A study of vowel-pair frequencies in Lozi: Implications for formal analyses

Stephen Nichols
University of Manchester

MANCHESTER FORUM IN LINGUISTICS

3 May 2019

Introduction

- A case study of vowel-pair frequencies in Lozi.
  - 🖫 Using data from Jalla (1982), a large digitised dictionary.
- The results have implications for any formal analysis of the language.
  - ☞ The only synchronic **phonotactic** vowel co-occurrence restriction in the language is a **prohibition** of the vowel pair **/o.u/** that is **blind to part of speech**.

Outline

1. Crash course in Bantu height harmony
2. Brief introduction to Lozi
3. Methodology
4. Results
5. Discussion: Synchrony
6. Discussion: Diachrony
7. Summary

Bantu height harmony

- Height harmony is **extremely common** in the Bantu languages and, in the vast majority of cases, is described as being **confined to verbs**.
  - See e.g. Clements 1991, Hyman 1999:§2, 2003, Odden 2015:§1.
- By far the commonest variety is the “canonical” asymmetric pattern.
  - This is found in, for example:
    Chichewa (N.31), Kinyarwanda (D.61), Luganda (E.15), Shona (S.11), Swahili (G.42).
- This has been the focus of almost all work on height harmony in Bantu.
  - Katamba (1984), Mtenje (1985), Moto (1989), Hyman (1991), Scullen (1992), Harris (1994, 1997), Beckman (1997), Downing (2010), Downing & Mtenje (2017).

Canonical Bantu height harmony

- Canonical height harmony is asymmetric w.r.t. rounding (and/or backness).
  - /i/ is lowered after both /e o/ whereas /u/ is lowered only after /o/.
  - This is both common currently and robust historically (Hyman 1999:238,245).
- Can, descriptively at least, be split into **front** and **back** height harmony.

Canonical five-vowel Bantu height harmony

(1) a. Front height harmony: **i** → **e** / {**e o**} (C) _

b. Back height harmony: **u** → **o** / **o** (C) _

- This is exemplified in the slides that follow with data from Bemba (M.42).

Front height harmony in Bemba

(2) **Unsuffixed:**

a. -bila
‘to sew’

b. -tunga
‘to thread’

c. -lemba
‘to write’

d. -longa
‘to pack’

e. -kaka
‘to tie’

(3) **Applicative suffix:**

a. -bilila
‘to sew for’

b. -tungila
‘to thread for’

c. -l**e**mb**e**la
‘to write to’

d. -l**o**ng**e**la
‘to pack for’

e. -kakila
‘to tie for’

(Hoch 1998: *sub vocibus*; own fieldwork)

Back height harmony in Bemba

(4) **Unsuffixed:**

a. -bila
‘to sew’

b. -tunga
‘to thread’

c. -lemba
‘to write’

d. -longa
‘to pack’

e. -kaka
‘to tie’

(5) **Reversive suffix:**

a. -bilulula
‘to unsew’

b. -tungulula
‘to unthread’

c. -lembulula
‘to rewrite/erase’

d. -l**o**ng**o**l**o**la
‘to unpack’

e. -kakulula
‘to untie’

(Hoch 1998: *sub vocibus*; own fieldwork)

Introducing Lozi

- Lozi is a Bantu language spoken mainly in Zambia.
  - Nearly 750,000 speakers (Eberhard et al. 2019).
  - Guthrie code: K.21 (Maho 2009).
  - But genetically Zone S (Gowlett 2003).
- It is descended from Sotho (S.33).
  - But has also been heavily influenced by Luyana (K.31).
  - As well as, to a lesser extent, Tswana (S.31).
    - 🕮 See Gowlett (1989) for a discussion of Lozi's history.

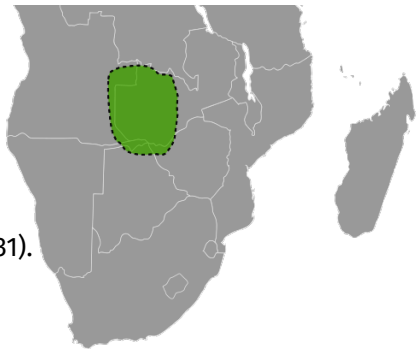


Image adapted from https://commons.wikimedia.org/wiki/File:Barotseland.svg

No front height harmony in Lozi

(6) **Unsuffixed:**

a. -kiya
‘to lock’

b. -luka
‘to weave’

c. -leka
‘to buy’

d. -longa
‘to pack’

e. -tama
‘to fold’

(7) **Causative suffix:**

a. -kiyisa
‘to make lock’

b. -lukisa
‘to make weave’

c. -lekisa
‘to sell’

d. -longisa
‘to make pack’

e. -tamisa
‘to make fold’

