Winfred Mkochi

Strong Accent Constituents in CiTonga (Bantu)

Universal Guidelines and Constraints

Thesis for the degree of Philosophiae Doctor

Trondheim, May 2014

Norwegian University of Science and Technology
Faculty of Humanities
Department of Language and Literature

NTNU – Trondheim
Norwegian University of
Science and Technology

Dedication

To Tamanda.

Acknowledgements

Abstract

This dissertation examines word-prosody of ciTonga, a Malawian Bantu language spoken by lakeshore people of northern Malawi. It is argued that the real word-prosody in this language (and perhaps many Bantu languages) revolves around the idea of Strong Accent Constituency, power relations between segments, syllables and between lower and higher prosodic categories as determined by Universal Guidelines such as SONORITY, FINALITY, EDGENESS and PROSODIC HIERARCHY as well as constraints which favour language- or context-specific Strong Accent Constituents (PROSODIC STEM, ACCENT FOOT, STEM-σ_1, PENULTIMATE-σ or FINAL-σ). Tone and prosodic morphemes such as Minimal Prosodic Words and Reduplicative Prosodic Morphemes also seem to be heavily regulated by Strong Accent Constituency. The empirical bases are three speech styles found in ciTonga (Nkhata-Bay Variety) namely, formal, common and elderly speech styles. It is one of several understudied and endangered languages in Malawi. This study therefore is in line with one of the goals of the University of Malawi's Centre for Language Studies, where this candidate serves as a member, which is to prioritize research activities on such languages. The candidate is a native speaker of ciTonga and, as such, he is *primary source* of most of the data. Other methods such as elicitation and focus group discussions were conducted with informants not only to get to the bottom of the matter, but to also understand social issues underlying language variation.

The dissertation has been presented in five chapters. Chapter 1 presents introductory remarks. These include the problem statement, a note on methodology, summary of findings, theoretical precedents, and, finally, organization of the dissertation. Chapter 2 presents basic facts about the language under study. These include language classification, previous works on ciTonga, speech sounds, the syllable, tone, as well as nominal and verbal morphology.

Chapter 3 presents a proposal for the theory of Strong Accent Constituency. It presents the data on vowel and consonant deletion which motivates this theory analysis. Then attempts are made to account for the facts in terms of stress-accent theory and Downing's (2006) Morpheme-Based Templates Theory both of which are found to be slightly problematic to account for segment deletion and preservation patterns exhibited in ciTonga. Finally, the chapter introduces the theory and accounts for the facts in terms of Strong Accent Constituent Theory.

Chapter 4 presents formally the theory of Interaction between Tone and Strong Accent Constituents. The chapter presents the data on tone assignment in basic verbs, simple past

tense verbs and present progressive aspect verbs which motivate this type of theory analysis. Attempts are then made to account for the facts in terms of tone alignment theory (as argued for by Mtenje 2006), autosegmental accent (Clements and Goldsmith 1984) and 'pitch-accent' or accentual properties of tone in Bantu languages (as hinted upon by Downing 2004). All these theoretical perspectives are found to be slightly inadequate to account for tone distribution patterns in ciTonga. On the other hand, a theory based on Interaction between Tone and Strong Accent Constituents is shown to account for the facts slightly better.

Chapter 5 presents formally proposals for Strong Accent Constituent-Based Templates as a theory of morphology-prosody interfaces in ciTonga and perhaps many other Bantu languages with a Strong Accent Constituent system. It presents the data on general phonological words, Minimal Prosodic Words and reduplicative Prosodic Stems. It then reviews two competing theories in literature within the Generalized Templates Theory namely, the Prosodic Hierarchy-Based Templates Theory and the Morpheme-Based Templates Theory both of which have a goal to account for morphology-prosody interfaces. Both these theories are shown to be slightly inadequate to account for parameters exhibited by phonological words in ciTonga. On the other hand, it is suggested that a theory of Strong Accent Constituent-Based Templates may account for the facts slightly better.

Chapter 6 summarizes and concludes the dissertation.

Table of contents

List of Universal Guidelines

EDGENESS

Edgemost syllables are stronger than non-edge syllables.

FINALITY

Non-final syllables are stronger than final syllables

LEFTEDGE

The left edge is stronger than the right edge

PITCH

Units with high pitch are stronger than units with low pitch.

PROSODIC HIERARCHY

Lower level prosodic constituents are stronger than higher level prosodic constituents.

RIGHTEDGE

The right edge is stronger than the left edge

SONORITY

Least sonorous consonants and most sonorous vowels are stronger than most sonorous consonants and least sonorous vowels.

STRESS

Stressed syllables are stronger than unstressed syllables.

SYLLABLE WEIGHT

Heavy syllables are stronger than light syllables.

VOWEL LENGTH

Long vowels are stronger than short vowels.

List of Constraints

ACFT/SAC

Accent Feet belong to Strong Accent Constituents

ALIGN – L (MStem, H)

(Simple past tense)

The left edge of the morphological stem in simple past tense verbs is aligned with a high tone (Mtenje 2006).

ALIGN (PRSTEM, L, FT, L)

Every Prosodic Stem begins with a foot (cf. Kager 2007:211).

ALIGN (PRSTEM, R, FT, R)

Every Prosodic Stem ends with a foot (cf. Kager 2007:211).

ALIGN-L (Mword, H)

(Present Progressive Aspect)

The leftmost edge of a Morphological Word in the Present Progressive Aspect is aligned with a high tone (Mtenje 2006).

ASS (TONE)

Tones are associated with tone bearing units (cf. Yip 2000:80).

BINARITY

A prosodic constituent contains minimally and maximally two of the units dominated by the constituent (i.e. Prosodic Word contains minimally and maximally two feet; Foot contains minimally and maximally two syllables or moras; syllable contains minimally and maximally two moras) (McCarthy and Prince 1993, Prince and Smolensky 2004, Orie 1997; cited in Downing 2006b:9).

OR

BINARITY

"Each daughter of a constituent must be adjacent to some edge of the constituent" (Downing 2006b:125).

*COMPLEX ONSET

No complex onsets (cf. Kager 1999:97)

FAITH-SAC

Strong Accent Constituents are preserved.

FTBIN

Feet are binary (at the level of the mora or the syllable) (Prince and Smolensky 2004:56).

HEADEDNESS

Any prosodic category C^i must dominate a C^{i-1} (e.g. Prosodic Word must dominate a foot) (Itô and Mester 1992, Orie 1997, Selkirk 1995; cited in Downing 2006b:37).

IDENT-HEAD PROSODIC PHRASE (IDENT-HPP)

Identity of the Head Syllable of a Prosodic Phrase (Stressed syllable) is preserved.

MAX-BR

"All the segments of the Base are contained in the RED" (McCarthy and Prince 1993, cited in Downing 2006b:13).

MAX-C

Input consonants are preserved in the output (cf. Kager 1999:181).

MAX-FOOT

Segments of the strong position Foot are preserved (ad hoc).

MAX-HV

Input high vowels are preserved in the output (ad hoc).

MAX-IO

Input segments are preserved in the output (cf. Kager 1999:67).

MAX-LIQ

Input liquid consonants are preserved in the output (ad hoc).

MAX-LOWV

Input low vowels are preserved in the output (ad hoc).

MAX-MIDV

Input mid vowels are preserved in the output (ad hoc).

MAX-NAS

Input nasal consonants are preserved in the output (ad hoc).

MAX-OBS

Input obstruent consonants are preserved in the output (ad hoc).

MAX-PROSODIC STEM

Segments of the strong position Prosodic Stem are preserved (ad hoc).

MINWRD≡MINSAC

A Minimal Prosodic Word is co-extensive with a Minimal Strong Accent Constituent.

No l (fin. μ)

No liquid before final mora (Mkochi 2007/08)

NOCODA

Syllables have no codas (cf. Kager 1999:94).

ONSET

Syllables have onset (Kager 1999:93).

PARSE-σ

Every syllable is contained inside a foot (Prince and Smolensky 1993).

PROSODICSTEM

Prosodic Stems are minimally binary (at the level of the syllable or the mora) (Downing 2006b).

REALIZEMORPHEME

"Every input morpheme must have [the appropriate] output realization" (Downing 2006b:132, adapted, Akinlabi 1996, Walker 2000).

RED≡MINSAC

The reduplicative prosodic morpheme is co-extensive with a Minimal Strong Accent Constituent.

SPREAD

Tones spread (cf. Mtenje 2006).

STRESS-TO-WEIGHT

If stressed, then heavy (Kager 1999:268).

TONE/SAC

Tone belongs to Strong Accent Constituents.

List of Abbreviations

AC	Accent Constituent
AP	Autosegmental Phonology
appl	Applicative
C	Consonant
caus	Causative
CV	Consonant and vowel sequence
dat	Dative
DFT	Distant Future Tense
DG	Determiner Guideline
dist fut	Distant Future Tense
dist past	Distant Past Tense
fv	Final vowel
GTT	Generalized Templates Theory
hab	Habitual
HPP	Head Prosodic Phrase
HV	High vowel
int	Intensive
IPA	International Phonetic Alphabet
LIQ	Liquid consonant
LOWV	Low vowel
LV	Liquid and vowel syllable

MBT	Morpheme-Based Templates Theory
MIDV	Mid vowel
n/a	Not available
NAS	Nasal consonant
near fut	Near Future Tense
NV	Nasal and vowel syllable
OBS	Obstruent consonant
OM	Object agreement marker
OT	Optimality Theory
OV	Obstruent and vowel syllable
pass	Passive
past	Simple Past Tense
PBT	Prosodic Hierarchy-Based Templates Theory
pen	Penultimate syllable
PL	Penultimate vowel lengthening
PREF	Prefix
Pre-nas	Prenasalized consonant
pres perf	Present Perfect Aspect
prog	Present Progressive Aspect
PRSTEM	Prosodic Stem
PRWRD	Prosodic Word
rec	Reciprocal
RED	Reduplicative morpheme

SAC	Strong Accent Constituent
SACT	Strong Accent Constituent Theory
SM	Subject agreement marker
SPE	The Sound Pattern of English
stat	Stative
subj	Subjunctive mood
TAM	Tense-Aspect-Mood
TBU	Tone Bearing Unit
TETU	The emergence of the unmarked
V	Vowel
VP	Verb phrase

List of Symbols

σ	Syllable
μ	Mora
*!	Fatal violation
☞	Winner candidate
*	Violation/ungrammatical
.	syllable boundary
-	Morpheme boundary
[	Onset of Morphological Stem
´	High Tone
'	(Primary) stress
^{h}	Aspiration
$^{j/w}$	slur/gliding/re-syllabified high vowel
< ' >	Syllable deleted
< >	Extrametrical
~	Nasalized vowel
⇒	becomes/becoming
σ_n	Final syllable
σ_1	Stem-initial syllable
( )	Accent Foot
{ }	Prosodic Stem
≡	Is co-extensive with

Chapter 1

Introduction

1.1   Introduction

The aim of this study is to investigate the word-prosody of ciTonga, a Malawian Bantu language spoken by the lakeshore people of northern Malawi. The main focus is on the verb word because "a major source of the peculiarly Bantu penchant for accentual reanalysis lies in the morphological make-up of the verb" (Clements and Goldsmith 1984:5). I will argue that the real word-prosody in this language (and perhaps many Bantu languages) revolves around the idea of Strong Accent Constituency, power relations between segments, syllables and between lower and higher prosodic categories as guided by Universal Guidelines such as SONORITY, FINALITY, EDGENESS and PROSODIC HIERARCHY as well as constraints which favour language- or context-specific Strong Accent Constituents (PROSODIC STEM, ACCENT FOOT, STEM-σ_1, PENULTIMATE-σ or FINAL-σ). Tone and prosodic morphemes such as Minimal Prosodic Words and Reduplicative Prosodic Morphemes also seem to be heavily regulated by Strong Accent Constituency. The empirical bases are three speech styles found in ciTonga namely, formal, common and elderly speech styles[1].

The following section defines the problem I am addressing: What is the real structure of the Prosodic Word in ciTonga and perhaps Bantu languages of its type? Section 1.3 presents a note on methodology. Section 1.4 presents the findings which point towards theories of Strong Accent Constituency, the Interaction between Tone and Strong Accent Constituents, and Accent Constituent-Based Templates as relevant vehicles to deal with word-prosody, the interaction of tone and Strong Accent Constituents, and the structure of Prosodic Words and Prosodic Stems, respectively. Section 1.5 reviews theoretical precedents. Finally, I present the organization of the dissertation and summary of this chapter in sections 1.6 and 1.7, respectively.

---

[1] There are no clear boundaries between one speech style and another. The classification into formal, common and elderly speech styles is therefore an adhoc one and it is intended purely for discussion purposes.

1.2   The problem

According to Downing (2004), many Bantu languages have been described as being tonal, that is, tone is part of the lexicon (and Proto-Bantu is reconstructed as having two tones, High and Low). However, almost all of them are also described as having accentual properties. Many of them are also reported to have stress-accent independent from tone. Findings in Downing's (2004) survey of Bantu languages confirm Hyman's (1977) earlier finding that (stem-) initial and penult are the most common positions, crosslinguistically, to be assigned main stress. She lists about 26 Bantu languages as having penultimate stress[2]. Only two languages in the list (western Lingala dialects and Luvale) have stem-initial stress. Luvale has stem-penult stress as well. The most commonly reported phonetic correlate of penultimate stress in both tonal and non-tonal languages is vowel lengthening[3].

The fact that stress, cued by penultimate vowel lengthening, falls on the penultimate syllable is further solidified by the fact that when the words are extended, such as when suffixes are added, penultimate lengthening shifts to new penultimate vowels, an indication that penultimate lengthening is something which is automatic. The foot type often cited is a syllabic trochee consisting of a heavy penultimate syllable and a light final syllable. These facts are exemplified by data from ciTonga given in (1) below[4].

---

[2] These languages are Bondei, Chichewa, Chimwi:ni, Chizigula, Digo, Giryama, Kinande, Kirufiji, Kishambaa, Véhiculaire Kituba, Eastern Lingala, Luvale, Makonde, Namwanga, Nguni [Zulu, Xhosa, Swati, Ndebele], Nyakyusa, Nyooro-Toro, Pogolo, Low Runyankore, Shona, Sotho Tswana, Swahili, Taita, Tsonga, Tumbuka, and Venda.

[3] The type of penultimate lengthening (PL) which I am concerned with in this case is what Hyman (2009:195, 196) describes as "typical" PL in many Bantu languages. In this type of PL, penultimate lengthening is believed to be an effect of stress and typical examples are ones "which unambiguously involve the addition of a mora" as opposed to mere "phonetic lengthening" (Hyman 2009:196).

[4] Although stress has not been formally presented in ciTonga (and many Bantu languages), its presence has often been recognized in various analyses. For instance, Mtenje (2006) writes: "In ciTonga, tone spreading appears to be constrained by two factors: Firstly, like in Chichewa, it does not apply when the vowel receiving the [spreading] tone is part of a **final foot**."

(1)   *Syllabic foot analysis of penultimate lengthening*

Verb   *English gloss*

be.('nee.k-a)   cover
cover-fv

be.ne.('k-aa.n-a)   cover each other
cover-rec-fv

le.('lee.s-a)   look
look-fv

le.le.('s-aa.n-a)   look at each other
look-rec-fv

('vii.n-a)   dance
dance-fv

vi.('n-ii.s-a)   cause to dance
dance-caus-fv

('bii.k-a)   cook
cook-fv

bi.k-i.('l-aa.n-a)   cook for each other
cook-appl-fv

βe.('lee.ŋg-a)   read
read-fv

βe.le.ŋg-e.('s-aa.n-a)   cause each other to read
read-caus-rec-fv

One crucial observation then would be that a disyllabic foot falls on the last two syllables of a Prosodic Stem/Word, with actual stress falling on the penultimate syllable. Another crucial observation is that the stressed syllables are heavy (i.e. bimoraic). These observations are familiar in many stress languages.

One piece of evidence often cited for stress in many languages is word minima (cf. McCarthy and Prince 1986). In cases where an input verb stem is monomoraic, some strategy has to be employed to satisfy the disyllabic foot minimality condition. For instance the ciTonga verb stems **-ba** 'steal', **-lja** 'eat', **-swa** 'break', **-fwa** 'die' and -**mwa** 'drink' are monomoraic. In citation form, an epenthetic vowel [i] is usually attached as a prefix[5]. This fact is illustrated in (2) below.

(2) *A Minimal word is disyllabic*

Input	Output	English gloss
/ba/	[ii.-ba]	steal!
/lja/	[ii.-lja]	eat!
/swa/	[ii.-swa]	break!
/fw-a/	[ii.-fwa]	die!
/mw-a/	[ii.-mwa]	drink!

This observation of a disyllabic minimal word is not unique because it has generally been observed since work by McCarthy and Prince (1986) that lexical words in unrelated languages spoken throughout the world are required to have a minimal size, typically two moras or two syllables. The analysis offered is that a minimal word matches with the size of a bimoraic foot or a disyllabic foot. Thus, the initial impression one gets in standard ciTonga, like in many Bantu languages, is that the size of a minimal word matches with a disyllabic foot as given in (3) below.

---

[5] There is evidence to show that these verbs are indeed monomoraic. For instance, they can take other types of prefixes as in **kuú-mwa** 'to drink' **kuú-fwa** 'to die' and **kuú-ba** 'to steal' in which case the inserted vowel [i] disappears (cf. Mtenje 2006).

(3)   *A minimal word is said to be a disyllabic Foot*

Input   *Output*   *English gloss*

/b-a/   [('ii.-ba)]   steal

/lj-a/   [('ii.-lja)]   eat

/sw-a/   [('ii.-swa)]   break

/fw-a/   [('ii-.fwa)]   die

/mw-a/   ('ii.-mwa)   drink

A word-stress analysis offered above, however, is problematic in a number of ways. To begin with, although uneven trochees (Heavy+Light) are reported in the literature (e.g. Jacobs 1990, 2000; Rice 1992; van der Hulst and Klamer 1996; and Mellander 2001, 2004), they are not supported by others (see Hayes 1995). Secondly, and most crucially, the use of the disyllabic foot analysis to account for minimal words such as the ones we have presented above runs into conflict with minimal words in most elderly speech styles of the same language which are simply *bimoraic* as illustrated in (4) below.

(4) *A minimal word is bimoraic in elderly speech styles*

Word   *English gloss*

baa   steal!

 ljaa   eat!

swaa   break!

fwaa   die!

 mwaa   drink!

A third and most serious problem for the stress analysis suggested above is that the domain for penultimate vowel lengthening (or stress) in ciTonga is not the Prosodic Word. The idea that the domain for penultimate lengthening in these languages is not a word is thoroughly investigated by Hyman (2009) and many others. According to Hyman (2009:198), for instance, two domains of penultimate lengthening (or penultimate stress) are ascertained in

Bantu languages which have the PL (penultimate vowel lengthening) phenomenon. PL may be utterance-penult (e.g. Sotho: Doke 1967:125) or phrase-penult (e.g. Chichewa: Kanerva 1990, Tumbuka: Downing 2006a, Makonde: Kraal 2005, and Matengo: Yoneda 2005). As the forms in (5) below indicate, the domain for PL in ciTonga seems to be the same as in neighbouring languages Chichewa and Tumbuka. Assuming that vowel lengthening is indeed a cue for stress, the domain of stress in ciTonga must be the Phonological Phrase. This is because lengthening usually shifts to the penultimate syllable of the rightmost Word in a Phonological Phrase.

(5)   *PL in ciTonga is phrase-penult*

Phrase   English gloss

$_{VP}$[le.lee.s-a]   look

$_{VP}$[le.le.s-a   vi-.βaa-.nt^{h}o]   look at bad people

$_{VP}$[zu.mbuu.w-a]   reveal

$_{VP}$[zu.mbu.w-a   ma-.yee.so]   reveal an exam

$_{VP}$[sa.mbii.z-a]   teach

$_{VP}$[sa.mbi.z-a   mu-.nt^{h}i.kaa.ze]   teach a woman

$_{VP}$[βe.lee.ŋga]   count

$_{VP}$[βe.le.ŋga   a-.nt^{h}u.luu.me]   count men

Thus, as Hyman (2009) observes, the attraction of length is across words and "we can hypothesize that penultimate prosody starts out as intonational and undergoes boundary narrowing." The problem question this study attempts to address then is: **What is the real prosodic structure of a Phonological Word in Bantu languages of ciTonga type?**

This problem is exacerbated by several other phonological phenomena such as vowel and consonant deletion, high tone assignment, word minima, size of the reduplicative prosodic morpheme, and blocking of general morpheme deletion processes within the Prosodic Word/Stem domain. The problem is to understand general principles which are responsible for these processes whose solutions, it is believed, can enhance our understanding of the

original problem: the real prosody of the phonological word in this language. These sub-problems are themselves complex in nature as outlined below.

1) Vowel deletion

In the formal speech style, all vowels of the Morphological Stem are preserved (e.g. **ti-.[to.nde.k-e.s-a.n-eé.ng-e** 'we should cause each other to fail'). In the common speech style, non-low vowels of the final syllable are usually deleted (e.g. formal **ti-.[to.nde.k-e.s-a.n-eé.ng-e ⟹ ti-.[to.nde.k-e.s-a.n-eé.ng**). In common elderly speech style, however, the only vowels that are preserved are those which are nuclei of the first two syllables and (bimoraic) penult syllables. Deletion of low vowels is not acceptable (e.g. formal **ti-.[to.nde.k-e.s-a.n-eé.ng-e ⟹ ti-.[to.nde.k-.s-a.n-eé.ng**). In rare elderly speech styles, however, pre-final low vowels may be deleted as well while the final low vowel is usually preserved (e.g. formal **to.nde.k-e.s-a.n-aá.ng-a ⟹ to.nde.k.-s-.n-aáng-a** 'be causing each other to fail').

2) Consonant deletion

All consonants of the Morphological Stem are preserved in the more formal speech style (e.g. **bi.k-ii.l-.a** 'cook for'). However, in the common speech style, liquid consonant onsets to final syllables are usually deleted (e.g. formal **[bi.k-ii.l-a ⟹[bi.k-ii-.ʲa**). Nasal and obstruent consonants are never deleted in the same position. In common elderly speech styles, the only liquid consonants that are preserved are those which are onsets of the first two syllables and the penultimate syllable of the Morphological Stem. Liquid consonants in all other positions are usually elided (e.g. formal **[tʰa.mba.l-i.l-aan-a. ⟹ [tʰa.mba.<'>.l-aa.na** 'stretch legs over each other'). Nasal and obstruent consonants are also never deleted anywhere. In addition to multifarious vowel and liquid consonant deletion, some elderly speakers in rare speech styles may also delete nasal consonant onsets (e.g. **ka.li.p-i.l-aa.n-a ⟹ ka.li.p-.l-ãã** 'reprimand each other with'). Obstruent consonants are usually preserved in same positions.

3) High tone assignment

An adequate characterization of ciTonga grammar should be able to account for the following high tone assignment facts. Tone in high-toned basic verbs (without prefixes) is restricted to the last two moras of a Prosodic Stem. It is usually assigned to the penultimate syllable/mora of the Morphological Stem (e.g. **kʰu.mbií.l-a** 'admire') when the Stem is in phrase-final

7

position. When the high-toned verb appears in a phrase-medial position the high tone falls on the final syllable/mora (e.g. **kʰu.mbi.l-á mu.ntʰu.luu.me** 'admire a man').

In the simple past tense, the high tone is restricted to the initial two syllables/moras of the Morphological Stem. It is assigned to the initial syllable when the Stem has two or three syllables (e.g. **ndi-.ŋgu-.[lé.lee.s-a** 'I saw'). However, the high tone is realized on the second syllable of Stems with four or more syllables (e.g. **ti-.ŋgu-.[tʰa.mbá.lí.l-aa.n-a** 'we stretched legs over each other'). The tone then spreads to following moras or syllables (if there are any), but it does not spread to the last two syllables.

Finally, *normal* verbs (where morphological stems have two or more syllables) in the present progressive aspect verbs have no morphological marker (e.g. **ndí-.[sá.mbii.z-a** 'I am teaching'). A high tone seems to spread from word-initial syllable to the antepenult syllable. The spreading tone does not penetrate the last two syllables. When a morphological stem involved is *monomoraic*, however, there appears a morpheme **-tu-** before the stem (e.g. **ndi-.túu-.[vwa** 'I am hearing'). This entity assigns the high tone to itself.

4) Size of Minimal Prosodic Words/Stems

A Prosodic Word in ciTonga can roughly be a Morphological Stem (e.g. **[βe.lee.ŋg-a** 'read'). It can also roughly be a Morphological Word consisting of a prefix string and the Morphological Stem (e.g. **ndi.ŋgu-mu-[βé.lee.ng-a** 'I read him'). In cases where the Morphological Stem is monomoraic (e.g. **[ba** 'steal'), there are two ways of achieving the Minimal Prosodic Word. In the formal and common speech styles an epenthetic vowel [i] is added as a prefix to the monomoraic verb stem (e.g. **ii-[ba**) in its citation form. However, in elderly speech styles the stem vowel is lengthened (e.g. **[b-aa**).

5) Reduplicative Prosodic Morphemes

Reduplicative facts also present another problem for the real prosody of the ciTonga Prosodic Word or Stem. In formal and common speech styles reduplication involves repetition of entire Morphological Stem (e.g. **ndi-.ŋgu-.βe.lé.βé.tá-[βe.le.βee.t-a**). In the elderly speech styles, however, the reduplicative prosodic morpheme may copy only two initial syllables of the Morphological Stem (e.g. **ndi-.ŋgu-.[βé.lé.-βe.le.βee.t-a**). The reduplicative prosodic morpheme also adheres to the disyllabic or bimoraic minimality condition. In the formal and common speech styles, the reduplicative prosodic morpheme maintains the epenthetic vowel of the Base when the Morphological Stem involved is monomoraic (e.g. **i.ba-[ii.ba** 'steal a

lot'). In elderly speech styles, however, the reduplicative prosodic morpheme maintains the bimoraic shape of the Base (e.g. **baa-[baa**).

6) Blocking of general morpheme deletion processes

An adequate characterization of ciTonga Prosodic Words should also be able to account for processes which block general morpheme deletion processes. For instance, a 'normal' Present Progressive Aspect verb (with two or more syllables) may be expressed without a morphological marker (e.g. **ndí-.[lé.lee.s-a** 'I am looking'). However, when the Morphological Stem involved is monomoraic, the grammar of ciTonga allows the morpheme **-tu-**[6] (most likely the morphological marker of the present progressive aspect) to appear before it (e.g. **ndi-.túu-.[lj-a** 'I am eating'). What is further puzzling is the fact that when an object marker is introduced in a normal present progressive aspect verb complex, **-tu-** surfaces again, but now before the (underlined) object marker (e.g. **ti-tú-ví-[βeléβeet-a** 'we are talking about them').

Another blocking of a general morpheme deletion process involves the infinitive marker **-ku-** when it follows the Distant Future Tense marker **-zamu-** and it precedes a monomoraic verb stem or a (monomoraic) object marker. Normally, the distant future tense marker **-zamu-** appears before the morphological stem and there is no infinitive marker in the verb morphology (e.g **ndi-.za.mu-.[lé.lee.s-a** 'I will look'). However, when the verb stem involved is monomoraic (e.g. **-ba** 'steal', **-lja** 'eat', etc), the deletion of **-ku-** is blocked (e.g. **ndi-.za.mu.-kuú-.[b-a** 'I will steal', vs. *ndi-.za.mu.-[b-a**). Similarly, infinitive marker **-ku-** does not appear before the object marker when the tense marker is that of the simple past tense **-ŋgu-** (e.g. **ndi-.ŋgu-.mu-.[lé.lee.s-a** 'I looked at him'). However, when the tense marker is that of the distant future tense **-zamu-**, **-ku-** appears before an object marker as well (e.g. **ndi-.za.mu-.ku-.mu-.[lé.lee.s-a** 'I will look at him'). What is it that is in **-zamu-**, and not past tense marker **-ŋgu-**, that warrants retention of **-ku-** before monomoraic verb stems and (monomoraic) object markers?

---

[6] Mtenje (2006, 1994/95) uses **-ti-** or **-tu-** before object markers and monomoraic verbs, respectively. But to my knowledge, the two are used interchangeably. I will use **-tu-** because it is more common than **-ti-**.

1.3   A note on methodology

Accent (or prominence) is not a physical phenomenon. Thus, as Hayes (1995:5) observes for stress-accent, any theory of accent is in an indirect relation with the facts that support it. Thus, the study of accent is not like the study of other branches of phonology where it is easy to establish when the observed facts confirm or falsify a hypothesis. The study of accent in Bantu languages would be better handled if there were a clear and unambiguous phonetic correlate of it. Unfortunately, this is not the case. Just like with the study of stress accent, however, the fact that accent does not have phonetic correlates does not mean that observations about it cannot be given a solid empirical base (cf. Hayes 1995:9-23). It is possible to study the accent system by carefully examining and comparing various phonological diagnostics for it. For ciTonga, I have relied on vowel and consonant deletion and preservation, high tone assignment, word minima, reduplication, and blocking of general morpheme deletion processes.

In view of the confusion which abounds in literature on prosody in African languages, I adopted for this study recent Grounded and Emergent Approaches in order to confront the problem of bias. Thus, I relied on prior knowledge from existing literature throughout the research structure while preserving the spirit that theory emerges from the data rather than using data to test theory (Glaser 1992, cited in Jaccard and Jacoby 2010). As a native speaker of ciTonga, *I am primary source of most of the data*. I tapped on my fellow native speakers' intuitions where my own judgments of grammaticality were fuzzy. These are traditional methods of data collection in linguistics (Chomsky 1957, 1964, 1965, 1968). Elicitation and focus group discussions were also used in order to obtain information which was either phonological or sociolinguistic in nature. Computer software called *Praat* developed by (Boersma and Weenink 2012) was used to verify some of the pitch contours. Where secondary data sources have been used, acknowledgement is given accordingly.

Having had a clearly defined set of data I was looking for, I wrote notes about ideas and insights which I got at particular moments. I normally consulted these ideas when analyzing my data. Although theoretical sampling is used largely in reference to objects of study other than lexical items, I used the idea to keep collecting words of varying lengths (in terms of syllable and mora count) and this diversity provided new information I never anticipated when I was starting this project. Through the continuous play of data collection and analysis I always noticed the kind of data I needed in order to proceed. As is the case in most grounded and emergent approaches, I combined insights gained during data collection, from reading

past literatures, field notes, and the information contained in the data to modify existing theories of prosody and language to suit the more complex data which have resulted in Guidelines and Constraints for Strong Accent Constituents Theory. The approach, therefore, was bottom-up in the sense that I used generalizations derived from a careful review of the data to derive more general themes.

1.4   Findings

The thesis I make is that the real prosody of a Bantu Prosodic Word is Strong Accent Constituency, power relations between segments, syllables, and between lower and upper prosodic categories as guided by Universal Principles such as SONORITY, EDGENESS and PROSODIC HIERARCHY as well as constraints which favour language- or context-specific Strong Accent Constituents. The Universal Guideline of SONORITY provides that least sonorous consonants and more sonorous vowels are stronger than more sonorous consonants and least sonorous vowels. EDGENESS provides that edge syllables are stronger than non-edge syllables. Finally, the Universal Guideline of PROSODIC HIERARCHY provides that lower prosodic constituents are stronger than higher prosodic constituents. The cut-off point of what Strong Accent Constituent is is left to language users themselves to decide although they all have access to the same Guidelines and Constraints. For instance, the Prosodic Stem and the low vowel are perceived as Minimal Strong Accent Constituents in formal and common speech styles of ciTonga while most elderly speech styles perceive the Accent Foot to be the Minimal Strong Accent Constituent, and the low vowel is not perceived as a Strong Accent Constituent by others.

Evidence for the phonology of Strong Accent Constituency in ciTonga has been drawn from vowel deletion, consonant deletion, high tone assignment, blocking of general morpheme deletion processes, word minima and reduplication. The results are three theoretical perspectives namely, Strong Accent Constituents Theory, the Interaction of Tone and Strong Accent Constituents, and Strong Accent Constituent-Based Templates Theory. Strong Accent Constituents Theory deals with power relations between segments, syllables and prosodic constituents within a Prosodic Word/Stem. The other two theories deal with the interaction between Strong Accent Constituents and tone, and morphology-prosody interfaces, respectively. The ultimate result is the theory of Universal Guidelines and Constraints for Strong Accent Constituents.

1.4.1 Strong Accent Constituents Theory

This theory is motivated by patterns of vowel and consonant deletion in formal, common and elderly speech styles of ciTionga. The observations point to the fact that Universal Principles such as Sonority, Edegeness, and Prosodic Hierarchy play a crucial role in the grammars of natural language. For instance, it is clear that there is a strong relationship between segment faithfulness, on the one hand, and segment sonority, edgeness and a level in which it appears in the Prosodic Hierarchy. I will argue that every prosodic unit (Phonological Phrase, Prosodic Word, Prosodic Stem, Foot, syllable, mora, segment, feature and any smallest atom of human speech) is an **Accent Constituent** (unit of prominence) and that what is preserved, especially in relation to Accent, is a **Strong Accent Constituent** of the relevant domain (Language, Utterance, Phrase, Word, Stem, or Accent Foot). What is Stronger Accent Constituent in one speech community or context may not necessarily be stronger in another context. At the heart of Strong Accent Constituent Theory, therefore, is the fact that a finite set of Universal Guidelines **EDGENESS**, **FINALITY**, **SONORITY**, **PROSODIC HIERARCHY**, **STRESS**, **SYLLABLE WEIGHT**, **PITCH**, and several others, play a crucial role in determining what is strong and not strong. What counts as Determiner Guidelines (DG) of strength in one language may not necessarily be so in another language. The Guidelines may jointly make a determination, relegating some Guideline in the process. Since language- or context-specific Strong Accent Constituents are the ones which are usually preserved, it appears that there is only one **FAITHFULNESS** constraint which is responsible for this status: **FAITH - STRONG ACCENT CONSTITUENT** (**FAITH-SAC**). What happens to weaker ACs, it seems, is of little or none of the accent grammar's business. Another constraint **ACFT/SAC**[7] ensures that Accent Feet are assigned to Strong Accent Constituents such as PROSODIC STEM, STEM-σ_1, PENULT-σ or FINAL-σ as determined and provided by the Universal Guidelines of PROSODIC HIERARCHY, EDGENESS, FINALITY and SONORITY. ACCENT FEET are themselves Strong Accent Constituents too.

1.4.2 The Interaction between Tone and Strong Accent Constituents

This theory is motivated by patterns of tone assignment in different types of verb complexes. The analysis suggests that the high tone belongs to Strong Accent Constituents. This generalization will be accounted for in terms of the constraint TONE/SAC (see footnote 7 in the preceding section) which requires high tones to be assigned to Strong ACs such as ACCENT FEET and PROSODIC STEMS as guided by the Universal Guideline of PROSODIC

---

[7] The constraint ACFT/SAC would be compared with Smith's (2000) augmentation constraints since tthe prominence of Strong Accent Constituents get enhanced by having Accent Feet and tone fall on them.

HIERARCHY. To account for tone-spreading, I have followed Mtenje's (2006) analysis where the constraint SPREAD requires high tones to spread to the next TBUs. Blocking of tone spreading by the last two syllables of the Prosodic Stem/Word/Phrase has been attributed to another constraint IDENT-HPP which requires identity of Head Prosodic Phrases (penultimate syllables) to be preserved. Deletion of the present progressive aspect marker **-tu-** before normal verb Stems (with two or more syllables) has been attributed to the constraint HEADEDNESS outranking FAITH-SAC. As a lexically-specified SAC (foot head), **-tu-** was supposed to be properly dominated by Prosodic Stem. Its deletion therefore satisfies the principle of HEADEDNESS better. Once deleted, the tone which was meant for the foothead **-tu-** is re-assigned according to rules of the Association Convention, requiring tones to be assigned from left to right in a one-to-one fashion. The fact that **-tu-** surfaces before monomoraic verbs would be indication that it is now parsed by Prosodic Stem and it no longer violates HEADEDNESS.

1.4.3 Strong Accent Constituent-Based Templates Theory

This theory has been motivated by the blocking of general morpheme deletion processes, the phonological patterns of general words, word minima and reduplicative prosodic morphemes. The fact that many Prosodic Words have one Prosodic Stem is accounted for in terms of non-violation of the principle of HEADEDNESS, requiring a prosodic constituent to contain at least one of the units it dominates. Prosodic Words must also satisfy MAXIMALITY CONDITION to ensure that they are minimally as large as possible, i.e. they must contain a minimum and maximum of two Prosodic Stems. Since Prosodic Words are first and foremost morphological categories, the idea of maximality is not always satisfied. The principle of BINARITY is thus largely violated because many Prosodic Words contain just one Prosodic Stem or more than two Prosodic Stems. This is not surprising because, as it has been argued repeatedly by others (e.g. Downing 2006b, and references cited therin), Prosodic Words are roughly Morphological Words and anything goes in Morphological categories.

The Prosodic Stem, however, is identified as the domain for accent and tone. By HEADEDNESS, it must contain at least one Accent Foot if it is to be properly parsed. Prosodic Stems must also satisfy the MAXIMALITY CONDITION to ensure that they are minimally as large as possible. By BINARITY, Prosodic Stems are expected to contain a maximum of two Accent Feet. It has been proved that the first two and last two moras or syllables of the Prosodic Stem are earmarked for Accent Foot. This is easy to tell when the corresponding Morphological Stem has four or more syllables. For the Prosodic Stem there is an additional

condition: It must contain minimally two syllables as required by Downing's (2006b) constraint of PROSODICSTEM. Blocking of general morpheme deletion processes before monomoraic stems provides evidence for such a demand.

The analysis of Minimal Prosodic Words in formal and in common speech styles is that it must be a Prosodic Stem while in the elderly speech styles it must be an Accent Foot. This is in tandem with the fact that elderly speakers take an Accent Foot as their Minimal Strong Accent Constituent while the Minimal Strong Accent Constituent in the formal and common speech styles is the Prosodic Stem. The prevailing Guideline is the PROSODIC HIERARCHY which provides that lower prosodic constituents such as the Prosodic Stem (in the formal and common speech styles) and the Accent Foot (in the elderly speech styles) are stronger than higher prosodic constituents. The generalization requiring explanation then is that Minimal Prosodic Words are co-extensive with a Minimal Strong Accent Constituent. The constraint I have suggested therefore is **MINWRD≡MINSAC**, requiring Minimal Prosodic Words to match with the size of a system's Minimal Strong Accent Constituent. Since Prosodic Stems are required to be minimally disyllabic (by PROSODICSTEM), the Minimal Prosodic Word in the formal and common speech styles is disyllabic. It is bimoraic in the elderly speech styles because the foot in ciTonga is essentially bimoraic (as required by FTBIN).

Similarly, the disyllabic reduplicative prosodic morpheme in the formal and common speech styles is co-extensive with the prosodic constituent Prosodic Stem which hitherto has proved to be a Minimal Strong Accent Constituent in these speech styles. In the elderly speech styles, however, reduplicative forms which have the size of a bimoraic Accent Foot are optimal. The choice of Accent Foot as the size of RED seems to be based on the fact that ACCENT FOOT, and not PROSODIC STEM, is the Minimal Strong Accent Constituent in these speech styles. The guiding principle in both cases is once again the Prosodic Hierarchy and the suggested constraint responsible for their shapes is **RED≡MINSAC**, requiring RED to be co-extensive with a Minimal Strong Accent Constituent of a system.

1.5 Theoretical Perspectives

This project is carried out under the school of Generative Linguistics initiated by Noam Chomsky and Morris Halle in the 1950s. A most popular theoretical framework within this school at the moment, especially in the branch of phonology, is Optimality Theory developed by Prince and Smolensky (1991, 1993) and extended by others. This theory is a culmination of a long project in generative phonology which has evolved from the linear mode articulated

in Chomsky and Halle's (1968) *SPE* (*The Sound Pattern of English)*. I outline briefly some of the landmarks in this project.

The **SPE** argued that phonological representations of utterances are organized in linear strings of segments each of which comprised unordered bundles of distinctive features describing the articulatory and acoustic properties of that segment. Chomsky and Halle (1968) had clarified the notions of underlying versus surface representations. They made a proposal of a series of extrinsically ordered rules which derived surface representations from abstract or remote representations. The units at either of these levels of representation constituted a linear sequence (of segments and boundaries), and phonological rules which were sensitive to context mapped the representations at the underlying level to those of the phonetic level in a sequential fashion (i.e. one applying to the output of the other and in a given fixed order). This mode of phonological representation necessitated it to be called Linear Phonology. The success of Linear Generative Phonology is measured in three ways: firstly, it showed that phonological alternations can be a result of some rule; hence the alternation is predictable or rule-governed. Secondly, it succeeded in reducing the redundancy of phonological representations. And finally, the rule formalism led to simplification of the grammar which made a robust account of language acquisition (Bensoukas 2004:233ff).

However, in 1976 Goldsmith demonstrated that the *SPE* mode or the linear mode of phonological representation was in principle not able to handle tonal phenomena commonly encountered in African tone languages. His proposal was that tone be separated from the segments that ultimately bear them and be represented on separate levels referred to as **Tiers**. Each tier is a linear sequence of phonological units or features which can be affected independently by rules applying specifically to that level. In this sense then, levels are related to each other but independent of each other. Segments on different tiers are linked to each other by association lines showing how they are to be co-articulated. This non-linear mode is usually referred to as **Autosegmental Phonology** (AP).

Autosegmental Phonology also demonstrated that independent universal principles and language-specific rules combine in order to determine how melody units are associated (or co-articulated) with slots on the skeletal tier. The linkage of the levels was governed by a more general principle, the Universal Association Convention which stated that "when un-associated vowels and tones appear on the same side of an association, they will be automatically associated in a one-to-one fashion, radiating outward from the association line"

(Goldsmith 1990:14). Autosegmental Phonology also postulated Well-formedness Conditions which guaranteed that such linkages did not cross association lines. One of the crucial insights of AP was the assumption that the tiers were to be treated as autonomous.

Another non-linear theory that was developed around the same time as Autosegmental Phonology was **Metrical Stress Theory**. This theory was a reaction to the inadequacy of the Linear Generative Phonology (SPE) account of stress. It was originally proposed by Liberman (1975) but more fully elaborated and refined by Liberman and Prince (1977), Hayes (1980, 1995), and Halle and Vergnaud (1987). The central claim in metrical theory is that stress is "the linguistic manifestation of rhythmic structure, and that the special phonological properties of stress can be explicated on this basis" (Hayes 1995:1). The fundamental insight of metrical theory, then, is that stress is best characterized not as a feature or property of a segment but as a relation of prosodic prominence between sister elements (such as syllables or moras) in a given domain.

Optimality Theory represents a third phase in phonological thinking usually called the phase of Constraints or Principles (Prince and Smolensky 1993; McCarthy and Prince 1993a, b). It embodies a conception of the notions of underlying and surface structures in derivational theories, but it discards a belief that a set of serially ordered rules apply one after another in order to derive a correct output. Classical OT recognizes only two stages namely, the input and the phonological output. A crucial conception about OT is that the relation between an input and an output is governed by the interaction of violable universal constraints on output well-formedness. Constraints interact at the output level, although some constraints (faithfulness constraints) can refer to the input level. The constraint set is universal. Variation between languages or dialects comes about because of differences in ranking of the constraints. Candidates which violate high-ranking constraints are disqualified and those which satisfy them are optimal (see Yip 2000, Bensoukas 2004 for slightly better summaries).

There are two types of constraints namely, **markedness** and **faithfulness** constraints. Markedness constraints evaluate the featural, segmental and prosodic well-formedness of output forms. Faithfulness constraints evaluate the extent to which input and output forms correspond. Thus, phonological processes such as insertion, deletion, featural change or change of association lines in autosegmental representations will be penalized by faithfulness constraints. Markedness and faithfulness constraints have conflicting requirements of well-

formedness. The former demands forms to surface in their unmarked structure while the latter "prohibits differences between input and output" structures (McCarthy 2008:13).

The interaction of constraints and how an optimal candidate is arrived at is displayed in a figure called a *tableau*. The constraints head the columns, with the highest ranked on the left. The candidates begin the rows, and the violations are shown by asterisks below the relevant constraints. An exclamation mark against a violation mark (!) indicates that this is a fatal violation which completely rules out the candidate. Shaded cells indicate that they do not matter because a decision has already been made by a high-ranking constraint. The symbol ☞ shows the winner and solid lines between constraints indicate crucial rankings while dashed lines indicate that the ranking is not (or not yet) crucial.

1.6   Organization of the dissertation

This dissertation is organized as follows-: The following chapter presents basic facts about ciTonga. Chapter 3 presents formally a proposal for a Strong Accent Constituents Model for word-prosody in ciTonga. Chapter 4 presents formally the theory of Interaction between Tone and Strong Accent Constituents. Chapter 5 deals with problems of prosodic morphology and it suggests that a Strong Accent Constituent-Based Templates Theory is slightly better placed to deal with the issues. Chapter 6 summarizes and concludes the dissertation.

1.7 Chapter summary

The aim of this chapter was to present an introduction to the study. In a nutshell, the chapter has presented the aim of the study, problem statement, a note on methodology, summary of findings, theoretical perspectives, and, finally, organization of the dissertation. In the following chapter, I present quickly basic facts about ciTonga.

Chapter 2

Basic Facts

2.1 Introduction

This chapter presents basic facts about Malawian Tonga (popularly known as ciTonga among its speakers). The issues discussed include language classification, previous works, speech sounds, the syllable, tone, as well as nominal and verbal morphology.

2.2 Language Classification

Guthrie (1948) classifies Malawian languages as Bantu languages belonging to zones M, N and P. CiTonga belongs to 'Zone N Group 10' together with neighbouring Chichewa and Tumbuka. Bryan (1959) puts ciTtonga and Tumbuka (Northern Malawi's lingua franca) in one group where Chichewa (or Nyanja, Malawi's lingua franca) is excluded. In colonial literature, the language is considered a dialect of Tumbuka. The language is mostly spoken in Nkhata-Bay, a district in Northern Malawi that covers 4089km^2, with a population of 213,779 people (according to 2008 population census). It is estimated that 1.7% of Malawi's 15,066,320 people speak the language in their homes. The language is also spoken in neighbouring parts of Mzimba, Rumphi, Likoma Island and Nkhota-kota Districts.

2.3 Previous Works

As Mkochi (2005) observes, very little has been published on ciTonga linguistics, let alone phonology. Perhaps the earliest work that can be termed a linguistic pursuit is that which was done by Turner (1952: Tumbuka-Tonga-English Dictionary). Serious linguistic studies published as journal articles include Mtenje (1994/95, 2006), Mkochi (2005, 2007/8, 2009) and copies of undergraduate and graduate dissertations found mostly in the departments of English and African Languages and Linguistics at Chancellor College, University of Malawi. Some of the works which are written in ciTonga are Chirwa (1932), *Mcapu wa Chitonga* (n.a) (1932), *Mazgu ghaku Chiuta*, *The Bible in Chitonga* (1986), and Mphande (2000). None of these works is linguistically important. Materials written in English about ciTonga and its people include MacAlpine (1905), Mary Tew (1950), Monica Wilson (1958), Jaap van Velsen (1959a,b, 1964), Banda (1985), Soko (1985), Mphande (1998), Msosa (1999) and Mphande (2000). Mphande (2000) and Vail (1972) show that the earliest books written in this language which were read in primary schools included *Mkwele, Chiswamsangu, Marko, Mcapu, Nthanu*, and others. But these are not available now.

2.4 Speech sounds

CiTonga has five phonemic vowels namely, low, central and unrounded vowel [a], mid, front and unrounded vowel [e], mid, back and rounded vowel [o], high, front and unrounded vowel [i], and high, back and rounded vowel [u]. Long vowels are encountered in limited and predictable positions. Length is not contrastive. Long vowels of Bantu languages are mostly represented by double letters, the method of indicating length advocated by the International Institute of African Languages and Cultures (e.g. **bii.k-a** 'cook'). The IPA uses (:) after the letter representing the long sound segment. In this dissertation, long vowels are represented by double letters.

The table in (1) represents consonant sounds found in this language.

(1) *ciTonga Consonants*

		BILABIAL	LABIO-DENTAL	ALVEOLAR	PALATAL	VELAR	GLOTAL	LAB. VELAR
STOP	Non-pre.nas	b p pʰ		d t tʰ		g k kʰ		
	Pre-nas	mb mpʰ		nd ntʰ		ŋg ŋkʰ		
AFFRIC	Non-pre-nas				ɟ c cʰ			
	Pre-nas				ɲɟ ɲcʰ			
NASAL		m		n	ɲ	ŋ		
FRIC		β	v f	z s		χ	h	
GLIDE					j			w
LAT				l				

2.5 The syllable

Like many Bantu languages, the structure of the syllable in ciTonga is essentially CV. Many of them are light. Heavy syllables are attested in the penultimate syllable position of many words. This length is generally believed to be a cue for stress in many Bantu languages. The forms in (2) below exemplify this fact. A period symbol '.' has been used to show syllable boundary.

