JOLLS: Journal of Language and Literature Studies
Website: https://journal-center.litpam.com/index.php/jolls/index
Email: jolls@litpam.com
DOI: https://doi.org/10.36312/jolls.v6i3.5106

September 2026 Vol. 6, No. 3
e-ISSN: 2808-1099
pp. 1039-1052

Introduction to Linguistics from a Phonological Perspective: The Final Vowel /a/ → [o] and Word-Initial Rhotic /r/ → [ɣ] Correspondences in the Batahan Dialect

[1] *Resdillapratiwi, [1]Lia Agustina Damanik

[1]English Study Program, STAIN Mandailing Natal, Jl. Prof. Dr. Andi Hakim Nasution, Komplek STAIN Madina, Panyabungan, Mandailing Natal, North Sumatra, Indonesia

*Corresponding author e-mail: resdillapratwi@stain-madina.ac.id

Received: May 2026; Revised: July 2026; Accepted: August 2026; Published: September 2026

Abstract

This study provides a preliminary phonological description of the Batahan dialect, a coastal Malay (Pesisir) variety spoken in Batahan District, Mandailing Natal Regency, North Sumatra. It focuses on two segmental correspondences between Standard Indonesian and Batahan: the word-final vowel correspondence, in which Standard Indonesian /a/ corresponds to Batahan [o] in open final syllables (e.g., *apa* /apa/ ~ [apo] 'what'; *mana* /mana/ ~ [mano] 'where'), and the word-initial rhotic correspondence, in which /r/ corresponds to a voiced dorsal fricative transcribed here as [ɣ] (e.g., *rumah* /rumah/ ~ [ɣumah] 'house'; *ramai* /ramai/ ~ [ɣamai] 'crowded'). Data were collected from ten native Batahan speakers in and around Pasar Batahan throuɣ word-list elicitation, carrier sentences, semi-structured interviews, and recorded conversation, and were analysed using auditory IPA transcription and a synchronic, correspondence-based comparative procedure. The findings indicate that the /a/ ~ [o] correspondence is regular in word-final open syllables, while the /r/ ~ [ɣ] correspondence is concentrated in word-initial onsets; both pattern Batahan with the west-Sumatran coastal Malayic area, alongside Minangkabau and other Sumatran Malay varieties. Because the study rests on a small purposive sample, targeted elicitation, and auditory (non-acoustic) transcription, the results are best read as documentation of recurring correspondences rather than as claims about historical derivation or articulatory ease. The study contributes an initial lexical record of Batahan sound correspondences and a basis for broader dialectological and sociophonetic investigation of under-documented varieties in North Sumatra.

Keywords: Dialect; Coastal Malay; Phonological correspondence; Language vowels; Rhotic variation; Sumatran dialectology

How to Cite: Resdillapratiwi, R., & Damanik, L.A. (2026). Introduction to Linguistics from a Phonological Perspective: The Final Vowel /a/ → [o] and Word-Initial Rhotic /r/ → [ɣ] Correspondences in the Batahan Dialect. *Journal of Language and Literature Studies, 6*(3), 1039-1052. doi: https://doi.org/10.36312/jolls.v6i3.5106

https://doi.org/10.36312/jolls.v6i3.5106

INTRODUCTION

Language is widely recognized as the primary medium of human communication through which individuals express ideas, emotions, intentions, and social identities As a symbolic and structured system, language enables members of a community to interact, transfer knowledge, preserve traditions, and maintain social cohesion (Wewalaarachchi & Sigh, 2020). In linguistic scholarship, language is not viewed merely as a collection of words, but as a rule-governed system consisting of interconnected levels such as sounds, word formation, sentence structure, and meaning. Because language is dynamic and continuously shaped by social interaction, geographical mobility, and historical development, variation is considered a natural characteristic of all speech communities. No language exists in a completely uniform form; instead, every language display

differences in pronunciation, vocabulary, grammar, and usage according to region, social group, and communicative context (van der Elst-Koeiman et al., 2022).

The scientific study of language is known as linguistics. Linguistics systematically investigates how language is structured, acquired, used, and changed over time. Among its major branches, phonology plays a crucial role because it focuses on the sound system of language and the ways sounds function to distinguish meaning. Phonology examines phonemes, sound patterns, syllable structures, stress, intonation, and sound changes that occur within a language or dialect. Unlike phonetics, which concerns the physical production and acoustic properties of speech sounds, phonology emphasizes the abstract organization of sounds in the mental grammar of speakers. Through phonological analysis, researchers are able to identify regular sound correspondences, recurring pronunciation patterns, and rule-governed variation that characterize a particular speech variety. Therefore, phonology is especially useful in investigating dialectal pronunciation differences that may appear irregular on the surface but are systematic within a local linguistic system (J.W.Y., 2020).

Dialect is commonly defined as a regional or social variety of a language distinguished by systematic differences in pronunciation, vocabulary, and grammar. In modern linguistics, dialect does not imply inferiority or deviation from a standard language; rather, every dialect is a legitimate linguistic system with its own internal rules (Tairova, 2024). Regional dialects are often shaped by geographical separation, local history, contact with neiɣboring communities, and patterns of migration (Sakirman, 2017). Pronunciation features are among the most salient markers of dialect identity because they are immediately audible and socially recognizable (Kulubekova, 2025). Consequently, phonological study of dialects contributes significantly to understanding language variation, language change, and identity construction. When speakers consistently pronounce certain vowels or consonants differently from speakers in neiɣboring areas, such patterns become important evidence of local linguistic innovation (Piepers et al., 2023).

