

Nominalization Patterns in Tombatu Language: A Generative Morphological Analysis of Affixal Noun Formation

Verra Ellen Manangkot¹, Jola Kristiani Liuw¹, Lucky Herliawan Yanuarsyah Amalputra²,
Metty Ribka Manila¹, & Fergina Lengkoan¹

¹ Universitas Negeri Manado, Tondano, Indonesia

² Universitas Pendidikan Indonesia, Bandung, Indonesia

verramanangkot@unima.ac.id

ARTICLE HISTORY

Received : 2026-04-08

Revised : 2026-07-13

Accepted : 2026-07-17

KEYWORDS

Nominalization

Tombatu language

Generative morphology

Austronesian linguistics

Minahasan languages

Morphological rules



ABSTRACT

Regional languages preserve complex grammatical systems that reveal how communities organize experience, identity, and cultural knowledge. However, nominalization in Tombatu, an underdescribed Austronesian language of Minahasa, has not been systematically examined through an explicit rule-based morphological framework. This study addresses this gap by investigating the structural patterns, semantic functions, and generative mechanisms of affixal noun formation in Tombatu. Employing a descriptive-taxonomic design, data were collected over six months in Tombatu Village from proficient speakers through naturally occurring speech, elicitation, recording, note-taking, and informant validation. The final dataset found unique nominalized forms. These forms were classified into three major processes comprising 11 noun-forming prefixes, one noun-forming suffix, and 14 affix combinations, yielding 26 identified affixal configurations. The analysis applied Word Combination Rules, Derivation Rules, the Boundary Insertion Convention, and relevant morphological constraints. The findings show that prefixation exhibits the greatest structural and semantic variation in the dataset, producing agentive, person-denoting, instrumental, object-related, and result-related nouns. For example, /tapə-/ combines with /luku?/ 'to drink' to form /tapəluku?/ [tapəluku?] 'drinker'. Suffixation is more restricted, with /-an/ forming locative nouns, as in /tawoi/ 'to work' plus /-an/, which produces /tawoian/ [tawoian] 'workplace'. Complex affix combinations form nouns denoting places, processes, objects, states, results, and entities. This study provides the first systematic generative account of Tombatu nominalization, expands empirical evidence for Austronesian word formation, strengthens regional-language documentation, and offers a linguistic basis for language maintenance and potential contrastive morphological activities in multilingual education.

1. Introduction

Every language carries the memory, identity, and worldview of the community that speaks it. Through its grammatical system, speakers classify people, actions, places, instruments, objects, events, and culturally meaningful experiences. When a regional language is used less frequently, the loss involves not only vocabulary but also the grammatical knowledge through which local meanings are constructed and transmitted. Tombatu, an Austronesian language spoken in the Minahasa region of North Sulawesi,

represents this linguistic and cultural heritage. Field observations show that Tombatu is still used in families, traditional markets, and community activities. However, fluency is mainly retained by speakers over 50, while younger speakers increasingly mix Tombatu with Manado Malay (Gosal et al., 2023). This shift requires urgent documentation, as over 70% of eastern Indonesian languages are shifting toward Malay or Indonesian, while eight Maluku languages with fewer than fifty speakers have become extinct or nearly extinct (Ulhaq et al., 2022).

Morphology, particularly nominalization, provides an important entry point for understanding the structure of Tombatu. Its word-formation system relies strongly on affixation, reduplication, and multiple verbal affix patterns, demonstrating the central role of bound morphemes in shaping lexical categories and grammatical meanings (Gosal et al., 2023). This characteristic corresponds with descriptions of Austronesian languages as having relatively transparent morpheme boundaries, extensive affixation, and a strong tendency toward prefixing morphology (Klamer, 2020;). Nominalization is especially significant because Austronesian affixes often perform, or have historically developed, noun-forming functions alongside other morphosyntactic roles (Beguš et al., 2023). Dedicated affixal constructions may also form event-denoting and participant-denoting nouns that occupy central syntactic positions (Jufrizal et al., 2024). In Tombatu, nominalization transforms actions, qualities, and states into nouns denoting agents, places, instruments, objects, processes, and results (Beguš et al., 2023; Gosal et al., 2023; Jufrizal et al., 2024). For example, /luku?/ [luku?] ‘to drink’ combines with /tapə-/ to form /tapəluku?/ [tapəluku?] ‘drinker’ and with /pa- -an/ to produce /paluku?an/ [paluku?an] ‘drinking place’. These forms demonstrate that Tombatu nominalization follows systematic structural and semantic patterns.

Existing linguistic scholarship provides important foundations for analysing these processes but has not adequately explained the Tombatu system. Structural and descriptive traditions regard language as an organized system of forms, meanings, functions, grammatical categories, and morphemic relationships (Saussure, 1973; Hockett, 1965; Martinet, 1980; Lyons, 1995; Robins, 1992; Pike, 1992; Verhaar, 2001; Nida, 1970; Kridalaksana, 1984, 1992). Derivational morphology views nominalization as a key process that converts verbal and adjectival bases into nouns through affixation and related morphemes (Escarza, 2023; Jufrizal & Refnita, 2021). Rule-based approaches explain word formation through structured operations, morphosyntactic compositionality, and form–meaning relations shaped by productivity, category selection, lexical restrictions, and semantic interaction (Brook & Blamire, 2023; Luuk, 2019; Marneffe et al., 2021; Naudé & Miller-Naudé, 2023). These frameworks support Tombatu analysis, while Universal Dependencies enables comparison with broader linguistic typologies and corpora (Arista, 2022; Marneffe et al., 2021).

The gap becomes clearer when Tombatu is considered within Indonesian, Minahasan, and Austronesian research. Earlier studies described important aspects of Indonesian and regional

phonology, grammar, and morphology (Dardjowidjojo, 1994; Lapoliwa, 1981; Samsuri, 1987;), while Sudaryanto (1993) established procedures for linguistic observation, listening, recording, and note-taking. Sudiroatmadja (1992) demonstrated the applicability of generative morphology to Javanese word formation. Within Minahasan linguistics, Sneddon (1978) examined Proto-Minahasa historically and comparatively, whereas Kalangi (1984) described selected features of Tombatu phonology and morphology. Nevertheless, these studies did not specifically explain how Tombatu noun-forming morphemes attach to roots, stems, bases, and existing words or how they produce particular semantic functions. Recent Austronesian studies have further shown that nominalization is relevant to alignment, historical reconstruction, diathesis, and affixal systems. Aldridge (2021) relates nominal structures to alignment change, Ross and Zeitoun (2024) emphasize Formosan evidence in reconstructing Proto-Austronesian morphology, and Kaufman (2026) discusses nominalization-derived structures in Austronesian diathesis. Yet these studies mainly address better-documented languages or broad comparative issues. Consequently, Tombatu remains underrepresented, particularly regarding noun-forming morphemes, affixation patterns, base selection, semantic outcomes, morphological restrictions, and rule-governed derivations.

This study addresses that gap by examining Tombatu nominalization through a generative morphological framework. Unlike previous research that discusses Minahasan languages broadly or describes general aspects of Tombatu phonology and morphology, this study focuses specifically on how Tombatu nouns are formed, which morphemes function as nominalizers, which lexical bases they select, and what meanings they produce. Its novelty lies in providing an explicit rule-based account of prefixation, suffixation, and complex affix combinations rather than treating noun-forming expressions as an unstructured lexical list. Word Combination Rules, Derivation Rules, Inflection Rules, the Boundary Insertion Convention, and relevant morphological constraints are applied to explain both acceptable formations and restrictions on affix attachment.

The study therefore contributes descriptively by documenting Tombatu nominalizers and formation patterns, theoretically by modelling their structures and semantic outcomes, and comparatively by incorporating Tombatu evidence into broader discussions of Austronesian morphology, derivation, lexical-category change, and linguistic diversity. This study is theoretically and practically significant because it integrates morphological analysis with the

documentation of an underdescribed language. Effective documentation should extend beyond lexical collections and recordings to include grammatical structure and theoretical interpretation. Kim (2020) emphasizes that documenting allomorphy and inflection requires extensive primary data and flexible, corpus-informed methods. Hemmings (2020) also stresses meta-documentation and accessible outputs for researchers and communities. These perspectives support a theory-informed framework for documenting Tombatu morphology. Accordingly, this study describes the rules, characteristics, patterns, and systems of Tombatu nominalization and identifies its noun-forming morphemes. It addresses three questions: (1) What are the characteristics and rules of Tombatu nominalization? (2) What patterns and systems characterize it? and (3) Which morphemes function as noun-forming elements? The findings clarify relationships among morphological form, lexical category, semantic function, and structural restrictions while supporting language maintenance, local-language instruction, material development, and comparative research on Minahasan and Austronesian languages.

Overall, nominalization is central to Tombatu because it links grammatical structure, lexical meaning, and culturally embedded knowledge. Through noun formation, speakers encode performers, social identities, instruments, places, actions, objects, states, and results relevant to their communicative environment. Rather than merely listing affixes, this study explains how noun forming morphemes combine with different bases, develop specific meanings, and permit or restrict particular formations. The findings show that Tombatu offers valuable empirical and theoretical insights into derivational morphology, affix ordering, base selection, semantic asymmetry, and lexical category formation. Thus, Tombatu is both a language requiring preservation and an important source of linguistic knowledge.

