

KARO DIALECTAL VARIATION IN TANAH KARO: A LEXICAL AND PHONOLOGICAL SOCIOLINGUISTIC ANALYSIS

Sukutenta br Sembiring, Efendi Barus, Ely Ezir

Universitas Islam Sumatera Utara, Medan, Indonesia

E-mail: sukutenta.sembiring@yahoo.com

Received: 2025-11-01

Accepted: 2025-11-16

Published: 2025-11-17

Abstract

Language holds the memory of a community and shapes how people understand themselves, their landscape, and their cultural belonging. This study investigates dialectal variation in the Karo language of Tanah Karo, North Sumatra, through a lexical and phonological sociolinguistic framework. Using a descriptive qualitative design supported by dialectological procedures, the research analyzes 200 Swadesh-based lexical items and natural speech data collected from six native speakers representing three major dialect regions: Julu (Tigapanah), Teruh Deleng (Tiganderket), and Singalor Lau (Tigabinanga). Structured interviews, direct observation, and audio-recorded elicitation sessions formed the basis of the dataset. Findings show that while the dialects remain mutually intelligible, they exhibit systematic micro-variations in vowel realization, lexical selection, and morphophonemic structure. Tigabinanga displays the most consistent vowel fronting and several unique lexical items, whereas Tigapanah and Tiganderket share closer phonological and lexical affinities. These patterns reflect the influence of geographical separation, inter-village communication, and long-standing cultural identity. The study contributes to the documentation of Indonesia's regional languages and underscores the importance of preserving dialectal diversity. Further research using acoustic phonetics or quantitative dialectometry is recommended to deepen the understanding of phonological distance within the Karo speech community.

Keywords: *dialects; Karo language variation, lexical and phonological variation, sociolinguistics, Tanah Karo*

1. Introduction

Language is more than a tool for conveying information; it is the medium through which humans construct meaning, negotiate identity, and sustain cultural memory. In multilingual nations such as Indonesia, language functions simultaneously as a communicative system and a marker of cultural belonging. The inseparable relationship between language and society, as emphasized by Halliday (1970, in Chaer, 1995), underscores how linguistic practices both reflect and shape social realities. Every

linguistic choice—whether lexical, phonological, or stylistic—signals a speaker's position within a network of cultural, geographical, and social relationships.

Within the complex linguistic landscape of North Sumatra, the Karo people have maintained a distinct linguistic heritage that continues to evolve across generations. The Karo language comprises several recognized dialects, including Julu, Teruh Deleng, Singalor Lau, Jahe, and Liang Melas (Robert, 2007; Prinst, 2002). Although these dialects share a common ancestral root, they exhibit notable differences in vocabulary, pronunciation, and local usage patterns. These variations arise from the interaction of geographical separation, social structure, and cultural identity, illustrating Sapir's (1965) assertion that linguistic diversity is a natural and expected characteristic of human societies.

Despite the recognition of Karo dialect diversity in earlier works (Sembiring, 2009; Sembiring, 2010; Pasaribu, 2013; Daulay, 2016), systematic comparative studies focusing specifically on lexical and phonological variation among the major dialects of Tanah Karo remain limited. Most existing studies provide broad classifications or descriptive overviews but have not examined in depth how micro-level lexical shifts and phonological correspondences emerge across neighboring dialect regions. This creates a scholarly gap, particularly in understanding how geography, social interaction, and intergroup contact contribute to the shaping of dialect boundaries within a relatively small and interconnected region.

This study responds to that gap by offering a focused sociolinguistic analysis of lexical and phonological variation among three central Karo dialects: Julu (Tigapanah), Teruh Deleng (Tiganderket), and Singalor Lau (Tigabinanga). The research aims to identify the nature of these variations, explore their patterns, and explain the factors that influence them. By drawing on sociolinguistic perspectives articulated by Fishman (1979), Fasold (1990), and Labov (2001), the study situates Karo dialect differences within a broader theoretical framework of language variation as a systematic and socially meaningful phenomenon.

