

MORPHOPHONEMIC AND STRUCTURAL PATTERNS IN THE KAMBERA NUMERAL SYSTEM

Raynesta Mikaela Indri Malo¹, Suryani Kurniawi K. L. Kapoe²

^{1,2}Universitas Kristen Wira Wacana Sumba

raynesta@unkriswina.ac.id

ABSTRACT

Numeral systems constitute an important component of linguistic structure because they encode not only quantitative meaning but also grammatical organization, morphophonemic processes, and language change. Cross-linguistically, numeral systems often preserve conservative features while remaining sensitive to internal restructuring and language contact. This study investigates the numeral system of Kampera, an Austronesian language spoken in East Sumba, East Nusa Tenggara, Indonesia, with the aim of describing its structural organization, identifying morphophonemic patterns, and examining possible contact-induced influences.

This research adopts a qualitative descriptive approach. Data were collected through structured and semi-structured interviews with native speakers of Kampera in Kampera District, East Sumba Regency. Structured elicitation was used to document numeral forms from units to millions, followed by in-depth interviews with key informants to validate the data and explain structural variation.

The findings show that Kampera employs a decimal (base-10) numeral system in which basic numerals are lexicalized and higher numerals are systematically formed through additive constructions. Structurally, the numeral system is predominantly characterized by an [n base] pattern, in which the unit numeral precedes the numerical base, while the [base n] pattern is restricted to the teen numerals (11–19). This distribution indicates a mixed structural system across different numerical domains. Larger numerals are consistently organized in a hierarchical left-branching structure, with higher-value elements preceding lower-value elements, demonstrating that numeral formation follows regular compositional principles rather than arbitrary lexical combinations.

The study also reveals systematic morphophonemic alternations. The numeral 'hau' (one) undergoes monophthongization when combined with higher numerical bases, producing contracted forms such as 'hangahu' (one hundred) and 'harata' (one thousand). This process simplifies pronunciation while preserving morphological structure. In addition, the independent form 'hambullu' (ten) surfaces as 'kambullu' in compound numerals, reflecting context-dependent morphophonemic restructuring associated with numeral compounding. Comparison with earlier descriptions by Klamer further suggests that these forms represent ongoing historical developments within the Kampera numeral system rather than isolated phonological changes.

The study also identifies lexical variation involving the numeral forms 'dua' and 'dambu' (two). The findings show that 'dua' is predominantly used in the tens, hundreds, and thousands (e.g., 'dua kambullu', 'dua ngahu', 'dua harata'), whereas 'dambu' is retained in the unit domain and in certain higher expressions such as 'dambu njuta' (two million). Rather than indicating straightforward lexical replacement, this distribution reflects functional differentiation within the numeral system.

Keywords: Kampera language; Numeral system; Morphophonemics; Language contact; Austronesian typology

ABSTRAK

Sistem bilangan merupakan komponen penting dalam struktur linguistik, karena sistem ini tidak hanya membawa makna kuantitatif, tetapi juga struktur gramatikal, proses fonologis, dan dinamika sosiolinguistik. Dalam berbagai bahasa, bilangan sering kali mempertahankan bentuk-bentuk kuno sekaligus berfungsi sebagai domain yang sangat rentan terhadap kontak bahasa dan penggantian. Penelitian ini mempelajari sistem bilangan dalam bahasa Kampera, sebuah bahasa Austronesia yang digunakan di Sumba Timur, Nusa Tenggara Timur, Indonesia, dengan tujuan mendeskripsikan pola strukturalnya serta mengidentifikasi perubahan morf fonemik dan perubahan yang dipicu oleh kontak bahasa.

Penelitian ini menggunakan pendekatan deskriptif kualitatif. Data dikumpulkan melalui wawancara terstruktur dan semi-terstruktur dengan penutur asli bahasa Kampera yang berdomisili di Kecamatan Kampera, Kabupaten Sumba Timur. Tahap awal dilakukan melalui elisitasi terstruktur untuk memperoleh bentuk-bentuk numeralia pada berbagai domain bilangan, termasuk satuan, puluhan, ratusan, ribuan, dan jutaan. Selanjutnya, wawancara mendalam dengan informan kunci dilakukan untuk memvalidasi data, memperjelas pola pembentukan, serta mengidentifikasi variasi yang dipengaruhi oleh bilingualisme dengan bahasa Indonesia.