(2)   *Heavy syllables are realized in word-penultimate position*

Chitonga   *English*

je.'ndee.s-a   drive

'lee.mb-a   write

βe.le.'βee.t-a   speak

to.'ndeé.k-a   fail

mwe.tu.'lii.j-a   smile

go.go.'tee.z-a   exagerate

2.6   The noun word

Like in many Bantu languages, all nouns in ciTonga belong to a noun class. The noun word may consist of a noun stem (noun word minus prefixes) and a noun prefix. The prefix expresses grammatically relevant information of noun class and number (plural or singular). It plays a crucial role in agreement between the nouns and other grammatical classes (such as adjectives, possessives, demonstratives and verbs) in a construction. According to my observation, there are about 18 noun classes in this language as given in the table in (3) below.

(3) *CiTonga noun classes*[8]

(SM = subject agreement marker, OM = object agreement marker, PREF = prefix)

CLASS	1	2	3	4	5	6	7	8	9	10	12	13	14	6	15	16	17	18
PREF	mu	βa	mu	mi	li	ma	ci	vi	*N	*N	ka	tu	u	ma	ku	pa	ku	mu
SM	wa	βa	u	i	li	ŋga	ci	vi	i	zi	ka	tu	u	ŋga	ku	pa	ku	mu
OM	mu	βa	u	i	li	ŋga	ci	vi	i	zi	ka	tu	u	ŋga	ku	pa	ku	mu

For example, the noun word **muu-nt^hu** [1-person] consists of the prefix **mu-** and the stem **-nt^hu** 'person'. The prefix shows that the noun is in singular (number) and belongs to Class 1.

---

[8] Nouns belonging to classes 9/10 do not have prefixes, and not all nouns have prefixes.

The prefix **βa-** in the noun **βaa-nt^{h}u** [2-person] shows that the noun word belongs to Class 2 and it is in singular form. A noun can have more than one prefix, in which case the initial prefix is relevant for classification and agreement purposes. For example, the prefix **ka-** in the noun word **ka-muu-nt^{h}u** [12-1-person] shows that the noun belongs to Class 12 of diminutive forms, and that the noun in diminutive form is a Class 1 noun. Note that it is the prefix **ka-** which will govern classification and agreement patterns in constructions. I illustrate in (4) below how concordial agreement between subject noun phrases and their predicates have been presented in this dissertation.

(4) *Concordial agreement between nouns and predicates*

(a)   [mu-nt^{h}u]   [mu-fiípa]   [wa-ŋgu-li-léemb-a]   [bukú]   [li-tuúβa]

      [1-person]   [1-black]   [1SM-past-5OM-write-fv]   [5book]   [5-white]

      'An African wrote (it) a holy book.'

(b)   [ka-mu-nt^{h}u] [ka-fîpa]   [ka-ŋgu-li-léemb-a]   [bukú] [li-tuúβa]

      [12-1-person] [12SM-black] [12SM-past-5OM-write-fv]   [5book] [5-white]

      'A small African wrote (it) a holy book.'

2.7   The Verb Word

I have tried as much as possible to reduce the ambiguity which comes with the use of the terms Stem and Word in African languages. For the purposes of this dissertation, the two most important constituents of a Morphological Word (whole word) are the **Prefix String** and the **Morphological Stem** (or simply Stem). The Morphological Stem, in most cases, can stand on its own (as a free morpheme), while the prefix string is largely dependent on the Morphological Stem. A Basic Morphological Stem consists of a **Root** and a **final vowel** (fv) as given in (5) below.

(5)   *Basic Morphological Stem*

Basic stem   *English gloss*

βeleβeet-a

speak-fv   speak!

biik-a

[cook-fv]   cook!

cʰimbiíj-a

run-fv   run!

As we have seen in many examples above, the commonest final vowel is **-a**. It expresses indicative or neuter and imperative moods. In certain types of verbs, the final vowel is **-e** and it expresses the subjunctive mood to represent obligation or necessity (e.g. **ti-tondek-eéng-e** [we-fail-hab-subj] 'we should be failing').

A Morphological Stem may also consist of the root plus various forms of Suffixation or Extensions preceding the final vowel. The relevant extensions in the present study include the reciprocal (rec), the applicative (appl/dat), the causative (caus), the stative (stat), the passive (pass), and the intensive (int). I illustrate this fact in (6) below.

(6)   *Extended verb stems*

Extended Stem   *English gloss*

βeleβet-aan-a

speak-rec-fv   talk about each other!

bik-iij-a

cook-appl-fv   cook for/with/on!

lemb-ees-a

write-caus-fv   cause to write!

lemb-eés-a

write-int-fv   write intensively!

cʰimbil-iík-a

run-pass-fv   be run on/after!

pum-iík-a

beat-stat-fv   be beaten!

The term Morphological Stem (or simply Stem) will be used to mean both Basic and Extended Stems.

The **prefix string** consists of a subject agreement marker (SM), followed by a member of TAM (tense-aspect-mood) and an object agreement marker (OM). The subject agreement marker occurs in the initial position of the verbal structure and it agrees with verb phrases or modifiers in terms of noun class gender, person and number. It is obligatory as given in (7) below.

(7) *The subject marker is obligatory*

[β-aána]   [βa-ŋgu-bík-á]   [muú-tu].

[2-child]   [2SM-past-cook-fv]   [3-head]

"Children cooked a head"

The object agreement marker follows a member of TAM and it occurs before onset of the Morphological Stem. It is optional and it agrees with the object noun phrase in terms of noun class gender, person and number as given in (8) below.

(8) *The object marker is optional*

[β-aána]   [βa-ŋgu-(wu)-bík-a]   [muú-tu].

[2-child]   [2SM-(3OM)-past-cook-fv] [3-head]

"Children cooked it (a head)"

Tense and aspect are some of the terms I will frequently refer to. They both relate to situations to time. Tense will refer to a technical term distinct from aspect. It is a deictic category that locates an event in time (past, present, future) "usually with reference to the present moment" (Comrie 1976:5, cited in Rose et al 2002:87). The time markers simple past, distant past, simple present, near future and distant future are considered to be tenses in this dissertation. The simple present tense is expressed without a morphological marker. These tenses in ciTonga are illustrated in (9) below.

(9)   *Tenses*

The simple past (PAST)

[mw-aána]   [wa-ŋgu-bík-á]   [mbohóole].

[1-child]   [1SM-past-cook-fv]   [10-potato]

"A child cooked potatoes."

The distant past (PAST)

[mw-aána]   [wa-ku-bík-á]   [mbohóole].

[1-child]   [1SM-dis past-cook-fv]   [10-potato]

"A child cooked potatoes long time ago."

The simple present (PRESENT)

[mw-aána]   [wá-bík-á]   [mbohóole].

[1-child]   [1SM-cook.fv]   [10-potato]

A child cooks potatoes.

The near future (FUTURE)

[mw-aána]   [wá-bík-éng-é]   [mbohóole].

[1-child]   [1SM-cook-near fut-sub]   [10-potato]

"A child cooks potatoes shortly."

The distant future (FUTURE)

[mw-aána]   [wa-zamu-bík-á]   [mbohóole].

[1-child]   [1-SM-dist fut-cook-fv]   [10-potato]

"A child will cook potatoes."

Aspects, on the other hand, "have to do, not with the location of an event in time, but with the temporal distribution or contour" (Hocket 1958:237, cited in Saeed 2003:126). Aspect therefore allows speakers to view an event in various ways such as whether it is completed,

25

incomplete, timeless, "as something stretched over a perceptible period, or as something repeated over a period" (Saeed, *Ibid*). Although I am mindful of the fact that tense and aspect cannot be discretely separated, for the purposes of this study I proceed to consider the present progressive, the past progressive, and the present perfect as being aspectual. These are illustrated in (10) below. The present progressive aspect is expressed without a morphological marker.

(10)   *Aspects in Chitonga*

The present progressive

[Mw-aána]   [wá-bík-á]   [mbohóole].

[1-child]   [SM-cook-fv] [potato]

"A child is cooking potatoes."

The past progressive/habitual

[Mw-aána]   [wá-bík-áng-á]   [mbohóole].

[1-child]   [1SM-cook-past prog-fv]   [10-potato]

"A child was cooking potatoes."

The present perfect

[Mw-aána]   [w-á-bík-á]   [mbohóole].

[1-child]   [1SM-pres perf-cook-fv]   [10-potato]

"A child has cooked potatoes."

The structure of the Morphological Word I will assume in this dissertation, therefore, is like the one I present in (11) below.

(11)   *The ciTonga verb structure*

```
                          Morphological Word
             Prefix String                    Morphological Stem
          SM      TAM   OM            Root    Extensions   Final vowel
E.g.     ndi-    ŋgu-   mu-           bík-       is-           a

         I-past-1OM-cook-caus-fv

         "I caused him to cook."
```
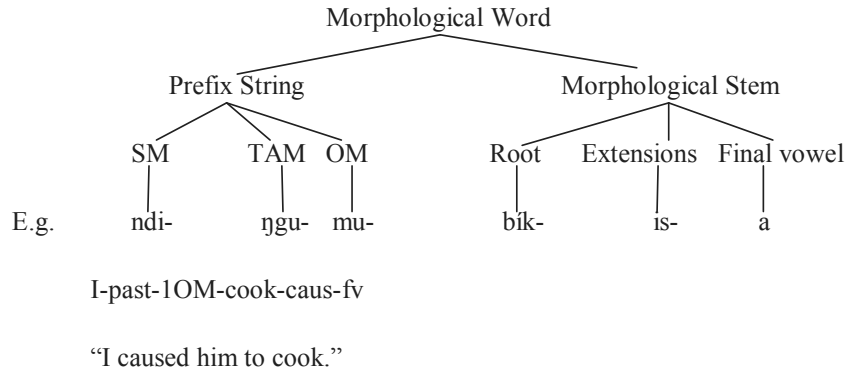
2.8   Tone

According to Mtenje (1994/95, 2006), ciTonga has two tones, namely High and Low. Basic verbs (without prefixes) are classified into High-Toned or Low-Toned depending on whether there is a High Tone in the string. Low-Toned verbs are lexically unmarked for tones while High-Toned verbs are marked with a High Tone on the final vowel (but retracted to the penultimate vowel)[9]. These facts are illustrated in (12) and (13) respectively[10].

(12)   *Low-Toned verbs (Mtenje 1994/95)*

Basic stem	*English gloss*
le.lee.s-a	look at/see
da.nii.k-a	answer
sa.mbii.z-a	teach
mbwii.k-a	jump

---

[9] According to Mtenje (1994/95, 2006), there is a tone retraction rule in ciTonga similar to that found in neighbouring languages like Chichewa and Chiyao (Mtenje 1986, 1993) which retracts a word-final High Tone leftwards to the penultimate vowel under varying conditions such as *the presence of a long penultimate syllable*.
[10] I have transcribed the data.

(13)   *High-Toned basic verb stems (Mtenje 2006)*

Basic stem	English gloss
cʰi.mbií.j-a	run
sa.mbií.l-a	learn
to.ndeé.k-a	fail
kʰu.mbií.l-a	admire
ba.ŋguú.l-a	shout

Generally, it is not always clear whether tones associate to segments, syllables or moras. Mtenje (1994/95, 2006) is not consistent either. He refers to the vowel at one point and the syllable or the mora at another point. I will assume that the TBU is the mora since tone can fall on either the initial or the second mora of bimoraic syllables (e.g. **sóo.mba** 'fish' vs. **guú.tu** 'ear'). Tone is contrastive in rare cases and this type of tone often associates with the penultimate mora of Morphological Stems (e.g. **dii.ka** 'spill' vs. **dií.ka** 'cover oneself', **da.nii.ka** 'answer' vs. **da.nií.ka** 'be invited'). In some cases, I refer to the syllable as the TBU. But this is strictly in cases where the syllables concerned are light and thus monomoraic.

2.9   Chapter summary

This chapter has presented basic facts about ciTonga. These have included language classification, previous works on ciTonga, speech sounds, the syllable, tone, as well as nominal and verbal morphology. The following chapter presents arguments for **Strong Accent Constituency** within the **Universal Guidelines and Constraints framework,** a very slight departure from classical OT.

Chapter 3

Strong Accent Constituents Theory

3.1   Introduction

This chapter presents formally proposals for a theory of Strong Accent Constituency for word-prosody in ciTonga and perhaps many other Bantu languages. The empirical base is found in patterns of vowel and consonant deletion in this language. The chapter is organized as follows: Section 3.2 presents the data on vowel and consonant deletion which motivates this theory analysis. Attempt is made in sections 3.3 and 3.4 to account for the facts in terms of Stress-Accent Theory and Downing's (2006) Morpheme-Based Templates Theory. Both of these theories are found to be slightly inadequate to account for the facts. Section 3.5 argues for and outlines the theory of Strong Accent Constituency. Section 3.6 presents a Strong Accent Constituent Analysis of the facts. In section 3.7 I summarize the chapter.

3.2.   The data

In this section, I am going to present the data on vowel and consonant deletion and preservation which motivate Strong Accent Constituent Analysis of the facts. I have further categorized the data in terms of three speech styles namely, formal, common and elderly speech styles. I use the symbol "[" to show where the Morphological Stem begins.

3.2.1   Vowel Deletion and preservation

3.2.1.1 Formal speech style

In the formal speech style, all vowels of the Morphological Stem are preserved. This fact is exemplified in (1) below.

(1) *No vowel deletion in Formal speech style*

Formal speech	*English gloss*

ti-[βe.le.βe.t-e.s-a.n-ee.ŋg-e
we-speak-caus-rec-hab-subj

we should be causing each other to speak

ti-[ŋa.na.mu.l-i.s-a.n-ee.ŋg-e
we-turn-caus-rec-hab-subj

we should be causing each other to turn

ti-[to.nde.k-e.s-a.n-eé.ŋg-e
we-fail-caus-rec-hab-subj

we should be causing each other to fail

ndi-[go.go.te.z-e.eŋg-e
I-exaggerate-hab-subj

I should be exaggerating

ti-[tʰa.mba.l-i.s-a.n-ee.ŋg-e
we-stretch-caus-rec-hab-subj

we should be stretching legs over one another

3.2.1.2 Common speech style

In the common speech style, non-low vowels of the final syllable are usually deleted. All prefinal vowels [σ_1-σ_{penult}] are preserved. Low vowels are always preserved (including in the final syllable position). Deletion of final non-low vowels is exemplified in (2) below[11].

---

[11] The final mid vowels [e] and [o] can also be realized as high vowels [i] or [u], respectively (e.g. **le.k-a.n-aa-.n-e** ⇒ **le.k-a.n-aa-.ni**, **ga.lí.moo.to** ⇒ **ga.lí.moo.tu** or (rarely) **ga.lí.moo.ti**). This observation can be explained in terms of both prominence reducing and vowel reduction and it may point to the fact that the final syllable is prosodically nonprominent (cf. Crosswhite 2004).

(2) *Non-low vowels of final syllables are deleted in the common speech style*

Formal speech   *Common speech*

ndi-.to.ndeé.k-e   **ndi-.to.ndeé.k**
I-fail-subj   "I should fail"

le.k-a.n-aa-.n-e   **le.k-a.n-aa-.n**
leave-rec-indic-hon   "leave each other"

mu-.to.ndoó.ne   **mu-.to.ndoó.n**
1-lizard   "lizard"

ga.lí.moo.to   **ga.lí.moo.t**
5-car   "car"

ti-[βe.le.βe.t-e.s-a.n-ee.ŋg-e   **ti-[βe.le.βe.t-e.s-a.n-ee.ŋg**
we-speak-caus-rec-hab-subj   "we should be causing each other to speak"

ti-[ŋa.na.mu.l-i.s-a.n-ee.ŋg-e   **ti-[ŋa.na.mu.l-i.s-a.n-ee.ŋg**
we-turn-caus-rec-hab-subj   "we should be causing each other to turn"

ti-[to.nde.k-e.s-a.n-eé.ŋg-e   **ti-[to.nde.k-e.s-a.n-eé.ŋg**
we-fail-caus-rec-hab-subj   "we should be causing each other to fail"

3.2.1.3 Elderly speech styles

In common elderly speech styles the only non-low vowels that are preserved are those of the first two syllables (σ_1-σ_2) and the bimoraic penultimate syllable of the Morphological Stem. The rest may be deleted. Deletion of low vowels is not acceptable. These facts are illustrated in (3) below.

(3)   *No deletion of vowels of first two syllables and the penultimate syllable;*
        Also, no deletion of low vowels in common elderly speech style.

Formal speech   *Elderly speech style*

ti-[βe.le.βe.t-ee.ŋg-e   **ti-[βe.le.β.t-ee.ŋg**
we-speak- hab-subj   "we should be speaking"

ti-.[zo.me.le.z-e.k-ee.ŋg-e   **ti-.[zo.me.l.z-.k-ee.ŋg**
we-accept-pass-hab-subj   "we should be acceptable"

ti-[to.nde.k-e.s-a.n-eé.ŋg-e   **ti-[to.nde.k-.s-a.n-eé.ŋg**
we-fail- caus-rec-hab-subj   "we should be failing each other"

ndi-[go.go.te.z-ee.ŋg-e   **ndi-[go.go.t.z-ee.ŋg-**
I-exagerate-hab-subj   I should be exagerating

ti-[βe.le.βe.t-e.s-a.n-ee.ŋg-e   **ti-[βe.le.β.t-.sa-.n-ee.ŋg**
we-speak-caus-rec-hab-subj   "we should be causing each other to speak"

ti-[tʰa.mba.l-i.s-a.n-ee.ŋg-e   **ti-[tʰa.mba.l-.s-a.n-ee.ŋg**
we-stretch-caus-rec-hab-subj   "we should be stretching legs over one another"

[βe.le.βe.t-e.s-a.n-aa.ŋg-a   **[βe.le.β.t-.sa-.n-aa.ŋga**
speak-caus-rec-hab-fv   "be causing each other to speak"

[to.nde.k-e.s-a.n-aá.ŋg-a   **[to.nde.k-.sa-.n-aá.ŋga**
fail-caus-rec-hab-fv   "be causing each other to fail"

[tʰa.mba.l-i.s-a.n-aa.ŋg-a   **[tʰa.mba.l-.sa-.n-aa.ŋga**
stretch-caus-rec-hab-fv   "be causing each other to stretch legs over one another"

In some rare cases, elderly speakers may delete low vowels as well. Vowels of the first two syllables [σ_1-σ_2] and the bimoraic penultimate syllable [σ_{penult}] of the Morphological Stem are

preserved as well. Low vowels of the final syllable are also preserved[12]. These facts are illustrated in (4) below.

(4)   *Elderly speakers may in some rare cases delete low vowels as well*

Formal speech   *Rare elderly speech style*

ti-[βe.le.βe.t-e.s-a.n-ee.ŋg-e   **ti-[βe.le.β.t-.s-.n-ee.ŋg**
we-speak-caus-rec-hab-subj   "we should be causing each other to speak"

ti-[to.nde.k-e.s-a.n-eé.ŋg-e   **ti-[to.nde.k-.s-.n-eé.ŋg**
we-fail-caus-rec-hab-subj   "we should be causing each other to fail"

ti-[tʰa.mba.l-i.s-a.n-ee.ŋg-e   **ti-[tʰa.mba.l-.s-.n-ee.ŋg**
we-stretch-caus-rec-hab-subj   "we should be stretching legs over one another"

[βe.le.βe.t-e.s-a.n-aa.ŋg-a   **[βe.le.β.t-.s-.n-aa.ŋga**
speak-caus-rec-hab-fv   "be causing each other to speak"

[to.nde.k-e.s-a.n-aá.ŋg-a   **[to.nde.k-.s-.n-aá.ŋga**
fail-caus-rec-hab-fv   "be causing each other to fail"

3.2.2   Consonant Deletion and preservation

3.2.2.1 Formal speech style

In formal speech style, there is no consonant deletion within the Morphological Stem. This fact is illustrated in (5) below.

---

[12] The low vowel may also be pronounced as an [i] or [e] (e.g. **to.nde.k-e.s-a.n-aá.ŋg-a** ⇒ **to.nde.ki-.si-.n-aá.ŋga/ to.nde.ki-.se-.n-aá.ŋga**).

(5)   *No consonant deletion in a more formal speech style*

Formal speech   *English gloss*

[tʰa.mba.lii.l-a

stretch-appl-rec-fv   stretch legs over each other

[se.k-e.l-e.s-aa.n-a

laugh-appl-caus-rec-fv   cause to laugh for each other

[zo.me.le.z-aa.n-a

agree-rec-fv   agree with each other

[vu.nu.li.l-a.n-aa.ŋg-a

steam-rec-hab-fv   be steaming each other

[pu.muu.l-a

rest-fv   rest

[su.k-ii.l-a

clean-appl-fv   clean for

3.2.2.2 Common speech style

In the common speech style, however, liquid consonant onsets to final syllables are usually deleted as given in (6) below[13]. Nasal and obstruent consonants are never deleted anywhere. When the liquid onset is deleted, the preceding high vowel syllabifies in the onset of the final syllable, hence the slur (gliding 'ʲ') referred to by Turner (1952). The symbol '<>' indicates that a whole syllable (made up of liquid onset and non-low vowel) has been elided.

---

[13] Recall that low vowels are preserved in *common speech style* as well as in common elderly speech styles.

(6) *Liquid onsets to final syllables are usually deleted in common speech style*

Formal   *Common speech*   *English gloss*

[bi.k-ii.l-a   **[bi.k-ii-.^ja**
cook-appl-fv   cook for

[pu.muu.l-a   **[pu.muu-.^wa**
rest-fv   rest

[su.k-ii.l-a   **[su.k-ii-.^ja**
clean-appl-fv   clean for

[t^ha.mba.li.l-ii.l-a   **[t^ha.mba.li.l-ii-.^ja**
stretch-appl-fv   stretch legs for

[mbe.lee.le   **[mbe.lee.<'>**
9/10-sheep   sheep

3.2.2.3 Elderly speech styles

In the common elderly speech style, however, the only liquid consonants that are preserved are those which are onsets to the first two syllables and the penultimate syllable of the Morphological Stem. Nasal and obstruent consonants are never deleted anywhere. The former fact is illustrated in (7) below. Once again, the symbol '<>' indicates that a whole syllable has been elided.

(7) *Liquid deletion in common elderly speech style*

Formal	*Elderly*	*English gloss*
[βe.le.βe.t-e.l-aa.n-a	**[βe.le.β.t-.l-aa.n-a**	
speak-appl-rec-fv		speak for each other
[ka.li.p-i.l-aa.n-a	**[ka.li.p-.l-aa.n-a**	
reprimand-appl-rec-fv		reprimand for each other
[tʰa.mba.l-i.l-aan-a.	**[tʰa.mba.<'>.l-aa.na.**	
[stretch-appl-rec-fv]		stretch legs over each other
[se.k-e.l-e.s-aa.n-a	**[se.ke-.<'>.s-aa.na**	
[laugh-appl-caus-rec-fv]		cause to laugh for each other
[zo.me.le.z-aa.n-a	**[zo.me.<'>.z-aa.na**	
[agree-rec-fv]		agree with each other
[vu.nu.li.l-a.n-aa.ŋg-a	**[vu.nu.<'>.-ʷa.n-aa.ŋg-a**	
[steam-rec-hab-fv]		be steaming each other

In rare cases, some elderly speakers may also delete nasal consonant onsets (e.g. **ka.li.p-i.l-aa.n-a ⇒ ka.li.p-.l-ãã**)[14]. This is over and above liquid and low vowel deletion. Deletion of obstruent consonants, however, is just not acceptable (e.g. formal **zo.me.lee.z-a ⇒ *zo.me.lee**). Just like we saw in the vowel deletion patterns, many speakers may fall in between these extremes.

Summary

Vowel deletion: In the formal speech style, all vowels of the Morphological Stem are preserved. In the common speech style, non-low vowels of the final syllable are usually deleted. Pre-final non-low vowels [σ_1-σ_{penult}], on the other hand, are usually preserved. Low vowels are also preserved in all positions. In common elderly speech style, however, the only non-low vowels that may be preserved are those which are nuclei of the first two syllables and stem-penult syllables. Deletion of low vowels is not permissible. In rare elderly speech styles, however, low vowels may be deleted as well, but not those belonging to the first two syllables

---

[14] Note that the feature [+nasal] itself does not delete. My analysis is that this feature is a Strong Accent Constituent in ciTonga.

and the penultimate and the final syllables. Many speakers, as I will show later, fall in between these speech styles and they are not easy to categorize. Furthermore, a vowel [i] or [e] can be heard in positions where a vowel is supposed to be deleted (see footnotes 11 and 12).

Consonant deletion: All consonants of the Morphological Stem are preserved in the more formal speech style. In the *common speech style*, liquid consonant onsets to final syllables are usually deleted. Nasal and obstruent consonants are never deleted anywhere. In the *common elderly speech style*, the only liquid consonants that are preserved are those which are onsets of the first two syllables and the penultimate syllable of the Morphological Stem. The rest may be preserved. Nasal and obstruent consonants are also never deleted anywhere. In addition to multifarious vowel and liquid consonant deletion, some elderly speakers in *rare speech styles* may also delete nasal consonant onsets. Obstruent consonants are however usually preserved in the same positions. Many other speakers fall in between these extremes.

3.3 Stress analysis

3.3.1 Vowel deletion and preservation

I will start my analysis by pointing out that, among others, vowel deletion has been presented as a hiatus resolution strategy in many languages. According to this view, vowel deletion comes as a result of a phonological rule that deletes an input vowel whenever there is another vowel approaching. The forms in (8) below illustrate vowel deletion in Luganda spoken in Uganda (Katamba 1989:171)[15].

(8) *Vowel deletion in Luganda (Katamba 1989:171)*

/ba-a-lab-a/   [ba:laba]   'they saw'

/ba-e-lab-a/   [be:laba]   'they see themselves'

/ma-a-to/   [ma:to]   'boats, canoes'

In the ciTonga data presented above, however, there is no any approaching vowel that would trigger deletion of the final or pre-penult vowels. The nature of vowel deletion exhibited in

---

[15] Note that only the second example exhibits a clear case of vowel deletion and it even retains the mora of the added vowel. In the other two examples, deletion of the prefix vowel is not so obvious since the vowels in the sequence are identical.

our data resembles processes in stress languages[16]. Many Bantu languages have been described as being tonal, meaning that tone is part of the lexicon (and Proto-Bantu is reconstructed as having two tones (High and Low). However, many of them are also reported to have stress-accent independent from tone (Downing 2004:120). Findings in Downing's (2004) survey of Bantu languages confirm Hyman's (1977) earlier finding that (stem-) initial and penult are the most common positions, crosslinguistically, to be assigned main stress. Downing (2004) lists about 26 Bantu languages as having penultimate stress. Only two languages in the list (western Lingala dialects and Luvale) have stem-initial stress. Luvale has stem-penult stress as well. The most commonly reported phonetic correlate of penultimate stress in both tonal and non-tonal languages is vowel lengthening.

Indeed, like many Bantu languages, ciTonga has penultimate vowel lengthening (PL). Thus, the fact that stress may fall on the penultimate syllable is further solidified by the fact that when the words are extended, such as when suffixes are added, penultimate lengthening shifts to new penultimate syllables, an indication that penultimate lengthening is something which is automatic and worth our attention. These facts are exemplified by data from the *formal* speech style of ciTonga given in (9) below.

---

[16] As I said in footnote 11, this observation can also be explained in terms of both prominence reducing and vowel reduction and it may point to the fact that syllables or units which preserve their input vowels are prosodically prominent (cf. Crosswhite 2004).

(9)   *Syllabic foot analysis of penultimate lengthening*

Basic verb   *English gloss*

[be.'nee.k-a   cover
[be.ne.'k-aa.n-a   cover each other

[le.'lee.s-a   look
[le.le.'s-aa.n-a   look at each other

['vii.n-a   dance
[vi.'n-ii.s-a   cause to dance

['bii.k-a   cook
[bi.k-i.'l-aa.n-a   cook for each other

[βe.'lee.ŋg-a   read
[βe.le.ŋg-e.'s-aa.n-a   cause each other to read

Let me now turn to the issue of vowel deletion. Two reasons often cited for stress-driven vowel deletion are that such processes either ensure that (1) the number of unparsed syllables is minimized or (2) a preceding stressed syllable is heavy (e.g. Gouskova 2003, citing Hill et al. 1998, Jeane 1978, 1982). In terms of OT, vowel deletion is accounted for by interaction of metrical constraints and FAITHFULNESS. Two metrical constraints relevant to this discussion are **STRESS-TO-WEIGHT** and **PARSE-σ**.

(10)   **PARSE-σ**
        Every syllable is contained inside a foot structure (Prince and Smolensky 1993).

(11)   **STRESS-TO-WEIGHT**
        "If stressed, then heavy" (Kager 1999:268).

However, vowel deletion as presented in ciTonga would not be triggered by the need to ensure that the stressed syllable is heavy. Among others, this is because final vowel deletion takes place even when the preceding penultimate syllable in output forms is heavy (e.g. formal **ti-[βe.le.βe.t-e.s-aa.n-e** ⇒ **ti-[βe.le.βe.t-e.s-aa.n**). Secondly, the principle of STRESS-TO-WEIGHT would not account for the fact that mid vowel may also be realized as a high vowel instead of it being deleted (e.g. **ti-[βe.le.βe.t-e.s-aa.n-e** ⇒ **ti-[βe.le.βe.t-e.s-aa.ni**).

Therefore, the constraint Parse-σ, requiring every syllable to be contained inside a foot structure, looks promising (cf. Prince and Smolensky 1993). Any segment or syllable which is not parsed by a foot incurs a violation mark for this constraint. To satisfy or to minimally violate it, syllables or segments which are outside the foot structure must be deleted, or they should be reduced in number, respectively.

If the constraint Parse-σ is indeed responsible for deletion of non-low final vowels (mid vowels), then the foot type in this language is not a syllabic trochee. The final syllable must first and foremost be outside the seemingly moraic foot structure in order for its vowel to undergo deletion. This argument is sensible for another reason. Uneven trochees (Heavy+Light) are reported in literature (e.g. Jacobs 1990, 2000; Rice 1992; van der Hulst and Klamer 1996; and Mellander 2001, 2004), but they are not supported by others (see Hayes 1995). It is most likely then that the foot type is a moraic trochee. Deletion of the final mid and high vowels, however, leads to violation of the constraint **Max-Mid Vowel** (**Max-MidV**) which requires input mid vowels to be preserved in the output. It also leads to violation of **Max-High Vowel** (**Max-HV**) which militates against deletion of high vowels in the output. The fact that a mid or high vowel is eventually deleted indicates that the constraint Parse-σ ranks above Max-MidV and Max-HV.

 Unfortunately, in our data, the number of syllables outside the foot structure does not reduce significantly since the consonant onsets (Cs) may not undergo the process. Perhaps (for argument's sake) this can be attributed to a faithfulness constraint **Max-C**, requiring input consonants to be preserved in the output. Furthermore, the surviving consonants are not even syllabified as codas perhaps due to a syllable structure constraint **NoCoda** (Kager 1999:122) which is highly ranked in many Bantu languages. These (stranded pre-final) consonants are also not re-syllabified in new onset positions perhaps due to another high-ranking constraint ***Complex Onset** which bans sequences of consonant onsets.

(12) **Max-MidV**

   Input mid vowels are preserved in the output.

(13) **Max-HV**

   Input high vowels are preserved in the output.

(14) **Max-C**

   Input consonants are preserved in the output.

(15) **NOCODA**

Syllables have no codas.

(16) ***COMPLEX ONSET**

No complex onsets

With these facts in mind, we can now attempt to account for the patterns of vowel deletion and preservation in the three speech styles of ciTonga.

3.3.1.1 Formal speech style

The observation in the formal speech style was that all vowels of the Morphological Stem are preserved. The analysis of this fact would be that outputs where all vowel contrasts are preserved are optimal. In terms of OT, a **MAX-IO** constraint ranks above the constraint PARSE-σ. The constraint MAX-IO requires that input segments be preserved in the output. The constraint PARSE-σ requires every segment or syllable in a Prosodic Word to be parsed by Feet. Any segment which is not parsed incurs a single violation mark for this constraint. A tableau in (18) below illustrates. Counting concerns segments of the Morphological Stem only (demarcated by the symbol "[").

(17) **MAX-IO**

Input segments are preserved in the output

(18) *MAX-IO ranks above PARSE-σ in formal speech style*

/ti-[to.nde.k-e.s-a.n-e.ng-e/	MAX-IO	PARSE-σ
☞ (a) ti-[to.nde.ke-.s-a.('n-ee).ŋg-e		10!
(b) ti-[to.nde.ke-.s-a.('n-ee).ŋg	*!	9

Although candidate (18a) has the largest number of violators of PARSE-σ, it is still optimal because it satisfies the high-ranking constraint MAX-IO which requires input segments to be preserved in the output. Candidate (18b) is ruled out for violating MAX-IO, although it minimally violates PARSE-σ.

3.3.1.2 Common speech style

Just like in the formal speech style, the vowel of the penultimate syllable in the common speech style is long. However, non-low final vowels are deleted (e.g. formal **ti-[βe.le.βe.t-e.s-**

aa.n-e ⇒ **ti-[βe.le.βe.t-e.s-aa.n**). The stress analysis would be that candidates which delete non-low final vowels are optimal. In terms of OT, it is most likely that the constraint PARSE-σ outranks MAX-MIDV. A tableau in (19) below illustrates.

(19) *Mid vowel deletion triggered by PARSE-σ; the foot is most likely moraic.*

/ti-[βeleβet-es-an-e/ 'we should cause each other to speak'	MAX-C	PARSE-σ	MAX-MIDV
(a) ti-[βe.le.βe.t-e.('s-aa).n-e.		10!	
(b) ti-[βe.le.βe.t-e.('s-aa).	*!	8	*
☞ (b) ti-[βe.le.βe.t-e.('s-aa).n.		9	*

Thus, candidate (19a) is ruled out because of its maximal violation of the constraint PARSE-σ which requires that every syllable in a Prosodic Word/Stem be parsed by feet. Candidate (19b) is unsuccessful because it violates a high-ranking constraint MAX-C which militates against deletion of consonant segments. Candidate (19c) is optimal because it satisfies MAX-C and it minimally violates PARSE-σ. It violates a low-ranking constraint MAX-MIDV which can be tolerated.

One 'challenge'[17] to the PARSE-σ theory, however, is the preservation of low vowels in the same final syllable position (e.g. [βe.le.βee.ta ⇒ [βe.le.βee.ta, *βe.le.βee.t). The fact that a final low vowel (LOWV) resists deletion would indicate that a constraint **MAX-LOWV** ranks above PARSE-σ in the common speech style. A tableau in (21) below illustrates.

---

[17] I am using the word 'challenge' with reference to the context of the present discussion.

(20) **MAX-LOWV**

Input low vowels are preserved in the output.

(21) *Low vowel resists deletion in the common speech style*

/bombone/ 'black ants'	MAX-LOWV	PARSE-σ
☞ (a) βe.le.('βee.)ta		***
(b) βe.le.('βee).t	*!	

Thus, candidate (21a) is optimal because it satisfies a high-ranking constraint MAX-LOWV. Candidate (21b) on the other hand is ruled out for violating this crucial constraint. Many *questions* are left unanswered by this nature of analysis. Is the moraic foot still on the penultimate syllable or because of the most sonorous vowel [a] it is now shifted to the edgemost position? Is there something that preserves the final low vowel which is at par with something (a prosodic constituent) that initially preserved pre-final non-low vowels?

Positional faithfulness: Indeed one other crucial observation was that non-low vowels of pre-final syllables (σ_1-σ_{penult}) are duly preserved. What easily comes to mind is the idea of positional faithfulness. The ranking of a positional faithfulness constraint "**MAX-[σ_1-σ_{penult}]**" above a markedness constraint PARSE-σ would precisely ensure that the contrast is maintained in the strong position [σ_1-σ_{penult}].[18] To put it in familiar words, the suggestion is that a markedness constraint PARSE-σ succeeds in neutralizing weaker positions such as the final syllable position (σ_n), but it fails to have any effect on the strong position (σ_1- σ_{penult}). A tableau in (22) below illustrates.

---

[18] Examples of strong positions where contrast is generally present include stressed syllables (Trubetskoy 1939; Selkirk 1994; Alderete 1995), initial syllable (Trubetskoy 1939; Selkirk 1994; Beckman 1995, 1998), Root (McCarthy and Prince 1995; Casali 1996) and nouns (Smith 2001). At first glance, one would say that it is most likely that the strong position in the common speech style of ciTonga is the Root. This, however, is not the case here.

(22) *Positional faithfulness in the common speech style*

MAX-$[\sigma_1\text{-}\sigma_{pen}]$ > PARSE-σ

/ti-[βeleβet-es-an-e/ 'we will cause each other to speak'	MAX-$[\sigma_1\text{-}\sigma_{pen}]$	PARSE-σ	MAX-MIDV
(a) ti-[βe.le.βe.t-e.('s-aa).n-e.		10!	
(b) (b) ti-[βe.le.β.t-.('s-aa).n.	*!**	7	***
☞ (b) ti-[βe.le.βe.t-e.('s-aa).n.		9	*

Candidate (22a) is disqualified because it has a maximal number of violations of the constraint PARSE-σ. Candidate (22b) is ruled out for violating a high-ranking positional faithfulness constraint MAX-$[\sigma_1\text{-}\sigma_{pen}]$ which requires preservation of segments of the strong position $[\sigma_1\text{-}\sigma_{pen}]$. Candidate (22c), on the other hand, is optimal because it satisfies MAX-$[\sigma_1\text{-}\sigma_{penult}]$ and it violates PARSE-σ only minimally. The *challenge* with this analysis is that we don't know what the constituent $[\sigma_1\text{-}\sigma_{pen}]$ is all about.

3.3.1.3 Elderly speech styles

In the common elderly speech style, the first observation was that the only vowels that are preserved are those which are nuclei of the first two syllables and the bimoraic penultimate syllable. Once again, low vowels are usually preserved. The stress analysis of the first observation would be that the two Stem-initial syllables and the bimoraic penultimate syllable are parsed by moraic Feet as given in (23) below.

(23) *Feet in common elderly speech*

Formal speech   *Common speech style*

ti-[βe.le.βe.t-e.s-a.n-ee.ŋg-e   **ti-[(βe.le).β.t-.sa-.('n-ee).ŋg**
we-speak-caus-rec-hab-fv   "we should be causing each other to speak"

ti-[to.nde.k-e.s-a.n-ee.ŋg-e   **ti-[(to.nde).k-.s-a.('n-ee).ŋg**
we-fail-caus-rec-hab-fv   "we should be causing each other to fail"

ti-[tʰa.mba.l-i.s-a.n-ee.ŋg-e   **ti-[(tʰa.mba).l-.s-a.('n-ee).ŋg**
we-stretch-caus-rec-hab-fv   "we should be stretching legs over one another"

[βe.le.βe.t-e.s-a.n-aa.ŋg-a   **[(βe.le).β.t-.sa-.('n-aa).ŋga**
speak-caus-rec-hab-fv   "be causing each other to speak"

[to.nde.k-e.s-a.n-aa.ŋg-a   **[(to.nde).k-.sa-.('n-aa).ŋga**
fail-caus-rec-hab-fv   "be causing each other to fail"

Some of the generalizations would be that Prosodic Stems begin and end with a Foot and that no Foot is placed on final syllables. Adapting Kager (2007:211) in a parallel analysis, the fixed Foot in common elderly speech styles requires stem-to-foot alignment, requiring that every Prosodic Stem begins or ends with a foot, as captured by the constraint pair below:

(24) **ALIGN (PRSTEM, L, FT, L)**

      Every Prosodic Stem begins with a foot.

(25) **ALIGN (PRSTEM, R, FT, R)**

      Every Prosodic Stem ends with a foot.

The well-known constraint NONFINALITY would ensure that no foot is on the final syllable.

A crucial generalization for our purpose, however, is that outputs which preserve segments of stress Feet are optimal. This fact, once again, can be accounted for in terms of positional faithfulness. It seems that the ranking of a positional faithfulness constraint 'MAX-FOOT above PARSE-σ' (the absurdity) ensures that vowel segments of the strong position stress Foot are preserved. A tableau in (27) below illustrates.

(26) MAX-FOOT

Segments of the strong position Foot are preserved.

(27) *Feet preserve their mid vowels*

MAX-FOOT >> PARSE-σ >> MAX-MIDV

/ti-[βeleβet-es-eng-e/ 'we should be causing to speak'	MAX-FOOT	PARSE-σ	MAX-MIDV
(a) ti-[(βe.le.)βe.t-e.('s-ee).ŋg-e		6!	
(b) ti-[(β.l).β.t-.(s.).ŋg	*!**	3	******
☞ (c) ti-[(βe.le).β.t-.('s-ee).ŋg		3	***

Thus, candidate (27a) is disqualified because it maximally violates the constraint PARSE-σ. Candidate (27b) is ruled out for violating a high-ranking positional faithfulness constraint MAX-FOOT which prohibits deletion of segments of the strong prosodic category Foot. On the other hand, the candidate (27c) is optimal because it satisfies MAX-FOOT and it violates PARSE-σ only minimally.

A serious challenge to a stress thory based on preservation of segments is brought about by faithfulness of the low vowel. It is not easy to tell where the foot falls. One observation in the common elderly speech style was that the low vowel is preserved, regardless of the position where it is found (e.g. formal **t-a-.[βe.lé.βé.t-é.s-á.n-aa.ŋg-a** ⇒ **t-a-.[βe.lé.β.t-.s-á.n-aa.ŋg-a**). A generalization would be that output forms which preserve input low vowels are optimal. In terms of OT, the analysis would be that the constraint PARSE-σ is outranked by MAX-LOWV as given in (28) below.

(28)   *Are all preserved low vowels also footed?*

/t-a-[βeleβet-es-an-aŋg-a/ 'We used to cause to speak to each other'	MAX-LOWV	PARSE-σ
☞ (a) t-a-.[(βé.lé).β.t-.s-a.(n-aa).ŋg-a		******
☞ (b) t-a-.[(βé.lé).β.t-.s-a.(n-aa.ŋg-a)		****
☞ (c) t-a-.[(βé.lé).β.t-.(s-a.n-aa.ŋg-a)		**
(d) t-a-.[(βé.lé).β.t-.s-.(n-aa).ŋg-	*!*	

The challenge lies in the fact that the stress analysis which relies on preservation of segments as a phonetic correlate would not only yield doubtful feet such as the right feet in (28b&c), but it would also be difficult to tell where the Foot falls exactly among the first three candidates (28a&b&c).

Another crucial observation in the elderly speech styles was that low vowels may be deleted as well in rare cases. Analysis of this fact would be that it is more important to have syllables parsed by feet than mere preservation of low vowels. In terms of OT, the constraint PARSE-σ would outrank the constraint MAX-LOWV in this speech style. A tableau in (29) below illustrates.

(29) *PARSE-σ may outrank MAX-LOWV in rare elderly speech style*

/ti-[βeleβet-es-an-eŋg-e/ 'we should make each other speak '	PARSE-σ	MAX-LowV	MAX-MIDV
(a) ti-.[(βe.le.)βe.t-e.s-a.('n-ee).ŋg-e	8!		
(b) ti-.[(βe.le.)β.t-.s-a.('n-ee).ŋg-e	6!		**
☞ (c) ti-.[(βe.le.)β.t-.s-.('n-ee).ŋg-e	5	*	**

47

Thus, candidates (29a&b) are disqualified because they have a maximal number of violations of the constraint PARSE-σ. On the other hand, candidate (29c) is successful since it has a minimal number of violations of PARSE-σ. It violates low-ranking constraints MAX-LowV and MAX-MIDV, but these violations can be tolerated for the sake of PARSE-σ. Although this last analysis looks neat, the stress analysis of vowel deletion and preservation faces more challenges than opportunities for Bantu languages like ciTonga as we will further see in forthcoming sections.

3.3.2 Consonant deletion and preservation

Observations about consonant deletion included the following facts: All consonants of the Morphological Stem are preserved in the more formal speech style (e.g. [tʰa.mba.li.l-ii.l-a 'stretch legs for'). In the *common speech style*, liquid consonant onsets to final syllables are usually deleted (e.g. [tʰa.mba.li.l-ii.l-a ⇒ [tʰa.mba.li.l-ii-.ʲa). Nasal and obstruent consonants are never deleted anywhere. In the *common elderly speech style*, the only liquid consonants that are preserved are those which are onsets of the first two syllables and the penultimate syllable of the Morphological Stem (e.g. [vu.nu.li.l-a.n-aa.ŋg-a ⇒ [vu.nu.<'>.-ʷa.n-aa.ŋg-a 'be steaming each other'). Nasal and obstruent consonants are also never deleted anywhere. In addition to multifarious vowel and liquid consonant deletion, some elderly speakers in *rare speech styles* may also delete nasal consonant onsets of the final syllable (e.g. [ka.li.p-i.l-aa.n-a ⇒ [ka.li.p-.l-ãã 'reprimand each other'). Obstruent consonants are usually preserved in the same position.

I will start my analysis by reviewing Turner's (1952) views and my own views in Mkochi (2007/08) regarding liquid consonant deletion, and final syllable slur or deletion in ciTonga. According to Turner (1952:i-ii), native speakers of ciTonga are rapid speakers and because of this they are fond of slurring (gliding) or deleting final syllables. Turner writes in an introduction to his *Tonga-Tumbuka-English Dictionary*:

> The Tonga column gives the words only when they are different from the Tumbuka – the words being in many cases the same. The Tonga folk, being rapid speakers, slur or elide the final syllable of many words: e.g. *kulira* becomes *kuliya*, *kukura* becomes *kukuwa*, *kutora* becomes *kuto'* – the accent in the last example remaining on the final syllable, indicating that the terminal *-ra* has been elided. In further inflections of the verb, however, the elided syllable is first restored, then the final *-ra* is again elided, so that *kutorera* becomes *kutore'*.

Turner was referring to the common speech style and I will transcribe his data as in (30) below. The slur is indicated by the symbol (ʲ) or (ʷ).

(30) *Final syllable slur and deletion in ciTonga (Turner 1952, transcribed)*

Formal	*Common*	*English gloss*
ku-.[líi.l-a inf-cry-fv	**ku-.[líi-.ʲa**	to cry
ku-.[kúu.l-a inf-grow up-fv	**ku-.[kúu-.ʷa**	to grow up
ku-.[tóo.l-a inf-take-fv	**ku-.[tóo**	to take
ku-.[tó.l-ee.l-a inf-take-appl-fv	**ku-.[tó.l-ee**	to take for

Turner makes a significant observation that final syllable slur and deletion are a common phenomenon in this language. However, his analysis that ciTonga speakers delete or slur final syllables because they are rapid speakers is less scientific. CiTonga speakers delete or slur final syllables even when they are at their slowest speech rate.

A crucial observation about final syllable deletion in ciTonga is made by myself in Mkochi (2007/08) where I observe that liquid onsets to light final syllables are deleted. I observe that liquids do appear everywhere, but not in the position before the final mora. A generalization I draw is that there is no liquid in the position before the final mora. In terms of OT, I suggest a constraint **No l (fin. μ)** which bans liquids from occurring before the final mora.

(31)   **No l (fin. μ)**

     No liquid before final mora (Mkochi 2007/08)

I also observe in Mkochi (2007/08) that a slur occurs only when the vowel of a syllable preceding the final syllable is a high vowel (e.g. formal [**bi.kii.la** ⇒ [**bi.kii.ʲa** 'cook for'). My suggestion is that once the liquid is deleted, the high vowel of a preceding syllable syllabifies in the onset of the following final syllable. In the forms where the final syllable is completely

49

elided, I observe that all the preceding vowels there are [-high] (e.g. [**pe.mbee.la** 'kindle for' ⇒ [**pe.mbee**; vs. ***pe.mbee.ya**). The analysis I make there is that non-high vowels, being more sonorous than high vowels, make bad onsets and as such they are not allowed in syllable margins in this language (cf. Zec 2007). Thus, since the liquid of the final syllable is deleted, the final vowel [**-a**] is left stranded (onset-less). I argue that a syllable without an onset violates the constraint **ONSET** which requires every syllable to have an onset. Since an onset cannot be obtained for the final syllable, the end result is that both the liquid and the low vowel of the final syllable are elided[19].

(32)   **ONSET**

    Syllables have onset.

This description manages to capture some crucial facts about liquid consonant and syllable deletion in ciTonga. However, it does not exhaust most of the necessary data on consonant deletion in this language as we do here. Secondly, the constraint **No l (fin. μ)** aids in capturing my claim, but the constraint itself is language- and context-specific. In other words, it makes little or no appeal to universality. It appears therefore that analysis based on stress may better account for the facts.

