

## The reconstruction of protolanguage Tidore and Tobelo

Sulastri T. Gani<sup>1</sup>, Vivit Zarqa Hi. Muhamad<sup>2</sup>, Nurull Rizma Muhiddin<sup>3</sup>, Agung<sup>4</sup>, \*Nur Fajrhi<sup>5</sup>

<sup>1-5</sup>English Literature Study Program, Faculty of Cultural Sciences, Universitas Khairun, Indonesia

### Abstract

*This article aims to reconstruct the proto of two local languages spoken in North Maluku, Tidore (Td) and Tobelo (Tb) languages. North Maluku province is linguistically unique as it is home to two language groups, Austronesian and Non-Austronesian (Papuan). Despite its shared linguistic features, linguists argue that it is difficult to study language in this region because previous researchers have used different linguistic classifications. Languages in this area might be influenced by social, political, and economic conditions, and the strong contact between Austronesian and non-Austronesian languages significantly influences language use and development. Both Td and Tb are classified as members of Papuan language groups, particularly the North Halmahera stock. This study employs a descriptive qualitative method. The data are obtained through interviews with selected speakers using the 200 Swadesh words list. Several steps are performed in the data analysis stage. The first step is to describe the phonological system of each language, including phonemes, allophones, and their distribution. The step is followed by arranging the sound correspondence of the languages. The results of sound correspondence are used in the phonological reconstruction process to determine the proto-forms of the related languages. then continues with a diachronic comparative analysis to examine the historical relationship. The result reveals that both languages have five vowels, /a/, /i/, /u/, /e/, /o/ and 19 consonants, /p/, /b/, /m/, /w/, /f/, /t/, /d/, /n/, /s/, /r/, /l/, /c/, /j/, /ɲ/, /y/, /k/, /g/, /ŋ/, and /h/. The vowels of both languages appear to be stable. Each vowel segment in Td corresponds identically to its counterpart in Tb, suggesting these sounds were retained directly from the protolanguage without changes. Although the Td and Tb languages currently have 19 consonants, the reconstruction results only establish 16 proto-sounds, /\*p/, /\*b/, /\*m/, /\*w/, /\*f/, /\*t/, /\*d/, /\*n/, /\*s/, /\*l/, /\*ɲ/, /\*y/, /\*k/, /\*g/, /\*ŋ/, and /\*h/. This is because the phonemes /c/, /j/, and /r/ in both languages do not show regular correspondence patterns.*

**Keywords:** reconstruction, Tidore, Tobelo, proto-phonemes, North Halmahera Stock.

**Article History:** Received 30 Apr 2026, Final revision 7 Jun 2026, Published 27 Jun 2026

### Introduction

Indonesia is home to two large language groups, the Austronesian and the non-Austronesian, or Papua, Group (Putrohari et al., 2015). Currently, North Maluku province preserves 19 local languages, which belong to the Austronesian and Non-Austronesian language families, that coexist. The Td and Tb languages are part of the Papuan language family, specifically the North Halmahera subgroup (Voorhoeve, 1988). The Tidore (Td) language is an endangered language in the Southern part of Halmahera Island and on some islands, including Td, Maitara, and Mare (Eberhard, Simons, & Robinson, 2026a; Voorhoeve, 1988). The number of Td speakers is estimated at around forty thousand, distributed across Td Island and its surrounding areas. Almost all speakers are also fluent in North Maluku Malay, making bilingualism commonplace. In fact, children usually acquire Malay before learning Td, at around ages 4 to 6 years (van Staden, 2006). This condition indicates a shift in language use patterns. Within this group are also found the Ternate, Sahu, and West Makian languages, along with several dialects scattered across northeastern Halmahera, such as Galela, Modole, Tobaru, and Pagu (Voorhoeve, 1988). Meanwhile, The Tobelo (Tb) language is an indigenous language spoken

\*Corresponding author: Nur Fajrhi, Address: English Dept., FIB, Universitas Khairun, Jalan Pertamina, Kampus II Gambesi, Ternate Selatan, Ternate, Maluku Utara. Email address: [nur.fajrhi@gmail.com](mailto:nur.fajrhi@gmail.com) ©2026 Leksika. All rights reserved.

across the North-east Halmahera Island and two-thirds of the Morotai Island (Eberhard, Simons, & Robinson, [2026b](#); Rugebregt et al., [1986](#)). There is no specific reference that clearly states the estimated number of speakers of the Tb language. However, *Ethnologue* classifies it as a stable language, meaning that although it is not used in formal institutions, it is still actively used in family settings and is still being learned and used by children.

The presence of both languages is threatened by the dominance of North Maluku Malay, which is known as the lingua franca in daily conversation (D. M. Eberhard et al., [2026](#)). Most community members choose to use this variant of Malay because it is considered more practical and has a broader social scope (Bowden, [2012](#)). Therefore, the use of language is declining, especially among younger generations, increasing the risk of language shift or, worse, extinction. Moreover, some shared upshots of these languages include Austronesian influence due to extensive contact in the region. One of the most common features is that both languages have the same word order. Both Td and Tb follow the same word order, Agent-Verb-Object (AVO) (Holton & Klamer, [2017](#)). A note is given that such word order only appears in the written form of Tb. Morphologically, both languages exhibit interesting characteristics. Both languages happen to have verbal prefixes that mark subjects (Holton, [2006](#); Lekawael, [2013](#); Abd Halil, [2022](#); Mahmud & Tamalene, [2025](#)). Moreover, the histories of these two languages cannot be separated from the strategic importance of North Maluku. Both languages have a very close connection to Ternate, which was a center of politics and trade, resulting in many loanwords from this connection. Evidence of loanwords can be seen in Old Javanese-influenced terminology related to trade, ritual objects, musical instruments, and non-local flora and fauna in both languages (Schapper & Zielenbach, [2025](#)).