In English Language Teaching, these findings support multilingual and culturally responsive instruction. Comparing nominalization in Tombatu, Indonesian, and English can strengthen learners' morphological and metalinguistic awareness of affixation, derivation, category change, agentive and locative meanings, and relationships between form and meaning. Learners' local, national, and international linguistic resources should therefore be treated as pedagogical assets rather than deficiencies (Goodman and Tastanbek, 2021; Prilutskaya, 2021; Cenoz and Gorter, 2025; Yun et al., 2025). Integrating Tombatu examples into ELT activities may promote cross linguistic noticing, vocabulary development, morphological analysis, understanding of English word formation, and affirmation of learners'

identities. This study consequently contributes to Tombatu documentation, morphological theory, multilingual competence, local knowledge, and cultural continuity.

2. Literature Review

2.1 Morphology, Word Formation, and Nominalization

Morphology is a central field of linguistic inquiry concerned with the internal structure of words, the combination of morphemes, and the formation of new lexical items. Words are therefore not merely isolated vocabulary units but structured grammatical forms composed of meaningful elements within an organized system of signs, forms, and functions (Saussure, 1973; Hockett, 1965; Martinet, 1980; Lyons, 1995; Robins, 1992; Pike, 1992; Verhaar, 2001). Morphological analysis examines roots, stems, bases, affixes, grammatical categories, word-formation rules, and semantic relations. In languages with productive affixation, word formation is particularly important because the combination of morphemes may create new lexical meanings and grammatical functions. Nida (1970) emphasizes that morphological analysis must identify both how morphemes combine and how these combinations affect meaning, while Kridalaksana (1984, 1992) relates word formation to grammatical category, lexical function, and semantic development. Accordingly, morphology cannot be separated from semantics because the formation of a word involves not only a formal modification of its base but also the creation or alteration of its meaning.

Nominalization is a major derivational process through which nouns are formed from verbs, adjectives, roots, stems, bases, or existing words. It enables actions to be conceptualized as entities, processes as concepts, qualities as abstract nouns, and activities as agents, places, instruments, objects, or results. Thus, nominalization functions not only as a morphological operation but also as a grammatical and cognitive mechanism for organizing human experience. Its significance is especially evident in agglutinative languages, where individual affixes generally carry identifiable grammatical and semantic functions (Keraf, 1990, 1991). Tombatu displays such characteristics through prefixes, suffixes, and affix combinations that systematically derive nouns from verbal and other lexical bases. These processes are predominantly derivational because they create new lexical items or change the semantic function and, in some cases, the grammatical category of the base, whereas inflection primarily expresses grammatical information without altering the core lexical category (Scalise, 1986; Katamba, 1993).

Recent Austronesian studies have further demonstrated the importance of nominalization for understanding morphology, syntax, semantics, alignment, and historical change (Aldridge, 2021; Ross and Zeitoun, 2024; Kaufman, 2026). Nevertheless, most contemporary studies address well-documented Austronesian languages or broad comparative questions, leaving the nominalization systems of underdescribed languages such as Tombatu insufficiently examined.

2.2 Generative Morphology as the Theoretical Framework

This study adopts generative morphology, which originated within the broader framework of Generative Grammar. Chomsky introduced the generative approach in *Syntactic Structures* (1957), expanded it in *Aspects of the Theory of Syntax* (2014). Its central assumption is that language constitutes a rule-governed system and that speakers possess internal linguistic competence enabling them to produce, interpret, and evaluate linguistic expressions. This competence is not restricted to memorized forms but includes productive knowledge of the rules through which new structures are generated (Katamba, 1993). Applied to morphology, this principle explains how speakers create and understand previously unencountered words and judge whether a derived form is structurally acceptable and semantically appropriate. Generative morphology locates word formation within the lexicon and describes it through explicit rules (Sudiroatmadja, 1992). Therefore, word-formation rules must operate on actual words recognized by speakers rather than on hypothetical forms, making the framework particularly appropriate for analysing authentic Tombatu data used by its speech community.

Within this framework, Tombatu nominalization can be described through Word Combination Rules, Derivation Rules, Inflection Rules, and the Boundary Insertion Convention (Scalise, 1986; Katamba, 1993). Word Combination Rules represent the linear relationship between morphological elements as $[\] X, [\] Y \rightarrow [[\] X Q [\] Y] Z$, where Q marks an internal boundary, while the Boundary Insertion Convention represents the resulting structure as $[\# [\] X Q [\] Y \#] Z$ (Scalise, 1986). Derivation Rules are especially relevant because Tombatu noun formation involves prefixes, suffixes, and combined affixes attached to roots, stems, bases, and existing words, as represented in patterns such as $[pre + [word]]$, $[pre + [stem]]$, $[[word] + suf]$, $[[stem] + suf]$, $[[pre + [word]] \# suf]$, and $[[pre + [stem]] \# suf]$ (Scalise, 1986). Inflection is positioned at the final stage because it does not normally alter lexical category, consistent

with the Extended Ordering Hypothesis, which places derivation before compounding and inflection (Scalise, 1986;). The analysis must also account for constraints that prevent unrestricted rule application, including the Unitary Base Hypothesis, Binary Branching Hypothesis, No Phrase Constraint, Adjacency Condition, and Atomic Condition (Aronoff, 1976). These principles explain why particular Tombatu affixes select specific bases, why certain combinations remain productive, and why other theoretically possible formations are blocked. More broadly, the systematic documentation of such rules contributes to the study of linguistic diversity and language endangerment by demonstrating how local morphological systems reflect wider social, cognitive, and communicative patterns (Lupyan & Dale, 2021; Bromham, 2023).

2.3 Previous Studies, Research Gap, Novelty, and Implications

Cross-linguistic research demonstrates the importance of systematic phonological, grammatical, and morphological analysis and confirms the productivity of derivational morphology in Austronesian languages (Beguš & Dąbkowski, 2025; Engelenhoven, 2021; Gosal et al., 2023; Hidayati et al., 2021; Klamer, 2020; Krohn, 2021; Suhandano, 2021). In Tombatu, affixation, reduplication, and inflection distinguish derivational from inflectional processes (Gosal et al., 2023). Related languages also form agentive, instrumental, locative, and other deverbal nouns through productive affixation (Engelenhoven, 2021; Hidayati et al., 2021; Klamer, 2020). Research on Austronesian voice, Indonesian affixation, and regional lexical-morphological parallels further supports a generative-morphological analysis of Tombatu (Beguš & Dąbkowski, 2025; Krohn, 2021; Suhandano, 2021). However, direct studies of Tombatu and related Minahasan languages remain limited. Although Sneddon (1978) examined Proto-Minahasa from a historical-comparative perspective and Kalangi (1984) described aspects of Tombatu phonology and morphology, neither study provided a focused, rule-based analysis of nominalization. Recent scholarship has reinforced the theoretical importance of nominalization in Austronesian alignment, diathesis, and historical morphology, but has not specifically addressed Tombatu (Aldridge, 2021; Ross and Zeitoun, 2024; Kaufman, 2026).

The novelty of the present study lies in its systematic analysis of Tombatu nominalization through a generative morphological framework. Unlike earlier research that discusses Minahasan or regional linguistic structures more broadly, this study identifies Tombatu noun-forming prefixes, suffixes, and affix combinations and explains how they attach

to roots, stems, bases, and words to produce specific nominal meanings. By applying Word Combination, Derivation, and Inflection Rules, the Boundary Insertion Convention, and morphological constraints, this study explains productive formations and structural restrictions in Tombatu (Gosal et al., 2023). It contributes to generative morphology and Austronesian word-formation research while documenting an underdescribed language.

Practically, the findings may support language preservation, teaching, dictionary and instructional-material development, and metalinguistic awareness (Salem & Mesra, 2020). Comparable Austronesian patterns of affixation and morphosyntactic variation further strengthen the study's cross-linguistic and methodological relevance (Klamer, 2020; Lumbantobing & Situngkir, 2026; Schapper, 2020; Suhandano, 2021; Wu, 2022). They may also contribute to English Language Teaching by enabling learners to compare nominalization patterns in Tombatu, Indonesian, and English. Such cross-linguistic comparison is consistent with multilingual education and translanguaging perspectives, which recognize learners' complete linguistic repertoires as resources for learning rather than as obstacles (Goodman and Tastanbek, 2021; Prilutskaya, 2021; Cenoz and Gorter, 2025; Yun et al., 2025).

3. Method

This study employed a descriptive-taxonomic method to examine the nominalization system of the Tombatu language. The descriptive orientation was used because the study aimed to describe linguistic facts as they are used by speakers, particularly the forms, meanings, and functions of noun formation in Tombatu. Meanwhile, the taxonomic orientation was applied to classify the data into systematic morphological categories, including prefixation, suffixation, and affix combinations. Through this design, the study did not merely list Tombatu noun forms, but organized them according to their morphological patterns, semantic functions, and word-formation rules.

The study was grounded in a generative morphological framework because Tombatu, as an agglutinative language, shows productive affixation processes that require rule-based explanation. This framework was considered appropriate because the research questions focus on the characteristics, rules, patterns, systems, and morphemes involved in Tombatu nominalization. The analysis followed the principles of word formation proposed by Aronoff, particularly the use of Word Combination Rules (WCR), Derivation Rules (DR), and Inflection Rules (IR). These rules were used to explain how Tombatu

roots, stems, bases, and words combine with bound morphemes to produce nouns.