Beyond its academic contribution, this research carries cultural significance. Local languages in Indonesia face increasing pressure from dominant national and global languages, which can lead to the erosion of dialectal richness and intergenerational linguistic knowledge. Documenting lexical and phonological variations within the Karo language therefore contributes not only to linguistic scholarship but also to cultural preservation. Understanding how dialects differ—and how they remain mutually intelligible—supports efforts to maintain the linguistic identity of the Karo community while recognizing its internal diversity.

In this spirit, the study seeks to deepen our understanding of how language, space, and community interact in Tanah Karo. The findings are expected to enrich existing documentation of Indonesian regional languages and to provide a sociolinguistic basis for future work in dialectology, language education, and cultural revitalization.

2. Literature Review

2.1 Language and Society

Language is inseparable from the social world in which it is used. It reflects relationships, power structures, cultural practices, and collective identities. This foundational principle, articulated in early linguistic scholarship by Sapir (1965) and later expanded by Chambers (2003), provides the cornerstone of sociolinguistic inquiry.

Sociolinguistics, as defined by Fishman (1979), examines how language operates within its social environment, while Fasold (1990) emphasizes that variation is not random but patterned according to social factors such as age, class, gender, and geography.

Halliday (1970) and Ferguson & Gumperz (1960) highlight that all languages contain inherent varieties shaped by communicative purpose and social context. Their work introduces the concepts of “lects” and “registers,” which describe how linguistic choices shift depending on participants, settings, and functions. This perspective positions variation as a natural and essential feature of language, enabling it to adapt to the complexities of social life.

Within this theoretical frame, the study of variation in the Karo language becomes part of a broader conversation about how communities negotiate identity through speech. The Karo-speaking population, embedded in the culturally diverse landscape of North Sumatra, exemplifies how language reflects social differentiation while simultaneously sustaining cultural unity.

2.2 Concept of Language Variation

Language variation is an intrinsic property of all human languages. Sapir’s (1921) early observation that “everyone knows that language is variable” continues to guide modern linguistic thought. Variation manifests across linguistic levels—phonological, morphological, lexical, and syntactic—and reveals the dynamic nature of language as it shifts through time, space, and social context.

Bayle (2007) classifies these levels as fundamental domains of linguistic analysis, each providing a lens for understanding how languages diverge and converge. In sociolinguistics, special attention is given to dialects, which emerge due to geographical separation, and to registers, which arise from differences in formality, profession, or communicative purpose.

The Karo language offers concrete illustrations of this variability. For instance, the lexical differences in naming common objects—such as the regional variation in the word for “coconut”—show how communities adapt their linguistic resources to local identity and interactional needs. Such examples demonstrate that language variation is not simply a matter of vocabulary choice but an expression of cultural distinctiveness shaped by lived experience.

2.3 Types and Causes of Language Variation

Sociolinguists generally agree that variation is structured and socially meaningful. Saragih (2014) identifies several causes of linguistic variation: regional boundaries, social class, age, gender, status, and geographical conditions. Regional dialects form when communities are separated by physical barriers, leading to distinctive phonological and lexical features. Social dialects emerge from differences in economic background, occupation, or educational level. These distinctions influence not only vocabulary and pronunciation but also preferred styles of communication. For example, younger speakers may gravitate toward expressive or innovative forms, while older speakers often retain more traditional ones.

Turell (2002) and Labov (1966, 2001) further argue that variation follows systematic patterns rather than spontaneous or isolated deviations. Linguistic choices are shaped by community norms and by the unconscious desire to express belonging or distinction. Such patterns reveal how language is tightly interwoven with identity and social structure.

In the case of the Karo language, internal factors such as phonological proximity interact with external factors such as geography, migration, and intergroup contact. The mountainous terrain of Tanah Karo, with its natural divisions between valleys and settlements, provides fertile ground for the development of micro-dialects that remain mutually intelligible yet distinct.