Hasil penelitian menunjukkan bahwa bahasa Kampera menggunakan sistem numeralia desimal (basis-10), di mana numeralia dasar bersifat leksikalisasi, sedangkan numeralia yang lebih tinggi dibentuk secara sistematis melalui konstruksi aditif. Secara struktural, sistem numeralia ini didominasi oleh pola [n basis], yaitu numeralia satuan mendahului basis numeralia, sedangkan pola [basis n] terbatas pada numeralia belasan (11–19). Distribusi tersebut menunjukkan bahwa bahasa Kampera memiliki sistem struktur campuran pada domain numeralia yang berbeda. Numeralia bernilai besar juga disusun secara konsisten dalam struktur hierarkis bercabang ke kiri (left-branching), dengan unsur bernilai lebih tinggi mendahului unsur bernilai lebih rendah. Hal ini menunjukkan bahwa

pembentukan numeralia mengikuti prinsip komposisional yang sistematis, bukan sekadar gabungan leksikal yang bersifat arbitrer.

Penelitian ini juga mengungkap adanya alternasi morfofonemis yang sistematis. Numeralia 'hau' (satu) mengalami proses monofongisasi ketika bergabung dengan basis numeralia yang lebih tinggi sehingga menghasilkan bentuk kontraksi seperti 'hangahu' (serratus) dan 'harata' (seribu). Proses ini menyederhanakan pengucapan tanpa menghilangkan struktur morfologisnya. Selain itu, bentuk mandiri 'hambullu' (sepuluh) muncul sebagai 'kambullu' dalam numeralia majemuk, yang mencerminkan restrukturisasi morfofonemis bergantung pada konteks pembentukan numeralia. Perbandingan dengan deskripsi terdahulu oleh Klamer menunjukkan bahwa bentuk-bentuk tersebut lebih mencerminkan perkembangan historis yang masih berlangsung dalam sistem numeralia bahasa Kambera, bukan sekadar perubahan fonologis yang terpisah.

Penelitian ini juga mengidentifikasi adanya variasi leksikal pada bentuk numeralia 'dua' dan 'dambu' (dua). Hasil penelitian menunjukkan bahwa 'dua' lebih dominan digunakan pada numeralia puluhan, ratusan, dan ribuan (misalnya 'dua kambullu', 'dua ngahu', 'dua harata'), sedangkan 'dambu' tetap dipertahankan pada domain satuan dan pada beberapa konstruksi numeralia yang lebih tinggi, seperti 'dambu njuta' (dua juta). Distribusi ini tidak menunjukkan adanya penggantian leksikal secara langsung, melainkan mencerminkan diferensiasi fungsi di dalam sistem numeralia bahasa Kambera.

Kata kunci: Bahasa Kambera; Sistem bilangan; Morfofonemik; Kontak bahasa; Tipologi Austronesia

INTRODUCTION

Language can be defined as a system of communication consisting of symbols, words, and rules that enable individuals to convey and interpret meaning. As a uniquely human capacity, language plays a fundamental role in social interaction and cognitive processes. It is estimated that there are approximately 7,170 languages spoken worldwide, with significant variation in the number of speakers across languages (Ethnologue, 2026). This linguistic diversity is shaped by a range of factors, including historical developments such as migration, colonization, and cultural interaction, as well as geographical isolation and socio-cultural influences that contribute to regional variation.

Within any language, words are organized into categories known as word classes or parts of speech, based on their grammatical functions and roles in sentence structure. These categories facilitate efficient communication by allowing speakers to understand syntactic relationships and semantic roles among linguistic elements. Generally, word classes include adverbs, determiners, nouns, prepositions, conjunctions, verbs, adjectives (Aarts & Haegeman, 2021), pronouns (Thornbury, 2017), and numerals ("Kamus Besar Bahasa Indonesia," 2016; Kridalaksana, 2009). While many languages share similar categories, variation exists depending on the grammatical system and structural characteristics of each language.

Numerals are a fundamental category in human language, universally present and essential for expressing exact quantities. They are typically defined as words or expressions used to denote precise numerical values, such as 'seven' or 'ninety-eight', in contrast to non-specific quantifiers like 'some' or 'many' (Barlow, 2024, 2025). As a core linguistic category, numerals play a crucial role in quantification, ordering, and the conceptualization of numerical relationships, thereby reflecting important aspects of human cognition and abstraction (Yiagnigni, 2026). In most linguistic descriptions, particular attention is given to cardinal numerals, which are used in counting, answering questions of quantity, and assigning exact numerical values to referents (Barlow, 2024).

