasan berharga tentang sejarah persebaran penutur bahasa Batak, menunjukkan bahwa proses diferensiasi antara bahasa Batak Karo dan Batak Toba telah berlangsung selama hampir satu milenium. Ini mencerminkan dinamika sosial, budaya, dan geografis yang kompleks di wilayah Sumatera Utara selama periode tersebut.

Tingkat kekerabatan 65% juga mengindikasikan adanya perbedaan yang cukup signifikan antara kedua bahasa, meskipun masih dalam satu keluarga. Ini menunjukkan pentingnya upaya pelestarian dan dokumentasi masing-masing bahasa, mengingat keunikan dan kekayaan linguistik yang dimiliki oleh setiap varian bahasa Batak. Penelitian ini memberikan kontribusi penting dalam studi linguistik historis komparatif, khususnya dalam konteks bahasa-bahasa Austronesia di Indonesia. Metode leksikostatistik dan glotokronologi yang digunakan dapat menjadi model untuk studi serupa pada bahasa-bahasa daerah lainnya di Indonesia.

Untuk penelitian lanjutan, disarankan untuk melakukan studi komparatif yang lebih komprehensif dengan melibatkan semua varian bahasa Batak (Simalungun, Angkola Mandailing, Pakpak) guna mendapatkan gambaran yang lebih lengkap tentang hubungan genetik dalam keluarga bahasa Batak. Penggunaan metode linguistik historis komparatif yang lebih mendalam, seperti rekonstruksi fonologis dan leksikal, dapat memperkuat temuan tentang hubungan genetik dan memungkinkan rekonstruksi

proto-bahasa Batak. Analisis fitur gramatikal dan sintaksis juga perlu dilakukan untuk melengkapi studi leksikostatistik, sehingga dapat memberikan pemahaman yang lebih holistik tentang hubungan antar bahasa Batak. Integrasi data linguistik dengan bukti arkeologis dan genetik juga dapat memberikan pemahaman yang lebih komprehensif tentang sejarah migrasi dan persebaran penutur bahasa Batak.

Dengan melakukan penelitian lanjutan seperti yang disarankan, diharapkan pemahaman kita tentang hubungan genetik bahasa-bahasa Batak akan semakin mendalam dan komprehensif. Agar pemahaman terhadap hubungan genetik bahasa Batak semakin dalam, serta mengingat bahwa penelitian ini hanya dilakukan terhadap 500 kosakata sesuai panduan Mahsun (1995), maka peneliti menyarankan adanya penelitian lanjutan dari penelitian ini dengan jumlah kosakata yang lebih banyak lagi, karena peneliti percaya bahwa semakin banyak jumlah kosakata yang diteliti, maka semakin banyak pula kesamaan yang akan ditemukan diantara kedua bahasa kerabat ini, dan semakin dalam pula pemahaman yang dapat diperoleh nantinya.

Penelitian pada bahasa Batak Karo dan Batak Toba ini bertujuan agar masyarakat melihat bahasa daerah sebagai ciri khas dan identitas bangsa Indonesia yang wajib dilestarikan. Upaya ini penting agar bahasa-bahasa daerah di Indonesia tetap ada dan tidak punah, serta mendorong generasi muda untuk tidak malu menggunakan bahasa daerahnya masing-masing. Pemerintah diharapkan melakukan revitalisasi untuk melestarikan bahasa daerah dengan mendorong satuan pendidikan memuat pelajaran bahasa daerah dalam kurikulum. Integrasi bahasa Batak dalam Kurikulum 2013 di sekolah dasar Kabupaten Toba, seperti yang diungkapkan oleh Yolanda dan Sinaga

(2023), memainkan peran penting dalam mendukung pembelajaran dan pengembangan karakter siswa. Kurikulum 2013 menekankan pada penguatan karakter dan kompetensi dasar siswa, di mana bahasa daerah seperti bahasa Batak menjadi media penting dalam menyampaikan nilai-nilai budaya dan moral.

5. DAFTAR PUSTAKA

- Afria, R., Sanjaya, D., dan Tiara, M. (2020). *Leksikostatistik dan Grotokronologi Bahasa Melayu Palembang, Basemah Lahat, Basemah Pagaralam, dan Kayu Agung: Kajian Linguistik Historis Komparatif*. *MADAH*, 11(1), 27–42.
<https://doi.org/10.31503/madah.v11i1.223>
- Chaer, A. (2014). *Linguistik Umum*. Jakarta: PT Rineka Cipta.
- Chaer, A. (2018). *Psikolinguistik*. Jakarta: PT Rineka Cipta.
- Damanik, R. A. P. (2019). *Leksikostatistik Bahasa Batak Simalungun dan Bahasa Batak Mandailing* [Skripsi]. Medan: Universitas Sumatera Utara.
- Dardanila, D. (2018, October). *Leksikostatistik Bahasa Karo dan Bahasa Gayo*. In *Talenta Conference Series: Local Wisdom, Social, and Arts (LWSA)* (Vol. 1, No. 1, pp. 185-191).
- Hutabarat, F. M., Ermanto, dan Juita, N. (2013). *Kekerabatan Bahasa Batak Toba Dengan Bahasa Batak Mandailing* [Skripsi]. Padang: Universitas Negeri Padang.
- Keraf, G. (2018). *Lingusitik Bandingan Historis*. Jakarta: PT Gramedia.
- Mahsun. (1995). *Dialektologi Diakronis*. Yogyakarta: Gadjah Mada University Press.
- Mualita, G. (2015). *Kekerabatan Bahasa Batak Toba dan Bahasa Batak Angkola Suatu Kajian Linguistik Historis Komparatif*. *Arkhai-*
Jurnal Ilmu Bahasa dan Sastra Indonesia, 6(1), 46-52.
- Napitupulu, L. H., dan Silaban, Y. N. (2020). *Kajian Leksikostatik Pada Bahasa Batak Toba Dan Batak Angkola Mandailing*. *Jurnal Bahasa Indonesia Prima (BIP)*, 2(2), 82–90.
<https://doi.org/10.34012/bip.v2i2.1336>
- Sugiyono. (2018). *Metode Penelitian Kuantitatif Kualitatif Dan R&D*. Bandung: Alfabeta.
- Sukapiring, P. (2019). *Pelajaran Bahasa Karo*. Medan: Universitas Sumatera Utara.
- Syafi'i, imam, dan Ibrahim, A. S. (2019). *Leksikostatistik Lima Bahasa Nusantara: Bahasa Jawa, Bahasa Madura, Bahasa Sunda, Bahasa gBali, dan Bahasa Indonesia*. *BASINDO: jurnal kajian bahasa, sastra Indonesia, dan pembelajarannya*, 3(1), 85–93.
<https://doi.org/10.17977/um007v3i12019p085>
- Yolanda, H., & Sinaga, W. (2023). *Signifikansi Bahasa Batak sebagai Faktor Kunci dalam Membentuk Karakter dan Mendukung Pembelajaran dalam Kurikulum 2013 di Sekolah Dasar Kabupaten Toba*. *Kompetensi*, 16(1), 45–58.
<https://doi.org/10.36277/kompetensi.v16i1.96>
- Wartono, W. (2018). *Leksikostatistik Dan Glotokronologi Bahasa Batak: Hubungan Kekerabatan Bahasa Batak Dialek Toba, Simalungun, Mandailing Dan Karo*. *MEDAN MAKNA: Jurnal Ilmu Kebahasaan dan Kesastraan*, 11(1), 61-75.