A Phonological Study of the Tegem Language

Ahmed Sosal Altayeb Mohammed Ali

SSLAHM002

Main supervisor: Prof. Rajend Mesthrie (University of Cape Town)

External co-supervisor: Prof. Dr. Anne Storch (University of Cologne)

A minor dissertation submitted in partial fulfillment of the requirements for the award of the degree of Master of Linguistics

Faculty of the Humanities

University of Cape Town

2018

DECLARATION

This work has not been previously submitted in whole, or in part, for the award of any degree. It is my own work. Each significant contribution to, and quotation in, this dissertation from the work, or works, of other people has been attributed, and has been cited and referenced.

Signature: Signed by candidate Date: 24 November 2018

Abstract

This study describes the phonological structure of the Tegem language, a little-known Niger-Kordofanian language spoken by around 2000 people in Sudan. The research follows the basic linguistic theory in identifying the segments, investigating their phonotactic pattern, and identifying their functional role in meaning distinction. The study is based on lexical items collected from two Tegem language speakers via wordlists elicitation sessions. That provides the core basis for a detailed foundational description of the phonetic and phonological features of consonants, consonant sequences, vowels, syllables, and tones in Tegem. The description includes a brief account of relevant morphophonemic phenomena such as the voicing assimilation, consonant labialization and noun class sound alternations.

Tegem consonants and vowels phonemes are categorized into two and three categories respectively. The consonants comprises of five obstruent and nine sonorant phonemes out of 20 phones. The vowels include four front, two central, and four back vowels phonemes out of 12 phones. Both the consonants and vowels are very common to occur in a phonologically (and morphologically) complex clusters. There are phonotactic constrains on such sequences conditioned by the environment where they occur. The study explored those sequences as bisegmental structures of adjacent segment sequences. The suprasegmental analysis found six closed and seven open syllables in Tegem where the monosyllabic lexemes of CVC and CVV are the most salient among its 13 syllable types. The research also recognized a pattern of backness ($\pm$back) vowel harmony in the disyllabic nouns and adjectives. The syllable is determined as the bearing unit of the lexical tone in Tegem, i.e. change in the syllable tone is contrastive. The lexical tones include two level tones: high (H) and low (L), and four contour tones: falling (F), rising (R), falling-rising (FR), and rising-falling (RF). The amount of the linguistic data in this study and its description form a solid foundation for further investigation of this poorly documented language.

Table of contents

List of tables and figures

Acknowledgements

I would like to express my gratitude for having had the opportunity to study the Tegem language and write my MA dissertation on it. I still remember the date of 27 August 2018 when for the first time I had a meeting with the Tegem community in the house of the Omar Ali Kunja in Omdurman city, Khartoum. That meeting happened after almost one year of search for Tegem speakers. Special thanks to the Tegem community for their cooperation and support since the beginning of data collection in 2016. I would also like to express my thanks to many people who have supported me and made any contribution to this study. The generous funding of the Mandela Rhodes Foundation MRF, specifically, the Mandela Rhodes Scholarship 2017, made it possible to pursue my MA studies and write this research as partial fulfilment for the degree.

I would first like to thank my supervisors Prof. Rajend Mesthrie (University of Cape Town) and Prof. Dr. Anne Storch (University of Cologne). The door to Prof. Rajend Mesthrie office was always open whenever I needed a consultation about my studies or had anything to discuss about my research which were offered on the spot when we met. My gratitude to him for his continuous support throughout my MA research, for his patience, motivation, and insightful comments. Prof. Anne's thorough and utterly valuable comments arrived in my e-mail with unbelievable swiftness. Very special thanks to her for the unwavering dedication, enthusiasm, and immense knowledge. Their guidance helped me in all the writing phases of this dissertation. I could not have imagined having better supervisors for my MA research.

Besides my supervisors, sincere thanks also goes to Dr. Abeer Bashir, Dr. Fatima Idris, and Dr. Matthias Brenzinger, for their insightful comments and encouragement during the early stages of the research and the during the field work. I would like to thank Prof. Ana Deumert for the helpful suggestions.

I thank my friends and colleagues in the Mandela Rhodes Foundation, the Department of Linguistics (University of Khartoum), the Linguistics Section, and the School of Education (University of Cape Town).

I would also like to thank the Tegem Mek in the Nuba Mountain: Al-Mak Hamdan for his support and instruction to the Tegem community administration in Khartoum to collaborate in the study. I am grateful to the Tegem community leaders and members in Khartoum, particularly, Khartoum State Tegem deputy Mek/ Omar Ali Kunja, Sheikh/ Atroon Mbella Omar, Khartoum State Tegem Mek/ Gadeem Alfaki Juju, Babikir Alkhalifah Abduun, Sluaiman Angatu Saeed, Younis Ali Kunja, and Ali Abass Kafi. I owe my most special thanks to Omar Ali Kunja for his patience during the long elicitation and recording sessions and his commitment to the Tegem people and the Tegem language. Without the precious support of the Tegem community it would not be possible to conduct this research.

Finally, I must express my very profound gratitude to my family: my parents and to my brothers and sisters for supporting me spiritually throughout my studies and my life in general.

Abbreviations and symbols

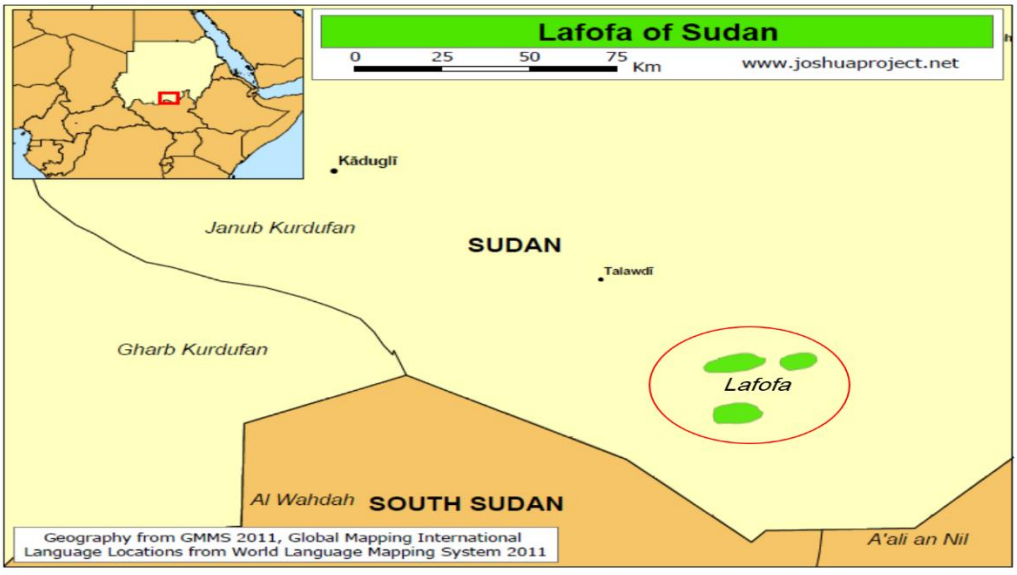
V	vowel
C	consonant (not to be confused with the unspecified phoneme C)
Ç	syllabic consonant
v	verb
n	noun
adj	adjective
adv.	adverb
pron.	pronoun
sg.	singular
pl.	plural
cf.	compare
[ ]	phonemic presentation
/ /	phonemic presentation
~	corresponds to/freely varies with
→	becomes
-	morphological boundary
.	syllable boundary
*	ungrammatical
!	imperative verb

1. Introduction

Tegem (or Lafofa as it is known in the literature) is a Kordofanian language spoken in Sudan in the Nuba Mountains, in the central Eliri range and two hills to the south and east (*see map 1 below*) (Grimes 2003:365-366). This section introduces the Tegem people, their settlements, and language.

1.1. Tegem settlements, neighbouring groups, and livelihood

Manger (1994:35) stated that Tegem people live in eight villages in the Eliri area, of which six are exclusive to them, including Jorrob, Romala, Debbe, Tokoi, Tela and the Lafofa village in the centre of the Eliri hills above Tungaru. There are still some Tegem people at their original settlement Jebel Tekeim[1] 25km to the east of Eliri (Norton & Alaki 2015:63). There are diaspora Tegem people in other towns in Sudan. Many of them were recently internally displaced due to conflict in South Kordofan and settled in regions such as Elobeid city in North Kordofan State and Khartoum State.



Map 1. Tegem (= Lafofa) location in Kordofan (Source: Joshua Project / Global Mapping International "Joshua Project" n.d.[2])

[1] The word Jebel means mountain in Arabic and Tekeim is another pronunciation for the word Tegem.
[2] https://joshuaproject.net/people_groups/12942/SU

In his book on the integration of the Lafofa Nuba into Sudanese Society, Manger (1994) provides quite valuable information about the different groups in the Southern region of the Nuba Mountains and Tegem people in particular. His study categorises the settlers of the Liri (or Eliri) area into four groups: (a) the indigenous Liri area group which is called the Liri Nuba, (b) two migrants Nuba groups including the Talasa and Tegem who moved to the area in 1825-30 and 1880s respectively, (c) then, the Arab groups of Hawazma, Kawahla, and the Jellaba Arabs, and (d) the West African Fellata (Fulani and Hausa) and the South Sudanese groups of Nuer, Dinka and Shilluk as the latest settlers.

In their old home village on Jebel Tekeim, the Tegem are always associated with another neighbouring group called El Amira which is considered by some scholars to be part of Tegem group[3] (Schadeberg & Blench 2013; Stevenson 1962). The main reason for the Tegem migration from the neighbouring Jebel Tekeim in the extreme south-east of the Nuba Mountains (Seligmann 1910:507; Stevenson 1964:84) to the Liri is the unrest in the southern Nuba Mountains during the Mahdia regime 1881-1899. The Talasa migration might be due to a drought in their home area in Tabuli. While the Arabs, West Africans and the South Sudanese groups moved to Eliri to make use of the economic opportunities the area offers.

The diversity of area groups means there are different languages spoken in the Liri. The Sudanese Arabic is spoken by members of all the groups, especially the young generations in the lowland villages and outside the Nuba Mountains region. The other languages include the Tegem language of Tegem group, the Liri people speak Nding language of Talodi-Masakin subgroup, and Talasa (Tumtum) people speak Talasa language which is part of the Kadugli-Korongo (Manger 1994:16; Norton & Alaki 2015:56). Then, the West African groups speak Fulfulde and the Southern Sudanese groups speak the Nilotic languages. It is possible that after the secession of South Sudan in 2011, the latter groups became a minority in the area.

[3] The Tegem people refer to El Amira speakers as El Amira people but they do not deny their old matrilineal relationship to hem which they know from the oral history from the elders.

Tegem livelihood is defined by their history of mobility since their settlement in the Liri area. In the late nineteenth century and early twentieth century, Tegem people home was the mountain village on the Liri mountain. They used to practice what Manger (1994:18) described as 'a quite intensive system of cultivation [of sorghum and peanuts, the latter as cash crops] with the multiple linkages between agriculture and livestock and labour demanding dry terracing'. Then, in 1930 during the Anglo-Egyptian Condominium (1899-1956), the Tegem have been in a conflict with the British and forced to move down and settle in several villages (five of them mentioned above) on the plain around the mountain. The down-movement resulted in a change from intensive hill cultivation to a more extensive cultivation and had greater economic significance for Tegem. After the end of that conflict, another upward movement happened to the mountain village but many stayed in the new villages on the plain (Manger 1994:18).

In the 1960s, the most important change in the Tegem livelihood is their involvement in labour migration to the country's economic center - Khartoum (Manger 1994:19) and later their displacement to Khartoum and North Kordofan areas since 2011 due to the renewed hostilities in South Kordofan region. In their new settlements, the majority of Tegem people work on casual and wage labour.

1.2. The Tegem language speakers

According to the 1956 National Population Census of Sudan, Stevenson (1984:28) estimated the number of Tegem speaking people to 5, 140. Schadeberg (1989) said that there are 3000 Tegem speakers. Currently, the community leaders (represented by their Khartoum State deputy Mek[4]) estimate the number of the speakers to around 2000 in the whole country. The majority are above 30 years of age and live in the Nuba Mountains.

There are no estimations for the total number of the Tegem community which is by far larger than the number of actual speakers since most of the youngsters and some groups speak Arabic only. There are even claims made by the language consultants Omar and Atroon (*see 3.1. below*) that some Tegem people identify themselves as members of the

[4] The Mek is the highest leadership (and administrative) authority in the Tegem (and Nuba) community and has official state recognition in Sudan.

Arab groups such as the Hawazma and Kawahla for different socioeconomic and political reasons, thus, abandon Tegem and adopt Arabic as native language for themselves and their children.

1.3. The language name

Stevenson called the language spoken by the Tegem people the Lafofa language, whereas Schadeberg (1981a) refers to it as Tegem. The reason for the different names is that Schadeberg used the indigenous endonym/glossonym, while Stevenson used the local administrative exonym/toponym and recognized them as names for the language spoken by the community. Norton & Alaki (2015:63) also explained that these two names are used in the literature because 'this language community is referred to either by the name of its central village, Lafofa, or its first village, Tegem [Tekeim]'. Other names used for Tegem are Kidie (Lewis et al. 2013) and Eliri (McDiarmid & McDiarmid 1931). These latter names has not been recognized by the language speakers at all and probably Eliri comes from the area name where they live. Most of the literature refer to the language using the exonym Lafofa (i.e. a name attributed by others).

The language community, represented in their leaders, prefer the native endonyms Tegem [t̪ìkêm] (the self-attributed name), which has no other meaning rather than referring to the community and the language. They consider Lafofa an Arabic name with derogatory connotations -roughly means 'wanderers' referring to the deliberate reluctance of one of the Tegem people called Kukudang[5] to guide the Mahdia delegates to Eliri area where the Tegem community live. The story says that after wondering for one day throughout the wood, he took them back to the starting point which is their camp. Kukudang's excuse was that no one can reach that area because of the strong witchcraft of the Tegem people and immediately the Mahdia leader called them the Lafofa –the people whose witchcraft make others wander. Though, Lafofa has more than

[5] Their oral history also reports that he belongs to a group called Sagarnyh which has kinship relationship with Tegem. They speak Tegem as well. Kukudang has other names: Nasraldin (Arabic) and Daldum (Mahdia leader used this name to refer to him because he was a short man).

one story indicating how it has been assigned to the community by the Arab groups, they consider this one to be more accurate as reported in their oral history. Therefore, Tegem will be used in this study to refer to both the language and the community.

1.4. Language situation and varieties

Persson (1984:1) claimed that a total of over 100 languages are spoken in Sudan (cited by Bashir 2003). One of these languages is Tegem. The lack of reliable data and the complexity of the linguistic situation in Sudan leads to inadequate classifications of the languages in this region. According to Tucker & Bryan (1966), systematic descriptive information on these languages is valuable for linguists, students, national administration, the publication of vernacular and literature textbooks, and education. However, Tegem state of vitality is very low (*see appendix 2*) because of the limited use of the language mostly among elders and some youngsters while most of the children speak Sudanese Arabic only. The language usage domains are very limited to the home and the occasional community/family gatherings.

An early study by MacDiarmid & MacDiarmid (1931) claimed that there are two dialects of Tegem: Eliri (Lafofa) and El Amira. The International Encyclopedia of Linguistics states that Tegem comprises three dialects: Jebel El Amira (El Amira), Jebel Tekeim (Jebel, Tekeim, Tegem), and Lafofa (Grimes 2003:365-366). Similarly, Stevenson (1964:84) mentions that there are variations between the Tegem language spoken at the Lafofa village and the other ones spoken at Jebel El Amira and Tekeim located far south-east of the Nuba Mountains region. Stevenson (1962:126) and Blench (2012:4), argued that El Amira has a reduced morphology and different lexical behavior which confirms that it is related but distinct language from Tegem. They conclude that they are not dialects of the same language as McDiarmid & McDiarmid (1931:154-155) claimed.

Proposals regarding the number of dialects of Tegem range from two to three, and they are mainly based on the areal distribution of the Tegem people and their associated group, the people of El Amira. The reported oral history on the matrilineal kinship

relationship with El Amira group suggests that their language is related to Tegem but it may possibly have evolved from more than one language at the same time as the language morphology suggests (Stevenson 1962; Blench 2012). MacDiarmid & MacDiarmid (1931) who considered El Amira as a Tegem dialect noticed resemblances[6] around 50 years after the majority of Tegem people left Jebel Tekeim (next to Jebel El Amira) to their new location 20km to the south. The information provided by Tegem speakers in a pilot study suggests that there is a degree of vocabulary resemblance, but no mutual intelligibility between Tegem and El Amira speakers.

The other two dialects, Tekeim and Lafofa are varieties of Tegem developed due to the areal distance between the two areas where their speakers are located i.e. Jebel Tekeim to the east and Lafofa village on Eliri hills[7]. The previous varieties has been noticed during the fieldwork for this study. One of the two main language consultants is from Tekeim village and the other is from the other Tegem villages on the plain next to the Lafofa village in Eliri. The variation is in not in the lexicon or grammar but in the pronunciation of some segments which is not a phonemic variation (free variation) and they attribute it to their geographical origins and age differences as well. The speaker from Tekeim is the older and his pronunciation is considered as the old way of speaking and the younger speaker from Eliri speaks the recent variety of Tegem (*see 4.1.1.2.4.* for further details and examples).

In conclusion, further investigation is needed to confirm or disregard the above hypothesis in the previous literature and in this study.

1.5. Language classification

Since most of the Kordofanian languages are poorly described phonologically and others have not been studied at all, more efforts are needed for comprehensive documentation and accurate classification of these languages. One of these languages is Tegem. It is

[6] Their assumption was based on a comparison between 20 words in Tegem (referred to as Lafofa in their study) and El Amira (cf. MacDiarmid & MacDiarmid 1931:154-155).

[7] This concurs with Norton & Alaki's (2015:63) statement referring to Jebel Tekeim variety as follows 'its autonomy as a dialect is suggested by its distance from the other settlements centred on Lafofa village'.

considered ambiguous in terms of genetic affiliation – whether it is a Kordofanian language of the Niger-Kordofanian phylum or not. Most of the previous studies tend to classify Tegem as a remote language from the other Talodi languages (Stevenson 1956 – 7, Tucker & Bryan 1966, Schadeberg 1981a). Norton & Alaki (2015) excluded Tegem from their comparative Kordofanian languages survey but they did not suggest any alternative affiliations. Blench (2012:4) reconstructed the whole Kordofanian family and argues that both Tegem and El Amira can be considered as a separate branch of Niger-Congo. Tegem varieties also are subject to further uncertainty. It is not confirmed whether they are separate languages or not due to the lack of information regarding their structure.

There have been attempts to classify the language in previous literature. Greenberg (1963) places Tegem within the Niger- Kordofanian language phylum which is the largest language family in Africa and the world (*see Appendix 1*), comprising about 1,400 languages spoken in Sub-Saharan Africa. Speakers of this phylum inhabit the largest geographical area compared to the other African language phyla (Olson 2004). Kordofanian languages were classified as a distinct family by Greenberg (1950), and he attached it to the Niger-Congo phylum, thereafter referred to it as Niger-Kordofanian to underscore the Kordofanian part. Thelwall & Schadeberg (1983:226) indicate that all five Kordofanian languages of the Niger-Kordofanian family in the Nuba Mountains (Talodi, Tegem, Heiban, Rashad, and Katla-Tima) descended from one language spoken by the original settlers, which split into four and then five groups, the last split being between Talodi and Tegem. Recent researches by Schadeberg and Blench (2013) argued that the genetic relations of the above mentioned five Kordofanian languages to each other and to Niger-Congo remains debatable.

Regardless of the previous controversy, the classification of Tegem adopted in this study follows Schadeberg (1981b). He classifies Tegem as Niger-Kordofanian, Kordofanian, Talodi as shown in figures 1 and 2.