In terms of stress, the suggestion would be that deletion of vowels and consonants which are outside the foot structure prevents or minimizes violation of the constraint PARSE-σ which requires every syllable or segment to be parsed by a Foot structure. Deletion of consonants and vowels to satisfy or to minimally violate the constraint PARSE-σ, however, leads to violation of another constraint MAX-IO which requires input segments to be realized in the output. The fact that some types of vowel and consonant segments are eventually deleted indicates that the constraint PARSE-σ outranks some constraints which specifically require preservation of these types of segments. In the following sub-sections, I attempt to analyse consonant deletion in terms of the interaction of the metrical constraint PARSE-σ and FAITHFULNESS constraints.

3.3.2.1 Formal speech style

The observation in the formal speech style was that all consonants (and vowels of course) of a Prosodic Stem are preserved. The generalization is that output forms which preserve all

---

[19] It is also possible, and I now believe, that the final vowel just undergoes vowel coalescence.

segments are optimal. Just like in our discussion of vowel deletion, this generalization may be linked to ranking of MAX-IO above PARSE-σ. A tableau in (33) below illustrates.

(33) *MAX-IO >> PARSE-σ in formal speech style*

/bik-il-a/	MAX-IO	PARSE-σ
☞ (a) [bi.('k-ii).l-a		****
(b) [bi.('k-ii). ‑ʲa	*!	***

Although candidate (33a) has the largest number of violators of PARSE-σ, it is still the optimal candidate because it satisfies the high-ranking constraint MAX-IO which requires input segments to be preserved in the output. Candidate (33b) is ruled out for violating MAX-IO.

3.3.2.2 Common speech style

In the common speech style, however, the observation was that liquid consonant onsets to final syllables are usually deleted. Nasal and obstruent consonants are not deleted in the same position. In addition, liquid consonants of the position $[\sigma_1\text{-}\sigma_{penult}]$ are also preserved. The generalization here is that outputs which delete liquid consonants, but preserve liquid consonants and any other consonant of the position $[\sigma_1\text{-}\sigma_{penult}]$ are optimal. In terms of stress, the trigger of liquid consonant deletion would be the need to minimize the number of syllables and segments which are outside the foot structure. And in terms of OT, the fact that liquid consonant onsets to final syllables are deleted would indicate that the constraint PARSE-σ outranks MAX-LIQ, requiring liquid consonants to be preserved in the output. On the other hand, the preservation of obstruent and nasal consonants can be attributed to the constraints MAX-OBS and MAX-NAS outranking PARSE-σ. Finally, the preservation of liquid consonants in the position $[\sigma_1\text{-}\sigma_{penult}]$ can, once again, be linked to the positional faithfulness constraint MAX-$[\sigma_1\text{-}\sigma_{penult}]$ outranking PARSE-σ. A tableau in (37) below illustrates. Since low vowels are usually preserved in this speech style, I include the constraint MAX-LowV in the tableau.

(34)   **MAX-LIQ**

Input liquid consonants are preserved in the output.

(35)   **MAX-NAS**

Input nasal consonants are preserved in the output.

51

(36)   **MAX-OBS**

Input obstruent consonants are preserved in the output.

(37) *Liquid consonant deletion triggered by PARSE-σ in the common speech style*

/pum-il-an-il-a/ 'hit for each other with'	MAX-LOWV	MAX-$[\sigma_1-\sigma_{penult}]$	PARSE-σ	MAX-LIQ
(a) [pu.m-i.l-a.('n-ii).l-a			8!	
(b) [pu.m-i.-ʲa.('n-ii).-ʲa		*!	6	**
☞ (c) [pu.m-i.l-a.('n-ii).-ʲa			7	*

All the candidates satisfy the constraint MAX-LOWV. However, candidate (37a) is disqualified because it has a maximal number of violators of PARSE-σ. Candidate (37b) is ruled out since it violates a high-ranking constraint MAX-$[\sigma_1-\sigma_{penult}]$ which requires segments of the strong position $[\sigma_1-\sigma_{penult}]$ to be preserved in outputs. The last candidate (37c) succeeds since it satisfies all the high ranking constraints and it minimally violates PARSE-σ as compared to candidate (37a). It violates a low-ranking constraint MAX-LIQ which militates against deletion of liquid consonants. This violation, however, can be tolerated.

3.3.2.3 Elderly speech styles

In common elderly speech styles, the observation was that liquid consonants are deleted, except when they are onsets to the first two syllables and the (bimoraic) penultimate syllable of the Morphological Stem. Nasal and obstruent consonants are never deleted anywhere. Two generalizations can be made out of these observations: Outputs which preserve obstruent and nasal consonants are optimal; and outputs which delete liquid consonants are optimal. Preservation of obstruent and nasal consonants would indicate that PARSE-σ ranks below the constraints MAX-OBS and MAX-NAS. The constraint ranking PARSE-σ >> MAX-LIQ, however, would ensure that liquid consonants are deleted.

The observation that liquid consonants belonging to the first two syllables and the bimoraic penultimate syllable of the Morphological Stem are preserved would be an indication that the

first two syllables and the bimoraic penultimate syllable of the Morphological Stem are parsed by moraic Feet. This fact, once again, is reminiscent of the theory of Positional Faithfulness. According to this theory, the ranking of a positional faithfulness constraint above a markedness constraint would precisely ensure that the contrast is maintained in the strong positions. In our case, contrast is preserved in the strong position [Foot]. The suggested analysis of liquid consonant deletion in the common speech style would be that the markedness constraint PARSE-σ succeeds in neutralizing weaker positions, but, obviously, it cannot have any effect on the strong stress position [Foot]. A positional faithfulness constraint MAX-FOOT, requiring input liquid consonants of the strong position [Foot] to be preserved in the output can be said to be responsible for the generalization that forms which preserve liquids in the first two syllables and the penultimate syllable are optimal. The fact that input liquid consonants of the strong position [Foot] are preserved in the output indicates that PARSE-σ ranks below the positional faithfulness constraint MAX-[FOOT]. A tableau in (38) below illustrates.

(38) *MAX-FOOT ranks above PARSE-σ in common elderly speech style*

/gululul-il-a/ 'cut grass for'	MAX-FOOT	PARSE-σ	MAX-LIQ	MAX-HV
(a) [(gu.lu).lu.('l-ii).l-a		****!		
(b) [(gu.ʷu).<'>.(-'ʷii).-ʲa	*!*	*	****	*
☞ (c) [(gu.lu).<'>.('l-ii).-ʲa		*	**	*

Thus, candidate (38a) is disqualified because it has the largest number of violators of the constraint PARSE-σ. Candidate (38b) is non-optimal because it violates a high-ranking positional faithfulness constraint MAX-[FOOT] which requires input segments of the strong position [Foot] to be preserved in the output. It violates this constraint twice, but an initial violation is enough to book the candidate. Candidate (38c) is optimal because it satisfies the constraint MAX-[FOOT] and it minimally violates PARSE-σ.

Another observation was that in rare cases some elderly speakers may also delete nasal consonant onsets over and above liquid deletion. Obstruent consonants are preserved. One

generalization falling out of these observations is that forms with liquid and nasal consonants in final syllable positions of outputs are not optimal. This would be an indication that the constraint PARSE-σ now ranks above both MAX-LIQ and MAX-NAS in this speech style. On the other hand, preservation of obstruent consonants is an indication that the constraint PARSE-σ continues to rank below MAX-OBS.

3.3.3 Problems with the stress analysis above

The most serious problem for the stress analysis suggested above is that the domain for penultimate vowel lengthening (or right stress foot) in ciTonga is not the Prosodic Word. Assuming that vowel lengthening is indeed a cue for stress, the domain of stress in ciTonga must be the Phonological Phrase. This is because lengthening usually shifts to the penultimate syllable of the rightmost Word in a Phonological Phrase as given in (39) below.

(39) *PL in ciTonga is phrase-penult*

VP[le.ˈ**lee**.s-a]   look

VP[le.le.s-a   vi-.ˈ**βaa**.ntʰo]   look at bad people

VP[zu.ˈ**mbuu**.w-a]   reveal

VP[zu.mbu.w-a ma.ˈ**jee**.so]   reveal an exam

VP[sa.ˈ**mbii**.z-a]   teach

VP[sa.mbi.z-a   mu.ntʰi.ˈ**kaa**.ze]   teach a woman

VP[βe.ˈ**lee**.nga]   count

VP[βe.le.nga   a.ntʰu.ˈ**luu**.me]   count men

As the examples above show, Prosodic Words are phonetically realized differently in different contexts (e.g. phrase medial vs. phrase final positions). In phrase-final position, Prosodic Words which I suspect are phrase-heads have penultimate lengthening. In phrase-medial positions, we have Prosodic Words with no penultimate lengthening or any other cue for stress[20]. The question remains then: What is the real prosody of the Prosodic Word in ciTonga and Bantu languages?

---

[20] One suggestion would be that vowel lengthening is reserved for the strong foot and that the right foot is the head (strong foot) of the relevant domain. However, vowel lengthening in ciTonga and many Bantu languages is

A second problem concerns the phonetic correlate for the Stem-initial Foot. There is no vowel lengthening whatsoever. One would suggest that preservation of vowel and liquid segments is a cue for stress. This would mean that every syllable with a preserved vowel or consonant (e.g. low vowels in almost all speech styles) bears stress. Such a conclusion would be absurd because preservation of low vowels and obstruent and nasal consonants in the examples from the elderly speech styles given in (40) below have nothing to do with metrical Feet.

(40) *Stress theory that solely relies on segment preservation lacks merit*

Verb word	*English gloss*
ti-.[(βe.le.)β.t-.s-a.(n-ee.)ŋg we-speak-rec-hab	we should be causing each other to speak
[(pa.la).pa.(s-aa).n-a cuddle-rec-fv	cuddle each other
[(βe.le.)β.t-.s-a.(n-aa).ŋg-a speak-caus-rec-hab-fv	be causing each other to speak

Expansion or reduction of the bracket for preserved vowels and liquids would also present another challenge for any stress analysis. For instance, the bracket for preserved vowels in the elderly speech community can expand to include three or more syllables on either side of the scale. If preservation of vowels were a phonetic correlate, we would have swinging Feet, expanding or reducing as given in (41) below.

a postlexical (phrasal) phenomenon and it is something beyond the scope of the present study which is concerned with the real prosody of the Word (with or without PL).

(41) *Speakers can expand the bracket of preserved vowels*

Formal speech   *Mixed speech style*

ti-[to.nde.k-e.s-a.n-eé.ŋg-e   **ti-[(to.nde).k-.(s-a.n-eé.ŋge)**
we-fail-caus-rec-hab-subj   "we should be causing each other to fail"

ti-[to.nde.k-e.s-a.n-eé.ŋg-e   **ti-[(to.nde.ke)-.s-.(n-eé).ŋg**
we-fail-caus-rec-hab-subj   "we should be causing each other to fail"

[βe.le.βe.t-e.s-a.n-aa.ŋg-a   **[(βe.le).β.t-.(sa-.n-aaŋga)**
speak-caus-rec-hab-fv   "be causing each other to speak"

[βe.le.βe.t-e.s-a.n-aa.ŋg-a   **[(βe.le.βe).t-.(sa-.n-aaŋga)**
speak-caus-rec-hab-fv   "be causing each other to speak"

[βe.le.βe.t-e.s-a.n-aa.ŋg-a   **[(βe.le.βe.te)-.s-.(n-aaŋga)**
speak-caus-rec-hab-fv   "be causing each other to speak"

As if this is not enough, speakers also have a tendency to pronounce mid and low vowels as high vowels. Low vowels can also be realized as mid vowels. These facts are illustrated in (42) below.

(42)   *Change of vowel quality in 'non-foot' positions*

Formal speech   *Elderly speech style*

ti-[βe.le.βe.t-e.s-a.n-ee.ŋg-e   **ti-[(βe.le).βi.ti-.si-.(n-ee).ŋgi**
speak-caus-rec-hab-fv   "we should be causing each other to speak"

ti-[to.nde.k-e.s-a.n-ee.ŋg-e   **ti-[(to.nde).ki-.si-.(n-ee).ŋgi**
fail-caus-rec-hab-fv   "we should be causing each other to fail"

ti-[tʰa.mba.l-i.s-a.n-ee.ŋg-e   **ti-[(tʰa.mba).li-.se-.(n-ee).ŋgi**
stretch-caus-rec-hab-fv   "we should be stretching legs over one another"

[βe.le.βe.t-e.s-a.n-aa.ŋg-a   **[(βe.le.)βi.ti-.se-.(n-aa.ŋga)**
speak-caus-rec-hab-fv   "be causing each other to speak"

The realization of mid and low vowels as high vowels or low vowels as mid vowels is well reported for stress languages. Vowels can become more sonorous in foot heads, while in foot non-heads and unstressed syllables they typically become less sonorous. In Chomoro, for instance, high vowels become mid vowels in stressed syllables (Crosswhite 1998 et eq; cited in de Lacy 2007:301). In the ciTonga data, however, the impression we get is that anything can happen to unfooted segments: optional preservation, deletion or change into another vowel quality, etc., as long as it does not involve strong segments or strong constituents. There is need for a more general principle which preserves all strong positions and at the same time allow anything to happen to weaker positions (except making them strong).

The other problem concerns unknown categories. It is easy to say that the first two syllables and the bimoraic penultimate syllables of the Prosodic Stem are parsed by Feet. This fact is backed by evidence: they preserve vowels and liquid consonants which are deleted outside these domains. The problem is that the Prosodic Hierarchy recognizes the Prosodic Word, Foot, Syllable and the Mora. In the absence of an intermediate category between Prosodic Word and Foot, we cannot account for the constituent $[\sigma_1\text{-}\sigma_{penult}]$ which equally preserves pre-final vowels and liquid consonants in the common speech styles. Is this also a Foot structure? Only a theory that recognizes a Prosodic Stem category, as Downing's (2006) Morpheme-Based Templates Theory does, would adequately account for it.

Lack of generalizations is another problem in the stress analysis given above. Edgemost syllables, most sonorous vowels, and least sonorous consonants are much more faithful than non-edge segments, least sonorous vowels and more sonorous consonants. Segments belonging to the Foot structure appear to be more faithful than those belonging to higher levels in the Prosodic Hierarchy. A missing generalization therefore is the fact that what is preserved is a strong constituent as defined by universal principles of Edgeness, Sonority and Prosodic Hierarchy. This obviates another challenge: How do we best represent Sonority, Edgeness and Prosodic Hierarchy in our grammar?

A final problem concerns non-universality of some OT constraints. It appears each speech group has its own sense of what is minimally strong and not strong. For instance, the Foot structure and obstruent consonants are undoubtedly strong constituents in the elderly speech styles, while a totally different constituent $[\sigma_1\text{-}\sigma_{penult}]$ as well as low vowels and nasal and obstruent consonants are strong constituents in the formal and common speech styles. Do we still need constraints such as MAX-OBS, MAX-LIQ, MAX-FOOT, or MAX-$[\sigma_1\text{-}\sigma_{penult}]$, MAX-

LOWV when all a speech community or group requires are universal principles such as Sonority, Edgeness and Prosodic Hierarchy to guide them on what ought to be strong or weak?

Summary

This section has attempted to present a stress analysis of vowel and consonant deletion in ciTonga. In the formal speech style, no vowel or consonant was deleted. The analysis has been that the constraint PARSE-σ is most likely outranked by MAX-IO.

Crucial observations in the common speech style included the following: Final non-low vowels and liquid onsets are deleted while low vowels, obstruents and nasals are preserved in the same position. In addition, non-low vowels and liquid consonants of the constituent [σ_1-σ_{penult}] are duly preserved. The analysis has been that a moraic foot falls on the penultimate syllable. The ranking of the constraint PARSE-σ above MAX-MIDV, MAX-HV and MAX-LIQ would result in deletion of non-low vowels and liquid consonants. The constraint PARSE-σ is satisfied or minimally violated if syllables outside the foot structure are deleted or minimized and vowel deletion is a step towards this direction. Preservation of low vowels and obstruent and nasal consonants in similar positions where non-low vowels and liquid consonants get deleted would indicate that the constraint PARSE-σ is outranked by MAX-LOWV, MAX-OBS and MAX-NAS. Preservation of non-low vowels and liquid consonants belonging to the strong position [σ_1-σ_{penult}] has been attributed to a positional faithfulness constraint MAX-[σ_1-σ_{penult}], which requires segments of the strong position [σ_1-σ_{penult}] to be preserved, outranking PARSE-σ.

One observation in the common elderly speech style was that the only vowels and liquid consonants that are preserved are those which are nuclei and onsets to the first two syllables and the bimoraic penultimate syllables. Low vowels and obstruent and nasal consonants are preserved as well. The analysis suggested has been that the first two moras of the Morphological Stem and the bimoraic penultimate syllable are parsed by Feet. Preservation of non-low vowels and liquids in these positions therefore would straightforwardly follow from a positional faithfulness constraint MAX-FOOT which requires segments of strong positions such as feet to be preserved. Preservation of low vowels, obstruents and nasals, once again, is attributed to ranking of MAX-LOWV, MAX-OBS and MAX-NAS above PARSE-σ. In rare elderly speech styles, however, low vowels not belonging to feet may be deleted as well, a sign that the constraints MAX-LOWV and PARSE-σ are re-ranked. Deletion of nasal

consonants in some rare elderly speech styles would indicate that PARSE-σ can in rare cases outrank MAX-NAS.

The word-stress analysis presented above, however, also presents us with a lot of challenges. To begin with, it is based on a wrong understanding of Prosodic Word. The domain for penultimate vowel lengthening in many Bantu languages, including ciTonga, is the phonological phrase. This fact obviates another problem: What is the real prosody of Prosodic Words (with or without PL)? Preservation of vowels as a cue for stress has also been found to be problematic. In particular, the phonetic correlate for word-stress is not as obvious as that of the phonological phrase. Furthermore, there are certain unknown prosodic constituents such as [σ$_1$-σ$_{penult}$] which do not fit any of the constituents of the Prosodic Hierarchy consisting of PrWord, Foot or Syllable. Most importantly, the metrification presented above fails to capture a crucial generalization that what MAX constraints preserve are strong constituents (including segments) as determined by Universal Principles of Sonority, Edge Distance or Prosodic Hierarchy. How best can we represent these Universal Guidelines in our grammar? And do we still need OT constraints such as MAX-OBS, MAX-LIQ, MAX-FOOT, or MAX-[σ$_1$-σ$_{penult}$] in an accent grammar when all a speech community or group requires are universal principles to guide them on what is strong and be preserved, or weak and be deleted or further weakened?

In the following section, I attempt to account for the facts using another persuasive theory developed by Downing (2006b), the Morpheme-Based Generalized Templates Theory (MBT), which has **Prosodic Stem** as a core prosodic unit in many Bantu languages including ciTonga.

3.4 MBT analysis

In this section, I present a Morpheme-Based Templates Theory (MBT) analysis of vowel and consonant deletion patterns in ciTonga. The central claim of Downing's (2006b) theory is that "the basic morphology-prosody correlation is between a single morpheme and a single syllable". A branching requirement on morphological heads such as Stems accounts for the tendency of prosodic morphemes towards having binary minimality. MBT then replaces Foot binarity as a motivation for the tendency of, for instance, reduplicative Prosodic Stems and Minimal Prosodic Words towards having binary minimality in many languages.

These two claims result into two principles namely, the MORPHEME-SYLLABLE CORRELATION and the HEADS BRANCH. The former requires that each morpheme contains one syllable, while the latter ensures that lexical heads such as Stems branch prosodically.

Since Stems are constituents minimally consisting of two morphemes, it is expected of them to have two syllables, one for each morpheme (by MORPHEME-SYLLABLE CORRELATION). A constraint born out of this is called **PROSODICSTEM** (Prosodic Stems are minimally binary). Since Affixes are non-heads, they are not bound to be branching.

The re-defined principle of **BINARITY** explains the binarity maximality which is typical of reduced prosodic morphemes such as reduplicants. Following works such as Harris (1994), Halle and Vergnaud (1987), and Ussishkin (2000:53), Downing proposes that the relevant adjacent elements for defining binarity are a constituent daughter and a constituent edge[21]. BINARITY is thus defined as in (44) below.

(43)   **PROSODICSTEM**

Prosodic Stems are minimally binary (Downing 2006b).

(44)   **BINARITY** (Downing 2006b:125)

Each daughter of a constituent must be adjacent to some edge of the constituent.

Downing (2006b) further argues that the markedness constraint BINARITY is outranked by FAITHFULNESS constraints in the regular vocabulary, as it appears to be rare for languages to require all Prosodic Stems to contain exactly two moras or two syllables. The unmarked status of binary constituents emerges only in certain morphological constructions such as reduplicants and this is attributed to TETU (the emergence of the unmarked) constraint ranking.

In the following sections I attempt to use Downing's Prosodic Stem Theory and I portray the fact that even normal Prosodic Stems can actively strive towards Stem Binarity. The analysis of vowel and consonant deletion within Prosodic Stems, therefore, is that this process is triggered by the need to satisfy the constraint BINARITY, requiring each daughter (syllable) of a Prosodic Stem to be adjacent to some edge of the constituent Prosodic Stem. As usual, I deal with the patterns of vowel deletion and preservation in the three speech styles namely, formal, common, and elderly speech styles.

---

[21] In Prosodic Hierarchy-Based Templates Theory, the binary maximality constraint is a result of the Prosodic Word-Stress Foot correlation: Stress Feet are minimally and maximally bimoraic or disyllabic by BINARITY.

3.4.1 Vowel deletion and preservation

3.4.1.1 Formal speech style

In the formal speech style, we had one observation regarding vowel deletion: None of the vowels is deleted. The analysis of this fact would be that outputs where all vowels of the Morphological Stem are preserved are optimal. In terms of OT, the suggestion is that MAX-IO ranks above BINARITY. A tableau in (45) below illustrates. Prosodic Stems are enclosed in curly brackets '{}'. Each segment of surplus syllables incurs one violation mark.

(45) *All segments are preserved in the formal speech style*

/ti-[tondek-es-ane-ŋg-e/ 'we should be causing each other to fail'	MAX-IO	BINARITY
☞ (a) ti-{to.nde.k-e.s-a.n-eé.ŋg-e}		8
(b) ti-{to.nde.k-e.s-a.n-eé.ŋg}	*!	7

Although candidate (45a) has the largest number of violators of BINARITY, it is still optimal because it satisfies the high-ranking constraint MAX-IO which requires input segments to be preserved in the output. Candidate (45b) is ruled out for violating MAX-IO.

3.4.1.2 Common speech style

One observation in the common speech style was that non-low vowels of final syllables are usually deleted while low vowels are preserved in the same position. In addition, non-low vowels of the position between the initial and the penultimate syllables ([σ_1-σ_{penult}]) of the Prosodic Stem are also preserved. The generalization here is that outputs which delete non-low vowels of the final syllable, but preserve both low and non-low vowels belonging to the position [σ_1-σ_{penult}], are optimal. Assuming the cause of vowel deletion to be the need to achieve a Prosodic Stem which has a maximum and minimum of two syllables or moras, this generalization can be accounted for by the constraint BINARITY, requiring each daughter of a Prosodic Stem to be adjacent to some edge of the constituent. Deletion of mid and high vowels therefore would straightforwardly follow from ranking BINARITY above MAX constraints (MAX-MIDV and MAX-HV). On the other hand, preservation of low vowels can be attributed to the constraint MAX-LOWV outranking BINARITY. The position [σ_1-σ_{penult}] is most likely a Prosodic Stem in which the final syllable (σ_n) is for some reason (to be discussed later) rendered extrametrical. Thus, preservation of non-low vowels in the position

[σ_1-σ_{penult}] can be linked to a positional faithfulness constraint **MAX-PROSODIC STEM** outranking BINARITY. A tableau in (47) below illustrates.

(46)   **MAX-PROSODIC STEM**

Segments of the strong position Prosodic Stem are preserved.

(47) *MBT analysis of mid vowel deletion in the common speech style*

/ti-[βeleβet-es-an-eŋg-e/ 'we should be causing each other to speak'	MAX-C	MAX-LOWV	MAX-PRSTEM	BINARITY	MAX-MIDV
(a) ti-{βe.le.βe.t-e.s-a.n-ee.}ŋge				10!	
(b) ti-{βe.le.βe.t-e.s-a.nee}	*!			8	*
(c) ti-{βe.l.β.t-.s-a.nee.}ŋg			*!**	6	****
☞ (d) ti-{βe.le.βe.t-e.s-a.nee.}ŋg				9	*

All the candidates satisfy the constraint MAX-LOWV. However, candidate (47a) is disqualified because it has a maximal number of violators of BINARITY. Candidate (47b) is ruled out because it violates a high-ranking constraint MAX-C (for argument's sake) which militates against deletion of consonant segments. Candidate (47c) violates another high-ranking constraint MAX-PRSTEM which requires vowel segments of the strong position Prosodic Stem to be preserved in outputs. The last candidate (47d) succeeds since it satisfies all the high ranking constraints and it minimally violates BINARITY as compared to candidate (47a). Explanation, however, is required as to why the Prosodic Stem includes the final syllable in the formal speech style while it avoids the final syllable in the common speech style especially when the final vowel is non-low. I will argue in section 3.6 that the final syllable is never extrametrical when the final vowel is a low vowel (Failure by final low vowel to undergo deletion even in the rare speech styles where pre-penult low vowels can be deleted makes me think that final syllables which have low vowel nuclei are parsed by Prosodic Stem).

3.4.1.3 Elderly speech styles
The observation in the common elderly speech style was that the only non-low vowels that are preserved are those which are nuclei of the first two syllables and the bimoraic penultimate

syllable. Low vowels are also not deleted anywhere. Three generalizations can be made from these observations: (a) input vowels of $[\sigma_1\text{-}\sigma_2]$ are preserved in the output; (b) input vowels of σ_{penult} are preserved in the output; and finally, (c) input low vowels are preserved in the output. Generalizations (a&b) are reminiscent of positional faithfulness theory (Trubetskoy 1939; Selkirk 1994; Beckman 1995, 1998). According to this theory, the ranking of a positional faithfulness constraint above a markedness constraint would precisely ensure that the contrast is maintained in the strong positions. Examples of strong positions where contrast is generally present in languages include stressed syllables (Trubetskoy 1939; Selkirk 1994; Alderete 1995), initial syllable (Trubetskoy 1939; Selkirk 1994; Beckman 1995, 1998), Root (McCarthy and Prince 1995; Casali 1996) and nouns (Smith 2001).

In our case, contrast is preserved in strong positions $[\sigma_1\text{-}\sigma_2]$ and $[\sigma_{penult}]$. The suggested analysis of vowel deletion in the common elderly speech style, therefore, is that the markedness constraint BINARITY succeeds in neutralizing weaker positions, but it fails to have any effect on the strong positions $[\sigma_1\text{-}\sigma_2]$ and $[\sigma_{penult}]$. Two positional faithfulness constraints MAX-$[\sigma_1\text{-}\sigma_2]$, requiring input vowels of the strong position $[\sigma_1\text{-}\sigma_2]$ to be preserved in the output, and MAX-$[\sigma_{penult}]$, requiring input vowels of the strong position $[\sigma_{penult}]$ to be preserved in the output, can be said to be responsible for generalizations (a) and (b). The fact that input vowels of strong positions $[\sigma_1\text{-}\sigma_2]$ and $[\sigma_{penult}]$ are preserved in the output indicates that BINARITY ranks below positional faithfulness constraints MAX-$[\sigma_1\text{-}\sigma_2]$ and MAX-$[\sigma_{penult}]$.

Generalization (c) above can be attributed to ranking of a general MAX constraint, MAX LOW VOWEL (MAX-LOWV), requiring input low vowels to be preserved in the output, above BINARITY. The constraints MAX-MID VOWEL (MAX-MIDV) and MAX-HIGH VOWEL (MAX-HV) would ensure that input mid and high vowels, respectively, are preserved in the output. But the fact that these vowels are easily eradicated in non-strong positions indicates that the constraint BINARITY ranks above MAX-MIDV and MAX-HV. A tableau in (48) below illustrates. I do not include the constraint MAX-C because it is not violated in all of the following tableaux of this section.

(48) *MBT analysis of vowel deletion in common elderly speech style*

MAX-[σ_1-σ_2], MAX-[σ_{penult}], MAX-LowV >> BINARITY >> MAX-MIDV

/ti-[βeleβet-es-an-eŋg-e/ 'we should be causing each other to speak'	MAX-LowV	MAX-[σ_1-σ_2]	MAX-[σ_{penult}]	BINARITY	MAX-MIDV
(a) ti-{βe.le.βe.t-e.s-a.n-ee.ŋg-e}				10!	
(b) ti-{βe.le.β.t-.s-.n-ee.ŋg}	*!			6	***
(c) ti-{βe.l.β.t-.sa-.n-ee.ŋg}		*!		6	****
(d) ti-{βe.le.β.t-.sa-.n-.ŋg}			*!	7	***
☞ (e) ti-{βe.le.β.t-.sa-.n-ee.ŋg}				7	***

Thus, candidate (48a) is disqualified because it has the largest number of violators of the markedness principle of BINARITY which requires each daughter (syllable) of a Prosodic Stem to be adjacent to some edge of the constituent Prosodic Stem. Candidates (48b-d) are ruled out for violating MAX-LowV, MAX-[σ_1-σ_2] and MAX-[σ_{penult}], respectively. The winning candidate (48e) satisfies all these constraints and it minimally violates BINARITY when compared with (48a).

One other crucial observation was that in rare cases most elderly speakers may delete low vowels as well. Low vowels belonging to the first and last two syllables are usually preserved. The analysis would be that keeping a Prosodic Stem at disyllabic size level is more important than preserving low vowels. Thus, the suggestion is that it is most likely in this speech style that MAX-LowV ranks below BINARITY. A tableau in (49) below illustrates.

(49) *Low vowels may be deleted in rare elderly speech styles*

(except those belonging to the first and last two syllables).

BINARITY >> MAX-LowV, MAX-MIDV

/ti-[to.nde.k-e.s-a.n-e.ŋg-e/ 'we will cause each other to fail'	BINARITY	MAX-LowV	MAX-MIDV
(a) ti-{to.nde.ke-.s-a.n-eé.ŋg-e}	8!		
(b) ti-{to.nde.k-.sa-.n-eé.ŋg}	6!		***
☞ (c) ti-{to.nde.k-.s-.n-eé.ŋg}	5	*	***

Thus, candidates (49a&b) are disqualified because they have a largest number of violators of the markedness constraint BINARITY which requires Prosodic Stems to be minimally and maximally binary. On the other hand, candidate (49c) is optimal because it violates BINARITY minimally. It violates low-ranking constraints MAX-LowV and MAX-MIDV, but these violations can be tolerated. The fact that low vowels belonging to the first and last two syllables are usually preserved would present a big challenge to MBT if Feet are not officially recognized by it. In other words, an adequate characterization of ciTonga grammar, it seems, would require both the **Prosodic Stem Theory** and a **Foot Theory** of some kind.

Finally, MBT can also account for the fact that many speakers fall in between the extremely formal speech style and extremely non-formal elderly speech styles. For instance, speakers in between may choose to expand the bracket of preserved non-low vowels to three or more syllables (e.g. {**βe.le.βe.t-.sa-.n-aaŋga**} 'be causing each other to speak'). MBT analysis of this observation would be that individual speakers are under pressure to minimize violation of general MAX constraints at the expense of BINARITY.

Elderly speakers may also reduce the bracket of preserved vowels to just initial and the penultimate syllable [e.g. formal **ti-[βe.le.βe.t-e.s-a.n-ee.ng-e ⇒ ti-{βe.l.β.t-.s-.n-ee.ŋg}**]. MBT would easily account for such problems by simply proposing that speakers are under pressure to reduce further the number of syllables. In this case, contrast is preserved in the strong positions [σ_1] and [σ_{penult}]. The suggested analysis of vowel deletion in this rare speech style, therefore, would be that the markedness constraint BINARITY succeeds in neutralizing

weaker positions, but it fails to have any effect on the strong positions $[\sigma_1]$ and $[\sigma_{penult}]$. Two positional faithfulness constraints MAX-$[\sigma_1]$, requiring input vowels of the strong position $[\sigma_1]$ to be preserved in the output, and MAX-$[\sigma_{penult}]$, requiring input vowels of the strong position $[\sigma_{penult}]$ to be preserved in the output, would be responsible for this status. The fact that input vowels of strong positions $[\sigma_1]$ and $[\sigma_{penult}]$ are preserved while the rest are falling indicates that BINARITY ranks below positional faithfulness constraints MAX-$[\sigma_1]$ and MAX-$[\sigma_{penult}]$. Thus, any syllable would survive, but survival of the initial and the penultimate syllables seem to be guaranteed by positional faithfulness constraints. A tableau in (50) below illustrates.

(50) *$[\sigma_1]$ and $[\sigma_{penult}]$ are the only strong positions in some rare elderly speech styles*

/ti-[to.nde.k-e.s-a.n-e.ŋg-e/	MAX-LowV	MAX-$[\sigma_1]$	MAX-$[\sigma_{penult}]$	BINARITY	MAX-MIDV
(a) ti-{to.nde.ke-.s-a.n-eé.ŋg-e}				8!	
(b) ti-{to.nd.k-.s-.n-eé.ŋg}	*!			4	***
(c) ti-{t.nd.k-.sa-.n-eé.ŋg}		*!		6	****
(d) ti-{to.nd.k-.sa-.n-.ŋg}			*!	4	****
☞ (e) ti-{to.nd.k-.s-a.n-eé.ŋg}				5	***

Candidate (50a) is disqualified because it has the largest number of violators of the principle of BINARITY which requires each daughter (syllable) of a Prosodic Stem to be adjacent to some edge of the constituent Prosodic Stem. Candidates (50b-d) are ruled out for violating MAX-LowV, MAX-$[\sigma_1]$ and MAX-$[\sigma_{penult}]$, respectively. The winning candidate (50e) satisfies all these constraints and it minimally violates BINARITY. I now turn to MBT analysis of consonant deletion and preservation.

3.4.2 Consonant deletion and preservation

3.4.2.1 Formal speech style

The observation in the formal speech style of ciTonga was that all consonants (and vowels of course) of a Morphological Stem are preserved in the more formal speech style. The generalization is that output forms which preserve all segments of a Morphological Stem are

optimal. Just like in our discussion of vowel deletion, this generalization may be linked to ranking of MAX-IO above BINARITY. In familiar terms, the constraint BINARITY ensures that a Prosodic Stem is minimally and maximally disyllabic or bimoraic. MAX-IO demands that input segments be preserved in the output. A tableau in (51) below illustrates.

(51) *All segments are preserved in the formal speech style*

/[bik-il-a/	MAX-IO	BINARITY
☞ (a) {bi.k-ii.l-a}		**
(b) {bi.k-ii.a}	*!	*******

Although candidate (51a) has the largest number of violators of BINARITY, it is still optimal candidate because it satisfies the high-ranking constraint MAX-IO which requires input segments to be preserved in the output. Candidate (51b) is ruled out for violating MAX-IO.

3.4.2.2 Common speech style

In the *common speech style*, the observation was that liquid consonant onsets to final syllables are usually deleted. Nasal and obstruent consonants are not deleted in the same position. In addition, liquid consonants of the position $[\sigma_1\text{-}\sigma_{penult}]$ are also preserved. The generalization here is that outputs which delete liquid consonants, but preserve liquid consonants and any other consonant of the position $[\sigma_1\text{-}\sigma_{penult}]$ are optimal. In MBT, the trigger of liquid consonant deletion would be the need to achieve a Prosodic Stem which must have a maximum and minimum of two syllables or moras. In terms of OT, the constraint BINARITY, requiring each daughter of a Prosodic Stem to be adjacent to some edge of the constituent, has been shown to be responsible for this status. Deletion of liquid consonants therefore would straightforwardly follow from ranking BINARITY above the constraint MAX-LIQ, requiring liquid consonants to be preserved in the output. On the other hand, preservation of obstruent and nasal consonants can be attributed to the constraints MAX-OBS and MAX-NAS outranking BINARITY. The position $[\sigma_1\text{-}\sigma_{penult}]$, once again, appears to be the Prosodic Stem and that the final syllable in this case is for some reason rendered extrametrical. Thus, preservation of liquid consonants in the position $[\sigma_1\text{-}\sigma_{penult}]$ can essentially be linked to the positional faithfulness constraint MAX-PRSTEM outranking BINARITY. A tableau in (52) below illustrates. Since low vowels are usually preserved in this speech style, I include the constraint MAX-LowV in the tableau.

(52) *Liquid consonant deletion triggered by BINARITY in the common speech style*

/[pum-il-an-il-a/ 'hit for each other with'	MAX-LOWV	MAX-PRSTEM	BINARITY	MAX-LIQ
(a) {pu.m-i.l-a.n-ii.l-a}			6!	
(b) {pu.m-i.-ʲa.n-ii.-ʲa}		*!	4	**
☞ (c) {pu.m-i.l-a.n-ii.-ʲa}			5	*

All the candidates satisfy the constraint MAX-LOWV. However, candidate (52a) is disqualified because it has a maximal number of violators of BINARITY. Candidate (52b) is ruled out since it violates a high-ranking constraint MAX-PRSTEM which requires segments of the strong position PRSTEM to be preserved in outputs. The last candidate (52c) succeeds since it satisfies all the high ranking constraints and it minimally violates BINARITY as compared to candidate (52a). It violates a low-ranking constraint MAX-LIQ which militates against deletion of liquid consonants. This violation, however, can be tolerated. However, explanation is required as to why the Prosodic Stem includes the final syllable in the formal speech style while it avoids the final syllable in the common speech style especially when the onset to the final syllable is a liquid. The final syllable is never extrametrical when the onset is a nasal or obstruent consonant (plus a final low vowel of course) as evidenced by their preservation in this position as well.

3.4.2.3 Elderly speech styles

In the common elderly speech style, the observation was that pre-final liquid consonants may be deleted as well, except when they are onsets to the first two syllables and the bimoraic penultimate syllable of the Morphological Stem. Nasal and obstruent consonants are never deleted anywhere. Two generalizations can be made out of these observations: Candidates which preserve obstruent and nasal consonants are optimal; and candidates which delete liquid consonants are optimal. Preservation of obstruent and nasal consonants indicates that BINARITY ranks below the constraints MAX-OBS and MAX-NAS. The constraint ranking BINARITY >> MAX-LIQ, however, would ensure that liquid consonants are deleted.

The observation that liquid consonants belonging to the first two syllables and the bimoraic penultimate syllable of the Prosodic Stem is, once again, reminiscent of the theory of

Positional Faithfulness. In our case, contrast is preserved in strong positions [σ_1-σ_2] and [σ_{penult}]. The suggested analysis of liquid consonant deletion in the common elderly speech style, therefore, is that the markedness constraint BINARITY succeeds in neutralizing weaker positions, but it fails to have any effect on the strong positions [σ_1-σ_2] and [σ_{penult}]. Two positional faithfulness constraints MAX-[σ_1-σ_2], requiring input liquid consonants of the strong position [σ_1-σ_2] to be preserved in the output, and MAX-[σ_{penult}], requiring input liquid consonants of the strong position [σ_{penult}] to be preserved in the output, can be said to be responsible for the generalization that forms which preserve liquids in the first two syllables and the penultimate syllable are optimal. The fact that input liquid consonants of strong positions [σ_1-σ_2] and [σ_{penult}] are preserved in the output indicates that BINARITY ranks below positional faithfulness constraints MAX-[σ_1-σ_2] and MAX-[σ_{penult}]. In other words, the constraint BINARITY cannot cause erasure of liquids belonging to strong positions MAX-[σ_1-σ_2] and MAX-[σ_{penult}]. A tableau in (53) below illustrates.

(53) *MBT analysis of liquid deletion patterns in common elderly speech styles*

MAX-[σ_1-σ_2], MAX-[σ_{penult}] >> BINARITY

/[gulumul-il-is-a/ 'cause to scrub with'	MAX-[σ_{penult}]	MAX-[σ_1-σ_2]	BINARITY	MAX-LIQ	MAX-HV
(a) {gu.lu.mu.l-i.l-ii.s-a}			8!		
(b) {gu.ʷu.mu.<'>.-ʷii.s-a}	*!	*	4	***	*
☞ (c) {gu.lu.mu.<'>.-lii.s-a}			6	*	***

Thus, candidate (53a) is disqualified because it has the largest number of violators of the constraint BINARITY. Candidate (53b) is non-optimal because it violates a high-ranking positional faithfulness constraint MAX-[σ_{penult}] which requires input segments of the strong position [σ_{penult}] to be preserved in the output. It also violates another high-ranking positional faithfulness constraint MAX-[σ_1-σ_2] which requires input segments of the strong position [σ_1-σ_2] to be preserved in the output. But a decision has already been made by MAX-[σ_{penult}]. Candidate (53c) is optimal because it satisfies the constraints MAX-[σ_{penult}] and MAX-[σ_1-σ_2] and it minimally violates BINARITY.

Another observation was that in rare cases some elderly speakers may also delete nasal consonant onsets over and above liquid deletion. Obstruent consonants are preserved. One generalization falling out of these observations is that forms with liquid and nasal consonants in outputs are not optimal in these styles. This would be an indication that the constraint BINARITY ranks above both MAX-LIQ and MAX-NAS. On the other hand, preservation of obstruent consonants is an indication that the constraint BINARITY continues to rank below MAX-OBS.

3.4.3   Problems with MBT analysis

One major contribution made by Downing's (2006b) MBT is expansion of the contents of morpho-prosodic constituents below the level of Prosodic Word to include at least Stem and Root. In fact, the need to recognize the Prosodic Stem as a distinct phonological domain from Prosodic Word has been argued for in numerous other Bantu languages (Inkelas 1989, 1993; Hyman 1993; Hyman and Mtenje 1999; Mchombo 1993; Myers 1987; and Mutaka 1994; cited in Downing 2006b). This formal recognition of the Prosodic Stem somehow simplifies the task at hand because most of the arguments I make revolve around the idea of Prosodic Stem. In fact, unlike in the stress **Foot Theory** analysis, the position **[σ_1-σ_{penult}]** of the Stem is dully recognized in MBT as a **Prosodic Stem**. One major problem with our stress analysis therefore was that it did not recognize the Prosodic Stem constituent as MBT does.

Ironically, MBT's first problem is its denial or downplaying of the constituent Foot as the Head of the Prosodic Stem in languages like ciTonga where Feet are not so obvious (Prosodic Stems are shown to dominate syllables and not Feet). This understanding of the structure of Prosodic Stems seems to be misguided and it is based on a not so correct assumption that accent is always signalled phonetically, such as by penultimate vowel lengthening in Bantu languages[22]. As we have seen in the analysis of vowel and consonant deletion, a serious challenge to MBT is to account for Prosodic Stems which contain strong positions such as [σ_{penult}] and [σ_1-σ_2] which are most likely accent Feet. Its core principles of PROSODICSTEM and BINARITY predict that a super Prosodic Stem will have two syllables or moras. MBT therefore cannot account for the fact that Prosodic Stems such as those in ciTonga actively preserve segments of a 'Foot' (in elderly speech styles) or/and Prosodic Stem (in formal and common speech styles).

---

[22] I will argue in the following sections that Accent Feet are phonetically realized only when they are a Minimal Strong Accent Constituent of a language system. In a system where the Prosodic Stem is a Minimal Strong Accent Constituent, the Accent Foot may be inert (especially in toneless words).

Furthermore, explanation is required as to why the Prosodic Stem includes the final syllable in the formal speech style while it sometimes avoids the final syllable in the common speech style. Even within the common speech style itself, final syllables whose onsets are liquid consonants or whose final vowels are non-low vowels are rendered extrametrical while they are never rendered extrametrical when their onsets are nasal or obstruent consonants and their final vowels are low vowels. The question is: what is it that determines metricality of the final syllable?

MBT, like many other theories of prominence, places itself among theories which refer to strong positions. For instance, Downing (2006b:121ff) argues that there are theoretical precedents which support the idea that phonological complexity correlates with head status. Some of the theoretical precedents she mentions are positional markedness (Beckman 1997, 1998; Harris 1990, 1994, 1997, 2004; Steriade 1994; and Barnes 2002), Head-Dependent Asymmetry (Dresher and van der Hulst's 1998) whose goal is to provide a general account of the correlation between prominent positions or heads and marked or complex structure. In addition, there are strong positions referred to by positional faithfulness theory such as stressed syllables (Trubetskoy 1939; Selkirk 1994; Alderete 1995), initial syllable (Trubetskoy 1939; Selkirk 1994; Beckman 1995, 1998), Root (McCarthy and Prince 1995; Casali 1996) and nouns (Smith 2001). The problem is that current constructions of both MBT and OT, to a greater extent, lack a universal principle that would put all the strong positions under one umbrella.

Another problem closely linked to the problem above is lack of generalization. In both MBT and stress analyses, MAX constraints preserve strong vowels and consonants in terms of Sonority (e.g. least sonorous obstruents and most sonorous low vowels), while positional faithfulness constraints also preserve strong positions STEM-σ_1, PENULTIMATE-σ or FINAL-σ, and PROSODIC STEM or ACCENT FOOT (as guided by Universal Principles of EDGENESS, FINALITY, SONORITY and PROSODIC HIERARCHY). The question is: Are general MAX constraints not doing the same thing as positional faithfulness constraints, preserving strong constituents[23]? A generalization of some sort is missing here.

Furthermore, it seems there is a problem of redundant constraints. It is apparent from our observations about vowel and consonant deletion that when given universal guidelines such as

---

[23] This may read as an attack on positional faithfulness theory. But I don't think so. All I am suggesting in the present study is that these strong positions are Strong Accent Constituents and they are the backbone of human language.

sonority and edge distance, individual speech groups or communities are able to determine what should constitute a minimal strong position. For example, it seems on the basis of sonority, the least sonorous obstruent consonants and most sonorous low vowels constitute stronger positions than the most sonorous liquids and least sonorous mid and high vowels. It seems the Prosodic Stem is a strong constituent in the formal and common speech styles and so segments belonging to it are preserved. In the elderly speech styles, on the other hand, the Foot structure seems to be stronger. Is it not redundant therefore to have non-informative constraints such as MAX-OBS, MAX-[σ_1], MAX-PRSTEM, MAX-FOOT in OT tableaux? Is not a set of Universal Guidelines such as EDGENESS, FINALITY, SONORITY and PROSODIC HIERARCHY, all that is needed for speech groups and communities to make their own choices (thresholds) of what is strong and not strong? A crucial question then would be: How best can we represent these and such other principles in our grammar?

Summary

In this section, I have attempted to account for optional vowel and consonant deletion in ciTonga in terms of MBT. The central claim is that vowel and consonant deletion are triggered by the need to satisfy the constraint BINARITY which requires Prosodic Stems to minimally and maximally contain two syllables or moras. Deletion of vowels and consonants would ensure that extra syllables are eradicated or minimized and that BINARITY is better satisfied or minimally violated. Due to FAITHFULNESS of consonants and some vowel segments, however, the constraint BINARITY is rarely satisfied.

In the formal speech style, for instance, none of the segments is deleted. The analysis there has been that the constraint MAX-IO which requires input segments to be preserved in the output completely outranks BINARITY.

One observation in the common speech style was that non-low vowels and liquid consonants of final syllables are usually deleted while low vowels and nasal and obstruent consonants are usually preserved in the same positions. In addition, non-low vowels and liquid consonants of the position PRSTEM are also usually preserved. The analysis of vowel and consonant deletion in this speech style has been that the constraint BINARITY manages to outrank the constraints MAX-MIDV, MAX-HV and MAX-LIQ, but it fails to outrank MAX-LOWV, MAX-NAS, MAX-OBS and a positional faithfulness constraint MAX-[PRSTEM].

In common elderly speech styles, one observation has been that non-low vowels and liquid consonants of the first two syllables and the bimoraic penultimate syllable are usually

preserved. Another observation is that low vowels and nasal and obstruent consonants are usually preserved too. In rare elderly speech styles, however, low vowels may be deleted as well, but not when they are nuclei to first two and last two syllables of the Prosodic Word/Stem. The analysis of vowel and consonant deletion and preservation patterns in the common elderly speech style has been that the constraint BINARITY is outranked by MAX-LowV, MAX-NAS, MAX-OBS and positional faithfulness constraints MAX-$[\sigma_1-\sigma_2]$ and MAX-$[\sigma_{penult}]$. In rare elderly speech styles, BINARITY manages to outrank MAX-LowV. In addition to multifarious vowel and liquid consonant deletion, some elderly speakers in *rare speech styles* may also delete nasal consonant onsets. Obstruent consonants are usually preserved. The analysis is that the constraint MAX-NAS in rare cases may rank below BINARITY while MAX-OBS continues to claim supremacy over BINARITY.