From a typological perspective, numeral systems exhibit both diversity and strong cross-linguistic regularities. Most languages organize numerals around a base, with higher numbers formed through arithmetic operations such as addition and multiplication, and in some cases exponentiation (Comrie, 2013, 2022). Among these systems, the decimal (base-10) system is by far the most widespread, dominating in nearly all regions of the world (Comrie, 2013; Yiagnigni, 2026). This prevalence reflects broader cognitive and cultural tendencies, as speakers typically construct larger numbers by combining smaller base units, most commonly based on ten (Harrison, 2007). Despite this general uniformity, numeral systems across languages display considerable variation in structure and formation.

In the Indonesian context, the decimal (base-10) system is the most widely used. However, previous studies have shown that numeral systems in Indonesian local languages are highly diverse, reflecting differences in cultural practices, geographical distribution, and historical development. Some languages employ base systems smaller than ten, such as quinary (base-5), while others use decimal systems as their primary structure (Mahsun, 2010; Winarti, 2017). In Papua, numeral systems can be particularly complex, incorporating combinations of binary (base-2), quinary (base-5), vigesimal (base-20), and body-part-based counting systems (de Vries, 1994). Similarly, languages in East Nusa Tenggara exhibit considerable variation in their basic numeral ranges, with some systems built on as few as five

core numerals and others extending to ten (Sanga, 2010). These findings highlight the richness and typological diversity of numeral systems in Indonesia.

In addition to their typological significance, numerals also provide insight into the interaction between lexical meaning and grammatical structure. As lexical items, numerals denote exact quantities; however, their formation and usage are often constrained by grammatical rules, including morphosyntactic ordering and structural composition. For instance, cross-linguistic research has shown that the internal structure of numerals (e.g., *[base n]* vs. *[n base]*) may correlate with syntactic patterns such as the ordering of numerals and nouns (Greenberg in Allasonnière-Tang & Her, 2020). This dual nature positions numerals at the intersection of lexicon and grammar, making them a valuable domain for linguistic analysis.

Kambera is an Austronesian language spoken primarily in the eastern part of Sumba Island, East Nusa Tenggara, Indonesia, and is classified within the Central Malayo-Polynesian subgroup (Klamer, 2005). Despite its significance, Kambera remains under-documented, particularly with regard to its morphosyntactic subsystems. While previous studies have addressed aspects of its phonology and lexicon, focused research on its numeral system is still limited. This study aims to address this gap by analyzing the Kambera numeral system, with particular attention to (1) the structural organization of numerals across numerical domains, (2) morphophonemic alternations in numeral formation, and (3) evidence of possible lexical influence of Indonesian. By situating Kambera within broader typological discussions, this study contributes to both Austronesian linguistics and language documentation efforts.

METHODOLOGY

This study employs a qualitative descriptive approach with an analytical orientation. Qualitative research aims to explore and provide in-depth understanding of a particular phenomenon from the participants' perspectives (Creswell, 2018). It is particularly suitable for examining linguistic structures as they are used and interpreted by native speakers. In this study, the qualitative approach allows for a detailed description of the Kambera numeral system, including its structural patterns and variations. Data collection was conducted through direct interaction with participants, as qualitative research emphasizes naturalistic inquiry and engagement with respondents (Rukajat, 2018).

Data were collected using elicitation and face-to-face interview techniques. Two types of interviews were employed: structured and semi-structured interviews (Herdiansyah, 2019). The structured interviews were conducted using a questionnaire distributed via Google Forms, which contained a list of numeral expressions in Indonesian. Participants were asked to provide their equivalents in Kambera. This stage functioned as a preliminary study to gather baseline data on numeral knowledge among speakers. A total of ten participants were involved in this phase.

The second stage involved semi-structured interviews with four key informants. These interviews allowed for more flexible yet guided data collection, enabling the researcher to validate the initial findings, clarify patterns of numeral formation, and explore variations influenced by language use. The data obtained from the structured interviews served as a reference point for comparison and deeper analysis in this stage. Overall, the data focused on how numerical expressions are formed in Kambera.