(8) **Applicative suffix:**

a. -kiyela
‘to unlock for’

b. -lukela
‘to weave for’

c. -lekela
‘to buy for’

d. -longela
‘to pack for’

e. -tamela
‘to fold for’

(Jalla 1982: *sub vocibus*; own fieldwork)

Back height harmony in Lozi

(9) **Unsuffixed:**

a. -kiya
‘to lock’

b. -luka
‘to weave’

c. -leka
‘to buy’

d. -longa
‘to pack’

e. -tama
‘to fold’

(10) **Reversive suffix:**

a. -kiyulula
‘to unlock’

b. -lukulula
‘to unweave’

c. -lekulula
‘to resell’

d. -l**o**ng**o**l**o**la
‘to unpack’

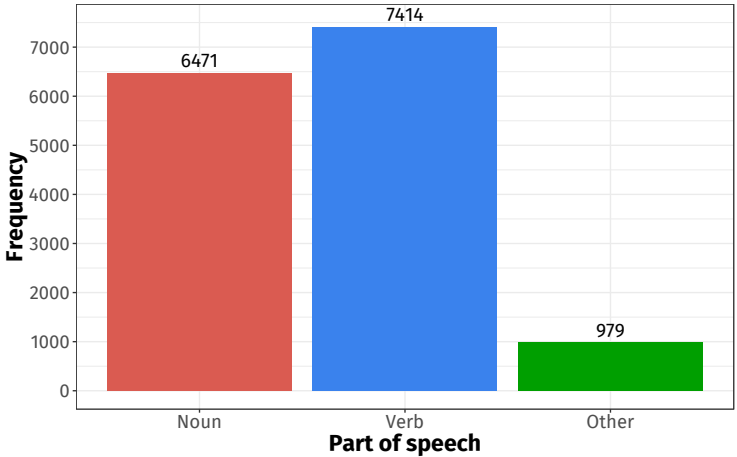
e. -tamulula
‘to unfold’

(Jalla 1982: *sub vocibus*; own fieldwork)

Methodology I

- The data come from Jalla (1982).
  - A print Lozi–English dictionary converted into text file.
  - Available on the Comparative Bantu Online Dictionary.
    - http://www.cbold.ish-lyon.cnrs.fr/
- After corrections and processing, there were a total of 24,238 entries.
  - Each individual entry was (already) tagged for part of speech.
- Perfective verb forms were then removed, giving a final total of **14,863**.
  - *-tamile* ← *-tama* 'to tie';
  - *-lekezi* ← *-lekela* 'to buy for';
  - *-lutuluzi* ← *-lutulula* 'to unthatch';
  - *-mizize* ← *-miza* 'to swallow'.