Although MBT and positional faithfulness theory exhibit some prowess in the analysis of vowel and consonant deletion and preservation patterns in ciTonga, the analyses have not been without challenges. MBT's position that the Prosodic Stem, especially with reference to Bantu languages, may contain nothing but syllables or moras, and not Feet or strong constituents, would lead to unknown categories such as $[\sigma_1-\sigma_2]$ and bimoraic $[\sigma_{penult}]$ which are most likely accent Feet of some kind in Bantu languages. Furthermore, explanation is required as to why the Prosodic Stem includes the final syllable in one context and excludes it in another context. MBT would also not be able to unite all the strong positions (segmental and non-segmental) in natural language. The analyses have also revealed lack of a generalization that what is preserved are strong constituents as guided by Universal Principles such as Sonority, Edgeness and Prosodic Hierarchy. The ensuing constraints therefore seem to be redundant since strong constituents to be preserved can easily be determined by following provisions of these Universal Guidelines. These problems, and perhaps many others omitted here, suggest that there must be a better way of accounting for the issues. In the following sections I will suggest that the most appropriate grammar of word prosody as presented in ciTonga is one based on **Universal Guidelines and Constraints for Strong Accent Constituents**.

3.5   Strong Accent Constituents Theory
In this section I argue and present the theory of **Strong Accent Constituency** within a larger framework of **Universal Guidelines and Constraints**. I will start once again by summarizing observations which were made regarding vowel and consonant deletion in ciTonga. In the formal speech style, all vowels of the Morphological Stem are preserved. In the common

speech style, non-low vowels of the final syllable are usually deleted. In common elderly speech styles, however, the only vowels that are preserved are those which are nuclei of the first two syllables and bimoraic penult syllables. Deletion of low vowels is not acceptable. In *rare elderly speech styles*, low vowels may be deleted as well.

Regarding consonant deletion, all consonants of the Morphological Stem are preserved in the more formal speech style. In the common speech style, however, liquid consonant onsets to final syllables are usually deleted. Nasal and obstruent consonants are never deleted anywhere. In the common elderly speech style, the only liquid consonants that are preserved are onsets of the first two syllables and the penultimate syllable of the Morphological Stem. Nasal and obstruent consonants are also never deleted anywhere. In addition to multifarious vowel and liquid consonant deletion, some elderly speakers in *rare speech styles* may also delete nasal consonant onsets. Obstruent consonants are usually preserved in the same positions. Many other speakers fall in between these extremes.

These observations bring to mind the question of the role of **sonority** and **edgeness** in the grammar of ciTonga. This question is not particularly new. The relevance of **sonority** in grammars of languages has been noted in a wide range of literature. Zec (2007), for instance, shows that there is a lot of literature written since 1881 which supports this fact especially with reference to syllabification. The influence of sonority on prosodic constituents higher than the syllable is also not a new thing (see de Lacy 2007).

According to Zec (2007), sonority of segments is represented by means of a scale such as one adapted in (54) below which corresponds to an ordering of segments ranging from those with the highest sonority (i.e. low vowels) to those with the lowest sonority (i.e. obstruents).

(54)   *Sonority scale (adapted from Zec 2007)*

Low vowels

Mid vowels

High vowels

Liquids

Nasals

Obstruents
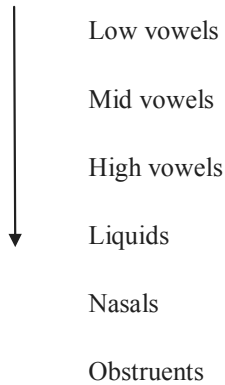
According to Zec (2007:178, and references cited therein), the most sonorous segments, such as vowels, occupy the nucleus position in syllables, while the less sonorous ones such as consonants occur towards the margins. The best margins are obstruents followed by nasals and liquids, with vowels being the worst margins. Thus, a combination of "OV is better than NV, which in turn is better than LV." Similarly, low vowels are best nucleus segments, followed by mid vowels and high vowels. Zec (2007) also observes that different languages have different sonority *thresholds* for the onset position and the nucleus position. For instance, some languages will allow nucleus segments to include nasal consonants while others will narrow the threshold to include only vowels.

Sonority has also been linked with higher prosodic structure. For instance, de Lacy (2007) reports that in Takia (spoken in North Guinea) sonority has influence on stress as follows:

(a) Stress the rightmost syllable with [a].
(b) Otherwise stress the rightmost syllable with [e] or [o].
(c) Otherwise stress the rightmost syllable.

Thus, like with other stress systems, stress in Takia is attracted by the edges. In a word where all of the vowels are of the same type, stress falls on the right edge (e.g. [ara'tam]). However, de Lacy (2007) argues that the most important factor for stress in Takia is sonority. The rule is that stress should be assigned to the most sonorous vowel available following the Takia sonority scale of |a > e, o > i, u|. According to de Lacy, the sonority requirements in Takia also override conditions on foot type. In words where vowels are the same (e.g. [ta'man]) the

foot type is iamb, while in words like ['abi] the foot type is trochee to ensure that the foot head is one with higher sonority.

De Lacy also reports that major theoretical approaches used to account for facts like these are either constraint-based or representational (Zec 2007). Zec indicates that most recent theories of sonority-driven stress (Kenstowicz 1997/2004, de Lacy 2002a, 2004, 2006) advocate the use of constraint interaction as a means of explanation. According to Zec (2007), the idea that constraint interaction can be used to account for sonority-driven stress is proposed in Kenstowicz (1997/2004) who advocates a fixed hierarchy of foot-head and non-head constraints. Kenstowicz's theory is said to relate directly to Prince and Smolensky's (2004) proposal about fixed ranking and the influence of sonority on syllable structure. The point I am trying to make is that there is already commitment to include the principle of Sonority in grammars of natural language.

Similarly, it is clear from our observations about ciTonga that there is a strong relationship between sonority and segment faithfulness[24]. The fixed rankings of faithfulness in (55) below would place obstruents as the least sonorous but most faithful consonants, while liquids are the most sonorous but the least faithful of the consonants. Of all the vowels, low vowels are the most sonorous and they are the most faithful too. On the other hand, high vowels are the least sonorous and least faithful of all the vowel types.

(55)   *Segment faithfulness hierarchy*

         (a) OBS > NAS > LIQ

         (b) LowV > MV > HV

In terms of Optimality Theory, the set of consonant faithfulness constraints with fixed ranking makes obstruent consonants the most faithful, and liquid consonants the least faithful. Similarly, a set of vowel faithfulness constraints with fixed ranking makes low vowels the most faithful and high vowels the least faithful. I illustrate these constraint rankings in (56) below.

---

[24] Beckman (1997) makes a similar observation.

(56) *Consonant and vowel faithfulness constraints with fixed ranking*

(a) MAX-OBS > MAX-NAS> MAX-LIQ

(b) MAX-LowV > MAX-MV > MAX-HV

Let me now turn to the idea of **Edgeness**. It plays a crucial role in the distribution of segmental and prosodic units in many languages. Prosodic units (such as feet or stress and tone) usually align with either the left or the right edge of morphological categories (Words, Stems, etc). In our example of Takia above (de Lacy 2007), edge-attraction is evident. In a word where all of the vowels are of the same type, stress falls on the right edge. Similarly, in ciTonga there is a strong relationship between Edgeness and segment faithfulness. Segments of edgemost syllables were the most likely to be preserved while segments of non-edge syllables were the most likely casualties. A set of faithfulness constraints with fixed ranking in (57) below would place segments of edgemost syllables as being more faithful than segments of syllables which stretch away from edges.

(57) MAX-EDGE-σ >> MAX-NONEDGE-σ

We can add the principle of **Prosodic Hierarchy** as one other idea that will determine the level of faithfulness of segments contained in a prosodic constituent. It seems, for instance, that segments belonging to Prosodic Stem and Foot structures will be more faithful than segments belonging to higher prosodic constituents. In terms of Optimality Theory, the set of positional faithfulness constraints with fixed ranking would make Accent Foot the most faithful, followed by Prosodic Stem and Prosodic Word. In cases where power relations between moras of a Foot is clear, segments of the Head Mora will be more faithful than segments of the Accent Foot due to high-ranking of MAX-HEAD MORA.

(58) MAX-HEAD MORA >> MAX-ACCENT FOOT >> MAX-PROSODIC STEM > MAX-PROSODIC WORD

A closer look at the positional faithfulness constraints and the general faithfulness constraints presented throughout this discussion indicates that they both preserve strong positions (e.g. edgemost syllables, most sonorous vowels, least sonorous consonants, lower prosodic constituents, etc). The strength of each position or constituent has been based on the ideas of Edgeness, Sonority, Prosodic Hierarchy, etc. A crucial point I want to make is that every prosodic unit (Phonological Phrase, Prosodic Word, Prosodic Stem, Foot, syllable, mora,

Segment, Feature or any smallest atom of human speech) is an **Accent Constituent** (unit of prominence) and that what is preserved, especially in relation to Accent, is a **Strong Accent Constituent** of the relevant domain (Language, Utterance, Phrase, Word, Stem, or Accent Foot). What is Strong Accent Constituent in one context may not necessarily be Strong AC in another context. In relation to prominence, both the traditional general MAX and positional faithfulness theories as presented above therefore miss a generalization that what are preserved are language- or context-specific Strong Accent Constituents. At the heart of Accent Constituents Theory, therefore, is the fact that a finite set of Universal Guidelines such as **Edgeness**, **Finality**, **Sonority**, **Prosodic Hierarchy**, **Stress**, **Syllable Weight**, **Pitch**, and several others play a crucial role in determining what is strong and not strong. The Guidelines have clauses which will guide speakers as to what is strong or weak as given in (59-68) below. What counts as Determiner Guideline (DG) of strength in one language may not necessarily be so in another language. For instance, heavy syllables might be Strong ACs in one language but ordinary ACs in another language. Another crucial point is that Guidelines may jointly make a determination, relegating some Guideline in the process.

(59)   **EDGENESS**

Edgemost syllables are stronger than non-edge syllables.

(60)   **RIGHTEDGE**

The right edge is stronger than the left edge

(61)   **LEFTEDGE**

The left edge is stronger than the right edge

(62)   **FINALITY**

Non-final syllables are stronger than final syllables

(63)   **SONORITY**

Least sonorous consonants and most sonorous vowels are stronger than most sonorous consonants and least sonorous vowels.

(64)   **PROSODIC HIERARCHY**

Lower level prosodic constituents are stronger than higher level prosodic constituents.

78

(65) **STRESS**

Stressed syllables are stronger than unstressed syllables.

(66) **SYLLABLE WEIGHT**

Heavy syllables are stronger than light syllables.

(67) **VOWEL LENGTH**

Long vowels are stronger than shorter vowels.

(68) **PITCH**

Units with high pitch are stronger than units with low pitch.

Since the cut-off point of what is strong and/or weak is left to individual languages and dialects themselves, arbitrary and redundant constraints such as FAITH-LIQ and MAX-OBS are relegated as universal constraints. And since language- or context-specific Strong Accent Constituents are the ones which usually survive deletion or alternation, it appears that there is only one **FAITHFULNESS** constraint which is responsible for this status: **FAITH - STRONG ACCENT CONSTITUENT** (**FAITH-SAC**). What happens to weaker ACs, it seems, is of little or none of the accent grammar's business.

(69) **FAITH-STRONG ACCENT CONSTITUENT**[25]

Strong Accent Constituents are preserved.

Finally, locating of a right edge Accent Foot in languages which have the property of phonological phrasing (phrase-penult vowel lengthening) can be a tricky thing. In languages like ciTonga, it is a bit easier to tell whether the Foot is on the bimoraic penultimate syllable or it is on the last two moras of the Prosodic Stem because of segment distribution. The Foot is on the bimoraic penultimate syllable when the final syllable is poor in terms of sonority (in which case the Guidelines followed are of EDGENESS and FINALITY), and it is on the last two moras when all segments of the final syllable are preserved or when the final syllable is satisfactory in terms of sonority (in which case the Guidelines of EDGENESS and SONORITY are followed). A crucial observation is that Accent Feet are assigned to Strong Accent

---

[25] The constraint may also apply to blocking of the insertion of Strong Accent Constituents. In this case a most suitable definition would have to be provided.

Constituents such as PROSODIC STEM, STEM-σ_1, and PENULTIMATE-σ or FINAL-σ as may be guided by Universal Guidelines of PROSODIC HIERARCHY, EDGENESS and FINALITY or SONORITY. The generalization therefore is that Accent Feet belong to Strong Accent Constituents. A constraint I suggest to be responsible for this status is **ACCENT FOOT/STRONG ACCENT CONSTITUENT (ACFT/SAC)**.

(70)   **ACFT/SAC**

   Accent Feet belong to Strong Accent Constituents

Summary

Most observations made about vowel and consonant deletion in ciTonga point to the fact that Universal Principles such as Sonority, Edgeness, Finality and Prosodic Hierarchy play a crucial role in the grammars of natural language. For instance, it is clear that there is strong relationship between segment faithfulness and segment sonority, edgeness and a level in which it appears in the Prosodic Hierarchy. Literature on prosody contains many other examples. What is at issue is how best we can represent these Universal Principles or Guidelines in our grammars. I have argued in this section that every prosodic unit (Phonological Phrase, Prosodic Word, Prosodic Stem, Foot, syllable, mora, segment, feature and any smallest atom of human speech) is an **Accent Constituent** (unit of prominence) and that what is preserved, especially in relation to Accent, is a **Strong Accent Constituent** of the relevant domain (Language, Utterance, Phrase, Word, Stem, or Accent Foot). What is Strong Accent Constituent in one speech community or context may not necessarily be strong in another context. At the heart of Accent Constituent Theory, therefore, is the fact that a finite set of Universal Guidelines such as **Edgeness**, **Sonority**, **Finality**, **Prosodic Hierarchy** and several others play a crucial role in determining what is strong and what is not. What counts as Determiner Guidelines (DG) of strength in one language may not necessarily be so in another language. The Guidelines may jointly make a determination, relegating some Guideline in the process. Since language- or context-specific Strong Accent Constituents are the ones which are usually preserved, it appears that there is only one **FAITHFULNESS** constraint which is responsible for this status: **FAITH - STRONG ACCENT CONSTITUENT (FAITH-SAC)**. It appears that what happens to weaker ACs is of little or none of the accent grammar's business. Finally, another constraint **ACFT/SAC** would ensure that Accent Feet are assigned to Strong Accent Constituents such as PROSODIC STEM, STEM-σ_1, and PENULTIMATE-σ or FINAL-σ as may be guided by Universal Guidelines of PROSODIC

HIERARCHY, EDGENESS and FINALITY or SONORITY. In the following sections, I present Strong Accent Constituent Theory analyses of vowel and consonant deletion processes in the various speech styles of ciTonga.

3.6 Strong Accent Constituent Analysis

3.6.1 Vowel deletion and preservation

3.6.1.1 Formal speech style

In the formal speech style, the observation was that all vowels of the Morphological Stem are preserved. Since the final vowel of the Morphological Stem is preserved as well, it seems, Strong Accent Constituents in this speech style include the PROSODIC STEM, STEM-σ_1 and FINAL-σ. Determiner Guidelines are PROSODIC HIERARCHY and EDGENESS. These two Guidelines provide that lower prosodic units such as the Prosodic Stem and edgemost syllables are Stronger Accent Constituents. Feet are required to be binary and due to this requirement an adjacent mora is included in the abstract foot structure. This foot analysis is illustrated in (71) below.

(71)   *Strong Accent Constituents such as PROSODIC STEM, STEM-σ_1 and FINAL-σ bear Accent Foot*

Formal speech   *English gloss*

ti-{(βe.le).βe.t-e.s-a.n-e(e.ŋg-e)}   we should be causing each other to speak
we-speak-caus-rec-hab-subj

ti-{(ŋa.na).mu.l-i.s-a.n-e(e.ŋg-e)}   we should be causing each other to turn
we-turn-caus-rec-hab-subj

ti-{to.nde).k-e.s-a.n-e(é.ŋg-e)}   we should be causing each other to fail
we-fail-caus-rec-hab-subj

ndi-{(go.go).te.z-e(e.ŋg-e)}   I should be exaggerating
I-exaggerate-hab-subj

ti-{(tʰa.mba).l-i.s-a.n-e(e.ŋg-e)}   we should be stretching legs over one another
we-stretch-caus-rec-hab-subj

Generalizations to be accounted for are that Accent Feet belong to Strong Accent Constituents and that the Feet are binary at the level of the mora. In terms of OT, the constraint ACFT/SAC and FOOT BINARITY (FTBIN) would ensure these statuses. The constraint ACFT/SAC ensures that Accent Feet fall on Strong Accent Constituents STEM-σ_1 and FINAL-σ of the Strong Accent Constituent PROSODIC STEM. The constraint FTBIN ensures that the Accent Feet are binary at the level of the mora. A tableau in (72) below illustrates. Once again, I use curly brackets '{}' for Prosodic Stem and braces '()' for Accent Foot.

(72) *EDGENESS as Determiner Guideline for right edge Foot*

SACs: PROSODIC STEM; STEM-σ_1; FINAL-σ

DGs: PROSODIC HIERARCHY; EDGENESS; EDGENESS.

/ti-[βeleβet-es-an-eŋg-e/ 'we should be causing each other to speak'	ACFT/SAC	FTBIN	FAITH-SAC
(a) ti-{(βe.le).βe.t-e.s-a.(n-ee).ŋge}	*!		
(b) ti-{(βe.)le.βe.t-e.s-a.n-ee.(ŋge)}		*!*	
☞ (c) ti-{(βe.le).βe.t-e.s-a.n-e(e.ŋge)}			

Thus, candidate (72a) is non-optimal because it violates a high-ranking constraint ACFT/SAC which requires Accent Feet to be assigned to Strong Accent Constituents such as edgemost syllables of a Prosodic Stem as provided by the universal Guidelines of PROSODIC HIERARCHY and EDGENESS. Candidate (72b) satisfies ACFT/SAC, but it is ruled out because it violates another constraint FTBIN which requires Accent Feet to be binary at the level of the mora. On the other hand, candidate (72c) wins because it satisfies both ACFT/SAC and FTBIN. None of the candidates violate FAITH-SAC.

3.6.1.2 Common speech style
One observation in the common speech style was that non-low vowels of final syllables are usually deleted. My suggested analysis of this fact is that the final syllable has lost the strong position status to the penultimate syllable and the right Accent Foot in this speech style falls on the penultimate syllable as given in (73) below. In other words, the Guideline EDGENESS might play a role, but it is not enough to determine what will be a Strong Accent Constituent

in this speech style. A SAC must also be guided by another principle of FINALITY, which provides that non-final syllables are stronger than final syllables, before we can consider issues of EDGENESS. An inventory of Strong Accent Constituents in this speech style and context therefore must include PROSODIC STEM, STEM-σ_1 and PENULTIMATE-σ. This status (SAC) qualifies the Prosodic Stem, stem-initial syllables and penultimate syllables to bear Accent Feet as given in (73) below. The final syllable is most likely extrametrical (not parsed by Prosodic Stem or Accent Foot) due to poor quality of its vowel.

(73) *Right Accent Foot falls on penultimate syllables in common speech styles*

Formal speech   *Common speech*

ti-{(βe.le).βe.t-e.s-a.n-e(e.ŋg-e)}   **ti-{(βe.le).βe.t-e.s-a.(n-ee)}.<ŋg>**
we-speak-caus-rec-hab-subj   "we should be causing each other to speak"

ti-{(ŋa.na).mu.l-i.s-a.n-e(e.ŋg-e)}   **ti-{(ŋa.na).mu.l-i.s-a.(n-ee)}.<ŋg>**
we-turn-caus-rec-hab-subj   "we should be causing each other to turn"

ti-{(to.nde).k-e.s-a.n-e(é.ŋg-e)}   **ti-{(to.nde).k-e.s-a.(n-eé)}.<ŋg>**
we-fail-caus-rec-hab-subj   "we should be causing each other to fail"

More important, however, is the fact that Accent Feet consistently belong to Strong Accent Constituents already determined by language users using Universal Guidelines of FINALITY and EDGENESS. This generalization, we said, can be accounted for in terms of the constraint ACFT/SAC. Furthermore, the accent grammar requires that segments of Strong Accent Constituents be preserved. The constraint FAITH-SAC is responsible for this task. Whatever happens to the non-prominent final vowel, it seems, is of little or none of the grammar's business. The constraint FTBIN in the tableau in (74) below ensures that Accent Feet are binary at all times.

(74) *FINALITY&EDGENESS identify penultimate syllable as SAC in the common speech style*

SACs: PROSODIC STEM; STEM-σ_1; PENULTIMATE SYLLABLE.

DGs: PROSODIC HIERARCHY; EDGENESS; FINALITY&EDGENESS.

/ti-[βeleβet-es-an-eŋg-e/ 'we will cause each other to speak'	FTBIN	ACFT/SAC	FAITH- SAC
(a) ti-{(βe.le).βe.te-.s-a.ne(e.ŋge)}		*!	
(b) ti-{(βe.le).β.t-.s-a.(n-ee)}.ŋg			*!*
(c) ti-{(βe).le.βe.te-.s-a.(nee)}.ŋg	*!		
☞ (d) ti-{(βe.le).βe.te-.s-a.(nee)}.ŋgi			
☞ (e) ti-{(βe.le).βe.t-e.s-a.(nee)}.ŋg			

All the candidates satisfy the constraint FTBIN. However, candidate (74a) is banned because it violates the constraint ACFT/SAC which requires Accent Feet to belong to Strong Accent Constituents such as STEM-σ_1 and PENULTIMATE-σ as determined by Universal Guidelines of EDGENESS and EDGENESS&FINALITY, respectively. The candidate violates this constraint by placing the right edge Accent Foot on the weak final syllable. Candidate (74b) is non-optimal since it violates the constraint FAITH-SAC, requiring elements of Strong Accent Constituents such as Prosodic Stems in the common speech style to be preserved. Candidate (74c) is ruled out because it violates the constraint FTBIN which requires Accent Feet to be binary at the level of the mora. Candidates (74d&e), however, win because they satisfy all the three constraints.

Another observation about vowel deletion in the common speech style was that low vowels are never deleted in the same final syllable position. My analysis of this fact is that a strong position shifts from the penultimate syllable to the final syllable when the vowel of the latter is the most sonorous low vowel. An Accent Foot also shifts to the final syllable as given in (75) below.

(75) *Accent Foot is rightmost when the final vowel is [+low]*

Formal speech   Common speech

{(βe.le).βe.t-e.s-a.n-a(a.ŋg-a)}   **{(βe.le).βe.t-e.s-a.n-a(a.ŋg-a)}**
speak-caus-rec-hab-fv   "be causing each other to speak"

{(βe.le).βe.t-e.s-a(a.n-a)}   **{(βe.le).βe.t-e.s-a(a.n-a)}**
speak-caus-rec-fv   "cause each other to speak"

{(ŋa.na).mu.l-i.s-a.n-a(a.ŋg-a)}   **{(ŋa.na).mu.l-i.s-a.n-a(a.ŋg-a)}**
turn-caus-rec-hab-fv   "be causing each other to turn"

{(to.nde).k-e.s-a.n-a(á.ŋg-a)}   **{(to.nde).k-e.s-a.n-a(á.ŋg-a)}**
fail-caus-rec-hab-fv   "be causing each other to fail"

Thus, what counts for an AC to be a Strong Accent Constituent is, first and foremost, for it to satisfy the Guidelines of SONORITY and EDGENESS. The Universal Guideline of FINALITY is suspended for the sake of SONORITY.

This analysis is not in conflict with our generalization that Accent Feet are assigned to Strong Accent Constituents. Edgemost and more sonorous low vowels can never be weaker to any other AC, at least for this speech style. In terms of OT, this foot assignment satisfies the constraint ACFT/SAC, requiring Accent Feet to be assigned to Strong Accent Constituents such as the PROSODIC STEM, STEM-σ_1 and the FINAL SYLLABLE. The constraint FAITH-SAC in the tableau in (76) below ensures that segments of these Strong Accent Constituents are preserved. The constraint FTBIN would ensure that the Accent Feet are binary at the level of the mora.

(76) *EDGENESS&SONORITY outrank FINALITY as DGs for SACs*

SACs:　　　　　PROSODIC STEM; STEM-σ_1; FINAL SYLLABLE

DGs:　　　　　PROSODIC HIERARCHY; EDGENESS; EDGENESS&SONORITY

/βeleβet-a/ 'speak'	FAITH-SAC	ACFT/SAC	FTBIN
(a) {(βe.le).(βee).t-a}		*!	
(b) {(βe).le.βee.(t-a)}			*!*
☞ (c) {(βe.le).βe(e.t-a)}			

Thus candidate (76a) is disqualified because it violates the constraint ACFT/SAC which requires Accent Feet to belong to Strong Accent Constituents such as the FINAL SYLLABLE as provided by EDGENESS&SONORITY. Candidate (76b) is nullified because it violates FTBIN which requires Feet to be binary at the level of the mora or the syllable. Candidate (76c) wins because it satisfies all the three constraints. We know that an Accent Foot is assigned to the final syllable because when the Stem appears in phrase medial position (that is, phrase stress is no longer on the penultimate syllable), only the final vowel [a] survives. Note that in most cases, especially in rare elderly speech styles, the vowel of the second syllable of the left Foot is usually deleted. Thus, the Prosodic Stem {(βe.le).βe(e.t-a)} 'speak' may become {(βe.l).(β.t-a)} in phrase medial position [e.g. ||{(βe.l).(β.t-a)} ukóongwa|| 'speak a lot'].

3.6.1.3 Elderly speech styles

One observation we made about the common elderly speech style was that the only non-low vowels that are preserved are those which are nuclei of the first two syllables and stem-penult syllables [e.g. formal **ti-{(βe.le).βe.t-e.s-a.n-e(e.ŋg-e)}** $\Rightarrow$ {**ti-(βe.le).β.t-.sa-.(n-ee).ŋg**}].. The analysis of this fact would be that ACCENT FOOT (as opposed to PROSODIC STEM), STEM-σ_1 and PENULTIMATE-σ are Strong Accent Constituents. The Determiner Guidelines are PROSODIC HIERARCHY, EDGENESS and EDGENESS&FINALITY. A crucial generalization, once again, is that segments of Strong Accent Constituents such as STEM-σ_1, PENULTIMATE-σ and ACCENT FEET are preserved and this fact can be accounted for by the constraint FAITH-SAC. The constraint ACFT/SAC in the tableau in (77) below ensures that the Accent Feet fall

on SACs (going by Guidelines of EDGENESS and FINALITY). The constraint FTBIN ensures that Accent Feet are binary at the level of the mora.

(77)   *EDGENESS & FINALITY are DGs for SACs in the common elderly speech style*

SACs: ACCENT FOOT; STEM-σ_1; PENULTIMATE-σ;

DGs: PROSODIC HIERARCHY; EDGENESS; EDGENESS&FINALITY

/ti-βeleβet-es-an-eŋg-e/ 'we should be teaching each other'	ACFT/SAC	FTBIN	FAITH-SAC
☞ (a) ti-{((βe.le.)βe.t-e.s-a.(n-ee.)}ŋg-i			
☞ (b) ti-{((βe.le.)β.t-.s-a.(n-ee.)}ŋg			
(c) ti-{((βe.l.)βe.t-e.s-a.(n-ee.)}ŋg-e			*!
(d) ti-{((βe.le.)βe.t-e.s-a.(n-.)}ŋg-e			*!

Thus, candidates (77a&b) are optimal because they satisfy FAITH-SAC which requires elements of Strong Accent Constituents such as Accent Feet in the common elderly speech styles to be preserved. The fact that non-low vowels outside the foot structure are deleted or altered reflects the property of an accent grammar of not paying much attention to what happens in weaker positions. Candidates (77c&d), on the other hand, are non-optimal because they violate FAITH-SAC.

Another observation was that in rare elderly speech styles low vowels may be deleted as well, except those ones belonging to the initial two and last two syllables of the Prosodic Stem. The analysis once again would be that STEM-σ_1 and FINAL SYLLABLE are Strong Accent Constituents. The Guidelines are EDGENESS and SONORITY, respectively. EDGENESS provides that the status of SAC be conferred on edgemost syllables; and SONORITY provides that a syllable with the most sonorous nucleus should be conferred the status of being SAC. The FINAL SYLLABLE with low vowel in the examples in (78) below better satisfy these guidelines than the PENALTIMATE SYLLABLE.

(78) *Only edgemost low vowels are SACs*

Formal speech	Rare elderly speech

{(tʰa.mba).l-i.s-a.n-a(a.ŋg-a)}
stretch-caus-rec-hab-fv

{(tʰa.mba).l-.s-.n-a(a.ŋga)}
"be causing each other to stretch legs"

{(βe.le).βe.t-e.s-a.n-a(a.ŋg-a)}
speak-caus-rec-hab-fv

{(βe.le).β.t-.s-.n-a(a.ŋga)}
"be causing each other to speak"

{(to.nde).k-e.s-a.n-a(á.ŋg-a)}
fail-caus-rec-hab-fv

{(to.nde).k-.s-.n-a(á.ŋga)}
"be causing each other to fail"

The fact that low vowels can now be deleted, even those of the Prosodic Stem, indicates that low vowels are no longer a Strong Accent Constituent at language level. However, they are SACs in a specific domain, the right edge where they are able to attract an Accent Foot. Thus their deletion outside the foot structures does not lead to violation of the constraint FAITH-SAC. A tableau in (79) below illustrates.

(79) *Final syllables with low vowel nucleus are Strong Accent Constituents*

SACs: ACCENT FOOT; STEM-σ₁; FINAL-σ

DGs: PROSODIC HIERARCHY, EDGENESS, EDGENESS&SONORITY

/βeleβet-es-an-aŋg-a/	ACFT/SAC	FTBIN	FAITH-SAC
☞ (a) -{(βe.le).β.t-.s-.na(a-.ŋg-a)}			
☞ (b) -{(βe.le).β.t-.s-a.n-a(a.ŋg-a)}			

Although candidate (79a) has the low vowel deleted, it does not violate FAITH-SAC because low vowels are no longer SAC at language level. The low vowel can be there or not because the grammar has little or no interest in weaker positions. The fact that the penultimate low vowel is also preserved can be attributed to the fact that the penultimate syllable is the Head of the Phonological Phrase and thus another SAC as guided by the Universal Guidelines of Stress as well as EDGENESS (RIGHTEDGE) and FINALITY (or something like that). I now turn to Strong Accent Constituent Theory (SACT) analysis of consonant deletion and preservation.

3.6.2 Consonant deletion and preservation

3.6.2.1 Formal speech style

One observation we made is that all consonants of the Morphological Stem, including liquid consonants in the final syllable position, are preserved in the more formal speech style. The fact that the liquid onset to the final syllable is also preserved point to the fact that the final syllable is one of the Strong Accent Constituents, regardless of the type of onset. The suggestion therefore is that a right Accent Foot in the formal speech style falls on the rightmost edge of another Strong Accent Constituent, Prosodic Stem, as illustrated in (80) below.

(80) *A right Accent Foot in formal speech style is edgemost*

Stem	English gloss
{(be.ne).ke.l-e(e.l-a)} cover-appl-fv	cover with
{(sa.mbi.)l-i(i.l-a)} learn-appl-fv	learn for
{(βe.le).βe.t-e(e.l-a)} speak-appl-fv	speak for
{(ka.li).p-i(i.l-a)} reprimand-appl-fv	reprimand

Thus, preservation of the liquid in the final syllable position can be explained on the basis of the strong position occupied by the final syllable as provided by EDGENESS. Two generalizations can be made from this analysis: Accent Feet belong to Strong Accent Constituents and segments of Strong Accent Constituents are preserved. In terms of OT, once again, the constraints ACFT/SAC and FAITH-SAC better account for the facts. The constraint FTBIN in the tableau in (81) below ensures that the Accent Foot is binary at the level of the mora.

(81) *EDGENESS is Determiner Guideline for SACs in the formal speech style*

SACs: PROSODIC STEM; STEM-σ_1; FINAL-σ

DGs: PROSODIC HIERARCHY; EDGENESS; EDGENESS

/kalip-il-a/ 'reprimand'	ACFT/ SAC	FTBIN	FAITH- SAC
(a) {(ka.li).(p-ii).l-a}	*!		
(b) {(ka.li).p-ii.(l-a)}		*!	
(c) {(ka.li).p-i(i.-ʲa)}			*!
☞ (d) {(ka.li).p-i(i.l-a)}			

Candidate (81a) is disqualified because it violates the constraint ACFT/SAC which requires Accent Feet to fall on Strong Accent Constituents like edgemost syllables of the Prosodic Stem as provided by the Determiner Guideline of EDGENESS. Candidate (81b) satisfies ACFT/SAC, but it is ruled out because it violates FTBIN which requires Feet to be binary at the level of the mora. Candidate (81c) is ruled out because it violates the constraint FAITH-SAC which militates against deletion of Strong Accent Constituents such as those of the Prosodic Stem and FINAL-σ. Candidate (81d), on the other hand, wins because it satisfies all the three constraints.

3.6.2.2 Common speech style

In the common speech style the observation was that liquid consonant onsets to final syllables are usually deleted, while liquid consonants of the Prosodic Stem (minus final syllables) are preserved. My suggested analysis is that the Strong Accent Constituents in this speech style do not include the final syllable. SACs include PROSODIC STEM, STEM-σ_1; and PENULTIMATE-σ following the Guidelines of PROSODIC HIERARCHY, EDGENESS and EDGENESS&FINALITY, respectively. The final syllable with liquid onsets is thus rendered extrametrical and whatever happens to the non-SAC liquid, the grammar has little or no control over it. This analysis is illustrated in (82) below.

(82) *Accent Feet avoid final syllables with liquid onsets*

Formal	*Common speech*	*English gloss*
{(tʰa.mba).li.l-i(i.l-a)} stretch-appl-fv	{(tʰa.mba).li.(l-ii).}-ʲa	stretch for
{(sa.mbi.)l-i(i.l-a)} learn-appl-fv	{(sa.mbi.)(l-ii).}-ʲa	learn for
{(gu.ndu.)mu(u.l-a)} destroy-fv	{(gu.ndu.)(muu).}ʷ-a	destroy
{(ka.li).p-i(i.l-a)} reprimand-appl-fv	{(ka.li).(p-ii).}-ʲa	reprimand
{(pu.m-i).l-a.n-i(i.l-a)} hit-appl-rec-appl-fv	{(pu.m-i).l-a.(n-ii).}-ʲa	hit for each other with

Two crucial generalizations about this analysis therefore are that Accent Feet fall on Strong Accent Constituents such as edgemost-but-non-final syllables. Thus the penultimate syllable is stronger than the final syllable. Since all segments of the Prosodic Stem are preserved, it is most likely that PROSODIC STEM is a minimal Strong Accent Constituent as guided by the principle of PROSODIC HIERARCHY. Another generalization therefore is that segments of Strong Accent Constituents such as PROSODIC STEM are preserved. These two generalizations are accounted for in terms of the constraints ACFT/SAC, requiring Accent Feet to belong to stronge Accent Constituents, and FAITH-SAC which requires Strong Accent Constituents to be preserved. A tableau in (83) below illustrates these facts. The constraint FTBIN ensures that Feet are binary at the level of the mora.

(83)   *EDGENESS&FINALITY are some of Determiner Guidelines for SACs in the common speech style*

SACs: PROSODIC STEM; STEM-σ_1; PENULTIMATE-σ

DGs: PROSODIC HIERARCHY; EDGENESS; EDGENESS&FINALITY

/pum-il-an-il-a/ 'hit for each other with'	FAITH-SAC	ACFT/SAC	FTBIN
(a) {(pu.m-i).-^{j}a.(n-ii)}.-^{j}a	*!		
(b) {(pu.m-i).l-a.n-i(i.-^{j}a)}		*!	
☞ (c) {(pu.m-i).l-a.(n-ii)}.-^{j}a			

Candidate (83a) is disqualified because it violates FAITH-SAC which requires segments of Strong Accent Constituents such as Prosodic Stems to be preserved. Candidate (83b) is ruled out because it violates ACFT/SAC which requires Accent Feet to fall on Strong Accent Constituents such as penultimate syllables as provided by EDGENESS&FINALITY. Candidate (83c) is optimal because it satisfies the constraints FAITH-SAC and ACFT/SAC.

Another observation that we made about consonant deletion in the common speech style was that nasal and obstruent consonants are never deleted in the same position. The analysis would be that NASAL CONSONANT and OBSTRUENT CONSONANT are Strong Accent Constituents in this speech style and they (plus low vowel) attract Accent Feet when they are edgemost. This analysis is illustrated in (84) below. The Determiner Guidelines of the right SAC are most likely EDGENESS&SONORITY. FINALITY, which provides that only non-final syllables are SACs, is suspended for SONORITY's sake[26].

---

[26] Just like in cases where the non-low final vowel is deleted, deletion of liquid onsets before final moras in the common speech style makes me think that the the final syllable of this quality is most likely not parsed by either the Accent Foot or the Prosodic Stem [e.g. formal {(t^{h}a.mba).li(i.l-a)} $\Rightarrow$ {(t^{h}a.mba).(lii-.)}^{j}a 'stretch one's legs'].

(84) *Sonorously strong final syllables attract Accent Feet*

Stem	*English gloss*
{(tha.mba).l-i.l-a(a.n-a)} stretch-appl-rec-fv	stretch legs over each other
{(se.k-e).l-e.s-a(a.n-a)} laugh-appl-rec-fv	cause to laugh for each other
{(zo.me).le.z-a(a.n-a)} agree-rec-fv	agree with each other
{(βe.le).βe(e.t-a)} speak-fv	speak
{(zo.me).le(e.z-a)} agree-fv	admit
{(ka.mbi).l-i(i.k-a)} speak-pass-fv	be spoken to

Once again, the two crucial generalizations are that Accent Feet belong to Strong Accent Constituents such as edgemost syllables with nasal or obstruent onsets (and of course most sonorous low vowel nuclei) (see footnote #26 on foot assignment when the onset of the final syllable is a liquid consonant). Another generalization is that Strong Accent Constituents such as nasal and obstruent consonants (and low vowels) are preserved. These generalizations, once again can be accounted for by the constraints ACFT/SAC and FAITH-SAC. A tableau in (85) below illustrates. The constraint FTBIN, as usual, ensures that Accent Feet are binary at the mora level.

(85)   EDGENESS&SONORITY relegate FINALITY as Determiner Guideline for SACs

SACs: PROSODIC STEM, STEM-σ_1, FINAL-σ

DGs: PROSODIC HIERARCHY, EDGENESS, EDGENESS&SONORITY

/kamb-il-ík-a/ 'be informed'	ACFT/SAC	FTBIN	FAITH-SAC
(a) {(ka.mb-i).(l-ií).}k-a	*!		
(b) {(ka.mb-i).l-ií.(k-a)}		*	
(c) {(ka.mb-i).l-i(í.-^{j}a)}			*!
☞ (d) {(ka.mb-i).l-i(í.k-a)}			

Candidate (85a) is disqualified because it violates the constraint ACFT/SAC which requires Accent Feet to be assigned to Strong Accent Constituents such as the final syllable as provided by the Guidelines EDGENESS&SONORITY in this speech style. Candidate (85b) is booked for violating FTBIN which requires Accent Feet to be binary at the level of the mora. Candidate (85c) is non-optimal for violating FAITH-SAC, requiring Strong Accent Constituents such as obstruent consonants and footed segments to be preserved. The last candidate (85d) wins because it satisfies all these constraints.

3.6.2.3 Elderly speech styles

In the *common elderly speech style* the only liquid consonants that are preserved are those which are onsets of the first two syllables and the bimoraic penultimate syllable of the Prosodic Stem (e.g. formal **gu.lu.lu.l-ii.l-a** $\Rightarrow$ **gu.lu.<'>.l-ii.-^{j}a**). The analysis would be that Accent Feet fall on STEM-σ_1, PENULTIMATE-σ while the final syllable is rendered extrametrical [[e.g. formal **{(gu.lu).lu.l-i(i.l-a)}** $\Rightarrow$ **{(gu.lu).<'>.(l-ii).}-^{j}a)**] as evidenced by deletion of the liquid onset of the final syllable. Furthermore, ACCENT FOOT, rather than PROSODIC STEM, is deemed a primary SAC by speakers of the common elderly speech style as evidenced by preservation of only those liquids which are footed.

Since penultimate syllables are the ones which have preserved liquid onsets, the direct determiners of right edge SAC seem to be EDGENESS&FINALITY. The former provides that edgemost syllables are stronger while the latter provides that non-final syllables are stronger

than final syllables. The penultimate syllable better follows these Guidelines and it comes out stronger. The generalization, once again, is that Accent Feet belong to Strong Accent Constituents such as the penultimate syllable in this case. The constraint ACFT/SAC seems to be responsible for this status. The constraint FAITH-SAC ensures that segments of the Strong Accent Constituent ACCENT FOOT and others like NASAL and OBSTRUENT CONSONANTS and LOW VOWELS are preserved. The constraint FTBIN ensures that the Accent Feet are binary at the level of the mora. A tableau in (86) below illustrates.

(86)   *EDGENESS&FINALITY as DGs for SACs in the common elderly speech style*

SACs: ACCENT FOOT; STEM- σ_1; PENULTIMATE-σ.

DGs: PROSODIC HIERARCHY; EDGENESS; EDGENESS&FINALITY

/gululul-il-a/ 'pull for'	FTBIN	ACFT/SAC	FAITH-SAC
(a) {(gu.lu).lu.l-i(i.l-a)}		*!	
(b) {(gu.lu).<lu>.(-ʷii).}-ʲa			*!
☞ (c) {(gu.lu).<'>.(l-ii).}-ʲa			

All the candidates satisfy FTBIN. However, candidate (86a) is disqualified because it violates ACFT/SAC which requires Accent Feet to fall on Strong Accent Constituents such as the penultimate syllable as provided by Determiner Guidelines EDGENESS&FINALITY. Candidate (86b) is non-optimal because it violates FAITH-SAC, requiring segments of Strong Accent Constituents such as footed segments to be preserved. Candidate (86c) wins because it satisfies all the three constraints in the tableau.

On the other hand, if the onset to the final syllable were a nasal or an obstruent, the right foot would be assigned to the edgemost syllable. The evidence for this analysis would be retention of the nasal or obstruent onset alongside the final low vowel [e.g. **{(to.nde).k.sa(a.na)}**]. An analysis of this fact would be that the Guidelines EDGENESS&SONORITY together consistently relegate FINALITY. Finally, in very-very rare cases nasal onsets may also be deleted by elderly speakers. I do not dwell on it here, but my analysis of this fact would still be that nasal

consonants are no longer Strong Accent Constituents in these very rare cases. Thus, their deletion would not lead to violation of FAITH-SAC.

Summary

In the formal speech style, all segments of the Prosodic Stem, including liquid consonants or non-low vowels in the final syllable position are preserved. The analysis suggested has been that PROSODIC STEM, STEM-σ_1 and FINAL-σ are Strong Accent Constituents following provisions of Universal Guidelines of PROSODIC HIERARCHY and EDGENESS.

In the common speech style, liquid consonant onsets or non-low vowels of final syllables are usually deleted while liquid consonant onsets or non-low vowels belonging to non-final syllables are preserved. The analysis has been that PROSODIC STEM (minus final syllable), STEM-σ_1 and the (bimoraic) PENULTIMATE-σ are Strong Accent Constituents as guided by principles of PROSODIC HIERARCHY, EDGENESS and EDGENESS&FINALITY. Another observation was that nasal and obstruent consonants are never deleted in the same position (final syllable). The analysis has been that Nasal and Obstruent Consonants are Strong Accent Constituents in this speech style and when the onset to the final syllable is a nasal or obstruent consonant (plus low vowel nucleus), the final syllable attains the SAC status. Thus relevant Strong Accent Constituents for foot assignment purposes in such constructions are PROSODIC STEM, STEM-σ_1 and FINAL-σ. The Determiner Guidelines are the PROSODIC HIERARCHY, EDGENESS and EDGENESS&SONORITY.

In elderly speech styles, the only liquid consonants or non-low vowels that are preserved are those which are onsets or nuclei of the first two syllables and of the bimoraic penultimate syllable of the Prosodic Stem. The analysis has been that ACCENT FOOT, rather than the PROSODIC STEM, is the minimal Strong Accent Constituent in this speech style as guided by the principle of PROSODIC HIERARCHY. Thus, an inventory of SACs in this speech style usually includes the ACCENT FOOT, STEM-σ_1 and the (bimoraic) PENULTIMATE-σ as guided by PROSODIC HIERARCHY, EDGENESS and EDGENESS&FINALITY. If the onset to the final syllable were a nasal or obstruent consonant (plus a low vowel nucleus), the final syllable would once again attain the SAC status on the basis of EDGENESS&SONORITY.

In all the cases above, the generalizations were that segments of Strong Accent Constituents are preserved and that Accent Feet belong to Strong Accent Constituents. The constraint FAITH-SAC would ensure that segments of the Strong Accent Constituents such as PROSODIC

STEM, ACCENT FOOT, STEM-σ_1, PENULT-σ or FINAL-σ, LOW VOWELS, and NASAL and OBSTRUENT CONSONANTS are preserved. The constraint ACFT/SAC would ensure that Accent Feet fall on Strong Accent Constituents of the Prosodic Stem which are STEM-σ_1, PENULT-σ or FINAL-σ. And the constraint FTBIN would ensure that the feet are binary at the level of the mora.

3.7   Chapter summary

The aim of this chapter was to present formally a proposal for the theory of Strong Accent Constituency within a larger framework of Universal Guidelines and Constraints. I began the chapter by presenting the data on vowel and consonant deletion which motivates this theory analysis. Then attempts were made to account for the facts in terms of stress-accent theory and Downing's (2006b) Morpheme-Based Templates Theory both of which are found to be slightly problematic to account for segment deletion patterns exhibited in ciTonga. The conclusion has been that a theory based on Universal Guidelines and Constraints for Strong Accent Constituents accounts for the facts slightly better.

The Interaction of Tone and Strong Accent Constituents

4.1   Introduction

The aim of this chapter is to present formally a theory of Interaction between Tone and Strong Accent Constituents as a slightly better way to account for distribution of tone in ciTonga and perhaps many other Bantu languages as well. The chapter is organized as follows: Section 4.2 presents the data on tone assignment in basic verbs, simple past tense verbs and present progressive aspect verbs which motivate this type of theory analysis. Sections 4.3, 4.4, and 4.5 attempt to account for the facts in terms of tone alignment theory (as argued for by Mtenje 2006), autosegmental accent (Goldsmith 1984, Clements and Goldsmith 1984), and pitch-accent (as hinted upon by Downing 2004), but all these theoretical perspectives have been found to be slightly problematic to account for tone distribution patterns as exhibited in ciTonga. Section 4.6 presents formally proposals for a theory of Interaction between Tone and Strong Accent Constituents.

4.2   The data

I will look at tone assignment in basic verbs, infinitive verbs and tensed verbs with emphasis on formal and common speech styles[27].

4.2.1   Tone in basic verbs (without prefixes)

Many basic verbs are toneless as given in (1) below.

---

[27] This has been a traditional approach to the study of tone in Bantu languages (i.e. looking at basic verbs and then tensed verbs). By basic verbs, I mean verb stems/words which have no prefixes. Infinitive verbs are those which have infinitive marker **ku-** prefixed to some morphological stem. And tensed verbs are verb complexes in various tense/aspect types (e.g. simple past tense, distant future tense, present progressive aspect, etc). In addition, I will assume that the tone bearing unit is the mora, but sometimes I use the syllable when it is light syllable.

(1)   *Low-Toned verbs*

Basic stem   *English gloss*

[le.lee.s-a   look at/see

[lee.mb-a   write

[da.nii.k-a   answer

[sa.mbii.z-a   teach

[βe.le.βee.t-a   speak

Many others, however, have a high tone on the penultimate mora as given in (2) below[28].

(2) *High-Toned basic verb stems*

Basic stem   *English gloss*

[cʰi.mbií.j-a   run

[dií.k-a   cover yourself

[da.nií.k-a   be invited

[sa.mbií.l-a   learn

[to.ndeé.k-a   fail

[kʰu.mbií.l-a   admire

[bií.j-a   cheat

When the high-toned verb appears in phrase-medial position the high tone falls on the final syllable/mora of the Prosodic Stem. Tone does not shift to the penultimate syllable of a next word as length does. This fact is illustrated in (3) below.

---

[28] Tone is phonemic in very rare cases.

(3) *Tone shifts to the final mora when the Stem is in phrase-medial position*

[cʰi.mbií.j-a		run
[cʰi.mbi.j-á	vi-.(βaa.)ntʰo	run away from bad people
[to.ndeé.k-a		fail
[to.nde.k-á	ma.jee.so	fail an exam
[sa.mbi.z-ií.j-a		help solve a clue
[sa.mbi.z-i.j-á mu.ntʰi.kaa.ze		help a woman solve a clue
[kʰu.mbií.l-a		admire
[kʰu.mbi.l-á	mu.ntʰu.luu.me	admire a man
[sa.mbií.l-a		learn
[sa.mbi.l-á	(ntʰaa.)ze	learn a lot
[(dií)k-a		cover
[di.k-á	bu.lá.(ŋgee.te)	cover with blankets

4.2.2   Tone in simple past tense verbs

The simple past tense in ciTonga is marked by the morpheme **-ŋgu-** and it normally precedes the verb stem. Where the verb stem has two or three syllables, the high tone is assigned to the initial syllable as given in (4) below.