In this study, nominalization was treated as a derivational morphological process because it produces new lexical items and may change the category or semantic function of the base. The method was therefore designed to identify not only the surface forms of affixes, but also the morphological mechanisms that allow those affixes to form nouns. This approach enabled the study to explain Tombatu noun formation as a systematic process governed by rules, constraints, and speaker acceptability.

3.1 Research Site and Data Sources

The research was conducted in selected Tombatu Village over a period of six months. This location was selected because Tombatu I is one of the speech communities where the Tombatu language is still used in daily communication, especially by older and active speakers. The setting was relevant to the purpose of the study because the research required authentic linguistic data from actual Tombatu usage rather than artificial or isolated examples. By collecting data in the speech community, the study was able to obtain forms of nominalization that reflect the living use of the language.

The data of this study consisted of Tombatu roots, stems, bases, words, and affixed forms that indicate noun formation. The main focus was on linguistic forms that show nominalization through prefixes, suffixes, and affix combinations. These data included noun forms derived from verbs, adjectives, roots, stems, bases, or existing words. Each form was examined in relation to its base category, morphological structure, and semantic function, such as agent, place, instrument, object, result, or entity made into something.

The data were obtained from spoken language use, informant consultation, and the researcher's linguistic competence as a native speaker. The researcher's native-speaker competence was used to support the preliminary identification of Tombatu forms, while informant consultation was used to verify the acceptability, meaning, and usage of the data. This step was important because, in generative morphology, the result of word formation must be recognized by speakers of the language concerned. Therefore, the study combined researcher intuition and speaker validation to ensure that the analyzed forms were linguistically acceptable in Tombatu.

3.2 Data Collection Techniques

The data were collected through several linguistic data collection techniques, namely tapping, skilled listening, recording, and note-taking. These

techniques followed the procedures proposed by [Sudaryanto \(1993\)](#), who emphasizes systematic observation and documentation in linguistic research. The tapping technique was used to obtain naturally occurring language data from speakers' communication. This technique helped the researcher identify forms of Tombatu nominalization as they appeared in actual use.

The skilled listening technique was applied to identify and select relevant linguistic forms, especially words that showed noun formation through affixation. During this process, the researcher listened carefully to spoken Tombatu data and identified roots, stems, bases, and affixed forms that functioned as nouns. Recording was used to preserve spoken data accurately, while note-taking was used to document important forms, meanings, contexts, and morphological patterns found during data collection. These techniques were useful for ensuring that the data could be reviewed and analyzed systematically.

3.4 Data Analysis Procedures

The data were analyzed using a rule-based morphological analysis within the framework of generative morphology. The analysis began by identifying each collected form according to its base category, such as root, stem, base, or word. After that, the forms were classified according to the type of nominalization process involved: prefixation, suffixation, or affix combination. This classification was necessary to determine the major patterns of Tombatu noun formation and to organize the findings according to the research questions.

After the classification stage, each nominalized form was analyzed to determine its morphological structure and semantic function. The analysis applied Word Combination Rules (WCR), Derivation Rules (DR), and Inflection Rules (IR) as proposed by Aronoff. WCR was used to explain the linear relationship between linguistic elements in complex word formation. The structure of WCR followed the rule. $[] X, [] Y \rightarrow [[] X Q [] Y] Z$, in which Q functions as an internal boundary. This rule was used to examine how Tombatu morphological elements combine in structured noun formation.

The internal structure of complex forms was further examined through the Boundary Insertion Convention (BIC) proposed by Boij, as cited in [Scalise \(1986:78\)](#), in the form $[# [] X Q [] Y \#] Z$. This convention was used to mark morphological boundaries within complex Tombatu noun forms. By applying BIC, the analysis could show where affixes attach to roots, stems, bases, or words, and how internal boundaries contribute to the formation of nouns.

The Derivation Rules (DR) were central to the analysis because Tombatu nominalization mainly involves derivational processes. The study followed the derivational structures ([Scalise, 1986](#)), including forms such as $[pre + [word]]$, $[pre + [stem]]$, $[[word] + suf]$, $[[stem] + suf]$, $[[pre + [word]] \# suf]$, and $[[pre + [stem]] \# suf]$. These structures were used to analyze nouns formed through prefixes, suffixes, and affix combinations. Meanwhile, Inflection Rules (IR) were placed at the final stage because inflectional processes do not normally change the lexical category of the base. This order followed the Extended Ordering Hypothesis proposed by [Scalise \(1986\)](#), which places derivation at Level I and Level II, compounding at Level III, and inflection at Level IV.

To prevent the analysis from becoming too broad or unrestricted, the study also applied the morphological constraints proposed by Aronoff, namely the Unitary Base Hypothesis (UBH), Binary Branching Hypothesis (BBH), No Phrase Constraint (NPC), Adjacency Condition (AdC), and Atomic Condition (AC). These constraints were used to determine whether a noun-formation pattern was structurally acceptable within the Tombatu morphological system. The final stage of analysis involved examining possible readjustment rules to account for phonological or morphological changes that occurred after affixation. Through these procedures, the study produced a systematic description of the characteristics, rules, patterns, systems, and noun-forming morphemes in the Tombatu language.

4. Results

This section presents the findings on nominalization in the Tombatu language and explains the principal morphological processes involved in noun formation. The analysis is organized around three major processes identified in the data: prefixation, suffixation, and affix combination. Together, these processes demonstrate that nominalization in Tombatu is productive, systematic, and governed by recognizable semantic patterns.

The findings further show that Tombatu nouns can be derived from different morphological bases, including roots, stems, bases, and existing words, through the attachment of bound morphemes. These derivational processes produce nouns with a range of semantic functions, such as agents, places, instruments, objects, results, and entities that are made or transformed into something. Overall, the data reveal a structured relationship between affix selection, base form, and the meaning of the resulting nominal expression.

The most dominant pattern found in the data is prefixation. Several prefixes function as noun-forming morphemes, including /tapə-/ , /ta-/ , /to-/ , /iwa-/ , /tapa-/ , /i-/ , /ko-/ , /do-/ , /so-/ , /tapo-/ , and /bɛwa-/. The second pattern is suffixation, represented by the suffix /-an/, which mainly forms locative nouns. The third pattern is affix combination, in which more than one bound morpheme attaches to a base to form a noun. These findings show that Tombatu nominalization is not limited to a single morphological pattern. Rather, the process operates through a layered morphological system in which each affix contributes a distinct structural role and semantic function to the formation of the derived word.

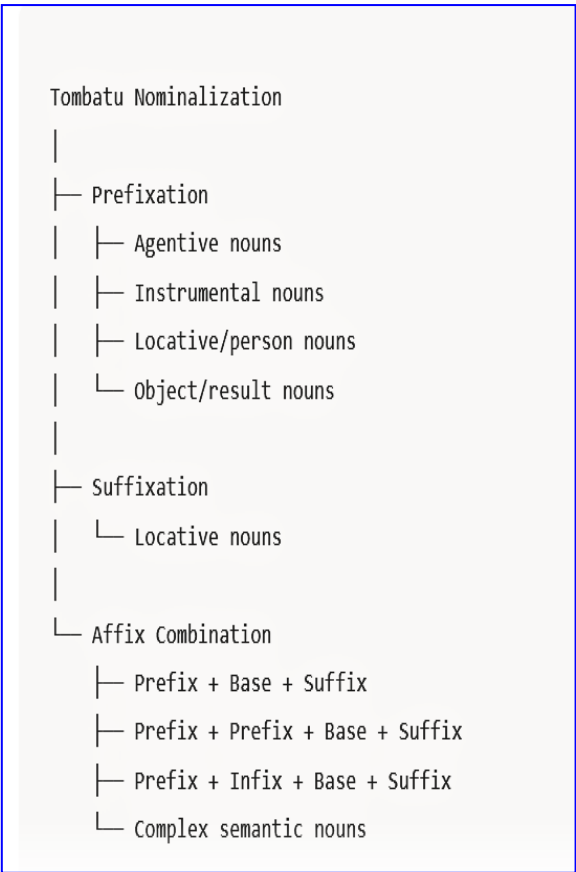


Figure 1. General System of Tombatu Nominalization

The figure illustrates that nominalization in Tombatu follows a hierarchical morphological structure. Among the identified processes, prefixation is the most diverse because it involves several noun-forming prefixes that contribute different semantic meanings to the derived forms. In contrast, suffixation is more restricted, as only one productive nominal suffix, /-an/, was identified in the data. Affix combination constitutes a structurally more complex morphological process because it enables two or more bound

morphemes to operate simultaneously within a single noun-forming construction. Through this process, multiple affixes interact with the base to produce a derived noun with a more complex structural pattern and semantic function.

4.1 Formation of Nouns through Prefixation

The first major finding is that Tombatu forms nouns by attaching prefixes to roots, stems, bases, or words. The basic rule of prefixal nominalization is as follows:

$$X \rightarrow \text{Prefix} + X$$

where **X** represents a root, stem, base, or word. In a more explicit morphological structure, the rule may be represented as:

$$X \rightarrow [\text{Prefix} + [\text{Base}]X]N$$

This rule shows that the prefix attaches to a base and produces a noun. The resulting form may retain part of the meaning of the base while adding a new nominal function. In Tombatu, prefixation is highly productive because different prefixes produce different types of nouns.