The investigation of language in the region of North Maluku is considerably challenging because the previous study adopted a different linguistic classification; language in the area might be influenced by non-linguistic aspects such as social, political, and economic conditions; and the strong contact between the Austronesian and non-Austronesian in the region (Putrohari et al., [2015](#)). However, several comparative studies in the area have been conducted previously. (Grimes & Grimes, [1984](#)) used the lexicostatistics method to measure differences between languages in North Maluku. The result showed that Td and Tb share 37% basic vocabulary, indicating that the languages belong to the same language stock. (Wada, [1980](#)) studied consonant correspondences among the North Halmahera Languages to support the reconstruction of Proto-North Halmahera (PNH) by comparing vocabularies from several languages, including Galela, Tobelo, and Tidore. However, Wada did not reconstruct PNH but provided a word list and sound correspondence that can be used for future reconstruction studies. (Chlenov, [1986](#)) argues that the internal classification of the North Halmahera languages remains unclear, as linguists have yet to agree on how to divide the languages into smaller groups. Chlenov then proposed three subgroups, the Halmahera subgroup, the Jailolo subgroup, and the Island subgroup. Still, he notes that the available evidence was not strong enough to fully confirm these groups because PNH had not been reconstructed. Therefore, a study of the relationship among the North Halmahera languages is needed.

This paper aims to reconstruct the proto-languages of Td-Tb spoken in North Maluku Province, Indonesia. The reasons for choosing Td and Tb are that both represent variations in North Halmahera: Td is strongly influenced by trade contacts (bilingualism with Malay, Old Javanese loanwords), while Tb retains authentic Papuan features despite being in the political orbit of Ternate-Td. This selection is crucial because proto-phonological documentation remains limited, and both are vulnerable to extinction due to a shift toward Malay; the reconstruction supports the preservation and understanding of the history of Papuan languages in Wallacea.

## Method

This study employs descriptive qualitative. The primary data were the 200-Swadesh wordlist from each language, collected through interviews from selected informants who followed the criteria from (Keraf, 1996) (Keraf, 1996). The first stage of data analysis applies a synchronic approach to provide detailed information for both languages' phonological systems. The synchronic approach examines linguistic structure and patterns at a specific point in time based on the language's internal features and rules (Anderson, 2016). This analysis warrants that the languages being compared are genuinely related, not just similar by coincidence (Dawson & Phelan, 2016). The second stage of data analysis is to trace the genealogical relationship between two languages using diachronic-comparative analysis. The stage includes identifying the cognates between the compared languages. The cognates are then being reconstructed to determine the proto-form of Td and Tb. This process uses a bottom-up approach, starting with direct observation of the data and gradually culminating in the proto-forms. The reconstruction process involves several phases. Firstly, all phonemic correspondences are classified based on their phonological position (initial, medial, final). Secondly, each correspondence is investigated to identify the most economical proto-phoneme supported by the majority of the cognate data. The next phase is to formulate the sound-change rules that explain how the proto-phonemes developed into their modern form. Lastly, proto-phonemes are reconstructed based on the result of sound correspondence and sound-change rules. Every phase is evaluated to ensure internal consistency and the absence of contradictions across phonological environments, as every sound change hypothesis should follow a regular pattern and be verifiable across the data (Hock & Joseph, 2009).

## Results and Discussion

This section presents results and discussions, which are overall divided into four parts. The first part provides a synchronic analysis of the Td language, including the inventory of vowels and consonants with their distribution. The second part describes the vowels and consonants of Tb with their distribution. This part is followed by an assessment of sound correspondence between Td and Tb to show how the sounds are related. The last section focuses on reconstructing the proto-phonemes of Td and Tb to see the historical connection between the two languages.

### Phonemes of Tidore Language

The Td language has five vowel phonemes. The phonemes include /a/, /i/, /u/, /e/, and /o/. The inventory of the phonemes is displayed in the following table.

Table 1. Vowels Inventory of Td Language

	Front	Central	Back
High	i		u
Middle	e		o
Low		a	

These vowels are identified through their contrastive distribution within their phonological environment. Each vowel occurs in the same environment, where replacing one with another may

change the meaning (Lodge, 2009). These contrasts prove the internal consistency of Td's vowel list (Maddieson, 2005). The following examples reveal the contrastive distribution of vowels in Td.

Table 2. Minimal Pair of Vowels Inventory in Td Language

Phoneme	words	meaning
/a/ : /o/	/goga/	cold
	/gogo/	hair
/u/ : /o/	/una/	to vomit
	/ona/	they
/e/ : /i/	/ake/	water
	/aki/	tongue

Additionally, the Td develops 19 consonant phonemes. These phonemes are categorized based on their place and manner of articulation, which vary from the labial plosives to the glottal fricative. The following table is an overview of the Td's consonant system.

Table 3. Consonants Inventory of Td Language

	Labial	Labiodental	Alveolar	Palatal	Velar	Glottal		
Plosive	p	b	t	d	c	ɟ	k	g
Nasal		m		n		ɲ		ŋ
Trill				r				
Fricative		f	s					h
Approximant		w			y			
Lateral approximant				l				

\* Symbols to the right in a cell are voiced, to the left are voiceless.

The inventories of these consonants are proved through their contrastive distribution. This approach The 19 previously inventoried consonants can be verified through contrastive distribution. This methodological approach grants the distinction of each phoneme by monitoring their behavior in the same phonological environments. Therefore, these lists are evidence that Td's consonant systems differ from those of other phonemes of the Td language.