2.4 The Karo Language and Its Dialects

The Karo language, belonging to the Austronesian family, is spoken by approximately 600,000 people in the highlands of North Sumatra. While closely related to Batak languages such as Simalungun, Gayo, and Alas, it maintains a unique phonological and lexical profile (Sembiring, 2009). Its dialectal diversity is shaped by centuries of cultural and geographical development, with the highland environment contributing significantly to patterns of linguistic differentiation.

Scholars generally classify Karo into five major dialects: Julu, Teruh Deleng, Singalor Lau, Jahe, and Liang Melas (Prinst, 2002; Robert, 2007). Although these dialects share structural similarities and high mutual intelligibility, they display subtle yet meaningful differences. Previous studies—such as those by Sembiring (2010), Pasaribu (2013), and Daulay (2016)—have catalogued specific features and confirmed that dialect variation in Karo is shaped by both geographical distribution and cultural interaction.

However, most existing works provide broad descriptive accounts and do not offer in-depth comparative analysis of lexical and phonological correspondences across the major dialect regions of Tanah Karo. This leaves a scholarly space for research that examines how vocabulary choices and vowel realizations differ systematically among dialects and what these differences reveal about linguistic evolution within the community.

Taken as a whole, the literature establishes a strong theoretical foundation for analyzing Karo dialectal variation. It also highlights the need for more focused studies that combine dialectological methods with sociolinguistic insights, particularly in regions where geography and social interaction continue to shape linguistic diversity. The present study builds on this foundation by offering a detailed analysis of lexical and phonological variation among three key Karo dialects.

3. Research Methodology

This study employed a descriptive qualitative design, an approach well suited for documenting linguistic variation within its natural social and geographical context. The analysis followed the dialectological procedures proposed by Mahsun (1995), which provide systematic steps for comparing language forms across regions while accounting for phonological and lexical correspondences.

3.1 Research Sites and Participants

The research was conducted in three major dialect regions of Tanah Karo: Tigapanah (representing the Julu dialect), Tiganderket (Teruh Deleng dialect), and Tigabinanga (Singalor Lau dialect). These locations were selected purposefully because they represent distinct dialect clusters and exhibit clear geographical separation.

Six native speakers participated in the study, two from each region. Informants were chosen based on the following criteria:

- (1) native fluency and lifelong residence in the dialect area,

- (2) age above forty to ensure linguistic stability and minimal influence from inter-dialectal mixing,
- (3) physical and mental fitness for providing reliable linguistic data.

These criteria align with standard dialectological practice for capturing stable and representative language forms.

3.2 Data Collection Techniques

Data were collected through structured interviews and direct observation. The structured interviews employed a Swadesh-based wordlist consisting of core vocabulary items commonly used in dialect comparison studies. While the exact number of items varied slightly across speakers due to differences in recall and natural usage, the list covered a broad range of lexical categories—including nouns, verbs, adjectives, and adverbs—sufficient to identify systematic phonological and lexical correspondences.

During the interviews, informants were asked to pronounce each lexical item and provide its dialectal equivalent. All sessions were audio-recorded to ensure accurate phonetic transcription. Observation was used to capture naturally occurring speech, providing additional phonological data and confirming lexical choices identified through elicitation.

3.3 Data Analysis Procedures

The analysis followed the interactive model of Miles, Huberman, and Saldana (2014). In the condensation stage, all lexical and phonological data were transcribed, categorized, and prepared for comparison. Items exhibiting variation were grouped and contrasted across the three dialect regions.

In the display stage, the data were organized into comparative tables, enabling a clear visualization of similarities and divergences. Patterns of vowel shifting, consonantal differences, and lexical replacement were identified through repeated comparison.

During the conclusion-drawing stage, the emerging patterns were interpreted in relation to sociolinguistic and dialectological theory. Particular attention was paid to how geographical separation, historical interaction, and community identity contributed to the observed variation.