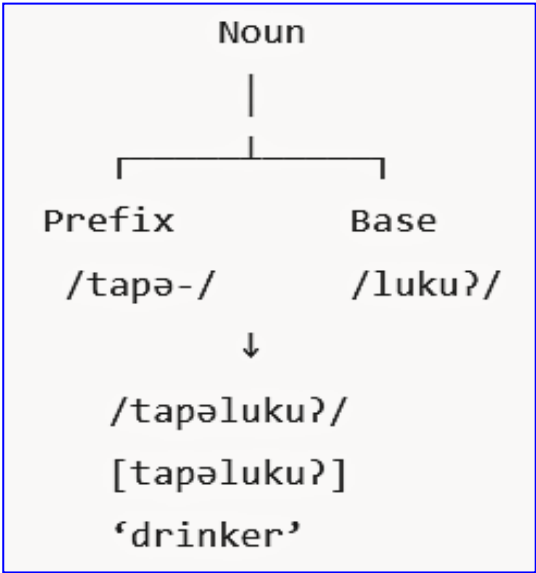


Figure 2. Prefixal Nominalization Pattern

This morphological structure demonstrates that the verbal base /luku?/, meaning ‘to drink’, is transformed into a noun through the attachment of the prefix /tapə-/. The resulting derived form, /tapəluku?/ [tapəluku?], means ‘drinker’. This derivational process is not merely a formal change in word class, as the prefix also contributes a clear agentive meaning. In other words, /tapə-/ identifies the person who performs the action expressed by the base verb.

Table 4.1 Noun-Forming Prefixes in Tombatu

No.	Prefix	Morphological pattern	Example	IPA form	Meaning of derived noun	Semantic function
1	/tapə-/	/tapə-/ + base	/tapəluku?/	[tapəluku?]	drinker	Agent
2	/ta-/	/ta-/ + base	/taluku?/	[taluku?]	person who drinks much	Agent
3	/to-/	/to-/ + base	/topandɛi/	[topandɛi]	clever person	Person/agent
4	/iwa-/	/iwa-/ + base	/iwahaliɖ/	[iwahaliɖ]	base/tool-related noun	Instrument/object
5	/tapa-/	/tapa-/ + base	/tapakelod/	[tapakelod]	sower	Agent
6	/i-/	/i-/ + base	/itiho?/	[itiho?]	something used for catching	Instrument/object
7	/ko-/	/ko-/ + base	/kokaiɦ/	[kokaiɦ]	rower/rowing tool	Agent/instrument
8	/do-/	/do-/ + base	/dodali/	[dodali]	something made reliable	Result/object
9	/so-/	/so-/ + base	/sosahul/	[sosahul]	spear tool	Instrument
10	/tapo-/	/tapo-/ + base	/tapolɛŋgo?/	[tapolɛŋgo?]	traitor	Agent/person
11	/bɛwa-/	/bɛwa-/ + base	/bɛwatiho?/	[bɛwatiho?]	catching tool	Instrument/object

Table 4.1 shows that Tombatu prefixal nominalization has a strong semantic orientation. The prefixes do not function as empty formal markers. Instead, each prefix contributes meaning to the derived noun. Some prefixes mainly produce agentive nouns, while others produce nouns referring to tools, objects, places, or results. This finding indicates that Tombatu prefixes are both structurally productive and semantically distinctive.

The prefix /tapə-/ forms nouns referring to agents or performers. It may carry either a positive or negative meaning depending on the base and context. For example, /tapə-/ + /luku?/ ‘to drink’ produces /tapəluku?/ [tapəluku?] ‘drinker’. The same prefix may attach to /tawoi/ ‘work’ to form /tapənawoi/ [tapənawoi], meaning ‘worker’ or ‘mystical healer’. This shows that /tapə-/ does not only mark a person who performs an action, but may also encode culturally specific roles.

The prefix /ta-/ also forms agentive nouns, but it tends to produce a more specific or evaluative meaning. For instance, /ta-/ + /luku?/ produces /taluku?/ [taluku?], meaning a person who drinks much. Compared with /tapəluku?/ [tapəluku?] ‘drinker’, the form /taluku?/ [taluku?] suggests a stronger habitual or characterizing meaning. This distinction shows that Tombatu has more than one

strategy for forming agentive nouns and that different prefixes may encode different shades of agency.

The prefix /to-/ produces nouns referring to persons or places. For example, /to-/ + /pandɛi/ ‘clever’ produces /topandɛi/ [topandɛi], meaning ‘clever person’. This shows that /to-/ may nominalize an adjectival base by turning a quality into a person-denoting noun. The derived noun does not simply describe the quality of being clever; it identifies a person associated with that quality. This pattern indicates that Tombatu nominalization can transform qualities into participant nouns.

The prefixes /iwa-/ , /i-/ , /ko-/ , /so-/ , and /bɛwa-/ are strongly associated with instrumental or object-related meanings. For example, /so-/ + /sahul/ produces /sosahul/ [sosahul], meaning ‘spear tool’, while /bɛwa-/ + /tiho?/ produces /bɛwatiho?/ [bɛwatiho?], meaning ‘catching tool’. These forms show that Tombatu nominalization is not limited to human agents. It also forms nouns referring to instruments and objects used for particular actions. This is important because it shows how Tombatu morphology encodes practical and material aspects of experience.

The prefix /tapo-/ forms nouns referring to persons characterized by an action, quality, or behavior. For example, /tapo-/ + /lɛŋgo?/ produces

/tapolɛŋgoʔ/ [tapolɛŋgoʔ], meaning ‘traitor’. This form suggests that /tapo-/ may express a person associated with a negative or socially marked quality. The meaning of the derived noun is therefore not only grammatical, but also evaluative. This finding shows that some Tombatu nominalizing prefixes carry social or moral interpretation.

Agent / Person	: /tapa-/ , /ta-/ , /to-/ , /tapa-/ , /tapo-/
Instrument / Object	: /iwa-/ , /i-/ , /ko-/ , /so-/ , /bewa-/
Result / Made entity	: /do-/ , /bewa-/
Place-related noun	: /to-/

Figure 3. Semantic Distribution of Prefixal Nominalization

Figure 3 illustrates that Tombatu prefixes are systematically distributed across several semantic domains. Among the prefixal formations identified in the data, agentive and person-denoting nouns occur most frequently, followed by nouns referring to instruments and objects. This distribution indicates

that prefixal nominalization in Tombatu is particularly productive in naming individuals according to their roles, activities, or characteristics, as well as in identifying tools and objects associated with particular actions. The pattern further suggests that Tombatu nominalization is closely related to the speakers’ everyday need to categorize participants, occupations, instruments, and material entities within their social and cultural environment.

4.2 Formation of Nouns through Suffixation

The second major finding is that Tombatu forms nouns through suffixation. In the data, the productive noun-forming suffix identified is /-an/. This suffix is attached after a root, stem, base, or word and mainly forms nouns indicating place. The basic rule of suffixal nominalization is as follows:

$$X \rightarrow X + \text{Suffix}$$

More specifically, the rule may be represented as:

$$X \rightarrow [[\text{Base}]X + /-an/]N$$

This structure shows that the suffix /-an/ attaches to a base and produces a noun. Unlike prefixation, which shows several semantic functions, suffixation in Tombatu appears to be more restricted. The suffix /-an/ mainly produces locative nouns.

Table 4.2. Noun Formation with the Suffix /-an/

No.	Suffix	Morphological pattern	Base meaning	Derived form	IPA form	Meaning of derived noun	Semantic function
1	/-an/	/tawoi/ + /-an/	work	/tawoian/	[tawoian]	workplace	Place
2	/-an/	/bɛlɛt/ + /-an/	lying	/bɛlɛtan/	[bɛlɛtan]	lying place	Place

The data in Table 4.2 show that /-an/ has a clear locative function. The base /tawoi/ ‘work’ becomes /tawoian/ [tawoian] ‘workplace’, while /bɛlɛt/ ‘lying’ becomes /bɛlɛtan/ [bɛlɛtan] ‘lying place’. In both cases, the suffix converts an activity or state into a place associated with that activity or state.

The interpretation of /-an/ as a locative nominalizer is supported by the consistency of its meaning. Unlike some prefixes that may express agentive, instrumental, or evaluative meanings, /-an/ is semantically focused. It does not generally form agentive nouns or instrument nouns in the data. Instead, it identifies a location where an action or state occurs. This makes /-an/ a specialized nominalizing suffix in Tombatu.

Another significant finding is that the data provide no evidence of productive double suffixation in Tombatu noun formation. This indicates that suffixal nominalization is structurally less complex than prefixation and affix combination, both of which allow a wider range of morphological configurations. The absence of double suffixation suggests that the suffixal system is relatively restricted but remains stable and functionally consistent. Rather than expressing multiple semantic categories, suffixation performs a more specialized role within the broader nominalization system, particularly in deriving nouns that denote places or locations associated with specific activities, objects, or functions.

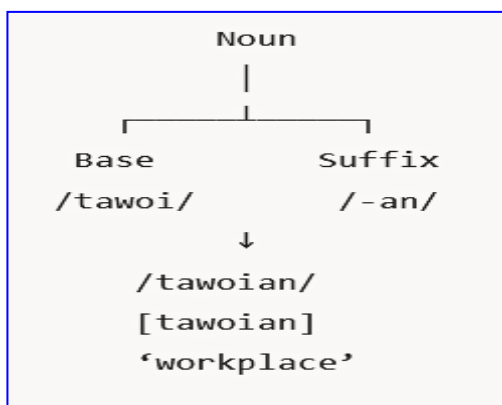


Figure 4. Suffixal Nominalization Pattern

Figure 4 illustrates that suffixation in Tombatu follows a right-attaching pattern. The base comes first, followed by the suffix /-an/. The result is a noun that refers to a place associated with the meaning of the base.