Table 4. Minimal Pairs of Consonants in Td Language

Phoneme	words	Meaning
/b/ : /f/	/yobo/	swollen
	/yofo/	smoke
/h/ : /y/	/yobo/	swollen
	/hobo/	wet
/p/ : /r/	/paha/	thighs
	/raha/	four

Phoneme	words	Meaning
/f/ : /m/	/ɲofa/	child
	/ɲoma/	star
/w/ : /r/	/waje/	sun
	/raje/	three
/t/ : /r/	/kotu/	black
	/koru/	to kill
/g/ : /d/	/gofu/	grass
	/dofu/	a lot
/c/ : /n/	/oca/	to hit
	/ona/	she, he, they
/ɣ/ : /ʒ/	/faya/	wife
	/faʒa/	dirty
/s/ : /r/	/haso/	heavy
	/haro/	to come
/k/ : /g/	/koko/	to stand
	/gogo/	hair
/l/ : /n/	/lamo/	big
	/namo/	bird
/ŋ/ : /ʒ/	/ŋau/	ear
	/ʒau/	to touch
/l/ : /ɲ/	/lao/	eye
	/ɲao/	fish

The number of consonants found in this study matches the previous study, although the types of consonants found failed to agree with the previous study. (Duwila, [2021](#)) suggested that the Td language includes a voiceless uvular plosive /G/. However, the data from the 200 basic vocabulary words do not show any evidence of this specific sound. Instead, this research identifies a voiceless velar nasal /ŋ/ within the language. Finally, the study provides strong evidence for this sound through minimal pairs. Syllables in the Td language follow several distinct patterns. They can consist of a single

V, such as /a/ in /ake/ 'water', or a CV structure like in /lamo/ 'big'. Other patterns include VC, as found in /ende/ 'that', and CVC, such as in /ɲun/ ('nose').

### Phonemes of Tobelo Language

The vowel inventory of the Tb language consists of five sounds: /a/, /i/, /u/, /e/, and /o/. The sounds can be mapped as follows:

Table 5. Vowels Inventory of Tb Language

	Front	Central	Back
High	i		u
Middle	e		o
Low		a	

These five phonemes can be demonstrated by their contrastive distribution, which means each phoneme plays an important role in distinguishing meaning when it occurs in the same environment (Dawson & Phelan, 2016). Several examples of minimal pairs are presented below:

Table 6. Minimal pair of Vowels in Tb Language

Phoneme	Words	Meaning
/a/ : /e/	/ame/	to kiss
	/eme/	mother
/a/ : /o/	/gahi/	salt
	/gohi/	egg
/a/ : /i/	/hoka/	leaf
	/hoki/	able
/u/ : /o/	/idu/	sleep
	/ido/	tuna fish

Moreover, the language has 19 consonant phonemes. The inventory can be seen below:

Table 7. Consonants Inventory of Tb Language

	Labial		Labiodental		Alveolar		Palatal		Velar		Glottal	
Plosive	p	b			t	d	c	j	k	g		
Nasal		m				n		ɲ		ŋ		
Trill						r						
Fricative			f		s						h	
Approximant		w						y				
Lateral approximant						l						

\* Symbols to the right in a cell are voiced, to the left are voiceless.

These consonant inventories can be established through their contrastive distribution as the primary evidence for categorizing sounds as separate functional units rather than mere variations of a single sound. The distribution can be seen below:

Table 8. Minimal Pairs of Consonants in Td Language

Phoneme	Words	Meaning
/p/ : /l/	/pehaka/	wet
	/pelaka/	to split
/m/ : /r/	/akeme/	meat
	/akere/	water
/f/ : /h/	/kafo/	ash
	/kaho/	dog
/k/ : /r/	/uku/	fire
	/uru/	mouth
/h/ : /n/	/ŋohi/	l
	/ŋoni/	we
/g/ : /ŋ/	/ŋohi/	l
	/gohi/	egg
/d/ : /h/	/duhu/	to hunt
	/huhu/	breast
/t/ : /r/	/nautu/	other
	/nauru/	man
/p/ : /b/	/peda/	sagoo
	/beda/	to bother

Phoneme	Words	Meaning
/h/ : /w/	/hane/	three
	/wane/	sun
/t/ : /c/	/jara/	horse
	/cara/	how
/s/ : /ŋ/	/sauku/	hot
	/ŋauku/	ear

The sounds /y/ and /ɲ/ are considered separate phonemes because they both appear at the beginning and in the middle of words. Since they can exist in the same positions, their appearance is not controlled by specific phonetic rules or nearby vowels (Baroni, 2023; Doucette, O'Donnell, et al., 2024). This unpredictability indicates that the sounds function as separate phonemes in the system (Donegan, 2015). Although the minimal pairs of the sounds are not found, the distribution of the voiced approximant palatal /y/ and voiced nasal palatal /ɲ/ indicates that they are different sounds. Both phonemes appear in the beginning and the middle of a word. The following data shows the specific environment in which each sound occurs.

Table 9. Position of /y/ and /ɲ/ in Td Language

Phoneme	Initial	Medial
/y/	/youku/ to burn	/iyata/ four
/ɲ/	/ɲi/ room	/ɲaɲi/ to sing

Moreover, most data indicate that Td is a vocalic language in which typical syllables follow the consonant-vowel (CV) pattern. Through the 200 Swadesh list, two words appear in CVS structure. Both sounds end with the phoneme /s/, such as the /s/ sound, as in /ramas/ 'to squeeze' and /napas/ 'to breathe, ' suggesting that both words are borrowed from other languages. Consequently, the close syllables are not an original characteristic of the Tb phonotactics.

### Correspondence Set of Td and Tb Language

The cognate data of Td and Tb indicate that all vowel phonemes of Td and Tb are reflected identically without any change. These consistencies are remarked through the following corresponding sets.