(4) *The simple past tense*

ndi-.ŋgu-.[víi.n-a   I danced
[I-past-dance-fv]

ndi-.ŋgu-.[sáa.mb-a   I bathed
[I-past-bathe-fv]

ndi-.ŋgu-.[tháa.β-a   I ran
[I-past-run-fv]

ndi-.ŋgu-.[lé.lee.s-a   I looked
[I-past-look-fv]

ndi-.ŋgu-.[βé.lee.ŋg-a   I read
[I-past-read-fv]

ndi-.ŋgu-.[tú.tuu.z-a   I pushed
[I-past-push-fv]

However, when the morphological stem has four or more syllables, the high tone is realized on the second syllable. The tone then spreads rightwards, but it does not spread to the last two syllables. These facts are exemplified in (5) below.

(5)   *Ideally tone is on the second mora of the morphological stem*

ndi-.ŋgu-.[tʰa.mbá.lii.-ʲa   I stretched (legs)
I-past-stretch-fv

ti-.ŋgu-.[le.lé.s-aa.n-a   we looked at each other
we-past-look-rec-fv

ndi-.ŋgu-.[βe.lé.βé.t-ee.s-a   I caused to speak
speak-caus-fv

ti-.ŋgu-.[vu.mbá.tí.l-aa.n-a   we hugged each other
we-past-hug-rec-fv

ti-.ŋgu-.[tʰa.mbá.lí.l-aa.n-a   we stretched legs over each other's
we-past-stretch-rec-fv

ti-.ŋgu-.[pa.kú.l-í.l-aa.n-a   we served for each other
we-past-serve-appl-rec-fv

4.2.3   Tone in the present progressive aspect verbs

Normal verbs (where morphological stems have two or more syllables) in the present progressive aspect have no morphological marker. A high tone seems to spread rightwards from word-initial syllable to the antepenult syllable. The spreading tone does not penetrate the last two syllables. These facts are illustrated in (6) below.

(6)   *Normal Present Progressive Aspect verbs have no morphological marker*

Verb   *English gloss*

ndí-.[lé.lee.s-a   I am looking
I-look-fv

ndí-.[zó.mee.l-a   I am admitting
I-admit-fv

ndí-.[sá.mbii.z-a   I am teaching
I-teach-fv

ndí-.[βé.lé.βee.t-a   I am speaking
I-speak-fv

ndí-.[vú.nú.lii.-ʲa   I am steaming (myself)
I-steam-fv

tí-.[tʰá.mbá.lí.l-aa.n-a   we are stretching legs over each another
we-stretch-rec-fv

When a morphological stem involved is monomoraic, however, there appears a morpheme **-tu-** before the stem. This entity assigns the high tone to itself. These facts are illustrated in (7) below[29].

---

[29] There is evidence to show that these verbs are indeed monomoraic. For instance, they can take other types of prefixes as in **kuú-.[mwa** 'to drink' **kuú-.[fwa** 'to die' and **kuú-.[ba** 'to steal'.

(7) *Present progressive verbs with morpheme -tu-*

Verb	English gloss
ndi-.túu-.[vw-a	I am listening
ndi-.túu-.[mb-a	I am singing
ndi-.túu-.[lj-a	I am eating
ndi-.túu-.[b-a	I am stealing
ndi-.túu-.[mw-a	I am drinking
ndi-.túu-.[sw-a	I am breaking

Summary

Basic verbs (without prefixes) in ciTonga are either toneless (low-toned) or high-toned. Tone is usually assigned to the penultimate mora (e.g. [**kʰu.mbií.l-a** 'admire'). When a high-toned verb appears in the phrase medial position the high tone falls on the final syllable of the Stem (e.g. [**kʰu.mbi.l-á mu.ntʰu.luu.me** 'admire a man').

The high tone in **simple past tense verbs** is assigned to the initial syllable of Morphological Stems with two or three syllables (e.g. **ndi-.ŋgu-.[lé.lee.s-a** 'I looked'). However, the high tone is realized on the second syllable of Morphological Stems with four or more syllables (e.g. **ti-.ŋgu-.[tʰa.mbá.lí.l-aa.n-a** 'we stretched legs over one another'). The tone then spreads to following moras or syllables (if there is any), but it does not spread to the last two syllables.

Normal verbs (where morphological stems have two or more syllables) in the **present progressive aspect verbs** have no morphological marker (e.g. **ndí-.[sá.mbii.z-a** 'I am teaching'). A high tone seems to spread from word-initial syllable to the antepenult syllable. The "spreading" tone does not penetrate the last two syllables. When a morphological stem involved is *monomoraic*, however, there appears a morpheme **-tu-** before the stem (e.g. **ndi-.túu-.[vw-a** 'I heard'). This entity assigns the high tone to itself.

In the following sections, I present three theoretical perspectives which have guided thinking in as far as tone assignment in Bantu languages is perceived. These thoughts include Tone Alignment Theory (e.g. Mtenje 2006 on ciTonga), Autosegmental Accent Approach to Tone (Clements' and Goldsmith 1984), and Accentual Properties of Tone in Bantu Languages (Downing 2004).

4.3 The tone alignment theory (Mtenje 2006)

4.3.1 Basic verbs

Crucial analyses of tone in ciTonga have been presented by Mtenje (1994/95, 2006). Mtenje (1994/95) observes that this language, like many Bantu languages, has two tones, high and low. He suggests that low-toned verbs (e.g. [tʰa.mba.lii.ja 'stretch legs') are lexically unmarked for tones, while high-toned verbs (e.g. [sa.mbi.zií.ja 'help out in solving a clue') are lexically specified for the high tone. What needs to be accounted for by the grammar of ciTonga therefore is the fact that the high tone in high-toned verbs is usually on the penultimate mora of the Prosodic Stem.

Mtenje's (1994/95) suggestion is that the target of the high tone is the final syllable, but there is a tone retraction rule similar to that found in neighbouring languages like Chichewa and Chiyao (Mtenje 1986, 1993) which retracts a word-final high tone leftwards to the penultimate 'vowel' under varying conditions. What is crucial in this analysis is the suggestion that the right edge of a morphological stem is the target for the high tone assignment. The only problem is the rule-based approach which, as it is widely believed today, cannot adequately explain why the high tone is located where it is.

Mtenje (2006) addresses this problem by proposing an OT alignment theory approach where the high tone in basic verbs aligns with the right edge of the Morphological Stem. However, he observes that it is common in languages (e.g. Chichewa and Chiyao) for high tones in verb stems to avoid stem-final vowels when the verbs are phrase-final. According to Mtenje (2006), this shows that the right edge of the morphological stem is not aligned with a high (H) tone and a constraint which prohibits H from being aligned with the right edge of the morphological stem, usually referred to as NONFINALITY in the literature, accounts for this tone fact. Mtenje follows Myers and Carleton (1996) in their formulation of this constraint (with a slight modification) as shown in (8) below.

(8)   **Stem Non-Finality (Stem Non-Fin)**

The right edge of the morphological stem is not aligned with a high tone.

However, as we also saw in our data, when the high-toned verb appears in phrase-medial positions the high tone is realized on the final mora (e.g. cʰi.mbi.j-á vi-.βaa.ntʰo 'run away from bad people'). The point I want to make is that there seems to be something in the *last two syllables* of the basic verbs which makes them the only domain for the high tone. This

fact calls for Universal Guidelines which must regulate power relations between the last two moras or syllables (when the final syllable is stronger than the penult and vice versa). My complaint is that constraints such as NONFINALITY or/and RIGHTMOST are not informative enough since they do not reflect the power shift that goes on between the last two syllables as Porosodic Words/Stems appear in different environments.

4.3.2   The simple past tense verbs

The facts requiring explanation in the simple past tense are as follows: (a) where the Morphological Stem has two or three syllables, the high tone is assigned to the initial syllable of the Stem (e.g. **ndi-.ŋgu-.[lé.lee.s-a**); (b) when the morphological stem has four or more syllables, the high tone is realized on the second syllable of the Stem (e.g. **ti-.ŋgu-.[tʰa.mbá.lí.l-aa.n-a**); and (c) the high tone spreads to following moras, but it does not spread to the last two syllables.

Both Mtenje (1994/95) and Mtenje (2006) address observation (a). Mtenje's (1994/95) analysis of tone assignment in the simple past tense is that the high tone is placed on the first vowel of the Morphological Stem where the high tone occurs after the tense marker **-ŋgu-**. As we can see, the analysis is purely descriptive and it lacks appeal to universality. Mtenje (2006) addresses this problem. Once again, he follows Myers and Carleton's (1996) analysis of Chichewa that the high tone can be accounted for by an alignment constraint that requires the high tone to align with the left edge of the Morphological Stem. The constraint is adapted in (9) below.

(9)   **ALIGN – L (MStem, L)**

        (Simple past tense)

        The left edge of the morphological stem in simple past tense verbs is aligned with a high tone (Mtenje 2006).

These analyses are not without challenges. For instance, both Mtenje (1994/95) and Mtenje (2006) make conclusions from inadequate data. Their assertion that the high tone is exclusively assigned to the root or stem-initial vowel is proved false by observation (b) above which shows that the high tone is realized on a second syllable or mora of the stem when the Morphological Stem involved has four or more syllables (e.g. **ti-.ŋgu-.[tʰa.mbá.lí.l-aa.n-a**). The problem we need to address is that the domain for the high tone in the simple past tense verbs seems to be *the initial* or *initial two syllables* of the Morphological Stem, depending on

the length of the Morphological Stem. What is so special about these syllables? The third observation (tone spread) is tackled in the following sub-section.

4.3.3   The present progressive aspect verbs

An adequate grammar of ciTonga should also be able to account for all the facts about tone assignment in the present progressive aspect verbs. The facts requiring explanation include the following:

(a)   *Normal* verbs (where morphological stems have two or more syllables) have no morphological marker (e.g. **ndí-.[sá.mbii.z-a** 'I am teaching').

(b)   A high tone seems to spread rightwards from word-initial syllable to the antepenult syllable (e.g. **ndí-.[βé.lé.βé.t-ee.s-a**).

(c)   The "spreading" tone does not penetrate the last two syllables.

(d)   When a morphological stem involved is *monomoraic*, there appears a morpheme **-tu-** before the stem (e.g. **ndi-.túu-.[vwa**).

(e)   The morpheme **-tu-** assigns the high tone to itself.

From this list of observations, Mtenje (2006) addresses observations (b-e). His analysis of observations (b&c) is that the present progressive aspect which is marked by a zero morpheme places a high tone on the first syllable of the entire verb word and this tone spreads to the next vowel(s). His OT analysis follows that adopted by Myers and Carleton (1996) for Chichewa. In the present progressive aspect, the left edge of the Morphological Word is aligned with a high tone and this tone spreads rightwards to the next vowel by tone spreading. The assignment of the high tone to the left edge of the Morphological Word is accounted for by the alignment constraint adapted in (10) below.

(10)   **ALIGN-L (Mword, H) (Align Mword)** (Mtenje 2006)

     (Present Progressive Aspect)

    The leftmost edge of a Morphological Word in the Present Progressive Aspect   is aligned with a high tone.

According to Mtenje (2006), the spreading high tone, like in many Bantu languages, is restricted not to occur where the affected TBU is in the final metrical foot (i.e. the last two syllables). He then proposes two constraints namely, **Spread**, requiring the tone to spread, and

Don't Spread, prohibiting tone from spreading to the right foot structure. Observations (d) and (e) can be taken together. Mtenje (2006) invokes a prosodic constituent called **Prosodic Stem** which is required to be minimally disyllabic. According to Mtenje, **-tu-** insertion ensures that this requirement is satisfied.

Mtenje's (2006) alignment theory manages to capture most of the facts about tone assignment in the present progressive aspect. However, his analysis is not without problems. Just like NONFINALITY, Mtenje's constraint **Don't Spread** is not informative enough: It does not, for instance, reflect the prominence associated with the heavy penultimate syllable which is believed to be the Head of the Phonological Phrase and one which resists neutralization. Mtenje (2006) would also not account for the fact that when an object marker is introduced in a normal present progressive aspect verb complex, the epenthetic **-tu-** surfaces again, but now before the object marker (underlined) as given in (11) below.

(11) **-tu-** *also surfaces before object markers*

Simple Past   *English gloss*

ndi-.tú-.<u>mú</u>-.[cʰi.mbí.l-ií.-ʲa   I am running after him
I-prog-1OM-run-appl-fv

ndi-.tú-.<u>ví</u>-.[sa.mbí.l-ií.-ʲa   I am learning for them
I-prog-8OM-learn-appl-fv

ti-.tú-.<u>ví</u>-.[to.ndé.k-aá.n-a   we are each other failing them
we-prog-8OM-fail-rec-fv

ndi-.tú-.<u>mú</u>-.[kʰú.mbií.l-a   I am admiring him
I-prog-1OM-admire-fv

In a footnote, Mtenje (2006) confesses that "the presence of the epenthetic **-tu-** in the long verbs and in monosyllabic verbs… is rather unexpected". Thus, an explanation is needed to account for observation (a) where the present progressive aspect is expressed without a morphological marker, as well as explanation of the presence of **-tu-** before both monomoraic verb stems and object markers. We also need to know why the prefix **-tu-** assigns the high tone to itself while the rest assign it to the well known Prosodic Stem.

Summary

This section has reviewed Mtenje's (1994/95, 2006) alignment-theoretic analysis of tone assignment in basic verbs, simple past tense verbs and present progressive aspect verbs of ciTonga. The observation in **basic verbs** was that the high tone is usually on the penultimate mora. His analysis is that the high tone is assigned to the rightmost syllable, but it falls on the penultimate syllable or mora due to the principle of NONFINALITY. One problem with this type of analysis is that constraints such as NONFINALITY and RIGHTMOST do not reflect the power shift that occurs between the penult and the final syllables as Prosodic Words/Stems appear in phrase-medial or phrase-final positions.

Coming to the **simple past tense**, one observation was that the high tone is assigned to the initial syllable or mora when the Morphological Stem involved has two or three syllables. Mtenje's analysis has been that there is an alignment constraint ALIGN-L (MSTEM, H), which requires the high tone in the simple past tense to align with the left edge of the Morphological Stem. One problem with this type of analysis is that it is based on inadequate data: The claim that the high tone is exclusively a property of the initial syllable of the Morphological Stem is not correct because the high tone falls on the second mora of the Morphological Stem when it (Morphological Stem) has four or more syllables. The puzzle we need to solve is that the domain of the high tone assigned by the simple past tense is the initial or initial two syllables/moras of the Morphological Stem, depending on the number of syllables therein. What is so special about this constituent? Another feature of tone in the simple past tense was tone spreading. This has been discussed together with observations in the **present progressive aspect**. In the present progressive aspect, the left edge of the Morphological Word (rather than Morphological Stem) is aligned with the high tone and a constraint responsible for this status is ALIGN-L (Mword, H). Tone spreading to following moras is accounted for by the constraint SPREAD. Another constraint DON'T SPREAD ensures that the tone does not spread into the right foot structure (last two syllables). Another crucial observation was that a morpheme **-tu-** is inserted before monomoraic verb stems and that this entity assigns the high tone to itself. Mtenje's analysis of **-tu-** insertion has been that it ensures that a constituent Prosodic Stem satisfies a minimal requirement of containing two syllables. The problems with these types of analyses include the postulation of constraints which are not informative enough of the power relations obtaining between categories (e.g. syllables). Mtenje would also not adequately account for the fact that the morpheme **-tu-** also

surfaces before object markers in this language and that it usually assigns the high tone to itself. What is so special about the morpheme **-tu-**?

4.4. Autosegmental accent analysis[30]

Within Autosegmental Phonology Theory tone and segmental material are represented in parallel tiers (Goldsmith 1976, and extended by himself and others). The notion of "Well-formedness Condition" ensures that all vowels are associated with at least one tone and that all tones are associated with at least one vowel. Association lines linking elements of the two tiers do not cross. The linkage of the levels is facilitated by a more general principle, the Universal Association Convention which states that "when unassociated vowels and tones appear on the same side of an association, they will be automatically associated in a one-to-one, radiating outward from the association line" (Goldsmith 1990:14). Association Conventions add or delete association lines in order to correct any representation that deviates from the neutral state defined by the Well-formedness Condition (Clements and Goldsmith 1984, citing Goldsmith 1979, among others).

Thus in a purely tonal language, such as the Hypothetical language given in (12) below, tones and vowel segments are supposed to be associated from left to right (12a). The surface form in (12b) satisfies the WFC in the sense that all the three tones are associated with at least a vowel and the remaining two vowels associate with the nearest Low Tone to ensure that every vowel is associated with a tone.

---

[30] Mtenje (2006) recognizes the fact that ciTonga is an accent language, but like many tonologists, he downplays it. Throughout his paper, he dedicates only one sentence to this issue: "In some studies of Bantu tone (cf. for instance Odden 1988; Hyman & Byarushengo 1984; Goldsmith 1984, among others), languages with one underlying high tone which is restricted to a predictable position like the one noted here have been referred to as predictable or *accentual systems*."

(12) *Association of tone and vowels in a Hypothetical language:*

sekupafana 'scratch'

    a. Input

```
se  ku pa fa  na        segmental tier
|   | |
L   H L                 Tone tier
```

    b. Output

```
se  ku pa fa  na        segmental tier
|   | |  ⟍  ⟍
|   | |   ⟍  ⟍
L   H L                 Tone tier
```

Languages using such a procedure include all those whose underlying tone melodies are said to be unrestricted, "**pure**" or "**true**" **tone languages** (Goldsmith 1984, Clements and Goldsmith 1984).

Some tonal languages of the world, however, do not fit this description because their surface tones are predicted on the basis of diacritic markings which designate certain vowels and tones as "**determinant**" and "**restricted**" as opposed to others which are unmarked or "neutral". According to this view, the diacritic marks are used to identify determinant vowels in a manner similar to the traditional notion of accent, except that they are purely abstract features, "permitting the prediction of observed tone patterns but having no intrinsic phonetic properties themselves" (Clements and Goldsmith 1984, and references cited therein). Languages in which surface tonal contours are determined at least partly by accentual information present in lexical representations are said to be "**pure**" **accent languages**. In an autosegmental approach, therefore, some languages were recognized to have initial tone association rules which assigned specific tones to specific tone-bearing units prior to the implementation of the WFC (Odden 1984, Kisseberth 1984, and Hyman and Byarushengo 1984).

Turning to ciTonga, it is clear in our data that high-toned basic and tensed verbs have HL and LHL Tone Melodies and the high tone appears to be pre-linked to an accent vowel or mora. In this case, ciTonga partly fits Clements' and Goldsmith's (1984) definition of tone-accent language where a "fixed Basic Tone Melody" can be specified for all words and in which "the

melody is partly assigned through a non-tonal element - *accent*- which 'sets the stage' for the tonal derivation to follow" (Clements and Goldsmith 1984)."

The distinction between tonal and accent languages, however, has been downplayed by many authors including by Goldsmith (1984) himself. Odden (1995:467) repeats this doubt: "…and if there is such a distinction, what the criteria are for treating a language tonally versus accentually – in fact there are languages which have been analysed as tonal and accentual by the same authors." The tone-accent typology is also avoided by many others (e.g. Downing 2004). Facing the facts, however, Goldsmith's accent theory has not been challenged in its fundamentals.

To sum up, this section has reviewed an accentual approach to the study of tone (Goldsmith 1984, Clements and Goldsmith 1984) in Bantu languages and it shows that there has been commitment to the study of the interaction between tone and prominent positions in Bantu languages. CiTonga partly satisfies the definition of an accent language since a fixed tone melody is specified for some (not all) words and in which "the melody is partly assigned through a non-tonal element of accent which sets the stage for the tonal derivation to follow" (Clements and Goldsmith 1984). The goal of the present study therefore is to understand better the fundamental abstract features referred to by Clements and Goldsmith (1984), and references cited therein, which "permit the prediction of observed tone patterns but having no intrinsic phonetic properties themselves."

4.5   Accentual Properties of Tone (Pitch-Accent)

In her 2004 survey of prosody in African languages, Downing focuses on "accentual" properties of Bantu tone systems (ignoring the more tonal properties) where she gives hints on a pitch-accent constituent analysis. It is important to bear in mind that Downing uses the term pitch-accent "to refer to particular stress-like properties of tone realization … culminativity, positional restrictions and tone-(stress) accent interactions – and not to any particular theory of *accent*" as the present study does. The concept of pitch-accent and language typology that follows it has also been doubted in the literature (see, for instance, Hyman 2007, van der Hulst 2011, etc) and I do not get distracted by these debates in this discussion.

According to Downing (2004), *culminativity* is a distinctive characteristic of stress-accent. Similarly, she notes that there are culminativity restrictions on High Tones within Stems or words that make Bantu tone systems resemble accent (stress-accent) systems. Just as in stress systems where only one main stress is found per word or stem, in many Bantu languages only

one High Tone is found per word or stem. In many Bantu languages (Philippson 1998), noun stems also have only one High Tone per stem. Furthermore, in some languages sequences of High Tones are eliminated following what is commonly known as Meeussen's Rule (Goldsmith 1984, Clements and Goldsmith 1984), a process which deletes all but one High Tone in a sequence. Following Cassimjee (1998), Cassimjee and Kisseberth (1998) and Kisseberth and Odden (2003), Downing (2004) argues that Meeussen's rule is one of a family of processes found in Bantu languages which are motivated by the Obligatory Contour Principle (OCP).

In ciTonga, however, Prosodic Words or Stems may have one or more tones as given in (13) below.

(13)   *Prosodic Words/Stems may have more than one tone*

Verb word   *English gloss*

[kʰu.mbi.l-aá.n-a   admire each other
Admire-rec-fv

ti-.ŋgu-.[sá.mbii.z-a   we taught
we-past-teach-fv

ndi-.ŋgu-.[βe.lé.βé.t-ee.s-a   I caused to speak
I-past-speak-caus-fv

ti-.ŋgu-.[to.ndé.k-aá.n-a   we failed each other
we-past-fail-rec-fv

ti-.ŋgu-.[kʰu.mbí.l-aá.n-a   we admired each other
we-past-admire-rec-fv

ti-.tú-.mú-.[kʰumbíl-aán-a   we are each other admiring him
we-prog-1OM-admire-rec-fv

ndi-.tú-.ká-.[sa.mbí.z-ií.j-a   I am aiding it to solve a clue
I-prog-12OM-teach-appl-fv

Tone in ciTonga and many other Bantu languages can also not be metrical because, as Downing (2004) also observes, most Bantu languages are imperfectly culminative since they

preserve a contrast between stems which have a High tone and those which are toneless. The questions requiring our attention therefore are: At what level is tone culminative in ciTonga? Which principles determine syllables to be linked with the tone first before it spreads rightwards to next tone bearing units?

Positional restriction is another property of stress: According to Downing (2004), main stress, cross-linguistically, tends to occur on syllables at the edge of a stem or word. In many Bantu languages, she argues, the target for tone spread or shift is the penult syllable. Like we saw in Mtenje's analysis of penultimate tone in ciTonga, Downing argues that one reason why the penultimate syllable is a likely target of tone spread or shift is that high tone realization is subject to the metrical principle of NONFINALITY. That is, the primary target is the ultimate syllable, but it is avoided, then the rightmost or the only high tone will surface on the penult. Downing (2004) also observes that besides the penult, the stem-initial is also a common target for high tone shift or spread. She gives an example of Giryama, where if the verb word contains two high tones, the rightmost one shifts to the penult while the leftmost shifts to the initial syllable.

Similarly in ciTonga, tones are restricted to the penultimate syllable/mora when the Word or Stem is in phrase-final position (e.g. [cʰi.mbi.lií.ja 'run after'). However, in phrase-medial position, the tone falls on the final syllable (e.g. [cʰi.mbi.li.já mu.nthi.kaa.ze 'run after a woman'). In most tensed verbs, the high tone falls on the second syllable when the Morphological Stem has four or more syllables (e.g. **ndi-.ŋgu-.[βe.lé.βee.ta** 'I spoke'). It falls on stem-initial syllable when the Morphological Stem has two or three syllables (e.g. **ndi-.ŋgu-.[βé.lee.nga** 'I read', **ndi-.ŋgu-.[víi.n-a** 'I danced'). In some cases, however, tone falls on the prefix string such as the present progressive aspect marker **-tu-** (e.g. **ti-.tú-.mú-.[kʰu.mbí.l-aá.n-a** 'we are each other admiring him'). It is evident that most of these positions where tone starts from are not known to have stress and the question is: What attracts tone in these positions? What is it with the first two syllables and the last two syllables of the Prosodic Stem that has affinity to tone? What is in the present progressive aspect marker **-tu-** that attracts tone?

According to Downing (2004), one other property of Bantu tone languages which lends them an accentual character is *the interaction of tone with stress-accent* (Downing 2004; see also de Lacy 2002b). She argues that the penultimate syllable is the most common locus for stress-accent in Bantu languages. In languages where tone shifts to the penultimate syllable, the

114

penult is also found to have stress-accent. Following Philippson (1998), Downing (2004) argues that one motivation for High Tone shift to the penult in some languages is that the High Tone is "attracted" to the stressed syllable. What is really strange, however, is the fact that the high tone can be so close to this so called word-penultimate stress, yet the stress cannot attract the tone which is right on its back as illustrated in (14) below.

(14) *Penultimate stress cannot attract tone of the preceding syllable*

Tensed verb	*English gloss*
ndi-.ŋgu-.[ká.ˈmbuu.l-a	I shouted
I-past-shout-fv	
ndi-.ŋgu-.[sá.ˈmbii.z-a	I taught
I-past-teach-fv	
ndi-.ŋgu-.[dá.ˈnii.k-a	I answered
I-past-answer-fv	
ti-.ŋgu-.[βe.lé.ˈβee.t-a	we spoke
I-past-speak	
ti-.ŋgu-.[βe.lé.βé.t-é.ˈs-aa.n-a	we caused to speak to each other
I-past-speak-caus-rec-fv	
ti-.ŋgu-.[βe.lé.ˈŋg-aan-a	we counted each other
we-past-count-rec-fv	

This is not to doubt the fact that there is stress on the penultimate syllable. But it belongs to the Phonological Phrase and the tone itself does not shift with it when its host is in phrase-medial position as penultimate lengthening does. The problem question is: what is it that is in the initial or the first two syllables of the Morphological Stem that hordes the tone and which in some cases prevents the 'stressed' penultimate syllable from getting this tone which is right on its back?

Summary

Three accentual properties of Bantu tone which are discussed by Downing (2004) include culminativity, positional restrictions, and tone-stress interactions. Much as tone in ciTonga

exhibits all of these properties, there are some challenges which would militate against a purely stress approach to the study of tone. For instance, a single Word in ciTonga may have two or three tones which obviate other problems: At what level is tone culminative in this language? Why is the high tone pre-linked to specific syllables and not others, not even those believed to have stress? Tone spreading also makes it skeptical that tone is indeed an important cue for stress. Furthermore, tone in this language cannot be metrical because it is imperfectly culminative since not all words have tone.

Tone is also not restricted to the stem-initial and penultimate syllables as presented in Downing's accentual properties of tone. In ciTonga, tone can be on a second mora of a Stem with four or more syllables and it can also fall on a final syllable when the Stem is in a phrase-medial position. Some prefixes such as the Present Progressive Aspect marker **-tu-** assign the high tone to themselves as well. These observations obviate other problems: What is it with the initial or the initial two syllables and the last two syllables of Stems that has affinity to the high tone? What is in the present progressive aspect marker **-tu-** that makes it keep the high tone to itself? What is so strange is the fact that a high tone can be just at the back of the so-called stressed penultimate syllable, yet it (the stressed syllable) cannot attract this tone. What is it that is in the preceding syllable which hordes the tone and prevents it from falling on a more salient penultimate syllable? Indeed, answers to all these problems seem to lie in the idea of the Interaction between Tone and Strong Accent Constituents, and not between tone and stress-accent per se.

4.6 The Interaction between Tone and Strong Accent Constituents

The contribution that this chapter makes is that tones belong to Strong Accent Constituents namely, PROSODIC STEM and ACCENT FOOT, which have been firmly established in the preceding chapter. A constraint **TONE/SAC**, requiring tones to belong to Strong Accent Constituents ensures this state of things. In addition, tone in ciTonga seems to be culminative at Accent Foot level. The possibility is that tone is culminative at Prosodic Stem level in other languages where only a single high tone is allowed per Prosodic Stem. Unless otherwise stated, my focus is on formal and common speech styles.

(15)   **TONE/SAC**

Tone belongs to Strong Accent Constituents.

I will begin my analysis with tone assignment in verb stems (without prefixes).

4.6.1   Tone in basic verbs (without prefixes)

The observation about tone assignment in basic verbs was that the High Tone is usually assigned to the penultimate mora of the Morphological Stem as given in (16) below.

(16)   *Tone falls on the penultimate mora*

High-toned basic verb	*English gloss*
[to.nde.k-a.n-aá.ŋg-a fail-rec-hab-fv	be failing each other
[cʰi.mbi.l-i.l-a.n-aá.ŋg-a run-appl-rec-hab-fv	be running after each other
[kʰu.mbi.l-aá.n-a admire-rec-fv	admire each other
[kʰu.mbi.l-a.n-ií.ŋg-e admire-indic-hon-hab-subj	be admiring
[cʰi.mbi.l-i.l-a.n-aá.ŋg-a run-appl-rec-hab-fv	be running after each other
[sa.mbi.z-i.l-a.n-aá.ŋg-a teach-appl-rec-hab-fv	be aiding each other in solving a clue

When these verbs appear in phrase-medial positions, tone is realized on the final mora as given in (17) below.

(17) *In phrase-medial positions, tone is realized on the final mora*

[to.nde.k-a.n-a.ŋg-á] [ma-jeeso]

[fail-rec-hab-fv]   [6-exam]

"be failing each other examinations"

[cʰi.mbi.l-i.l-a.n-a.ŋg-á]   [ukóoŋgwa]

[run-appl-rec-hab-fv]   [a lot]

"be running after each other a lot"

[kʰu.mbi.l-a.n-á]   [ma-laája]

[admire-rec-fv]   [6-cloth]

"admire each other's clothes"

The suggested analysis would be that tone in basic verbs is assigned to a Strong Accent Constituent within the Prosodic Stem as guided by Universal Principles of EDGENESS and FINALITY in more formal speech styles. EDGENESS provides that edgemost moras are stronger than non-edgemost moras. The Guideline of FINALITY provides that non-final moras are stronger than final moras. The PENULTIMATE-μ in phrase-final verbs above better follows these Guidelines and it is thus a SAC. This analysis is formalized in (18) below.

(18)   *Tone falls on the penultimate mora*

High-toned basic verb   *English gloss*

{**to.nde.k-a.n-aá.ŋg-a**}   be failing each other
fail-rec-hab-fv

{**cʰi.mbi.l-i.l-a.n-aá.ŋg-a**}   be running after each other
run-appl-rec-hab-fv

{**kʰu.mbi.l-aá.n-a**}   admire each other
admire-rec-fv

{**kʰu.mbi.l-a.n-ií.ŋg-e**}   be admiring each other
admire-indic-hon-hab-subj

{**cʰi.mbi.l-i.l-a.n-aá.ŋg-a**}   be running after each other
[run-appl-rec-hab-fv]

{**sa.mbi.z-i.l-a.n-aá.ŋg-a**}   be aiding each other in solving a clue
teach-appl-rec-hab-fv

A crucial generalization then is that tones belong to Strong Accent Constituents such as the PENULTIMATE-μ of Prosodic Stems which are found in phrase-final positions. The constraint **TONE/SAC** would ensure this status. A tableau in (19) below illustrates.

(19) *EDGENESS and FINALITY provide PENULTIMATE-µ as a SAC*

SACs: PROSODIC STEM; PENULTIMATE-µ

DGs: PROSODIC HIERARCHY; EDGENESS&FINALITY

/to.nde.ka.na.ŋga/ H	TONE/SAC
(a) {to.nde.k-a.n-aa.ŋg-á}	*!
(b) {to.nde.k-a.n-áa.ŋg-a}	*!
☞ (c) {to.nde.k-a.n-aá.ŋg-a}	

Candidates (19a&b) are non-optimal because they violate the constraint TONE/SAC which requires high tones to fall on Strong Accent Constituents such as the penultimate mora of the Prosodic Stem as guided by the universal principles of EDGENESS&FINALITY. On the other hand, candidate (19c) wins because it satisfies TONE/SAC.

4.6.2 Tone in the simple past tense verbs

When toneless basic verbs appear in tensed verbs, tone is assigned to one of the first two moras of the Morphological Stem. The observation in the simple past tense verbs was that the high tone is assigned to the initial mora of Morphological Stems with two or three syllables (e.g. **ndi-.ŋgu-.[lé.lee.s-a**). However, the high tone is realized on the second syllable of Morphological Stems with four or more syllables. The tone then spreads if conditions permit, but it does not spread to the last two syllables (e.g. **ti-.ŋgu-.[tʰa.mbá.lí.l-aa.n-a**).

My suggested analysis of tone assignment in the simple past tense is that tone is assigned to the initial SAC (or initial Accent Foot) of a Prosodic Stem as illustrated in (20) below.

120

(20) *Tone is assigned to the left Accent Foot in the simple past tense*

Past tense verb	*English gloss*
ndi-.ŋgu-.{(lé).le(e.s-a)} I-past-look-fv	I looked
ndi-.ŋgu-.{(βé).le(e.ŋg-a)} I-past-read-fv	I read
ndi-.ŋgu-.{(tú).tu(u.z-a)} I-past-push-fv	I pushed
ndi-.ŋgu-.{(βe.lé.)βe(e.t-a)} I-past-speak-fv	I spoke
ndi-.ŋgu-.{(βe.lé.)βé.t-e(e.s-a)} I-past-speak-caus-fv	I caused to speak
ti-.ŋgu-.{(vu.mbá).tí.l-a(a.n-a)} we-past-hug-rec-fv	we hugged each other
ti-.ŋgu-.{(tʰa.mbá).lí.l-a(a.n-a)} we-past-stretch-rec-fv	we stretched legs over each other's

What is surprising, however, is the fact that the Accent Foot which is assigned tone is not the rightmost one (as would be guided by the principle of RIGHTEDGE in this language, which provides that the right edge of a Prosodic Stem is stronger than the left edge). This contradiction with tone assignment in basic verbs can be understood better when we consider the fact that tones play a morphological role. In this case, a high tone assigned to the initial Accent Foot marks an important boundary between the prefix string and some Strong Accent Constituent, the Stem. In other words, *Morphology wins over Phonology*. What is crucial is the fact that the constraint TONE/SAC is still satisfied since tone falls on Strong Accent Constituents PROSODIC STEM, STEM-σ_1 and ACCENT FOOT, following Universal Guidelines of EDGENESS and PROSODIC HIERARCHY. The Guideline of FINALITY, which provides that non-final syllables/moras are stronger than final syllables/moras, is also dully followed. A tableau in (21) below illustrates.

(21) *Tone belongs to initial SAC in simple past tense verbs*

SACs: PROSODIC STEM &ACCENT FOOT; STEM-σ_1

DGs: PROSODIC HIERARCHY; EDGENESS

/ndi-ŋgu-tutuz-a/ H	TONE/SAC
(a) ndi-.ŋgú-.{**(tu)**.tu(u.z-a)}	*!
(b) ndi-.ŋgu-.{**(tu)**.tú(u.z-a)}	*!
☞ (c) ndi-.ŋgu-.{**(tú)**.tu(u.z-a)}	

Candidates (21a&b) are disqualified because they violate the constraint TONE/SAC which requires tone to belong to Strong Accent Constituents such as the PROSODIC STEM, ACCENT FOOT and STEM-σ_1 as guided by universal principles of EDGENESS and PROSODIC HIERARCHY. Candidate (21c), on the other hand, wins because it satisfies the constraint TONE/SAC[31].

The fact that tone is realized on a rightmost mora when the left Accent Foot is bimoraic, such as when the Stem involved has four or more syllables, would indicate that the Universal Guideline of RIGHTEDGE still has some influence concerning the preferred edge for prosodic features such as tone [e.g. **ndi-.ŋgu-.{(βe.lé.)βe(e.t-a)}**, *****ndi-.ŋgu-.{(βé.le.)βe(e.t-a)}**]. We also need to revise our analysis of tone assignment in the basic high-toned verbs to take into account the fact that even the right edge tone, much as it is assigned to the Prosodic Stem, it still belongs to an Accent Foot of a kind firmly established in the preceding chapter. This fact is illustrated in (22) below.

---

[31] A higher-ranked morphological constraint would ensure that the tone assigned by the simple past tense falls on the left and not right SAC or (monomoraic) Accent Foot [e.g. *****ndi-.ŋgu-.{(tu).tu(ú.z-a)}**].

(22)   *Right edge tone also belongs to an Accent Foot of some kind*

High-toned basic verb   *English gloss*

{(to.nde).k-a.n-a(á.ŋg-a)}   be failing each other
fail-rec-hab-fv

{(cʰi.mbi).l-i.l-a.n-a(á.ŋg-a)}   be running after each other
run-appl-rec-hab-fv

{(kʰu.mbi).l-a(á.n-a)}   admire each other
admire-rec-fv

{(kʰu.mbi).la-.ni(í-.ŋg-e)}   be admiring
admire-indic-honor-hab

{(cʰi.mbi).l-i.l-a.n-a(á.ŋg-a)}   be running after each other
run-appl-rec-hab-fv

{(sa.mbi).z-i.l-a.n-a(á.ŋg-a)}   be aiding each other in solving a clue
teach-appl-rec-hab-fv

The rest of the facts about Guidelines such as EDGENESS and FINALITY which determine the penultimate mora as a Strong Accent Constituent remain intact.

4.6.2.1 Tone spreading

Once the tone is linked to an appropriate Accent Foot, it has to spread. I follow Mtenje (2006) in using the constraint **SPREAD** which requires tones to spread rightwards to following moras as a tableau in (24) below illustrates. Note that although tone spreads outside the foot structure, it does not violate TONE/SAC because the spreading tone is still within the boundaries of the PROSODIC STEM which is the Minimal Strong Accent Constituent in formal and common speech styles of ciTonga. This observation makes sense because in elderly speech styles, where the SAC is the Accent Foot, tone would not spread outside the foot structure [e.g. formal **ta-.{βe.lé.βé.t-é.s-á.n-aa.ŋg-a}** ⇒ **ta-.{(βé.l).β.t-.s-a.n-a(a.ŋg-a)}, *ta-.{(βé.l).β.t-.s-á.n-a(a.ŋg-a)}**].

(23) **SPREAD**

Tones spread (Mtenje 2006).

(24) *Tone must spread*

SACs: PROSODIC STEM &ACCENT FOOT; STEM-σ_2

DGs: PROSODIC HIERARCHY; EDGENESS

/ti-ŋgu- pa kul-il-a n-a/ H	SPREAD	TONE/SAC
(a) ti-ŋgu-{(pa.kú).l-i.l-a(a.n-a)}	****!	
☞(b) ti-ŋgu-{(pa. kú).l-í.l-a (a. n-a)}	***	

Both candidates satisfy the constraint TONE/SAC which requires the tone to be assigned to Strong Accent Constituents such as the left ACCENT FOOT of the constituent PROSODIC STEM (in the simple past tense). However, candidate (24a) is non-optimal because it maximally violates the constraint SPREAD which requires tones to spread. On the other hand, candidate (24b) is successful because it minimally violates SPREAD.

One crucial observation in the winning candidate in the tableau above is that tone does not spread into the long penultimate syllable or the last two syllables. To account for this type of tone blocking, Mtenje (2006) suggests a constraint **Don't Spread**, requiring tones not to spread to the right Foot structure. However, this constraint, as I pointed out earlier on, is not very informative because it does not reflect the power relations between the syllables where tone can spread to and the syllables where neutralization is blocked. It is evident that what is at issue is the fact that the penultimate syllable (believed to be the anchor of stress in many Bantu languages – see Downing 2004) constitutes Head Prosodic Phrase (HPP) as evidenced by penultimate vowel lengthening (PL). A generalization therefore would be that forms which preserve the identity of Head Prosodic Phrases are optimal. I suggest that there is a constraint **IDENT-HEAD PROSODIC PHRASE (IDENT-HPP)** which requires identity of the Head Syllable of the Prosodic Phrase to be preserved (cf. Kager 2007). A tableau in (26) below illustrates.

(25) **IDENT-HEAD PROSODIC PHRASE (IDENT-HPP)**

Identity of the Head Syllable of a Prosodic Phrase (Stressed syllable) is preserved.

(26) *Head Prosodic Phrases resist neutralization*

SACs: PROSODIC STEM&ACCENT FOOT; STEM-σ_1; FINAL-σ

DGs: PROSODIC HIERARCHY; EDGENESS; EDGENESS&SONORITY

/ti-ŋgu-pakul-il-an-a/ H (a) ti-ŋgu-(pa.kú).l-í.**l-á(á**.n-a) ☞(b) ti-ŋgu-(pa.kú).l-í.**l-a(a**.n-a)	IDENT-HPP	SPREAD	TONE/SAC
(a)	*!*	*	
(b)		***	

Although candidate (26a) minimally violates the constraint SPREAD, it is ruled out for violating the high-ranking constraint IDENT-HPP, requiring identity of Head Prosodic Phrases such as the penultimate syllable of a Phonological Phrase to be preserved. Candidate (26b), on the other hand, is successful because it satisfies IDENT-HPP.

4.6.2.2 Tone in tensed disyllabic and monomoraic verbs

Recall that trisyllabic prosodic Stems were assigned two Accent Feet as repeated in (27) below.

(27) *Trisyllabic Prosodic Stems were assigned two Accent Feet*

Tensed Stem   *English gloss*

ndi-.ŋgu-.{(lé).le(e.s-a)}   I looked
I-past-look-fv

ndi-.ŋgu-.{(βé).le(e.ŋg-a)}   I read
I-past-read-fv

ndi-.ŋgu-.{(tú).tu(u.z-a)}   I pushed
I-past-push-fv

The basis of this reasoning is that this kind of footing better explains the presence of tone on the initial syllable which is undoubtedly a Strong Accent Constituent in Bantu languages like ciTonga. It also explains why the tone is not assigned to the long penultimate syllable which is believed to be the bearer of phrase stress. Disyllabic Stems, however, present a big challenge to most of the facts we have presented above. When the simple past tense marker is prefixed to toneless disyllabic verbs, for instance, tone is on the initial mora and outside of an expected Accent Foot. This fact is illustrated in (28) below.

(28) *Tone is on the initial mora of an Accent Foot when disyllabic Stems are involved*

Tensed Stem   *English gloss*

ndi-.ŋgu-.{ví(i.n-a)}   I danced   *ndi-.ngu-.{(vií.)n-a}
I-past-dance-fv

ndi-.ŋgu-.{sá(a.mb-a)}   I bathed
[I-past-bathe-fv]

ndi-.ŋgu-.{tʰá(a.β-a)}   I ran
[I-past-run-fv]

On the other hand, the tone is assigned according to expectation when the Stem involved is monomoraic. The high tone is realized on the penultimate mora as shown in (29) below.

(29) *Tone falls on tense marker **-ngu-** when the verb stem involved is monomoraic*

Tensed Stem	English gloss
ndi-.{ŋgu(ú.-lj-a)}	I ate
I-past-eat-fv	
ndi-.{ŋgu(ú-.mw-a)}	I drunk
I-past-drink-fv	
ndi-.{ŋgu(ú-.fw-a)}	I died
I-past-die-fv	
ndi-.{ŋgu(ú-.b-a)}	I stole
I-past-steal-fv	
ndi-.{ŋgu(ú-.sw-a)}	I broke
I-past-break-fv	
ndi-.{ŋgu(ú-.vw-a)}	I heard
I-past-hear-fv	

The problem is to account for the fact that the tone falls on the initial mora of the Prosodic Stem (penultimate mora) in verbs based on disyllabic Stems (28), while it falls on a second mora of the Prosodic Stem in tensed verbs based on monomoraic Stems (29). To make matters worse, why should the tone in (28) fall on the initial mora of the bimoraic syllable instead of the usual second mora and, most importantly, inside a foot structure of the disyllabic Stem? To solve this puzzle, I would like to propose that in a strong-accent language a single heavy syllable can belong to two different Accent Feet. In the disyllabic verb Stems in (28), for instance, the initial mora constitutes an Accent Foot on its own and the last two moras form an Accent Foot of their own. Thus, the simple past tense assigns the high tone to the left edge Accent Foot as given in (30) below.

(30) *Footing in disyllabic verbs*

Tensed Stem	English gloss
ndi-.ŋgu-.{(ví)(i.n-a)} I-past-dance-fv	I danced
ndi-.ŋgu-.{(sá)(a.mb-a)} I-past-bath-fv	I took a bath
ndi-.ŋgu-.{(tʰá)(a.β-a)} I-past-run-fv	I ran away
ndi-.ŋgu-.{(bí)(i.k-a)} I-past-cook-fv	I cooked
ndi-.ŋgu-.{(sé)(e.ng-a)} I-past-cut-fv	I cut (my hair)

This sounds strange, but it cannot be completely senseless in a language where the mora, and not the syllable, is the unit of foot analysis (and BINARITY is not always satisfied). There is another good reason for the left tone **NOT** to be assigned to the second mora of the penultimate syllable where Stems involved are disyllabic. It appears the penultimate mora is reserved for the *right* tone. Recall that in basic verbs the penultimate mora is the bearer of the right tone. This mora also bears the only tone which is contrastive in this language. These facts are illustrated in (31) below.

(31) *Phonemic tone restricted to the penultimate mora*

Minimal pair	*English gloss*
{dii.k-a}	spill
{dií.k-a}	cover oneself
{bii.j-a}	steal for
{bií.j-a}	cheat
{pʰaa.t-a}	shake off (dust)
{pʰaá.t-a}	be stuck
{da.n-ii.k-a}	answer
{da.n-ií.k-a}	be invited

Thus, in addition to accent-bearing units smaller than the syllable, it should also be a recognized fact that rules of foot construction in a strong accent language may split syllables; The first part of a heavy syllable may belong to one foot while the second part may belong to the next as shown in (30) above. Further research, however, is required to establish whether a single syllable can indeed belong to two different Accent Feet in 'pitch-accent' languages (where 'accent' can be assigned to units smaller than the syllable) as I am suggesting here[32].

Finally, the fact that the high tone does not fall on the initial mora of the past tense marker **-ŋgu-** in verb forms based on monomoraic Stems is an indication that the first mora of the simple past tense marker is not footed. However, it is important to note that **-ŋgu-** provides a proper onset and syllable required by the Prosodic Stem [e.g. **ndi-.{ŋgu(ú.-lja)}** 'I ate'].

4.6.3   Tone in the present progressive aspect verbs

Five observations were made about tone assignment in the present progressive aspect. (a) *Normal* verbs (where morphological stems have two or more syllables) in the present progressive aspect have no morphological marker (e.g. **ndí-.[sá.mbii.z-a** 'I am teaching'). (b) A high tone seems to spread from word-initial syllable to the antepenult syllable. (c) The "spreading" tone does not penetrate the last two syllables. (d) When a Morphological Stem involved is *monomoraic*, however, there appears a morpheme **-tu-** before the stem (e.g. **ndi-.túu-.vwa** 'I am hearing'). (e) The entity **-tu-** assigns the high tone to itself.

---

[32] Splitting of a single syllable into two feet is not permissible in stress languages (see Hayes 1995:50).

My suggested analysis of observation (a) is that **-tu-** is a bonafide morphological marker[33] for the present progressive aspect and that it is lexically marked as a Strong Accent Constituent and automatically a phonological Foot Head. From the foregoing discussions, Foot Heads have consistently appeared inside a Prosodic Stem. The generalization therefore is that Foot Heads belong to the Prosodic Stem. Since **-tu-** is outside the Prosodic Stem structure in the normal Prosodic Stems given in (32) below, it is a violation of the principle of **HEADEDNESS** which requires Accent Feet to be realized within the Prosodic Stem structure. In other words, the morpho-prosodic constituent Prosodic Stem dominates the metrical constituent Accent Foot. If the lexical Accent Foot Head **-tu-** is not properly headed, it must perish.

(32)   *Normal Present Progressive Aspect verbs have no morphological marker*

Verb   *English gloss*

*ndi-.(tú.)-{(le).le(e.s-a)}   I am looking
I-prog-look-fv

*ndi-.(tú.)-{(zo).me(e.l-a)}   I am admitting
I-prog-admit-fv

*ndi-.(tú.)-{(sa).mbi(i.z-a)}   I am teaching
I-prog-teach-fv

*ndi-.(tú.)-{(βe.le).βe(e.t-a)}   I am speaking
I-prog-speak-fv

Deletion of **-tu-**, however, leads to violation of the constraint FAITH-SAC which requires Strong Accent Constituents to be preserved. The fact that the morpheme **-tu-** is eventually deleted indicates that the constraint HEADEDNESS outranks the constraint FAITH-SAC[34].