3.4 Establishing Trustworthiness

Trustworthiness was ensured through methodological triangulation—using multiple data sources (elicitation and natural speech) and multiple informants from each dialect region. Transferability was strengthened by detailed descriptions of the research sites and participants. Dependability and confirmability were maintained through advisor validation, systematic documentation of analytical steps, and an audit trail of transcription and coding decisions, following the principles of Lincoln and Guba (1985).

4. Findings and Discussion

This section integrates the presentation and interpretation of the data gathered from the three Karo dialect regions: Tigapanah (Julu), Tiganderket (Teruh Deleng), and Tigabinanga (Singalor Lau). The data display precedes interpretation so that each linguistic pattern can be examined in light of sociolinguistic and dialectological theory.

To accommodate international readership, each entry is presented in the following order: dialect form → IPA transcription → English meaning.

4.1 Core Lexical and Phonological Variants

The following table presents fundamental lexical and phonological differences identified across the three dialects. These items were drawn from a Swadesh-based list and natural speech.

No	Tigapanah	Tiganderket	Tigabinanga	English Meaning
1	<i>biang</i> /biyang/	<i>biang</i> /biang/	<i>biang</i> /biang/	Dog
2	<i>bombang</i> /bombang/	<i>bombang</i> /bombang/	<i>bombang</i> /bɔmbang/	Wave / tide
3	<i>tutung</i> /tutung/	<i>tutung</i> /tutung/	<i>tutong</i> /tutɔng/	To burn / roast
4	<i>mbue</i> /mbuai/	<i>melala</i> /mɛlala/	<i>mbue</i> /mbuɛi/	Many / a lot
5	<i>taka</i> /taka/	<i>taka</i> /taka/	<i>taka</i> /takæ/	To split / divide
6	<i>bunga</i> /bunga/	<i>bunga</i> /bunga/	<i>bunga</i> /bungæ/	Flower
7	<i>ndabuh</i> /ndabuh/	<i>naktak</i> /naktak/	<i>nampok</i> /nampɔk/	To fall
8	<i>daging</i> /daging/	<i>jukut</i> /jukut/	<i>jukut</i> /jukut/	Meat
9	<i>uai</i> /uwai/	<i>ue</i> /ue/	<i>uei</i> /uei/	Yes (affirmative)

Table 4.1. Core Lexical and Phonological Variants

The data show that all three dialects remain mutually intelligible, demonstrating a strong lexical core. However, systematic variations emerge that shed light on the internal structure of the Karo language.

1. Vowel Shifts in Tigabinanga

The Tigabinanga dialect consistently performs vowel fronting and lowering, for instance:

- /a/ → /æ/: *taka* → *takæ*
- /u/ → /ɔ/: *tutung* → *tutɔng*

This pattern aligns with Labov's (1994, 2001) principles of chain shifting in regional dialects. Peripheral communities, particularly those geographically distant from linguistic centers, tend to innovate more rapidly.

2. Lexical Replacement

In items such as:

- *mbue*, *melala* ('many')
- *ndabuh*, *naktak*, *nampok* ('to fall')

We see lexical divergence rather than mere phonological variation. This suggests micro-regional innovation influenced by inter-village cultural identity.

3. Morphophonemic Variation

Forms such as *daging* vs *jukut* illustrate the coexistence of inherited and innovative lexemes. This aligns with Mahsun’s (1995) theory that dialects in Indonesia often reflect older historical layers preserved differently across regions.

4.2 Additional Phonological and Lexical Differences

No	Tigapanah (Form + IPA)	Tiganderket (Form + IPA)	Tigabinanga (Form + IPA)	English Meaning
1	<i>ndehər</i> /ndəhər/	<i>ndehər</i> /ndəhər/	<i>ndihər</i> /ndihər/	Near
2	<i>medem</i> /mədəm/	<i>medem</i> /mədəm/	<i>tayang</i> /tajan/	To sleep
3	<i>bisul</i> /bisul/	<i>bisul</i> /bisul/	<i>bareh</i> /bareh/	Boil (skin)
4	<i>mehuli</i> /məhuli/	<i>meradal</i> /məradal/	<i>merandal</i> /mərandal/	Good
5	<i>nipai</i> /nipai/	<i>nipe</i> /nipe/	<i>nipe</i> /nipe/	Snake

Table 4.2. Additional Variants

1. Predictable Phonological Patterns

The difference between *ndehər* and *ndihər* exemplifies Tigabinanga’s tendency to front central vowels, strengthening its profile as the most phonetically innovative dialect.