Table 10. Vowels Correspondence of Td and Tb

Correspondence	Gloss	TD	TB
TD /a/ : TB /a/	2. water	<i>ake</i>	<i>akere</i>
	6. child	<i>ɲofa</i>	<i>ɲohaka</i>

	129. tongue	<i>aki</i>	<i>akiri</i>
TD /e/ : TB /e/	47. to kiss	<i>hame</i>	<i>ame</i>
	140. died	<i>sone</i>	<i>honejen</i>
	82. liver	<i>gate</i>	<i>gate</i>
TD /i/ : TB /i/	4. I	<i>ηori/ηoti</i>	<i>ηohi</i>
	73. salt	<i>gasi</i>	<i>gahi</i>
	129. tongue	<i>aki</i>	<i>akiri</i>
TD /o/ : TB /o/	6. child	<i>ηofa</i>	<i>ηohaka</i>
	28. swollen	<i>yobo</i>	<i>yoboho</i>
	138. eye	<i>lao</i>	<i>lako</i>
TD /u/ : TB /u/	10. fire	<i>uku</i>	<i>uku</i>
	3. root	<i>mutu</i>	<i>ηutuku</i>
	157. stomach	<i>our</i>	<i>pokoru</i>

The identical correspondence of vowels in Td and Tb suggests that these phonemes are retained directly from their proto-language without changes. The consistency may signal that the proto-vowel system has been actively preserved in both languages (Doucette, O'donnell, et al., 2024). The absence of vowel change reflects a highly stable phonological development within the North Halmahera language stocks. Some words show some structural change, such as sound additions at the end of the word in Tb. However, these changes do not change the quality of the vowels as the correspondences are mostly identical, leading to the reconstruction of the proto-word as the simpler words are considered to reflect the ancient sounds perfectly (Ohala, 2017).

Although the vowel phonemes seem stable in both languages, the consonant correspondences show more variation and changes. The synchronic analysis of Td and Tb consonants has shown that both languages possess the same type and number of consonants. However, their correspondence is not always as consistent as seen in vowels. Some consonants remain unchanged, while some other consonants undergo adjustments in distribution. The difference highlights that consonant phonemes in both languages have developed more dynamically than the vowel phonemes. The examples below show how some consonants remain the same, while others change in certain words.

Table 11. Unchanged Consonants Correspondence of Td and Tb

Correspondence	Gloss	TD	TB
TD /p/ : TB /p/	154. short	<i>podo</i>	<i>kopodo</i>
TD /b/ : TB /b/	28. swollen	<i>yobo</i>	<i>yoboho</i>
	134. knee	<i>bukubuku</i>	<i>buku</i>

<b>Correspondence</b>	<b>Gloss</b>	<b>TD</b>	<b>TB</b>
TD /m/ : TB /m/	21. new	<i>suŋ / masuŋ</i>	<i>mahunji</i>
	33. big	<i>lamo</i>	<i>mapako/lamoko</i>
	47. to kiss	<i>hame</i>	<i>ame</i>
	92. mother	<i>em/yaya</i>	<i>eme</i>
	131. five	<i>matoha/</i>	<i>matoa</i>
TD /w/ : TB /w/	35. animal	<i>haiwan</i>	<i>haiwan</i>
	139. sun	<i>waje</i>	<i>waje</i>
	162. navel	<i>wotu</i>	<i>wotiti</i>
TD /t/ : TB /t/	3. root	<i>utu</i>	<i>ŋutuku</i>
	30. to walk	<i>tagi</i>	<i>majobo/hotagi</i>
	51. lake	<i>talaga</i>	<i>taaga</i>
	131. five	<i>matoha</i>	<i>matoa</i>
	162. navel	<i>wotu</i>	<i>wotiti</i>
TD /n/ : TB /n/	91. he/she	<i>una</i>	<i>una</i>
	107. you	<i>ŋona</i>	<i>ŋona</i>
	109. because	<i>karna</i>	<i>karna</i>
	126. man	<i>nanau/nau</i>	<i>nauru</i>
	142. they	<i>ona</i>	<i>onaŋa</i>
TD /l/ : TB /l/	33. big	<i>lamo</i>	<i>mapako/lamoko</i>
	46. worm	<i>kaltidi</i>	<i>kulubati</i>
	62. cold	<i>alo/goga</i>	<i>ma'alo</i>
	77. to bite	<i>goli</i>	<i>goli</i>
TD /y/ : TB /y/	12. smoke	<i>yofu</i>	<i>yoho</i>
	28. swollen	<i>yobo</i>	<i>yoboho</i>
	104. foot	<i>yohu</i>	<i>you</i>
TD /k/ : TB /k/	2. water	<i>ake</i>	<i>akere</i>
	8. dog	<i>kaso</i>	<i>kaho</i>

Correspondence	Gloss	TD	TB
	10. fire	<i>uku</i>	<i>uku</i>
TD /g/ : TB /g/	30. to walk	<i>tagi</i>	<i>tagi</i>
	51. lake	<i>talaga</i>	<i>taaga</i>
	77. to bite	<i>goli</i>	<i>goli</i>
	73. salt	<i>gasi</i>	<i>gahi</i>
	82. liver	<i>gate</i>	<i>gate</i>
TD /ŋ/ : TB /ŋ/	4. I	<i>ŋori</i>	<i>ŋohi</i>
	6. child	<i>ŋofa</i>	<i>ŋohaka</i>
	21. new	<i>sung / masuŋ</i>	<i>mahuŋi</i>
	71. you	<i>ŋona</i>	<i>ŋona</i>

Based on the data, 11 consonant sounds were found to occupy conservative positions. Those sounds include /p/, /b/, /m/, /w/, /t/, /d/, /n/, /l/, /y/, /k/, /g/ and /ŋ/. This phenomenon indicates that both languages still retain the original features of their proto-language in those specific segments. As a result of this sound conservation, the level of formal similarity between the two languages remains remarkably high and clearly visible. Therefore, this facilitates the identification of cognates without the need for complex reconstruction procedures.