4.3 Formation of Nouns through Affix Combination

The third major finding is that Tombatu forms nouns through affix combinations. Affix combination refers to the use of more than one bound morpheme in one word-formation process. In Tombatu, these combinations may involve prefixes and suffixes, and in some cases, more complex structures involving more than one prefix or an infix. The general rule may be represented as follows:

$$X \rightarrow \text{Prefix} + X + \text{Suffix}$$

A more explicit structure is:

$$X \rightarrow [\text{Prefix} + [\text{Base}]X + \text{Suffix}]N$$

The affix combinations found in the data include /pa- -an/, /pa- -en/, /pewe- -an/, /pewe- -en/, /po- -an/, /ka- -in/, /ke- -an/, /ko- -an/, /ko- -en/, /ko- -ai/, /wa- -an/, /do- -an/, /be- -um- -an/, and /pah- -in- -an/.

Table 4.3 Affix Combinations in Tombatu Nominalization

No.	Affix combination	Structural pattern	Type of formation	General semantic function
1	/pa- -an/	Prefix + Base + Suffix	Circumfixal nominalization	Place/process noun
2	/pa- -en/	Prefix + Base + Suffix	Circumfixal nominalization	Object/result noun
3	/pewe- -an/	Prefix + Base + Suffix	Circumfixal nominalization	Place/result noun
4	/pewe- -en/	Prefix + Base + Suffix	Circumfixal nominalization	Object/result noun
5	/po- -an/	Prefix + Base + Suffix	Circumfixal nominalization	Place/process noun
6	/ka- -in/	Prefix + Base + Suffix	Circumfixal nominalization	Abstract/result noun
7	/ke- -an/	Prefix + Base + Suffix	Circumfixal nominalization	State/place noun
8	/ko- -an/	Prefix + Base + Suffix	Circumfixal nominalization	Result/place noun
9	/ko- -en/	Prefix + Base + Suffix	Circumfixal nominalization	Object/result noun
10	/ko- -ai/	Prefix + Base + Suffix	Circumfixal nominalization	Result/entity noun

No.	Affix combination	Structural pattern	Type of formation	General semantic function
11	/wa- -an/	Prefix + Base + Suffix	Circumfixal nominalization	Place/result noun
12	/do- -an/	Prefix + Base + Suffix	Circumfixal nominalization	Place/result noun
13	/be- -um- -an/	Prefix + Infix + Base + Suffix	Complex nominalization	Complex result noun
14	/pah- -in- -an/	Prefix + Infix + Base + Suffix	Complex nominalization	Complex result/place noun

Table 4.2 shows that affix combination is the most structurally complex type of Tombatu nominalization. Unlike simple prefixation or suffixation, affix combinations involve multiple morphological positions. The base may be

surrounded by a prefix and suffix, or it may receive additional internal modification through an infix. This pattern shows that Tombatu nominalization can operate through layered morphology.

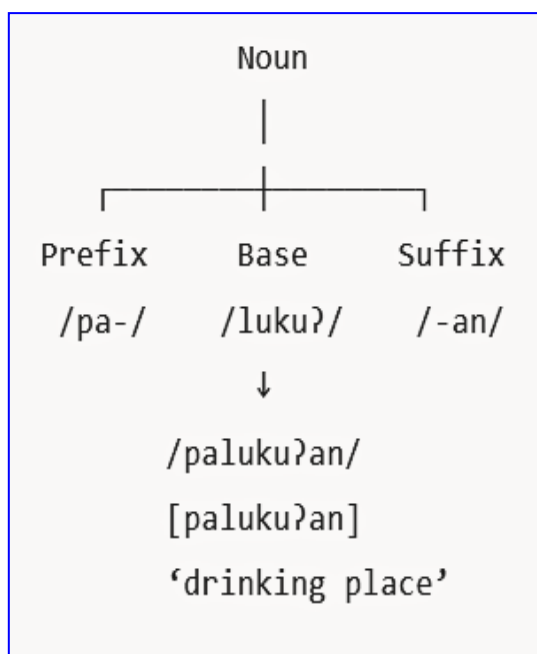


Figure 5. Affix Combination Pattern

Figure 5 illustrates the formation of /paluku?an/ [paluku?an] 'drinking place'. The base /luku?/ 'to drink' receives the prefix /pa-/ and the suffix /-an/. The resulting noun refers to a place associated with the action of drinking. This structure shows that the meaning of the derived noun emerges from the interaction between the base and both affixes.

Affix combinations play an important role in Tombatu nominalization because they broaden the semantic range of noun formation. Whereas prefixation commonly produces agentive or instrumental nouns and suffixation primarily forms locative nouns, combinations of affixes can generate nouns with more complex and specific meanings. These derived meanings may refer to a place, result, object, state, process, or entity. This pattern indicates that affix combination functions as a productive morphological strategy for forming nouns whose meanings require a more detailed semantic interpretation.

Another significant finding is that affix combination reflects the layered nature of Tombatu morphology. A derived noun may be formed not merely by attaching a single affix to a base, but by embedding the base within a broader morphological framework. This framework may consist of a prefix and a suffix or, in more complex formations, a prefix, an infix, and a suffix. Such patterns demonstrate that Tombatu word formation possesses an internal hierarchical organization. The affixes do not simply occur as a linear sequence of separate elements; rather, they interact systematically to form a structurally unified noun with a coherent meaning.

4.4 Summary of Key Findings

The results show that Tombatu nominalization operates through three major processes: prefixation, suffixation, and affix combination. Prefixation is the most diverse process because it involves eleven noun-forming prefixes and produces several semantic types of nouns. Suffixation is more limited because only one noun-forming suffix, /-an/, was identified, but this suffix has a clear and stable locative function. Affix combination is the most complex process because it involves multiple bound morphemes and produces nouns with layered meanings.

Table 4.4 Summary of Tombatu Nominalization Patterns

Nominalization process	Main form	Number of identified forms	Structural pattern	Main semantic function
Prefixation	/tapə-/ , /ta-/ , /to-/ , /iwa-/ , /tapa-/ , /i-/ , /ko-/ , /do-/ , /so-/ , /tapo-/ , /bɛwa-/	11 prefixes	Prefix + Base	Agent, person, instrument, object, result
Suffixation	/-an/	1 suffix	Base + Suffix	Place
Affix combination	/pa- -an/ , /pa- -en/ , /pewe- -an/ , /pewe- -en/ , /po- -an/ , /ka- -in/ , /ke- -an/ , /ko- -an/ , /ko- -en/ , /ko- -ai/ , /wa- -an/ , /do- -an/ , /be- -um- -an/ , /pah- -in- -an/	14 combinations	Prefix + Base + Suffix; Prefix + Infix + Base + Suffix	Place, process, object, result, state, entity

The results also reveal three original interpretations. First, Tombatu nominalization is semantically organized, not merely formally constructed. Each affix contributes a specific meaning to the derived noun. Second, Tombatu noun formation is asymmetrical: prefixation is semantically rich, suffixation is functionally specialized, and affix combination is structurally complex. Third, Tombatu nominalization reflects a close relationship between grammar and cultural experience. The language uses noun-forming morphology to identify people, tools, places, actions, and results that are meaningful in the speech community.

Overall, the findings indicate that Tombatu has a productive and systematic nominalization system. Nouns are formed through identifiable morphological rules, and the resulting forms are semantically interpretable. The evidence demonstrates that Tombatu nominalization is not simply a list of affixed words, but a structured system of word formation that reveals the grammatical complexity of the language.

5. Discussion

Three principal findings constitute the central contribution of this study. First, the dataset reveals three strategies of Tombatu nominalization, namely prefixation, suffixation, and affix combination. These strategies comprise 11 noun-forming prefixes, one noun-forming suffix, and 14 complex affixal configurations. Second, the three strategies exhibit a clear morphological and semantic asymmetry. Prefixation displays the widest range of nominal meanings, suffixation performs a more specialized locative function in the examined forms, and affix combinations exhibit the highest degree of structural complexity. Third, the relationships among affix form, lexical base, morphological position, and nominal meaning are sufficiently regular to be formulated

through explicit morphological rules rather than treated as an unstructured inventory of lexical items (Gosal et al., 2023). Collectively, these findings demonstrate that Tombatu nominalization is systematic within the boundaries of the present dataset. At the same time, the restrictions observed in the data indicate that the three strategies should not be regarded as equally productive or freely applicable.

The first key finding concerns the semantic differentiation of prefixal nominalization. In line with Gosal et al. (2023), the present study identifies 11 noun-forming prefixes: /tapə-/ , /ta-/ , /to-/ , /iwa-/ , /tapa-/ , /i-/ , /ko-/ , /do-/ , /so-/ , /tapo-/ , and /bɛwa-/ . These prefixes are not semantically interchangeable because each tends to select particular bases and produce particular categories of nominal meaning. The contrast between /tapəluku?/ ‘drinker’ and /taluku?/ ‘a person who drinks excessively’, for instance, shows that closely related agentive prefixes may encode different degrees of habituality, evaluation, or social characterization. The prefix /so-/ in /sosahul/ ‘spear tool’, by comparison, produces an instrumental interpretation rather than a human-agent interpretation. These patterns indicate that Tombatu prefixes contribute both grammatical category and lexical-semantic information, which is consistent with descriptions of agglutinative morphology in which affixes carry identifiable grammatical and semantic content (Gosal et al., 2023).