Data collection was conducted in Kambera District, East Sumba Regency, East Nusa Tenggara. The participants were native speakers of Kambera who are members of the local community and are proficient in both Kambera and Indonesian. Their bilingual competence was essential for ensuring accurate elicitation and interpretation of numeral forms.

RESULT AND DISCUSSION

Descriptive Overview of the Kambera Numeral System

Table 1 below presents the data obtained in this study, including numeral forms in Kambera from one to ten, as well as teens, tens, hundreds, thousands, and millions. The data show that the basic numerals from one to ten are expressed as *hau* 'one', *dambu* 'two', *tailu* 'three', *patu* 'four', *limma* 'five', *nomu* 'six', *pihhu* 'seven', *walu* 'eight', *hiwwa* 'nine', and *hombullu* 'ten'.

Table 1. Numerals in Kambara

1 (satu)	Hau
2 (dua)	Dambu
3 (tiga)	Tailu
4 (empat)	Patu
5 (lima)	Limma
6 (enam)	Nomu
7 (tujuh)	Pihhu
8 (delapan)	Walu
9 (sembilan)	Hiwwa
10 (sepuluh)	Hombullu
11 (sebelas)	Hombullu hau
20 (dua puluh)	Dua kambullu
21 (dua puluh satu)	Dua Kambullu Hau
30 (tiga puluh)	Tailu Kambullu
31 (tiga puluh satu)	Tailu Kambullu Hau
40 (empat puluh)	Patu Kambullu
41 (empat puluh satu)	Patu Kambullu Hau
50 (lima puluh)	Limma Kambullu
51 (lima puluh satu)	Limma Kambullu Hau
60 (enam puluh)	Nomu Kambullu
61 (enam puluh satu)	Nomu Kambullu Hau
70 (tujuh puluh)	Pihu Kambullu
71 (tujuh puluh satu)	Pihu Kambullu Hau
80 (delapan puluh)	Walu Kambullu
81 (delapan puluh satu)	Walu Kambullu Hau
90 (sembilan puluh)	Hiwa Kambullu
91 (sembilan puluh satu)	Hiwa Kambullu Hau
100 (seratus)	Hangahu
101 (seratus satu)	Hangahu Hau
130 (seratus tiga puluh)	Hangahu Tailu Kambullu
131 (seratus tiga puluh satu)	Hangahu Tailu Kambullu Hau
200 (dua ratus)	Dua Ngahu
1.000 (seribu)	Harata
2.000 (dua ribu)	Dua rata
1.000.000 (satu juta)	Hau njuta
2.000.000 (dua juta)	Dambu njuta
10.000.000 (sepuluh juta)	Hambullu Njuta

The findings indicate that Kambara belongs to the group of languages that employ a base-ten numeral system, as identified in previous studies of Indonesian languages (Mahsun, 2010; Winarti, 2017). In this system, higher numbers are constructed by combining basic numerals. For example, teen numerals (11–19) are formed by combining the numeral for ten (*hombullu*) with a unit numeral. Thus, *hombullu hau* means ‘eleven’ (10 and 1), and *hombullu limma* means ‘fifteen’ (10 and 5).

Tens are formed by combining a unit numeral with the term for ten. For instance, *patu kambullu* (four and ten) means ‘forty’, while *tailu kambullu* means ‘thirty’. Similarly, other tens include *dambu kambullu* ‘twenty’, *limma kambullu* ‘fifty’, *nomu kambullu* ‘sixty’, *pihu kambullu* ‘seventy’, *walu kambullu* ‘eighty’, and *hiwa kambullu* ‘ninety’. For compound numbers such as 41, the tens element precedes the unit, as in *patu kambullu hau* (four ten and one) ‘forty-one’.

Hundreds are expressed using the term *ngahu* ‘hundred’, with *hangahu* representing ‘one hundred’. Larger numbers are formed by combining hundreds, tens, and units in descending order. For example, *hangahu tailu kambullu hau* (one hundred – three ten and one) means ‘one hundred thirty-one’, while

patu ngahu pihu kambullu walu (four hundred – seven ten and eight) means ‘four hundred seventy-eight’. Similarly, *hangahu hau* represents ‘one hundred one’.