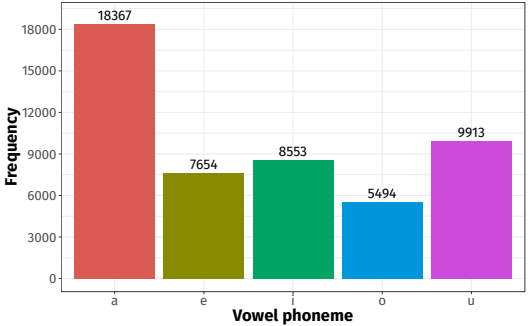
Methodology II

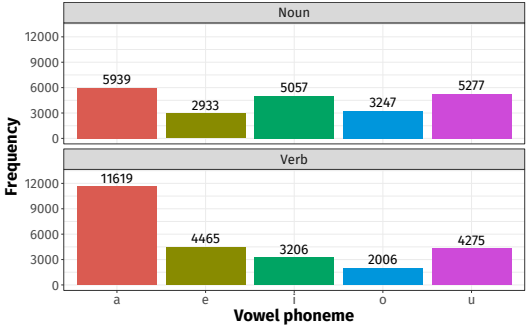


Methodology III

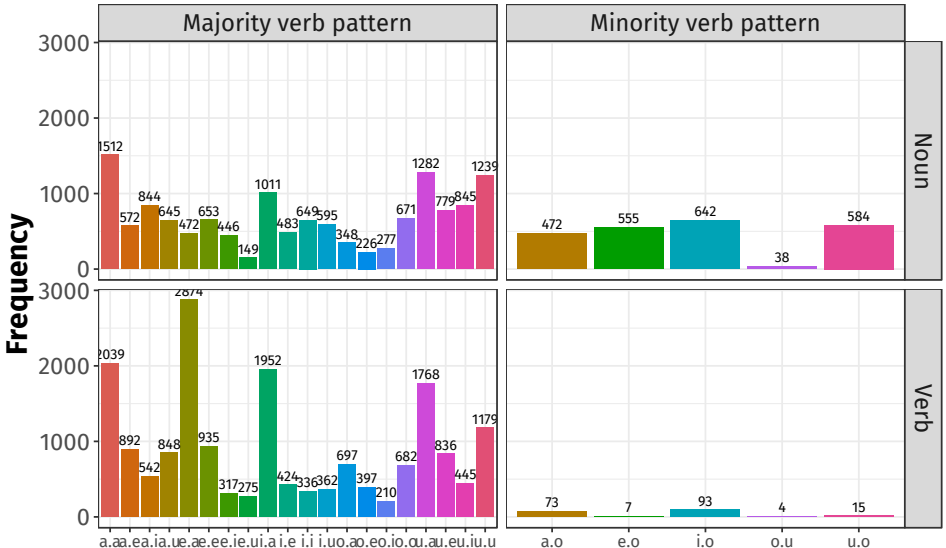
- Orthographic long vowels were reduced to short vowels and intervening consonants were skipped over.
  - This means that VCV sequences were treated the same as VV sequences.
- Both the **observed** and **expected frequencies** of all 25 possible vowel pairs were calculated.
- As were the corresponding **observed–expected ratios**.

Results I

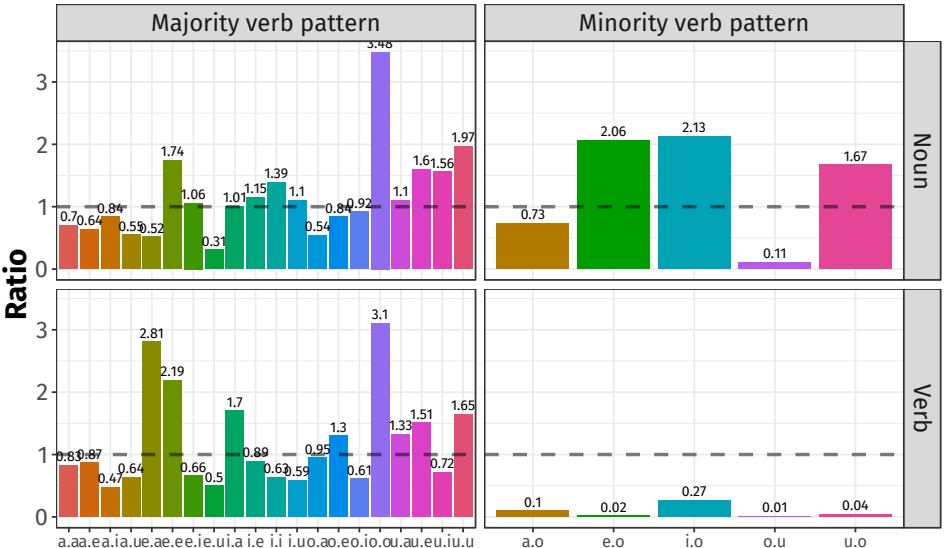

18367
7654
8553
5494
9913
18000
15000
12000
9000
6000
3000
0
Frequency
a
e
i
o
u
Vowel phoneme


Noun
5939
2933
5057
3247
5277
Verb
11619
4465
3206
2006
4275
12000
9000
6000
3000
0
Frequency
a
e
i
o
u
Vowel phoneme