My analysis of observation (b) is that, since the tone bearing Accent Foot is not realized, tone is assigned according to the Association Convention, associating tones and moras in a one-to-

---

[33] Note that Mtenje (2006) is aware of the fact that **-tu-** could as well be a morphological marker for the Present Progressive Aspect, as opposed to being an epenthetic form. Unlike Mtenje (2006), I am going to treat this process as involving **-tu-** deletion rather than insertion. My conviction is that -tu- is a bonafide morphological marker for the Present Progressive Aspect (and not an epenthetic syllable as Mtenje suggests). The basis for this position is that **-tu-** is confined to particular morphological constructions. Like most tense/aspect markers in this language, **-tu-** in the examples given above consistently appears before an object marker and a (monomoraic) verb stem. It would thus be more sensible to regard **-tu-** as some kind of a prefix which for some reason (unknown at this point) undergoes deletion in the normal Present Progressive Aspect verbs.
[34] Deletion of **-tu-** can also be a violation of the constraint REALIZEMORPHEME which requires segmental properties of a morpheme to be realized in the output (Akinlabi 1996, Walker 2000, cited in Downing 2006:132).

one fashion from left to right (Goldsmith 1976, 1984). Thus the high tone docks on the initial mora of the Prosodic Word from which it spreads rightwards (by SPREAD) from word-initial syllable to the antepenult syllable. The spreading tone, according to observation (c), does not penetrate the last two syllables once again due to the high-ranking constraint IDENT-HPP which requires identity of Head Prosodic Phrases to be preserved. Thus, the tone is able to spread to the stem-initial Accent Foot, but it is blocked from spreading to the second Accent Foot which is housed by the HPP (so to speak). A tableau in (35) below illustrates all these issues. The constraint ASS (TONE) in the tableau ensures that tone is associated with moras.

(33) **HEADEDNESS**

Prosodic constituents are properly headed (cf. Itô and Mester 1992, Orie 1997, Selkirk 1995).

(34) **ASS (TONE)**

Tones are associated with tone bearing units.

(35) *Foot Heads must belong to a Prosodic Stem*

SACs: **-tu-**; PROSODIC STEM; STEM-σ_1; FINAL-σ

DGs: MORPHOLOGY; PROSODIC HIERARCHY; EDGENESS; EDGENESS&SONORITY

/ndi-(tú)-[βeleβet-a/	ASS	HEADED NESS	IDENT-HPP	SPREAD	FAITH-SAC
(a) ndi-(tú.)-{(βé.lé).**βe**(**e**.t-a)}		*!		***	
(b) ndí-{(βé.lé).**βé**(**e**.t-a)}			*!	**	*
☞ (c) ndí-{(βé.lé).**βe**(**e**.t-a)}				***	*

Thus, candidate (35a) is ruled out because it violates a high ranking constraint HEADEDNESS which requires Accent Feet to be properly dominated (i.e. by Prosodic Stems). Candidate (35b) is disqualified since it violates another high-ranking constraint IDENT-HPP which requires identity of Head of the Phonological Phrase to be preserved. On the other hand, candidate (35c) is successful because it satisfies both HEADEDNESS and IDENT-HPP. It

violates low-ranking constraints SPREAD and FAITH-SAC. These violations, however, can be tolerated for the sake of IDENT-HPP and HEADEDNESS.

One other observation about the present progressive aspect verb formation is that when the Morphological Stem involved is monomoraic, the grammar of ciTonga allows a morpheme -**tu-** to appear before it as repeated in (36) below.

(36) *Tu- surfaces before Monomoraic Morphological Stems*

(ndi = 1ˢᵗ pers. Sing.)

Pres. Prog	English gloss
ndi-.túu-.[vw-a	I am listening
ndi-.túu-.[mb-a	I am singing
ndi-.túu-.[lj-a	I am eating
ndi-.túu-.[b-a	I am stealing
ndi-.túu-.[mw-a	I am drinking
ndi-.túu-.[sw-a	I am breaking

I follow Mtenje's (2006) account by invoking a prosodic constituent **Prosodic Stem** which is required to be minimally disyllabic (see also Downing 2006b). The forms in (36) above can therefore be represented as in (37) below.

(37) *Prosodic Stem is minimally disyllabic*

Pres. Prog   *English gloss*

ndi-.{túu-.vw-a}   I am listening

ndi-.{túu-.mb-a}   I am singing

ndi-.{túu-.lj-a}   I am eating

ndi-.{túu-.b-a}   I am stealing

ndi-.{túu-.mw-a}   I am drinking

ndi-.{túu-.sw-a}   I am breaking

To cut the long story short, **-tu-** surfaces before monomoraic verb stems because it is this time around parsed by the Prosodic Stem. Since the high tone falls on the initial syllable of the Prosodic Stem (i.e. on **-tu-** itself), the left edge Accent Foot must be monomoraic. This fact is illustrated in (38) below.

(38) ***-tu-*** *before monomoraic verb stems is parsed by Prosodic Stem*

ndi-.{(tú)(u-.vw-a)}   I am listening

ndi-.{(tú)(u-.mb-a)}   I am singing

ndi-.{(tú)(u-.lj-a)}   I am eating

ndi-.{(tú)(u-.b-a)}   I am stealing

ndi-.{(tú)(u-.mw-a)}   I am drinking

 ndi-.{(tú)(u-.sw-a)}   I am breaking

This analysis shows that the constraint HEADEDNESS is now satisfied since the foothead (tu) is now properly dominated by the Prosodic Stem. A tableau in (39) below illustrates.

(39) *-tu- is not deleted when* HEADEDNESS *is satisfied*

/ndi-**tú**-lj-a/	HEADEDNESS
☞ (a) ndi-.{(tú)(u-.lj-a)}	
(b) ndi-.(tú){(u-.lj-a)}	*!

Thus, candidate (22a), where **-tu-** surfaces before monomoraic verb Stems, is acceptable because the Foot Head is now properly dominated by the Prosodic Stem. On the other hand, candidate (39b) is ruled out because the foothead (tu) falls outside the Prosodic Stem. In the following chapter, we will look at further cases of blocking of general morpheme deletion before monomoraic verb stems and object markers. I argue there that this would be an indication that both verb stems and the object markers are parsed by Prosodic Stem.

Summary

This section has presented a Strong Accent Constituent Analysis of tone assignment in basic verbs, the simple past tense and the present progressive aspect verb complexes. The main point made is that tone belongs to Strong Accent Constituents. This generalization has been accounted for in terms of the constraint TONE/SAC which requires high tones to be assigned to Strong ACs. To account for tone-spreading, I have followed Mtenje's (2006) analysis where the constraint SPREAD requires high tones to spread to next TBUs. Blocking of tone spreading by the last two syllables of the Prosodic Stem/Word/Phrase has been attributed to another constraint IDENT-HPP which requires identity of Head Prosodic Phrases (penultimate syllables) to be preserved. Deletion of the present progressive aspect marker **-tu-** before normal verb Stems (with two or more syllables) has been attributed to the constraint HEADEDNESS outranking FAITH-SAC. The former constraint requires Accent Feet (SACs) to belong to the Prosodic Stem. Once deleted, the tone which belonged to foot head **-tu-** is re-assigned according to rules of the Universal Association Convention, requiring tones to be assigned from left to right in a one-to-one fashion. The fact that **-tu-** surfaces before monomoraic verbs would be indication that it is now parsed by Prosodic Stem and it no longer violates HEADEDNESS.

4.7   Chapter summary

The aim of this chapter was to present formally a theory of Interaction between Tone and Strong Accent Constituents. I began this chapter by presenting the data on tone assignment in basic verbs, simple past tense verbs and present progressive aspect verbs which motivate this type of theory analysis. Attempt was then made to account for the facts in terms of Tone Alignment Theory (as argued for by Mtenje 2006), autosegmental accent (Goldsmith 1984, Clements and Goldsmith 1984) and pitch-accent (as hinted upon by Downing 2004). All these theoretical perspectives have been found to be slightly problematic to account for tone distribution patterns in ciTonga. On the other hand, a theory based on Interaction between Tone and Strong Accent Constituents has been shown to account for the facts slightly better.

Chapter 5

Strong Accent Constituent-Based Templates

5.1   Introduction

This chapter presents formally proposals for Strong Accent Constituent-Based Templates as a slightly better way to account for morphology-prosody interfaces in ciTonga and perhaps many other Bantu languages. The chapter is organized as follows: Section 5.2 presents the data on General Phonological Words, Minimal Prosodic Words and Reduplicative Prosodic Stems. Section 5.3 reviews the Generalized Templates Theories namely, the Prosodic Hierarchy-Based Templates Theory and the Morpheme-Based Templates Theory both of which have a goal to account for morphology-prosody interfaces. Section 5.4 presents an MBT analysis of the facts. Like PBT, MBT has been found to be slightly inadequate to account for parameters exhibited by Phonological Words in ciTonga. Section 5.5 presents the theory and it analyses the data in terms of Strong Accent Constituent-Based Templates. Section 5.6 presents chapter summary.

5.2   The data

I am going to present the data on general Prosodic words, Minimal Prosodic Words and Reduplicative Morphemes.

5.2.1   General Phonological Words

Like many Bantu languages, a Phonological Word in ciTonga can be roughly considered as a Morphological Stem (i.e. Root + Suffixes) as given in (1) below. Once again, I have used the symbol "[" to indicate that the constituent to its right is a Morphological Stem (without prefixes).

(1)   *Phonological Words may consist of Morphological Stem alone*

Verb   *English gloss*

[bii.k-a
cook-fv   cook

[be.nee.k-a
cover-fv   cover

[βe.lee.ŋg-a
read-fv   read

[be.ne.k-aa.n-a
cover-rec-fv   cover each other

[bi.k-i.l-aa.n-a
cook-appl-rec-fv   cook for each other

[βe.le.ŋg-e.s-aa.n-a
read-caus-rec-fv   cause each other to read

A Phonological Word can also roughly be a Morphological Word with two important parts namely, the **Prefix String** and the **Morphological Stem**. Among others, the prefix string may comprise of the subject agreement marker, the tense/aspect marker and the object marker, in that order. These facts are exemplified in (2) below.

(2)   *Phonological Words may consist of Morphological Stem and Prefixes*

Verb   *English gloss*

ndi-.ŋgu-mu-.[bíi.k-a   I cooked him
I-past-1OM-cook-fv

ndi-.ŋgu-.vi-.[βé.lee.ŋg-a   I read them
I-past-8OM-read-fv

ndi-.ŋgu-.βe.lé.βee.t-a   I spoke
I-past-speak-fv

ti-.ŋgu-.[βe.lé.ŋg-é.s-aa.n-a   we caused each other to read
we-past-read-caus-rec-fv

5.2.2   Minimal Prosodic Words

In cases where the Morphological Stem is monomoraic (e.g. **-ba** 'steal', **-lja** 'eat', **-swa** 'break', **-fwa** 'die' and **-mwa** 'drink'), there are two ways of achieving the Minimal Prosodic Word, depending on speech styles. In formal and common speech styles an epenthetic vowel [i] is added as a prefix to the monomoraic verb stem. This fact is illustrated in (3) below.

(3) *Minimal Prosodic Words are disyllabic in formal and common speech styles*

Verb   *English gloss*

ii.-[b-a   steal!   [*ba]

ii.-[lj-a   eat!

ii.-[sw-a   break!

ii.-[fw-a   die!

ii.-[mw-a   drink!

In elderly speech styles, however, word minimality is not achieved through vowel insertion. Rather, the stem vowel is long (bimoraic). In other words, the Minimal Prosodic Word is bimoraic[35]. This fact is illustrated in (4) below.

---

[35] The minimal word is bimoraic even in phrase medial positions (e.g. **[lj-aa ukóoŋgwa** 'eat a lot', ***lj-a ukóoŋgwa**). This is indication that the length is not just mere phonological phrasing.

(4) *A Minimal Word is bimoraic in elderly speech style*

Verb	*English gloss*
[z-aa	come
[lj-aa	eat
[mb-aa	sing
[mw-aa	drink
[vw-aa	hear

5.2.3 Reduplication

Another important prosodic feature in Bantu languages is reduplication. In ciTonga, the reduplicative morpheme can be a total copy of the Base Stem, or just partial, again depending on speech styles. In formal and common speech styles reduplication involves repetition of a whole Morphological Stem as given in (5) below[36].

---

[36] The reduplicative morpheme is prefixed to the base. Evidence for this analysis is based on the fact that in partial reduplication found in elderly speech styles, it is the first chunk which is reduced.

(5)   *Reduplication of an entire Morphological Stem*

Reduplication   *English gloss*

ndi-.ŋgu-.ví.ná.-[vii.n-a   I danced a lot
I-past-RED-[dance-fv

ndi-.ŋgu-.kú.mbú.ká.-[ku.mbuu.k-a   I remembered a lot
I-past-RED-[dance-fv

ndi-.ŋgu-.βé.lé.ŋgá.-[βe.lee.ŋg-a   I read a lot
I-past-RED-[read-fv

ndi-.ŋgu-.vu.nú.lí.já-[vu.nu.lii.-ʲa   I steamily myself a lot
I-past-RED-[steam-fv

ndi-.ŋgu-.βe.lé.βé.tá-[βe.le.βee.t-a   I spoke a lot
I-past-RED-[speak-fv

ti-.za.mu-.le.lé.sá.ná-[le.le.s-aa.n-a   we will look at each other a lot
we-dist fut-RED-[look-rec-fv

ti-.za.mu-.zo.mé.lé.zá.ná-[zo.me.le.z-aa.n-a   we will agree with each other a lot
we-dist fut-RED-[admit.rec-fv

ti-.ŋgu-.tʰa.mbá.lí.lá.ná-[tʰa.mba.li.l-aa.n-a   we stretched legs over each other a lot
we-past-RED-[stretch-rec-fv

In elderly speech styles, however, the reduplicative prosodic morpheme copies only two initial syllables of the Morphological Stem as given in (6) below[37].

---

[37] Unless otherwise indicated, I do not delete vowels in elderly speech styles in this chapter since this is an optional process and there is no elderly person who exclusively speaks like this.

(6)   *Partial reduplication in common elderly speech style*

Reduplication   *English gloss*

ndi-.ŋgu-.ⱱí.ná.-[vii.n-a   I danced a lot
I-past-RED-[dance-fv

ndi-.ŋgu-.kú.mbú.-[ku.mbuu.k-a   I remembered a lot
I-past-RED-[dance-fv

ndi-.ŋgu-.βé.lé.-[βe.lee.ŋg-a   I read a lot
I-past-RED-[READ-fv

ndi-.ŋgu-.ⱱú.nú.-[vu.nu.lii.j-a   I steamed myself a lot
I-past-RED-[steam-fv

ndi-.ŋgu-.βé.lé.-[βe.le.βee.t-a   I spoke a lot
I-past-RED-[speak-fv

ti-.za.mu-.lé.lé.-[le.le.saa.n-a   we will look at each other a lot
we-dist fut-RED-[look-rec-fv

ti-.za.mu-.zó.mé.-[zo.me.le.zaa.n-a   we will agree with each other a lot
we-dist fut-RED-[admit.rec-fv

ti-.ŋgu-.tʰá.mbá.-[tʰa.mba.li.laa.n-a   we stretched legs over each other a lot
we-past-RED-[stretch-rec-fv

The reduplicative prosodic morpheme adheres to the minimality condition, just like in Minimal Prosodic Words. In the formal and common speech styles, the reduplicative prosodic morpheme maintains the epenthetic vowel when the Morphological Stem involved is monomoraic. This fact is illustrated in (7) below.

(7) *Reduplicative string is minimally disyllabic*

Verb   *English gloss*

i.za-ii-.[z-a   come a lot

i.lja-ii-.[lj-a   eat a lot

i.fwa-ii-.[fw-a   die a lot

i.mwa-ii-.[mw-a   drink a lot

i.mba-ii-.[mb-a   sing a lot

In elderly speech styles, however, the reduplicative prosodic morpheme maintains the bimoraic shape of the Base Prosodic Stem. This fact is illustrated in (8) below.

(8) *Reduplicative string is minimally bimoraic in elderly speech styles*

RED+Base   *English gloss*

zaa-[z-aa   come a lot

ljaa-[lj-aa   eat a lot

fwaa-[fw-aa   die a lot

mwaa-[mw-aa   drink a lot

mbaa-[mb-aa   sing a lot

5.2.4 Mixed speech styles

Many other speakers fall in between the above extremes. It is not strange, for instance, to hear an elderly speaker expanding RED up to the pre-final syllable as illustrated in (9) below.

(9) *RED in elderly speech style can be expanded to the pre-final syllable*

ndi-.ŋgu-.<u>vú.núlí.</u>-[vu.nu.lii.-ʲa   I steamed myself a lot

ndi-.ŋgu-.<u>βé.lé.bé</u>-[βe.le.βee.t-a   I spoke a lot

ti-.za.mu-.<u>lé.lé.sá</u>-[le.le.saa.n-a   We will look at each other a lot

ti-.za.mu-.<u>zó.mé.lé</u>-[zo.me.le.z-aa.n-a   we will agree with each other a lot

ti-.za.mu-.<u>zo.mé.lézá</u>-[zo.me.le.z-aa.n-a   we will agree with each other a lot

ti-.ŋgu-.<u>tʰá.mbá.lí</u>-[tʰa.mba.li.l-aa.n-a   we stretched legs over each other a lot

ti-.ŋgu-.<u>tʰa.mbá.lí.lá</u>-[tʰa.mba.li.l-aa.n-a   we stretched legs over each other a lot

Speakers 'in between' can also combine minimal bimoraic and disyllabic reduplicative shapes of the elderly speech styles and the formal and common speech styles. This fact is illustrated in (10) below.

(10) *Reduplicative string is minimally bimoraic in elderly speech styles*

RED+Base	*English gloss*
<u>zaa</u>-ii.[za	come a lot
<u>ljaa</u>-ii.[lj-a	eat a lot
<u>fwaa</u>-ii.[fw-a	die a lot
<u>i.mwa</u>-[mw-aa	drink a lot
<u>i.mba</u>-[mb-aa	sing a lot
<u>i.lja</u>-[lj-aa	eat a lot

Summary

A Prosodic Word in ciTonga can be roughly a Morphological Stem (e.g. [βe.lee.ŋg-a 'read'). It can also be roughly a Morphological Word consisting of a prefix string and the Morphological Stem (e.g. **ndi.ŋgu-.mu-[βé.lee.ŋg-a** 'I read him'). In cases where the Morphological Stem is monomoraic (e.g. **-ba** 'steal'), there are two ways of achieving the Minimal Prosodic Word. In the formal and common speech styles an epenthetic vowel [i] is added as a prefix to the monomoraic verb stem (e.g. **ii-.[ba**) in its citation form. In elderly speech styles the stem vowel is mostly lengthened (e.g. [**baa**).

Another important prosodic feature in Bantu languages is reduplication. In formal and common speech styles reduplication involves repetition of entire Morphological Stem (e.g. **ndi-.ŋgu-.[βe.lé.βé.tá-[βe.le.βee.t-a** 'I spoke a lot'). In the elderly speech styles, however, the reduplicative prosodic morpheme copies only two initial syllables/moras of the Morphological Stem (e.g. **ndi-.ŋgu-.βé.lé.-[βe.le.βee.t-a**). The reduplicative prosodic morpheme also adheres to the disyllabic or bimoraic minimality condition. In the formal and common speech styles, the reduplicative prosodic morpheme maintains the epenthetic vowel of the Base when the Morphological Stem involved is monomoraic (e.g. **i.za-ii.[za** 'come a lot'). In elderly speech styles, however, the reduplicative prosodic morpheme maintains the bimoraic shape just like the Base Stem (e.g. **zaa-[zaa**).

5.3 Generalized Templates Theory

Within Optimality Theory, issues which concern the nature of correlation between Morphological Words and Phonological Words, especially prosodic morphemes, such as Minimal Prosodic Words and reduplicative Prosodic Stems, have been approached through what is called Generalized Templates Theory (GTT). According to this theory, the constant size of prosodic morphemes are caused by general theoretical principles correlating particular morphological categories such as Stem, Root and Affix with particular prosodic constituents and from a constraint grammar defining the canonical shapes as unmarked. The theory predicts that all prosodic morphemes of the same morphological category will have identical constraints defining their canonical shape. For example, it is typical for affixes to be monosyllabic or monomoraic, and for Stems to be disyllabic. Two competing versions of GTT have emerged namely, the Prosodic Hierarchy-Based Generalized Templates Theory (McCarthy and Prince 1994a, 1994b, 1995a, 1995b, 1998, 1999; Urbanczyk 1996, 2000) and more recently the Morpheme-Based Generalized Templates Theory (Downing 2006b). In following sections, I attempt to account for the facts using Downing's Morpheme-Based Templates Theory (MBT). I will argue that analysis based on Strong Accent Constituency accounts for the facts slightly better than does MBT. I will begin by outlining ideas contained in the Prosodic Hierarchy-Based Templates Theory (PBT) which are under dispute in Downing's (2006b) proposal for MBT.

5.3.1 Prosodic Hierarchy-Based Generalized Templates Theory (PBT)

The Prosodic Hierarchy-Based Templates Theory (PBT) presented here contains ideas as adapted and put together by Downing (2006b). Within PBT, Phonological Words are derived from Morphological Words mainly by default. Prosodic Words are derived through two

families of theoretical principles correlating different morphological categories with different prosodic constituents. These principles are the **Prosodic Hierarchy** and the STEM → PROSODIC WORD HOMOLOGY and jointly they establish a correlation between the morphological categories Word and Stem and the prosodic category Foot. The correlation between the morphological categories Root and Affix and the prosodic constituent Syllable are defined by markedness constraints as well as constraint rankings.

Beginning with work by McCarthy and Prince (1986), prosodic morphology became an important source of evidence for the Prosodic hierarchy. McCarthy and Prince (1986) observe, for instance, that the possible shape of the reduplicative morpheme in many languages is restricted to prosodic categories found in the Prosodic Hierarchy. These constituents include the Syllable, the Foot and the Prosodic Word as shown in (11) below.

(11)   *Prosodic hierarchy* (McCarthy and Prince 1986, 1993; Inkelas 1989; Nespor and Vogel 1986; Selkirk 1978/81, 1984, 1995)

Prosodic Word
|
Foot
|
σ
|
μ
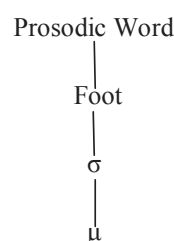
According to PBT, the **Strict Layer Hypothesis** would ensure that the constituents in the prosodic hierarchy are properly parsed (Selkirk 1978/81, 1984, 1995; Nespor and Vogel 1986; cited in Downing 2006b:7-8). One crucial clause of the hypothesis is the requirement of HEADEDNESS (Itô and Mester 1992, Orie 1997, Selkirk 1995; cited in Downing 2006b). This requirement demands that each constituent should be properly headed by containing not less than one of the units at the next level: Prosodic Words must contain stress feet, stress feet must contain syllables and syllables must contain moras.

Prosodic constituents are also required to be minimally as large as possible (e.g. Feet are required to have a maximal two syllables or moras). This condition falls out from a general prosodic principle of MAXIMALITY CONDITION (Halle and Vergnaud 1987, Itô 1989, cited in Downing 2006b:9). This condition and the Locality Principle together motivate the well known BINARITY constraint which defines the maximal and minimal size for all prosodic constituents as consisting of two of the units dominated by the constituent. It is asserted that

in the absence of the MAXIMALITY CONDITION, maximal structures would never be realized. In terms of Optimality Theory, the MAXIMALITY CONDITION is formalized by two constraints namely, a faithfulness constraint MAX-BR, optimizing segmental identity of the Base and the reduplicative string, and a MARKEDNESS constraint BINARITY requiring a prosodic category to be minimally and maximally binary at an applicable level of analysis (Downing 2006b:13).

(12)   (a)   **HEADEDNESS** (Itô and Mester 1992, Orie 1997, Selkirk 1995; cited in
               Downing 2006b:37)

               Any prosodic category C^i must dominate a C^{i-1} (e.g. Prosodic Word must
               dominate a foot).

        (b)   **MAXIMALITY CONDITION** (Downing 2006b:9, citing Itô 1989:219, Prince
               1985)

               Units are of maximal size, within the other constraints on their form.

        (c)   **BINARITY** (McCarthy and Prince 1993, Prince and Smolensky 2004, Orie
               1997; cited in Downing 2006b:9)

               A prosodic constituent contains minimally and maximally two of the units
               dominated by the constituent (i.e. Prosodic Word contains minimally and
               maximally two feet; Foot contains minimally and maximally two syllables or
               moras; syllable contains minimally and maximally two moras).

The Prosodic Hierarchy given in (11) and the constraints given in (12) would account for Phonological Words (including RED and Minimal Words) by establishing a necessary correlation between the morphological category Word and the prosodic category stress Foot. A syllogism that follows then is that a Prosodic Word must dominate a stress Foot as required by the principle of HEADEDNESS, and Phonological Words will have a minimal and maximal two feet, by BINARITY (McCarthy and Prince 1994a, 1994b, 1995a, 1995b, 1999; McCarthy 2000; cited in Downing 2006b:38).

The motivation to parse morphological words/stems into Prosodic Words is based on two principles namely, **PROSODIC MORPHEME** and the **STEM → PRWORD HOMOLOGY**. The first one requires that every morpheme be assigned a morphological category (e.g. Stem, Root, or Affix). The second principle ensures that the canonical shape of the Stem is co-

extensive with Prosodic Word (McCarthy and Prince 1994b, 1999:262, McCarthy 2000:169; cited in Downing 2006:38).

(13) (a) PROSODIC MORPHEME (e.g. RED) = MCAT (STEM, ROOT, AFFIX) (as adapted by Downing 2006b:38)

(b) STEM → PRWORD HOMOLOGY: Stem ≈ PrWord

Align the left and right edges of every Stem with the left and right edges of some Prosodic Word.

The principles of HEADEDNESS and BINARITY, when used together with the principles of PROSODIC MORPHEME and STEM → PRWORD HOMOLOGY define the correlation between the morphological category Stem and the prosodic category Foot. Prosodic morphemes which are categorized as Stems are also required by these laws to be minimally Foot-sized because both Words and Stems are parsed into a Prosodic Word. The fact that only Stem prosodic morphemes (e.g. reduplicative morphemes) which are reduced into the canonical foot size, and not the Base, is accounted for by the general constraint ranking schema which makes it optimal for prosodic morphemes to have less marked structure than corresponding Base morphemes (McCarthy and Prince 1994a, 1994b, 1999; cited in Downing 2006b:40).

5.3.2 Problems with PBT

Downing (2006b) finds PBT to be problematic because some of its claims are not supported by facts on the ground. For instance, Downing finds that contrary to what PBT claims, templates for root-and-pattern morphology do not match stress Feet; minimal words in many languages are not minimal stress Feet; all words are not subjected to the same minimality condition; truncations are not identical to minimal words; not all Stems are Prosodic Words or stress Feet; and that not all stress domains are Prosodic Words. I will elaborate two of the problems which are more relevant to the facts I am going to discuss here: Minimal words are not minimal stress feet in many languages; and not all stems are Prosodic Words or Stress Feet in many languages.

Minimal words are not minimal stress Feet: In PBT word minimality is explicated on the basis of correlating morphological Words and Stems with the Prosodic Word category. Since (by HEADEDNESS) Prosodic Words are deemed to dominate stress Feet, it is predicted that a minimal word will be foot-sized, that is a disyllabic or bimoraic minimal word (as required by BINARITY). Downing (2006b:94-100) argues, however that in many languages minimal words

147

are not minimal stress feet. She cites, among others, Hayes's (1995), Garrett's (1999) and Gordon's (1999) comprehensive cross-linguistic surveys of the correlation between minimal stress foot and minimal word requirements all of which show that there is in fact no strong correlation. Hayes (1995:88-89) shows that 30 of the 70 languages he surveyed showed a mismatch between minimal word and minimal stress foot size. Majority of 70 cases studied by Garret (1999) showed that minimal word is not connected either to foot structure or to stress patterns.

In Gordon's survey of the weight properties of some 344 languages, only 158 (46%) require minimal words to be larger than a light (CV) monosyllable. In languages where the minimal word is disyllabic or bimoraic in Gordon's survey, there is no evidence of stress or minimum stress foot size. For instance, about 50% of the languages which have a disyllabic minimal word requirement do not have the quantity insensitive stress systems that should correlate with this requirement. As if this is not enough, about 50% of the languages with a bimoraic (CVV) minimal word requirement have either a quantity insensitive stress system or no stress.

Downing (2006b:96) also argues that there are other languages where minimal words can be required to be larger than the minimal stress Foot. She cites Uradhi and Yidiɲ (Kager 1995) where the minimal word is required to be disyllabic, while the minimal stress foot is a bimoraic monosyllable. She also cites another Australian language (Alyawarra: Downing 1998, Goedemans 1996) where the minimal word is disyllabic while the minimal stress foot is a monomoraic monosyllable. Buller et al. (1993; cited in Downing 2006b:96) also report that Banawá (an Arawakan language spoken in Brazil) has monomoraic stress Feet but a bimoraic minimal word requirement.

Not all Stems are Prosodic Words or stress Feet: In PBT reduplicative morphemes are prosodically categorized as Prosodic Words (by STEM → PRWORD HOMOLOGY) and as such they must be a stress domain containing a stress Foot and other phonotactic requirements for Prosodic Words. Downing (2006b:104-107), however, argues that in many languages foot-sized reduplicative morphemes do not satisfy all of these requirements. She mentions a case of Fox disyllabic reduplication pattern (citing Dahlstrom 1997) where there is phonotactic evidence for Prosodic Word status of reduplicative morphemes, but no information on stress is provided.

Downing also observes that in many Bantu languages disyllabic reduplicative morphemes are not always Prosodic Words. Using examples from Swati, she demonstrates that the verbal

reduplicative string contains exactly two syllables although the Base may be longer than this. This fact is illustrated in (14) below.

(14)   *Swati verbal reduplication* (Downing 2006b:105 and references cited therein. Stem
        follows '=' and reduplicative morpheme is underlined)

Verb stem	*Gloss*	*X here and there; from time to time*
(a) ba-yá=li:ma	'they plough'	ba-ya-<u>limá</u>=li:ma
(b) ba-ya=líme:la	'they plough for'	ba-ya-<u>lime</u>=líme:la
(c) ba-ya=hlábe:la	'they sing'	ba-ya-<u>hlabe</u>=hlábe:la
(d) ba-ya=hlabela:na	they sing for e.o.	ba-ya-<u>hlabe</u>=hlabéla:na

PBT would account for disyllabic minimality of the reduplicative morphemes by giving them a Prosodic Word status and in turn impose a binary foot size on them. Citing Downing (1999), Downing (2006b:105) contends that the disyllabic reduplicative morphemes in (14) above cannot be Prosodic Words for two reasons: firstly, high tones in Swati never cross Prosodic Word boundaries as is the case here. Note that the high tone consistently shows up on an antepenultimate syllable. If the reduplicant were a Prosodic Word, she argues, the expectation would be that it would block the high tone shift. A second argument is that stress in Swati is assigned to the penultimate syllable, indicated by length on this syllable. However, the so-called reduplicative Prosodic Word is never realized with a lengthened vowel, "even though we would expect this to be possible if it were parsed as a Prosodic Word."

It is against such a background that Downing (2006b) develops another GTT theory, the Morpheme-Based Templates Theory outlined in the following section.
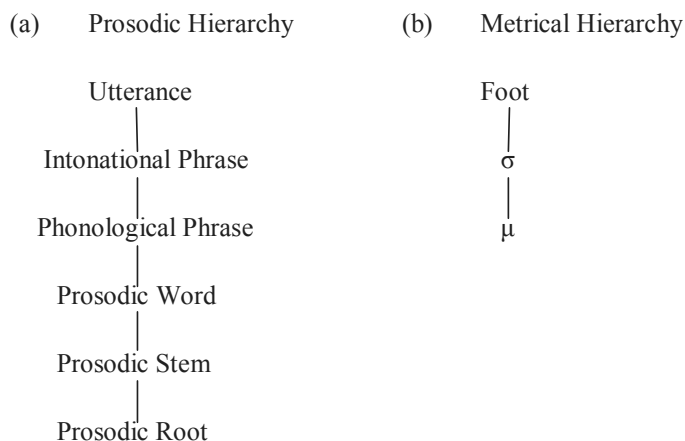
5.3.3   Morpheme-Based Templates Theory

Like PBT, MBT is a conception of Generalized Templates Theory (GTT) of prosodic morpheme shapes. As the name suggests, the motivation for canonical shape is independent of the Prosodic Hierarchy. MBT builds on Dresher and van der Hulst's (1998) proposal that canonical morpheme shape follows from a correlation between morphological complexity and phonological complexity: "Lexical morphemes meet minimality requirements, not because they contain a stress Foot, but rather because they are heads and license complex phonological structure" (Downing 2006b:111).

In Prosodic Hierarchy-Based Templates (PBT), canonical morpheme shapes such as those of the RED are explicated on the basis of the correlation between the morphological constituent Stem and the prosodic constituent Foot. This correlation is accounted for through the interaction of two constraints namely, the STEM⇒PRWORD HOMOLOGY which optimizes parsing Stems as Prosodic Words, and the principle of HEADEDNESS which requires Prosodic Words to minimally contain one stress Foot, a constituent dominated by Prosodic Word in the Prosodic Hierarchy. Downing (2006b) argues that these two constraints which formalize the core claims of PBT are empirically inadequate if we consider the problems mentioned above. She argues further that the source of the problem is scarcity of non-phonological sublexical prosodic constituents. Downing therefore suggests that there is need to expand the prosodic hierarchy to include constituents which are smaller than a Prosodic Word but also distinct from the metrical prosodic constituents dominated by Prosodic Word.

Downing (2006b) presents arguments in favour of expanding the contents of morpho-prosodic constituents below the level of Prosodic Word to include at least Stem and Root. She cites, among others, an example of Swati (Bantu) where the domain for tone realization is Prosodic Word while reduplication takes Stem as the Base domain. Citing Downing (1999), Downing (2006b) argues that other phonological processes in Swati such as labial dissimilation in the passive also takes the Stem as their domain. According to Downing, the need to recognize the Prosodic Stem as a distinct phonological domain from Prosodic Word has been argued for in numerous other Bantu languages (She cites Hyman 1993, Hyman and Mtenje 1999, Mchombo 1993, Myers 1987, and Mutaka 1994).

Following Inkelas (1989, 1993), Downing (2006b) then proposes to modify the Prosodic Hierarchy as shown in (15a) where Prosodic Stem and Prosodic Root are included as universally available sublexical morpho-prosodic constituents which are the prosodic equivalents of the morphological constituents, bimorphemic Stem and monomorphemic Root.

(15)   *Morphological and metrical prosodic hierarchies* (Downing 2006b:115; adapting
Inkelas 1989:46)

(a)   Prosodic Hierarchy   (b)   Metrical Hierarchy

Utterance   Foot
| |
Intonational Phrase   σ
| |
Phonological Phrase   μ
|
Prosodic Word
|
Prosodic Stem
|
Prosodic Root

The revised Prosodic Hierarchy places the metrical constituents (Foot, syllable, and mora) in their own Metrical Hierarchy in order to eliminate the hierarchical connection between Prosodic Word and Stress Foot that crucially motivates minimality conditions in PBT.

The central claim of Downing's morpheme-based GTT (MBT) is that "the basic morphology-prosody correlation is between a single morpheme and a single syllable". This theory also seeks to replace Foot binarity as a motivation for the tendency of prosodic morphemes towards having binary minimality. According to MBT, a branching requirement on morphological heads accounts for this tendency. Downing then develops each of these proposals.

MBT replaces STEM ⇒ PRWORD HOMOLOGY with the minimal **Morpheme-Syllable correlation**. This line of thinking, she argues, is found in OT literature dealing with the correlation between morphological structure and prosodic constituents namely, that the minimal morphology-prosody correlation is between a single morpheme and a single syllable (Feng 2004, McCarthy and Prince 1994b, Russel 1997, and Urbanczyk 1996). Downing (2006b:119-120) then formalizes this correlation as given in (16) below.

(16)   **MORPHEME-SYLLABLE CORRELATION** (**MORPH-SYLL**, adapted, Russel 1997:121)

Each morpheme contains exactly one syllable.

Following McCarthy and Prince (1999), Downing (2000, 2006b) argues that constraints like MORPH-SYLL which evaluate the prosodic weight of a string can be considered a variety of correspondence constraints, establishing a relationship between the segments and prosody of a single morpheme. Following Oostendorp (2004), Downing also assumes that constraints like MORPH-SYLL which define correspondence between a string and a syllable are only satisfied if some element of the string which realizes the morpheme is associated with the head (nucleus) of a syllable.

In MBT, any tendency for (prosodic) morphemes to satisfy a binary minimality condition falls out from Dresher and van der Hulst's (1998) proposal that "there is a correlation between morphological complexity and phonological complexity" (Downing 2006b:121). Thus, Roots, which are lexical heads, meet minimality requirement "not because they contain a stress Foot, but rather because heads require branching phonological structure" (Ibid). According to Dresher and van der Hulst (cited by Downing 2006b), a branching requirement on heads is one way of enforcing a "Head-Dependent complexity asymmetry", a property of linguistic systems cross-linguistically (Ibid, citing Anderson and Ewen 1987).

Downing (2006b:122) formalizes the branching principle motivating binary minimality as given in (17), and branching is defined as in (18).

(17)   HEADSBRANCH (Downing 2006b:122; adapted Dresher and van der Hulst 1998)

Lexical heads (Roots) must prosodically branch.

(18)   PROSODIC BRANCHING (Downing 2006b:122; adapted Ussishkin 2000:43)

A constituent branches *iff* it or its daughter contains more than one daughter

All of the three representations given in (18) below satisfy HEADSBRANCH. A head may contain two syllables (18a) or two moras (18b). In (18c), the head dominates a mora with two daughters.

(18)   Heads branch (Downing 2006b:122)

(a) Head   (b) Head   (c) Head
     /\   /\ |
    σ σ   μ μ   μ
                                        /\
                                       V   C
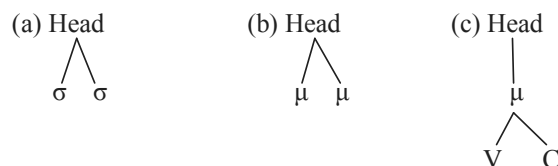
Thus, Roots, as monomorphemic heads, are predicted to be monosyllabic by MORPH-SYLL and to optimally satisfy branching by matching (18b) or (18c). Like Roots, Affixes as monomorphemic non-heads are also predicted to be monosyllabic by MORPH-SYLL, but they are not required to branch. Since stems are constituents minimally consisting of two morphemes, it is expected of them to have two syllables, one for each morpheme. This is predicted by the MORPHEME-SYLLABLE CORRELATION principle. The constraint in (19), a corollary of MORPH-SYLL, formalizes this disyllabic minimality requirement on Prosodic Stem, the morpho-prosodic constituent which corresponds to the morphological Stem.

(19)   PROSODICSTEM   (Downing 2006b:124)

$$\text{PROSODICSTEM} \quad \sigma \quad \sigma$$

Downing (2006) further observes that the branching constraints (HEADSBRANCH, PROSODIC BRANCHING and PROSODICSTEM) require optimal (unmarked) lexical constituents to be minimally binary branching, but they are not enough to define the binary maximality requirement which is typical of reduced prosodic morphemes like reduplicants. Downing (2006b:125) notes that a more general motivation for binary maximality comes from the Locality Principle, "which limits all phonological processes to a binary window: 'a special element and […] a structurally adjacent element and no other' (McCarthy and Prince 1986:1)." Following work by Harris (1994), Halle and Vergnaud (1987), and Ussishkin (2000:53), Downing proposes that the relevant adjacent elements for defining binarity are a constituent daughter and a constituent edge. BINARITY is thus defined as in (20) below.

(20)   BINARITY (Downing 2006b:125)

        Each daughter of a constituent must be adjacent to some edge of the constituent.

Downing (2006b) further argues that the markedness constraint BINARITY is outranked by FAITHFULNESS constraints in the regular vocabulary, as it appears to be rare for languages to require all Stems to contain exactly two moras or two syllables. The unmarked status of binary constituents emerges only in certain morphological constructions such as reduplicative prosodic morphemes and this is attributed to TETU (the emergence of the unmarked) constraint ranking, the same one which optimizes maximally binary canonical shape in PBT.

Summary
Within OT, morphology-phonology interfaces have been approached through what is called Generalized Templates Theory (GTT), requiring the constant size of prosodic morphemes to

follow from general principles correlating particular morphological categories with particular morphological categories such as Stem, Root and Affix with particular prosodic constituents and from a constraint grammar defining the canonical shapes as unmarked. The prediction made by the theory is that all prosodic morphemes of the same morphological category will have identical constraints defining their canonical shapes. Prosodic Hierarchy-Based Templates Theory (PBT) and Morpheme-Based Templates Theory (MBT) are two competing theories within the GTT. In PBT, Prosodic Words are derived through the principles of Prosodic Hierarchy and the STEM-PROSODIC WORD HOMOLOGY which jointly establish a correlation between the morphological categories Word and Stem and the prosodic category Foot (by HEADEDNESS). Downing (2006b) finds PBT to be problematic on a number of grounds such as the fact that minimal words are not usually minimal stress feet in many languages and that not all Stems are Prosodic Words or stress Feet. She therefore proposes to account for prosodic morphemes such as Minimal Prosodic Words and reduplicative Prosodic Stems in terms of Morpheme-Based Templates.

As the name suggests, the motivation for canonical shapes in MBT is independent of the metrical Prosodic Hierarchy. A Morpho-Prosodic Hierarchy eliminates the hierarchical connection between Prosodic Word and stress Foot which crucially motivates minimality conditions in PBT. MBT makes two central claims. Firstly, the basic morphology-prosody correlation is between a single morpheme and a single syllable. Secondly, a branching requirement on morphological heads replaces Foot binarity. These two claims result into two principles namely, the MORPHEME-SYLLABLE CORRELATION and the HEADS BRANCH. The former requires that each morpheme contains one syllable, while the latter ensures that lexical heads such as roots branch prosodically. Since Stems are constituents minimally consisting of two morphemes, it is expected of them to have two syllables, one for each morpheme (by MORPHEME-SYLLABLE CORRELATION). A constraint born out of this is called PROSODICSTEM (Prosodic Stems are minimally binary). The re-defined principle of BINARITY explains the binarity maximality which is typical of reduced prosodic morphemes such as reduplicants. Since Affixes are non-heads, they are not bound to be branching. The following sections attempt to analyze the size of general Prosodic Words, Minimal Prosodic Words, and Reduplicative Prosodic Morphemes in ciTonga in terms of MBT.

5.4 MBT analysis

5.4.1 General Prosodic Words

One observation we made about Prosodic Words in ciTonga is that they can be roughly a Morphological Stem without prefixes (e.g. [**βe.lee.ŋg-a** 'read']). They can also be roughly a Morphological Word consisting of a prefix string and the Morphological Stem (e.g. **ndi.ŋgu-mu-.[βé.lee.ŋg-a** 'I read him'). Just like in PBT, Morphological Words are said to map into Prosodic Words by default (Inkelas 1989, Prince and Smolensky 2004). This position is widely accepted and it makes sense because as much as it is a phonological category, a Prosodic Word is also a morphological category. It denotes a correlation between two structures: phonological structure and morphological structure. As Downing (2006b:8) puts it, the category Prosodic Word "is a phonological domain roughly equivalent to the morphological category, Word."

The constraints PROSODICSTEM and BINARITY, however, predict that Prosodic Stems are going to be at most disyllabic or bimoraic. Many Prosodic Stems in ciTonga, even without suffixes, are polysyllabic, and they satisfy the constraint PRSTEM which requires Prosodic Stems to be minimally binary at the level of the syllable or the mora. They, however, clearly violate BINARITY, requiring daughters (syllables) to be adjacent to some edge of the constituent, or in other words, to be minimally and maximally binary. Explanation for this would be that the constraint MAX-IO, requiring preservation of input segments, ranks above BINARITY. A tableau in (21) below illustrates.

(21) *Many Prosodic Stems violate BINARITY for the sake of MAX-IO*

/[βeleβeet-a/ 'speak'	PRSTEM	MAX-IO	BINARITY
☞ (a) {βe.le.βee.t-a}			**
(b) {βe.le}		*!***	

Both candidates satisfy the constraint PRSTEM which requires Prosodic Stems to contain minimally two syllables. However, candidate (21a) is optimal because it satisfies the high-ranking constraint MAX-IO which requires input segments to be preserved in the output. On the other hand, candidate (21b) is ruled out for violating MAX-IO. It violates this constraint four times although one violation would be enough to dismiss the candidate.

5.4.2 Minimal Prosodic Words

Two other observations were made about Minimal Prosodic Words in ciTonga. In cases where the Morphological Stem is monomoraic (e.g. [**ba** 'steal'), there are two ways of achieving the Minimal Prosodic Word. In the *formal and common speech styles* an epenthetic vowel [i] is added as a prefix to the monomoraic verb stem in its citation form (e.g. **ii-.[ba**). In elderly speech styles, the stem vowel is lengthened (e.g. [**baa**). Both observations indicate that a binary minimal word is the target. A generalization requiring explanation is that Minimal Prosodic Words are binary (at the level of the syllable or the mora). MBT would account for this generalization as a case of ranking PRSTEM above DEP-IO, forbidding segment or mora insertion. This fact is illustrated in the tableau in (22) below.

(22) *Minimal disyllabic words accounted for by PROSODICSTEM*

[lj-a 'eat'	PRSTEM	DEP-IO
(a) {lj-a}	*!	
(b) {ii-.lj-a}/{lj-aa}		**/*

Thus, candidate (22a) is disqualified because it violates a high-ranking constraint PROSODICSTEM which requires Prosodic Stems to have a minimum of two syllables or moras. On the other hand, candidates (22b) are optimal because they satisfy PROSODICSTEM. They violate a low-ranking constraint DEP-IO which militates against insertion of vowels or moras. A challenge for PBT and MBT would be to explain the variation of Minimal Prosodic Word between the formal/ common speech styles, on the one hand, and the elderly speech styles on the other. PBT would suggest that the language system has two foot types (syllabic and moraic). MBT would suggest that the Prosodic Stem in this system can be syllabic or moraic. Both these ideas are not insightful enough because they do not give a convincing reason why a single language should have two different Foot or Prosodic Stem types, respectively. In other words, a general principle of some sort is lacking.

5.4.3 Reduplication

Besides general and minimal words, MBT should also be able to account for the following reduplication facts. (a) In formal and common speech styles reduplication involves repetition of a whole Morphological Stem (e.g. **ndi-.ŋgu-.βe.lé.βé.tá-[βe.le.βee.t-a** 'I spoke a lot'). (b) In elderly speech styles, however, the reduplicative prosodic morpheme copies only two

initial syllables of the Morphological Stem (e.g. **ndi-.ŋgu-.βé.lé.-[βe.le.βee.t-a**). (c) The reduplicative prosodic morpheme also adheres to the disyllabic or bimoraic minimality condition. In formal and common speech styles, the reduplicative prosodic morpheme maintains the epenthetic vowel [i] when the Morphological Stem involved is monomoraic (e.g. **i.za-ii.[za** 'come a lot'). In elderly speech styles, however, the reduplicative prosodic morpheme has the bimoraic shape like that of the Base Prosodic Stem (e.g. **zaa-[zaa**).

According to Downing (2006b:149), the reduplicative morpheme in Bantu languages is a Stem and that reduplication is a form of Stem-Stem compounding[38]. The point she is driving at is that with the Stem status, the canonical form of a reduplicative prosodic morpheme has to satisfy the constraint PROSODICSTEM which requires Prosodic Stems to be branching (to be minimally binary). According to observation (a), the RED is more than two syllables and it thus satisfies PROSODICSTEM. It, however, violates BINARITY, requiring Prosodic Stems to be minimally and maximally binary. This means that BINARITY ranks below MAX-BR which requires all segments of the morphological stem to be 'copied' by RED. A tableau in (24) below illustrates.

(23)   **MAX-BR**

"All the segments of the Base are contained in the RED" (McCarthy and Prince 1993, cited by Downing 2006b:13).

(24) *RED copies entire morphological stem in the formal and common speech styles*

/[RED~Stem~-βeleβeet-a/	PRSTEM	MAX-BR	BINARITY
(a) {βe.le.}-{βe.le.βee.t-a}		*!***	-**
☞ (b) {βe.le.βe.ta}-{βe.le.βee.t-a}			**-**

Both candidates do not violate PRSTEM because they meet the minimal requirements of two syllables or two moras per Prosodic Stem. However, candidate (24a) is disqualified because it violates a high-ranking constraint MAX-BR which requires segments of the Base to be preserved in the RED. It violates this constraint twice (for two syllables not copied by RED).