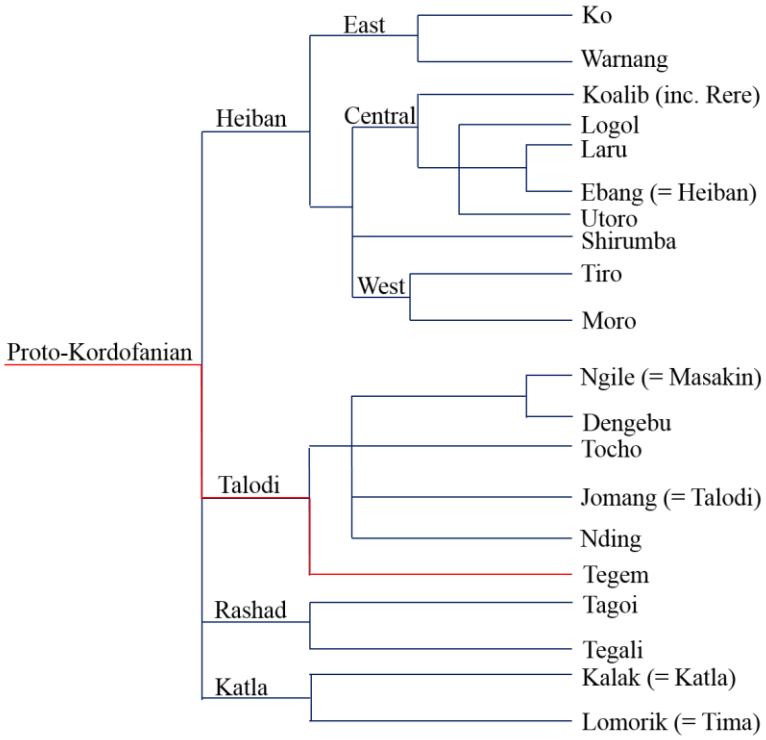


Figure 1. *Proto-Kordofanian Family tree (Schadeberg 1981b).*

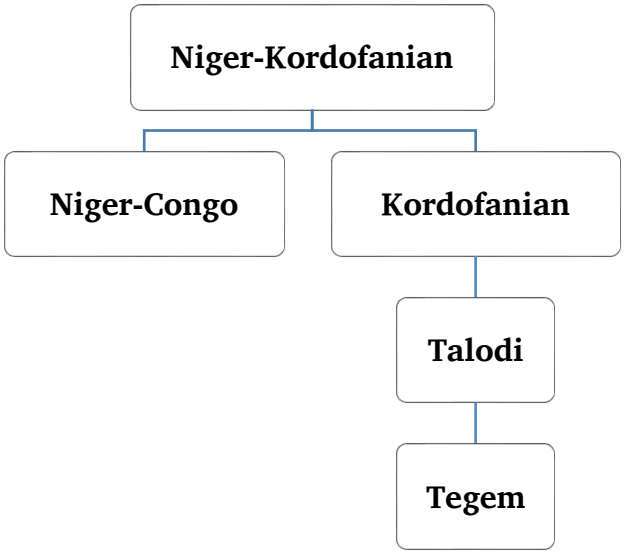


Figure 2. *Tegem, sub-branch within the Niger-Kordofanian family tree* (Williamson 1989:8).

2. Literature review

This section includes information and review of the available sources on Tegem and other Kordofanian and Sudanese languages. Then, some comparative data from Norton & Alaki (2015) is presented to show a degree of resemblance on segmental level between Tegem and a reconstructed structures of what they called a proto-Talodi language. The last two sections briefly explain the research questions, objectives, and significance of this study.

2.1. Sources on Tegem and other Nuba Mountains languages

Brenda Seligmann's 1910/11 ethnographic studies on the Nuba Mountains include the earliest data on Tegem. In 1915-19 Meinhof's re-presented that data in his work. During a three months trip to the Nuba Mountains between 1930 - 1931, the missionary couple Phoebe and Donald McDiarmid collected, as part of a linguistic survey, a concise list of what they called Eliri (referring to Tegem) and El Amira which they considered as a 'dialect' of Eliri (MacDiarmid & MacDiarmid 1931).

Stevenson (1956-7) documented, published and collected some materials on the languages of the Nuba Mountains, although the Tegem data is considerably less than what has been collected for the other languages, such as Koalib, Otoro, Tira, Moro, Mesakin, Katcha, Kadugli, Miri, Krongo, Liguri, Temein, Katla, Nyimang, Kadaru, and Ghulfan. Stevenson's data on Tegem is part of extensive files published by Blench (1997).

The only more recently published data and descriptive notes on Tegem are in Schadeberg (1981a), which contains the author's word list (about 200 items) collected during late 1974 and early 1975. It contains an additional list collected by Robin Thelwall which differs from it in some transcription aspects. In the second volume of his Survey of Kordofanian, Schadeberg (1981a) provided phonological and morphological notes on six Talodi group languages: Jomang, Nding, Dengebu, Ngile, Tocho, and Tegem for comparative purposes. His phonological notes on Tegem will be utilized in this study in order to understand the segmental and morphophonemic patterns of the language.

Norton & Alaki (2015) comparative-historical analysis of Talodi languages includes Schadeberg's (1981a) data on Tegem and other nine Talodi languages: Torona[8], Lumun, Tocho, Acheron, Tuwal (Dengebu)[9], Dagik, Daloka (Ngile), Tasomi (Jomang), Nding. The result of their lexicostatistic comparison of the various Talodi languages shows that the lowest similarity score, ranged between 23-27%, is for Tegem in comparison to the similarity between the other nine languages where an average of minimum 52% has been recorded. Norton & Alaki attempt to reconstruct a proto-Talodi language from the existing wordlists and data they collected for the study. Though, their work excludes Tegem from the Talodi language, it provides important sources for the Tegem neighbours' languages.

Apparently, with reference to that literature, further studies are essential to explore the phonological structure of Tegem. Several reasons show the significance of such research. There are few publications on the grammars of other relevant languages. The existing literature comprises of word lists and surveys and there is no detailed phonological description for the language. The existing data sometimes lacks in quality and insight because some of the early word lists were collected by researchers who are not linguists. Blench (2012) described some of those word lists on Tegem as constituting an 'ill-transcribed database'. In accordance with linguistic theory, when linguists try to describe a language, one of the very first priorities is to determine the phonological inventory of the language segments. This requires linguistic research objectively devoted to explore the phonological structure of the language. Therefore, the current study is an attempt to provide an informed analysis to fill some of what Stevenson (1962:127) called, many remaining gaps, especially with reference to the Tegem language which lacks even the very basic description.

[8] Norton & Alaki (2015) identified this language as extinct and the data for their study was collected from the language last speaker in 2012 in Khartoum.

[9] The names in brackets corresponds to the ones used to refer to the same languages by Schadeberg (1981b).

2.2. Thesis statement

There is insufficient linguistic research in Sudan when we take into account the number of languages spoken languages there, and their development and endangerment status (*see Appendix 2*) (Lewis et al. 2015). Some of the early works on the Sudanese languages include Bender (1975, 1983, 1989), Bendor-Samuel and Hartell (1989), Doornbos and Bender (1983), Schadeberg and Bender (1981), and Stevenson (1984). This study describes some of the main aspects of the Tegem language grammar. This description aims at providing an analysis of the phonological system of Tegem, which would help to set up foundational basis for further description and documentation of the language.

Since Schadeberg's (1981a) survey is the only descriptive work on the language, it is convenient to consider the other phonological studies that have been done on some Niger-Kordofanian languages. This applies particularly to the Talodi languages, as they belong to the same language group with shared features. Knowledge of previous studies will benefit the present analysis.

When compared with Talodi languages (to which Tegem belongs) there is evidence of sound correspondences. Table 1 shows the recurring sound correspondences between proto-Talodi and Tegem roots (Norton & Alaki 2015:70).

Table 1. *Sound correspondences between proto-Talodi and Tegem (roots) (Norton & Alaki 2015:70).*

i:i (±ATR?)	eye fire tail water fly sew wash/wet black clean rough
ɛ:ɛ	bird ear mother skin thing give red two
ə:ɛ	knee leg neck snake tooth
ə:i	bird lake tongue eat
a:a	finger head name road root woman wide
ɔ:ua	mountain word/speech wet stand
u:u (±ATR?)	bone dog feather fog guts hand:wing river squeeze warm white

Another example for the relations between Talodi languages and Tegem is the proto-Talodi medial [a] ~ final [ɛ] alternation comes from items with word-final [ɛ] that have a corresponding [a] in Tegem language (Norton & Alaki 2015:106). See table 2 below:

Table 2. *Final medial-vowel alternation in proto-Talodi and Tegem (Norton & Alaki 2015:106).*

Gloss	proto-Talodi	Tegem
'cow'	*w-aɪ/*k-ie	(aɟi/k-aɟi)
'tongue'	*t̪ʊ -ləŋɛ /lə-	(líáŋ-i, sg = pl)
'river'	*t-ʊwɛ /n-	(t̪-uwaa-jt̪ /r-)
'bone'	s-əmɛ /m- (Dagik)	(t-ʊam /m-)
'split'	gitti-d̪ɛ (suffix Tasomi)	(ridiɛn-tạ̪ŋ)

Therefore, the existing literature provides a secondary data source, besides this research word list, on phonology (Schadeberg 1981a), comparative wordlists (Schadeberg 1981a & Blench 1997), Talodi languages phonology (Schadeberg 1981a, Norton & Alaki 2015), Sudanese languages classification and the scarce available written information on Tegem in general.

2.3. Research questions and objectives

This research aims to answer one question: What are the general characteristics of the phonological system of Tegem? In order to answer this question, the current study intends to fulfil the following objectives:

(a) To provide a basic description of the phonology of Tegem. Hence, it gives a systematic description of the language segments (consonants and vowels), syllable structure, surface tone system, and some morphophonological features.

(b) To contribute to the existing literature on the Nuba Mountains languages with further linguistic data from a less know language.

2.4. Significance of the research

The subject of this research is the phonology of the Tegem language. The principal aim is to provide an informed description of Tegem phonology: sound system and phonemic structure.

The Tegem language includes several recently noted linguistic properties that are distinct from the other Nuba Mountains languages. Better knowledge of these noteworthy patterns would therefore contribute to the existing linguistic data sources, support the efforts of documenting the Sudanese languages, and contribute to historical classification.

This study's significance is patent through the following facts:

(a) This research is the first linguistic work devoted principally to study the phonology of Tegem.

(b) This study provides a substantial preliminary basis for further grammatical analysis of Tegem.

(c) It contributes to the debate on the Tegem language classification and dialects by presenting an informed work on its phonological structure.

3. Fieldwork and methodology

This part of the dissertation discusses the study approach, provides information on the language consultants and the fieldwork, and the phonological analysis criteria used in the study.

The present study adopts a descriptive-analytical method. It utilizes informed linguistic concepts referred to in the literature as "Basic (descriptive) Linguistic Theory" (Dixon 2009 & 2012). These concepts are used to describe and analyze the phonological system and identify the segmental phonemes and their variants in Tegem. The key result is an initial phonological description, which contribute to our understanding of the phonemic patterns and behavior of Tegem linguistic structures.

The description in this study is synchronic describing the language phonological characteristics at this point in time. However, in few instances there are references to data from previous studies and analysis for clarification purposes.

3.1. Language consultants

The main language consultants (informants) are two adult males: Omar (47) and Atroon (55). The major contribution was from Omar[10]. There are few contributions from the following Tegem speakers: Babikir (48), Younis (31 years of age), Slyman (32), Gadeem (62), and Ali (65). They are native speakers of Tegem. All the language consultants were born in the Nuba Mountains, but are now resident in Khartoum. They frequently visit and receive their relatives from the Eliri hills of the Nuba Mountains (in the Tegem area). Although they speak Arabic, which is the contact language, they continue using Tegem at home, in family gatherings, and ceremonies in Khartoum.

The linguistic data collected from these language consultants is by means of interviews through elicitation and sound recording sessions in Khartoum. These consultants were

[10] He is also the Deputy Mek of the Tegem community in Khartoum and the Omda (chief) of the community in the area where he lives in Omdurman (*these information disclosed upon his consent*).

nominated by their community as the best speakers of Tegem in Khartoum and had the consent and recommendation of their Meks in Khartoum and the Nuba Mountains.

3.2. Research ethics clearance

This study has been approved by the Research Ethics Committee at the Linguistics Section. The Tegem leadership represented in their two Meks in the Nuba Mountains and Khartoum gave their approval to do a research on the Tegem language. The ethical clearance consent form has been translated and the participants understood the informed consent statement in Arabic before their representative sign it. The language consultants gave their consent to voluntarily participate in the study. They agree that their responses can be used for research and education.

3.3. Fieldwork and instruments

This study uses word lists and, with less proportion, naturally occurring reordered utterances (during the elicitation sessions) as its data sources.

The data was collected for this research during a period of about four months from 2016 -2018 at the Department of Linguistics of the University of Khartoum. The first elicitation sessions took place in the period between September and October 2016. Then, another interview was held in January 2017 and the final fieldwork trip[11] was in April and May 2018.

The following word lists are used as main instrument for data elicitation: Swadesh wordlist, categorized by semantic field (Bowern 2007), SIL comparative African wordlist SILCAWL (Snider & Roberts 2006), Schadeberg's 1981a word list, and Ibadan word list of 400 basic items (Akinkugbe 1978). More than 900 items (excluding the utterances) has been collected. All the elicitation sessions are recorded. Then the elicited items are recorded again in recording only sessions. They comprise of about 37 hours recordings

[11] That third trip happened during my MA studies at the University of Cape Town, South Africa (2017-2018) and the purpose was to collect more data for the MA dissertation completion.

for the 14 elicitation sessions and 7 hours audio recording of the 3 recording sessions. All the participating speakers have been recorded in the elicitation session and only the main consultant Omar is recorded in the recording sessions. The collected data is sorted, described and analyzed. The research utilized for sorting and data management purposes include Toolbox, Phonology Assistant, Praat and other computer audio software.

3.4. Analysis approach

The analysis criteria adopted in this research explore the phonetic, functional, and distributional aspects of the language phonological segments. Taking into account the study objectives, the research (a) identifies the speech sounds and the sound structure (b) makes generalizations about the distribution of the sounds in the language, and (c) determines its function as a distinctive segment in contrasts to other similar segments (*see 4.1.1.*). Thus, patterns of contrast (in both syntagmatic and paradigmatic dimensions), alternations, and phonotactics are taken into consideration.

4. Data Analysis

4.1. Tegem Consonants

4.1.1. The consonant inventory

Tegem has 20 consonants, 19 consonant sequences, and 11 geminated consonants phones (*see Table 3 below*). They are divided into two major classes, obstruents and sonorants (Schadeberg 1981a:76). The obstruents class includes stops and a fricative and the sonorants includes nasals and approximants (consisting of liquids and glides). In Tegem, in the former class there are nine stops and one fricative. Voicing variation is not phonemic in Tegem.

Table 3. *Consonant inventory of Tegem*

		bilabial	**dental**	**alveolar**	**retroflex**	**palatal**	**velar**	**glottal**
stops		p, b	t̪, d̪		ʈ	c, ɟ	k, g	
fricative								h
nasals		m		n		ɲ	ŋ	
liquids	central			r	ɽ			
	lateral			l, ɬ				
glides		w				j		
stops + glide		bw, pw	t̪w		ʈw	ɟw, cw	kw, gw	
nasal + glide		mw						
liquids + glide				rw, lw				
nasals + stops		mb	nd̪		nɖ	nɟ	ng, ŋg	
nasal + stop + glide		mbw			nɖw			
geminates		bb, mm	t̪t̪, d̪d̪	rr, ll, ɬɬ	ʈʈ	cc, ɟɟ	kk	

Three descriptive-analytical phonological criteria are utilized in describing the consonants and analyzing their phonemic patterns. They are constructed on the phonetic, the distributional, and the functional aspects of the sounds (cf. Berg 1993). The phonetic

criterion determines the distinctive phonetic characteristics in which the consonants varies. The distributional criterion identifies the occurrence and distribution pattern of the phones across the three word positions: initial, medial (intervocalic), and final. The functional criterion refers to the meaning contrastive function of the sounds which is important aspect for phonemic distinction.

In the following paragraphs, the consonants are divided into two groups: the obstruent consonants and the sonorant consonants. Each section provides detailed description of every consonant, its distribution, combination with other consonants, and its phonemic pattern.

4.1.1.1. Obstruent consonants

4.1.1.1.1. The bilabial stops [b] and [p]

(a) The voiced bilabial stop [b]

The voiced bilabial stop [b] has been attested in all three positions - initial, medial, and final, e.g. *bír* 'bow', *bìáábúú* 'noon', *ʈɔ̂bˀ* 'stick/pen'. Word-finally, it occurs as an unreleased bilabial stop [bˀ] without any conditioning from the preceeding vowel, e.g. *ʈ̪wœ́bˀ* 'road', *mɔ̌bˀ* 'animal pl.'. There are no restrictions on the preceeding and/or the following vowels, i.e. it may occur adjacent to any of the Tegem vowels.

The bilabial [b] may occur geminated in the medial position [bb], e.g. *bíɛ́bbúlíl* 'long', *bébbú* 'dog'. It may occur as a cluster with [m] in the initial and medial positions [mb], e.g. *mbìì* 'pull', *ímbíí* 'brother in law', *bíámbál* 'big'. That cluster may be followed by the glide [w] initially and medially as [mbw], e.g. *ḿbwʌ́kˀ* 'suckle', *íbɛ̂mbwɛ̌ŋ* 'child'. The stop consonant [b] may also form a cluster when followed by the glide [w], e.g., *œ́bwɛ̌nɖɛ́h* 'who', *bìʈûbwǐr* 'blacksmith', *bwʌ̌h* 'bark (dog)'. Further analysis for these sequences, their variants, and phonemic status is in *4.1.1.3.2., 4.1.1.3.3., 4.5.2.1., 4.5.2.2.,* and *4.5.2.2* below.

(b) The voiceless bilabial stop [p]

Although, it is most frequently attested in the initial position, the voiceless bilabial stop may occur in all three positions - initial, medial, and final- *pwà* 'sugar cane', *pèppwʌ̀* 'goat (young male)', and *ʧɔ̂pˀ* 'stick/pen'. The voiceless bilabial [p] varies freely with its voiced counterpart [b] in all these positions, e.g., *píbírbwɛ̌ŋ* ~ *bípírpwɛ̌ŋ* 'narrow', *kɛ̌pˀ* ~ *kɛ̌bˀ* 'wall'. Similar to [b], word-finally it occurs as unreleased voiceless bilabial stop [pˀ] without any conditioning from the preceeding vowel.

Unlike [b], [p] is not attested as a sequence with [m]. Although, it is not common, there is no contrast when [mb] is replaced by [mp], e.g., *ímbíŋ* ~ *ímpíŋ* 'niece pl.'. The preference of the former sequence [mb] may be due to their shared voicing feature as voiced bilabials [b] and [m] (*cf. 4.5.1.1.*).

The consonant cluster [bw] ~ [pw] occurs word-initially and intervocalically and the voiceless variant [pw], is commonly attested in the initial position, e.g., *pwà* ~ *bwà* 'sugar cane'. They are free variants.

The term underspecification suggests 'that information should be omitted from underlying phonological representations' in phonological features theories, i.e. the features of a segment can be minimally specified or underspecified (Crystal 2008:501-502). An example of the underlying phonological representations could be the voicing feature. Thus, the voicing feature of [b] and [p], e.g., *kɛ̌pˀ* ~ *kɛ̌bˀ* 'wall', is underspecified in Tegem and should be referred to using an unspecified symbol. Following Reh (1996), the unspecified phoneme /B/ will be used to represent the underspecified voiced and voiceless phones [b] and [P] in Tegem.[12]

[12] In Schadeberg's (1981a) notes, the voiced [b] varies with a voiced fricative [β] in the intervocalic position. Such a variation has not been attested in the present data.

4.1.1.1.2. The dental stops [t̪] and [d̪]

(a) The voiceless dental stop [t̪]

The voiceless dental stop [t̪] occurs in all the three positions and may be preceded and/or followed by a wide range of vowels. In the final position, it is unreleased [t̪˺] and tends to be preceded by one of the four vowels [i], [ɪ], [e], or [ʌ], e.g. *t̪úàŋít̪˺* 'nose bridge', *lìgìt̪˺* 'face', *t̪ɪ̀ɪ̀ŋèt̪˺* 'heavy', and *kǎt̪˺* 'cave'. In the medial and final positions, the preceding environment is restricted to the following four vowels: front [i], [ɪ], [ɛ], and central vowel [æ].

In Tegem, the voiceless dental stop [t̪] may be geminated in the medial position. It is most frequent for the geminated [t̪t̪] to be preceded by the vowels [e], [ɛ], or [æ] (short and long) such as in *ndwèt̪t̪àŋ* 'cut! (with a knife)', *mìét̪t̪éémʌ́ʌ̀* 'kind of beans pl.', and *lwɑ̂t̪t̪ɑ̀k˺* 'mat'. The non-geminated [t̪] is always preceded by one of the four short front vowels [i], [ɪ], [ɛ], or [e], e.g., *kít̪ɛ̂ɲ* 'laugh', *kít̪ít̪˺* 'earth', *bít̪ílìl* 'green', *pɛ̀t̪ɛ́ɛ́rín* 'two', and *lét̪ʌ̀ɬǐh* 'bring (from far)'.

The dental stop [t̪] forms a cluster with the glide [w] → [t̪w]. The sequence [t̪w] occurs most frequently in the initial position, e.g. *t̪wàà* 'river' and found medially in few words: *t̪íít̪wʌ̀d̪ɔ̀ŋ* 'rotten' and *bìɛ̌t̪wìl* 'ant'. This cluster is followed by the central short vowels [æ], [a], [ʌ] or the short and long front variants of [i], [ɪ], [ɛ] (*see 4.1.1.3.5.* & *4.5.2.4.*).

(b) The voiced dental stop [d̪]

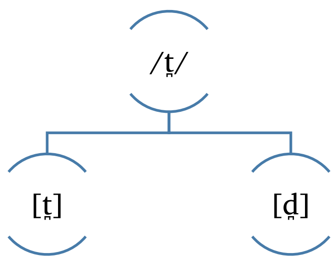
The voiced dental stop [d̪] occurs intervocalically and word-finally. It can only be preceded by a nasal consonant, back vowel(s), or central vowel(s) only e.g., *t̪ìnd̪ìlìɛ́* 'pierce', *kúd̪úk˺* 'go!', and *lûɬɬíd̪˺* 'very close'.