Thousands are expressed using the term *rata* ‘thousand’, with *harata* representing ‘one thousand’. Larger numbers follow the same additive principle, combining thousands, hundreds, tens, and units. For instance, *harata hangahu tailu kambullu hau* means ‘one thousand one hundred thirty-one’, while *pihu rata dua ngahu limma kambullu tailu* means ‘seven thousand two hundred fifty-three’.

For larger units, Kambera uses *njuta* ‘million’. Numerals such as *hau njuta* ‘one million’, *dambu njuta* ‘two million’, and *hambullu njuta* ‘ten million’ follow the same pattern, with the unit numeral preceding *njuta*.

Structural Patterns in Kambera Numerals

The descriptive data indicate that Kambera employs a decimal (base-10) numeral system. Basic numerals are lexicalized, while higher numerals are formed through additive constructions. The numeral system is predominantly characterized by an [n base] structure, in which the unit numeral precedes the base element. This pattern is clearly observed in the formation of tens, as illustrated by the following examples:

- *dambu kambullu* ‘twenty’ (‘two ten’)
- *tailu kambullu* ‘thirty’ (‘three ten’)
- *walu kambullu* ‘eighty’ (‘eight ten’)

In these forms, the unit numerals (*dambu*, *tailu*, *walu*) occur before the base *kambullu* ‘ten’, demonstrating that Kambera constructs higher numerals using a “unit + base” pattern. This indicates that the [n base] structure is both productive and systematic in the language.

In contrast, the [base n] structure (i.e., ‘ten + unit’) is restricted to a specific numerical domain, namely the teen numerals (11–19). For example:

- *hombullu hau* ‘eleven’ (‘ten one’)
- *hombullu walu* ‘eighteen’ (‘ten eight’)

This distribution of the [n base] and [base n] structure shows that Kambera exhibits a mixed structural system, where different numerical domains follow different formation patterns.

In larger numerals, elements are consistently ordered from higher to lower values (e.g., thousands → hundreds → tens → units), reflecting a hierarchical and left-branching structure. For example, the numeral 1.131 is constructed by sequencing *harata* ‘one thousand’, *hangahu* ‘one hundred’, *tailu kambullu* ‘thirty’, and *hau* ‘one’, in strict descending order.

- *harata hangahu tailu kambullu hau*
1.131 (1000 → 100 → 30 → 1)

This ordering indicates that each component occupies a fixed structural position within the numeral phrase. The left-branching nature of the construction means that higher-value elements precede lower-value elements, forming a layered structure in which each successive unit modifies or adds to the preceding value. Such consistency suggests that numeral formation in Kambera is governed by systematic compositional rules rather than arbitrary combinations of lexical items. It also reflects a cognitive strategy of numerical representation, where speakers process and produce quantities by decomposing them into hierarchically organized units.

Morphophonemic Alternations in Numeral Formation

A key finding of this study is the presence of systematic morphophonemic alternations involving the numeral expressing ‘one’. In the contemporary data, the numeral appears as *hau*, which undergoes monophthongization when combined with higher numerical units, resulting in contracted forms such as *hangahu* (*hau* & *ngahu*) ‘one hundred’ and *harata* (*hau* & *rata*) ‘one thousand’. Monophthongization refers to the change of a diphthong into a monophthong, a phonological process that commonly occurs to reduce articulatory complexity, facilitates easier and more efficient pronunciation. Thus, the diphthong /au/ in *hau* is simplified to /a/, yielding forms that are easier to pronounce while preserving their grammatical function. These reduced forms have subsequently become lexicalized as stable elements of the contemporary Kambera numeral system.

A historical perspective, however, suggests that these forms are part of a much older morphological system. According to Klammer (1998), Kambera distinguishes between counting and quantification. The

form *diha* functions only in counting sequences and is derived from the verb *diha* ‘to count’, as illustrated by the imperative expression *diha-ha da hapi!* ‘Count the cows!’. It is not used to quantify noun phrases. Instead, quantification requires either the individualizing prefix *ha-* or classifier-based forms. Klamer argues that *hau* ‘one (object)’ and *heu* ‘one (animal)’ are historically merged forms consisting of the prefix *ha-* and a classifier, while *dambu* (‘two’) originated from the merger of *dua* and the classifier *mbua* (classifier for spherical objects). Thus, in traditional Kambera, *hau* was not originally a simple numeral but a morphologically complex quantifying form.