2. Verb Divergence

The divergence in the term for ‘sleep’—*medem* vs *tayang*—is particularly striking. Verbs are known in sociolinguistics (Trudgill, 1983) to carry cultural and experiential nuance, making them more susceptible to dialectal specialization.

3. Adjectival Innovation

The variation among *mehuli*, *meradal*, and *merandal* reveals a process of gradual phonological erosion. This supports Mahsun’s model that dialect change in Indonesia often proceeds via small shifts accumulating over generations.

4. Retention of Core Nouns

The relative stability of nouns such as *nipai/nipe* suggests that referential vocabulary—especially fauna and flora terms—evolves more slowly.

4.3 Semantic Categorization

A. Nouns

Tigapanah	Tiganderket	Tigabinanga	English Meaning
<i>nipai</i>	<i>nipe</i>	<i>nipe</i>	Snake
<i>bunga</i>	<i>bunga</i>	<i>bunga</i>	Flower

B. Verbs

Tigapanah	Tiganderket	Tigabinanga	English Meaning
<i>medem</i>	<i>medem</i>	<i>tayang</i>	To sleep
<i>ndabuh</i>	<i>naktak</i>	<i>nampok</i>	To fall

C. Adjectives

Tigapanah	Tiganderket	Tigabinanga	English Meaning
<i>mehuli</i>	<i>meradal</i>	<i>merandal</i>	Good
<i>melket</i>	<i>melket</i>	<i>melket</i>	Dirty

Table 4.3. Semantic Categories by Dialect

The semantic categorization reveals a hierarchy of variability:

1. Nouns are the most stable, reflecting their foundational role in reference and communication.
2. Verbs show the greatest divergence, indicating sensitivity to cultural practices, metaphorical usage, and experiential semantics.
3. Adjectives occupy a middle tier, particularly those that express evaluation or moral judgment (e.g., ‘good’).

This distribution mirrors cross-linguistic patterns identified in semantic typology.

4.4 Sociolinguistic Interpretation

The dialectal variation observed in Karo is a product of multiple interacting sociolinguistic factors:

1. Geography and Mobility

Mountainous terrain separates villages in Karo, limiting contact and encouraging phonological drift in peripheral areas such as Tigabinanga.

2. Identity and Ancestry

Speakers interviewed noted that dialect reflects village history, reinforcing Barth’s (1969) concept of linguistic boundaries as cultural boundaries.

3. Communication Networks

Tigapanah and Tiganderket share stronger economic and social ties, explaining their closer linguistic resemblance.

4. Oral Tradition

Karo communities maintain rich oral practices, which help preserve older lexical forms in some regions while accelerating innovation in others.

4.5 Integrated Summary

The findings show that:

1. The three dialects are mutually intelligible yet structurally distinct.
2. Tigabinanga is the most innovative dialect, especially in vowel quality and verb lexicon.
3. Tigapanah and Tiganderket form a dialect continuum, reflecting greater interaction.
4. Variations align with geographic, social, and historical factors, supporting classical sociolinguistic theory.

5. Conclusion

This study examined lexical and phonological variation across three dialect regions of the Karo language—Tigapanah (Julu), Tiganderket (Teruh Deleng), and Tigabinanga (Singalor Lau)—through a sociolinguistic and dialectological lens. The

findings demonstrate that while the three dialects maintain a high degree of mutual intelligibility and share a strong lexical core, each region exhibits distinctive patterns that reflect its unique sociocultural and geographical setting. The Tigabinanga dialect shows the highest level of phonological innovation, notably through consistent vowel fronting and lowering, while Tigapanah and Tiganderket maintain more conservative phonological features and a higher degree of mutual similarity.