Indonesia is one of the most linguistically diverse countries in the world, with hundreds of local languages and numerous dialect varieties spread across its archipelagic territory. This diversity reflects centuries of migration, trade networks, political interaction, and cultural exchange (Nugroho & Setyaningsih, 2019). However, despite such richness, many local dialects remain insufficiently documented in academic literature (Hutapea et al., 2023). Research attention often concentrates on national languages or major regional languages, while smaller community varieties receive limited scholarly examination. At the same time, processes such as urbanization, formal education, media exposure, and language shift increasingly affect local speech practices (Ambalegin & Arianto, 2018). For this reason, documentation and analysis of local dialects are urgently needed, both for theoretical linguistics and for cultural preservation (Moskovitz et al., 1989). Each undocumented dialect potentially contains valuable evidence about sound change, contact phenomena, and local identity formation (Le et al., 2024).

One of the regional speech varieties that deserves closer attention is the Batahan Dialect spoken in the Batahan area of Mandailing Natal Regency, North Sumatra. This region occupies a linguistically strategic area influenced by interactions among Malay varieties, Batak languages, Minangkabau-related speech communities, and Indonesian. Such multilingual environments often produce distinctive dialect features shaped by inheritance and contact (Hadi & Mahdi, 2023). Preliminary observation indicates that Batahan speakers exhibit characteristic pronunciation patterns that differ from surrounding varieties. Among the most noticeable features are the shift of the vowel /a/

to /o/ and the change of consonant /r/ to /ɣ/ in everyday speech (van Pinxteren, 2023). For instance, the form *kemana* may be realized as *kamano*, *hendak* as *hondak*, and *rautan* as *ɣautan*. These examples suggest that the Batahan Dialect possesses systematic phonological traits rather than random pronunciation deviations (Saada, 2023).

Recent studies have shown growing interest in phonological variation within Indonesian regional languages and dialects. Several researchers have examined vowel alternation patterns in Malayic dialects, demonstrating that vowel quality often changes according to syllable position, stress, or regional innovation (van der Elst-Koeiman et al., 2022). Other studies on Sumatran speech communities have documented consonant variation, including rhotic weakening, replacement, and alternative articulations influenced by contact and local identity (van der Elst-Koeiman et al., 2022). In sociolinguistic research, pronunciation features have also been shown to function as markers of solidarity, locality, and group membership. Furthermore, studies in dialectology consistently reveal that local sound patterns can provide evidence for historical relationships among neiɣboring varieties. These findings collectively confirm that phonological variation is a productive field of inquiry in Indonesian linguistics.

Althouɣ previous studies have contributed substantially to the understanding of regional phonology, most of them focus on larger languages or more widely recognized dialects (Omar & Alhassan, 2020). Research on smaller local varieties in peripheral or transitional regions remains limited. In the context of North Sumatra, available studies often prioritize Batak languages, mainstream Malay varieties, or broader regional comparisons, while the Batahan Dialect has received little specific attention. Existing literature rarely provides a focused description of its vowel system, consonant correspondences, or dialect-specific sound changes. More importantly, no study has comprehensively examined the co-occurrence of the final vowel shift /a/ to /o/ and the consonant change /r/ to /ɣ/ as core phonological characteristics of the Batahan Dialect. As a result, important empirical data concerning this local variety remain underrepresented in Indonesian linguistic scholarship (Sperlich et al., 2021).

The present study is positioned within that research gap. Its scientific novelty lies in several aspects. First, this study provides a focused phonological description of the Batahan Dialect, a local variety that has rarely been discussed in formal academic publications (Motlaɣ et al., 2023). Second, unlike earlier works that typically examine either vowel change or consonant variation separately, this research analyzes two prominent sound processes simultaneously, namely the shift /a/ → /o/ and /r/ → /ɣ/, in order to identify their distributional patterns within one dialect system. Third, this study interprets these sound changes not only as structural phonological phenomena but also as socially meaningful markers that distinguish Batahan speakers from neiɣboring communities. Fourth, by documenting naturally occurring speech data from an under-researched region, this article contributes new empirical evidence to Indonesian dialectology and phonological studies (Liu et al., 2023).

From a theoretical perspective, the vowel shift /a/ to /o/ may indicate a patterned process of vowel modification associated with syllable-final environments or local phonotactic preferences. Similar changes have been discussed in phonological theory as vowel backing, rounding, or quality adjustment triggered by prosodic structure. Meanwhile, the change of /r/ to /ɣ/ may represent a rhotic transformation involving articulatory relocation from an alveolar trill or tap toward a fricative or uvular realization. Cross-linguistically, rhotic sounds are well known for their instability and wide range of dialectal realizations (van der Elst-Koeiman et al., 2022) .Therefore, the Batahan data provide a relevant case for examining how local communities reorganize sound systems while maintaining communicative efficiency and social distinctiveness. If

these patterns occur consistently across speakers and lexical items, they can be interpreted as regular phonological rules rather than incidental pronunciation variation (Syafiq & Hafiz, 2023).

Based on the background and research gap described above, the problems addressed in this study are formulated as follows: (1) how does the vowel /a/ shift to /o/ in the Batahan Dialect; (2) how does the consonant /r/ change to /ɣ/ in the Batahan Dialect; (3) in what phonological environments do these changes occur; and (4) what linguistic significance do these sound patterns hold within the Batahan speech community? These questions guide the analysis toward both structural explanation and sociolinguistic interpretation. Accordingly, the objective of this article is to analyze the phonological characteristics of the Batahan Dialect by describing the final vowel shift /a/ to /o/ and the consonant change /r/ to /ɣ/, identifying their distribution patterns, and explaining their significance as distinctive features of local linguistic identity. Throuɣ this investigation, the study seeks to enrich the documentation of Indonesian regional dialects and contribute to broader discussions in phonology, dialectology, and language variation studies.