---

[38] Downing cites Downing (2000, 2003), Hyman et al. (1999), and Inkelas and Zoll (2005) as containing detailed arguments supporting a Stem analysis of the reduplicative morpheme.

Candidate (24b) is optimal because it satisfies MAX-BR. It violates the low-ranking constraint BINARITY four times (divided between RED and the Base).

Observation (c) indicates that reduplicative Prosodic Stems are minimally binary (at the level of the syllable or the mora). Once again, MBT would account for this generalization as caused by PRSTEM which requires Prosodic Stems to be minimally binary (at the level of the mora or the syllable). A tableau in (25) below illustrates.

(25) *RED is subjected to Prosodic Stem minimality condition*

RED$_{Stem}$-lja 'eat'	PRSTEM
(a) {lja}-{ii.lj-a}	*!
{lja}-{lj-aa}	
☞ (b) {i.lja}-{ii.lj-a}	
{ljaa}-{lj-aa}	

Candidates (25a) are ruled out because they violate the high-ranking constraint PROSODICSTEM (1 violation by each RED). Candidates (25b) are optimal because they satisfy this constraint. The challenge for both MBT and PBT, once again, would be accounting for two different shapes of the minimal size of the reduplicative Prosodic Stem in a single system (disyllabic or bimoraic). There must be a single better and larger principle that can account for the variation rather than explanation based on different types of feet or different types of Prosodic Stems (i.e. syllabic vs. moraic).

The fact that the reduplicative prosodic morpheme copies only two initial syllables of the Morphological Stem in elderly speech styles (e.g. **ndi-.ŋgu-.βé.lé.-[βe.le.βee.t-a**) is reminiscent of what is called the Emergence of the Unmarked (TETU): a marked structure that is optimal in the Base is found to be non-optimal in the RED (Alderete et al. 1999; McCarthy and Prince 1994a, 1994b, 1999; and Steriade 1988; all cited in Downing 2006b:41). In other words, the size of the reduplicative Prosodic Stem in the elderly speech styles represents a canonical Prosodic Stem shape. The fact that only Stem prosodic morphemes (e.g. reduplicative morphemes) are reduced into the canonical two-syllable size,

and not general Stems like the Base Prosodic Stem, is accounted for by the general constraint ranking schema which makes it optimal for prosodic morphemes to have less marked structure than corresponding Base morphemes (McCarthy and Prince 1994a, 1994b, 1999; cited in Downing 2006b:40). Within OT, this tendency is formalized by a constraint ranking schema given in (26) below.

(26) *TETU constraint ranking* (adapted by Downing 2006b:41 from McCarthy and Prince 1999:261)

        IO-FAITHFULNESS >> MARKEDNESS CONSTRAINTS >> B-R FAITHFULNESS

This constraint ranking may tolerate marked structure in the output of the Base (and unreduplicated forms) since the constraint FAITH-IO, which requires input structure to be realized in the output, ranks above (some) markedness constraints. The very same marked structure will be non-optimal in the reduplicative morpheme since some markedness constraints outrank the constraint MAX-BR which requires total copy of the Base by RED.

Two important markedness constraints in our case are PROSODICSTEM and BINARITY. Unmarked disyllabic Prosodic Stems are possible if PROSODICSTEM and BINARITY outrank MAX-BR to give the TETU ranking in (27).

(27) *TETU ranking defining a single stress foot as the 'unmarked Prosodic Word*

        IO-FAITHFULNESS >> PROSODICSTEM & BINARITY >> MAX-BR

This analysis is illustrated in the tableau in (28) below.

(28) *Partial reduplication in elderly speech styles*

/RED_{Stem}-vunuliil-a/	MAX-IO	PROSODIC STEM	BINARITY	MAX-BR
(a) {vu.nu.li.ʲa}-{vu.nu.lii.-ʲa}			*!*	
(b) {vu}.-{vu.nu.lii.-ʲa}		*!	*	******
☞ (c) {vu.nu}-{vu.nu.lii.-ʲa}				****
(d) {vu.nu}-{vu.nu}	*!***			****

159

Candidate (28a) is ruled out because it violates a high-ranking constraint BINARITY which requires each daughter of a constituent to be adjacent to some edge of the constituent. Candidate (28b) is nullified because it violates another high-ranking constraint PROSODICSTEM, requiring Prosodic Stems to be minimally disyllabic. It also violates another high-ranking constraint BINARITY, but a decision has already been made by PROSODICSTEM. The optimal candidate (28c) satisfies both PROSODICSTEM & BINARITY. It violates MAX-BR four times (one violation mark for each segment not copied by RED), but these violations can be tolerated for the sake of BINARITY. Candidate (28d) is disqualified for violating MAX-IO which militates against deletion of general input segments in the output.

One last observation we made was that many speakers fall in between the formal or common and the elderly speech styles. For instance it is not strange to hear speakers expanding RED up to the pre-final syllable [e.g. **ti-.za.mu-.<u>zo.mé.lé</u>-[zo.me.l-e.z-aa.n-a** 'we will agree with each other a lot']. A suggested analysis would be that the continuum between formal and extreme informal elderly speech styles is influenced by the need to minimize the number of violations of either the BINARITY constraint side or the MAX-BR side. For instance, increasing the number of preserved syllables in the reduplicative Prosodic Stem ensures that serious violations of the constraint MAX-BR are averted. On the other hand, reducing the number of the syllables or segments ensures that BINARITY is minimally violated. Similarly, reduplicative Prosodic Stems based on monomoraic verbs can combine minimal forms found in the formal and common speech styles. The analysis once again is that these forms reflect the different means to satisfying the constraints PROSODICSTEM and BINARITY in this language.

5.4.4 Problems with MBT

The first problem concerns understanding of Prosodic Word especially in Bantu languages. Downing (2006b:105) observes that a Prosodic Word in Bantu languages is known for its penultimate stress, cued by vowel lengthening. She uses this phenomenon, for instance, to argue against the fact that the reduplicative morpheme is a Prosodic Word as presented by, among others, McCarthy and Prince (1993b). Using examples from Swati, she demonstrates that the verbal reduplicative string contains exactly two syllables although the Base may be longer than this. Citing Downing (1999), Downing (2006b:105) contends that the disyllabic reduplicative morphemes cannot be Prosodic Words because, among others, they do not have

penultimate vowel lengthening (stress) as it is expected of Prosodic Words in this language. Downing's use of penultimate stress and lengthening as a benchmark for ruling out reduplicative strings as having Prosodic Word status, however, is problematic because Prosodic Words are most likely not domains for penultimate vowel lengthening in Bantu languages like Swati and ciTonga. As I have already highlighted in the case of stress analysis in the preceding chapter, the idea that the domain for penultimate lengthening in these languages is not a Prosodic Word is thoroughly investigated and confirmed (e.g. by Hyman 2009 and, ironically, by Downing herself). According to Hyman (2009), two domains of penultimate lengthening (or penultimate stress) are ascertained in Bantu languages which have the PL (penultimate lengthening) phenomenon. PL may be utterance-penult (e.g. Sotho: Doke 1967:125) or phrase-penult (e.g. Chichewa: Kanerva 1990, Tumbuka: Downing 2006a, Makonde: Kraal 2005, and Matengo: Yoneda 2005). Thus, as Hyman (2009) observes, the attraction of length is across words. This is not a fundamental problem, that is, one which is unique for MBT. However, it becomes an MBT problem because Downing (2006b) alludes to it as one ground for rejecting Prosodic Hierarchy-Based Templates Theory.

One fundamental problem of MBT is that the theory would not give adequate insight regarding variation in terms of Minimal Prosodic Word and minimal reduplicative Prosodic Stem as presented in the formal/common speech styles and the elderly speech styles of ciTonga. MBT would suggest that both disyllabic and bimoraic Minimal Prosodic Words and RED satisfy the constraints PRSTEM and BINARITY. These two constraints are not insightful enough as they don't really say why the Minimal Prosodic Word should be disyllabic and bimoraic (in the same or different systems). Similarly, MBT accounts for total copy of the Base by RED in the formal and common speech styles on the basis of ranking MAX-BR above BINARITY, and partial reduplication in the elderly speech styles on the basis of BINARITY outranking MAX-BR. Whichever ranking, MBT fails to reveal a common ground for both sizes of RED which are well defined bimoraic constituents in elderly speech styles and Prosodic Stems in the formal/common speech styles. What is lacking is one general principle which captures the variation.

Another problem is related to the preceding challenge. MBT informs us that different systems have different realizations of RED because in one system it is an Affix (syllable), in another system a Root (thus CVC) and still in another Stem. Where RED is disyllabic, MBT, just like in PBT, we are informed that this is a case of Emergence of the unmarked (TETU) Prosodic Stem (by BINARITY or PROSODICSTEM). This might be correct. The core problem however is

that, in the absence of a single principle underlying all realizations of RED (which are well defined prosodic constituents), MBT would find it hard to explain why RED is a syllable in one system, Prosodic Root in another, a disyllabic (TETU) Prosodic Stem or indeed a total copy of the Base Prosodic Stem in other systems. In other words, there is nothing that binds all these prosodic/morphological categories together. The TETU explanation is not informative enough: Why should disyllabic Reduplicative Prosodic Stems be unmarked in Swati and total copy of the Base which is a wellformed Prosodic Stem and found in all the 14 languages in Malawi be described as having a marked shape? Something just doesn't add up. Both PBT and MBT would not tell us with precision what the common basis is.

I have already alluded to the following problem in chapter 3. MBT suggests that Prosodic Stems in Bantu languages like ciTonga can contain nothing but syllables and moras, with no any Foot of some kind. This understanding of the structure of Prosodic Stems is misguided and it is based on a wrong assumption that accent is always felt phonetically like in stress accent, such as by vowel lengthening[39]. As we saw in the analysis of vowel and consonant deletion, a serious challenge to MBT is to account for Prosodic Stems which contain strong positions and Foot-like structures on two initial light syllables and the bimoraic penultimate syllable, which are most likely Accent Feet. MBT does not envisage anything of this sort. Its core principle of PROSODICSTEM together with BINARITY predicts that a super Prosodic Stem in languages with no stress will have two syllables or moras. An adequate characterization of ciTonga grammar, therefore, should be able to account for the fact that a minimal Prosodic Stem is required to be disyllabic while at the same time containing a Foot structure of some kind.

Finally, MBT, like many other theories of prominence, places itself among theories which refer to strong positions such as stressed syllable, initial syllable, and so on (such as those referred to by Beckman 1995, 1997, 1998; Harris 1990, 1994, 1997, 2004; Steriade 1994; Barnes 2002; Dresher and van der Hulst's 1998; Trubetskoy 1939; Selkirk 1994; Alderete 1995; McCarthy and Prince 1995a; Casali 1996; Smith 2001). Strong positions may also include moras, feet, and segments (and features) as we saw in preceding chapters. MBT, however, would not spell out a universal principle underlying all these strong constituents and how they are arrived at in different systems. Furthermore, the fact that RED copies strong

---

[39] I suggested in chapter 3 and I do the same in the following sections that there may be a phonetic cue of some kind for Accent Foot only when it is a Minimal Strong Accent Constituent of a relevant language system or style. Where the SAC is a Prosodic Stem, Accent Foot may sometimes just be inert (such as in formal and common speech styles).

positions such as initial syllable, or first two syllables, or indeed a Prosodic Stem, needs a general statement that enhances prominence, rather than one that demotes them to things like affix (for one syllable RED).

Summary

In MBT, non-reduplicative Prosodic Stems can exceed the maximum two-syllable size due to ranking of FAITH-IO, requiring input segments to be preserved in the output, above the constraint BINARITY which requires Prosodic Stems to be minimally and maximally binary. Minimal Words, however, are usually binary due to the high-ranking constraint PROSODICSTEM which demands that prosodic and morphological heads such as Stems be branching. The reduplicative morpheme, assumed to be a Stem, has to satisfy the constraints PROSODICSTEM which requires Prosodic Stems to be minimally binary, and BINARITY which ensures that the size of reduced forms such as RED is not less than and does not exceed two syllables. In the formal and common speech styles, however, RED could exceed two syllables. The interpretation of this fact would be that the constraint MAX-BR which requires all segments of the Base to be copied by RED ranks above the constraint BINARITY which demands that the minimal and maximal size of RED be two syllables. In elderly speech styles, where only two initial syllables of the Base are copied by RED, the interpretation would be that BINARITY ranks above MAX-BR following strictly the TETU constraint ranking schema which makes marked structures which are optimal for the Base unoptimal for the reduplicative Prosodic Stem.

These analyses however have not been without challenges. To begin with, MBT is most likely built on inadequate understanding of the Prosodic Word concept if the example given above is anything to go by. The theory also lacks crucial generalizations. There is lack of a single universal principle that would precisely account for all realizations of the prosodic Reduplicative Prosodic Morpheme and the Minimal Prosodic Word across the speech styles and systems. Although I agree with the idea of a PROSODICSTEM constraint, requiring Prosodic Stems to be minimally disyllabic (not bimoraic), I find MBT's understanding of the structure of Prosodic Stem to be slightly problematic especially with reference to Bantu languages like ciTonga. The theory suggests that, by and large, (minimal) Prosodic Stems may contain nothing but syllables, no foot of some kind. This idea is misguided and it is most likely based on wrong assumption that accent is always stress-accent and that it is always cued phonetically. Finally, MBT, like many other theories of prominence, places itself among

theories which refer to strong positions, but there is no attempt to try to understand a basic principle underlying all strong positions as I do in the present study.

5.5   Strong Accent Constituent-Based Templates Theory

I have maintained the PBT idea where metrical constituents (Foot, Syllable and Mora are not separated from morpho-prosodic categories. However, I have included Prosodic Stem in the Prosodic Hierarchy as argued for in MBT and by others elsewhere. Prosodic Root would be added in languages where it is of phonological importance.

(29)   *The Prosodic Hierarchy*

Prosodic Word
|
Prosodic Stem
|
Accent Foot
|
Syllable
|
Mora
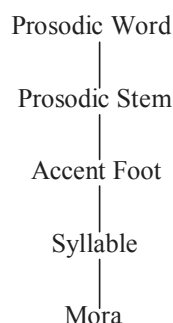
I have maintained the PBT idea of Strict Layer Hypothesis especially its crucial clause of HEADEDNESS which requires constituents in the prosodic hierarchy to be properly parsed (Sekirk 1978/81, 1984, 1995; Nespor and Vogel 1986; Mester 1992, Orie 1997, Selkirk 1995). Following this idea, each Prosodic Word is supposed to be properly headed by containing at least a Prosodic Stem. A Prosodic Stem contains at least an Accent Foot, Accent Foot at least a Syllable and Syllable contains at least a Mora. In addition, a Prosodic Stem must be minimally disyllabic as predicted by the MORPHEME-SYLLABLE CORRELATION PRINCIPLE and required by Downing's (2006b) constraint PROSODICSTEM.

Prosodic constituents must also satisfy MAXIMALITY CONDITION to ensure that they are minimally as large as possible. I have maintained Downing's (2006b) sense of BINARITY in which the relevant adjacent elements for defining binarity are a constituent daughter and a constituent edge. Prosodic Words are therefore expected to contain a minimum and maximum of two Prosodic Stems. Prosodic Stems must contain a minimum and maximum of two Accent Feet, and Accent Feet a minimum and maximum of two moras. At the heart of Strong Accent Constituent-Based Templates Theory, however, is the fact that Prosodic Words and Stems must contain at least a Strong Accent Constituent of a particular language system as

guided by the Universal Principle of PROSODIC HIERARCHY. PROSODIC HIERARCHY provides that lower level prosodic constituents such as ACCENT FOOT and PROSODIC STEM are stronger than higher prosodic constituents.

5.5.1 The Prosodic Word and the Prosodic Stem

The analysis of Prosodic Words in ciTonga would be that many of them have one Prosodic Stem as illustrated in (30) below.

(30)   *Prosodic Words made up of Prosodic Stem alone*

Verb   *English gloss*

{bii.k-a}
cook-fv   cook

{be.nee.k-a}
cover-fv   cover

{βe.lee.ŋg-a}
read-fv   read

{be.ne.k-aa.n-a}
cover-rec-fv   cover each other

{bi.k-i.l-aa.n-a}
cook-appl-rec-fv   cook for each other

{βe.le.ŋg-e.s-aa.n-a}
read-caus-rec-fv   cause each other to read

This state of affairs is not in violation of the principle of HEADEDNESS, requiring a prosodic constituent to contain at least one of the units it dominates. Prosodic Words, however, must also satisfy MAXIMALITY CONDITION to ensure that they are minimally as large as possible, i.e. they must contain a minimum and maximum of two Prosodic Stems. Since Prosodic Words are first and foremost morphological categories, the idea of maximality is not always satisfied. The principle of BINARITY is thus largely violated because many Prosodic Words contain just one Prosodic Stem. Prosodic Words in ciTonga may also contain two or more Prosodic Stems. Evidence for this claim is based on blocking of general morpheme deletion processes such as those of present progressive aspect marker **-tu-** and the infinitive marker **-**

ku- before monomoraic verb Stems and object markers. The evidence points to the fact that object markers and the distant future tense marker **-zamu-** have a Stem status in this language which qualifies them for Accent Foot and Prosodic Stem parsing.

5.5.1.1 Blocking of -tu- deletion

As I have shown in the preceding chapter, a 'normal' present progressive aspect verb (with two or more syllables) in ciTonga is expressed without a morphological marker. This fact is illustrated in (31) below.

(31)   *Normal present progressive aspect verbs have no morphological marker*

Verb   *English gloss*

ndí-.[lé.lee.s-a   I am looking
I-look-fv

ndí-.[zó.mee.l-a   I am admitting
I-admit-fv

ndí-.[sá.mbii.z-a   I am teaching
I-teach-fv

ndí-.[βé.lé.βee.t-a   I am speaking
I-speak-fv

ndí-.[vú.nú.lii.-ʲa   I am steaming myself
I-steam-fv

tí-.[tʰá.mbá.lí.l-aa.n-a   we are stretching legs over each another
I-stretch-rec-fv

However, when the Morphological Stem involved is monomoraic, the grammar of ciTonga allows the morpheme **-tu-** to appear before it as given in (32) below.

166

(32) *Tu- surfaces before monomoraic Morphological Stems*

Pres. Prog   *English gloss*

ndi-.túu-.[vw-a   I am listening

ndi-.túu-.[mb-a   I am singing

ndi-.túu-.[lj-a   I am eating

ndi-.túu-.[b-a   I am stealing

ndi-.túu-.[mw-a   I am drinking

ndi-.túu-.[sw-a   I am breaking

Mtenje (2006) accounts for **-tu-** insertion in (32) above by invoking a prosodic constituent called **Prosodic Stem** which is required to be minimally disyllabic. Thus the forms in (32) above can be represented as in (33) below.

(33) *Prosodic Stem is minimally disyllabic*

Pres. Prog   *English gloss*

ndi-.{túu-.vw-a}   I am listening

ndi-.{túu-.mb-a}   I am singing

ndi-.{túu-.lj-a}   I am eating

ndi-.{túu.b-a}   I am stealing

ndi-.{túu-.mw-a}   I am drinking

ndi-.{túu-.sw-a}   I am breaking

However, what puzzles Mtenje is the fact that when an object marker is introduced in a normal Present Progressive Aspect verb complex, an epenthetic **-tu-** surfaces again, but now before the object marker. I illustrate this fact in (34) below. Object markers are underlined.

(34)   *Epenthetic /tu/ also surfaces before object markers in longer verb words*

Present progressive verb	*English gloss*
ndi-.tú-.<u>kú</u>-.[ka.mbúu.l-a	I am shouting to you
I-prog-you-shout-fv	
ti-.tú-.<u>ví</u>-.[βe.lé.βee.t-a	we are talking about them
we-prog-8OM-speak-fv	
ti-.tú-.<u>mú</u>-.[βe.lé.ŋg-aa.n-a	we are each other reading him
we-prog-1OM-read-rec-fv	
ti-.tú-.<u>mú</u>-[ka.mbú.l-aa.n-a	we are each other shouting at him
we-prog-1OM-shout-rec-fv	

In a footnote, Mtenje (2006) confesses that "the presence of the epenthetic **-tu-** in the long verbs and in monosyllabic verbs… is rather unexpected".

Unlike Mtenje (2006), I treat this process as involving **-tu-** deletion rather than insertion. My conviction is that **-tu-** is a bonafide morphological marker for the present progressive aspect (and not an epenthetic syllable)[40]. The basis of this reasoning is that **-tu-** is confined to particular morphological constructions. Like most tense/aspect markers in this language, **-tu-** in the examples given above consistently appears before an object marker or a (monomoraic) verb stem. It would thus be more sensible to regard **-tu-** as some kind of a prefix which for some reason undergoes deletion in the normal present progressive aspect verbs. My suggestion then is that by adopting the same Prosodic Stem theory, Mtenje's puzzle (**-tu-** being inserted before object markers as well) is easily solved. Object markers in this language seem to have a Stem status which is coterminant with a Prosodic Stem. This analysis is illustrated in (35) below.

---

[40] Note that Mtenje (2006) is aware of the fact that **-tu-** could as well be a morphological marker for the present progressive aspect, as opposed to being an epenthetic form.

(35) **-tu-** *and the object marker constitute a proper Prosodic Stem*

ndi-{tú-.kú.}-{ka.mbúu.l-a}   I am shouting to you
I-prog-you-shout-fv

ti-{tú-.ví.}-{βe.lé.βee.t-a}   we are talking about them
we-prog-8OM-speak-fv

ti-{tú-.mú.}-{βe.lé.ŋgaa.n-a}   we are each other reading him
we-prog-1OM-read-rec-fv

ti-{tú-.mú.}-{ka.mbú.laa.n-a}   we are each other shouting at him
we-prog-1OM-shout-rec-fv

The fact that the object marker has a Stem status finds further evidence in discourse: In conversation, ciTonga speakers have a tendency to elide the verb stem and the whole meaning is borne out by the subject marker, morpheme **-tu-**, and the object marker itself. A dialogue between speaker A and speaker B in (36) below illustrates.

 (36) *The object marker is coterminant with a Prosodic Word*

A: [u-<u>tú-mú</u>-ziíβa]   [pamweŋga]   [u-<u>tú-mú</u>-zíβá]   [cáá]?

       [you-prog-1OM-know] [or]   [you-prog-1OM-know][not]

       "Do you know him or not?"

B: [*ndi-túu-mu*].

       [I-prog-1OM]

       "I do."

To cut the long story short, it appears **-tu-** surfaces before the object marker and the monomoraic verb stem because it is now parsed by Prosodic Stem in both cases. The generalization, therefore, is that Prosodic Stems are minimally disyllabic. In terms of OT, Downing's (2006b) PROSODICSTEM constraint seems to be responsible for this status. Tableaux in (37) and (38) below illustrate.

(37)   *Monomoraic Stems are parsed by Prosodic Stem*

/ndi-tú-[lj-a/	PRSTEM
(a) ndí-{lj-a}	*!
☞ (b) ndi-{túu-lj-a}	

Candidate (37a) is ruled out for violating the constraint PROSODICSTEM which requires Prosodic Stems to contain at least two syllables. On the other hand, candidate (37b) is optimal because it satisfies this constraint by blocking of **-tu-** deletion.

(38) *Object markers are parsed by Prosodic Stem*

ti-tú-[vi-[βeléβet-a	PRSTEM
(a) tí-.{ví-.}{βé.lé.βee.t-a}	*!
☞ (b) ti-.{tú-.ví-.}{βe.lé.βee.t-a}	

Candidate (38a) is nullified since the monomoraic 'OM Stem' violates the constraint PROSODICSTEM. Candidate (38b) is optimal because it satisfies this constraint by blocking of **-tu-** deletion. Thus the blocking of **-tu-** deletion ensures that a proper Prosodic Stem has two proper syllables.

As a matter of interest, in most elderly speech styles non-low vowels of the object marker are optionally deleted [e.g. formal **ti-.{tú-.ví-.}{βe.lé.βee.t-a** $\Rightarrow$ **ti-.{tú-.v-.}{βe.lé.βee.t-a}** 'we are speaking them'], an indication that the target is not the same with that of the formal and common speech styles. It would appear that a minimal Prosodic Stem in the elderly speech styles is coextensive with a Strong Accent Constituent of their system which is the bimoraic Accent Foot. The vowel [u] of the morpheme **-tu-** may also undergo deletion in elderly speech styles if the vowel of the object marker is more sonorous than that of **-tu-**. When this happens, tone then docks on the most sonorous vowel of the object marker. These facts are illustrated in (39) below.

(39) **-tu-** *and the object marker form a binary Accent Foot in elderly speech styles*

Formal & common speech   *Elderly speech*

ti-{tú-.ká.}-{βe.lé.βee.t-a}   **ti-{(t-.ká.)}-{βe.lé.βee.t-a}**
we-prog-12OM-speak-fv   "we are talking about it"

ti-{tú-.ngá.}-{βe.lé.ŋg-aa.n-a}   **ti-{(t-.ngá.)}-{βe.lé.ŋg-aa.n-a}**
we-prog-6OM-read-rec-fv   "we are to each other counting them"

ti-{tú-.pá.}-{ka.mbú.l-aa.n-a}   **ti-{(t-.pá.)}-{ka.mbú.l-aa.n-a}**
we-prog-shout-rec-fv   "we are to each other shouting at it (place)"

The fact that a minimal Prosodic Stem in elderly speech styles is coextensive with the Strong Accent Constituent of their system which is the bimoraic Accent Foot is given more credence as we discuss the structure of the reduplicative Prosodic Stem and the Minimal Prosodic Word later in this chapter.

5.5.1.2 Blocking of -ku- deletion in Distant Future Tense verbs
The infinitive marker **-ku-** is usually prefixed to a Morphological Stem as given in (40) below.

(40) *Infinitive verb formation*

Inf. verb	*English gloss*
kuú-.[b-a inf-steal-fv	to steal
kuú-.[lj-a inf-eat-fv	to eat
kuú-.[mw-a inf-drink-fv	to drink
ku-.[lé.lee.s-a inf-look-fv	to look
ku-.[tú.tuu.z-a inf-push-fv	to push
ku-.[zó.mee.l-a inf-admit-fv	to admit
ku-.[zú.mbuu.-ʷa inf-reveal-fv	to reveal

Normally, infinitive marker **-ku-** is deleted when it follows the distant future tense marker **-zamu-**. This fact is illustrated in (41) below.

(41)   *Infinitive marker ku- deletes when it follows a tense marker*

Distant future verb   *English gloss*

ndi-.za.mu-.[lé.lee.s-a   I will look   *ndi-.za.mu-.**ku-**.lé.lee.sa
I-dist fut-look-fv

ndi-.za.mu-.[tú.tuu.z-a   I will push
I-dist fut-push-fv

ndi-.za.mu-.[zó.mee.l-a   I will admit
I-dist fut-admit-fv

ndi-.za.mu-.[zú.mbuu.-ʷa   I will reveal
I-dist fut-reveal-fv

ndi-.zamu-.[tha.mbá.lii.-ʲa   I will stretch
I-dist fut—stretch-fv

ndi-.za.mu-.[zo.mé.lee.z-a   I will agree
I-dist fut-agree-fv

ti-.za.mu-.[le.lé.s-aa.n-a   we will look at each other
we-dist fut-look-rec-fv

However, when the verb stem involved is monomoraic (e.g. -ba 'steal', -lja 'eat', etc), the deletion of -**ku-** is blocked as given in (42) below.

(42)   *Deletion of infinitive marker ku- is blocked when it is before monomoraic*
        verb stems

DFT verb   *English gloss*

ndi-.za.mu.-kuú-.[b-a   I will steal

ndi-.za.mu-.kuú-.[lj-a   I will eat

ndi-.za.mu-.kuú-.[mw-a   I will drink

ndi-.za.mu-.kuú-.[vw-a   I will listen

ndi-.za.mu-.kuú-.[sw-a   I will break

173

A discrepancy requiring explanation is that the infinitive marker -**ku**- can be deleted before normal Morphological Stems (with two or more syllables) but the same process is blocked before monomoraic verb stems. A possible analysis, once again, is that the blocking of -**ku**-deletion before monomoraic verb stems is triggered by the need to achieve a proper **Prosodic Stem** which contains at least two syllables as required by the Downing Constraint of PROSODICSTEM (Downing 2006b). This analysis is formalized in (43) below.

(43) -**ku**- *forms Prosodic Stem with monomoraic verb stem*

DFT verb	English gloss
ndi-.za.mu.-{kuú-.b-a}	I will steal
ndi-.za.mu-.{kuú-.lj-a}	I will eat
ndi-.za.mu-.{kuú-.mw-a}	I will drink
ndi-.za.mu-.{kuú-.vw-a}	I will listen
ndi-.za.mu-.{kuú-.sw-a}	I will break

What is surprising, however, is that -**ku**- is maintained while there is a distant future tense marker -**zamu**- which could have played the same role as -**ku**- in ensuring that the Morphological Stem is properly parsed by Prosodic Stem. Indeed, -**ku**- does not surface when there are other tense markers such as the simple past tense marker -**ŋgu**-, a signal that the Prosodic Stem minimal requirements are achieved by adding -**ŋgu**- to its fold as illustrated in (44) below.

(44) *There is no* -**ku**- *after simple past tense marker* -*ŋgu*-

Simple past	English gloss	
ndi-.{ŋguú-.ba}	I stole	*ndi-.ngu.-{kuú-.ba}
ndi-.{ŋguú-.lja}	I ate	
ndi-.{ŋguú-.mwa}	I drunk	
ndi-.{ŋguú-.vwa}	I listened	
ndi-.{ŋguú-.swa}	I broke	

174

The point I am trying to make is that there is something that blocks **-zamu-** from forming a Prosodic Stem with the monomoraic verb Stem as does the simple past tense marker **-ŋgu-**. The reason seems to be that **-zamu-** itself has a Stem status which qualifies it for Prosodic Stem parsing. This analysis is illustrated in (45) below.

(45) *DFT marker **-zamu-** seems to have Prosodic Stem status*

DFT verb   *English gloss*

ndi-.{za.mu}.-{kuú-.b-a}   I will steal

ndi-.{za.mu}-.{kuú-.lj-a}   I will eat

ndi-.{za.mu}-.{kuú-.mw-a} I will drink

ndi-.{za.mu}-.{kuú-.vw-a} I will listen

ndi-.{za.mu}-.{kuú-.sw-a} I will break

Similarly, the infinitive marker **-ku-** does not appear before the object marker when the tense marker is that of the simple past tense **-ŋgu-** as illustrated in (46) below.

(46) ***ku-** is deleted after tense markers and before object markers*

Dist Fut verb	*English gloss*	
ndi-.ŋgu-.<u>mu</u>-.[lé.lee.s-a I-past-inf-1OM-look-fv	I looked at him	*ndi-.ŋgu-.**ku**-.<u>mu</u>-.[lé.le.s-a
ndi-.ŋgu-.<u>vi</u>-.[tú.tuu.z-a I-past-inf-8OM-push-fv	I pushed them	
ndi-.ŋgu-.<u>ti</u>-.[zó.mee.l-a I-dist past-inf-13OM-accept-fv	I admitted them	
ndi-.ŋgu-.<u>ci</u>-.[zú.mbuu.-ʷa I-dist past-inf-7OM-reveal-fv	I revealed it	

However, when the tense marker is the distant future tense **-zamu-**, **-ku-** appears before an object marker as well. This fact is illustrated in (47) below.

(47) *ku-deletion is blocked after distant future tense marker* **-zamu-** *when it precedes an object marker*

DFT verb	*English gloss*
ndi-.za.mu-.ku-.<u>mu</u>-.[lé.lee.s-a	I will look at him
I-dist fut-inf-1OM-look-fv	
ndi-.za.mu.-ku-.<u>vi</u>-.[tú.tuu.z-a	I will push them
I-dist fut-inf-8OM-push-fv	
ndi-.za.mu-.ku-.<u>ti</u>-.[zó.mee.l-a	I will admit them
I-dist fut-inf-13-admit-fv	
ndi-.za.mu-.ku-.<u>ci</u>-.[zú.mbuu.-ᵂa	I will reveal it
I-dist fut-inf-7OM-reveal-fv	

My analysis of this situation once again is that both the Distant Future Tense marker and the Object Marker are morphologically marked for Stem status which qualifies them for Prosodic Stem parsing in this language. Blocking of deletion of the infinitive marker **-ku-**, therefore, ensures that the 'object marker' Prosodic Stem (so to speak) is minimally disyllabic (by PROSODICSTEM). This analysis is illustrated in (48) below.

(48) *Both the distant future tense marker and the object marker are parsed by Prosodic Stem.*

DFT verb	*English gloss*
ndi-.{za.mu.}-{ku-.<u>mu</u>.}-{lé.lee.s-a}	I will look at him
I-dist fut-inf-1OM-look-fv	
ndi-.{za.mu.}-{ku-.<u>vi</u>.}-{tú.tuu.z-a}	I will push them
I-dist fut-inf-8OM-push-fv	
ndi-.{za.mu.}-{ku-.<u>ti</u>.}-{zó.mee.l-a}	I will admit them
I-dist fut-inf-13-admit-fv	
ndi-.{za.mu.}-{ku-.<u>ci</u>.}-{zú.mbuu}.-ᵂa	I will reveal it
I-dist fut-inf-7OM-reveal-fv	

The forms above indicate that a Prosodic Word can contain three Prosodic Stems which is in violation of BINARITY. This is not surprising because, as it has been argued repeatedly, Prosodic Words are roughly Morphological Words and anything goes in complex morphological categories. For record, I close this section by suggesting that the simple past tense marker **-ŋgu-** has no Stem status and for this reason it can supply a second mora needed by the OM 'Prosodic Stem' as given in (49) below or before monomoraic verb stems as I have shown elsewhere above.

(49) ***ku-*** *is deleted after tense markers and before object markers*

DFT verb	*English gloss*
ndi-.{ŋgu-.<u>mu.</u>}-{lé.lee.s-a}	I looked at him
I-past-inf-1OM-look-fv	
ndi-.{ŋgu-.<u>vi.</u>}-{tú.tuu.z-a}	I pushed them
I-past-inf-8OM-push-fv	
ndi-.{ŋgu-.<u>ti.</u>}-{zó.mee.l-a}	I admitted them
I-dist past-inf-13OM-accept-fv	
ndi-.{ŋgu-.<u>ci.</u>}-{zú.mbuu}.-ʷa	I revealed it
I-dist past-inf-7OM-reveal-fv	

This discussion would not be complete without reiterating the prosodic structure of a normal Prosodic Stem. As we saw in chapters 3&4, the Prosodic Stem is the domain for accent and tone. By HEADEDNESS, it must contain at least one Accent Foot if it is to be properly parsed. Prosodic Stems must also satisfy MAXIMALITY CONDITION to ensure that they are minimally as large as possible. By BINARITY, Prosodic Stems are expected to contain a maximum of two Accent Feet. It has been proved that the first two and last two moras or syllables of the Prosodic Stem are earmarked for (an abstract) foot parsing. This is easy to tell when the corresponding Morphological Stem has four or more syllables as illustrated in (50) below.

(50)   *The first and last two syllables/moras of the Prosodic Stem are parsed by Accent Feet*

PrStem   *English gloss*

{(βe.le).βe(e.t-a)}   speak
speak-fv

{(βe.le).βe.t-e.s-a(a.n-a)}   cause each other to speak
speak-caus-rec-fv

{(tʰa.mba).(lii)}.-ʲa   stretch (legs)
stretch-fv

{(tʰa.mba).li.l-a.n-a(a.ŋg-a)}   be stretching legs with each another
stretch-rec-hab-fv

However, when the Morphological Stem involved has three syllables, it is not easy to tell. As we saw in the preceding chapter, however, there seems to be some evidence from tone which points to the fact that the initial syllable in trisyllabic Stems forms a unary foot as well. The forms in (51) below illustrate foot parsing in three-syllable words.

(51) *Trisyllabic Stems have two Accent Feet*

{(βe.)le(e.ŋg-a)}   read
read-fv

{(sa.)mbi(i.z-a)}   teach
teach-fv

{(le.)le(e.s-a)}   look
look-fv

{(le.)mb-e(e.s-a)}   cause to write
write-caus-fv

{(to.)lo(o.m-a)}   reprimand
reprimand-fv

The evidence can be given as follows. When toneless verbs such as those in (51) above occur in tensed verbs such as the simple past tense, tone falls on the initial syllable (mora) and not

on the heavy penultimate syllable. What hordes the tone in the initial syllable is most likely the strong accent on the initial syllable of the trisyllabic Prosodic Stem. This fact is illustrated in (52) below.

(52)   *Tone is assigned to the left monomoraic Accent Foot (first syllable) in three-syllable word*

ndi-ŋgu-{(βé.)le(e.ŋg-a)}   I read
I-past-read-fv

ndi-ŋgu-{(sá.)mbi(i.z-a)}   I taught
I-past-teach-fv

ndi-ŋgu-{(lé.)le(e.s-a)}   I looked
I-past-look-fv

ndi-ŋgu-{(lé.)mbe(e.s-a)}   I caused to write
I-past-write-caus-fv

ndi-ŋgu-{(tó.)lo(o.m-a)}   I reprimanded
I-past-reprimand-fv

Unary feet in (52) above satisfy the principle of HEADEDNESS, requiring that an Accent Foot contains at least a mora. They however violate BINARITY, requiring that Feet realize their minimal and maximal size requirement by containing two syllables or two moras. Violation of BINARITY by many Accent Feet does not come as a strange thing. Other prosodic constituents such as Prosodic Words and Prosodic Stem do the same. What is more important for a prosodic constituent, therefore, is for it to contain one of the elements it dominates (a Prosodic Word contains a Prosodic Stem, a Prosodic Stem an Accent Foot and an Accent Foot must contain a Mora). One important feature of a strong-accent language therefore is that Accent Feet can be unary. Another important feature is that the Feet are strictly located on the edges of Prosodic Stems (STEM-σ_1, FINAL-σ or PENULTIMATE-σ as guided by relevant Universal Principles). Normal disyllabic Stems also seem to have two Accent Feet, one on the initial mora and the second one on the last two moras, thereby splitting a single heavy syllable into two Accent Feet (see section 4.6.2.2 in the preceding chapter for all details).

5.5.2 Minimal Prosodic Word

We made two observations about the structure of Minimal Prosodic Words in ciTonga. To begin with, an epenthetic vowel [i] is added as a prefix before monomoraic verb stems in citation forms of formal and common speech styles. This fact is illustrated again in (53) below.

(53) *A Minimal Prosodic Word in formal and common speech styles is disyllabic*

Verb	*English gloss*
ii.-[ba	steal
ii.-[lja	eat
ii.-[swa	break
ii.-[fwa	die
ii.-[mwa	drink

In elderly speech styles, however, the Minimal Prosodic Word formed from the same Stems is simply bimoraic as repeated in (54) below[41].

(54) *A Minimal Prosodic Word in elderly speech styles is bimoraic*

Verb	*English gloss*
[baa	steal
[ljaa	eat
[swaa	break
[fwaa	die
[mwaa	drink

I will start with PBT account of Minimal Prosodic Words. The analysis in PBT is that Morphological Words and Stems map into Prosodic Words (by STEM → PROSODIC WORD HOMOLOGY). Since Prosodic Words must contain at least a Foot (by HEADEDNESS), the

---

[41] The minimal word is bimoraic even when the word is in phrase medial position (e.g. **ljaa ukóongwa** 'eat a lot' as opposed to ***lja ukóongwa**). Thus, phonological phrasing as a cause of vowel lengthening can be temporarily suspended.

prediction is that a minimal word will be disyllabic in stress languages where the level of analysis is the syllable and bimoraic in languages where the level of analysis is the mora. In other words, PBT would suggest that the Feet in formal and common speech styles are disyllabic trochee while in the elderly speech styles Feet are bimoraic. This is not insightful enough. There is no evidence that Feet in this language are disyllabic.

MBT would not solve matters either. It would suggest that the Prosodic Stems in formal and common speech styles are minimally disyllabic while in elderly speech styles they are minimally bimoraic. In other words, the two speech styles differ in terms of levels of analysis: the syllable in the former, and the mora in the latter. The challenge therefore is to unearth the principle underlying both disyllabic and bimoraic Minimal Prosodic Words in a single language system.

My suggested analysis of the variation between the formal and common speech styles, on the one hand, and the elderly speech styles on the other is that the Minimal Prosodic Word in the former is a Prosodic Stem while in the latter it is an Accent Foot. This analysis is consistent with the analysis of segment deletion in chapter 3 where it was firmly established that elderly speakers take an Accent Foot as their Minimal Strong Accent Constituent while the Minimal Strong Accent Constituent in the formal and common speech styles is the Prosodic Stem. The prevailing Guideline is the PROSODIC HIERARCHY which provides that lower prosodic constituents such as the Prosodic Stem (in the formal and common speech styles) and the Accent Foot (in the elderly speech styles) are stronger than higher prosodic constituents. The two speech communities therefore differ in terms of their choice of what should be a SAC, although both of them have access to same Universal Guidelines. A generalization requiring explanation therefore is that Minimal Prosodic Words are co-extensive with a Minimal Strong Accent Constituent of a system as guided by the Universal Principle of PROSODIC HIERARCHY. I suggest that a constraint **MINWRD≡MINSAC** is responsible for this status. Thus, a Minimal Prosodic Word can be a Prosodic Stem, an Accent Foot, or an Accent Mora/Syllable, depending on the language system. Since Prosodic Stems are required to be minimally disyllabic (by PROSODICSTEM), the Minimal Prosodic Word in the formal and common speech styles is disyllabic. It might be bimoraic in the elderly speech styles because the foot in ciTonga is essentially bimoraic (as required by FTBIN). Thus, candidates which satisfy MINWRD≡MINSAC and PROSODICSTEM, or MINWRD≡MINSAC and FTBIN, are optimal in the respective speech styles. Tableaux in (56) and (57) below illustrate. The

constraint FAITH-SAC in the tableau would as well ensure that no Strong Accent Constituent (e.g. low vowel) is inserted[42].

(55) **MINWRD≡MINSAC**

A Minimal Prosodic Word is co-extensive with a Minimal Strong Accent Constituent.

(56) *Minimal Prosodic Word in formal and common speech styles*

MINSAC: PROSODIC STEM

DG: PROSODIC HIERARCHY

/b-a/ 'steal'	MINWRD ≡MINSAC	PRSTEM	FAITH-SAC
(a) (b-aa)	*!		
(b) {b-aa}		*!	
☞ (c) {ii-.b-a}			

Candidate (56a) is disqualified because it violates the constraint MINWRD≡MINSAC which requires a Minimal Prosodic Word to match with the size of a Minimal Strong Accent Constituent such as PROSODIC STEM in the formal and common speech styles as guided by the Universal Principle of PROSODIC HIERARCHY. Candidate (56b) is banned because it violates the constraint PRSTEM which requires a canonical Prosodic Stem to contain at least two syllables (not moras). On the other hand, candidate (56c) is optimal because it satisfies both MINWRD≡MINSAC and PRSTEM.

---

[42] As a matter of interest, insertion of the high vowel [i] does not lead to violation of FAITH-SAC since high vowels are not among Strong Accent Constituents in this language.

(57) *Analysis of Minimal Prosodic Word in elderly speech styles*

MɪNSAC: Aᴄᴄᴇɴᴛ Fᴏᴏᴛ

DG: Pʀᴏsᴏᴅɪᴄ Hɪᴇʀᴀʀᴄʜʏ

/-[b-a/ 'steal'	MɪNWʀᴅ≡ MɪNSAC	FᴛBɪɴ	PʀSᴛᴇᴍ
(a) {ii-.ba}	*!		
☞ (b) (b-aa)			

Candidate (57a) is disqualified since it violates the constraint MɪNWʀᴅ≡MɪNSAC which requires a Minimal Prosodic Word to match with the size of a Minimal Strong Accent Constituent such as the Accent Foot in the elderly speech styles as guided by the Universal Principle of Pʀᴏsᴏᴅɪᴄ Hɪᴇʀᴀʀᴄʜʏ. On the other hand, candidate (57b) wins because it satisfies MɪNWʀᴅ≡MɪNSAC. I now turn to reduplication.

5.5.3 Reduplication

5.5.3.1 Formal and common speech styles

In the formal and common speech styles reduplication involved roughly repetition of an entire Morphological Stem (e.g. **ndi-.ŋgu-.βe.lé.βé.tá-[βe.le.βee.t-a**). My analysis of this fact would be that the reduplicative morpheme in ciTonga is a Stem and it is co-extensive with the prosodic constituent Prosodic Stem which hitherto has proved to be a Strong Accent Constituent in these speech styles. A generalization to be accounted for therefore is that the size of the reduplicative Prosodic Morpheme is co-extensive with a Minimal Strong Accent Constituent. A constraint I suggest to be responsible for this status is **Rᴇᴅ≡MɪNSAC**. A tableau in (59) below illustrates[43].

(58) **Rᴇᴅ≡MɪNSAC**

    The reduplicative prosodic morpheme is co-extensive with a minimal Strong Accent Constituent.

---

[43] Several violations of the constraint Fᴀɪᴛʜ-SAC are committed by some of the candidates in the following tableaux. These are omitted.

(59) *RED is co-extensive with Prosodic Stem in formal and common speech styles*

MINSAC: PROSODIC STEM

DG: PROSODIC HIERARCHY

/[RED-βeleβet-a/	RED≡ MINSAC
(a) (βe.le.)-{βe.le.βee.t-a}	*!
☞ (b) {βe.le.βe.ta}-{βe.le.βee.t-a}	

Thus, candidate (59a) is non-optimal because it violates the constraint RED≡MINSAC which requires the size of RED to match with the minimal Strong Accent Constituent such as the PROSODIC STEM as guided by the Universal Principle of PROSODIC HIERARCHY in formal and common speech styles.

A second observation was that when the morphological stem involved is monomoraic (e.g. [-**za** 'come'), RED maintains the inserted vowel [i] as in the Base (e.g. **i.za-[ii.za** 'come a lot'). My suggested analysis of this observation is that RED is co-extensive with Prosodic Stem which is a Strong Accent Constituent in formal and common speech styles as guided by the Universal Principle of Prosodic Hierarchy. The constraint RED≡MINSAC would ensure this status. In addition, Prosodic Stems are required to be minimally disyllabic (by PRSTEM). Hopefully, the constraint FAITH-SAC in the tableau in (60) below ensures that no Strong Accent Constituent is inserted.

(60) *RED is subjected to Prosodic Stem minimality condition*

MINSAC: PROSODIC STEM

DG: PROSODIC HIERARCHY

RED-lj-a 'eat'	RED≡ MINSAC	PRSTEM	FAITH- SAC
(a) (ljaa)-{ii-.lj-a}	*!		
☞ (b) {i.lja}-{ii-.lj-a}			

Candidate (60a) is ruled out because it violates the constraint RED≡MINSAC which requires reduplicative Stems to be co-extensive with a Strong Accent Constituent such as a PROSODIC STEM in the formal and common speech styles as guided by the Universal Principle of PROSODIC HIERARCHY. Candidate (60b), on the other hand, is optimal because it satisfies both RED≡MINSAC and PRSTEM.

5.5.3.2 Elderly speech styles

The observation about reduplication in the elderly speech styles was that most elderly speakers have a tendency to copy just the first two syllables of the Base Stem (e.g. **ndi-.ŋgu-.βé.lé.-[βe.le.βee.t-a**). In most cases, the second vowel of the partial RED is also elided (e.g. **ndi-.ŋgu-.βé.l.-[βe.le.βee.t-a**). The analysis I suggest is that reduplicative forms which have the size of an Accent Foot are optimal in the elderly speech styles. The choice of Accent Foot as the size of RED seems to be based on the fact that a bimoraic ACCENT FOOT, and not PROSODIC STEM, is the Minimal Strong Accent Constituent in these speech styles. The guiding principle is, once again, the PROSODIC HIERARCHY. A generalization requiring explanation therefore is that RED is co-extensive with a Minimal Strong Accent Constituent such as an Accent Foot. A constraint responsible for this status, once again, is RED≡MINSAC. A tableau in (61) below illustrates.

185

(61) *RED is co-extensive with Accent Foot in elderly speech styles*

MᵢɴSAC: Aᴄᴄᴇɴᴛ Fᴏᴏᴛ

DG: Pʀᴏsᴏᴅɪᴄ Hɪᴇʀᴀʀᴄʜʏ

/RED-βeleβet-a/ 'speak'	RED≡ MᵢɴSAC	Fᴛʙɪɴ
(a) {βe.le.βe.ta}-{βe.le.βee.t-a}	*!	
(b) {βe.le}.-{βe.le.βee.t-a}	*!	
(c) {βe.l}.-{βe.le.βee.t-a}	*!	
(d) (βe.)-{βe.le.βee.t-a}		*!
☞ (e) (βe.le).-{βe.le.βee.t-a}		
☞ (f) (βe.l).-{βe.le.βee.t-a}		

Candidates (61a-c) are disqualified because they violate the constraint RED≡MᵢɴSAC which requires the size of RED to match with a system's Minimal Strong Accent Constituent such as the Aᴄᴄᴇɴᴛ Fᴏᴏᴛ in the elderly speech styles as guided by the Universal Principle of Pʀᴏsᴏᴅɪᴄ Hɪᴇʀᴀʀᴄʜʏ. Candidate (61d) is ruled out since it violates the constraint Fᴛʙɪɴ which requires Accent Feet to be binary at the level of the mora in this system. On the other hand, candidates (61e-f) are both attested in elderly speech styles and they win because they satisfy both RED≡MᵢɴSAC and Fᴛʙɪɴ[44].