The geminated consonant [d̪d̪] occurs in intervocalic position after back vowels, e.g., *ɬɔ́d̪d̪ɔ́* 'underneath'. It may be combined with [n] to form the consonant sequence [nd̪] in

the initial and medial positions, such as in *nd̪úlúh* ‘smell!’ and *ŋ́ɟɛ̌nd̪ʌ̂ŋ* ‘sing!’ (*see 4.1.1.3.6. & 4.5.2.5.*).

Two sounds are said to be in complementary distribution when the environment in which they occur is mutually exclusive, i.e. their environment is never the same. It appears from the above analysis that [t̪] and [d̪] are in complementary distribution - they cannot be preceded by the same vowel. Intervocalicaly, after the back vowels and the central vowels, /t̪/ is realized as voiced dental stop [d̪]. Elsewhere, it is realized as voiceless dental stop [t̪]. Since [t̪] is more frequent and its distribution is less restricted, we may conclude that /t̪/ is considered to be the phoneme and it has two allophones as follows:

Rule (1)



/t̪/ → [d̪] / [+back V] or [+central V] _ [V], otherwise [t̪]

4.1.1.1.3. The retroflex stop [ʈ]

The retroflex stop [ʈ] occurs in the initial and final positions, e.g., *ʈóók*̚ ‘arrows bag’. In the final position it is pronounced as unreleased consonant, e.g., *ŋʌ̌ʈ*̚ ‘dirt/stealing’.

In the medial position it only occurs as geminated [ʈʈ] and is preceded by one of the long front vowels [ii], [ee], or short [ɛ], e.g., *ʈííʈʈwʌ̀h* ‘many’, *kìlíɛ̌góm ʈɪ̀ʈéèʈʈìlít̪ílìt̪*̚ ‘six’, and *ʈɛ̌ʈʈɛ̀ɛ́* ‘calabash’.

It may be combined with the glide [w] → [ʈw] and this is the only environment where [ʈ] may occur word ininitially, e.g., *ʈwàáì* ‘tying’ and *ʈwíí* ‘forehead’. In the medial position, the first segment [ʈ] of the sequence [ʈw] may be geminated, e.g., *ʈííʈʈwʌ̀h* ‘many’ (*see 4.1.1.3.8. & 4.5.2.7.*).

The retroflex stop [ʈ] may come after the voiced alveolar nasal [n] in the initial and intervocalic positions. Its voicing feature assimilates to the voiced [n] and becomes [ɖ] to form the consonant sequence [nɖ] (*see 4.5.1.1.*). The [nɖ] sequence may be followed by [w] in the initial position to form the cluster [nɖw]. There are more details about these sequences in sections *4.1.1.3.9.* and *4.5.2.8.* below. The language consulatants Omar and Atroon report that there is no contrast between the sequence [nɖ] and the devoiced single phone [ʈ] in the word initial position, *nɖɛ́bɛ́r* ~ *ʈɛ́bɛ́r* 'cut! (with hand)'. Likewise, the cluster [nɖw] and the devoiced sequence [ʈw], e.g., *nɖwáàí* ~ *ʈwáàí* 'tie!', and *nɖwèn̈ʈ̈ʌ̀ŋ* ~ *ʈwèn̈ʈ̈ʌ̀ŋ* 'cut! (with knife)'. However, it is likely that a detailed morphological analysis may show the function of [nɖ] and [nɖw] as imperative prefixes in contrast to [ʈ] and [ʈw] sequences when they occur in an utterance (*cf. 4.5.1.1. & 4.5.2.8.*).

Unlike the other Tegem stops, the retroflex /ʈ/ does not have voiced counterpart and it is a distinct phoneme.

4.1.1.1.4. The palatal stops [c] and [ɟ]

(a) The voiceless palatal stop [c]

Although it is more frequent word-initially, it may also occur intervocalically. The voiceless palatal stops [c] varies freely with its voiced couterpart [ɟ] in these positions, but more frequently word-initially, e.g., *cîĭ* ~ *ɟîĭ* 'wire' and *ɟìír* ~ *cìír* 'bast of a tree', *cɔ́cɔ́ŋ* ~ *cɔ́ɟɔ́ŋ* 'story' (*see 4.1.1.3.11.*).

The palatal stop [c] may be combined with the glide [w] in the initial and medial positions and becomes [cw]. Again [c] and [ɟ] varies freely when they are part of the cluster [ɟw] ~ [cw], such as in *rwʌ̂ŋ kɔ́łɔ́ɔ̀cwìk*˥ ~ *rwʌ̂ŋ kɔ́łɔ́ɔ̀ɟwìk*˥ 'think!' and *cwîû* ~ *jwîû* 'rubbish'. The consonant [c] may also be geminated in the medial position such as in *cɔ́ccɔ́ŋ* 'sixth child'.

(b) The voiced palatal stop [ɟ]

The palatal stop [ɟ] occurs in the initial and intervocalic positions. It has been attested

unreleased word-finally in one word: *kwɨ́ɟ̚* ‘room’. It can be followed or preceded by a wide range of front and back vowels.

The palatal [ɟ] can be preceded by a nasal in the initial and medial positions [nɟ] (*see 4.1.1.3.12.*). This sequence is not allowed with the voiceless palatal stop [c]. In the medial position, the sequence [nɟ] is a result of sound alternation process between the consonants [ɲ] and [n]. When the phones [ɲ] and [ɟ] are combined, the palatal nasal [ɲ] changes to the alveolar [n], i.e. /ɲC/[13] → [nɟ] *kɛ̌ɲ + ɟéŋ → kɛ̌nɟéŋ* ‘give him/her’, and *bɔ̌ɲ + ɟéŋ → bɔ̌nɟéŋ* ‘give me’. This cluster has been attested only in the imperative verbs. The sequence [nɟ] is analysed as a syllabic nasal followed by a stop in section *4.3.5.* below. The consonant [ɟ] may also form a cluster when it is followed by the bilabial glide [w], such as in *ngúɟwìn̪ʌ̀ŋ* ‘wash’ (*see 4.1.1.3.11.*).

The palatal stops [c] and [ɟ] are in free variation in both the initial and medial positions. While [c] has been attested more frequently in the initial position, [ɟ] has been attested in both initial and intervocalic positions. The unspecified uppercase symbol /C/ is considered as the phoneme representation of the two underlying voiced and voiceless representations [ɟ] and [c].

4.1.1.1.5. The velar stops [k] and [g]

(a) The voiceless velar stop [k]

The voiceless velar stop [k] occurs word-initially and finally. In the final position, it only occurs as unreleased stop [k̚] without any conditioning from the preceding vowel (cf. Schadeberg 1981a).

In phonology, free variation refers to the substitutability of one phoneme for another in specific environment, without change in the meaning- allophonic variantaion (Crystal 2008:20-198). In some Tegem words there is free variation between the unreleased velar stop [k̚] and the unreleased dental stop [t̪̚] in the final position, such as in *ɬùùɬûk̚* ~ *ɬùùɬût̪̚* ‘cotton’ and *ʈɨ̀ɲùk̚* ~ *ʈɨ̀ɲùt̪̚* ‘small bat’. Since [k̚] and [t̪̚] does not contrast word-

[13] The uppercase C refers to the phonemic representation of [ɟ] as explained in the following paragraphs.

finally, they are allophones in that position. The dental stop is more salient in this position and it will be considered as the phoneme.

The velar stop [k] only occurs geminated [kk] intervocalically and may be preceded and followed by different front, central, or back vowels e.g., *kìkkùɟùr* 'narrow', *kǐkkáàlɛ́* 'wrong', and *kêkkáh* 'four'. It has not been attested as a non-geminated in the intervocalic position.

It occurs as a sequence when followed by the bilabial glide [w]. This consonant cluster [kw] occurs in the initial position only such as in *kwáá* 'sand/soil' and *kwʌ́* 'beard' . This sequence can be followed by any vowel except the close back vowels [u] and [ʊ] which may only occur after the single phone [k].

(b) The voiced velar stop [g]

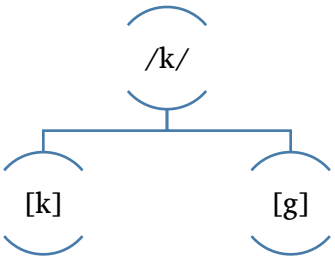
The velar stop [g] occurs in intervocalic position only and cannot be geminated, e.g., *lìgìt̪˥* 'face' and *kìgɛ̌l* 'cover'. It is not restricted to a specific set of following or preceding vowels.

The voiced velar stop cluster [gw] occurs intervocalically only, for example, *ɛ́gwɛ́h* 'bark of a tree pl.' and *kògwɛ́ɛ̀* 'wing'. The consonant cluster [kw] occurs word-initially and [gw] medially. Sections *4.1.1.3.14.* and *4.5.2.9.* below provide further analysis on these sequences.

The velar [g] forms a cluster in the initial position when preceded by the alveolar nasal [n] → [ng] and intervocalically by the velar nasal [ŋ] → [ŋg], such as in *ngêɲɟìk˥* 'spit!', and *cèŋgìmèè* 'important person'. This consonant cluster has not been attested word-finally (*see 4.1.1.3.15. & 4.5.2.10*).

The velar stops [k] and [g] occur in mutually exculsive environments. They are in complimentary distribution because the voiced [g] occurs intervocalically whereas the voiceless [k] occurs elswhere. Therefore, we may conclude that the voiceless /k/ is the phoneme with two allophonic realisations as follows:

Rule (2)



/k/ → [g] /[V]_[V], otherwise [k]

4.1.1.1.6. The glottal fricative [h]

The glottal fricative [h] occurs in the final position only and its absence is not contrastive, e.g., *múh* ~ *mú* 'bowl pl.' and *ʈùh* ~ *ʈù* 'pod lake', *ɟìh* ~ *ɟì* 'eye pl.'. It is deleted when a suffix or another word is attached to a word ending in [h], e.g., *mû-ì* 'these are bowls'; *ʈŭ-ì* 'this is a pond/lake'; and *ɟìì-lɨ̌h* 'eyebrow pl.'. It does not form a cluster with any of the consonants in Tegem.

The glottal fricative [h] is restricted to the final position and if dropped it does not change the meaning of the word, e.g., *ríh* ~ *rí* 'branch pl.'. It is deleted when followed by another lexeme, such as in *ríh* 'branch pl.' when followed by the demonstrative lexeme *gɔ́* → *rí-gɔ́* 'these branches'. Thus, as a non-contrastive (optional) consonant it is not part of Tegem phonemes.

4.1.1.2. Sonorant consonants

4.1.1.2.1. The bilabial nasal [m]

The bilabial nasal [m] is the most frequently used consonant besides [l] in Tegem. It has been attested in all three word positions. It may be preceded and followed by various vowels intervocalically and finally.

The consonant [m] may be geminated intervocalically, for example, *mĕmmɛ̀ɛ́* 'calabash pl.', *púmmíl* 'all', and *ŋ́gèmmì* 'throw away'. This bilabial consonant may form a cluster

with [w] in the initial position [mw], e.g., *mwʌ̀kˀ* 'ankle pl.' and *mwíí* 'forehead pl.' (*see 4.1.1.3.17. & 4.5.2.7.*).

The bilabial nasal /m/ is a phoneme, because it is widely distributed across all three word positions and contrasts with other phonemes in identical environments, e.g. *ìmĭ̀l* 'my husband', *ìlĭ̀l* 'our husbands', *kwʌ́m* 'Basham tree', *kwʌ́l* 'air', and *ʈwʌ̀kˀ* 'ankle', *ʈwʌ̀m* 'bone' (*see Table 14*).

4.1.1.2.2. The alveolar nasal [n]

The alveolar nasal [n] only occurs in word initial, medial and (semi)final positions. In the final position it varies freely with [l] or [ŋ] (cf. Schadeberg 1981a). For example, *bŏnɟéŋ* ~ *bŏnɟén* 'give me', *pɛ̀ʈ̤ɛ́ɛ́rín* ~ *pɛ̀ʈ̤ɛ́ɛ́ríŋ* 'two', *pʌ̀ɖ̤ʌ̂ʌ́l* ~ *pʌ̀ɖ̤ʌ̂ʌ́n* 'three'. The nasal [n] has never been attested before any vowel and can only be followed by a stop consonant in the initial and medial positions but there are no restrictions on the preceding vowel, e.g., *ńɟɔ̀ɔ́* 'stab!', *ʈ̤ûnɖ̤ʌ̂* 'waste dump', and *ʈ̤ìngìmɛ̂ɟí* 'bury!'.

It is always part of a cluster sequence with the stops, i.e. [nɖ̤], [nɖ], [nɟ], and [ng] in the initial and medial positions. The sequence [nɖ] is a combination of the alveolar [n] and the voiceless retroflex stop [ʈ] where the latter is replaceed by [ɖ] after the voiced nasal [n] as a result of voicing assimilation process from [nʈ] to [nɖ] (*see 4.5.1.1. & 4.1.1.3.*).

In the final position, the main two language consultants prefer the variant [l] when it varies freely with [n] and [ɲ], and they think it is the most common at the present. The lateral [l] is also more salient in the collected data in (semi)final[14] position than [n], however, their phonemic status word-finally is determined below (in *4.1.1.2.4.*) as part of the velar nasal [ŋ] analysis as it varies with [n], [l], and [ɲ] in the final position. The alveolar nasal [n] may occur in all three word positions and it contrasts with other consonants in the initial and medial positions in near minimal pairs, e.g., *ńɖʌ́ʌ̀* 'take!' and *ʈʌ́ʌ̀* 'bean'; *ʈ̤ûnɖ̤ʌ̂* 'waste dump' *ʈ̤úúɖ̤ʌ̀* 'summer' . Though, it is unique as it is the only

[14] This is the position just before the final segment in the long words.

conosonants that does not allow vowels after it, /n/ is a distinctive phoneme word-initially and medially and occurs only in a sequence before stops.

4.1.1.2.3. The palatal nasal [ɲ]

The palatal nasal [ɲ] occurs in all three positions. In the intervocalic position it occurs after the two short open vowels [ɪ] and [ɛ] and the back close rounded vowel [u]. In the final position it is either preceded by the short vowels [e], [ɛ] or the vowels sequences [iie] and [ɔɔe]. In the final position it varies freely with [n] and [ŋ]. For example, *ńd̈ɛ̌ɲ* ~ *ńd̈ɛ̌ŋ* 'look here.'. However, it does not form a cluster with any consonant.

Although, its distribution intervocalically is restricted and it is a free variant finally, the palatal nasal [ɲ] has been attested in all positions. It contrasts with other consonants in identical environments at least word-initially. For instance, in the following minimal pairs: *ɲûl* 'clay' and *mûl* 'pottery bowel pl.'; *ɲúù* 'swell' and *kúù* 'light'; *kɛ̌ɲ* 'mouth' and *kɛ̌pˀ* 'wall'. Therefore, the palatal nasal /ɲ/ is considered as a phoneme.

4.1.1.2.4. The velar nasal [ŋ]

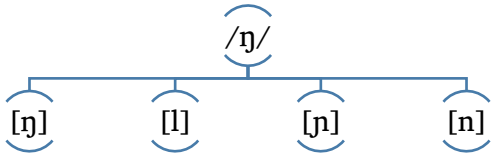
In Tegem, the velar nasal [ŋ] occurs in all three positions across different phonetic environments without restrictions, e.g., *ŋ̀ɣ̂m* 'salt', *t̪ìŋɛ̂* 'crying', and *ʈát̪émmìlìŋ* 'reply'.

In Tegem, [ŋ] occurs as part of the consonant(s) cluster [ŋg] in the intervocalic position (*see 4.1.1.3.15.* & *4.5.2.10.*). The consonant sequence of liquid and nasal [rŋ] has been attested in two words only in the collected data, i.e., *ŋérŋɛ̀él* 'tickle', and *ŋèrŋéɛ̀l* 'sharp'. Thus, the sequence [rŋ] is a marginal sequence and not considered further in this study because of its very low distribution frequency in the available data.

In Tegem, the velar nasal [ŋ] is in free variation with [n] ~ [ɲ] ~ [l] in the final position (in some words). The distribution patterns predict they are phonemically distinct phonemes in the other positions, initially and medially, but behave as allophones word-finally. The velar nasal [ŋ] distribution is more frequent and less resticted in comparison

to the other variants in all three word positions. Thus, /ŋ/ is the phoneme word-finally with four different realisations as shown below:

Rule (3)



/ŋ/ → [ŋ] ~ [n] ~ [ɲ] ~ [l] /___#

/ŋ/ → [ŋ] / ____V,C

The speakers associate the [n] and [ŋ] variants with the people of Tekeim village on Jebel Tekeim and consider it as an older variety of Tegem. The other variant [l] is described as the current common variant among the younger generation[15]. This variation was noticeable during the data elicitation sessions between the two main language consultants. Atroon (aged 55) who is originally from the Tekeim village, frequently, uses the [n] and [ŋ]. Omar (aged 47) who is originally from Al-Khor village on the plains next to Jebel Lafofa (*see 1.1.*) uses the other variant [l].

4.1.1.2.5. The alveolar central liquid [r]

The alveolar central liquid [r] has been attested in all three positions - word-initially, intervocalically and finally, e.g., *rìɟìŋ* 'say', *ńɟìrôm* 'sneeze', and *ɟǎŋér* 'buffalo'. There are no restrictions on the vowels before or after the alveolar [r].

The alveolar liquid [r] may be geminated in the medial position more frequently between the close front vowels. It also occurs in clusters with [w] in the initial and intervocalic positions: [rw],e.g. *rwʌ̂ŋ* 'word', *kórwʌ́ŋ* 'name'. This cluster is discussed further in *4.1.1.3.19.* and *4.5.2.11* below.

[15] This is possibly an arial feature developed over time as part of the Tegem people history of mobility and resettlement (cf. Manger 1994).

The alveolar central liquid /r/ is a phoneme due to its high distribution frequency and distinctive feature as it contrasts with other phones in identical environments. For example, *rîî* 'well pl.', *lîî* 'grass'; *rʌ̀h* 'tobacco', *ɬʌ́h* 'hair pl.'; *kílé* 'bring this!', *kìré* 'eating'; *ɟîr* 'queue', *ɟɨ̈l* 'deer pl.'; *kɔ́rɔ̂ŋ* 'young child', *kɔ́ɽɔ́ŋgèmɛ́ɛ̀* 'shef'.

4.1.1.2.6. The retroflex central liquid [ɽ]

The retroflex central liquid [ɽ] only occurs in word-medial position, e.g., *píɽíɛ̌* 'bird', *áɽɔ̌k*˥ 'frog pl.', and *ŋʌ́ɽìŋ* 'dirty'. It may not be preceded with the long front close vowel [ii] or the short open-mid vowel [ɛ], which is an exclusive environment for the geminated [ʈʈ].

The following sequences [mɽ] and [gɽ]: *kàmɽílʌ̀h* 'liver' and *mí-gɽìɛ́-góm mɛ́ɛ̀* 'hundred'. Each occured once only in the above words so they are not included in the consonant sequence analysis in *4.1.1.3.* below.

The retroflex central liquid [ɽ] occurs in a mutually exclusive distribution environment with the non-geminated retroflex stop [ʈ]. While the non-geminated stop [ʈ] occurs word-initially, the liquid [ɽ] occurs intervocalically. The intevocalic geminated [ʈʈ] occurs in a restricted intervocalic environment where it is preceded by [ɛ] or [ii], whereas the liquid [ɽ] occurs elsewhere intervocalically.

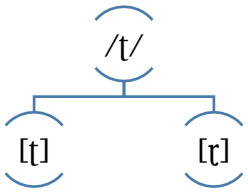
When a suffix is attached to a stem beginning with the retroflex stop [ʈ], its allophone [ɽ] occurs intervocalically, such as the suffix *ʈ̈ì* 'neutral 3sg. pronoun' before the retroflex stop [ʈ] in *ʈʌ̌ɽìéh* 'cutting' and the sequence of retroflex [ʈ] and glide [w] in *ʈwàáì* 'tying'.

ʈ̈ì-ɽʌ̌ɽìéh
'someone is cutting'.

ʈ̈ì-ɽwàáì
'someone is tying'.

Based on the previous examples, [ɽ] is in complementary distribution with [ʈ] and /ʈ/ is the phonemic realisation of two allophones:

Rule (4)



/ɽ/ → [ɾ] /[ii] or [ɛ] ___V,C

/ɽ/ → [ɽ] / ____V,C

Schadeberg (1981a:84) and Stevenson (1964:84) transcribes the retroflex stop [ɽ] as an alveolar stop [t] and they propose that the central linquid retroflex [ɾ] belongs to the phoneme /t/. Unlike the previous analysis, Schadeberg says that the alveolar /t/ varies with the retroflex flap [ɾ] in the intervocalic position and in certain consonant sequences, such as in *Kuku ḍɛ́ɛḍí ɲii ta ~ Kuku ḍɛ́ɛḍí ɲii ɾa*[16] 'Kuku took the water'.

4.1.1.2.7. The alveolar lateral liquid [l] and velarised [ɬ]

(a) The alveolar lateral liquid [l]

The alveolar lateral liquid [l] has been attested in all three word positions such as in *lùúɾíl* 'fog', *bíṭílìl* 'green', and *kíɲíl* 'fat person'. It occurs next to a wide range of vowels without certain restriction pattern.