METHOD

Research Design

This study adopts a descriptive-qualitative design with a synchronic, dialectological orientation. The approach is documentary rather than experimental: its purpose is to record and describe sound patterns as they occur in the everyday speech of native speakers, and then to interpret their distribution and significance (Chambers & Trudgill, 1998; Mahsun, 2005). Analysis is grounded in articulatory-descriptive phonetics with International Phonetic Alphabet (IPA) transcription, and in a correspondence-based comparative procedure in which Standard Indonesian forms serve as a reference point for identifying Batahan reflexes. Standard Indonesian is used here strictly for synchronic comparison and is not assumed to be the historical source of the Batahan forms.

Research Site and Participants

Data were collected in Batahan District, Mandailing Natal Regency, North Sumatra, where the dialect is in daily use. Recording was carried out primarily in Pasar Batahan and adjacent settlements within the district; because this concentrates the sample in one part of the dialect area, the geographical representativeness of the data is treated as a limitation rather than a settled matter. Ten native speakers were selected throuɣ purposive sampling, using three criteria: having been born and raised in the Batahan area, using the dialect actively in daily interaction, and consenting to be interviewed and recorded. To allow a first look at inter-speaker consistency, the sample included women and men across several age groups. The full demographic profile of the participants is summarised in Table 1.

Table 1. The Demographic Profile of the Participants

Code	Age	Sex	Village / Settlement	Occupation	Education	Batahan Use
P1	62	Male	Batahan I	Farmer	Elementary School	Very frequent (daily)
P2	56	Female	Batahan II	Housewife	Junior Hiɣ School	Very frequent (daily)
P3	48	Male	Batahan III	Fisherman	Senior Hiɣ School	Frequent
P4	43	Female	Kampung Kapas	Teacher	Bachelor's Degree	Frequent

Code	Age	Sex	Village / Settlement	Occupation	Education	Batahan Use
P5	39	Male	Pasar Baru Batahan	Trader	Senior Hiɣ School	Moderate
P6	35	Female	Batahan IV	Civil Servant	Bachelor's Degree	Moderate
P7	31	Male	Batahan V	Entrepreneur	Diploma	Moderate
P8	27	Female	Kubangan Tompek	University Student	Senior Hiɣ School	Occasional
P9	24	Male	Batahan I	Private Employee	Bachelor's Degree	Occasional
P10	21	Female	Batahan II	University Student	Senior Hiɣ School	Occasional

Instruments

Four instruments supported data collection. First, a word list consisting of 80 lexical items was compiled to include words containing the target environments, namely word-final /a/ and word-initial /r/, ensuring systematic representation of the variables under investigation. Second, a set of carrier sentences was used to elicit the target words in connected speech. The carrier sentences included *"Say it again, please," "I heard the word yesterday,"* and *"Please pronounce clearly."* Third, a semi-structured interview guide covered local vocabulary, everyday routines, traditional activities, and speakers' attitudes toward the Batahan dialect in order to encourage spontaneous speech. Fourth, field notes documented contextual information, including the recording setting, participant behaviour, and any relevant sociolinguistic observations.

All speech data were recorded using a Zoom H1n Handy Recorder with its built-in stereo microphones. Recordings were saved in uncompressed WAV format (48 kHz, 24-bit) to preserve phonetic detail. During recording, the microphone was positioned approximately 20–30 cm from the participant's mouth in a quiet indoor environment to minimise background noise and maintain consistent recording quality. The recordings were transcribed using the International Phonetic Alphabet (IPA, 2015 revision). Each token was subsequently entered into a coding sheet that included the participant code, lexical item, phonological environment (word-final /a/ or word-initial /r/), elicitation type (word list, carrier sentence, or interview), phonetic realisation, and additional remarks. The coding sheet was designed to facilitate systematic comparison across speakers and linguistic contexts.

Data Collection

Data collection combined four techniques whose outputs were kept distinct in the corpus. Controlled elicitation consisted of 80 lexical items and 10 carrier sentences for each participant, yielding a total of 800 elicited word tokens and 100 carrier-sentence productions across the ten participants. These tasks produced comparable tokens of each target form under controlled conditions. Semi-structured interviews lasted approximately 20–25 minutes per participant, resulting in a total of about 220 minutes of recorded interview data. The interviews elicited extended speech on local vocabulary, everyday routines, and speakers' attitudes toward the Batahan dialect. In addition, natural conversations were recorded in homes, local markets, and community gathering places for approximately 15–20 minutes per recording session, producing a total of about 170 minutes of spontaneous conversational data. These recordings provided naturally occurring speech for cross-checking the forms obtained throuɣ elicitation. The quantity and duration of each source are summarised in the corpus synopsis at the beginning of the Results section, allowing elicited and spontaneous evidence to be analysed separately.

Data Analysis

Analysis proceeded from transcription to interpretation. Recordings were transcribed in the International Phonetic Alphabet (IPA), and a total of 900 elicited tokens (800 word-list tokens and 100 carrier-sentence tokens) were extracted for quantitative analysis. Tokens obtained from the interview and spontaneous conversation were analysed separately as supporting qualitative evidence. Every token containing a target segment was classified according to its phonological environment (word-initial, medial, or final; open or closed syllable) and coded for its phonetic realisation. For the word-final vowel /a/, variants were classified as [a], [ə], [ɔ], or other realisations, whereas the word-initial rhotic /r/ was classified as [r], [ɾ], [ɣ]/[ʁ], or other variants. The frequency of each variant was tabulated by participant, lexical item, and phonological environment to determine the regularity of each correspondence rather than relying on isolated examples.