In addition to the eleven conservative sounds, several other sounds were found to undergo regular correspondence within the data. This regular correspondence indicates a consistent and systematic pattern of sound change between the two languages. This phenomenon proves that the phonetic differences are not coincidental, but rather the result of measurable phonetic evolution. These regular patterns can be clearly observed through the following series of cognate comparison.

Table 12. Change Consonants Correspondence of Td and Tb

Correspondence	Gloss	TD	TB
TD /f/ : TB /h/	6. child	<i>ŋofa</i>	<i>ŋohaka</i>
	12. smoke	<i>yofa</i>	<i>yoho</i>
	72. to dig	<i>fai</i>	<i>haiti</i>
TD /ŋ/ : TB /n/	93. fish	<i>ŋao</i>	<i>naoko</i>
	108. right	<i>kaŋira</i>	<i>niraka</i>
TD /h/ : TB /∅/	35. animal	<i>haiwan</i>	<i>haiwan</i>

Correspondence	Gloss	TD	TB
	47. to kiss	<i>hame</i>	<i>ame</i>
	70. four	<i>raha</i>	<i>iyata</i>
	104.foot	<i>yohu</i>	<i>you</i>
	124.wide	<i>woha</i>	<i>iwoata</i>
	131.five	<i>matoha/</i>	<i>matoa</i>
TD /s/ : TB /h/	8. dog	<i>kaso</i>	<i>kaho</i>
	21. new	<i>sung / masung</i>	<i>mahungi</i>
	73. salt	<i>gasi</i>	<i>gahi</i>
	74. to scratch	<i>koraso</i>	<i>rangho</i>
	140. to die	<i>sone</i>	<i>honengen</i>
	150.hot	<i>sahu</i>	<i>hauku</i>

The correspondence between TD /f/ and TB /h/ indicates that the /f/ sound is the proto-form that underwent a process of sound weakening known as lenition (Campbell, 2013). Fricative sounds /f/, which require stronger articulation, often weaken into the laryngeal sound /h/ (Crowley & Bower, 2010). This is a classic example of sound lenition, which shows that the oral sound is lost because /h/ is easier to pronounce (O'Brien, 2012). This direction of change is universal and unidirectional, making it nearly impossible for an aspirated sound to strengthen back into a fricative. Thus, /f/ in TD reflex the original phonetic features from the proto language, while TB /h/ has experienced phonetic innovation.

The correspondence between TD /ɲ/ and TB /n/ demonstrates palatalization. The idea of sonority hierarchy shows that sounds can be ordered by their strength (Crowley & Bower, 2010). The sounds /n/ and /ɲ/ are both voiced nasals, but /n/ is simpler to produce than /ɲ/, which needs more complex tongue movement. Simpler sounds are more likely to be proto-sounds. Therefore, /n/ is more likely the proto sound, while /ɲ/ is a later change from palatalization. This result is supported by previous research by (Chlenov, 1986), who states that the sound /ɲ/ may be a phonological innovation resembling a sound found in Austronesian languages. Also, in the reconstruction of the proto-subgroup languages in the bird heads of Papua, the phoneme /ɲ/ is not found (Klamer & Holton, 2018).

The correspondence between TD /h/ and TB /∅/ is an example of syncope, which is the loss of segments in the middle of words (Crowley & Bower, 2010). This process results in the total disappearance of /h/ articulation in TB. The loss of the /h/ sound in Tb shows Tb has a more advanced sound change in some aspects compared to Td. When compared with Galela, which belongs to the same language group, Tb also shows similar changes. Some sounds that still exist in Galela, including /h/, are lost in Tobelo (Wada, 1980). The language that retains the /h/ sound (TD) is considered more conservative than the language that has lost it (TB). As a result of this deletion, the phonotactic structure in TB undergoes simplification, often followed by the merging of two vowels that were

previously separated. The correspondence between TD /s/ and TB /h/ is one of the best-known cases of sound weakening in human language because of its high frequency (Wada, 1980). The more complex fricative /s/ tends to decay into the laryngeal /h/ because it is easier to articulate (Crowley & Bower, 2010).

Aside from regular correspondence data, some sounds lacked sufficient evidence to form a consistent correspondence pattern. These sounds appear to have irregular correspondence throughout the compared cognate data. This sporadic pattern often results from outside influence, such as borrowing or incomplete analogical change (Gardani, 2020; Weiß, 2024). Due to this limited data, the historical change patterns for these sounds are extremely difficult to track systematically. Therefore, these sporadic data points are categorized separately as they cannot serve as a robust basis for determining regular sound changes (Miceli & Round, 2022). The first non-representative set can be seen through the voiced alveolar trill phoneme /r/ in the Td language that does not have any regular correspondence to the Tb language below:

Table 13. Irregular Consonants Correspondence of Td and Tb

Correspondence	Gloss	TD	TB
TD /r/ : irregular TB	70. four	<i>raha</i>	<i>iyata</i>
	109.because	<i>karna</i>	<i>karna</i>
	146.name	<i>roŋa</i>	<i>romaŋa</i>
	166.one	<i>remoi</i>	<i>moi</i>
	186.fly	<i>soro</i>	<i>hoho</i>
	191.three	<i>raŋe</i>	<i>haŋe</i>
	194.to blow	<i>lru</i>	<i>lru</i>

In gloss 70, phoneme /r/ in TD corresponds with phoneme /y/ in TB. Phoneme /r/ in TD did not undergo any sound change in TB in gloss 109 and 146. Moreover, in gloss 166, phoneme /r/ in TD, along with phoneme /e/, are deleted; therefore, only the sound /moi/ in TB remains for that gloss. In gloss 186 and 191, phoneme /r/ in TD changed into phoneme /h/ in TB. Lastly, phoneme /r/ in TD change into /t/ in TB. The lack of regular patterns makes these sound changes hard to track. Additional data is needed to map the phonetic correspondence of /r/ between TD and TB.