This analysis is further supported by Kapita’s (1982) Kambera–Indonesian Dictionary. Kapita records *diha* with two meanings: ‘one’ and ‘to count’, confirming that both functions coexisted historically. He also identifies *ha-* as a prefix denoting ‘one’ and explains that *hakambulu* represents the earlier form for ‘ten’, which was subsequently reduced to *hambulu*. This historical record suggests that several present-day numeral forms preserve traces of earlier morphological constructions that have undergone phonological reduction over time.

The present-day Kambera data indicate that this older distinction has largely disappeared. Speakers no longer use *diha* as a numeral in ordinary counting or quantification; instead, it survives primarily as an imperative verb meaning ‘to count’. The form *hau*, originally associated with classifier-based quantification, has become generalized as the default numeral for ‘one’ across the numeral system. This development represents a process of historical reanalysis in which a formerly morphologically complex classifier construction has been reinterpreted as an independent numeral.

From this perspective, the contraction of *hau* in forms such as *hangahu* and *harata* is not merely a phonetic adjustment, but the continuation of a historical process involving grammaticalization, lexicalization, and phonological reduction. What appears synchronically as monophthongization may therefore reflect remnants of an earlier morphological system in which numerals, classifiers, and the individualizing prefix *ha-* were more transparently related.

Another important morphophonemic alternation involves the numeral *hambullu* ‘ten’, which surfaces as *kambullu* in compound numerals such as *tailu kambullu* ‘thirty’. Klamer (1998) reconstructs the historical form as *ha-kambulu*, where *ha-* functions as the individualizing prefix and *mbulu* refers to the numerical base ‘ten’. Kapita (1982) similarly records *hakambulu* as the earlier form, noting that it was later reduced to *hambulu*. The contemporary alternation between *hambullu* and *kambullu* therefore appears to represent different stages of historical reduction and restructuring within the numeral system rather than an isolated sound change.

From a synchronic phonological perspective, the replacement of the weak glottal fricative /h/ by the stronger velar stop /k/ can be interpreted as consonantal strengthening, increasing articulatory clarity during rapid pronunciation of frequently used compound numerals. At the same time, the alternation also functions as a morphological marker, signalling that the numeral has entered a compound construction. Consequently, speakers can more easily distinguish the independent numeral *hambullu* ‘ten’ from the bound numeral base *kambullu* used in expressions such as *dua kambullu* ‘twenty’ and *tailu kambullu* ‘thirty’.

Importantly, these alternations occur only in specific structural environments. The forms *hau* and *hambullu* remain unchanged when used independently, but undergo systematic modification when incorporated into larger numeral constructions. This indicates that morphophonemic processes in Kambera are conditioned not only by phonological constraints but also by morphological structure and the historical development of the numeral system. Together, the evidence from present-day speakers, Klamer’s grammatical description, and Kapita’s dictionary suggests that the modern Kambera numeral system preserves traces of an earlier classifier-based counting system whose original morphological composition has gradually become opaque through lexicalization and language change.

It should be noted that Klamer’s (2011) description is based primarily on data collected in Kataka, a Kambera-speaking variety located approximately 70 km from the research site of the present study. This geographical difference is important because the two varieties exhibit certain lexical and morphological differences. In the Kataka variety, the form *hakambulu* (‘ten’), consisting of the individualizing prefix *ha-* and *kambulu* ‘tens’, is still actively used. In contrast, the speakers investigated in the present study consistently use the reduced form *hambullu* as the independent numeral for ‘ten’. This comparison suggests that the modern Kambera spoken in the present research area represents a later stage of phonological reduction and lexicalization, while the Kataka variety preserves an older morphological structure more transparently.

Lexical Replacement and Language Contact

Another important observation in the Kambera numeral system is the presence of lexical variation involving the numeral expressing ‘two’. In the present data, *dambu* is consistently used as the numeral ‘two’ in the unit domain and in expressions such as *dambu njuta* ‘two million’. However, speakers frequently use *dua* in compound numerals referring to tens, hundreds, and thousands (e.g., *dua kambullu* ‘twenty’, *dua ngahu* ‘two hundred’, and *dua rata* ‘two thousand’).

Rather than treating *dua* as a straightforward borrowing from Indonesian, the historical description provided by Klamer (2011) suggests a different interpretation. Klamer reconstructs *dambu* as a merged form derived from *dua* and the classifier *mbua* (a classifier spherical objects). This analysis indicates that *dua* may represent an older element within the Kambera numeral system itself, while *dambu* developed through classifier incorporation. The coexistence of *dua* and *dambu* in present-day Kambera may therefore reflect the preservation of different historical forms rather than simple lexical replacement through language contact.