Segmental identifications followed standard auditory phonetic criteria. To minimise transcriber bias, 20% of the analysed tokens (180 tokens) were independently re-transcribed by a second linguist with postgraduate training in phonetics and phonology who was not involved in the initial transcription. Agreement between the two transcribers was examined by comparing the IPA transcription assigned to each token. Any disagreements were reviewed jointly while listening to the original recordings until a consensus transcription was reached, and the agreed form was entered into the final dataset. Consistency of a realisation across speakers and lexical items was interpreted as evidence of a stable sound correspondence, whereas inconsistent patterns were treated as variable forms.

Ethical Considerations

Participation was voluntary. Before each session, speakers were informed of the purpose of the study, and oral consent to participate and to be recorded was obtained and documented prior to data collection. Participants were also informed that their involvement was entirely voluntary and that they could withdraw from the study at any stage without consequence. Identities were kept confidential by replacing personal names with participant codes (P1–P10), and the data were used solely for academic and research purposes. The study did not undergo formal institutional ethical review; therefore, the ethical conduct of the research was based on the informed oral consent of all participants, together with the principles of voluntary participation, confidentiality, and anonymity.

RESULTS AND DISCUSSION

Results

The analysed corpus comprises 80 target lexical items elicited from ten participants, yielding a total of 900 elicited tokens (800 word-list tokens and 100 carrier-sentence tokens), supplemented by approximately 390 minutes of interview and spontaneous conversational recordings. Of the elicited tokens, 450 contained the target environment of word-final /a/, whereas 450 contained the target environment of word-initial /r/. Among the eligible word-final /a/ tokens, 426 tokens (94.7%) were realised as [o], while the remaining 24 tokens (5.3%) retained [a] or showed other variant realisations. For the word-initial /r/ environment, 418 of the 450 eligible tokens (92.9%) were realised as the dorsal rhotic [ɣ], whereas the remaining 32 tokens (7.1%) were produced as alveolar [r], tap [ɾ], or other variants. The frequencies of these correspondences, together with their

distribution across speakers and lexical items, are presented below and, for inter-speaker patterns.

The word-final vowel correspondence /a/ ~ [o]

The most regular pattern in the corpus is the correspondence between Standard Indonesian word-final /a/ and Batahan [o] in open final syllables. Representative forms are given in Table 2.

Table 2. The word-final vowel correspondence /a/ ~ [o] (Standard Indonesian ~ Batahan)

No	Standard Indonesian	Batahan	IPA (SI ~ Batahan)	Gloss	Environment
1	*apa*	*apo*	/apa/ ~ [apo]	what	word-final open syllable
2	*mana*	*mano*	/mana/ ~ [mano]	where	word-final open syllable
3	*siapa*	*siapo*	/siapa/ ~ [siapo]	who	word-final open syllable
4	*bawa*	*bawo*	/bawa/ ~ [bawo]	bring; carry	word-final open syllable
5	*di mana*	*di mano*	/dimana/ ~ [dimano]	where (at)	word-final open syllable

In every item in Table 2, the vowel that alternates is the one in the final open syllable: Standard Indonesian /a/ corresponds to Batahan [o] while the remainder of the word is unchanged (e.g., /apa/ ~ [apo], /mana/ ~ [mano], /bawa/ ~ [bawo]). The correspondence is therefore positionally defined—it targets the word-final open syllable specifically—and it recurs across a range of common lexemes rather than being confined to a few items, which indicates a productive pattern rather than isolated lexical substitution.

Two forms that are sometimes cited in connection with this pattern are excluded from Table 2 for analytical clarity. In *kemana* ~ [kamano] 'where to', the final /a/ ~ [o] correspondence is accompanied by a separate change in the pretonic syllable (/kə/ ~ [ka]); the form therefore involves more than one correspondence and cannot serve as a clean example of the final-vowel pattern alone. In *hendak* ~ [hondak] 'want', the vowel that differs is in the first syllable (/ɛ/ ~ [o]), and the final syllable is closed and unaffected; this form reflects a different correspondence altogether and is not an instance of word-final /a/ ~ [o]. Keeping these cases separate ensures that the final-vowel correspondence is defined precisely. Phonetically, the Batahan reflex involves backing and rounding of the low vowel: /a/, a low unrounded vowel, corresponds to [o], a mid back rounded vowel, in the relevant environment.

The word-initial rhotic correspondence /r/ ~ [ɣ]

The second pattern is the correspondence between Standard Indonesian word-initial /r/ and a Batahan voiced dorsal fricative, transcribed here as [ɣ]. Representative forms are given in Table 3.

Table 3. The word-initial rhotic correspondence /r/ ~ [ɣ] (Standard Indonesian ~ Batahan)

No	Standard Indonesian	Batahan	IPA (SI ~ Batahan)	Gloss	Environment
1	*rautan*	*ɣautan*	/rautan/ ~ [ɣautan]	(pencil) sharpener	word-initial onset
2	*rumpun*	*ɣumpun*	/rumpun/ ~ [ɣumpun]	cluster; clump	word-initial onset

3	*racikan*	*ɣacikan*	/ratʃĭkan/ ~ [ɣatʃikan]	concoction; mixture	word-initial onset
4	*rumah*	*ɣumah*	/rumah/ ~ [ɣumah]	house	word-initial onset
5	*ramai*	*ɣamai*	/ramai/ ~ [ɣamai]	crowded; lively	word-initial onset

In addition to the systematic raising of the word-final low vowel /a/ to the mid-back rounded [o], the Batahan dialect presents another equally robust, categorically applied, and phonologically significant process, namely the spirantization and dorsalization of the alveolar rhotic /r/ into a voiced velar fricative [ɣ] when it occurs in the absolute word-initial prevocalic onset position, as compellingly evidenced by the consistent correspondences *rautan* → [ɣautan], *rumpun* → [ɣumpun], *racikan* → [ɣatʃĭkan], *rumah* → [ɣumah], and *ramai* → [ɣamai], thereby indicating that the dialect's phonological grammar operates on two distinct yet equally deterministic rules that target the polar extremes of the prosodic word—the left edge for the rhotic and the right edge for the vowel.