The liquid [l] may only be geminated in intervocalic position and when preceded and followed by same front vowel. It may also occur as part of the consonants cluster [lw] in the initial and medial positions, e.g. *lwìì* 'okra' and *ɲìlwɛ́* 'life' (*see 4.1.1.3.21.* & *4.5.2.12.*).

[16] The example is from Schadeberg (1981a:84); where the low tone is not placed on the vowel.

(b) **The velarised alveolar lateral liquid [ɫ]**

The velarised alveolar lateral liquid [ɫ] has been found word-initially, intervocalically, and word-finally, e.g., *ɫʌ́d̪*˺ 'here', *jèɫǽ̂ɫ* 'big bed bug pl.', and *kʌ̀d̪ʌ̌ɫ* 'tortoise'. It has been attested either followed or preceded by a back vowel or the central vowels [æ] and [a].

It may only be geminated in intervocalic position and when preceded and followed by a back/central vowels such as in *lûɫɫód̪*˺ 'bow' and *t̪ʌ̂ɫɫɔ́* 'above'. It has not been attested as part of any other consonant sequence, except when it is followed by an extra-short vowel and the retroflex [ɽ] then it may functions as syllabic [l̩ɽ], e.g., *l̩ɽăbìnd̪ɔ́ŋ* 'freeze' (*cf. 4.3.5.*).

The velarized alveolar lateral liquid [ɫ] tends to be restricted to the environment where there is following and/or preceding back vowels [u], [ʊ], [o], [ɔ] and the central vowels [ʌ], [a], [æ]. Its counterpart, the alveolar lateral liquid [l], occurs with front, central and other back vowels. The velarized /ɫ/ and lateral liquid /l/ are distinct phonemes for their high distribution and occurrence and because they are in contrast with each other. For example, *t̪ŭɫ* 'Al-Haluug (fruit) ' and *t̪ŭl* 'cough'*; jèlààl* 'old woman pl.' *jélǽ̆ɫ* 'big fish pl.'; *ɫʊ́ʊ̀* 'buttock pl.' and *lʊ̀ʊ̀* 'flu'.

4.1.1.2.8. The bilabial glide [w]

The bilabial glide [w] occurs word-initially, e.g., *wăméh* 'if sg.', *wìl* 'ostrich', and *wíígɔ́* 'nostril pl.'. It has not been attested in an intervocalic or word final positions but it occurs after consonants word-medially and forms a consonant cluster, e.g., *kʊ́gwááró* 'armpit', *t̪ʊ̥̀rwʌ̌h* 'crow', and *kʊ́mwʌ̀t̪*˺ 'locust'.

It may be preceded by many consonants and consonant sequences including: stops [bw], [pw], [t̪w], [ʈw], [nɖw], [mbw], [ɟw], [cw], [kw], [gw], nasal [mw], and the liquids [rw] and [lw] (*see 4.1.1.3.* below).

The bilabial glide [w] as a single consonant occurs word-initially and in consonant clusters. This consonant forms clusters with 8 different phonemes in Tegem. It contrasts

with other cononants in the initial position such as in the following minimal pairs: *wìì* 'Jummaiz tree', *lìì* 'bunch', *mìì* 'sacral (body part)'; and *wìl* 'ostrich', *mìl* 'heart pl.'. Thus, as a distinct contrastive bilabial glide /w/ it is a phoneme in Tegem.

4.1.1.2.9. The palatal glide [j]

The palatal glide [j] occurs word-initially only and finally in few words, e.g., *júl* 'elephant', *jîr* 'queue', and *ṭ̈ìlílíàj* 'closing'. It does not form any sequence with other consonant. In the final position, it is also pronounced as the close front vowel [i] when added as a suffix to a word such as *ɲíí* 'water'; *ɲíí-ì ~ ɲíí-j* 'this water'.

It is predictable, according to the available data, that the palatal glide [j] is a plural gender class marker (*see 4.5.*), e.g. *ɟîr* 'queue sg.', *jîr* 'queue pl.', *cúr* 'feather sg.', *júr* 'feather pl.'. Therefore, /j/ is a contrastive phoneme with limited distribution to word-initial position only as the above minimal pairs indicate.

To conclude, the previous analysis include detailed description of each sound in the consonant inventory of Tegem (see *Table 3*). Hence, there are five obstruent phonemes (/B/, /ṭ/, /ʈ/, /C/, /k/), out of the ten phones and nine sonorant phonemes (/m/, /n/, /ɲ/, /ŋ/, /r/, /l/, /ɬ/, /w/, /j/) out of the ten phones (*see 5.1.* below). The previous part of the study also provides a brief account of the possible consonant combinations with the each one of the phones. The consonant sequences are complex structures in Tegem, so the following sections *4.1.1.3.* and *4.5.* give more detailed description and anlysis of those structures.

4.1.1.3. Consonant sequences

This section studies the consonant sequences, their distribution, environment conditions, and possible contrast with other sequences or single unit consonants. The consonant sequences can be categorized into three groups: (1) *Cw* (stop + *w*, liquid + *w*, nasal + *w*, and nasal + stop + *w*), (2) N stop, and (3) geminates. Those categories are discussed blow but not in the above order because of the interrelationship between all the different

groups. The *C* above refers to consonant, *w* the glide consonant [w], and *N* to a nasal consonant.

4.1.1.3.1. [bb] and [bbw]

The geminated sequences [bb] and [bbw] occurs only in word-medial position. The consonant [b] may only be lengthened when it is preceded by a sequence consisiting of one of [b] ~ [p] and the vowels [ɛ], [e], [iɛ], or [æ], e.g., *bĕbbwéŋ* 'small', *bèbbìí* 'calf (male 2-3 years old)', *pĕbbàlìl* 'wide', and *pèppwʌ̀* goat (male young). No minimal pairs have been attested for these clusters.

The clusters [bb] and [bbw] occur in a restricted environment - preceeded by a syllable with the bilabial [b] as its onset. Hence, this cluster is a geminated variant of the non-geminated [b] conditioned by that specified environment.

4.1.1.3.2. [mb] and [mbw]

The bilabial consonant combination with the voiced bilabial stop [m] is only possible in combination with the voiced bilabial stop [b] → [mb]. It has not been attested as a cluster with its allophonic voiceless variant [p]. It occurs in initial and intervocalic positions, e.g., *ḿbíŋ* 'hit' and *ǎmbúgbúkᵀ* 'lung pl.'. The sequence [mb] may be followed and/or preceded by different front, central and back vowels. The consonant sequence [mb] or one of its phones may contrast with other consonants in identical environments, for example *mbìì* 'pull!', *bìì* 'thirst', *mwìì óó* 'Kujur pl. (traditional healer)', and *míí* 'town'; *ḿbér* 'uproot' and *bwèr* 'jump'; *ímbíŋ* 'niece' and *ílíŋ* 'niece pl.'. Though, in imperative verbs, the bilabial [b] in the initial position may occur without the preceding [m] without any contrast, e.g., *ḿbûlí* ~ *bûlí* 'throw!'. Similarly, when a prefix is attached to an imperative verb beginning with [mb], the consonant [b] may occur without [m], e.g., *ḿbîŋ* 'hit!' *kɛ́-bíŋ* 'hit him!'. The bilabial sequence [mbw] is not common in Tegem and has been attested in the following two words only: *ḿbwʌ́kᵀ* 'suckle!' and *íbɛ̂mbwɛ́n* 'child'.

4.1.1.3.3. [bw] ~ [pw]

The consonant sequence [bw]~[pw] occurs in the initial and medial positions e.g., *bwʌ̌r* ~ *pwʌ̌r* 'fear' and *pɛ̀ppwʌ̀* ~ *bɛ̀bbwʌ̀* 'goat (young & male)'. The sequence can be preceded and followed by a wide range of front, central and back vowels. The sequence components may contrast with other phones, e.g., *pʌ́ʌ̀l* 'fish' and *pwʌ̂ʌ́l* 'wet'; *bʌ̀r* 'rabbit' and *bwʌ̌r* 'fear'; *wîîr̪ìt̪˺* 'elbow pl.' and *bwîîrɛ́r* 'learn'; *mwʌ̂* 'egg pl.' and *pwʌ̀* 'goat (old & male)'; in *kwɛ̀r* 'tear/cry' and *bwɛ̀r* 'jump'.

4.1.1.3.4. [t̪t̪] and [d̪d̪]

The geminated consonant sequence [t̪t̪] occurs in intervocalic position after front/central vowels only, e.g., *t̪wìt̪t̪ê* 'lying' and *lwǽt̪t̪ǽk˺* 'mat'. The sequence [d̪d̪] occurs in word medial position and similar to the allophone [d̪] it is attested after the back and central vowels in two words only, e.g., *t̪ʌ̀ʌ́d̪d̪úɬǒ* 'that', and *ɬɔ́d̪d̪ɔ́* 'underneath'. There are no minimal pairs attested for the sequence. The geminated consonant sequence [t̪t̪]/[d̪d̪] is considered an environment conditioned lengthened variant of the non-geminated consonant [t̪].

4.1.1.3.5. [t̪w]

The consonant sequence [t̪w] occurs in the initial position and less frequently intervocalically. It can only be followed by the front vowels or the central vowel [ʌ], e.g., *t̪wɛ̀ɛ́r̪ɛ̀* 'friend' and *t̪wʌ̀h* 'home backyard/ garden'. The cluster [t̪w] phones may be in contrast with other consonants/consonant clusters in identical environments: *t̪wàà* 'river', *t̪àà* 'firewood', and *kàà* 'body/border pl.'; *t̪ì* 'it (pronoun)' *t̪wɛ́r* 'throat' and *kwɛ́r* 'animal farm'; *t̪wʌ́m* 'bone splint' and *kwʌ́m* 'Basham (kind of tree)'.

4.1.1.3.6. [nd̪]

This consonant cluster has been attested in the initial and intervocalic positions. It is more common in the medial position and occurs usually after [i] and less frequently after [ɛ] and [u]. The sequence [nd̪] may occur in near minimal pairs such as in the following examples: *t̪ûnd̪ʌ̂* 'waste dump' and *t̪úúd̪ʌ̀* 'summer'; *ńd̪ɛ̌ɲ* 'look! (here)' and *kwɛ̂ɲ* 'scorpion'; *ɲìnd̪î* 'donkey' and *ɲííɟì* '1st person pronoun I'.

4.1.1.3.7. [ʈʈ] and [ʈʈw]

The geminated sequence [ʈʈ] and [ʈʈw] occur in intervocalic position and come after the front vowels [ii], [ee], and [ɛ] only, e.g., *ʈííʈʈwàh* 'many', and *ʈìɽéèʈʈì* 'one', and *ʈèʈʈêm* 'very short'. The cluster [ʈʈw] has been attested once only in the previous example, hence it is a marginal sequence. The consonant [ʈʈ] is a lengthened (geminated) variant of the non-geminated phoneme [ʈ] conditioned by the environment.

4.1.1.3.8. [ʈw]

The sequence [ʈw] occurs in the initial position only. It can be followed by both front, central and back vowels. When compared to other similar sequences, the consonant [ʈ] in the sequence may contrast with other unit: [c] in [cw], and [b] in [bw], e.g., *ʈwêl* 'arm accessories' and *cwêl* 'waist accessories'; *ʈwìr* 'Kujur (traditional healer)' and *bwìr* 'fear'.

4.1.1.3.9. [nɖ] and [nɖw]

The consonant cluster [nɖ] occurs in the initial and medial positions. It can be preceded by front vowels word-medially. In the initial position [ɖ] varies freely with the voiceless retroflex [ʈ] in some verbs and [n] occurrence is optional, e.g., *nɖâbín* ~ *ʈâbín* 'cut! (with an axe)'. The sequence [nɖw] has been attested in the initial position in two words only. Similar to [nɖ], [ɖ] varies freely with [ʈ], and [n] occurrence is optional, i.e., the sequence become [ʈw], e.g., *nɖwáàí* ~ *ʈwáàí* 'tie!', and *nɖwèʈ̈ʈ̈ʌ̀ŋ* ~ *ʈwèʈ̈ʈ̈ʌ̀ŋ* 'cut! (with knife)'. No minimal pairs has been attested for these sequences.

4.1.1.3.10. [cc] and [ɟɟ]

The geminated sequences [cc] and [ɟɟ] are in free variation and they occur in the medial position. The consonant [c] occurs in the medial position as geminated [cc] only. When they are geminated, [c] and [ɟ] are preceded and followed by any two indentical short vowels. These clusters have been attested in three words only and they can be pronounced non-geminated as well, e.g., *cɔ̀ccɔ̆ŋ* ~ *cɔ̀cɔ̆ŋ* 'story', *cɔ́ccɔ́ŋ* ~ *cɔ́cɔ́ŋ* 'sixth

child', and *ʈ̪ɛ̀ɟɟɛ̀h* ~ *ʈ̪ɛ̀ɟɛ̀h* 'this'. Thus, [cc] and [ɟɟ] are lengthened free variants of the phoneme /C/ in the medial position.

4.1.1.3.11. [ɟw] and [cw]

The sequence [ɟw] has been attested in word medial position and [cw] in the initial position. Their occurrence frequency is very low in the available data. They may occur in minimal pairs with other consonants (or sequences), i.e., *cwéè* 'drink!', *ʈéè* 'warts', and *kwéé* 'kind of tree'; *cwêl* 'waist accessories', *ʈwêl* 'arm accessories', *kwěl* 'moon'.

4.1.1.3.12. [nɟ] and [nɟw]

The sequence [nɟ] occurs in the initial and middile positions, e.g., *nɟúúɖ́ŋ* 'dig!' and *bŏnɟéŋ* 'give me!'. The sequence [nɟw] has been found word-medially once only in *ʈ̪ínɟwì* 'find'. There are no minimal pairs attested for these consonants as sequences.

4.1.1.3.13. [kk]

The geminated consonant [kk] occurs in intervocalic position. It is always preceded by a syllable with a velar stop [k] as its onset and a short vowel as a peak, e.g., *kêkkáh* 'four', *kíkkáh* 'with', *kúkkíl* 'far', and *kèkkwèè* 'Raika cover (basket)'. The velar stop [k] (including the cluster [kw] → [kkw]) can occur in word medial position only as geminated consonant. Thus, it is a lengthened variant of [k] restricted to the environment mentioned above.

4.1.1.3.14. [kw] and [gw]

The consonant sequence [kw] occurs in the initial position whereas, [gw] ocurs in the intervocalic position. The consonant [k] as part of [kw] sequence may contrast with other consonants, e.g., *kwěl* 'moon', *ʈwěl* 'calabash bowel', and *ʈ̪wěl* 'dance'; *kwɛ̀ɛ̀* 'sword' and *ʈɛ̀ɛ̀* 'water tough'; *kwɑ́m* 'Basham (kind of tree)' and *ʈ̪wɑ́m* 'bone splint'. The velar stops [g] and [k] are in complementary distribution and they act similarly when they are

part of the sequences [kw] and [gw], i.e., [k] occurs word initially followed by the glide [w] and medially [g] occurs before [w].

4.1.1.3.15. [ng] and [ŋg]

The sequence [ng] occurs in the initial position. The cluster [ŋg] have been attested in intervocalic positions and in few words in the initial position (verbs and nouns) where [n] and [ŋ] varies freely before their cluster component [g], e.g., *ŋgáh* ~ *ngáh* 'hitting' and *ngênɲìk*˥ ~ *ŋgênɲìk*˥ 'spit'.

4.1.1.3.16. [mm]

The geminated consonant sequence [mm] occurs in an intervocalic position only. It is preceded by either on of the short vowels: [ɛ], [ɔ], or [u] and followed by various vowels, e.g., *múŋúmɛ̀mmôl* 'cooking pot pl.', *mó̜rɔ́mmʌ̂* 'head/skull pl.', . There are no minimal pairs attested for [mm] in that position. This consonant sequence is a lengthened variant of the single unit [m] and conditioned by the environment.

4.1.1.3.17. [mw]

The consonant sequence [mw] occurs word-initially but it has also been attested in intervocalic position in one word- *kúmwʌ̀t̪*˥ 'locust'. It may be followed by front, central and back vowels. The consonant [m] in this cluster may contrast with other consonants such as *mwíí* 'forehead pl.', *kwíí* 'hoof', and *míí* 'town'; *mwʌ̀m* 'bone pl.' and *kwʌ̂m* 'mountain'.

4.1.1.3.18. [rr]

The alveolar consonant [r] may be geminated in the medial position. The geminated [rr] occurs after the short front vowels [i] and [ɪ]. However, this gemination is phonetic and not contrastive phoneme, i.e., there is not contrast between the geminated [rr] in intervocalic position and the non-geminated variant [r]. Such as in the following examples *t̪ìrríṯíél* ~ *t̪ìríṯíél* 'divide' and *t̪írrɛ́h* ~ *t̪írɛ́h* 'burn'.

4.1.1.3.19. [rw]

The consonant sequence [rw] occurs in the initial and medial positions of the word. It is found always before the central vowel [ʌ] and in the intervocalic position. There are no minimal pairs for this cluster but there are near minimal pairs where the contrast may correspond to the consonant [r] only, for example *rwâŋ* 'word/speech' and *bʌ́ŋ* 'sheep'; *ʈ̪ṵ̀rwʌ̌h* 'crow' and *ʈ̪ûʈ̪ʌ̌h* 'baldness/grey hair'; *rwâ̰rʌ̀ʌ́* 'wet wood pl.', *rŏ̰rʌ̀ʌ́* 'dry wood pl.'.

4.1.1.3.20. [ll] and [ɬɬ]

The geminated consonants [ll] and [ɬɬ] occur in the word medial position only. The consonant [ll] has been attested preceded by one syllable consisting of an onset consonant and/or one of the following vowels as a peak: [i], [ɪ], [e], [ɛ] or [u] such as in *ʈ̪ìllí* 'go', *ímíllí* 'possessive pronoun (mine)', *éllàh* 'back', *jéllál* 'big pl.', and *rûllʌ̂* 'waste dump pl.'. The environment where it occurs is restricted but its non-geminated variant [l] may be preceded by more than one syllable and may occur in an environment similar to [ll] as well, e.g., *ʈ̪ìlɨ̀ì* 'hitting a drum', *ílíŋ* 'niece pl.', *jélɛ̌ɬ* 'big fish pl.', *bɛ́líl* 'new', and *ʈ̪úlúh* 'smelling'. The sequence [ɬɬ] also has been attested in a restricted environment and may occur in an environment similar to the non-geminated [ɬ]. The cluster [ɬɬ] always occurs preceded by an onset consonant and one of three vowels [i], [u], or [ʌ] as a peak and followed by a back/central vowel, e.g., *kʌ̌ɬɔ́* 'in the caves', *ńd̪íɬɬɔ̌* 'found', and *lûɬɬód̪˥* 'bowing'. Whereas, the non-geminated [ɬ] occurs in those environments and elsewhere, e.g., *kʌ̂ɬɔ́*[17] 'rain', *pàɬáá* 'hunger', *ɬùùɬṵ̂k˥* 'cotton', and *jèɬáá* 'old man pl.'. The geminated consonants [ll] and [ɬɬ] are geminated due to phonetic environment conditions.

4.1.1.3.21. [lw]

The consonant sequence [lw] occurs in word initial and medial positions. It may be preceded by one of the following vowels [i], [e], [æ], [a], or [u]. The consonant [l] is in contrast with other consonants in identical environments for instance, *lwʌ̀* 'net', *pwʌ̀* 'male goat', and *kwʌ̀* 'tree'; *lwèè* 'hole', and *kwèè* 'Raika (kind of basket)'.

[17] There is tonal contrast with the previous *kʌ̌ɬɔ́* 'in the caves'.

The previous paragraphs shows that consonant sequences are very frequent in Tegem but they are restricted to specific consonant combinations. Stevenson (1964:84) states that nasal consonants can be combined with the glide *-w* but not the liquids. However, Table 4 below includes 8 different stops, nasals and liquids before the glide *-w*, 3 different phonemes consonants after the nasals *n-* and *m-*, and 5 geminates.

Table 4. *Consonant sequences in Tegem*

Description	**Consonant seq.**	**Example**	**Gloss**
stops + glide	bw	*ébwɛ̌ndɛ́h*	who
	t̪w	*bǐɛ̌t̪wǐl*	ant
	ʈw	*ʈwʌ̀m*	bone
	ɟw	*ngúɟwɨ̌t̪ʌ̀ŋ*	wash!
	cw	*cwĩ̃i*	rubbish
	kw	*kwʌ́l*	air
	gw	*mùgwɛ̀r*	beer
liquids + glide	lw	*lwʌ́r*	kidney
	rw	*kúrwʌ́ŋ*	name
nasal + glide	mw	*mwɑ̂*	egg pl.
nasals + stop	mb	*bîámbʌ́ŋ*	big sheep
	nd̪	*ŋɲɛ́nd̪ɑ̂ŋ*	sing!
	nɖ	*ʈínɖɛ̀è*	full
	nɟ	*nɟɔ́ɔ́*	stab!
	ng	*ngúɟwɨ̌t̪ʌ̀ŋ*	wash!
	ŋg	*cɛ̀ŋgɪ̀mɛ̀è*	important person
nasals + stops + glide	mbw	*mbwʌ́k˥*	suckle!
	nɖw	*nɖwáàî*	tie!
geminates	bb	*bébbú*	dog
	mm	*mo̰rɔ́mmɑ̂*	head pl.
	t̪t̪	*ʈɪɽéɛ̀t̪t̪í*	one
	ʈʈ	*ʈɛ̌ʈʈɛ̀é*	calabash
	kk	*kíkkɑ̂*	correct

The distribution of the above consonant sequences is restricted as shown in the following table 5. It shows that no consonant sequence is allowed word-finally in Tegem. The majority occurs in the initial and medial positions.