The semantic role of Tombatu prefixes is also supported by cross-linguistic evidence. In Yoruba, prefixation contributes identifiable meaning to nominal derivation (Eleshin, 2021). In Dura, nominal morphology derives both agentive and non-agentive nouns from verbal bases, thereby demonstrating the role of nominalizers in lexical-category change (Keraf, 1990, 1991; Pons, 2021). Tombatu extends this comparative evidence by showing that nominalizing

prefixes do not merely convert a verbal or adjectival base into a noun. They also classify the resulting referent as a participant, instrument, object, or socially evaluated individual. The distinction among these meanings suggests that nominalization participates in the linguistic categorization of social roles, behavioural characteristics, and material objects within the Tombatu speech community.

The second key finding concerns the functional specialization of suffixation. Within the present dataset, the suffix /-an/ forms locative nouns such as /tawoian/ ‘workplace’ and /beletan/ ‘lying place’. In both examples, the suffix transforms an activity or state into a noun referring to the location associated with that activity or state. This specialized locative function contrasts with the broader semantic range of prefixation and indicates a division of morphological functions within the Tombatu nominalization system (Gosal et al., 2023). Prefixes distribute nominal meanings across several semantic domains, whereas /-an/ performs a comparatively focused role in the data examined.

This interpretation is typologically plausible. Suffixes frequently form cohesive morphological and phonological units with the bases to which they attach (Missud & Villoing, 2020). Nese similarly employs the suffix /-ian/ in nominalization (Takau, 2023), while comparative research on noun and verb morphology confirms that suffixes can maintain relatively stable functions in derivational relationships (Koenig & Michelson, 2020). Nevertheless, the Tombatu dataset contains only a limited number of forms involving /-an/. The strongest empirically defensible interpretation is therefore that /-an/ functions consistently as a locative nominalizer in the forms documented in this study. A broader claim concerning its general productivity would require additional corpus evidence, elicited forms, and acceptability judgements from speakers.

The third key finding concerns the layered organization of affix combinations. Tombatu constructions involving a prefix and suffix, or a prefix, infix, base, and suffix, demonstrate that several bound morphemes may be organized within a single derived noun. The infixes /-in-/ and /-um-/ occur in complex configurations such as /be- -um- -an/ and /pah- -in- -an/. However, the present evidence indicates that their role in nominalization is more reliably structural than independently semantic (Gosal et al., 2023). They contribute to the internal organization of complex derived forms, but the dataset does not establish a consistent nominal meaning that can be attributed exclusively to either infix.

This structural function differentiates the infixes from prefixes that express relatively transparent agentive or instrumental meanings and from the suffix /-an/, which expresses a locative meaning in the examined forms (Gosal et al., 2023). Cross-linguistic findings nevertheless demonstrate that infixes may

perform more explicit semantic functions in other languages. Hindko, for example, uses /-o-/ to express location or direction, while Movima employs infixation within a highly complex morphological system (Haude, 2023). The Tombatu evidence therefore contributes to a broader typological understanding of infixation by showing that an infix may participate in nominal derivation without necessarily carrying a stable and independently identifiable nominal meaning.

An additional analytical question concerns whether the Tombatu prefix and suffix configurations should be classified as genuine circumfixes or as instances of sequential affixation. The occurrence of one affix before a base and another after it does not automatically establish circumfixation. A circumfix analysis requires evidence that the two elements function together as a single discontinuous morpheme and that neither element independently produces the same derivational meaning. Until such evidence is established through distributional analysis, speaker judgements, and comparison with independently occurring affixes, the more cautious interpretation is to describe the structures as complex affix combinations.

The findings also substantiate the relevance of generative morphology to the analysis of Tombatu. The attested nominal forms can be represented as recurring relationships among lexical bases, affixes, morphological boundaries, output categories, and semantic interpretations. Aronoff’s requirement that word-formation rules be grounded in actual words recognized by speakers is particularly relevant because the Tombatu patterns are based on attested speech and speaker-validated meanings rather than hypothetical combinations (Aronoff, 1976). The rules therefore capture morphological generalizations that emerge from authentic language use.

Generative morphology also provides a means of explaining speakers’ implicit knowledge of affix selection and base compatibility. Tombatu speakers are able to recognize the relationship between a derived noun and its lexical base, interpret the semantic contribution of its affixes, and judge whether a particular combination is acceptable. Such knowledge can be represented through derivational schemas that predict recurring structural and semantic patterns. The framework consequently offers both descriptive adequacy and a degree of predictive capacity for analysing Tombatu noun formation. This interpretation is consistent with the broader generative conception of linguistic competence. Chomsky’s *Generative Grammar*, in *Syntactic Structures*, conceptualizes linguistic knowledge as a system of rules that enables speakers to produce and interpret expressions beyond a memorized inventory. Tombatu speakers’ ability to relate derived nouns to recurring combinations of bases and affixes is compatible with this theoretical principle. However, the productivity of each proposed rule should eventually be tested through novel-form experiments rather than inferred solely from attested lexical items.

This interpretation aligns with the broader generative view of linguistic competence. In *Syntactic Structures*, (Chomsky, 1957) Generative Grammar defines linguistic knowledge as a rule-governed system that enables speakers to produce and interpret expressions beyond a memorized set of forms. Similarly, Tombatu speakers appear able to associate derived nouns with recurring combinations of bases and affixes. Nevertheless, the productivity of the proposed rules should be verified through experiments involving novel forms rather than inferred solely from attested lexical items.

The Tombatu evidence is predominantly derivational rather than inflectional. The processes identified in the study create new lexemes, modify lexical categories, or introduce meanings involving participants, locations, instruments, objects, states, processes, and results. This interpretation is consistent with Scalise (1986) and Katamba (1993), who emphasize that derivational analysis must account for the input base, the affix, the output category, the resulting meaning, and the restrictions governing the formation. In Tombatu, the attachment of nominalizing morphology frequently produces a new noun with a semantic function that differs from that of its base.

Within this framework, Word Combination Rules can represent the linear ordering of morphemes, Derivation Rules can model the formation of new nominal lexemes, and the Boundary Insertion Convention can make the internal organization of complex forms explicit (Gosal et al., 2023). Inflection Rules are less central to the present findings because the study does not independently demonstrate an inflectional process responsible for creating the identified nominal lexemes. Inflection is therefore more appropriately regarded as a possible later component of the grammatical system rather than as the primary mechanism of Tombatu nominalization. The broader applicability of explicit rule-based analysis is supported by research on Maltese /-ata/ nominalization (Caruana, 2023), adjectival and nominal typology (Mitrović & Panagiotidis, 2020), Polish complex noun formation (Kolbusz-Buda, 2024), and morphosyntactic processes in Albanian and Cinyanja (Deda & Hoshafi, 2023; Njobvu, 2020).

A central theoretical implication of the findings is that Tombatu nominalization is constrained rather than mechanically unlimited. The Unitary Base Hypothesis, Binary Branching Hypothesis, No Phrase Constraint, Adjacency Condition, and Atomic Condition provide relevant analytical diagnostics for explaining why particular affixes select particular bases and why some formally conceivable combinations may be unacceptable (Aronoff, 1976). These constraints indicate that the nominalization system is governed by lexical, categorial, structural, and semantic conditions. The semantic asymmetry among the three nominalization strategies is therefore theoretically significant. Agentive and person-denoting prefixes,

instrumental and object-related prefixes, locative /-an/, and structurally complex affix combinations occupy different functional domains. The nominalization system is consequently not composed of freely interchangeable affixes. Instead, each affix or affixal configuration operates within a particular relationship among base category, morphological position, and semantic output. Within Austronesian linguistics, this lexical evidence complements research showing that nominalization can interact with alignment, voice, category formation, and diachronic development (Aldridge, 2021). Although the present study does not provide direct evidence of alignment change, it establishes that nominalization plays a central role in organizing Tombatu lexical categories and semantic relations. This provides an empirical foundation for future morphosyntactic and historical comparison.

The study also holds significant documentary value. Declining language transmission may lead not only to vocabulary loss but also to the erosion of grammatical distinctions and culturally embedded systems of meaning. This concern aligns with the broader threat to linguistic diversity identified by Bromham et al. (2022). Language endangerment should therefore be investigated through systematic analysis of grammatical knowledge. By documenting how Tombatu morphology encodes agents, social identities, instruments, places, processes, objects, and results, this study preserves an organized component of the language's grammatical system. It also shows that an underdescribed regional language can contribute valuable evidence to broader morphological theory.

Several empirical and analytical gaps remain. Earlier studies established important historical and descriptive foundations for Minahasan linguistics. Sneddon (1978) contributed a historical-comparative account of Proto-Minahasa, while Kalangi (1984) described aspects of Tombatu phonology and morphology. However, these studies did not provide a focused analysis of nominalization or formulate explicit rules governing Tombatu noun formation. The present study addresses this historical gap by documenting nominalizing morphemes and explaining their structural and semantic functions.