In formal generative terms, this word-initial rhotic shift can be explicitly formulated as a rule whereby the underlying coronal trill /r/ undergoes a radical place and manner reassignment, changing its primary articulator from the alveolar ridge to the velar region while simultaneously converting its intermittent airstream mechanism (characteristic of a tap or trill, typically specified as [-continuant] or with a trill feature) into a continuous turbulent frication (specified as [+continuant]) in the environment of a word boundary immediately preceding any vocalic segment, and this transformation entails a multi-specification feature alteration: the input consonant, which is Coronal, [+sonorant], and [-continuant] (in the case of a trill), surfaces as Dorsal, [-sonorant] (given that voiced fricatives are obstruents), and [+continuant], while its laryngeal feature of voicing remains faithfully preserved as [+voice], thus yielding the voiced velar fricative [ɣ] that is perceptually salient and articulatorily stable. When subjected to the constraint-based framework of Optimality Theory, this dialectal divergence from Standard Indonesian can be elegantly modeled through a demotion of faithfulness constraints in favor of a highly ranked markedness constraint that specifically prohibits a coronal rhotic from occupying the onset of a prosodic word; specifically, in the standard language, the constraint IDENT(place) and IDENT(continuant), which mandate the preservation of the input's coronal place and non-continuant manner, are ranked inviolably above the structural constraint *r/ONS_INIT, which marks as ill-formed any alveolar trill appearing at the commencement of a phonological word, thereby permitting the faithful realization [rumah] despite the inherent articulatory complexity of initiating an utterance with a lingual apex trill.

Conversely, in the Batahan dialect, the constraint hierarchy is completely reranked such that *r/ONS_INIT outranks both IDENT(place) and IDENT(continuant), compelling the grammar to prioritize the avoidance of word-initial rhotic markedness at the expense of featural faithfulness, thus forcing the optimal output [ɣumah] from the underlying /rumah/, as the dorsal fricative represents a less marked onset consonant that does not require the intricate ballistic lingual movements and precise aerodynamic conditions necessary for sustaining an alveolar trill at the very start of an articulatory prosodic domain. The articulatory and acoustic motivation for this specific change is deeply grounded in the universal phonetic instability of the rhotic class, particularly the alveolar trill, which demands a rapid, cyclic vibration of the tongue tip against the alveolar ridge sustained by a delicate balance of intraoral pressure and muscular tension,

a configuration that is notoriously susceptible to reduction, backing, or lenition across numerous language families worldwide.

To determine the distribution of the rhotic correspondence, all lexical items containing Standard Indonesian /r/ were examined across word positions. In the present corpus, the correspondence /r/ → [ɣ] is restricted to word-initial onset position. No instances of [ɣ] were found in word-medial or word-final position. Instead, medial and final rhotics were realised as an alveolar rhotic [r], indicating that the dorsal fricative reflex is positionally conditioned rather than a general replacement of /r/ throuɣout the lexicon.

Representative medial forms include *beras* /bəras/ → [bəras], *kering* /kəriŋ/ → [kəriŋ], and *jarum* /dʒarum/ → [dʒarum] (examples from the corpus). Likewise, word-final rhotics such as *air* /air/ → [air] retain an alveolar rhotic rather than showing dorsal frication. The distribution therefore suggests that the correspondence is limited to the initial onset.

Discussion

The two correspondences documented above can be interpreted at several distinct levels, which it is important to keep separate: their phonetic realisation, their phonological distribution, their status as dialect correspondences relative to other varieties, and—only tentatively—their historical dimension. The findings speak most directly to the first three; claims about the fourth require comparative and instrumental evidence beyond the present data.

The final-vowel correspondence in areal perspective

The regularity of /a/ ~ [o] in word-final open syllables aligns Batahan with a recognisable west-Sumatran Malayic type. Whereas the two best-known Malay vowel types realise word-final ⟨a⟩ as either [a] ('a-varieties') or [ə] ('schwa-varieties') (Asmah Haji Omar, 1977; Adelaar, 2005), a further type realises it as [o], most prominently in Minangkabau (Almos, 2012) and in several coastal and Jambi Malay varieties (Krulikowska et al., 2020). Batahan patterns with this [o] type. The fact that the correspondence is confined to the final open syllable, and does not apply to /a/ in other positions, is consistent with the cross-linguistic tendency for word-final syllables to be favoured sites of vowel modification, and it invites future analysis in terms of syllable structure and prosodic position rather than lexical idiosyncrasy. Comparison with neiɣbouring Minangkabau would help establish whether the Batahan pattern is inherited, contact-induced, or both.

The rhotic correspondence in typological perspective

Describing the Batahan onset rhotic in IPA rather than as ⟨ɣ⟩ makes its typological place clear. The Malayic rhotic is among the most variable segments in the family, ranging from an alveolar tap or trill to velar and uvular fricatives across Sumatran and peninsular varieties (McDowell & Anderbeck, 2020); dorsal-fricative realisations of /r/ are documented, for example, in conservative southern-Sumatran Malay ([ɣ] or [x], with [ʁ] in some rural varieties). The Batahan reflex thus falls squarely within the attested range of Malayic dorsal rhotics, and is better understood as a regional realisation of a hiɣly variable segment than as an anomaly. Cross-linguistically, dorsal rhotics are common—French and much of German use a uvular fricative or approximant—which further supports treating [ɣ]/[ʁ] as a natural rhotic outcome. Determining the exact place of articulation, and whether it is uniform across speakers and positions, requires acoustic measurement.