A similar observation concerns the numeral *njuta* ‘million’. Klamer (1998) notes the form *hau njuta* and remarks that it is comparable to Indonesian ‘juta’. While this similarity may indicate contact with Indonesian, the available evidence is insufficient to conclude that *njuta* is necessarily a direct borrowing. Given that Kambera and Indonesian both belong to the Austronesian language family, the resemblance may also reflect a broader historical relationship or parallel lexical development. Further comparative and historical research is needed to clarify its origin.

CONCLUSION AND SUGGESTION

This study has examined that Kambera employs a decimal (base-10) numeral system in which basic numerals are lexicalized and higher numerals are systematically formed through additive constructions. Structurally, Kambera numerals are predominantly characterized by an [n base] pattern, especially in the formation of tens (e.g., *tailu kambullu* ‘thirty’), whereas the [base n] pattern is restricted to the teen numerals (e.g., *hombullu hau* ‘eleven’). This distribution demonstrates a mixed structural system across different numerical domains.

The study also reveals systematic morphophonemic alternations. The numeral *hau* ‘one’ undergoes monophthongization when combined with higher numerical bases, producing contracted forms such as *hangahu* ‘one hundred’ and *harata* ‘one thousand’. This process simplifies pronunciation while preserving morphological structure. In addition, the form *hambullu* ‘ten’ surfaces as *kambullu* in compound numerals, reflecting context-dependent morphophonemic restructuring associated with numeral compounding. Comparison with the description provided by Klamer (2011) suggests that these forms represent historical developments within the Kambera numeral system rather than isolated phonological changes. While Klamer's data from Kataka area still preserve forms such as *hakambulu* ‘ten’, the present study documents different realizations in the Kambera District, indicating dialectal variation and possible ongoing language change.

The findings further reveal functional differentiation involving the numeral forms *dua* and *dambu*. The form *dua* is predominantly used in the tens, hundreds, and thousands, whereas *dambu* is retained in the unit domain and in expressions such as *dambu njuta* ‘two million’. In light of Klamer's analysis, *dambu* itself is historically a merged form derived from *dua* and the classifier *mbua*. Therefore, the observed distribution is better interpreted as functional specialization within the numeral system rather than straightforward lexical replacement. Likewise, although *njuta* resembles Indonesian ‘juta’, the present data are insufficient to determine whether it represents direct borrowing or reflects shared Austronesian inheritance, and this question requires further historical-comparative investigation.

Taken together, these findings demonstrate that the Kambera numeral system is not a static set of lexical items but a dynamic system shaped by the interaction of internal grammatical organization and external sociolinguistic pressures. More broadly, this study contributes to Austronesian linguistics and linguistic typology by providing detailed documentation of an underdescribed numeral system and by illustrating how universal tendencies, such as decimal structure and additive formation, interact with language-specific features and contact phenomena.

Future research may expand on this study by exploring variation across speakers and generations, examining the role of education and language use in shaping numeral choice, or comparing Kambera with other languages in the region to further understand patterns of typological variation and contact-induced change.