Table 14. Addition in Tb

Correspondence	Gloss	TD	TB
Prothesis in TB	30. to walk	<i>tagi</i>	<i>majobo/hotagi</i>
	62. cold	<i>alo/goga</i>	<i>maalo</i>
	142.they	<i>ona</i>	<i>naona</i>
	154.short	<i>podo</i>	<i>kopodo</i>
	157.stomach	<i>oru</i>	<i>pokoru</i>

Correspondence	Gloss	TD	TB
Epenthesis in Tb	146 name	<i>ronja</i>	<i>romanja</i>
	138 eye	<i>lao</i>	<i>lako</i>
Paragoge in TB	3. root	<i>mutu</i>	<i>ηutuku</i>
	6. child	<i>ηofa</i>	<i>ηohaka</i>
	28. swollen	<i>yobo</i>	<i>yoboho</i>
	33. big	<i>lamo</i>	<i>mapako/lamoko</i>
	76. tooth	<i>lη</i>	<i>injiri</i>
	93. fish	<i>nyao</i>	<i>naoko</i>
	129. tongue	<i>aki</i>	<i>akiri</i>
	145. to vomit	<i>una</i>	<i>unanja</i>

In addition to the /r/ phoneme, there is a phenomenon of syllable addition from TD to TB. This addition can occur in the word-initial (prothesis), medial (epenthesis), or final (paragoge) positions. The addition of a sound at the initial position of a word without altering its core meaning is known as prothesis (Campbell, 2013). To date, no regular pattern has been identified for these additions. Therefore, the phenomenon remains categorized as sporadic or unsystematic change.

Table 14 also illustrates the presence of sound addition in the medial position of words in TB compared to TD which is known as epenthesis (Campbell, 2013). The insertion of a sound into a word to facilitate articulatory transitions or to satisfy specific syllabic structures. However, as of now, again, no regular pattern has been identified for its occurrence. These additions are still considered sporadic sound changes. Moreover, Table 14 also reveals sound addition in the word-final position in TB compared to TD through the phenomenon of paragoge, the adjustment of words to phonotactic rules without altering core meanings (Campbell, 2013). The irregular pattern in these additions makes this phenomenon a sporadic change within the dataset. The West Papuan typology regularly applies paragoge to maintain the open-syllable system within its structures (Fatkhayati, 2023). This phonetic element in TB provides a key indicator for tracing a morphological history that is more complex than that of its sister languages.

### The Reconstruction of Proto-Td Tb

This section portrays the results of the proto-phoneme reconstruction derived from the Td and Tb languages. Through the comparative method, sound correspondences are analyzed to determine the most phonetically plausible proto-forms. This reconstruction covers both vowel and consonant phonemes, illustrating the genetic relationship between the two languages.

The following are the results of the proto-vowel phoneme reconstruction for the Td and Tb languages. Based on the correspondence data, the vowel system at the proto-level demonstrates a high degree of stability. The list below details these vowel phonemes along with their distribution patterns in both languages.

Table 15. Proto vowels TdTb

	Front	Central	Back
High	*i		*u
Middle	*e		*o
Low		*a	

The Proto Td-Tb (TdTb) vowel system is reconstructed with five standard vowel phonemes: *\*a*, *\*i*, *\*u*, *\*e*, and *\*o*. The vowel is symmetrically distributed across front, central, and back positions at high, mid, and low heights. The inventory of Vowels with 5 members is classified as average vowel quality, making 51.2 % of the world's language (Maddieson, 2013). This five-vowel system is a common feature throughout the North Halmahera Stock, including the Galela and Sahu (Chlenov, 1986; Voorhoeve, 1988; Wada, 1980). This uniformity suggests that languages within this group are highly conservative in maintaining the vowel structure of their proto-language. In world language typology as well as among Papuan languages, the five-vowel system is considered the most stable and commonly encountered configuration (Lynch, 2020). The identical vowel patterns between TdTb and other North Halmahera members reinforce the evidence of a close genetic relationship within the stock.

Based on the comparative analysis, the reconstruction of Proto Td-Tb (PTdTb) consonant phonemes yields a total of 15 phonemes. These seventeen proto-phonemes are shown in the following table.

Table 16. Proto Consonants of TdTb

	Labial		Labiodental		Alveolar		Palatal		Velar		Glottal	
Plosive	*p	*b			*t	*d			*k	*g		
Nasal		*m				*n				*ŋ		
Trill												
Fricative			*f		*s							*h
Approximant		*w						*y				
Lateral approximant						*l						

\* Symbols to the right in a cell are voiced, to the left are voiceless.

The reconstruction of 15 proto-consonants for Proto TdTb produces an inventory that is consistent with the typology of the North-Halmahera stocks. The phonological system of the North Halmahera languages is relatively simple, with open syllables, which is aligned with the consonant system derived here. This finding also resonates with studies from (Klamer & Holton, 2018) and Wada (1980). Their work includes broader comparative studies addressing consonant correspondence in the North Halmahera languages. Even so, Wada (1980) did not propose any explicit list of proto-phonemes of the North Halmahera languages. Chlenov (1986) proposed 18 proto-consonants of the North Halmahera language. The present reconstruction of 15 consonants did not contradict Chlenov, it suggests proto-consonant variation across the subgroup.