On articulatory-ease explanations

Earlier informal accounts of both patterns appealed to ease of articulation—that [o] is a 'smoother' or 'more stable' final vowel than [a], or that a back fricative is 'less effortful' than an alveolar trill. Such claims are not supported by the present data and should be treated with caution: articulatory economy is one hypothesis among several in the phonological literature, and testing it would require articulatory or acoustic measurement that this study did not undertake. The regularity and environmental conditioning of the correspondences are established here; their articulatory motivation is left as a question for instrumental research.

Contact, history, and identity

Because Batahan sits in a contact zone with Minangkabau and other coastal Malay varieties, both correspondences may reflect a combination of inherited Malayic features and areal convergence, but disentangling inheritance from contact requires systematic comparison with neiɣbouring varieties collected under the same conditions (Crowley & Bowern, 2010). Similarly, the observation that salient pronunciation features often index local belonging is well founded in sociolinguistics (Labov, 1972), and Batahan forms such as [apo] and [ɣumah] are plausible identity markers; substantiating that function, however, calls for attitudinal and interactional data, which the interviews can supply in a dedicated follow-up analysis. These interpretations are therefore offered as directions consistent with the data rather than as established conclusions.

Situated against recent Indonesian dialect phonology—contrastive descriptions of Javanese dialects (Mukaromah et al., 2025), variationist work on Sumatran varieties such as Acehnese Daya (Alamsyah et al., 2024), and articulatory accounts of vowel and consonant place adaptation elsewhere (Povilonis, 2024)—the Batahan findings reinforce a common theme: regional varieties maintain organised, environmentally conditioned sound systems, and their documentation refines the comparative picture. Batahan's specific contribution is the pairing of an [o]-type final-vowel correspondence with a dorsal-rhotic correspondence in a previously undescribed coastal variety.

Several limitations bound these claims. The sample is small and purposive, and is concentrated in Pasar Batahan rather than spread across the dialect area. The reliance on targeted elicitation may over-represent the target forms relative to fully spontaneous speech. Transcription is auditory rather than acoustic, so fine phonetic detail, including the velar-versus-uvular question, remains unresolved. And the comparison with neiɣbouring varieties draws on the published record rather than on first-hand data collected for this study. The findings should therefore be read as preliminary documentation.

CONCLUSION

This study described two recurrent segmental correspondences in the Batahan dialect and situated them within a broader comparative Malayic perspective. In the analysed corpus of 900 elicited tokens, comprising 450 tokens in each target environment, Standard Indonesian word-final /a/ corresponded to Batahan [o] in 426 of 450 tokens (94.7%), while Standard Indonesian word-initial /r/ corresponded to a voiced dorsal fricative, transcribed as [ɣ], in 418 of 450 tokens (92.9%). These findings demonstrate that the vowel correspondence is hiɣly regular in word-final open syllables, whereas the rhotic correspondence is concentrated in word-initial onset position. The hiɣ frequencies of these correspondences indicate that they are systematic phonological features of the analysed corpus rather than isolated lexical variants. Both correspondences place Batahan within the west-Sumatran coastal Malayic area, alongside Minangkabau and

other Sumatran Malay varieties, while their co-occurrence in the Batahan dialect provides an additional comparative data point for Malayic dialectology.

At the same time, the features appear to be socially salient and plausibly contribute to the local linguistic identity of Batahan speakers, althouɣ their indexical and attitudinal significance remains to be established throuɣ dedicated sociophonetic research. These conclusions are necessarily constrained by the study's limited purposive sample and by the reliance on auditory rather than instrumental acoustic analysis, and no claims are made regarding the historical origin or direction of these sound correspondences. Accordingly, the present study should be regarded as preliminary documentation of two recurrent phonological correspondences in the Batahan dialect. It provides an empirical foundation for future research employing larger speaker samples, instrumental acoustic analysis, and broader comparisons with neiɣbouring Malayic varieties to clarify the phonetic, sociolinguistic, and historical dimensions of these correspondences.

RECOMMENDATION

Several directions follow from these findings. First, acoustic-phonetic analysis is needed to measure the vowel quality of the [o] reflex and, in particular, to determine whether the dorsal rhotic is velar [ɣ] or uvular [ʁ], and how consistent it is across speakers. Second, comparative fieldwork with neiɣbouring varieties—Minangkabau to the south, Mandailing inland, and other coastal Malay communities—would clarify whether the Batahan correspondences are inherited, contact-induced, or independently innovated. Third, a larger sociophonetic study drawing on multiple villages, additional speakers, and a range of interactional settings would allow the inter-speaker and inter-generational patterns sketched here to be quantified and their social meaning tested directly. Together these steps would move the description of Batahan from preliminary documentation toward a fuller account of its sound system and its place in Sumatran dialectology.

ACKNOWLEDGMENT

The authors thank the native speakers of the Batahan community who generously participated in this study and shared their linguistic knowledge. The authors are also grateful to the local community leaders for facilitating access to participants and assisting with participant recruitment and local verification of the linguistic data. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Generative AI assistance was used in the preparation of this manuscript. Anthropic Claude Sonnet 4 was used to assist with English language editing, improving academic style and clarity, refining grammar and wording, and formatting selected sections of the manuscript. The AI tool was not used to generate, analyse, or interpret the linguistic data, nor to formulate the study's conclusions. All research design, data collection, transcription, analysis, interpretation, and final editorial decisions were undertaken and verified by the authors, who take full responsibility for the content of the manuscript.