Table 5. *Distribution of consonant sequences*

	consonant	**word-initial**	**intervocalic**	**word-final**
stops + glide	bw	+	+	-
	t̪w	+	+	-
	ʈw	+	+	-
	ɟw	+	+	-
	cw	+	+	-
	kw	+	-	-
	gw	-	+	-
liquids + glide	lw	+	+	-
	rw	+	+	-
nasal + glide	mw	+	-	-
nasal + stop	mb	+	+	-
	nd̪	+	+	-
	nɖ	+	+	-
	nɟ	+	+	-
	ng	+	-	-
	ŋg	+	+	-
nasals + stops+ glide	mbw	+	+	-
	nɖw	+	-	-
geminates	bb, mm, t̪t̪, ʈʈ, kk	-	+	-

In Table 5, the sequence *kw-* occurs word-initially and *–gw-* occurs word-medially which are exclusive environments because they are in complimentary distribution. There are other sequence in the table which have not been attested word-medially, including *mw-*

, *ng-*, and *nḍw-*; they have low occurrence frequency in the available data. The geminates position is always conditioned to the intervocalic (medial) position only.

The consonant sequences in Tegem are investigated via a bisegmental analysis. In this type of analysis, the consonant sequence is considered as a cluster of two distinct phonemes (Clements 2000:145). The study follows this analysis because Tegem has diffierent affixes attached to the stem/root, there are syllabic nasals, and cross boundary consonants combinations. The further analysis below (*4.3.4., 4.3.5.* & *4.5.*) also proves that each consonant in a sequence behaves differently in various morphological environments. According to this analysis, they are not prenasalised or labialised consonants.

4.2. Tegem vowels

4.2.1. The vowel inventory

The phonetic inventory of Tegem vowels consists of 12 vowels, as shown in table 6. Schadeberg's (1981a) transcription distinguishes 7 vowels and the added vowels in this study are [æ], [e], [ʌ], [o], and [ɑ]. All the short vowels in Tegem can occur as long vowels and some of them may even occur as what Schadeberg called triple-long vowels (extreme vowel length). The following sections include details of the vowel system in Tegem.

Table 6. *Vowel inventory of Tegem*

	Front	**Central**				**back**
close	i					u
	ɪ					ʊ
Close-mid	e					o
Open-mid	ɛ					ɔ
Open		æ	a	ʌ	ɑ	

4.2.1.1. The front vowels [i], [ɪ], [ɛ], and [e]

(a) The close front vowels [i] and [ɪ]

The two close front vowels [i] and [ɪ] are highly frequent but the lax vowel [i] is more frequent than the tense [ɪ] and all the rest of the vowels. They may occur in all three word positions and next to both obstruent and sonorant consonants: *ííɲì* 'I', *bíṭ̩ílìl* 'green', *bwɛ̀lì* 'send'; *ímbŭkᵀ* 'wife', *bìr* 'rabbit', *kwì* 'breath'. They also contrast among themselves and with the other vowels as follows:

Example (1)

[i] – [ɪ]		[i] – [ɛ] – [æ] – [ʌ]		[ɪ] – [a] – [u] – [ʊ]	
ṯí	'grinding stone'	*rìh*	'secondary root pl.'	*íh*	'ox pl.'
ṯì	'it'	*rɛ̀h*	'skin'	*áh*	'finger pl.'
		rɑ̀h	'tobacco'	*úh*	'blow (with mouth)'
		rʌ̀h	'backyard'	*óh*	'horse pl.'

According to their frequent unrestricted distribution and contrast with other vowels, the close front vowels /i/ and /ɪ/ are distinct phonemes in Tegem.

(b) The mid front vowels [e] and [ɛ]

The mid front vowels [e] and [ɛ] occur in initial, medial , and final word positions, e.g., *ép*˥ 'earth worm pl.', *kílék*˥ 'move', *ímbáábé* 'paternal uncle'; *ɛ́ɲíṯ*˥ 'body front pl.', *íbɛ́ɛ́lɛ́ɲ* 'child pl.', *ṯímălwɛ́ɲɟɛ́* 'gather'. They may be preceded and followed by both obstruent and sonorant consonants. They also contrast with each other and the other vowels, as in the following examples:

Example (2)

[e] – [ɛ] – [u]		[ɛ] – [ɑ]		[ɛ] – [ɔ]	
mèr	'aromatic grass pl.'	*kwɛ̀r*	'tear/cry'	*bɛ̌m*	'Kujur pl.'
mɛ́r	'toe pl.'	*kwàr*	'leopard'	*bǒm*	'gay'
mûr	'belly pl.'				

Thus, the minimal contrast and distribution frequency of the mid front vowels /e/ and /ɛ/ shows that they are distinct phonemes in Tegem.

4.2.1.2. The back vowels [u], [ʊ], [o], and [ɔ]

(a) The close back vowels [u] and [ʊ]

The close back vowels [u] and [ʊ] have been attested in all three word positions next to various obstruent and sonorant consonants, e.g., *úḍì* 'they (masculine)', *ṯúlúh* 'smelling'

éérú ‘dog pl.’; *úh* ‘horse pl.’, *mûl* ‘fruit pl.’, *pùmʌ̀bù* ‘woman’. These two vowels are contrastive such as:

Example (3)

[u] – [ʊ] – [ɑ]	[ʊ] – [ʌ]	[u] – [ɔ]
úh ‘blow (with mouth)’	*ʈù* ‘swimming’	*ɬú* ‘kind of stick’
úh ‘horse pl.’	*ʈʌ̂* ‘head’	*ɬʊ̌* ‘animal dung’
àh ‘fly pl.’		

The above distribution pattern and contrastive nature of the close back vowels /u/ and /ʊ/ show that they are two distinct phoneme.

(b) The mid back vowels [o] and [ɔ]

The close-mid and open-mid back vowels [o] and [ɔ] are similar to the previous vowels as they occur in all three word positions next to obstruents and sonorants: *óm* ‘hand pl.’, *bór* ‘snail’, *kúgwáárʊ́* ‘armpit’; *ór* ‘fox pl.’, *bógʊ̌r* ‘hen’, *ʊ̌rʊ́* ‘country pl.’. They contrast with other vowels as follows:

Example (4)

[o] – [ɔ]	[ɔ] – [ʊ]	[o] – [a]
ʈóm ‘small mountain’	*rô* ‘horn pl.’	*ʈ̈ɪ̀ʈ̈ʌ̆bʌ̆h* ‘breaking’
ʈɔ́m ‘rope’	*rʊ̌* ‘primary root pl.’	*ʈ̈ɪ̀ʈ̈ʊ̌bʌ̆h* ‘splitting’

The back vowels /o/ and /ɔ/ have no distribution restrictions and they are in contrast with the rest of the vowels. Hence, they are two separate phonemes in Tegem.

4.2.1.3. The central vowels [æ], [a][18], [ʌ], and [ɑ]

In this study all the vowels [æ], [ʌ], and [ɑ] corresponds to [a] in Schadeberg’s (1981a) transcription. The complexity of these vowels in Tegem come from the fact that they are

[18] This symbol may display as [ɑ] in some words because of the font (Charis SIL) but that does not affect the meaning because [ɑ] and [a] are not contrastive in Tegem.

not conditioned by environment, there is no contrast between some of them, and there is no specific pattern when they are not in contrast. Therefore, to add additional evidence to support the final decision on their phonemic status, the morphophonemic behaviour of the vowel in two-gender classes (*see 4.5.* below) will be considered.

(a) The central vowels [æ] and [ʌ]

Tegem open-mid central vowels [æ] and [ʌ] may occur word-initially, medially and finally. The vowel [æ] has been attested in various environments but more frequent after consonant sequences and the sonorants [r] and [l] and it is in free variation with [ʌ] in these positions, e.g., *ʈwœ́bˀ* ~ *ʈwʌ́bˀ* 'road/path', *rœ́h* ~ *rʌ́h* 'thorn pl.', *lɛ̀r* ~ *lʌ̀r* 'cloth pl.'. The vowel [ʌ] is not restricted to specific following or preceding consonant(s) e.g., *lʌ́r* 'cloth pl.', *ʈʌ̀ŋʌ̌r* 'thigh', *kúmwʌ̀ʈˀ* 'locust'. Both [æ] and [ʌ] occur word-initially as a result of a deletion and/or alternation of the initial phones in the plural gender, e.g., *kílbúkˀ* (sg.) and *œ́lbúkˀ* (pl.) 'robe/gown'; *bɔ́gɔ̌r* (sg.) and *ʌ́gɔ̌r* (pl.) 'hen'.

The vowels [æ] and [ʌ] do not contrast with each other but they contrast with other vowels in Tegem, e.g.,

Example (5)

[æ] - [a]	[ʌ] – [u]	[æ] – [o]	[ʌ] – [a]
bíámbœ̂l 'big bed bug'	*lʌ́r* 'kidney pl.'	*œ́l* 'bed bug pl.'	*pwʌ̀* 'goat (male old)'
bíámbʌ́l 'big	*lûr* 'cloud'	*ɔ̌l* 'machete pl.'	*pwà* 'sugar cane'

The two vowels [æ] and [ʌ] may occur in both plural and singular gender classes in Tegem. The vowel [æ] occurs in plural gender more than the singular and it may occur in one gender words (verbs, uncountable nouns, etc.) as well. Though, [ʌ] frequently occurs in different classes and genders. When they occur in two-gender pairs they are usually retained in the plural gender[19], e.g., *ʈwʌ̀kˀ* (sg.) and *mwʌ̀kˀ* (pl.) 'ankle'; *pʌ́l* (sg.) and *œ́l* (pl.) 'bed bug'. When they undergo an alternation process, they are changed from [æ] or [ʌ] in singular to [ɔ] in plural or from [ʊ] in singular to [æ] or [ʌ] in the plural gender as in the following examples:

[19] That is if we assume the plural is derived from the singular.

Example (6)

[æ] & [ʌ] → [ɔ]			**[æ] & [ʌ] → [ʊ]**		
singular	**plural**	**gloss**	**plural**	**singular**	**gloss**
pɛ̀ppwʌ̀	*ɛ̀ɛ̀rɔ̌*	'goat (male young)'	*ɬʌ́ʈʌ̂r*	*lʊ́ʈwʌ̂r*	'thread'
kwʌ̀	*bɔ̀ɔ̀*	'tree'	*ágwɛ́h*	*kúgwɛ́h*	'bark of tree'
ʈwɛ̀ɛ̀	*rɔ̀ɔ̀*	'arm'	*ɛ̀gwɛ̀h*	*kúgwɛ̀h*	'tree monkey'

The previous paragraphs show that [æ] is in free variation with [ʌ], they contrast with other vowels, they act similarly in two-class gender, and [ʌ] occurrence is more frequent and less conditined than [æ]. Accordingly, the open-mid central vowel /ʌ/ is the phoneme and [æ] is one of it is realisations in Tegem.

(b) The central vowels [a] and [ɑ]

The open central vowels [a] and [ɑ] have been attested in all three word positions. The vowel [a] occurs very frequently and it is not conditioned to specific environment, e.g., *ándì* 'why', *lágáàbík*ˀ 'rub the body', *ʄíɲá* 'people'[20]. Whereas, [ɑ] found in few words in Tegem and it only occurs word-initially or after the bilabial stop [b], e.g., *àɖúɽwìm* 'mosquito pl.' and *ʈìɽăbăh* 'breaking'; the vowel [a] occurred preceded by [b] in one word only: *bíámbál* 'big'. These two vowels do not contrast with each other but they have minimal pairs with other vowels such as:

Example (7)

[a] – [ɔ]		[a] – [æ]		[ɑ] – [ɪ]	
káh	'finger'	*ʈwáá*	'Aradeeb (Tamarind) sg.'	*báh*	'uproot!'
kɔ́h	'tree cavity'	*ʈwɛ̀ɛ̀*	'arm'	*bíh*	'ox'

When they occur in two-gender classes [a] and [ɑ] usually retained in the plural if they were there in the singular (see the first column of *Example 8* below). Though, in some singular-gender classes the vowels: [i], [u], [ʊ], and [ɔ] are changed to [a] in the plural, e.g.,

[20] The used font (Charis SIL) may diplay both [a] and [ɑ] as [ɑ] when italicised but they are different in the phonetic transcription of the examples provided above.

Example (8)

[ɑ] – [a]			**[i], [u], [ʊ], & [ɔ] → [a]**		
singular	**plural**	**gloss**	**singular**	**plural**	**gloss**
kàrǎŋ	*àrǎŋ*	'nose'	*kíʈíàà*	*áʈíàà*	'hen (wild)'
ʈíàɖ̈	*míàɖ̈*	'road/path'	*kúgwʌ̀*	*ágwʌ̀*	'cock'
bàɖ̈úʈwìm	*àɖ̈úʈwìm*	'mosquito'	*kúgwááró*	*ágwááró*	'armpit'
bàh	*àh*	'house fly'	*kɔ́ɖ̈ɔ́*	*áɖ̈ɔ́*	'son/daughter'

The vowels [a] and [ɑ] occur in almost mutually exclusive environments (except the example mentioned above), [a] is more frequent and not conditioned by specific environment, they contrast with other vowels, and both [a] and [ɑ] may be retained in plural gender. Therefore, the central vowels [a] and [ɑ] in Tegem are in complenetary distribution after the bilabial stop [b] and /a/ is the phoneme.

4.2.2. Vowel length

All the vowels occur lengthened word-initially, medially, and finally except the variants [ææ] and [ɑɑ] which have not been attested lengthened in the initial position. There are no restrictions on the preceding and/or following consonants occurring adjacent to the long vowels. Long vowels may occur next to the obstruents and the sonorants. Some long vowels such as *aa, ʊʊ, oo* occur as a lexeme without any consonant but they may optionally be preceded by a glottal stop [ʔ], e.g., *áà ~ ʔáà* 'cow pl.', *ùú ~ ʔùú* 'mouse pl.', *mwìì óó ~ mwìì ʔóó* 'Kujur pl.'. The long vowels may contrast with both short and other long vowels in identical environments such as in the following table 7:

Table 7. *Long vowel minimal pairs*

(near) minimal contrast	**example**	**gloss**	**example**	**gloss**
/ii/ : /ɪɪ/	*bìì*	thirst	*bīī*	boa
/ɪɪ/ : /ii/	*kwìì*	nerve	*kwīì*	Jummaiz tree
/ee/ : /ɛɛ/	*méé*	heel pl.	*mɛ́ɛ́*	sesame pl.
/ɛɛ/ : /ʌʌ/	*kùgwɛ́ɛ̀*	wing	*kúgwʌ́ʌ́*	farming tool
/aa/ : /ee/	*kwáá*	sand/soil	*kwéé*	Habeel tree

/ʌʌ/ : /ʊʊ/	*ʌ̀ʌ́*	monkey	*ʊ̀ʊ́*	mouse pl.
/ɯɯ/ : /oo/	*mûú*	ant heap pl.	*môó*	molar tooth pl.
/ɯɯ/ : /ɔɔ/	*úúɲ̀ì*	they (masculine)	*ɔ́ɔ́ɲ̀ì*	you pl.
/ʊʊ/ : /ɯɯ/	*rʊ̀ʊ́*	navel pl.	*rúú*	leg pl.
/oo/ : /aa/	*ʈì̀ɲóò*	baobab fruit shell	*ʈíɲáà*	straight
/ɔɔ/ : /ʊʊ/	*ɔ̀ɔ̀*	shoe pl.	*ʊ̀ʊ́*	mouse pl.
/ii/ : /i/	*rî*	well pl.	*rí*	branch pl.
/aa/ : /a/	*káá*	kind of vegetable	*ká*	finger
/ɯɯ/ : /u/	*múú*	charcoal pl.	*mú*	bowl pl.
/ʊʊ/ : /ʊ/	*ʊ̀ʊ́*	mouse pl.	*ʊ́*	horse pl.
/ʌʌ/ : /ʌ/	*mʌ́ʌ̀*	bean pl.	*mʌ̂*	head pl.
/ɔɔ/ : /ɔ/	*rɔ́ɔ́*	river/sea pl.	*rɔ̌*	primary root pl.

The long vowels distribution is unconditioned across the three word positions and they are in contrastive opposition with other vowels as seen in the above (near-) minimal pairs. Thus, the vowel length is a contrastive feature in Tegem.

4.2.3. Extra-long/short vowels

In Schadeberg's (1981a:79) phonological notes, the extreme vowel length condition has been explained as a result of adding the suffix *–i* which sometimes occurs as the glide *–j* (used as a demonstrative marker) to nouns ending in long vowels, e.g., *ʈ̈íbìí* → *ʈ̈íbìí-ì* 'spear'; *ɲìh* → *ɲ-îî-ì* 'blood', *ŋéé* → *ŋéée̋j* 'fat'. However, such a triple-long vowel has been attested in one word only referring to 'dark' and the variants derived from it as follows: *kîîí* 'dark', *kîîíʈ̈* 'night', and *kîîíʈʈɔ́* 'morning'. This morphologically complex unit cannot be analyzed at this stage of the research and it may include other lexemes ending in long vowels with falling rising tones such as, *ʈîî* 'pile of soil', *cîî* 'wire'. In the latter two examples it is not clear if [i] is a long vowel (as transcribed) or extra-long vowel (*ʈîîí* and *cîîî*). The available examples also suggests that this phenomenon may be restricted to the close front vowels only.

Some vowels in Tegem can be extra short as well. This non-contrastive feature occurs on the central and back vowels [ɐ̯], [æ̯], [ɑ̯], and [ɔ̯] when followed and/or preceded by the

bilabials [w], [m], [b] and the liquid [r], e.g., *mɔ̥́rɔ́mmʌ̂* ‘head pl.’, *ḿbwʌ̥́k˺* ‘suckle!’, *ʈwʌ̥̀k˺* ‘ankle’, *kʌ̥̀mɽíʌ̀h* ‘liver’, *kə̥̀rʌ̆ŋ* ‘nose pl.’, *rʌ̆mʌ̥̀lʌ́ŋ* ‘knife pl.’, *t̪wæ̥̀b˺* ‘farming tool’, *ʈṳ̀rwʌ̆h* ‘crow’. These short vowels may correspond to a very short schwa (cf. Schadeberg 1981a:78), such as in *bərʌ̂líl* ~ *bʌ̥́rʌ̂líl* ‘thin’ and *ləmááɽì* ~ *lə̥̀mááɽì* ‘walking’. These extra short vowels are not distinct phonemes in Tegem.

4.2.4. Vowel sequences

Vowel sequences in Tegem are most frequent in the medial position, less frequent in the final position, and not allowed in the initial position. There are 22 possible vowel sequences, 16 of these sequences are combinations of the long and short close front vowel [i]/[ii] and another vowel as follows:

[i] [ii]	Medial	**ie**	**iɛ**	ia	iæ	iʌ	**iie**	**iaa**	**aai**	**ɔɔi**	iiɛ	iɛɛ	iɔɔ	ooi
	Final	**ie**	**iɛ**	iee	iiɔ	iiaa[21]		**iie**	**iaa**	**aai**	**ɔɔi**[22]			

Medial	[u]	ue	uɛ	ua	aau
	[ɔɔ]	ɔɔɛ			
	[æ]	æee			

The most common possible clusters are between the front vowels and front and central vowels. The rest of the sequences has not been attested more than 1-5 times in the available data and some of them occurs due to combination of two lexemes and in phrases, e.g.,

Example (9)

[iiɛ] in *ɟìì-ɛ́ɲ*
eye pl.-child
‘eye pupil’

[æee] in *pú-ræ̀-èélì*
you-future marker-do[23]
‘What do you want to do?’

[21] This sequence of two long vowels sequences has been attested in one word only in the available data.
[22] In the sequence [ɔɔi], the vowel [i] varies freely with the glide [j] word-finally, e.g., *bὲbbɔ́ɔ̀-í* ~ *bὲbbɔ́ɔ̀-j* ‘heifer’.
[23] This glossing is approximate.

[ɔɔɛ] in *bɔ̀ɔ̀-ɛ́ɲ*
tree pl.-child
'fruit'

[ɔɔi] in *ɛ́ɛ́rɔ́ɔ̀-í*
heifer pl. (2-3 years old)-this
'These are heifers.'

If we exclude those marginal sequences (occurred less than four times) the number could be only five frequent vowels clusters: [iɛ], [ie], [ia], and [iaa]. The first three occur in both word medial and final positions, but [ia] occurs in the medial position only, e.g., *ɟíér* 'hunt', *ɟìɛ̀* 'razor'; *líéh* 'foot', *ʈìèé* 'chest'; *bíámbál* 'big'; *bìáábúú* 'noon', *líllíàà* 'open'.

The vowel sequence consisting of a short and a long vowel is possible as shown above. There is only one vowel cluster with two long vowels [iiaa], which has been attested in one word only- *ʈ̈ìiáá* 'fry'.