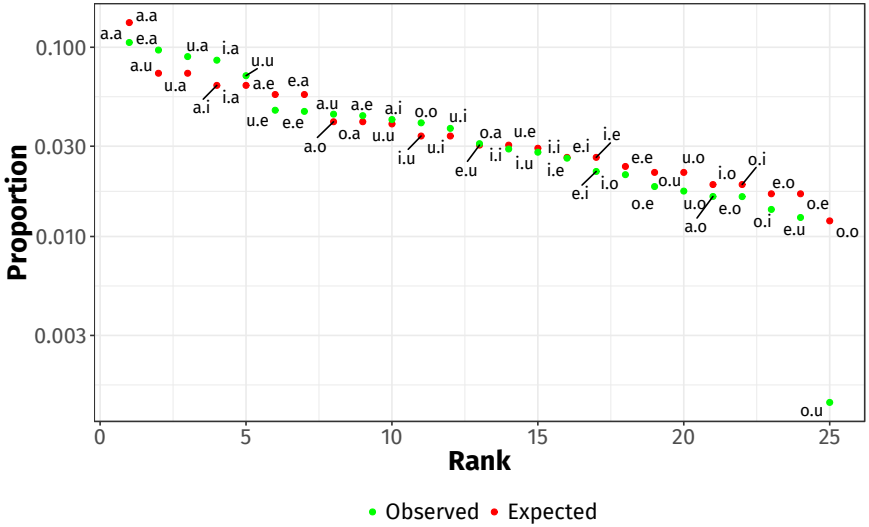
Results II


Majority verb pattern
Minority verb pattern
Noun
Verb
Frequency
3000
2000
1000
0
1512
572
844
645
472
653
446
149
1011
483
649
595
348
226
277
671
1282
779
845
1239
472
555
642
38
584
2039
892
542
848
2874
935
317
275
1952
424
336
362
697
397
210
682
1768
836
445
1179
73
7
93
4
15
a.a a.e a.i a.u a.æ e.e e.i e.u i.a i.e i.i i.u o.a o.e o.i o.u u.a u.e u.i u.u
a.o
e.o
i.o
o.u
u.o

Results III


Majority verb pattern
Minority verb pattern
Noun
Verb
Ratio
0
1
2
3
0.7
0.64
0.84
0.55
0.52
1.74
1.06
0.31
1.01
1.15
1.39
1.1
0.54
0.88
0.92
3.48
1.1
1.6
1.56
1.97
0.73
2.06
2.13
0.11
1.67
0.83
0.87
0.47
0.64
2.81
2.19
0.66
0.5
1.7
0.89
0.63
0.59
0.95
1.3
0.61
3.1
1.33
1.51
0.72
1.65
0.1
0.02
0.27
0.01
0.04
a.a
a.e
a.i
a.u
e.a
e.e
e.i
e.u
i.a
i.e
i.i
i.u
o.a
o.e
o.i
o.o
u.a
u.e
u.i
u.u
a.o
e.o
i.o
o.u
u.o

Results IV


Proportion
0.100
0.030
0.010
0.003
0
5
10
15
20
25
Rank
a.a
a.a
e.a
a.u
u.a
u.a
i.a
a.i
i.a
u.u
a.e
e.a
u.e
e.e
a.u
a.e
a.o
o.a
u.u
a.i
i.u
o.o
u.i
u.i
o.a
u.e
e.u
i.i
i.u
i.i
i.e
e.i
i.e
i.o
e.e
e.i
o.e
u.o
o.u
i.o
o.i
u.o
a.o
e.o
e.o
o.i
o.e
e.u
o.o
o.u
Observed
Expected

Discussion: Synchrony I

- There is a **near-total absence** of /o.u/ in both **verbs** and **nouns**:
  - Suggests an active **phonotactic** vowel co-occurrence **restriction** against **/o.u/**;
  - This applies regardless of part of speech.
- In this case, the reversive suffix would underlyingly be /-**u**l**u**l-/.
  - Undergoes a phonotactically-governed change to [-**o**l**o**l-] following /**o**/;
  - But elsewhere surface faithfully as [-**u**l**u**l-].

Discussion: Synchrony II

- Exceptions in the data set:
  - 4 instances occur across boundaries in compounds or with reduplication (discussed around (16) later on).
  - 4 are loan words from English:

    (11) a. bish**o**p**u** ‘bishop’
    b. sit**o**f**u** ‘stove’
    c. ingil**o**p**u** ‘envelope’
    d. w**o**l**u**pulete ‘wall plate’

  - 9 are identified as loans from Luyana, e.g.:

    (12) a. nj**o**p**u** ‘damp, dewy place’
    b. mal**o**p**u** ‘beer’
    c. nd**o**p**u** ‘elephant’
    d. s**o**p**u** ‘fine grass growing in fertile soil’

Discussion: Synchrony III

- In addition, the intervening segments are not random.
  - 63% have an intervening labial (see also (11) and (12)):

    (13) a. bub**o**f**u** ‘blindness’
    b. siy**o**p**u** ‘hut used for ritual confinements’

  - 16% have an intervening lateral, e.g.:

    (14) a. lub**o**l**u** ‘double chin’
    b. muh**o**l**u** ‘stomach, tripe’

- 60% occur word-finally.
  - In an ongoing production study, I find that, for some speakers, final /u/ is deleted or devoiced word-finally and, when retained, often appears to be phonetically lower after /o/ than elsewhere.