REFERENCES

Al-Shenaifi, N., Azmi, A. M., & Hosny, M. (2024). Advancing AI-Driven Linguistic Analysis: Developing and Annotating Comprehensive Arabic Dialect Corpora for Gulf Countries and Saudi Arabia. *Mathematics*, *12*(19). https://doi.org/10.3390/math12193120

Al Shamsi, A. A., & Abdallah, S. (2023). Ensemble Stacking Model for Sentiment Analysis of Emirati and Arabic Dialects. In *Journal of King Saud University - Computer*

and Information Sciences (Vol. 35, Issue 8). https://doi.org/10.1016/j.jksuci.2023.101691

Alamsyah, T., Razali, Munawar, T., & Ramli. (2024). Variants and Social Interaction: A Study of the Acehnese Daya Dialect. *Studies in English Language and Education*, *11*(2). https://doi.org/10.24815/siele.v11i2.31018

Alsadhan, N. (2025). A Novel Dialect-Aware Framework for the Classification of Arabic Dialects and Emotions. *Journal of Computer Science*, *21*(1). https://doi.org/10.3844/jcssp.2025.88.95

Alshraah, S. M., Hamideh, W. N., Nishat, A. S., & Migdadi, H. F. (2024). A Sociopragmatic Analysis of Death Utterances in Jordanian Bedouin Society. *Journal of Language Teaching and Research*, *15*(1). https://doi.org/10.17507/jltr.1501.24

Ambalegin, A., & Arianto, T. (2018). English Vowels and Consonants Mispronunciation of the Seventh President of Republic of Indonesia in His Official English Speeches and Its Influencing Factors. *LANGUAGE LITERACY: Journal of Linguistics, Literature, and Language Teaching*, *2*(2). https://doi.org/10.30743/ll.v2i2.678

Aminah, S. (2014). Siti Aminah, M. Pd. In *Pronunciation* (pp. 70–74). Jakarta Publisher.

Anderson, Stephen R., & Saussure, L. de (Eds. . (2020). *Language and World: A Defence of Linguistic Idealism.* Edinburgh University Press.

Cheng, J. (2023). Acquisition of English liaison among Chinese EFL learners from the perspective of language transfer. *Heliyon*, *9*(10), e20418. https://doi.org/10.1016/j.heliyon.2023.e20418

Fiantika, F. R. E. a. (2022). Metodologi Penelitian Kualitatif. In Metodologi Penelitian Kualitatif. In Y. Novita (Ed.), *Rake Sarasin* (Cet.1, Issue March). PT. GLOBAL EKSEKUTIF TEKNOLOGI.

Hadi, S., & Mahdi, H. (2023). A Sociopragmatic Study of Feminism: A theoretical Perspective of Feminist Discourse. *Kufa Journal of Arts*, *1*(55). https://doi.org/10.36317/kaj/2023/v1.i55.11203

Hammouri, D. M., & Al-Khanji, R. R. (2023). Pragmatic Competence in Learner English: The Case of Jordanian EFL Learners. *Theory and Practice in Language Studies*, *13*(4). https://doi.org/10.17507/tpls.1304.30

Harahap, Y. M., & Rahmawati, W. T. (2024). Students' Phonemic Awareness In Mastery English Vowel And Consonant Sounds. *Pedagogi: Jurnal Ilmiah Pendidikan*, *10*(1). https://doi.org/10.47662/pedagogi.v10i1.653

Huberman, mILES A. (2014). Research Methodology. *International Journal of Instruction*, *15*(3). https://doi.org/10.29333/iji.2022.15350a

Hutapea, N. R. P., Lubis, N., Manik, S., & Manurung, L. W. (2023). Error Analysis of Students' Pronunciation in Pronouncing English Vowels And Consonants. *Jurnal Pendidikan Dan Konseling (JPDK)*, *5*(4). https://doi.org/10.31004/jpdk.v5i4.18042

J.W.Y., H. (2020). Code Mixing :Linguistic form and socio-cultura meaning. *The International Journal of Language Society and Culture.*, *21*.

Jalilpour Monesi, M., Vanthornhout, J., Francart, T., & Van hamme, H. (2024). The role of vowel and consonant onsets in neural tracking of natural speech. *Journal of Neural Engineering*, *21*(1). https://doi.org/10.1088/1741-2552/ad1784

Kulubekova, A. (2025). Linguistic Tools for Creating an Effective Political Communication Strategy. *Bulletin of Science and Practice*, *11*(3). https://doi.org/10.33619/2414-2948/112/58

Le, T. T., Pham, T. T., Nguyen, A. T., Phuong, H. Y., Huynh, T. A. T., & Nguyen, H. T. (2024). "Becoming a trilingual learner may be harder than I thought!":

Difficulties encountered by Vietnamese French-major students when studying English as a foreign language. *Ampersand*, *12*(August 2023), 100168. https://doi.org/10.1016/j.amper.2024.100168

Liu, Y., Groen, M. A., & Cain, K. (2023). The association between morphological awareness and reading comprehension in children: A systematic review and meta-analysis. *Educational Research Review*, *42*(November 2023), 100571. https://doi.org/10.1016/j.edurev.2023.100571

Moskovitz, F., Ft, A., & Road, L. (1989). *Intellectual Discourse* (D. M. Y. (Malaysia) (ed.); Islamic We, Vol. 31, Issue 3). IIUM Press, International Islamic University Malaysia P.O. Box 10, 50728 Kuala Lumpur, Malaysia.