To conclude, vowel sequence is one of the vowel features in Tegem including some less frequently occurring sequences which are part of the lexeme root, such as the following clusters which occurred only once: [iʌ] *kʌ̀mʈ́ʌ̀h* 'liver', [iæ] *líǽŋ* 'tongue', [iiɔ] *ɲîîɔ̌* 'well-adjusted', [iɔɔ] *píɔ́ɔ́l* 'cold', [uɛ] *búɛ̀ɲ* 'rat', and [ue] *búèɲ* 'monitor lizard'. Therefore, as Schadeberg (1981a:79) mentioned vowel sequences are frequent in Tegem with any vowel excluding the vowels [ɪ], [ʊ], or [ɑ].

4.3. Tegem syllable structure and surface tone

4.3.1. Syllable structure

Tegem has both open and closed syllables, which may have a heavy or light mora. The syllable consists of a consonant or consonant sequence margins (onset/coda) referred to as C or CC (where the coda CC can be a geminated consonant) and nucleus vowel, vowel sequence, or syllabic consonant (peak) referred to as V, VV, VVV, or C̩ (where VV(V) can be a long vowel).

4.3.2. Syllable types

The following table 8 shows the syllable inventory of Tegem, i.e. 13 syllable types divided into 7 open and 6 closed syllables. The syllable type CVVVC has been attested only once in a monosyllabic word i.e. *kîíṯ*˥ 'night' and in multisyllabic word. The vowels sequence VVV belongs to more than one syllable (*see 3.1.3. below*), so it will not be considered as one of Tegem syllables.

Table 8. *Tegem syllable types*

Syllable type	example	gloss
V	*ù*	strength
VV	*áà*	cow pl.
CV	*mò*	Al-hemaidh fruit
CVV	*ɲíí*	water
CVVV	*ɲíìé*	dough
CCVV	*mwíí*	forehead pl.
CCV	*kwì*	breath
C̩	*ń̩.ɖɛ́h*	what
VC	*ǒl*	machete pl.
VVC	*ʌ́ʌ̀l*	fish pl.
CVC	*ɲɔ́ŋ*	death
CVVC	*líǽŋ*	tongue
CCVC	*ṯwʌ́m*	bone splint

Table 8 also sets the possible segments combinations in the syllable onsets, coda(s), and peak according to the language phonotactic. In the syllable onset, not more than two consonant sequences are allowed in Tegem, e.g., CVVC in *pìɛ̂l* 'red', CCVV.CV in *ʈwɛ̀ɛ́.ɽɛ̀* 'friend', and CVC.CCV in *ʈín.ɟwì* 'find', C.CCVC.CVC in *ɳ̣.ɖwɛ̀ʈ.ʈʌ̀ŋ* 'cut!'. In the syllable coda, only single coda consonant is allowed in all the 6 closed syllables, e.g., VC in *ép*˥ 'earth worm pl.', VVC in *ɔ́ɔ́l* 'baboon pl.', and CVVC in *búɛ̀n* 'rat'. In the nucleus of the syllable, both long and short vowels are allowed as well as vowel sequences, such as V.CVC *í.bɛ́l* 'thing', CVV.CVVC *láá.ɖíɛ́l* 'come from far pl.', and CV.VVC in *pí.ɔ́ɔ́l* 'cold'.

The most common word syllable structure in Tegem consist of 1-3 syllables, e.g., CVC in *jîr* 'queue pl.', CCVV.CV in *kwíí.ʈɔ́* 'nostril' and CV.CVV.CV in *là.máá.ɽì* 'walking'. Then, less frequently, words with 4-5 syllables for example, CV.CV.CV.CVC in *pì.mì.ɟú.líl* 'smooth pl.', CVV.CV.CVV.CVC in *pìɛ̌.pú.lúù.rìl* 'orange', and CV.CVC.CV.CVC.CV in *mì.ʈɛ̀m.mɔ́.rɔ́m.mʌ̂* 'pickaxe'. Though, the maximum may be up to six syllables in rare examples such as the second unit CV.CVVC.CV.CV.CV.CVC in *kìlíɛ̌góm ɽì.ɽéɛ̀ʈ.ʈì.lí.ʈí.lìʈ*˥ 'six'. The complex sequence of consonant gemination [C.C] is addressed in the next paragraphs.

4.3.3. Consonants and vowels distribution patterns

There are no specific segmental categories specified for each syllable position for most of the above syllables. Though, there are frequent occurring segments in five of Tegem syllable types. The coda in the syllable types VC and VVC is commonly nasal or liquid sonorant and occurs in monosyllabic plural gender words only, e.g., *ém* 'goat pl.', *ɔ́r* 'fox pl.', and *ɛ̌ŋ* 'mouth pl.'. The coda in the syllable types CCVC is also frequently sonorant, but similar to the VC and VVC, sometimes an unreleased obstruent may occur as well, e.g., *kwʌ̂l* 'wind', *ɽwʌ̀m* 'bone', *lwʌ̂r* 'writing' and *rwʌ̂ŋ kɔ́ɬ.ɬɔ́ɔ̀.ɟwìk*˥ 'think'. In the three-vowel sequence syllable CVVV, one of the vowel sequence components is always the (short/long) close front vowel [i] in the syllable peak, e.g., *ɟíɛ̀ɛ́* 'bite/year', *ɲûîɔ̌* 'well-adjusted', *ɽwàáì* 'tying'. Then, in Ç.CCVC, the peak is either one of the front or central vowels, e.g., *ɱ̀.bwʌ́k*˥ 'suckle!' and *ɳ.ɖwɛ̀ʈ.ʈʌ̀ŋ* 'cut!'.

Tegem has a morphologically complex word structure since most of the verbs, adjectives, pronouns, and nouns consist of more than one lexeme: compounding and/or affixation. Therefore, it is not easy to determine the root/stem of the word especially for the verbs and the adjectives. When observing some of the disyllabic nouns and adjectives which are apparently made of roots only, there is a pattern of vowel harmony based on vowel backness (±back): front, central, and back vowels. The harmony patterns of these vowel sets are presented in the following table 9:

Table 9. *Tegem vowel distribution in CVCV(C) roots of nouns and some verbs*

		First syllable (V_1)									
		i	ɪ	e	ɛ	a	ʌ	u	ʊ	o	ɔ
Second syllable (V_2)	i	X			X						
	ɪ		X								
	e	X				X					
	ɛ	X			X						
	a	X				X					
	ʌ						X				
	u							X			
	ʊ										
	o							X	X	X	
	ɔ							X			X

This pattern is not necessary imposed on suffixes: *kòɽǒŋ* → *kòɽǒŋ-í* 'worm' and *kàrǎŋ* → *kàrǎŋ-í* 'nose'.

Table 9 shows how frequently the disyllabic roots tend to comprise either front, back, or central vowels, i.e., no front and back vowels are allowed to co-occur in the two syllables of the same root.

4.3.4. Segment sequences in a syllable

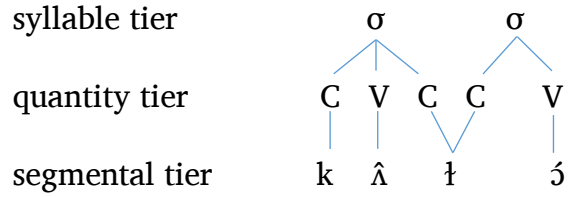
Tegem speakers prefer open syllables word-medially, therefore consonants in the intervocalic position belong to the onset of the rightward syllable such as: CV.CVVC in

pʌ̀.gʌ̀ʌ́l ‘few’ and CV.CV.CVC in *pí.rí.lʉ̀l* ‘pure/clean’. If there are consonant sequences intervocalically consisting of two consonants (excluding the geminates and three consonant sequences), the whole cluster is usually syllabified to the onset of the rightward syllable conforming to the phonotactic of the language which allows consonant clusters in the initial position only, e.g., [ɽw] in V.CV.CCVC *à.n̰dú.ɽwìm* ‘mosquito’ and CV.CCVC *n̰ù.gwʌ̀r* ‘scratch’.

The geminate consonants which occur in intervocalic positions in Tegem belongs to both leftward and rightward syllables at the same time i.e., they are part of both codas and onsets of the two syllables. For instance, in the disyllabic unit *kʌ̂ɬɬɔ́* ‘rain’ the geminated consonant [ɬɬ] functions as coda in the initial syllable [kʌ̂ɬ-], and as onset in the final syllable [-ɬɔ́] as explained in the following diagram (1):

Diagram (1)

kʌ̂ɬɬɔ́



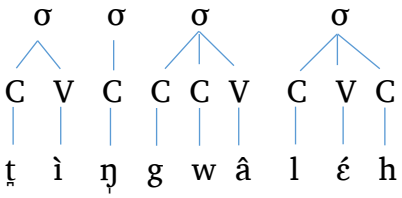
The sigma σ above refers to the syllable. The quantity tier is linked to a single unit when it is short (non-geminated) and to two quantity units when it is lengthened (geminated). In the diagram the V functions as the peak of the syllable and C functions as its margins (onset/coda). Such an intervocalic geminated consonant in Tegem has a dual function – coda in syllable 1, and onset in syllable 2. This may be represented in the following rule: $[_1$kʌ̂ $[_2$ɬɬ] $_1$ɔ́$]_2$ (σ1 = [kʌ̂ɬ], σ2 = [ɬɔ́]). Therefore, geminates in Tegem act as ambisyllabic/bidirectional consonants.

In case of more than three intervocalic consonants in a sequence [-ŋgw-], the initial consonant in the cluster is always a syllabic nasal and the other two consonants are the coda of the leftward syllable [-n̩.gw-]. For example in the four syllable unit *n̰ù.ŋ.gwâ.lɛ́h* ‘lose’ where *n̰ùŋ-* is the first syllable, *-ŋgw-* is an intervocalic three-consonant sequence in

which –*ŋ̩*- is a syllabic consonant. That word syllabification is shown in the following diagram (2):

Diagram (2)

t̪ìŋ̩.gwâ.lɛ́h

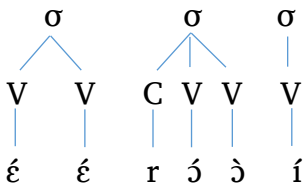


Thus, *-ŋ̩-* functions as a syllabic consonant and forms the second syllable. The sequence *-gw-* is the onset of the following open syllable *-gwâ-*.

The front close vowel [i] in three-vowel sequences in final open syllable sometimes functions as a separate syllable when it is not part of the lexeme root, e.g., *ɛ́ɛ́.rɔ́ɔ̀.í* 'These are heifers.'

Diagram (3)

ɛ́ɛ́.rɔ́ɔ̀.í



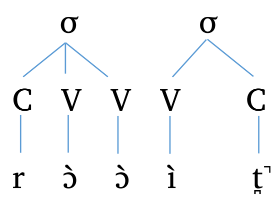
In the word medial position, the three-vowel sequences usually occur as a combination of long vowel and a short vowel: -iɔɔ-, -ɔɔi-, -ɔɔɛ-, -aai-, and -iiɛ-. These sequences occur in disyllabic word in the medial position where the long vowel belongs to one syllable and the short vowel belongs to the other, i.e., they are VV.V or V.VV not VVV, e.g., *rɔ̀ɔ̀.ì̪t̪*[7] 'water stream pl.' (cf. *Diagram 3* above). This does not include the open syllable monosyllabic words with three vowel sequence when they are in compound words or followed by the fricative [h], e.g., *ŋì.ɲíté.ŋǎl* (gloss: prefix - dough - homeware) 'iron' and *ƭíɛ́ɛ̀h* 'bird pl.'[24] where [h] is an optional segment and not a phoneme in Tegem. The

[24] This is another plural form for 'bird' in addition to *æɾíɛ̆*'.

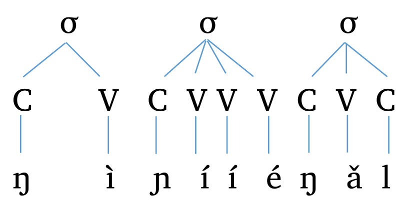
following diagrams (4a, 4b, & 4c) provides further explanation for the syllables of the previous words:

Diagram (4)

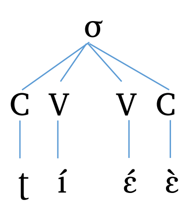
(a) *rɔ̀ɔ̀.ìt̪˺*



(b) *ŋì.ɲíí é.ŋǎl*



(c) *t̪íɛ́ɛ̀h*



The above examples shows that the three-consonant sequence syllable can be part of disyllabic (4a), multisyllabic (4b), or monosyllabic (4c) words.

4.3.5. Syllabic consonants

The syllable onset consonant sequence with a nasal [n], [m], [l], and [ɬ] results from the loss of a following vowel. Thus, it may retain some of the characteristics of the missing vowel such as the tone, as in *ḿbíŋ* ‘hit!’, *ńɟɔ̀ɔ́* ‘stab!’, *ĺ̥ɽíɟʎ́ʎ̀* ~ *lɽíɟʎ́ʎ̀* ‘straight’, and *ɬ̥àɽǎbìnd̪ɔ́ŋ* ~ *ɬɽǎbìnd̪ɔ́ŋ* ‘freeze’. The latter two examples can be pronounced with an extra short vowel. Alternatively, the vowel can be dropped and the two sonorants will be pronounced as a cluster without any tone on the laterals. The initial nasals [m] and [n] bear the syllabicity and tone of the elided vowel. This is a common feature of some African languages such as Engenni (an Edoid language of Nigeria) (Clements 2000:140).

One of the syllable types found in several African languages are the syllabic nasals which are derived from a full underlying NV syllable (N = nasal & V = verb) (Clements 2000:140), such as in the above examples. Childs (2003:67) described that process as an alternation from an onset consonant to a tone bearing unit when the vowel is lost, i.e., it becomes a prominent phone assuming the vowel property. In Tegem, nasals in [mb], [nɟ], [nɖ] and [nd̪] in syllable onsets are syllabic consonants and bear the absent vowel tone. The laterals in the clusters [lɽ] and [ɬɽ] are either seprated from [ɽ] with an extra-short vowel [lɪ̥́ɽ]/[ɬa̰ɽ] or become syllabic laterals in consonant cluster without bearing the tonal property of the lost vowel. In the nasals, it occurs in the onset cluster of imperative verbs and the laterals are not restricted. The following are the syllabic consonants attested in Tegem:

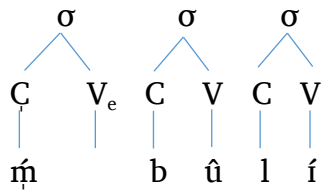
Example (10)

m̩ in *ḿ̩bûlí* 'throw!'
n̩ in *ń̩ɖʒ́gìt̪̀ì* 'wipe!'
n̩ in *ń̩d̪ɛ̌ŋ* 'look here!'
n̩ in *ń̩ɟɔ̀ɔ́* 'stab!'
l̩ in *lɪ̥́ɽɟʌ́ʌ̀* ~ *l̩ɽɟʌ́ʌ̀* 'straight',
ɬ̩ in *ɬa̰ɽăbìnd̪ɔ́ŋ* ~ *ɬ̩ɽăbìnd̪ɔ́ŋ* 'freeze'.

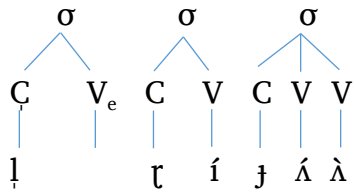
The syllable diagram for syllabic nasals may look as follows:

Diagram (5)

(a) *ḿ̩bûlí*



(b) *l̩ɽ́ɟʎ̀ʌ̀*



In the above diagrams V_e refers to the elided vowel after the syllabic lateral [l̩] and the nasal [m̩] and the diagram can be drawn with the syllabic nasal node only excluding the empty vowel node.

4.4. Tone

Schadeberg (1981a:79) identified four prominent tones types in Tegem: high (H), low (L), falling (F), and rising (R), such as (H) in *kúú* (L) in *kò* 'kind of insect', (F) in *kúù* 'light', and (R) in *kùú* 'grave'. He also mentioned that there is a falling-rising (FR) tone which occur due to affixation or complex morphological conditions[25] on the triple-long vowels, e.g., *lîî-í* 'this is grass' *kîîíʈʈɔ́* 'morning' (*cf. 4.2.3.*). In addition to those tones, the current study also distinguishes a rising-falling (RF) phonetic tone melody conforming to the same above conditions, e.g., *ɲìí-ì* 'spear'. There are two contour tones and both of them may occur on long and short vowels (Schadeberg 1981a:79). Thus, when the above mentioned FR and RF tones occur on both triple-long and long vowels as well, e.g., *mîí* 'a pile of soil pl.', *cîí* 'wire', *bîì* 'bee'. Schadeberg (ibid) also noted that in some falling-rising variants there is a downstep of falling-high when there is an intervening consonant e.g., *rúùí* ~ *rúú*ꜜi 'guts', *ʈ̪úùrí* ~ *ʈ̪úúr*ꜜi 'belly'.

Table 10. *Tegem lexical tone inventory*

Tone pattern	**example**	**gloss**	**example**	**gloss**
high (H)	*óm*	hand pl.	*kúú*	Dileeb tree frond
low (L)	*òm*	Gumgum (bird) pl.	*kùù*	heat
falling (F)	*ɔ̂m*	mountain pl.	*kúù*	light
rising (R)	*ɔ̌m*	gay pl.	*kùú*	grave
falling-rising (FR)	*ʈîí*	pile of soil	*ʈíěh*	ear
Rising-falling (RF)	*ʈì̂l*	cowbird	*ʈìêh*	shoulder

The lexical tones in Table 10 indicate that Tegem is a tone language, i.e., tones are used to create lexical distinctions while affixation is used for grammatical distinction. The language most frequently uses the two level tones (H/L) and then the two contour tones (F/R). The other two contour tones (FR/RF) are used occasionally and they contrast with

[25] Referring to the complexity of the morphological structure of nouns such as *kîîʈ̪ 'night'*.

the other two contour tones, e.g., FR in *ʈîĭ* 'pile of soil' and F in *ʈî* 'kind of grass'; RF in *ʈîl* 'cowbird' and R in *ʈìĭl* 'acacia wood'.

The tone domain (TBU: tone-bearing unit) in Tegem is the syllable as change of one syllable tone changes the word meaning, i.e., 'in a tone language each syllable of a word will have its own tone.' (Dixon 2009:279). However, the tone level/contour as well as the vowel length may change when the suffixes *-gɔ́* and *–í* (meaning 'this') attached to the word, e.g., *ríh* → *ríí-gɔ́* 'branch pl.' *rɔ̀ɔ̀* → *rɔ̀ɔ́ɔ́-í* 'arm pl.', *ɟîr* → *ɟiìr-í* 'queue', *pòlĭl* → *pòlíll-í* 'bad'. This change in tone pattern also suggests that the contour tones in Tegem behave as a combination of independent level tones such as HL → F and LH → R, e.g., *ímbŭkˀ* 'wife' → *ímbùúg-ì* 'this is my wife'. The FR and RF also can be considered combinations of three level tone: HLH → FR and LHL → RF.

Table 11. *Tegem tone association on disyllabic noun roots*

Tone association	**example**	**gloss**
RH	*kŏrɔ́*	country
HR	*pɔ́ʈɔ̌kˀ*	frog
FH	*bûʈúm*	lizard
HF	*múŋûr*	Al-madeika tree pl.
LL	*mìrìŋ*	trachea
HH	*kílér*	food
LF	*mùʈɔ̂ŋ*	day
LR	*mùŋŭr*	cheek pl.

According to the tone association pattern in Table 11, disyllabic CVCV(C) noun roots do not allow two contour tones and no two different level tones are allowed on the two syllables. The association either include contour tone and a level tone , e.g., RH or two similar level tones, e.g., LL. There are no restrictions on the verbs and adjectives tone assciations because they usually occur with affixation or as combined units.

4.5. Morphophonological alternations

This section discusses some morphophonological alternations in Tegem consonants with specific focus on the consonant sequences and geminates. It also identifies their occurrence across different word categories, mainly, the nouns, adjectives, and verbs singular/plural forms.

Schadeberg (1981a) distinguishes between what he called two-class genders (singular/plural) and one-class gender noun classes. There are morphophonological processes in Tegem especially in singular/plural gender distinction including consonant alternation and deletion and vowel quality change and lengthening. One or more phones in the consonant (or consonant cluster) in the singular gender may be deleted and/or changed in the plural gender. These processes occur with both consonants and vowels. Though, they are frequent word-initially, they may also be attested in medial and final segment(s).

Alternation refers to a variation in the phonological realisation of one or more segments of a word. Each segement variant is referred to as an alternant (cf. *4.5.1.1.* below) (Crystal 1996:21). The alternation can be conditioned by the phonological and/or morphological environment.

4.5.1. Phonologically conditioned alternations

Phonologically conditioned alternations are characterised by the phonological environment that the segment occurs in when it takes on one or more feature of the neighbouring segment. Two phonologically conditioned alternations discussed in this section are: voicing assimilation and labialization.

4.5.1.1. Voicing assimilation

In Tegem voiceless consonants assimilate to the preceding voiced consonant and become voiced such as in the following examples:

Example (11)

nʈ → nɖ *ʈwáàj* → *nɖwáàj* 'tie!'

nt̪ → nd̪ *t̪úlúh* → *nd̪úlúh* 'smell!'

nc → nɟ *cúú-ɟʌ́ŋ* → *nɟúúɟʌ́ŋ* 'dig!'

nk → ng ~ ŋg *kɔ̌bér* → *ngɔ̌bér* 'break it!'