Another observation about reduplication in the elderly speech styles was that when a morphological stem involved is monomoraic, there is no insertion of [i]. Instead, the verb stem has a long vowel (e.g. **za** $\Rightarrow$ **z-aa**). When reduplicating such verbs, the vowel of the RED Stem is long as well (e.g. **zaa-[z-aa**). Since reduplicative Stems are co-extensive with Accent Feet in these speech styles, the analysis of this observation is that a minimal reduplicative Prosodic Morpheme is Foot-sized. Since Accent Feet in this language are required to be

---

[44] Expansion of the size of RED to three or more syllables (e.g. βe.le.βe.-{βe.le.βee.t-a} is most likely influenced by the need to minimize violation of the constraint MAX-BR requiring the reduplicative Prosodic Stem to preserve the entire Base Prosodic Stem.

bimoraic, the bimoraic reduplicative morpheme satisfies this requirement. Once again, the generalization is that reduplicative Prosodic Morphemes are co-extensive with a Minimal Strong Accent Constituent. The constraint RED≡MINSAC seems to be responsible for this status. A tableau in (62) below illustrates (I assess RED only).

(62) *RED is size of an Accent Foot*

MINSAC: ACCENT FOOT

DG: PROSODIC HIERARCHY

RED-z-a 'come'	RED ≡ MINSAC	FTBIN
(a) (za)-(z-aa)		*!
(b) {i.za}-(z-aa)	*!	
☞ (c) (zaa)-(z-aa)		

Candidate (62a) is disqualified because it violates the constraint FTBIN which requires Accent Feet to be binary at the level of the mora. Candidate (62b) is ruled out for violating the constraint RED≡MINSAC which requires RED to be coextensive with a Minimal Strong Accent Constituent such as an ACCENT FOOT in the elderly speech styles as guided by the Universal Principle of PROSODIC HIERARCHY. Candidate (62c) wins since it satisfies both RED≡MINSAC and FTBIN. A crucial fact is that a bimoraic RED Stem in the elderly speech is not achieved through insertion of a syllable as it is the case in the formal and common speech styles. This would be a clear indication that disyllabicity (or proper Prosodic Stem) is not actively strived for in this speech style.

Summary

This section has indicated that the Strong Accent Constituent-Based Templates Theory maintains the PBT set-up where metrical constituents (Foot, Syllable and Mora) are not separated from morpho-prosodic categories. However, it also includes the Prosodic Stem in the Prosodic Hierarchy as argued for in MBT and others. Strong Accent Constituent-Based Templates also maintain the PBT idea of Strict Layer Hypothesis especially its crucial clause

of HEADEDNESS which requires constituents in the prosodic hierarchy to be properly parsed (Sekirk 1978/81, 1984, 1995; Nespor and Vogel 1986; Mester 1992, Orie 1997, Selkirk 1995). Following this idea, each Prosodic Word is supposed to be properly headed by containing at least a Prosodic Stem. A Prosodic Stem contains at least an Accent Foot, Accent Foot at least a Syllable and Syllable contains a Mora. Prosodic constituents must also satisfy MAXIMALITY CONDITION to ensure that they are minimally as large as possible.

Many Prosodic Words have one Prosodic Stem. This state of affairs is not in violation of the principle of HEADEDNESS, requiring a prosodic constituent to contain at least one of the units it dominates. Prosodic Words must also satisfy MAXIMALITY CONDITION to ensure that they are minimally as large as possible, i.e. they must contain a minimum and maximum of two Prosodic Stems. Since Prosodic Words are first and foremost morphological categories, the idea of maximality is not always satisfied. The principle of BINARITY is thus largely violated because many Prosodic Words may contain just one Prosodic Stem or more than two Prosodic Stems. This is not surprising because, it has been repeatedly argued, Prosodic Words are roughly Morphological Words.

This section has also shown that the Prosodic Stem is at the centre of Strong Accent Constituency. By HEADEDNESS, it must contain at least one Accent Foot if it is to be properly parsed. Prosodic Stems must also satisfy MAXIMALITY CONDITION to ensure that they are minimally as large as possible. By BINARITY, Prosodic Stems are expected to contain a maximum of two Accent Feet. It has been proved that in normal circumstances (with Stems of four or more syllables) the first and last two syllables are earmarked for Accent Feet. This satisfies the MAXIMALITY CONDITION and the principle of BINARITY. However, there is an additional condition for the Prosodic Stem: It must also contain minimally two syllables (by PROSODICSTEM) as well supported by the principle of MORPHEME-SYLLABLE CORRELATION PRINCIPLE in Downing's (2006b) MBT.

This section has also presented a SAC account of the disyllabic Minimal Prosodic Word in formal and common speech styles on the one hand, and the bimoraic Minimal Prosodic Word in the elderly speech styles on the other. A suggested analysis has been that the Minimal Prosodic Word in the former is a Prosodic Stem while in the latter it is an Accent Foot. Elderly speakers take an ACCENT FOOT as their minimal Strong Accent Constituent while the minimal Strong Accent Constituent in the formal and common speech styles is the PROSODIC STEM. The prevailing Guideline is the PROSODIC HIERARCHY which provides that lower

prosodic constituents such as the Prosodic Stem (in the formal and common speech styles) and the Accent Foot (in the elderly speech styles) are stronger than higher prosodic constituents. The generalization requiring explanation was that Minimal Prosodic Words are co-extensive with a Strong Accent Constituent as guided by the Universal Principle of PROSODIC HIERARCHY. The constraint **MINWRD≡MINSAC** was suggested as being responsible for this status. Since Prosodic Stems are required to be minimally disyllabic (by PROSODICSTEM), the Minimal Prosodic Word in the formal and common speech styles is disyllabic. It is bimoraic in the elderly speech styles because the foot in ciTonga is essentially bimoraic (as required by FTBIN).

Finally, this section has also presented a SAC analysis of reduplication facts. The observation in the formal and common speech styles was that reduplication involves roughly repetition of an entire Morphological Stem (e.g. **ndi-.ŋgu-.βe.lé.βé.tá-[βe.le.βee.t-a**). The analysis of this fact has been that the reduplicative prosodic morpheme is a Stem and it is co-extensive with the prosodic constituent Prosodic Stem which hitherto has proved to be a Strong Accent Constituent in these speech styles. In the elderly speech styles, however, only the first two syllables of the Base (e.g. **ndi-.ŋgu-.βé.lé.-[βe.le.βee.t-a**) may be copied. The suggested analysis has been that reduplicative forms which have the size of an Accent Foot are optimal in the elderly speech styles. The choice of Accent Foot as the size of RED seems to be based on the fact that a bimoraic ACCENT FOOT, and not PROSODIC STEM, is the Minimal Strong Accent Constituent in these speech styles. The guiding principle in both cases is the PROSODIC HIERARCHY and the constraint responsible for their shapes is RED≡MINSAC, requiring RED to be co-extensive with a Minimal Strong Accent Constituent of a system.

5.6   Chapter summary

The aim of this chapter was to present formally proposals for Strong Accent Constituent-Based Templates as a slightly better theory to account for morphology-prosody interfaces in ciTonga and perhaps many other Bantu languages with a Strong Accent system. I began the chapter by presenting the data on general phonological words, Minimal Prosodic Words and reduplicative Prosodic Stems. I then reviewed two theories within the Generalized Templates Theory namely, the Prosodic Hierarchy-Based Templates Theory and the Morpheme-Based Templates Theory both of which have a goal to account for morphology-prosody interfaces. Both these theories have been found to be slightly inadequate to account for parameters exhibited by phonological words in ciTonga. On the other hand, a theory of Strong Accent Constituent-Based Templates has been found to account for the facts slightly better.

Chapter 6

Summary and Conclusions

6.1   Introduction

The aim of this study was to investigate word-prosody in ciTonaga, a Malawian Bantu language. This chapter presents summary and conclusions of the study.

6.2   Summary

The dissertation has been presented in five chapters. Chapter 1 presented introductory remarks. These included the problem statement, a note on methodology, summary of findings, theoretical precedents, and, finally, organization of the dissertation. Chapter 2 presented basic facts about the language under study (ciTonga). These included language classification, previous works on ciTonga, speech sounds, the syllable, tone, as well as nominal and verbal morphology.

Chapter 3 presented a proposal for the theory of Strong Accent Constituency. It presented the data on vowel and consonant deletion which motivates this theory analysis. Attempts were made to account for the facts in terms of stress-accent theory and Downing's (2006b) Morpheme-Based Templates Theory both of which were found to be slightly problematic to account for segment deletion and preservation patterns exhibited in ciTonga. Finally, the chapter introduced the theory and accounted for the facts in terms of Strong Accent Constituency.

Chapter 4 presented formally the theory of Interaction between Tone and Strong Accent Constituents. The chapter began by presenting the data on tone assignment in basic verbs, simple past tense verbs and present progressive aspect verbs which motivated this type of theory analysis. Attempt was made to account for the facts in terms of tone alignment theory (as argued for by Mtenje 2006), autosegmental accent (Clements and Goldsmith 1984) and pitch-accent (as hinted upon by Downing 2004). All these theoretical perspectives were found to be slightly inadequate to account for tone distribution patterns in ciTonga. On the other hand, a theory based on Interaction between Tone and Strong Accent Constituents was shown to account for the facts slightly better.

Chapter 5 presented formally proposals for Strong Accent Constituent-Based Templates as a theory of morphology-prosody interfaces in ciTonga and perhaps many other Bantu languages

with a Strong Accent Constituent system. The chapter began by presenting the data on general phonological words, Minimal Prosodic Words and reduplicative Prosodic Stems. It then reviewed two theories within the Generalized Templates Theory namely, the Prosodic Hierarchy-Based Templates Theory and the Morpheme-Based Templates Theory both of which have a goal to account for morphology-prosody interfaces. Both these theories were found to be slightly inadequate to account for parameters exhibited by phonological words in ciTonga. On the other hand, it was suggested that a theory of Strong Accent Constituent-Based Templates would account for the facts slightly better.

6.3   Conclusions

Many conclusions would be made out of this study, but I will stick to issues which have been presented. In general, the present study makes three important conclusions regarding the prosodic structure of phonological words in ciTonga. The general conclusion is that the idea of **Universal Guidelines and Constraints for Strong Accent Constituents** accounts for the facts slightly better. Primarily, these facts included vowel and consonant deletion and preservation, high tone assignment and morphology-prosody interfaces. Three sub-theories, namely, Strong Accent Constituents Theory, Interaction between Tone and Strong Accent Constituents, and Strong Accent Constituent Templates, have been shown to account for these facts slightly better.

6.3.1   Strong Accent Constituents Theory

One generalization from the observations made about vowel and consonant preservation and deletion in ciTonga is that Universal Principles such as SONORITY, EDGENESS and PROSODIC HIERARCHY play a crucial role. For instance, it was clear that there was strong relationship between segment faithfulness, on the one hand, and segment sonority, edgeness and a level in which it appears in the Prosodic Hierarchy, on the other hand. The conclusion has been that every prosodic unit (Phonological Phrase, Prosodic Word, Prosodic Stem, Foot, syllable, mora, segment, feature and any smallest atom) in a particular language is an Accent Constituent (unit of prominence) and that what is preserved, especially in relation to Accent, is a **Strong Accent Constituent** of the relevant domain (Language, Utterance, Phrase, Word, Stem, or Accent Foot).

Another conclusion has been that what is Strong Accent Constituent in one speech community or context may not necessarily be stronger in another context. At the heart of Strong Accent Constituent Theory, therefore, is the fact that a finite set of Universal Guidelines such as

EDGENESS, FINALITY, SONORITY, PROSODIC HIERARCHY, STRESS, SYLLABLE WEIGHT, PITCH, and several others, play a crucial role in determining what is strong and not strong and that what counts as Determiner Guideline (DG) of strength in one language may not necessarily be so in another language. The Guidelines may jointly make a determination, relegating some Guideline in the process.

Since language- or context-specific Strong Accent Constituents are the ones which are usually preserved, it appears that there is only one **FAITHFULNESS** constraint which is responsible for this status in as far as the grammar of Strong Accent Constituency is concerned: **FAITH - STRONG ACCENT CONSTITUENT (FAITH-SAC)**, requiring that Strong Accent Constituents be preserved in the output. What happens to weaker ACs, it seems, is of little or none of the accent grammar's business. Another constraint **ACFT/SAC** has been shown to be responsible for the fact that Accent Feet belong to Strong Accent Constituents PROSODIC STEM, ACCENT FOOT, STEM-σ_1, PENULTIMATE-σ or FINAL-σ as determined and provided by the said Universal Guidelines.

6.3.2 The Interaction between Tone and Strong Accent Constituents

The conclusion that can be made about the interaction between tone and Strong Accent Constituency is that tone belongs to Strong Accent Constituents such as PROSODIC STEM, ACCENT FOOT, STEM-σ_1, PENULTIMATE-σ or FINAL-σ as determined by the Universal Guidelines of PROSODIC HIERARCHY, EDGENESS, FINALITY and SONORITY, among others. This generalization has been accounted for in terms of the constraint TONE/SAC which requires high tones to be assigned to Strong ACs.

Furthermore, the constraint SPREAD ensures that high tones spread to next TBUs. However, the study has established that blocking of tone spreading by the last two syllables of the Prosodic Stem/Word/Phrase is attributed to another constraint IDENT-HPP which requires identity of Head Prosodic Phrases (penultimate syllables) to be preserved. Deletion of the present progressive aspect marker **-tu-** before normal verb Stems (with two or more syllables) has been attributed to the constraint HEADEDNESS outranking FAITH-SAC. The principle of HEADEDNESS ensures that Accent Feet belong to the Prosodic Stem (since **-tu-** is a lexically specified SAC, and thus automatically qualifies for foot parsing – kind of PARSE-SAC constraint at play). Once deleted, the tone which was meant for the foot head **-tu-** is re-assigned following requirements of the Association Convention that tones should be assigned from left to right in a one-to-one fashion. The fact that **-tu-** surfaces before monomoraic verbs

would be indication that it is now parsed by Prosodic Stem and it no longer violates HEADEDNESS.

6.3.3 Strong Accent Constituent-Based Templates Theory

Several observations were made about the structure of Phonological Words and Stems in ciTonga. One conclusion has been that many Prosodic Words satisfy the principle of HEADEDNESS by containing at least one Prosodic Stem. However, most Prosodic Words fail to satisfy the principle of BINARITY because many of them may contain just one Prosodic Stem or more than two Prosodic Stems. This, it has been argued, is not surprising because Prosodic Words are roughly Morphological Words and anything goes in Morphological categories.

Another crucial fact is that the Prosodic Stem seems to be at the centre of Strong Accent Constituency. By HEADEDNESS, it must contain at least one Accent Foot if it is to be properly parsed. Prosodic Stems must also satisfy MAXIMALITY CONDITION to ensure that they are minimally as large as possible. By BINARITY, Prosodic Stems are expected to contain a maximum of two Accent Feet. It has been suggested that the first two and last two moras or syllables of the Prosodic Stem are earmarked for (abstract) Accent Feet. However, there is an additional condition for the Prosodic Stem: It must contain minimally two syllables (by Downing's PROSODICSTEM).

Another crucial point which has been made about Prosodic Words and Stems is that the Minimal Prosodic Word in formal and common speech styles is co-extensive with a Prosodic Stem while in elderly speech styles it is co-extensive with an Accent Foot. Since Prosodic Stems are required to be minimally disyllabic (by PROSODICSTEM), the Minimal Prosodic Word in the formal and common speech styles is disyllabic. It is bimoraic in the elderly speech styles because the foot in ciTonga is essentially bimoraic (as required by FTBIN). This variation is due to the fact that elderly speakers take an Accent Foot as their Minimal Strong Accent Constituent while the Minimal Strong Accent Constituent in the formal and common speech styles is the Prosodic Stem. The prevailing Guideline is the Prosodic Hierarchy which provides that lower prosodic constituents such as the Prosodic Stem (in the formal and common speech styles) and the Accent Foot (in the elderly speech styles) are stronger than higher prosodic constituents. The generalization has been that Minimal Prosodic Words are co-extensive with a Strong Accent Constituent of a system as determined by the Universal

Principle of Prosodic Hierarchy. The constraint **MINWRD≡MINSAC** has been suggested as being responsible for this status.

Finally, the reduplicative prosodic stem in formal and common speech styles is co-extensive with the prosodic constituent Prosodic Stem which hitherto has proved to be a Strong Accent constituent in these speech styles. In the elderly speech styles, however, reduplicative forms which have the size of an Accent Foot are optimal. The choice of Accent Foot as the size of RED seems to be based on the fact that a bimoraic ACCENT FOOT, and not PROSODIC STEM, is the Minimal Strong Accent Constituent in the elderly speech styles. The guiding principle in both cases is the Prosodic Hierarchy and the constraint responsible for their shapes is **RED≡MINSAC**, requiring RED to be co-extensive with a Minimal Strong Accent Constituent of a system.

6.3.4 Stress Accent vs. Strong Accent

Although this is not very obvious, distinctions can also be made between Stress Accent and Strong Accent as exhibited in ciTonga. One of them concerns their domains of operation. Stress Accent in ciTonga belongs to the Phonological Phrase and it is assigned to the most salient position of the phrase, the phrase-penultimate syllable (somehow guided by the Universal Guidelines of EDGENESS and FINALITY). On the other hand, Strong Accent Feet belong exclusively to the Prosodic Stem.

One other crucial distinction between a stress accent Foot and a Strong Accent Foot therefore might be that Strong Accent may be sensitive to Sonority while Stress Accent Foot, being assigned postlexically, is not. Thus, while a Strong Accent Foot shifts to the right most edge due to sonority properties of the final syllable, the Stress Foot (in bold and cued by PL) remains on the penultimate syllable. These facts are illustrated in (1) below. A right Strong Accent Foot is on the penultimate syllable, rather than the final syllable, only when the final syllable has a consonant onset or a vowel nucleus which is of poor quality in terms of sonority (e.g. liquids and non-low vowels). A strikethrough on a segment (e.g. 'e̶') indicates that the segment may be deleted in common speech styles and the syllable in which it appears is most likely extrametrical.

(1) *Stress Accent and Strong Accent have no effect on each other*

Word/Stem	English gloss
{(βe.le).'**βe**(**e**.t-a)}	speak
{(sa).'**mbi**(**i**.z-a)}	teach
{(tʰa.mba).li.l-a.'**n-a**(**a**.ŋga)}	be stretching legs over each other
ti-.{(le.le).('**s-aa**).ne̱}	we should look at each other
ndi-{(βe.le).('**βee**).te̱}	I should speak
{(tha.mba).('**lii**).-ɬa}	stretch

Another crucial distinction between Stress Accent and Strong Accent concerns the Accent Bearing Unit (ABU) and phonetic correlate. The Stress Accent Bearing Unit is undisputably the syllable and it is cued by vowel lengthening. On the other hand, any unit of language such as Prosodic Stem, Accent Foot, Syllable, Segment, Mora and Features can bear Strong Accent. There is no phonetic correlate save the fact that Strong Accent Constituents are preserved when everything else seems to be falling. These facts have been adequately illustrated in the dissertation, except facts about Features and smallest atoms of human language which may also bear Strong Accent. These would require separate treatment.

Finally, Stressed Syllables and Stress Feet can be Strong Accent Constituents since they are units of great prominence. However, it appears that not all Strong Accent Constituents or Accent Feet have stress.

6.3.5 Tone and Stress Accent

Crucial conclusions can also be drawn about the interaction between Tone and Stress Accent in ciTonga. One of them is that the domain for tone is the Prosodic Stem. On the other hand, the domain for Stress Accent is the Phonological Phrase. To this effect, Stress Accent and Tone have no effect on each other. Thus a high tone can be right on the back of the stressed syllable, but it cannot be attracted to it [e.g. ndi.ngu.(βé).'**le**(**e**nga) 'I read'].

However, the right edge tone may coincide with stress on the penultimate syllable [e.g. *sa.mbi.('ʑii).ʲa* 'help solve a clue']. This is just due to the fact that both Stress Accent and Tone are required to be assigned to a Strong Penultimate Syllable as guided by Universal

Guidelines of EDGENESS&FINALITY, but not necessarily due to the fact that Stress Accent attracts tone. This conclusion helps us understand better the fact that tone will shift to the final syllable/mora of the same Stem when it (the Stem) is in phrase-medial position, while stress completely abandons the Stem to dock onto a new Stem in the phrase-final position, as illustrated in (2) below.

(2) *Tone does not shift to new penultimate syllable Stems as Stress Accent does*

Word/Phrase		*English gloss*
{sa.mbi.'**zií**.ja}		reveal answers
{sa.mbi.zi.já}	mu.nthi.'**kaa**.ze	reveal answers to a woman
{chi.'**mbií**.ja}		run
{chi.mbi.já}	wu.'**kóo**ngwa	run a lot
{to.nde.'**kaá**.na}		fail each other
{to.nde.ka.ná}	wu.'**heé**.ni	fail each other badly

Thus the right edge high tone and the phrase-penult Stress Accent have no effect on each other, save the fact that they both adhere to a constraint that requires both of them to fall on a Strong Accent Constituent of the appropriate domain.

Another crucial conclusion to draw concerns the Tone- and Stress Accent- Bearing Units. Unlike Stress Accent which takes the syllable as its bearing unit, the tone bearing unit, it seems, is the mora. Thus, the right edge tone may fall on the penultimate mora of the Prosodic Stem following Universal Guidelines of EDGENESS&FINALITY. Like Stress Accent Feet, Tone is never attracted by sonority as do Strong Accent Feet.

6.4   For further research and discussion

The present study draws conclusions mainly from one word class, the verb, of one language only, ciTonga. There is need for further research to establish whether other word categories and other languages too take Strong Accent as their mode of word-prosody. Further research is also needed to establish if certain distinctive features also fit the mode of Strong Accent Constituency.

A more crucial point which is still under debate in Bantu linguistics is the relationship between phrasing and focus. As Downing (2007) observes, the prosodic behavior of languages like Chichewa, Tumbuka and ciTonga makes them "relevant for investigating the question of whether prosodic prominence or phrasing is the primary correlate of focus." However, much as sentiments made in this study would have a bearing on phonological phrasing and its relationship with focus, or the other way round, I leave this issue for further research and discussion.

Finally, I have not shown how the idea of Strong Accent Constituency would help solve the problem of Opacity which continues to be a subject of debate in Optimality Theory. It is my hope that the direction taken in this dissertation will help ease the tensions. Very crucially, we need to do more research to establish firmly whether Strong Accent is indeed distinct from Stress Accent.

References

Akinlabi, Akibinyi (1996) "Featural affixation." *Journal of Linguistics* 32: 239-89.

Alderete, J. (1995) Faithfulness to prosodic heads. Ms., University of Massachusetts, Amherst.

Alderete, John, Beckman, Jill, Benua, Laura, Gnanadesikan, Amalia, McCarthy, John, and Urbanczyk, Suzanne (1999) "Reduplication with fixed segmentism." *Linguistic Inquiry* 30: 327-64.

Anderson, John M. and Ewen, Colin J. (1987) *Principles of Dependency Phonology*. Cambridge: Cambridge University Press.

Banda, Joseph T.K. (1985) "Proverbs in Tonga-proverbe N.M3574O" Unpublished.

Barnes, Jonathan (2002) Positional neutralization, a phonologization approach to typological patterns. PhD dissertation, University of California, Berkeley.

Beckman, J. (1995) "Shona height harmony: Markedness and positional identity." In J. Beckman, L. Walsh-Dickey, and S. Urbanczyk (eds.). *University of Massachusetts Occasional Papers in Linguistics 18*, , 53-75. Amherst: Graduate Linguistics Students Association of Massachusetts Amherst.

Beckman, Jill N. (1997) "Positional faithfulness, positional neutralization and Shona vowel harmony." *Phonology* 14, 1-46.

Beckman, J. (1998) Positional Faithfulness. Ph.D dissertation, University of Massachusetts, Amherst.

Bensoukas, K. (2004) "Has there been a shift in Generative Phonology? The case of Optimality Theory." In: *Le éléments Paradigmatiques Thematiques et Stylistiques dans la Pensée Scientifique*, 231-262. Faculté de Letters, Rabat.

Boersma, P. and Weenink, D. (2012) *Praat: doing phonetics by computer* (Computer Program). www.praat.org.

Bryan, M.A. (1959) *The Bantu Languages of Africa*. London, New York, Capetown: Oxford University Press.

Buller, Barbara, Buller, Ernest, and Everett Daniel L. (1993) "Stress placement, syllable structure and minimality in Banawa." *International Journal of Ammerican Linguistics* 59: 280-293.

Casali, R. (1996) Resolving Hiatus. Ph.D dissertation, University of California at Los Angeles.

Cassimjee, Farida (1998) *IsiXhosa Tonology: An Optimality Domains Theory. LINCOM Studies in African Linguistics 29.* Munich: LINCOM EUROPA.

Cassimjee, Farida & Charles Kisseberth W. (1998) "Optimality domains theory and Bantu tonology: a case study from Isixhosa and Shingazidja." In Larry M. Hyman and Charles W. Kisseberth (eds.), *Theoretical Aspects of Bantu Tone*, 33-132. Stanford, Calif.: Centre for the Study of Language and Information.

Chirwa, Filemon K. (1932) *Nthanu za Chitonga*. Mzuzu: Livingstonia Mission.

Chomsky, N. (1957) *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. (1964) *Current Issues in Linguistic Theory.* The Hague, Paris: Mouton.

Chomsky, N. (1965) *Aspects of the Theory of Syntax*. Cambridge, Massachusetts.: MIT Press.

Chomsky, N. and Halle, M. (1968) *The Sound Pattern of English*. New York: Harper and Row.

Clements, G.N. and Goldsmith, J. (eds.) (1984) *Autosegmental Studies in Bantu Tone*. Dordrecht: Foris Publications, 281-308.

Comrie, B. (1976) *Aspect*. Cambridge: Cambridge University Press.

Crosswhite, Katherine M. (1998) Vowel reduction in Optimality Theory. Doctoral dissertation, University of California at Los Angeles.

Crosswhite, Katherine M. (2004) "Vowel reduction." In Hayes, Bruce, Robert Kirchner and Donca Steriade (eds.), *Phonetically Based Phonology*, 191-231. Cambridge: Cambridge University Press.

Dahlstrom, Amy (1997) "Fox (Mesquakie) reduplication." *International Journal of American Linguistics* 63: 205-261.

De Lacy, Paul (2002a) The formal expression of markedness. Doctoral dissertation, UMass. [Also ROA 542].

De Lacy, P (2002b) "The interaction of tone and stress in optimality theory." *Phonology* 19: 1-32.

De Lacy, Paul (2004) "Markedness conflation in Optimality Theory." *Phonology* 21: 145 199.

De Lacy, Paul (2006) *Markedness: Reduction and Preservation in Phonology*. Cambridge: Cambridge University Press.

De Lacy, Paul (2007) "The interaction of tone, sonority, and prosodic structure." In Paul de Lacy (ed.), *The Cambridge Handbook of Phonology*, 281-308. Cambridge: Cambridge University Press.

Doke, Clement M. (1967) *The Southern Bantu Languages*. London: Dawsons of Pall Mall.

Downing, Laura J. (1998) "Prosodic misalignment and reduplication." In Geert Booij and Jaap van Marle (eds.) *Yearbook of Morphology 1997*. Dordrecht: Kluwer Academic Publishers: 83-120.

Downing, Laura J. (1999) "Prosodic Stem ≠ Prosodic Word in Bantu." In Hall, T. Alan and Kleinhenz, Ursula (eds.), *Studies on the Phonological Word*. Amsterdam: John Benjamins: 73-98.

Downing, Laura J. (2000) "Morphological and prosodic constraints on Kinande verbal reduplication." *Phonology* 17: 1-38.

Downing, Laura J. (2003) "Compounding and tonal non-transfer in Bantu languages." *Phonology* 20: 1-42.

Downing, Laura J. (2004) "What African languages tell us about accent typology." *ZAS Papers in Linguistics* 37: 101-136.

Downing, L.J. (2006a) "The prosody and syntax of focus in Chitumbuka." In Laura J. Downing, Lutz Marten & Sabine Zerbian (eds.), *Papers in Bantu Grammar and Description,* 55-79. Berlin: ZAS Papers in Linguistics 43.

Downing, L.J. (2006b) *Canonical Forms in Prosodic Morphology*. Oxford: Oxford University Press.

Downing, L.J. (2007) Focus prosody divorced from stress and intonation in Chichewa, Chitumbuka and Durban Zulu.
www.linguistics.ucla.edu/people/jun/workshop2007ICPhS/Papers/DowningICPhS_2007.PDF
Downloaded on 1st December 2013.

Dresher, B. Elan and Van der Hulst, Harry (1998) "Head-dependent asymmetries in phonology: complexity and visibility." *Phonology* 15: 317- 52.

Feng, G. Bella (2004) "Phonological restrictions on morphemes in shaping reduplicated words." Presented at the Workshop on Word Domains: Theory and Typology, University of Leipzig, 7-8 April 2004.

Garrett, Edward (1999) "Minimal Words aren't minimal feet." *UCLA Working Papers in Linguistics* 1: 68-105.

Glaser, B. (1992) *Basics of Grounded Theory Analysis*. Mill Valley, CA: Sociology Press.

Goedemans, Rob (1996) "An Optimality account of Onset sensitivity in quantity-insensitive languages." *Linguistic Review* 13: 33-48.

Goldsmith, John (1976) Autosegmental phonology. Ph.D dissertation, MIT. Reproduced by Indiana University Linguistics Club.

Goldsmith, John (1979) "The aims of autosegmental phonology." In D. Dinnsen (ed.), *Current Approaches to Phonological Theory*, 202-22. Bloomington: Indiana University Press.

Goldsmith, John (1984) "Tone and accent in Tonga." In G.N Clements and J. Goldsmith (eds.), *Autosegmental Studies in Bantu Tone*, 105-182. Dordrecht: Foris.

Goldsmith, John (1990) *Autosegmental and Metrical Phonology*. Cambridge, Mass.: Basil Blackwell.

Gordon, Matthew (1999) "Syllable weight: phonetics, phonology and typology," Ph.D dissertation, UCLA.

Gouskova, Maria (2003) Deriving economy: Syncope in Optimality Theory. Ph.D dissertation, University of Massachusetts Amherst.

Guthrie, M. (1948) *The Classification of the Bantu Languages*. London, New York, Toronto: Oxford University Press.

Halle, M. and Vergnaud, J.R. (1987) *An Essay on Stress*. Cambridge, Mass.: MIT Press.

Haraguchi, S. (1977) *The Tone Pattern of Japanese: An Autosegmental Theory of Tonology*. Tokyo: Kaitakusha.

Harris, John (1990) "Segmental complexity and phonological government," *Phonology* 7: 255-300.

Harris, John (1994) *English Sound Structure*. Oxford: Blackwell.

Harris, John (1997) "Licensing inheritance: an integrated theory of neutralization." *Phonology* 14: 315-370.

Harris, John (2004) "Release the captive coda: the foot as a domain of phonetic interpretation." In J. Local, R. Ogden, and R. Temple (eds.), *Phonetic Interpretation: Papers in Laboratory Phonology 6*, 103-29. Cambrdge: Cambridge University Press.

Hayes, Bruce (1980) A metrical theory of stress rules. Ph.D dissertation, MIT [Published 1985, New York: Garland.]

Hayes, B. (1995) *Metrical Stress Theory*. Chicago: The University of Chicago Press.

Hill, Kenneth C., Emory Sekaquaptewa, H. Andrew Black, Ekkehart Malotki, & Michael Lomatuway'ma (1998) "A sketch of Hopi grammar." In *Hopi Dictionary, Hopiikwa Lavaytutuveni. A Hopi-English-Dictionary of the Third Mesa Dialect.* Tucson: University of Arizona Press.

Hocket, Charles F. (1958) *A Course in Modern Linguistics*. New York: McMillan.

Hyman, Larry M. (1977) "On the nature of linguistic stress." In L.M. Hyman (ed.) *Studies in Stress & Accent. Southern California Occasional Papers in Linguistics 4*: 37-82.

Hyman, Larry M. (1993) "Conceptual issues in the comparative study of the Bantu verb stem." In S.S. Mufwene and L. Moshi (eds.), *Topics in African Linguistics*, 3-34. Amsterdam: John Benjamins.

Hyman, Larry M. (2007) "How not to do phonological typology: The case of Pitch-Accent." UC Berkeley Phonology Lab Report (2007).

Hyman, Larry M. (2009) "Penultimate lengthening in Bantu." University of California Berkeley Phonology Lab Annual Report (2009). University of California, Berkeley. [Presented at the Third Conference on Bantu Linguistics, Tervuren, March 25-27, 2009].

Hyman, Larry, and Ernest Byarushengo (1984) "A model of Haya tonology." In G.N. Clements and J. Goldsmith (eds.), *Autosegmental Studies in Bantu Tone*, 19-51. Dordrecht: Foris.

Hyman, Larry M., Inkelas, Sharon, and Sibanda Galen (1999) "Morpho-syntactic correspondence in Bantu reduplication." MS, University of California, Berkeley.

Hyman, Larry M. and Mtenje, Al (1999) "Prosodic morphology and tone: the case of Chichewa." In Kager, Renè, van der Hulst, Harry, and Zonneveld, Wim (eds.), *The Prosody-Morphology Interface*, 90-133. Cambridge: Cambridge University Press.

Inkelas, Sharon (1989) "Prosodic constituency in the lexicon," Ph.D dissertation, Stanford University.

Inkelas, Sharon (1993) "Deriving cyclicity." In Sharon Hargus and Ellen M. Kaisse (eds.), *Studies in Lexical Phonology. Phonetics and Phonology 4*: 75-100. San Diego: Academic Press.

Inkelas, Sharon, and Zoll, Cheryl (2005) *Reduplication: Doubling in Phonology*. Cambridge: Cambridge University Press.

Itô, Junko (1989) "A prosodic theory of epenthesis." *Natural Language and Linguistic Theory* 7: 217-59.

Ito, J. and Mester, A. (1992) Weak layering and word binarity. Ms. University of California, Santa Cruz. Linguistics Research Centre Report No. 92-09.

Jaccard, J. and J. Jacoby (2010) *Theory Construction and Model-Building Skills: A Practical Guide for Social Scientists*. New York, London: The Guilford Press.

Jacobs, Haike (1990) "On markedness and bounded stress systems." *The Linguistic Review* 7: 81-119.

Jacobs, Haike (2000) "The revenge of the uneven trochee: Latin main stress, metrical constituency, stress related phenomena and OT." In Lahiri, Aditi (ed.) *Analogy, Leveling, Markedness*, 333-352. Berlin, New York: Mouton de Gruyter.

Jacobson, Roman (1931) "Die Betonung und ihre Rolle in der Wort- und Syntagma-Phonologie." Travaux du cercle linguistique de Prague 4. Reprinted in R. Jacobson (1962) *Seletcted Writings I: Phonological Studies*, 117-136. The Hague: Mouton.

Jeanne, La Verne (1978) Aspects of Hopi Grammar. Ph.D dissertation, Linguistics Department, MIT.

Jeanne, La Verne (1982) "Some phonological rules of Hopi." *International Journal of American Linguistics* 48: 245-270.

Kager, R. (1995) "On foot templates and root templates." In Marcel den Dikken and Kees Hengeveld (eds.), *Linguistics in the Netherlands 1995*: 125-138. Amsterdam: John Benjamins.

Kager, R. (1999) *Optimality Theory*. Cambridge: Cambridge University Press.

Kager, R. (2007) "Feet and metrical stress." In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*, 195-227. Cambridge: Cambridge University Press.

Kanerva, J. (1990) *Focus and Phrasing in Chichewa Phonology*. New York: Garland.

Katamba, F. (1989) *An Introduction to Phonology*. Essex: Longman Group UK.

Kenstowicz, Michael (1997) "Quality-sensitive stress." *Revista di Linguistica* 9.1: 157-188. [also ROA 33]. Also Kenstowicz, Michael (2004) "Quality-sensitive stress [abridged]". In John McCarthy (ed.) *Optimality Theory in Phonology: A Reader*, 191-201 Oxford: Blackwells.

Kisseberth, Charles (1984) "Digo tonology." In G.N Clements and J. Goldsmith (eds.),

Autosegmental Studies in Bantu Tone, 105-182. Dordrecht: Foris.

Kisseberth, Charles W. and David Odden (2003) "Tone." In Derek Nurse & Gérard Philippson (eds.), *The Bantu Languages*, 59-70. London: Routledge.

Kraal, Peter (2005) A grammar of Makonde. Ph.D dissertation, University of Leiden.

Liberman, M. (1975) The intonational system of English. Ph.D dissertation, MIT.

Liberman, M. and Prince, A. (1977) "On stress and linguistic rhythm," *Linguistic Inquiry* 8: 249-336.

MacAlpine, A.G. (1905) "Tonga beliefs and customs." *Aurora* 8, Livingstonia Mission.

Mazgu Ghaku Chiuta, The Bible in Chitonga (Malawi) (c) (1986), United Bible Societies, UBS-EPF 1987-5M-D53.

McCarthy, John J. (2000) "The prosody of phase in Rotuman." *Natural Language and Linguistic Theory* 18: 147-197.

McCarthy, John, J. (2008) *Doing Optimality Theory: Applying Theory to Data*. Malden, MA., Oxford, Victoria: Blackwell Publishing.

McCarthy, J. & Alan Prince (1986) *Prosodic Morphology.* Rutgers Technical Report TR-32. New Brunswick, Rutgers University Center for Cognitive Science.

McCarthy, J. and Prince, A. (1993a) "Generalized Alignment." In G.E. Booi and J. van Marle (eds.), *Yearbook of Morphology 1993*, 79-153. Dordrecht: Kluwer.

McCarthy, J. and A. Prince (1993b) *Prosodic morphology I: constraint interaction and satisfaction.* Report no. RuCCS-TR-3. New Brunswick, NJ: Rutgers University Centre for Cognitive Science. (Rutgers Optimality Archive #482-1201.)

McCarthy, J. & A. Prince (1994a) "The emergence of the unmarked: Optimality in prosodic morphology." *Proceedings of the North East Linguistic Society* 24: 333-79.

McCarthy, J. & A. Prince (1994b) "Two lectures on Prosodic Morphology." Presented at the University of Utrecht, 22 and 24 June 1994 (ROA NO. 59-0000).

McCarthy, J. & A. Prince (1995a) "Faithfulness and reduplicative identity." In J. Beckman, L. Walsh-Dickey, and S. Urbanczyk (eds.). *UMOP 18*: 249-384. Amherst: GLSA.

McCarthy, J. & A. Prince (1995b) "Prosodic Morphology." In John A. Goldsmith (ed.), *The Cambridge Handbook of Phonological Theory*, 318-366. Cambridge, Mass.: Blackwell.

McCarthy, J. & A. Prince (1998) "Prosodic Morphology." In A. Spencer and A.M. Zwick (eds.), *The Handbook of Morphology*, 283-305. Oxford: Blacwell Publishers Ltd.

McCarthy, J. & A. Prince (1999) "Faithfulness and Identity in Prosodic Morphology." In Kager, René, van der Hulst, Harry, and Zonneveld, Wim (eds.), *The Prosody-Morphology Interface*, 218-309. Cambridge: Cambridge University Press.

Mcapu wa Chitonga (n.a) (1932).

Mchombo, Sam A. (1993) "Reflexive and reciprocal in Chichewa." In Sam A. Mchombo (ed.), *Theoretical Aspects of Bantu Grammar*. Stanford, Carlif.: CSLI: 181-207.

Mellander, Evan (2001) "Quantitative processes in trochaic systems." In K. Megerdoomian & L.A. Bar-el (eds.) *West Coast Conference on Formal Linguistics 20*: 414-427 Somerville, MA: Cascadilla Press.

Mellander, Evan (2004) "(HL)-creating processes in a theory of foot structure." *Linguistic Review 20.2*: 243-280.

Mkochi, W. (2005) An Optimality-theoretic account of some aspects of Chitonga Phonology. MA Thesis, University of Malawi, Chancellor College, Zomba, Malawi.

Mkochi, W. (2007/8) Liquid/glide/zero alternation in Chitonga: an optimality analysis. *Journal of Humanities* 21: 72-84.

Mkochi, W. (2009) Bimoraic word minimality condition in Chitonga: *OTanalysis. SKY Journal of Linguistics* 22:277-285.

Mphande, David K. (1998) On the use of Tonga myths, folktales and proverbs in moral education. Ph.D. dissertation, University of Malawi.

Mphande, David K (2000) *Nthanthi za Chitonga Zakusambizgiya ndi Kutauliya*. Blantyre: CLAIM/MABUKU.

Msosa, W. (1999) "Fishery culture and origins of the Tonga people of Lake Malawi." In H. Kawanabe, G.W. Couter and A.C. Roosevelt (eds.) *Ancient Lakes: Their Cultural and Biological Diversity*, 271-280. Belgium: Kenobi Publications.

Mtenje, Al (1986) Issues in the non-linear phonology of Chichewa. Ph.D dissertation, University College, London.

Mtenje, Al (1993) "Verb structure and Tone in Chiyao." In Mufwene, S. and L. Moshi (eds.) *Topics in African Linguistics*, 179-192. Amsterdam: John Benjamins Publishing Company.

Mtenje, Al (1994/95) "Tone in Malawian Tonga." *Journal of Humanities*, 8/9: 65-72.

Mtenje, A. (2006) "Alignment theory and prosody in Malawian ciTonga." *Linguistic Analysis* 32:327-365.

Mutaka, Ngessimo M. (1994) *The Lexical Tonology of Kinande*. Munich: LINCOM EUROPA.

Myers, Scott (1987) Tone and the structure of words in Shona. Ph.D dissertation, University of Massachusetts, Amherst.

Myers, S. and T. Carleton (1996) "Tonal transfer in Chichewa." *Phonology* 13: 39-72.

Nespor, M. and I. Vogel (1986) *Prosodic Phonology*. Dordrecht: Foris.

Odden, David (1984) "Stem tone assignment in Shona." In G.N Clements and J. Goldsmith (eds.), *Autosegmental Studies in Bantu Tone* (pp. 255-280). Dordrecht: Foris.

Odden, David (1988) "AntiAntigemination and the OCP." *Linguistic Inquiry* 19: 451-475.

Odden, David (1995) "Tone: African languages." In John A. Goldsmith (ed.), *The Handbook of Phonological Theory*, (pp. 444-475). Cambridge, Massachusetts: Blackwell.

Oostendorp, M. van (2004) "Crossing morpheme boundaries in Dutch." *Lingua* 114: 1367-1400.

Orie, O.O. (1997) *Benue-Congo Prosodic Phonology and Morphology in Optimality Theory*. Munich: LINCOM EUROPA.

Philippson, Gerárd (1998) "Tone reduction vs. metrical attraction in the evolution of Eastern Bantu tone systems." In Larry M. Hyman & Charles W. Kisseberth (eds.), *Theoretical Aspects of Bantu Tone*, (pp. 315.329). Stanford, Calif.: CSLI.

Prince, Alan (1985) "Improving tree theory." *Proceedings of BLS* 11: 471-90.

Prince, A.S. and P. Smolensky (1991) "Optimality." Talk presented at Arizona Phonology Conference 3, University of Arizona, Tucson.

Prince, A. and P. Smolensky (1993) *Optimality Theory: Constraint Interaction in Generative Grammar*, RuCCs Technical Report NO. 2, Rutgers University Centre for Cognitive Science, Piscataway, N.J.

Prince, A. and P. Smolensky (2004) *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, MA: Blackwell Publishing.

Rice, Curtis (1992) Binarity and ternarity in metrical theory: Parametric extensions. Doctoral dissertation, University of Texas, Austin.

Rose, Sarah, Christa, Beaudoin-Lietz and Derek Nurse (2002) *A Glossary of Terms from Bantu verbal categories, with special emphasis on tense-aspect*. Munich: LINCOM EUROPA.

Russel, Kevin (1997) "Optimality theory and morphology." In Diana Archangeli and D. Terence Langendoen (eds.), *Optimality Theory: An Overview*, 102-133. Oxford: Blackwell.

Saeed, John I. (2003) *Semantics*. Oxford, Malden, MA: Blackwell Publishing.

Selkirk, E. (1978/81) "On prosodic structure and its relation to syntactic structure." In T. Fretheim (ed.) *Nordic Prosody II*, 111-40. Trondheim: Tapir.

Selkirk, E. (1984) *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: The MIT Press.

Selkirk, E. (1994) Optimality Theory and featural phenomena. Lecture notes, Linguistics 730, University of Massachusetts, Amherst.

Selkirk, E. (1995) "The prosodic structure of function words." In Beckman, J., L.W. Dickey, and Suzanne Urbanczyk (eds.) *Papers in Optimality Theory*. *University of Massachusetts Occasional Papers in Linguistics* 18:439-69. Amherst, Mass.: Graduate Linguistic Student Association.

Smith, Jennifer L. (2000) "Prominence, augmentation, and neutralization in Phonology." *Proceedings of Berkeley Linguistics Society* 26: 247-257.

Smith, Jennifer L. (2001) "Lexical category and phonological contrast." In R. Kirchner, J. Pater, and W. Wikely (eds.). *PETL 6: Proceedings of the Workshop on the Lexicon in Phonetics and Phonology*, 61-72. Edmonton: University of Alberta.

Soko, B.J. (1985) Tonga proverbs. Unpublished collection.

Steriade, Donca (1988) "Reduplication and syllable transfer in Sanskrit and elsewhere." *Phonology* 5: 73-155.

Steriade, Donca (1994) Positional neutralization and the expression of contrast. MS, UCLA.

Tew, Mary (1950) *The Peoples of the Lake Nyasa Region*. Passim.

Trubetskoy, N. (1939 [1969]) *Principles of Phonology*. C. Baltaxe, trans. Berkeley/Los Angeles: University of California Press.

Turner, Y. (1952) *Tumbuka-Tonga-English Dictionary*. Blantyre: Hetherwick Press.

Urbanczyc, Suzanne (1996) Patterns of reduplication in Lushootsheed. Ph.D dissertation, University of Massachusetts-Amherst.

Urbanczyc, Suzanne (2000/2006) "Reduplicative Form and the Root-Affix asymmetry." *Natural Language and Linguistic Theory* (2006) 24: 179-240.

Ussishkin, Adam P. (2000) The emergence of fixed prosody. PhD dissertation, University of California-Santa Cruz.

Vail, Hazen Leroy (1972) Aspects of the Tumbuka Verb. PhD. Thesis, University of
Wisconsin.

Van der Hulst, Harry (2011) "Pitch Accent Systems." In Marc van Oostendorp, Colin J.
Ewen, Elizabeth V. Hume &Keren Rice (eds.), *The Blackwell Companion to
Phonology*, 1003-1026. Oxford: Blackwell-Wiley.

Van der Hulst, Harry & Marian Klamer (1996) "Major class and manner features." In P.M.
Bertinetto, Michael Kenstowicz & M. Loporcaro (eds.) *Certamen Phonologicum II:
Papers from the 1990 Cortona Phonology Meeting*, 19-41. Torino, Rosenburg and
Selier.

Velsen, Jaap van (1959a) "The missionary factor among the lakeside Tonga of Nyasaland,"
Rhodes-Livingstone Journal, XXVI: 1-22.

Velsen, Jaap van (1959b) "Notes on the history of the lakeside Tonga of Nyasaland," *African
Studies*, XVIII, No. 3: 105-117.

Velsen, Jaap van (1964) *The Politics of Kinship: A study in Social Manipulation among the
Lakeside Tonga of Nyasaland*. Manchester: Manchester University Press.

Walker, Rachel (2000) "Nasal reduplication in Mbe affixation." *Phonology* 17: 65-115.

Wilson, Monica (1958) *The Peoples of the Tanganyika-Nyasa Corridor*. Cape Town. Passim.

Yip, Moira (2002) *Tone.* Cambridge: Cambridge University Press.

Yoneda, Nobuko (2005) "Tone patterns of Matengo nouns." In Shigeki Kaji (ed.),
*Crosslinguistic studies of tonal phenomena: Historical Development, Tone-Syntax
Interface, and Descriptive Studies*, 393-409. Tokyo: ILACC.

Zec, Draga (2007) "The Syllable." In Paul de Lacy (ed.), *The Cambridge Handbook of
Phonology*, 162-194. Cambridge: Cambridge University Press.