Lexical variation appears most prominently in verbs and evaluative adjectives, indicating that semantic domains connected to daily activities and subjective judgments are more susceptible to dialectal differentiation. Nouns, particularly those related to flora, fauna, and essential objects, remain relatively stable across regions, supporting cross-linguistic typologies of semantic resilience. These findings confirm that variation in Karo is systematic rather than random and is shaped by long-standing patterns of settlement, mobility, communication networks, and inter-village identity formation.

The interplay between geography, ancestry, and linguistic practice underscores the strong relationship between language and cultural belonging in the Karo highlands. The results also reinforce established theories of language variation, including Labov's model of regional differentiation, Mahsun's dialectological principles, and Fasold's insights into lexical sociocultural anchoring. As such, the study contributes to the ongoing documentation and preservation of local linguistic heritage, an increasingly urgent task in the face of sociocultural change, mobility, and language shift in Indonesia.

Further research using acoustic phonetics, quantitative dialectometry, or ethnographic interviews with younger and more mobile speakers would enrich the current findings and provide deeper insight into ongoing language change within the Karo speech community. This expanded inquiry would support a more comprehensive understanding of how dialectal variation evolves in multilingual, multicultural, and geographically diverse regions such as North Sumatra.

References

- Afriyanti, D. (2006). *Qualitative research and trustworthiness*. Jakarta: Universitas Terbuka.
- Bayle, E. (2007). *Principles of linguistic variation*. London: Routledge.
- Bogdan, R., & Biklen, S. (1992). *Qualitative research for education: An introduction to theory and methods*. Boston, MA: Allyn & Bacon.
- Chaer, A. (1995). *Sociolinguistik: Perkenalan awal*. Jakarta: Rineka Cipta.
- Chambers, J. K. (2003). *Sociolinguistic theory: Linguistic variation and its social significance*. Oxford: Blackwell.
- Daulay, R. (2016). *Variasi dialek bahasa Mandailing di Kabupaten Mandailing Natal*. Medan: Universitas Sumatera Utara.
- Fasold, R. (1990). *The sociolinguistics of language*. Oxford: Blackwell.
- Fishman, J. A. (1979). *Sociology of language*. Rowley, MA: Newbury House.
- Halliday, M. A. K. (1970). Language structure and language function. In J. Lyons (Ed.), *New horizons in linguistics* (pp. 140–165). London: Penguin.
- Labov, W. (1966). *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, W. (2001). *Principles of linguistic change*. Oxford: Blackwell.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.

- Mahsun, M. (1995). *Dialektologi diakronis*. Yogyakarta: Gadjah Mada University Press.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Thousand Oaks, CA: Sage.
- Pasaribu, H. (2013). *A brief description of the dialect of Toba: Variations in the region of Tobasa, Samosir, and Tapanuli Utara*. Medan: USU Press.
- Prinst, D. (2002). *Bahasa Karo dan dialeknya*. Medan: Universitas Sumatera Utara Press.
- Robert, P. (2008). *Language and society*. London: Routledge.
- Robert, S. (2007). *Karo dialects and their phonological characteristics*. Medan: Balai Bahasa Sumatera Utara.
- Saragih, A. (2014). *Variasi dan fungsi bahasa*. Medan: Unimed Press.
- Sapir, E. (1965). *Language: An introduction to the study of speech*. New York: Harcourt, Brace & World.
- Sembiring, B. (2009). *Loyalitas bahasa Karo*. Medan: Universitas Negeri Medan Press.
- Sembiring, R. (2010). *Variasi dialek bahasa Karo di Kabupaten Karo, Deli Serdang, dan Langkat*. Medan: USU Press.
- Turell, M. T. (2002). *Multilingualism and variation*. Amsterdam: John Benjamins.