The initial consonant agrees in voicing with the attached imperative nasal consonant prefix and becomes voiced. The alternants *ng* and *ŋg* are free variants but *ng* tends to occur more frequently in the initial position and *ŋg* occurs word-medially, e.g., *ʈáŋgîn* 'lizard'. This assimilation is more common in imperative verbs when a nasal prefix is added to a verb stem beginning with a consonant as shown above.

4.5.1.2. Labialization

Labialization means a secondary articulation which results in a noticeable lip-rounding (Crystal 1996:264). In Tegem, labialization take place on any consonant followed by the labial glide consonant [w] but it does not occur on consonants followed by any vowel. The following are examples of labialization encountered in Tegem:

Example (12)

mʷwɛ̂rɛ̀ 'run'

t̪ʷwìt̪̪ê 'lying'

ʈṵ̀rʷwʌ̌h 'crow'

ɲìlʷwɛ́ 'life'

ʈʷwěl 'calabash bowl'

t̪ìbʷwàh 'barking'

The previous consonant labialization is not phonemic in Tegem; rather, labialized consonants are in free variation with their non-labialized counterparts (*cf. 5.2.*). Though, it can be phonemic in other African languages such as in, Kalabarj Ijo, Twi and Lumun. Kalabarj Ijo has labialized velars that contrast with the non-labialized ones (Clements 2000:129).

4.5.2. Morphologically conditioned alternations

The morphologically conditioned alternations discussed in this section are mainly in the singular plural alternations with consonant sequences in Tegem.

4.5.2.1. [bb] and [bbw]

These two sequences has only been attested in the names of the domestic animals and in adjectives. In the two-class gender, the geminated consonant [-bb-] in the singular are replaced by the consonant [-r-] in the plural gender, e.g.,

Example (13)

bébbú (sg.) *éérú* (pl.) 'dog'
bɛ̀bbɔ́ɔ̀-í (sg.) *éérɔ́ɔ̀-í* (pl.) 'heifer pl. (2-3 years old)'
pèppwʌ̀ (sg.) *ɛ̀ɛ̀rɔ̌* (pl.) 'goat (male young)'

These examples shows that there are three alternation processes that take place to form the plural gender for those words:

(a) singular prefix consonant deletion *bé-bbú* → **ébbú,*
(b) singular prefix vowel lengthening *é* → **éébbú,* and
(c) geminated consonant change *-bb-* → *éérú.*

In the third example, [w] is deleted and [-bb-] replaced by [-r-]. Usually in Tegem the non-geminated variant [b] is retained word-medially in the plural gender such as in *ṯíbìí* (sg.) and *ríbìí* (pl.) 'spear'; *ƭúbúƭéɲ* (sg.) and *kúbúgèɲ* (pl.) 'tusk'.

4.5.2.2. [mb] and [mbw]

The consonant sequence [mb] has been attested in the imperative verbs (word-initially only (*see 4.1.1.*)), the kinship nouns (always preceded by the high front vowel [í]), and in the compound nouns (adjective + noun).

In the two-gender class, there is deletion and replacement in plural gender for the sequence [mb], e.g.,

Example (14)

ímbŭk˥	(sg.)	*ílŭk*˥	(pl.)	'wife'
bíámbál	(sg.)	*jɛ́llál*	(pl.)	'big'
pɛ̀ɽɛ̂mbn̂	(sg.)	*kɛ́ɛ̀ln̂*	(pl.)	'big ox'
íbɛ̂mbwɛ́ŋ	(sg.)	*íbɛ́ɛ́lɛ́ŋ*	(pl.)	'child'

There are 2-3 sound alteration processes involved in the formation of plurals with the sequence [mb] in nouns and adjectives- deletion, change and vowel lengthening. In the first example there are two processes (a) the consonant in the possessive marker[26] [ìm-] has been replaced by [l] in the plural → [ìl-], and (b) the first consonant of the noun [-bŭk˥] has been deleted (which is common in Tegem noun class system) in the plural gender → [ílŭk˥]. In the third example, where [bí] 'ox' preceded by the prefix [pɛ̀-] and adjective [-ɽɛ̂m-], there are three alternation processes:

(a) the final sound [m][27] in the adjective is replaced by [l] and the initial [ɽ-] by [k-] → **pɛ̀-kɛ̂l-bn̂*
(b) the initial consonants in the noun [b-] and the singular prefix [pɛ̀-] are deleted in the plural gender → **kɛ̂l-n̂*, and
(c) the adjective short vowel [-ɛ̂-] is lengthened in the plural gender where the contour tone spread into two level tones → *kɛ́ɛ̀ln̂*.

[26] This marker has been identified by Schadeberg (1981a:80)
[27] It may also be a separate syllabic m̩ not part of the adjective.

Then, in the sequence [mbw] in *ɪ̂bêm-bwéɲ* (sg.) *ɪ̂béél-éɲ* (pl.) 'child'- [m] changed to [l], both [b] and [w] has been deleted, and the vowel lengthened in the plural.

4.5.2.3. [bw] ~ [pw]

The sequence [bw]~[pw] occurs in two-gender classes in Tegem. In the plural form the sequence is deleted such as in the following word:

Example (15)

pi̯-ɽɪ̌-bwéɲ
prefix-eye sg.-child sg.
'eye pupil sg.'

ɟìì-éɲ
eye pl.-child pl.
'eye pupil pl.'

The above example (15) is a two-gender compound noun with the gloss and translation provided where (sg.) refers to the singular, (pl.) is the plural. There are three alternation processes applied to get the plural gender:

(a) the deletion of the singular prefix *pi̯-* and the consonant sequence *-bw-* → **ɽɪ̌-éɲ,*
(b) the replacement of the singular consonant *ɽ-* in *ɽɪ̌* 'eye sg.' with the plural consonant *ɟ-* → *ɟɪ̌* 'eye pl.' → **ɟɪ̌-éɲ*, and
(c) the vowel of the noun *-ɪ̌* is lengthened in the plural → *ɟìì-éɲ*

In Tegem, the single consonant [b] ~ [p] is deleted word-initially in plural gender, e.g., *bóɽwìì bóó* (sg.) and *mwìì óó* (pl.) 'Kujur (traditional healer)'. The glide [w] in [kw] cluster in singular is retained in plural gender, e.g., *kwíɪ̤tɔ́* (sg.) and *wíígɔ́* (pl.) 'nostril'*; kwɲ́* (sg.) and *wɲ́* (pl.) 'room'. Thus, the two consonants of a cluster are not always deleted.

4.5.2.4. [t̤w]

The consonant sequence [t̤w] occurs in two-gender class nouns and adjectives only. The two cluster phones in the singular undergo alternation of [t̤] to [r] and deletion of [w]

in the plural. The following examples are of two-gender nouns where [t̪w] has been attested word-initially:

Example (16)

t̪wèɛ́ɽɛ̀ (sg.) *rèɛ́rɛ̀* (pl.) 'friend'
t̪wèh (sg.) *rèh* (pl.) 'skin'
t̪wààìt̪˥ (sg.) *rɔ̀ɔ̀ìt̪*˥ (pl.) 'water stream'

In each example in (16), there are two alternation processes, such as in the first example:

(a) deletion of the second consonant in the cluster *-w* → **t̪èɛ́ɽɛ̀*, and
(b) change of the cluster initial consonant *t̪-* to *r-* and the second syllable onset from *ɽ-* to *r-* →**rèɛ́rɛ̀*.

It is common in Tegem singular/plural genders that the single unit dental stop *t̪* changes to *r* in the plural such as in *t̪àà* (sg.) and *ràà* (pl.) 'firewood'; *t̪íbìí* (sg.) and *ríbìí* (pl.) 'spear'. Though, when the sequence *t̪w* occurs in the medial position such as in *lúṯ̪wɅ̂r* (sg.) 'thread', *t̪* is retained without change while, *w* is deleted in the plural gender *ɬɅ́t̪Ʌ̂r* (pl.) 'thread', i.e., *w* is always deleted in the plural gender. Thus, when *t̪w* occurs in the initial position two change applies to the cluster to form the plural whereas, intervocalically only one change apply.

4.5.2.5. [nd̪]

The sequence [nd̪] occurs mostly in verbs and sometimes in nouns, adjectives and pronouns. In some verbs it may be an imperative prefix, e.g., *t̪úlúh* 'smelling' and *nd̪úlúh* 'smell!' (*see 4.1.1.*). In the two-gender nouns, it has been attested in one word where [nd̪] in the singular gender changes to a geminated [ll] in the plural as follows: *t̪ûnd̪Ʌ̂* (sg.) and *rûllɅ̂* (pl.) 'waste dump'. In two-gender nouns, when not in a cluster, the consonant

[ḓ] in the singular is retained in the plural, e.g., *kɔ́ḓɔ̀ɔ́* (sg.) and *áḓɔ̀ɔ́* (pl.) ‘son/daughter’; *kúḓǒm* (sg.) and *áḓǒm* (pl.) ‘door’.

4.5.2.6. [ʈʈ] and [mm]

The consonant sequence [ʈʈ] occurs in nouns and adjectives. In the two-gender plural form, the single unit [ʈ] (including its allophone [ɽ]) in the singular is changed to [m] and the same apply for the geminated one [ʈʈ] → [mm]. Such as in the following examples: *ʈúŋúɽèʈʈôl* (sg.) and *múŋúmèmmôl* (pl.) ‘cooking pot’; *ʈěʈʈèɛ́* (sg.) *měmmèɛ́* (pl.) ‘calabash’.

The sequence [mm] occurs in verbs, nouns, and adjectives. Thus, it has been found in two-gender (singular/plural) classes. The geminated bilabial [mm] occuers in the plural gender and always corresponds to the geminated retroflex consonant [ʈʈ] in the singular as shown in the previous examples.

4.5.2.7. [mw] and [ʈw]

The consonant cluster [mw] occurs in plural and mass uncountable nouns, with only two exceptions, i.e. the singular noun *kúmwʌ̀ʈ̚* ‘locust’ and the verb *mwêrɛ̀* ‘run’. Thus, when one considers the two-class gender pattern of sound change for [mw], it may correspond to the sequence [ʈw] which has been attested in singular nouns only as shown in example 17 below,

Example (17)

ʈwʌ̀m (sg.) *mwʌ̀m* (pl.) ‘bone’
ʈwʌ̀k̚ (sg.) *mwʌ̀k̚* (pl.) ‘ankle’
ʈwíí (sg.) *mwíí* (pl.) ‘forehead’

In example (17), there is an alternation in the plural gender of *ʈ* to *m* while *w* is retained. Thus, the consonant sequence [mw] is almost limited to a plural gender noun class where

it occurs in word-initial position. The sound alternation pattern resonates with the single consonant [ʈ] → [m] change pattern for singular/plural gender distinction.

4.5.2.8. [nɖ] and [nɖw]

Both the consonant clusters [nɖ] and [nɖw] occur most frequently in the imperative verbs such as *ńɖʌ́ʌ̀* 'take!' and *kɛ̀ɛ́nɖàá* 'turn back!' (*see 4.1.1.*). It can also occur in adjectives, nouns, and adverbs. Both the nouns and adjectives have two genders but the cluster [nɖ] occurs usually in verbs so there is only one adjective with two-gender class in the available data - *ʈínɖèè* (sg.) and *míndɛ̀è* (pl.) 'it is full'. In the previous example the cluster [nɖ] is retained in the plural gender word-medially.

4.5.2.9. [kw] and [gw]

These two consonant clusters occur in nouns and sometimes in verbs. In two-gender nouns, the word-initial sequence [kw] occurs in singular gender only. In the plural gender there is a deletion of the velar [k], e.g., *kwɨ́ɲ*⁷ (sg.) and *wɨ́ɲ*⁷ (pl.) 'room'; *kwɛ̌l* (sg.) and *wɛ̌l* (pl.) 'moon', while [w] is retained after the front vowels. Two or three alternation processes occur in the plural when the cluster [kw] is followed by a central vowel[28] in the singular gender, such as in the following instances:

Example (18)

kwʌ̂m (sg.) *ɔ̂m* (pl.) 'mountain'
kwʌ̀ (sg.) *bɔ̀ɔ̀* (pl.) 'tree'

There is consonants deletion and vowel replacement in the previous examples:

(a) the central vowel *ʌ* quality changes and replaced by the back vowel *ɔ* in the plural → **kwɔ̂m / kwɔ̀*
(b) the cluster kw- does not allow a following back vowel so the cluster is deleted in the first example → *ɔ̂m*

[28] It has never been attested followed by a back vowel.

(c) In the second example the cluster is deleted in the plural and a plural prefix *p* is inserted instead then the vowel lengthened[29] → *bɔ̀ɔ̀*

The intervocalic variant [gw] occurs in both singular and plural gender as a sequence without deletion but there is prefix vowel alternation (*i* → *a*) and consonant deletion, e.g.,

Example (19)

pɪ̰́tɪ́gwʌ́r (sg.) *á̰tógwʌ́r* (pl.) 'king'
kʌ̌lúgwʌ́ʌ̀ (sg.) *ʌ̌lúgwʌ́ʌ̀* (pl.) 'womb'

4.5.2.10. [ng] and [ŋg]

These clusters occur as a result of the addition of either [n] or [ŋ] before imperatives and compound nouns with initial velar stop [g], as in the following example (20):

Example (20)

ń-góm-lɔ́gɛ́
imperative prefix-hand-put in
'put in your hand (upward)'

kì-ɽì-ŋgĕn
prefix-heart-mouth
'upper part of the belly'

In the previous phrase and compound noun, the consonant [n] ~ [ŋ][30] is attached to the first syllable of nouns roots for morphological purposes in *kóm* 'hand' and *kĕn* 'mouth'.

[29] The lengthening may be because it is an open vowel. Addtionally, one may argue that there is an insertion process as well in this example.
[30] An alternative analysis is possible where [n] and [ŋ] are not free variants when in a cluster. The consonant clusters preceded by [n] tend to occur in the onset of imperative verbs which supports a hypothesis that it is part of an imperative prefix (*cf. 4.1.1., 4.5.1.1., 4.5.2.5* & *4.5.2.8*).

In two-gender nouns the sequence is found in both singular and plural genders of the noun without any alterantion or deletion, e.g., *ʈáŋgèè* (sg.) and *máŋgèè* (pl.) 'gecko'*; ʈáŋgîn* (sg.) and *máŋgîn* (pl.) 'kind of lizard'.

4.5.2.11. [rw]

The sequence [rw] occurs in two-gender nouns. In the plural gender there is no alternation or deletion for any of the consonant cluster segments [rw], e.g., *kùrwʌ̀ŋ* (sg.) and *ʌ̀rwʌ̀ŋ* (pl.) 'name'. The consonant [r] as single phoneme also retained in the plural gender such as in *kàrăŋ* (sg.) and *àrăŋ* (pl.) 'nose'. The second cluster consonant [w] is deleted intervocalically in some clusters' plural gender as in *lúʈwîr* (sg.) and *ɬáʈîr* (pl.) 'thread' and retained in others, e.g., *kúmwʌ̀ʈ˺* (sg.) and *ámwʌ̀ʈ˺* (pl.) 'locust'.

4.5.2.12. [lw]

The sequence [lw] occurs in nouns and less frequently in verbs. In two-gender words, it behaves differently in the initial and medial position. In the initial position, there is a deletion for the glide [w] while, in intervocalic position the sequence is retained in the plural gender, e.g.,

Example (21)

lwèè (sg.) *lèè* (pl.) 'hole'
lwœ̀r (sg.) *lœ̀r* (pl.) 'cloth'
búlwàk (sg.) *álwàk* (pl.) 'Kujur[31] / religious leader'
púlwàr pɲ̀nór (sg.) *ɲúlwàr ɲór* (pl.) 'tiger'.

When not in a cluster, the consonants [w] and [l] are most common in the plural gender than in the singular, e.g., *kwì* (sg.) and *wì* (pl.) 'breath'; *ʈéél* (sg.) and *méél* (pl.) 'neck'; *ìmĭl* (sg.) and *ìlĭl* (pl.) 'husband'. Sometimes [w] is changed to another consonant in the plural, e.g., *wɛ́l* (sg.) and *kɛ́l* (pl.) 'snake Kujur'.

[31] *Kujur* refers to traditional healers and spiritual men in the Nuba Mountains.

5. Conclusions and findings

This section includes summaries, discussions, and results of the analysis in the previous section of the study. It focus on Tegem phonological characteristics the study found.

5.1. Tegem consonants phonemes

In Tegem, consonants distribution is restricted i.e. some consonants does not occur in all three word positions. The items in table 12 show the distribution of the consonants, including both the phones and phonemes, in word-initial, intervocalic and word-final positions. The broken dash in the table indicates that the consonant is not attested in that position. The consonants are organized according to their place and manner of articulation.

Table 12. *The distribution of consonants within Tegem words*

	word initial	**gloss**	**intervocalic**	**gloss**	**word final**	**gloss**
b	*bɛ́lȋl*	new	*t̪ɪ̀ɽŏbăh*	splitting	*t̪wǽb˺*	road
p	*pàgààl̂ɪ̂l*	other	*pɪ̏ɛ̆púlúùrɪ̀l*	orange	*kép˺*	worm
t̪	*t̪wɛ́r*	throat	*kȋt̪ɛ̂ɲ*	laugh	*l̏ɪgɪ̏t̪˺*	face
d̪	-	-	*kʌ̀d̪ʌ̆ɬ*	tortoise	-	-
ʈ	*ʈɪ́ɽɪ́ŋɔ̌*	thick	*kɪ̃̏ɪ̂ʈʈɔ́*	morning	*ŋʌ̌ʈ˺*	dirt
c	*cwéè*	drink!	*cɔ́cɔ́ŋ*	sixth child	-	-
ɟ	*ɟɪ̏ɽêl*	knee	*t̪ʊ́ŋgȋɟóŋ*	fight!	-	-
k	*kɪ̏lɪ̂ɛ̆góm*	five	*kȋkkáh*	with	*kȋlék˺*	move!
g	-	-	*kúbʊ́gèɲ*	tusk	-	-
h	-	-	-	-	*mȋɛ̆h*	earl pl.
m	*mɛ́mmɛ́m*	short pl.	*kɪ̏mèè*	man	*kóm*	hand
n	-	-	*ìmɪ̆n-í*	husband	*nɖâbȋn*	cut (axe)
ɲ	*ɲɛ̀ɛ̀t̪ɪ̏*	because	*ɲúùɲú*	spit	*ńd̪ɛ̆ɲ*	look here
ŋ	*ŋɪ̀ɟɪ̂m*	salt	*mɪ̏ŋɛ́ɛ́máà*	peanut pl.	*rwʌ̂ŋ*	word
r	*rɪ́h*	branch pl.	*bȋrɪ́àl*	flying	*bɪ̀r*	rabbit

ʈ	-	-	*pîʈîĕ*	bird	-	-
l	*lḭ̈tílḭ̈tˀ*	together	*bîṭ̈tîlïl*	green	*kél*	snake pl.
ɬ	*ɬáḍ̈ˀ*	here	*ɬùùɬûkˀ*	cotton	*jélœ̌ɬ*	big fish pl.
w	*wíígɔ́*	nostril pl.	*ʈíʈwʌ̀h*	many	-	-
j	*júl*	elephant	-	-	-	-

Table 12 also shows that all stops are unreleased word-finally in Tegem while sonorants are not. Three consonants occur in a mutually exclusive environment: [ḍ̈], [g], and [ʈ] as variants of [ṭ̈], [g], and [ʈ] in the medial position. Then, the only fricative in Tegem [h] occurs word-finally only and the glide [j] occurs in the initial position only.

Table 13 below summarizes the distribution of consonants (phones and phonemes) in word-initial, intervocalic, and word-final positions. The plus indicates the presence of a consonant in the specified position, and the minus indicates the absence of a consonant in that position.

Table 13. *Distribution of consonants*

	consonant	**word-initial**	**intervocalic**	**word-final**
stops	p	+	+	+
	b	+	+	+
	ṭ̈	+	+	+
	ḍ̈	-	+	-
	ʈ	+	+	+
	c	+	+	-
	ɟ	+	+	+
	k	+	+	+
	g	-	+	-
fricative	h	-	-	+
nasals	m	+	+	+
	n	+	+	+
	ɲ	+	+	+
	ŋ	+	+	+

central liquid	r	+	+	+
	ɾ	-	+	-
lateral liquids	l	+	+	+
	ɬ	+	+	+
glides	w	+	+	-
	j	+	-	-

To determine the phonemic status of the above consonant inventory, table 14 below lists further minimal or near minimal pairs of Tegem consonant phonemes.