REFERENCES

- Aarts, B., & Haegeman, L. (2021). English Word Classes and Phrases. In B. Aarts, A. McMahon, & L. Hinrichs (Eds.), *The Handbook of English Linguistics* (pp. 103–129). John Wiley & Sons Ltd. Published.
- Allasonnière-Tang, M., & Her, O.-S. (2020). Numeral Base, Numeral Classifier, and Noun: Word Order Harmonization. *Language and Linguistics*, 21(4), 511–556. Retrieved from (10.1075/lali.00069.all). (hal-03018203)%0AHAL
- Barlow, R. (2024). Numerals. In *Reference Collection in Social Sciences* (pp. 1–5). Elsevier. <https://doi.org/10.1016/B978-0-323-95504-1.00005-3>
- Barlow, R. (2025). The Notion of the Numeral Base in Linguistics. *Philosophical Transactions B*, 380(20240208). <https://doi.org/https://doi.org/10.1098/rstb.2024.0208>
- Comrie, B. (2013). Numeral Bases. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Zenodo. <https://doi.org/https://doi.org/10.5281/zenodo.13950591>
- Comrie, B. (2022). The Arithmetic of Natural Language : Toward a Typology of Numeral Systems. *Macrolinguistics*, 10(1), 1–35. <https://doi.org/10.26478/ja2022.10.16.1>
- Creswell, J. (2018). *Research Design: Pendekatan Metode Kualitatif, Kuantitatif dan Campuran*. Yogyakarta: Pustaka Pelajar.
- de Vries, L. (1994). Numeral Systems of the Awyu Language Family of Irian Jaya. *Bijdragen Tot de Taal-, Land- En Volkenkunde*, 150(3). Retrieved from https://babsys-almaprimo.hosted.exlibrisgroup.com/permalink/f/r6qg78/TN_cdi_webofscience_primary_A1994QK50700004CitationCount%0Ahttps://drive.google.com/file/d/1ZWYV2pIVaYMeL79I5mdf8RRKO Cq7zidx/view?usp=sharing
- Ethnologue. (2026). How Many Languages Are There in the World? Retrieved from SIL International website: <https://www.ethnologue.com/insights/how-many-languages/>
- Harrison, K. D. (2007). Endangered Number Systems : Counting to Twenty on Your Toes □. In K. D. Harrison (Ed.), *When Languages Die: The Extinction of the World's Languages and the Erosion of Human Knowledge* (pp. 167–206). Oxford University Press.
- Herdiansyah, H. (2019). *Metodologi Penelitian Kualitatif untuk Ilmu-Ilmu Sosial* (2nd ed.). Jakarta: Salemba Empat.
- Kamus Besar Bahasa Indonesia. (2016). In *kbbi.kemendikbud.go.id*. Badan Pengembangan dan Pembinaan Bahasa, Kementerian Pendidikan dan Kebudayaan Republik Indonesia. Retrieved from <https://kbbi.kemendikbud.go.id>
- Kapita, O. H. (1982). *Kamus Sumba/Kambera-Indonesia*. Waingapu: Gereja Kristen Sumba.
- Klamer, M. (1998). *A Grammar of Kambera*. Berlin, New York: Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110805536>
- Klamer, M. (2005). Kambera. In A. Adelaar & N. Himmelmann (Eds.), *The Austronesian Languages of Asia and Madagascar* (pp. 709–734). London & New York: Routledge.
- Klamer, M. (2011). A grammar of Kambera. In G. Bossong & B. Comrie (Eds.), *A Grammar of Kambera*. Berlin: Mouton Grammar Library. <https://doi.org/10.2307/417682>
- Kridalaksana, H. (2009). *Kamus Linguistik*. Jakarta: Gramedia Pustaka Utama.
- Mahsun. (2010). *Genolinguistik: Kolaborasi Linguistik dan Genetika dalam Pengelompokan Bahasa dan Populasi Penuturnya*. Yogyakarta: Pustaka Pelajar.
- Rukajat, A. (2018). *Pendekatan Penelitian Kualitatif*. Yogyakarta: Deepublish.
- Sanga, F. (2010). Sistem Bilangan Pokok Tradisional dalam Masyarakat Nusa Tenggara Timur: Sebuah Kajian Etnografis. *Linguistika*, 17(33).
- Thornbury, S. (2017). Word Classes and Phrases. In *About Language: Tasks for Teachers of English* (pp. 258–260). Cambridge: Cambridge University Press.
- Winarti, S. (2017). Sistem Bilangan Beberapa Bahasa di Wilayah Papua, NTT, dan Maluku Utara. *Ranah: Jurnal Kajian Bahasa*, 6(2), 235–257. <https://doi.org/https://doi.org/10.26499/rnh.v6i2.450>
- Yiagnigni, A. N. (2026). Morphological and Syntactic Insights into Shupamem Numerals. *ATRAS*, 07(1), 199–216. <https://doi.org/https://doi.org/10.70091/Atras/vol07no01.13>

CURRICULUM VITAE

Complete Name	Institution	Education	Research Interests
Raynesta Mikaela Indri Malo	Universitas Kristen Wira Wacana Sumba	Magister Kajian Bahasa Inggris Universitas Sanata Dharma	Linguistics
Suryani Kurniawi K. L. Kapoe	Universitas Kristen Wira Wacana Sumba	Magister Kajian Bahasa Inggris Universitas Sanata Dharma	Language