The exclusion of /t/, /c/, and /r/ from the proto TdTb involves theoretical grounding. The comparative method holds that a proto-phoneme can be reconstructed when cognates exhibit regular correspondences across daughter languages (Campbell, 2013; Crowley & Bowerman, 2010). The absence of such regularity for these phonemes indicates either independent innovation or borrowing after the proto-phonemes period. The fact that those phonemes have minimal pairs in their synchronic analysis may strengthen this hypothesis. Moreover, the palatal nasal /ɲ/ in Td was resolved by reconstructing /\*n/, as Chlenov (1986) notes that /ɲ/ only appears in Loloda, Sahu, Ternate, and Td. Reconstructing /\*n/ over /\*ɲ/ is typologically justified as dental and alveolar nasals are universally attested, while palatal nasals are represented as a less stable category (Ladefoged & Maddieson, 1996).

One of the most important findings in this study is the merger of proto phonemes /\*f/ and /\*s/ into a single sound /h/ in Tb. The phenomenon is called a merger. A merger occurs when two different sounds in an older language become one single sound in later languages (Campbell, 2013). Once two sounds have merged, it is difficult to tell them apart in modern language. This study found this phenomenon only after comparing Td and Tb. In this case, Tidore still maintains /f/ and /s/ as separate phonemes, providing evidence for the reconstruction of two distinct sounds. The choice of /\*f/ and /\*s/ over /\*h/ is not unusual. Examples from linguistic evidence, such as the change of Latin /f/ to /h/ in Spanish, confirm it. The process is driven by articulatory weakening, in which a consonant with specific oral places of articulation is reduced to a laryngeal sound. The case of /h/ in Tb fits this phonology tendency, and the reconstruction of /\*f/ and /\*s/ as the proto form is supported by internal evidence and broad linguistic patterns.

The fact that Td has a more complex sound system than Tb tells us something important about how the two languages developed after their separation from their common ancestors. Td retains more of the original sounds, while Tb went through more changes. The 15 proto-consonants and the five-vowel system found in this study are a useful starting point for Td and Tb with other North Halmahera languages such as Galela, Tobaru, Sahu, or Ternate.

## Conclusion

The study of Td and Tb languages shows that both keep a very stable five-vowel system from their ancestor. This vowel pattern is a common feature in the North Halmahera group and indicates a close family relationship between them. Moreover, both languages have 19 consonants in their modern form. The results of sound correspondence show that both languages retain the same vowel phonemes as their ancestors. However, some sound changes were identified during the reconstruction, especially in the Tb language. The consonant reconstruction identified 15 proto-consonants, excluding /t/, /c/, and /r/ because they do not follow any of the regular correspondences. The proto consonants also revoked the palatal nasal /ɲ/ in Td, as it is considered weaker than /n/ in Tb. In addition, Td is reflecting a more archaic form, as it is considered simpler than Tb.

These findings provide a clearer map of the historical development of languages in the North Halmahera Stocks or in the West Papuan Group. Future research that applies the same methods to North Halmahera languages could help build a more complete picture of the Proto North Halmahera language and answer whether the simplification seen in Tb is local change or part of a bigger pattern across the stock.

## References

- Abd Halil, M., & Hilmi, H. S. (2022). Jenis dan Bentuk Afiksasi dalam Bahasa Tidore (Tinjauan Morfologi). *NUSRA: Jurnal Penelitian Dan Ilmu Pendidikan*, (1). <https://doi.org/DOI:https://doi.org/10.55681/nusra.v3i1.158>
- Anderson, S. R. (2016). Synchronic Versus Diachronic Explanation and the Nature of the Language Faculty. In *Annual Review of Linguistics* (Vol. 2, pp. 11–31). Annual Reviews Inc. <https://doi.org/10.1146/annurev-linguistics-011415-040735>
- Baroni, A. (2023). On the phonological interpretation of |A|, |I|, and |U|. In *Primitives of Phonological Structure* (pp. 72–85). Oxford University Press, Oxford. <https://doi.org/10.1093/oso/9780198791126.003.0003>
- Bowden, J. (2012). Local Languages, Local Malay, and Bahasa Indonesia; A case study from North Maluku. *Wacana, Journal of the Humanities of Indonesia*, 14(2). <https://doi.org/10.17510/wacana.v14i2.65>
- Campbell, Lyle. (2013). *Historical Linguistics : An Introduction* (Third Edition). Edinburgh University Press.
- Chlenov, M. A. (1986). North Halmahera: A Problem in the Internal Classification. *New Guinea Linguistics*, 24, 39–44. <https://doi.org/10.15144/PL-A70.39>
- Crowley, Terry., & Bower, Claire. (2010). *An Introduction to Historical Linguistics* (Fourth Edition). Oxford University Press.
- Dawson, H. C., & Phelan, M. (Eds.). (2016). *Language Files: Materials for an Introduction to Language and Linguistics* (Twelve Edition). The Ohio State University.
- Donegan, P. (2015). The Emergence of Phonological Representation. In *The Handbook of Language Emergence* (pp. 33–52). Wiley. <https://doi.org/10.1002/9781118346136.ch1>
- Doucette, A., O'donnell, T. J., & Sonderegger, M. (2024). Investigating the Universality of Consonant and Vowel Co-occurrence Restrictions. *Glossa: A Journal of General Linguistics*, 9(1), 1–9. <https://doi.org/https://doi.org/10.18452/1379>
- Doucette, A., O'Donnell, T. J., Sonderegger, M., & Goad, H. (2024). Investigating the universality of consonant and vowel co-occurrence restrictions. *Glossa: A Journal of General Linguistics*, 9(1). <https://doi.org/10.16995/glossa.9373>
- Duwila, E. (2021). Perbandingan Ciri-Ciri Linguistik Bahasa Ternate dengan Bahasa Tidore. *Tekstual*, 19(02).
- Eberhard, D. M., Simons, G. F., & Robinson, A. J. (2026a). Tidore Language. In *Ethnologue: Languages of the World*. SIL Global. <https://www.ethnologue.com/language/tvo/>
- Eberhard, D. M., Simons, G. F., & Robinson, A. J. (2026b). Tobelo Language. In *Ethnologue: Languages of the World* (Twenty-ninth edition). SIL Global. <https://www.ethnologue.com/language/tlb/>
- Eberhard, D. M., Simons, G. F., & Robinson Alison J. (2026). North Moluccan Malay. In *Ethnologue: Languages of the World* (Twenty-ninth edition). SIL Global.
- Fatkhiyati, N. (2023). Morphological process found in Fore language in Papua New Guinea. *Lililacs Journal: English Literature, Language, and Cultural Studies Journal*, 3(1), 15–23. <https://doi.org/10.21009/lililacs.031.02>