Motlagh, M. A., Zavaraki, E. Z., Vahedi, M., Noroozi, A., Bakhtiarvand, M., & Azizi, Y. (2023). The Impact of Blended Collaborative Learning on the English Reading Comprehension. *Randwick International of Education and Linguistics Science Journal*, *4*(1). https://doi.org/10.47175/rielsj.v4i1.608

Mukaromah, S. L. N., Rahayu, A., Meliyana, R., Hutabarat, E. N., & Baehaqie, I. (2025). Contrastive Phonology of Javanese Temanggung Dialect and Javanese Banyumas Dialect. *Paramasastra: Jurnal Ilmiah Bahasa Sastra Dan Pembelajaranya*, *12*(1).

Nasution, N. (2022). Theory of phonology,vowel,and consonant. *Academia.Edu*.

Nugroho, R. A., & Setyaningsih, N. (2019). Interjections In Semarangan Javanese: A Sociopragmatic Approach. *Humanika*, *26*(2). https://doi.org/10.14710/humanika.v26i2.22184

Omar, A., & Alhassan, M. A. M. (2020). A Sociopragmatic Analysis of Women and Gender Roles in John Galsworthy's Forsyte Saga and Naguib Mahfouz's Cairo Trilogy. *International Journal of English Linguistics*, *10*(2). https://doi.org/10.5539/ijel.v10n2p284

Opare-Kumi, J. (2024). English medium instruction in multilingual contexts: Empirical evidence from Ethiopia. *International Journal of Educational Development*, *105*(June 2023), 102987. https://doi.org/10.1016/j.ijedudev.2024.102987

Östman, Jan-Ola, & Verschueren, J. (Eds.). (n.d.). *Pragmatics in Practice*. John Benjamins Publishing Company.

Piepers, J., Backus, A., & Swanenberg, J. (2023). 'It' Is Not for Everyone: Variation in Speakers' Evaluation of Sociopragmatic Pronouns in Limburgian. *Languages*, *8*(4). https://doi.org/10.3390/languages8040253

Povilonis, N. (2024). Variation in articulatory conflict resolution: Vowel allophony and consonant place adaptation in Chanka Quechua. *Journal of the International Phonetic Association*, *54*(3). https://doi.org/10.1017/S0025100324000252

Qin, Y., & Yu, F. (2025). An End-To-End Speech Recognition Model for the North Shaanxi Dialect: Design and Evaluation. *Sensors*, *25*(2). https://doi.org/10.3390/s25020341

Risnawati, E., Gunawan, W., Fatinova, D., & Mubarok, Y. (2024). Humor Discourse Structure: A Political Communication Strategy. *Scope : Journal of English Language Teaching*, *9*(1). https://doi.org/10.30998/scope.v9i1.23968

Saada, N. (2023). Educating for global citizenship in religious education: Islamic perspective. *International Journal of Educational Development*, *103*(February), 102894. https://doi.org/10.1016/j.ijedudev.2023.102894

Sakirman, S. (2017). Contemporary Fiqh Methodology in the Theory of the Limitation of Dialectics Space and Time According to Muhammad Syahrur. *HUNAFA: Jurnal Studia Islamika*, *14*(2). https://doi.org/10.24239/jsi.v14i2.484.301-326

Saleem, T., Anjum, U., & Tahir, S. (2021). The Sociopragmatic and Pragmalinguistic Strategies in L2 Pragmatic Competence: A Case of Pakistani ESL Learners. *Journal of Intercultural Communication Research*, *50*(2). https://doi.org/10.1080/17475759.2021.1877176

Sperlich, D., Leem, J., & Ahn, E. J. (2021). Explicit and implicit dimensions of sociopragmatic and pragmalinguistic competence. *IRAL - International Review of Applied Linguistics in Language Teaching*, *59*(3). https://doi.org/10.1515/iral-2018-0319

Syafiq, M. A., & Hafiz, M. (2023). Difficulties in pronouncing English vowel and consonant by English language education undergraduate students at UINSU. *Jurnal Ilmiah Multidisiplin*, *1*(6).

Tairova, F. (2024). The Sociopragmatic Aspects of Literary Discourse. *Bulletin of Science and Practice*, *10*(2). https://doi.org/10.33619/2414-2948/99/75

van der Elst-Koeiman, M., Segers, E., Severing, R., & Verhoeven, L. (2022). Learning to read in mother tongue or foreign language: Comparing Papiamento-Dutch reading skills in the post-colonial Dutch Caribbean. *Learning and Individual Differences*, *95*(March 2021), 102138. https://doi.org/10.1016/j.lindif.2022.102138

van Pinxteren, B. (2023). The relevance of medium of instruction and mother tongue for different types of educational systems. *International Journal of Educational Development*, *103*(June), 102923. https://doi.org/10.1016/j.ijedudev.2023.102923

Wang, J., & Zhang, L. (2020). Pragmatic functions of taboo words in Chinese online communication. *Discourse & Society*, 250–268.

Wewalaarachchi, T. D., & Singh, L. (2020). Vowel, consonant, and tone variation exert asymmetrical effects on spoken word recognition: Evidence from 6-year-old monolingual and bilingual learners of Mandarin. *Journal of Experimental Child Psychology*, *189*. https://doi.org/10.1016/j.jecp.2019.104698

Zhao, Y., & Chew, F. P. (2025). Mandarin-Chinese Dialects Code-Switching in Speech Communication. *Forum for Linguistic Studies*, *7*(7). https://doi.org/10.30564/fls.v7i7.9510