Despite this contribution, the current study remains limited by its localized data source, the absence of frequency-based corpus measurements, the small number of suffixal examples, and the uneven exemplification of the 14 affix combinations. The analysis does not test the productivity of each affix with unattested lexical bases, fully distinguish allomorphic variants from separate morphemes, or model morphophonological readjustments in sufficient detail. It also does not conclusively determine whether the documented prefix and suffix combinations should be classified as circumfixes or sequential affixes. Furthermore, the synchronic design does not compare speakers from different age groups, discourse genres, villages, or neighbouring Minahasan language

communities. Consequently, the social distribution, degree of vitality, historical development, and broader productivity of the identified patterns remain open to further investigation.

The principal novelty of this study is its integration of detailed Tombatu documentation with a generative account of noun formation. Beyond cataloguing affixes, it identifies systematic semantic differences among three nominalization strategies and links them to base selection, affix position, structural organization, and morphological restrictions. Using Word Combination Rules, Derivation Rules, the Boundary Insertion Convention, and relevant constraints, the analysis explains both licensed and restricted forms. This combination of description, formal modelling, and semantic interpretation extends generative morphology to an underrepresented Minahasan language and clarifies how nominalizing morphemes interact with lexical bases to produce distinct noun categories. It also contributes to research on Austronesian word formation, grammatical categories, and historical development (Aldridge, 2021).

The study further provides a documentation framework covering lexical bases, grammatical categories, segmentation, phonological forms, glosses, derivational structures, semantic functions, distribution, and acceptability. This is crucial because declining intergenerational transmission may endanger distinctive grammatical and semantic systems (Bromham et al., 2022; Bromham, 2023). The findings can support dictionaries, morphological databases, learning resources, and community documentation programmes. Comparisons among Tombatu, Indonesian, and English nominalization may also enhance learners' morphological and metalinguistic awareness by treating Tombatu as a linguistic resource, consistent with translanguaging and multilingual education perspectives (Cenoz, 2025; Goodman and Tastanbek, 2021; Prilutskaya, 2021; Yun et al., 2025).

Future research should develop a searchable, speaker-validated Tombatu corpus representing different villages, age groups, social settings, and communicative genres to examine the frequency, productivity, and distribution of nominalizers. Elicitation, acceptability-judgement, and novel-word formation tasks should determine whether speakers apply the identified rules to unfamiliar bases. Comparative studies should investigate cognate nominalization patterns in other Minahasan languages and their historical links to Austronesian alignment, voice, and diathesis systems. Further analysis is needed on morphophonological alternations, allomorphy, affix ordering, base-selection restrictions, circumfixation

versus sequential affixation, and the semantic contributions of /-in-/ and /-um-/. Interdisciplinary research should also assess how dictionaries, digital corpora, community archives, and multilingual teaching materials support language maintenance, intergenerational transmission, and culturally responsive education.

6. Conclusion

This study demonstrates that affixal noun formation in the Tombatu language constitutes a systematic, semantically organized, and rule-governed nominalization system. Within the present dataset, 26 affixal configurations were identified, comprising 11 noun-forming prefixes, the suffix /-an/, and 14 affix combinations. These configurations represent three distinct but interconnected patterns: prefixation displays the greatest structural and semantic variation by forming agentive, person-denoting, instrumental, object-related, and result-related nouns; suffixation is more restricted, with /-an/ forming locative nouns; and affix combinations produce structurally complex nouns denoting places, processes, objects, states, results, and entities. The analysis further shows that nominal meanings are determined by affix position, base selection, morphological structure, and restrictions on affix attachment, as explained through Word Combination Rules, Derivation Rules, the Boundary Insertion Convention, and relevant morphological constraints.

The principal novelty of the study lies in providing the first systematic generative account of Tombatu nominalization that moves beyond listing affixes to explain how acceptable noun forms are derived and why particular combinations are restricted. The findings extend empirical evidence for Austronesian word formation, strengthen the grammatical documentation of an underdescribed Minahasan language, and provide a linguistic foundation for Tombatu language maintenance, instructional-material development, and contrastive morphological activities involving Tombatu, Indonesian, and English. Nevertheless, these conclusions should be interpreted within the scope of the linguistic forms and speakers represented in the current dataset. This limitation is particularly relevant to suffixal nominalization, as the analysis is based on a relatively small number of examples and may not fully capture its broader patterns of use in Tombatu.

References

- Aldridge, E. (2021). Two types of alignment change in nominalizations: Austronesian and Japanese. *Diachronica*, 38(1), 1–39. <https://doi.org/10.1075/dia.19044.ald>
- Aronoff, M. (1976). *Word formation in generative grammar*. MIT Press. <https://doi.org/10.2307/412595>

- Austin, P. K. (2010). Current issues in language documentation. *Language Documentation and Description*, 7, 12–33. <https://doi.org/10.25894/ldd225>
- Beguš, G., & Dąbkowski, M. (2025). *The origins of the Austronesian voice system and subject-only restriction* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/c9765_v2
- Beguš, G., Dąbkowski, M., & Drummond, E. (2023). *The origins of the Austronesian voice system and subject-only restriction* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/c9765>
- Bromham, L. (2023). Language endangerment: Using analytical methods from conservation biology to illuminate loss of linguistic diversity. *Cambridge Prisms: Extinction*, 1, Article e3. <https://doi.org/10.1017/ext.2022.3>
- Bromham, L., Dinnage, R., Skirgård, H., Ritchie, A., Cardillo, M., Meakins, F., Greenhill, S., & Hua, X. (2022). Global predictors of language endangerment and the future of linguistic diversity. *Nature Ecology & Evolution*, 6(2), 163–173. <https://doi.org/10.1038/s41559-021-01604-y>
- Brook, M., & Blamire, E. (2023). Language play is language variation: Quantitative evidence and what it implies about language change. *Language*, 99(3), 491–530. <https://doi.org/10.1353/lan.2023.a907010>
- Caruana, S. (2023). -ata nominalisations in Maltese: A language contact perspective. *Atti del Sodalizio Glottologico Milanese*, 16, 17–39. <https://doi.org/10.54103/1972-9901/20526>
- Cenoz, J., & Gorter, D. (2025). The potential of pedagogical translanguaging in English language and in English-medium content classes. *Linguistics and Education*, 86, Article 101399. <https://doi.org/10.1016/j.linged.2025.101399>
- Chomsky, N. (1957). *Syntactic structures*. Mouton. <https://doi.org/10.1515/9783112316009>
- Deda, A., & Hoshafi, M. (2023). Word formation head and the typology analysis of the word formation corpus in Albanian. *Journal of Educational and Social Research*, 13(6), 255–265. <https://doi.org/10.36941/jesr-2023-0162>
- Eleshin, B. (2021). Typology of headedness in Yorùbá nominal derivations. *Journal of Universal Language*, 22(1), 91–125. <https://doi.org/10.22425/jul.2021.22.1.91>
- Goodman, B., & Tastanbek, S. (2021). Making the shift from a codeswitching to a translanguaging lens in English language teacher education. *TESOL Quarterly*, 55(1), 29–53. <https://doi.org/10.1002/tesq.571>
- Gosal, N., Manangkot, V. E., & Liuw, J. K. (2023). Verbalization of Tombatu language. In R. H. E. Sendouw, T. Pangalila, S. Pasandaran, & V. P. Rantung (Eds.), *Proceedings of the Unima International Conference on Social Sciences and Humanities (UNICSSH 2022)* (Vol. 698, pp. 1517–1523). Atlantis Press. <https://doi.org/10.1002/tesq.571>
- Gosal, N., Manangkot, V. E., & Liuw, J. K. (2023). Verbalization of Tombatu language. In *Proceedings of the Unima International Conference on Social Sciences and Humanities (UNICSSH 2022)* (pp. 1517–1523). Atlantis Press. https://doi.org/10.2991/978-2-494069-35-0_182
- Gromann, D., Apostol, E.-S., Chiarcos, C., Cremaschi, M., Gracia, J., Gkirtzou, K., Liebeskind, C., Mockienė, L., Rosner, M., Schuurman, I., Sérasset, G., Silvano, P., Spahiu, B., Truică, C.-O., Utkā, A., & Valūnaitė Oleškevičienė, G. (2024). Multilinguality and LLOD: A survey across linguistic description levels. *Semantic Web*, 15(5), 1915–1958. <https://doi.org/10.3233/SW-243591>
- Haude, K. (2023). Movima. In P. Epps & L. Michael (Eds.), *Amazonian languages: Language isolates II: Kanoé to Yurakaré: An international handbook* (Vol. 2, pp. 807–850). De Gruyter Mouton. <https://doi.org/10.1515/9783110432732-004>
- Hemmings, C. (2020). Methods in language documentation and description: A guide to the Kelabit documentation project. *Journal of Modern Languages*, 30(1), 6–76. <https://doi.org/10.22452/jml.vol30no1.1>
- Hidayati, A. F., Nurhayani, I., & Choyimah, N. (2021). Madurese deverbal nominalization process. *Prosodi*, 15(2), 113–120. <https://doi.org/10.21107/prosodi.v15i2.12177>
- Hockett, C. F. (1965). *A course in modern linguistics*. Macmillan. <https://a.co/d/0d6MBSVE>
- Jufrizal, J., & Refnita, L. (2021). Gerund in English: A morpho-semantic note for EFL learners in Indonesia. *Lingua Didaktika: Jurnal Bahasa dan Pembelajaran Bahasa*, 15(1), 83–97. <https://doi.org/10.24036/ld.v15i2.114512>
- Jufrizal, J., Refnita, L., & Arianto, M. A. (2024). Ergative and antipassive construction in Minangkabau: Are there? *Linguistik Indonesia*, 42(1), 1–20. <https://doi.org/10.26499/li.v42i1.585>
- Kalangi, F. J. (1984). *Menapak bahasa daerah Tondanow di Minahasa Sulawesi Utara tentang fonologi dan morfologi* [Unpublished manuscript].
- Katamba, F. (1993). *Morphology*. Macmillan Press.
- Kaufman, D. (2026). A global view on the *Ausrichtung* affix and the deep differences