- Grimes, C. E., & Grimes, B. D. (1984). Languages of the North Moluccas: A Preliminary Lexico-Statistic Classification. In E. K. M. Masinambow (Ed.), *Buletin Leknas: Maluku dan Irian Jaya: III* (Number 1). Lembaga Ekonomi dan Kemasyarakatan Nasional, LIPI.
- Hock, H. H., & Joseph, B. D. (2009). *Language History, Language Change, and Language Relationship: An Introduction to Historical and Comparative Linguistics*. Mouton de Gruyter.
- Holton, G. (2006). The Relational Noun Marker in Tobelo (Northeast Halmaheran). In *Tenth International Conference on Austronesian Linguistics Linguistic Society of the Philippines 2006*. Linguistic Society of the Philippines and SIL International.
- Holton, G., & Klamer, M. (2017). The Papuan Languages of East Nusantara and the Bird's Head. In *The Languages and Linguistics of the New Guinea Area: A Comprehensive Guide* (pp. 569–640). De Gruyter. <https://doi.org/10.1515/9783110295252-005>
- Keraf, G. (1996). *Linguistik Bandingan Historis*. Gramedia Pustaka Utama.
- Klamer, M., & Holton, G. (2018). The Papuan Languages of East Nusantara and the Bird's Head. In *The Languages and Linguistics of the New Guinea Area: A Comprehensive Guide* (pp. 569–640). De Gruyter. <https://doi.org/10.1515/9783110295252-005>
- Ladefoged, P., & Maddieson, I. (1996). *The Sounds of the World's Languages*. Blackwell Publisher.
- Lekawael, R. F. J. (2013). Grammatical Relations of Pronominal Prefixes in Tobelo Language. *Jurnal Tahuri*.
- Lodge, K. (2009). *Fundamental Concepts in Phonology: Sameness and Difference*. Edinburgh University Press.
- Lynch, John. (2020). Chapter 5 Sound Systems. In *Pacific Languages* (pp. 75–99). University of Hawaii Press. <https://doi.org/10.1515/9780824842581-009>
- Maddieson, I. (2005). Diversity of vowel systems. *The Journal of the Acoustical Society of America*, 117(4\_Supplement), 2398–2398. <https://doi.org/10.1121/1.4785970>
- Maddieson, I. (2013). Vowel Quality Inventories. In M. S. Dryer & M. (eds.) Haspelmath (Eds.), *WALS Online (v2020.4) [Data set]*. Zenodo. <https://doi.org/https://doi.org/10.5281/zenodo.13950591>
- Mahmud, A. F., & Tamalene, M. N. (2025). Language Identity of Vulnerable Group: Tracing the Origins of Tobelo Dalam Tribe from a Linguistic Perspective. *Lire Journal (Journal of Linguistics and Literature)*, 9(2), 326–341. <https://doi.org/10.33019/lire.v9i2.455>
- Miceli, L., & Round, E. (2022). Where have all the sound changes gone? Examining the scarcity of evidence for regular sound change in Australian languages. *Linguistics Vanguard*, 8(s5), 509–518. <https://doi.org/10.1515/lingvan-2021-0094>
- O'Brien, J. (2012). *An Experimental Approach to Debuccalization and Supplementary Gestures* [University of California Santa Cruz]. <https://escholarship.org/uc/item/1cm694ff>
- Ohala, J. J. (2017). Phonetics and Historical Phonology. In *The Encyclopedic Dictionary of Applied Linguistics: A Handbook for Language Teaching* (pp. 667–686). Blackwell Publishing Ltd. <https://doi.org/10.1002/9781405166201.ch22>
- Putrohari, R. D., Simanjunta, T., Sudoyo, H., Lauder, M. R., Lauder, A. F., Neonbasu, P. G., Poelinggimang, E. L., Usmany, D. P., Yapsenang, Y., & Probonegoro, N. K. (2015). *Diaspora*

*Melanesia Di Nusantara*. Direktorat Sejarah dan Nilai Budaya, Direktorat Jenderal Kebudayaan, Kementerian Pendidikan dan Kebudayaan.

Rugebregt, E. J. F., Kakerisa, O., Pentury, J., Seumahu, C. C., & Anaotta, J. J. (1986). *Struktur Bahasa Tobelo*.

Schapper, A., & Zielenbach, M. (2025). Reframing Pre-modern Language Contact through Trade in Eastern Indonesia: Javanese Linguistic Influence in the Moluccas. *Journal of the Southeast Asian Linguistics Society JSEALS*, 18, 1–22. <http://hdl.handle.net/52542>

van Staden, M. (2006). The body and its parts in Tidore, a Papuan language of Eastern Indonesia. *Language Sciences*, 28(2–3), 323–343. <https://doi.org/10.1016/j.langsci.2005.11.013>

Voorhoeve, C. L. (1988). The Languages of the North Halmaheran Stock. *New Guinea Linguistics*, 181–209. <https://doi.org/10.15144/PL-A76.181>

Wada, Y. (1980). Correspondence of Consonants in and the Conservation of Archaic North Sounds Halmahera Languages in Galela. In *The Galela of Halmahera: A preliminary survey* (Number 7). National Museum of Ethnology.