Table 14. *Consonant minimal pairs in Tegem*

(near) minimal contrast	**example**	**gloss**	**example**	**gloss**
/B/ : /m/	*Búh*	horse	*múh*	breast pl.
/ʈ/ : /ʧ/	*ʈúú*	leg	*ʧúú*	charcoal
/ʧ/ : /l/	*ʧíěh*	ear	*líéh*	foot
/C/ : /ɲ/	*Cɨh*	eye pl.	*ɲɨh*	blood
/C/ : /l/	*ɔ́ɔ́Cɨ*	you pl.	*ɔ́ɔ́l*	baboon pl.
/k/ : /ŋ/	*kǎʈ˥*	cave	*ŋǎʈ˥*	dirt / stealing
/n/ : /ʈ/	*nɖúlúh*	smell!	*ʈúlúh*	smelling
/r/ : /l/	*júr*	feather pl.	*júl*	elephant
/ɬ/ : /l/	*ʈǔɬ*	kind of fruit	*ʈǔl*	cough
/w/ : /m/	*wɨ̃i*	Jummaiz tree	*mɨ̃i*	sacrum pl. (body)
/j/ : /ɲ/	*júl*	elephant	*ɲûl*	clay
/j/ : /C/	*jɨɛˆɔ̌l*	red monkey pl.	*Cɨɛ̀*	razor

The previous tables 12, 13, and 14 summarises the obstruents and sonorants discussion in the data analysis section. Hence, having described Tegem consonants, their distribution within a word, and contrast with each other, one may draw the conclusion that Tegem has 14 consonantal phonemes, as illustrated in table 15 below.

Table 15. *Tegem consonant phonemes*

		bilabial	dental	(post)alveolar	retroflex	palatal	velar
stops		B	t̪		ʈ	C	k
nasals		m		n		ɲ	ŋ
liquids	central			r			
	lateral			l, ɬ			
glides		w				j	

5.2. Tegem consonant sequences phonemic pattern

The consonant sequences comprising of *Cw* and *NC* can be analysed as prenasalised and labialized consonants respectively, otherwise as sequences of two consonants. Sections *4.1.1.3., 4.3.5* and *4.5.* describe the consonant sequences distributional behaviour within the word and their syllabic and morphophonemic patterns. Accordingly, the phonemic status of the above consonant sequences is determined as a cluster of two distinct consonants.

Their distribution is restricted to the word initial and medial positions only and they do not behave as the single unit phones in these environments (*see 4.1.1.3.)*. For example in the minimal pairs *kwěl* 'moon', *ʈwěl* 'calabash bowel', and *t̪wěl* 'dance', the contrast is between the initial consonants in the cluster *k*, *ʈ* and *t̪* independently from their following glide *w*.

The *NC* sequences consist of a syllabic nasal followed by stops, so the *N* constitute a syllable by itself and the *C* act as the onset of the next vowel (*see 4.3.5)*. Therefore, they do not behave as a single prenasalised segment, such as in the imperative verb *ṅ̩.ɖɔ́.gì.t̪ì* 'wipe!' where there is a syllable boundary between the *N* [ṅ̩] and the *C* [ɖ].

The morphophonemic analysis shows that in the alternation processes the *Cw* sequences does not follow the same pattern of alternation (*see 4.5.*). It is common that the *C* is deleted in the plural form while the *w* is not, e.g., *kwíít̪ɔ́* (sg.) and *wíígɔ́* (pl.) 'nostril'. If we assume that *kw* is a labialized consonant in the previous example, both *k* and *w* are

supposed to be deleted in the plural. Thus, the *Cw* sequences are clusters of two consonants (*cf. 4.5.1.2.*).

The above arguments concerning the consonant sequences phonemic status are based on the analysis findings of this phonological study. To provide a more accurate identification for the nature of those sequences, a detailed morphophonological investigation is needed.

5.3. Tegem vowel phonemes

The vowels in Tegem may be divided into three groups of 4 vowels each: front vowels [i], [ɪ], [ɛ], and [e]; central vowels [æ], [a], [ʌ] and [ɑ]; back vowels [u], [ʊ], [o] and [ɔ]. The distinction between the front and back vowels is apparent from the minimal contrast (*see Table 16 below*) but the phonemic distinction among central vowels is complex. They freely vary with each other in pairs: [æ] ~ [ʌ] and [ɑ] ~ [a] while the first pair contrasts with the second.

Table 16. *Vowels minimal pairs in Tegem*

(near) minimal contrast	**example**	**gloss**	**example**	**gloss**
/i/ : /ɪ/	*kíṯə̱ɲ*	laugh	*kíṯə̱ɲ*	look (there)
/i/ : /e/	*kíkkáh*	with	*kêkkáh*	four
/ i/ : /ɛ/	*bwìr*	fear	*bwɛ̀r*	jump
/ɪ/ : /ʌ/	*kwì*	breath	*kwʌ̀*	tree/stone
/ɪ/ : /u/	*ʈíl*	heart	*ʈúl*	seed/relatives
/e/ : /ɛ/	*búèɲ*	monitor lizard	*búɛ̀ɲ*	rat
/ɛ/ : /a/	*lɛ́h*	where	*láh*	direction
/ɛ/ : /ʌ/	*ṯwɛ̀h*	skin	*ṯwʌ̀h*	home backyard
/a/ : /ʌ/	*pwà*	sugar cane	*pwʌ̀*	goat (male old)
/a/ : /u/	*áḏ̱ì*	they (feminine)	*úḏ̱ì*	they (masculine)
/a/ : /ɔ/	*káh*	finger	*kɔ́h*	tree cavity
/ʌ/ : /o/	*lûłłʌ́ḏ̱ˀ*	very close	*lûłłóḏ̱ˀ*	bowing
/ʌ/ : /ɔ/	*łʌ́ḏ̱ˀ*	here	*łɔ́ḏ̱ˀ*	hanging down

/u/ : /ʊ/	*ʈúh*	bowel	*ʈʊ́h*	breast
/ʊ/ : /o/	*mồl*	fruit pl.	*mồl*	bowel pl.
/ʊ/ : /ʌ/	*mʊ̀*	Al-hemaidh fruit	*mʌ̂*	head pl.
/o/ : /ɔ/	*bór*	snail	*bɔ́r*	fox
/ɔ/ : /u/	*ɬɔ̌*	animal dung	*ɬú*	kind of stick
/ɔ/ : /e/	*ɲûɔ̌*	well-adjusted	*ɲííé*	dough

According to the vowels analysis in the previous section, there are 10 vowel phonemes out of 12 phones in Tegem (*see Table 17*). All the vowels may occur in word initial, medial and final positions. Every vowel in Tegem may occur as short and/or long and in vowel sequence.

Table 17. *Tegem vowel phonemes*

	front	**central**		**back**
close	i			u
	ɪ			ʊ
close-mid	e			o
open-mid	ɛ			ɔ
open		a	ʌ[32]	

Usually the vowels occur word-initially in the plural gender nouns as a result of a deletion of the first consonants in some Tegem plural gender classes. In this case the vowel may optionally be preceded by a glottal stop [ʔ], e.g., *àh* ~ *ʔàh* 'fly pl.', *ɔ́r* ~ *ʔɔ́r* 'fox pl.', *épˀ* ~ *ʔépˀ* 'earth worm pl.'.

5.4. Tegem syllable patterns

Tegem's 13 syllable structures can be summarized as follows:

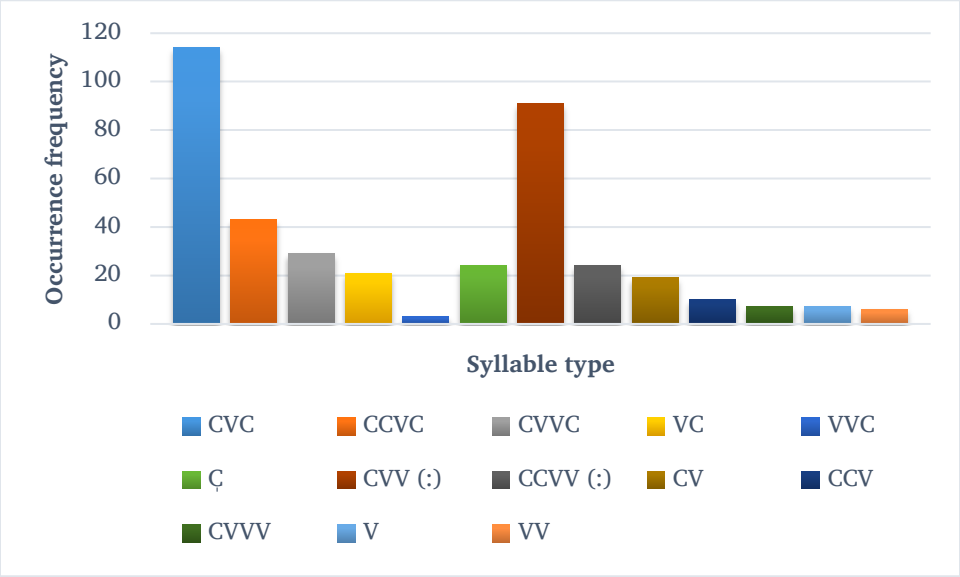
Open syllables: (C)V, (C)VV(V), CCV(V)

Closed syllables: (C)VC, (C)VVC, CCVC, Ç

[32] Following the same pattern throughout the study, when two phones are free variants, the less restricted in distribution and the most frequent in occurrence is determined as the phoneme. Hence, [ʌ] is determined as the representative symbol above (*see 4.2.1.3.a*). It is also used as a central vowel in the literature of the Nuba mountain languages (*cf. Bashir 2015*).

To determine the most salient syllable type, Figure 3 below displays the frequency of different syllable types in monosyllabic words of Tegem:

Figure 3. *Monosyllabic word structure*



The horizontal axis of figure 3 consists of two-group: 6 closed and 7 open syllables where the vertical axis represents each syllable's occurrence frequency. The colon (:) after VV indicates that only the long vowels occur in that syllable. That excludes the other vowel sequences such as *ie* and *ai*. The frequency of open syllables is around 56% and closed syllables makes about 44% of the monosyllabic words. The figure also shows that the most common monosyllabic word structures are CVC with around 30% and CVV with approximately 24%.

5.5. Tegem tone types

Tone as a distinctive feature is a characteristic of many African languages (Clement & Rialland 2008:39). Tone serves lexical and/or grammatical function in those languages. In Tegem, there are lexical tones. For a detailed account of tones in Tegem a post-lexical phonological study is required, i.e. to determine tonal behaviour beyond word level

which also gives further information on Tegem's segmental and supra-segmental phonology.

5.6. Some morphophonological features

Morphology imposes constraints on phonological patterns. Thus, an in-depth investigation of the phonological alternations in Tegem is essential to give full account for the morphophonemic alternations. The analysis (*in 4.5. above*) indicates that there are two kinds of alternations: regular alternations and irregular alternations and they happen across different word categories.

In Tegem, the compound nouns singular gender can have a singular prefix, e.g., *pìɽĭbwɛ́ɲ* 'eye pupil sg.' but the plural does not as in *ɟììɛ́ɲ* 'eye pupil pl.'. This example in the singular it comprises of singular prefix + noun (eye sg.) + noun stem (child sg.) → *pì-ɽĭ-bwɛ́ɲ*. In the plural it consists of noun (eye pl.) + noun stem (child pl.) → *ɟìì-ɛ́ɲ*.

The geminated consonants [t̪t̪] and [d̪d̪] occur in few verbs and nouns, e.g., *mìɛ́t̪t̪ɛ́ɛ́mʌ́ʌ̀* 'kind of beans pl.' and *lwɑ̂t̪t̪ɑ̀kˀ* 'mat sg.'. The dental [t̪] legthening also occurs after syllable with consonant sequence and a short vowel such as *t̪wìt̪t̪ê* 'lying' and *nɖwèt̪t̪ʌ̀ŋ* 'cut! (with a knife)'. The cluster [d̪d̪] may be used in Tegem to express emphasis, e.g., *ɬɔ́d̪d̪ɔ́* 'underneath' and *t̪ʌ̀ʌ́d̪d̪úɬǒ* 'that'. In the first example, the germination gives the impression of being further down and the second to emphasis that something is in a further distant from the speaker.

5.7. The importance of the study

This research provides the first detailed phonological study of Tegem language. The study contributes to the existing literature and linguistic data on the Nuba Mountains languages. An in-depth phonetic analysis will provide a valuable follow-up for this phonological study on both the segmental and supra-segmental levels. I hope this preliminary research stimulate others to pursue further studies on Tegem. In the next phase of the research, I hope to study the morphophonemic features, noun class system,

and syntactic structure of Tegem to write an informed description of this language's grammar.

References

Akinkugbe, O. Olujumi (1978). A comparative phonology of Yoruba dialects, Isekiri and Igala. Unpublished Ph.D. Thesis, University of Ibadan.

Bashir, Abeer (2015). Morphophonological Alternations in Tima. *Occasional Papers in the Study of Sudanese Languages,* No. 11 pp. 31-46.

Bashir, Abeer M.A (2003). Phonological Analysis of Jumjum Language: Phonemic and Generative Approaches. Unpublished M.A. Thesis, Khartoum: University of Khartoum.

Blaho, Sylvia (2008). The syntax of phonology: A radically substance-free approach. Ph.D. dissertation, University of Tromsø.

Blench, Roger (2012). Tegem-Amira: a previously unrecognised subgroup of Niger-Congo. Unpublished manuscript.

Bowern, Claire (2007). Swadesh wordlist, categorized by semantic field. *Linguistic Fieldwork: A Practical Guide*. UK: Palgrave Macmillan.

Clements, G. N. & Rialland, Annie (2008). Africa as a phonological area, In Heine & Nurse (ed.). *A Linguistic Geography of Africa*. UK: Cambridge University Press.

Clements, G. N. (2000). Phonology. In B. Heine & D. Nurse (ed.) (2000) *African Languages: an Introduction*. UK: Cambridge University Press.

Crystal, David (2008). *A Dictionary of Linguistics and Phonetics* (6th ed.). Oxford: Blackwell Publishing Ltd.

Dixon, Robert M. (2012). *Basic Linguistic Theory Volume 3: Further Grammatical Topics*. Oxford University Press.

Dixon, Robert M. (2009) *Basic Linguistic Theory Volume 1: Methodology*. Oxford University Press.

Greenberg, Joseph H. (1950). Studies in African linguistic classification: VII. Smaller families; index of languages. *Southwestern Journal of Anthropology* 6, 388-398.

Greenberg, Joseph H. (1963). *The languages of Africa*. Berlin: Mouton.

Grimes, B. (2003). Kordofanian Languages. In J. William Frawley (ed.). *International Encyclopedia of Linguistics*, (2nd ed.) Vol. 1., 365-366. New York: Oxford University Press.

Hyde, Brett (2002). A restrictive theory of metrical stress. *Phonology* 19, 313–339. Joshua Project. (n.d.). *Lafofa in Sudan*. Retrieved on 28/04/2016 from: https://joshuaproject.net/people_groups/12942/SU

Jurgec, Peter (2010). Feature Spreading 2.0: A Unified Theory of Assimilation. Ph.D. dissertation, University of Tromsø.

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig, (eds.). (2013). *Ethnologue: Languages of the world. Dallas: SIL Internationa*l. (17th ed.) Online version http://www.ethnologue.com

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). (2015). *Ethnologue: Languages of the World.* (18th ed.). Dallas: SIL International. Online version: http://www.ethnologue.com

MacDiarmid, P.A. and D.N. MacDiarmid (1931). The languages of the Nuba Mountains. *Sudan Notes and Records* 14, 149-162.

Manger, Leif O. (1994). *From the mountains to the plains: The integration of the Lafofa Nuba into the Sudanese society.* Uppsala: The Nordic Institute of African Studies.

Morén, Bruce (1999). Distinctiveness, coercion and sonority: A unified theory of weight. PhD thesis, New York: Routledge.

Morén, Bruce (2006). Consonant-vowel interactions in Serbian: Features, representations and constraint interactions. *Lingu*a 116, 1198–1244.

Myers, Scott (1997). OCP effects in Optimality Theory. *Natural Language and Linguistic Theory* 15, 847–892.

Norton, Russell & Alaki, Thomas Kuku (2015). The Talodi Languages: A Comparative-Historical Analysis. *Occasional Papers in the Study of Sudanese Languages,* No. 11 pp. 47-157

Olson, Kenneth S. (2004). An evaluation of Niger-Congo classification. *SIL Electronic Working Papers*.

Persson, Andrew (1984). The relationship of the language of Sudan. *Occasional Papers in the Study of Sudanese Languages*, No. 3 pp. 1.

Schadeberg, Thilo. C. (1981a). *A Survey of Kordofanian: The Talodi Group,* Volume 2. Hamburg: Helmut Buske.

Schadeberg, Thilo C. (1981b). Das Kordofanische. In: B. Heine, T.C. Schadeberg & E.Wolff (eds.), *Die Sprachen Afrikas, Band I: Niger-Kordofanisch.*, 117-128. Hamburg: Helmut Buske.

Schadeberg, Thilo C. (1989). Kordofanian. In: John Bendor-Samuel (ed.), *The Niger-Congo Languages*. 66-80. Lanham: University Press of America.

Schadeberg, Thilo C. & Blench, Roger. (2013). The Nuba Mountain Languages: An introduction. In Roger Blench & Thilo Schadeberg (eds), *Nuba Mountain Language Studies*. Cologne: Rüdiger Köppe. pp.1-24.

Snider, Keith & Roberts, James (2006). *SIL comparative African wordlist SILCAWL*. SIL International.

Seligmann, C. G. (1910). The Physical Characters of the Nuba of Kordofan. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, Vol. 40, pp. 505-524. http://www.jstor.org/stable/2843269

Stevenson, Roland C. (1964). Linguistic research in the Nuba Mountains II. *Sudan Notes and Records* Vol. 45, 79-102. Khartoum: University of Khartoum.

Stevenson, Roland C. (1962-1964). Linguistic research in the Nuba Mountains. *Sudan Notes and Records* Vol. 43 (1962), 118-130.

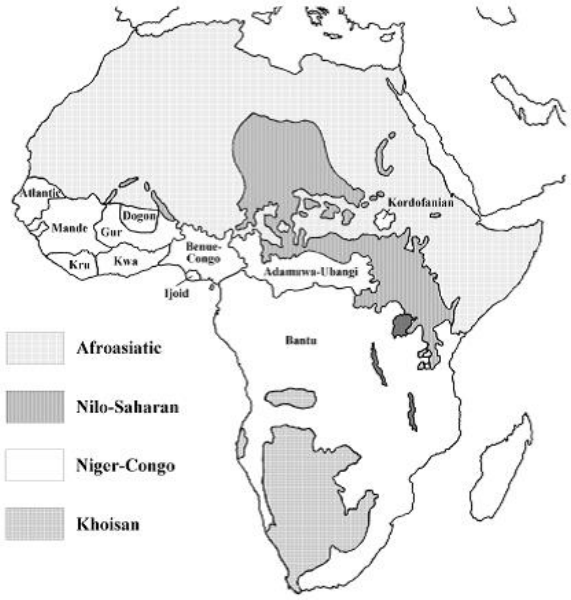
Stevenson, Roland C. (1984). *The Nuba people of Kordofan province: An ethnographic survey*. (Graduate College Publications, Monograph 7.) Khartoum: University of Khartoum.

Thelwall, Robin, and Thilo C. Schadeberg (1993). The linguistic settlement of the Nuba Mountains. *Sprache und Geschichte in Afrika* 5:219-232.

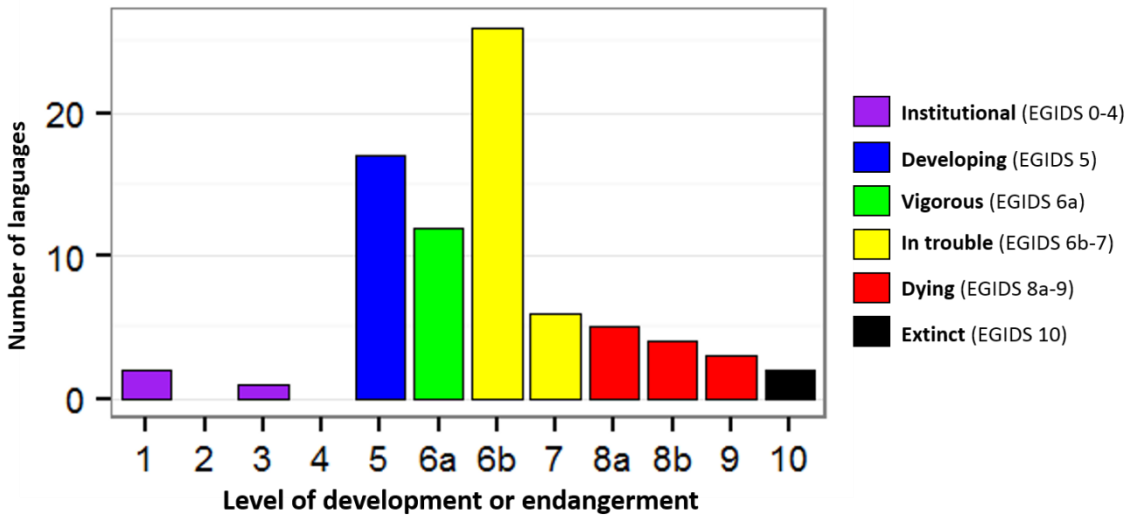
Tucker, A. N., & M. A. Bryan. (1966). *Linguistic analyses: The non-Bantu languages of north-eastern Africa.* London: Oxford University Press for the International African Institute.

Williamson, Kay (1989). Niger-Congo overview. In John Bendor-Samuel (ed.) *The Niger-Congo Languages*. Lanham, USA: University Press of America.

Appendices



Appendix 1. *Classification of African languages (Olson 2004:3).*



Appendix 2. *Sudan Language Status Profile (Lewis et al. 2015).*