- between nominalization-derived versus applicative-derived diathesis in Austronesian. In S. Riesberg, U. Reinöhl, & B. Hellwig (Eds.), *The documentarist turn: From observable linguistic behaviour to typological generalizations* (pp. 518–571). John Benjamins Publishing Company. <https://doi.org/10.1075/slcs.240.20kau>
- Keraf, G. (1990). *Linguistik bandingan tipologis*. Gramedia.
- Keraf, G. (1991). *Linguistik bandingan historis*. Gramedia.
- Kim, Y. (2020). Morphology and language documentation. In *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. <https://doi.org/10.1093/acrefore/9780199384655.013.518>
- Klamer, M. (2020). From Lamaholot to Alorese: Morphological loss in adult language contact. In D. Gil & A. Schapper (Eds.), *Austronesian undressed: How and why languages become isolating* (pp. 339–367). John Benjamins Publishing Company. <https://doi.org/10.1075/tsl.129.07kla>
- Klamer, M. (2020). From Lamaholot to Alorese: Morphological loss in adult language contact. In D. Gil & A. Schapper (Eds.), *Austronesian undressed: How and why languages become isolating* (pp. 339–368). John Benjamins. <https://doi.org/10.1075/tsl.129.07kla>
- Koenig, J.-P., & Michelson, K. (2020). What does being a noun or a verb mean? *Proceedings of the 27th International Conference on Head-Driven Phrase Structure Grammar*, 106–126. <https://doi.org/10.21248/hpsg.2020.6>
- Kolbusz-Buda, J. (2024). Between compounding and phrasal derivation: Polish complex nouns in sam(o)-. *Word Structure*, 17(3), 112–133. <https://doi.org/10.3366/word.2024.0240>
- Krauska, A., & Lau, E. (2023). Moving away from lexicalism in psycho- and neuro-linguistics. *Frontiers in Language Sciences*, 2, Article 1125127. <https://doi.org/10.3389/flang.2023.1125127>
- Kridalaksana, H. (1984). *Kamus linguistik*. Gramedia.
- Kridalaksana, H. (1992). *Pembentukan kata dalam bahasa Indonesia*. Gramedia Pustaka Utama.
- Krohn, H. S. (2021). Indonesian vs. Bribri: Striking lexical similarities in two unrelated languages. *Jurnal Bahasa dan Sastra*, 9(1), 1–9. <https://doi.org/10.24036/jbs.v9i1.111535>
- Lapoliwa, H. (1981). *A generative approach to the phonology of Bahasa Indonesia* (Pacific Linguistics, Series D, No. 34). Department of Linguistics, Research School of Pacific Studies, Australian National University. <https://doi.org/10.25501/SOAS.00029045>
- Li, L.-Z., & Yang, T. (2020). A cognitive study of nominalization in Mandarin. *Journal of Literature and Art Studies*, 10(9), 835–847. <https://doi.org/10.17265/2159-5836/2020.09.013>
- Lumbantobing, A. B., & Situngkir, H. (2026). *Tokenizations for Austronesian language models: Study on languages in Indonesia Archipelago* [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2602.06998>
- Lupyan, G., & Dale, R. (2021). Why are there different languages? The role of adaptation in linguistic diversity. *Trends in Cognitive Sciences*, 25(1), 5–16. <https://doi.org/10.1016/j.tics.2020.10.005>
- Luuk, E. (2019). A type-theoretical reduction of morphological, syntactic and semantic compositionality to a single level of description. In R. Mitkov & G. Angelova (Eds.), *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019)* (pp. 664–673). INCOMA Ltd. https://doi.org/10.26615/978-954-452-056-4_078
- Lyons, J. (1995). *Pengantar teori linguistik* (I. Soetikno, Trans.). Gramedia Pustaka Utama. (Original work published 1968).
- Martinet, A. (1980). *Éléments de linguistique générale*. Armand Colin.
- Missud, A., & Vilhoing, F. (2020). The morphology of rival -ion, -age, and -ment selected verbal bases. *Lexique*, (26), 29–52. <https://doi.org/10.54563/lexique.755>
- Mitrović, M., & Panagiotidis, P. E. (2020). Adjectives exist, adjectivisers do not: A bicategorical typology. *Glossa: A Journal of General Linguistics*, 5(1), Article 58, 1–28. <https://doi.org/10.5334/gjgl.940>
- Naudé, J. A., & Miller-Naudé, C. L. (2023). Generative linguistics as a theoretical framework for the explanation of problematic constructions in Biblical Hebrew. In W. A. Ross & E. Robar (Eds.), *Linguistic theory and the biblical text* (pp. 6–66). Open Book Publishers. <https://doi.org/10.11647/obp.0358.01>
- Nida, E. A. (1970). *Morphology: The descriptive analysis of words*. University of Michigan Press.
- Njobvu, N. (2020). *Two types of phrasal compound in Cinyanja* [Doctoral dissertation, University of KwaZulu-Natal]. ResearchSpace.
- Pike, K. L. (1992). *Konsep linguistik* (K. Gunawan, Trans.). Gelora Aksara Pratama.
- Pons, M.-C. (2021). [Review of the book *The Dura language: Grammar and phylogeny*, by N.

- Schorer]. *Himalayan Linguistics*, 20(1), 116–132.
<https://doi.org/10.5070/H920155279>
- Prilutskaya, M. (2021). Examining pedagogical translanguaging: A systematic review of the literature. *Languages*, 6(4), Article 180.
<https://doi.org/10.3390/languages6040180>
- Robins, R. H. (1992). *Linguistik umum: Sebuah pengantar* (S. Djajanegara, Trans.). Kanisius.
- Ross, M., & Zeitoun, E. (2024). Formosan languages and Proto-Austronesian morphology. In P. J.-K. Li, E. Zeitoun, & R. De Busser (Eds.), *Handbook of Formosan languages: The indigenous languages of Taiwan* (Part 2, pp. 134–182). Brill.
https://doi.org/10.1163/2772_5766_COM-103030
- Salem, V. E. T., & Mesra, R. (2020). Preservation of local language culture in Toundanouw Village, Southeast Minahasa Regency. In *Proceedings of the 3rd International Conference on Social Sciences (ICSS 2020)* (pp. 175–177). Atlantis Press.
<https://doi.org/10.2991/assehr.k.201014.039>
- Samsuri. (1987). *Aliran transformasi 1957–1965*. Lembaga Bahasa Universitas Katolik Atma Jaya.
- Schapper, A. (2020). Introduction to *The Papuan languages of Timor, Alor and Pantar: Volume 3*. In A. Schapper (Ed.), *The Papuan languages of Timor, Alor and Pantar: Volume 3: Sketch grammars* (pp. 1–52). De Gruyter Mouton.
<https://doi.org/10.1515/9781501511158-001>
- Sneddon, J. N. (1978). *Proto-Minahasan: Phonology, morphology, and wordlist* (Pacific Linguistics, Series B, No. 54). Department of Linguistics, Research School of Pacific Studies, Australian National University. <https://doi.org/10.15144/PL-B54.1>
- Sudaryanto. (1993). *Metode dan aneka teknik analisis bahasa: Pengantar penelitian wahana kebudayaan secara linguistik*. Duta Wacana University Press.
- Sudiroatmadja, M. H. (1992). *Javanese word formation: A generative grammar approach* [Doctoral dissertation, Georgetown University].
- Suhandano. (2021). Actor and patient in Javanese and Indonesian imperative clauses. In *Proceedings of the International Congress of Indonesian Linguistics Society (KIMLI 2021)* (pp. 99–106). Atlantis Press.
<https://doi.org/10.2991/assehr.k.211226.022>
- Takau, L. G. (2023). *A grammar of Nese*. ANU Press.
<https://doi.org/10.22459/GN.2023>
- Ulhaq, H. D., Suryana, Y., Asikin, N. A., & Fadhly, F. Z. (2022). A morpho-syntactic error analysis of university students' argumentative writing. *Indonesian Journal of Learning and Instruction*, 5(2), 23–28.
<https://doi.org/10.25134/ijli.v5i2.6844>
- Verhaar, J. W. M. (2001). *Asas-asas linguistik*. Gadjah Mada University Press.
- Wu, J. (2022). [Review of the book *Austronesian undressed: How and why languages become isolating*, edited by D. Gil & A. Schapper]. *Oceanic Linguistics*, 61(2), 791–796.
<https://doi.org/10.1353/ol.2022.0026>
- Yun, S., Choi, S. S., Olğün-Baytaş, M., Lyu, S. R., & La Croix, L. (2025). Enhancing learning for emergent bi/multilingual children through translanguaging pedagogies: A systematic literature review from 2010 to 2023. *Contemporary Issues in Early Childhood*. Advance online publication.
<https://doi.org/10.1177/14639491251338422>