Discussion: Synchrony IV

- Nevertheless, in the majority of cases where /o.u/ might occur (e.g. as epenthesis in loan words), /o.o/ is found instead, e.g.:

(15)	a.	lubot**o**l**o**	‘bottle’
	b.	kop**o**l**o**	‘corporal’
	c.	mabasik**o**l**o**	‘bicycle’
	d.	-p**o**l**o**fita	‘to prophesy’
	e.	sin**o**d**o**	‘synod’
	f.	sitol**o**p**o**	‘strap (for inspanning oxen)’
	g.	c**o**k**o**	‘chalk’
	h.	d**o**k**o**ta	‘doctor’
	i.	n**o**t**o**	‘musical note’

Discussion: Synchrony V

- Places where the /o.u/ restriction seems not to apply:

(16) Across boundaries in compounds and with reduplication:

a.	kutwel**o**-b**u**tuku	'pity, compassion'	(Jalla 1982: *sub voce*)
b.	maful**o**-f**u**lo	'eagerness, zeal'	(Jalla 1982: *sub voce*)
c.	muly**o**l**u**mbo	'senior person'	(Mwisiya 1977:7)
	↬ muly**a** **u** l**u**mbo		

(17) Between prefix and root or between two prefixes:

a.	ne-ni-ta-t**o**-k**u**ta	'I was going to have my hair cut'	(Gowlett 1967:249)
b.	aba-t**o**-l**u**-tusa	'they are not coming to help us'	(Gowlett 1967:272)
c.	B**o**-M**u**wae	'Honourable Princess'	(Fortune 2001:12)
d.	k**o**-k**u**-mezi	'at a wet place'	(Fortune 2001:33)

Discussion: Synchrony VI

- An explanation though may be found with reference to prosodic structure.
- It is not uncommon for compounds to be different prosodic words.
- Similarly, base and reduplicant are often separate prosodic words.
- But what about prefixes?

Discussion: Synchrony VII

- Certain speakers may prefer to write prefixes as **separate individual words**, especially with verbs (own fieldwork).
  - 🕮 Gorman (1950), a pedagogical resource for English speakers, does likewise.
- Similarly, in speaking, speakers sometimes insert a **pause** between prefixes and the **rest of the word**; again, most often with verbs (own fieldwork).
- It is often argued that, in the Bantu verb word, at least, the **root** and **derivational suffixes** are part of the **derivational stem**, a lower level than prefixes (Downing 1998a,b, 1999).

Discussion: Diachrony I

- Modern Sotho does not have the same harmony system as Lozi:
  - It has **regressive** height harmony (Parkinson 1996:§3.1.6, Gowlett 2003:§3.3).
- Sotho also has a larger vowel system than five-vowel Lozi.
- However, back height harmony in Lozi **does not emerge** as a simple consequence of the historical mergers:
  - Sotho /bɔf-ʊlʊl-a/ [b**o̲**fʊlʊla] → Lozi [b**o̲**f**o̲**l**o̲**la] ‘to outspan’, not *[bofulula]
  - But: Sotho [-tsʊma] → Lozi [-zuma] ‘to hunt’, not *[-zoma]

Discussion: Diachrony II

- Gowlett (1989:141) suggests that Luyana has **back** height harmony.
  - At least as far as the reversive suffix is concerned.
- Givón (1970), a sketch grammar of Luyana:
  - Makes no explicit mention of height harmony.
  - Sporadic examples show there is **no front** height harmony.
  - But there is no evidence either way regarding back height harmony.

Discussion: Diachrony III

- Jacottet (1896:84f) explicitly mentions a **lack** of **front** height harmony in verbs with the applicative suffix.

L'*harmonie vocalique* [...] ne se fait pas sentir en Louyi [Louyana].
Le suffixe directif [applicatif] est toujours *-ela* ou *-ena* (jamais *-ila* ou *-ina*)

Vowel harmony [...] is not found in Luyi [Luyana].
The directive [applicative] suffix is always *-ela* or *-ena* (never *-ila* or *-ina*)

Discussion: Diachrony IV

- Jacottet (1896) says nothing about **back** height harmony.
- But Jacottet (1901:220ff) contains examples **suggestive** of **back** height harmony in verbs with alternations in derivational suffixes:
  - *l**o**nga* ‘être plein’ — *l**o**ng**o**la* ‘sortir (de la nourriture d’un pot)’
  - *p**u**mba* ‘mettre de la terre’ — *p**u**mb**u**la* ‘déterrer’
- It is not clear whether back height harmony is found in nouns.
  - ☞ Though based on the Luyana loans in Lozi in (12), it is possible it is not.

Summary

- I have presented you with vowel-pair frequency data from Lozi.
- And argued that these suggest an active ban only on the vowel pair /o.u/ and that this is blind to part of speech.
- Something that any formal account of height harmony in Lozi must reflect.

Acknowledgements

Thanks to members of the **University of Manchester Phonetics Lab** for helpful discussion, **Ricardo Bermúdez-Otero** and **Wendell Kimper** for supervision and audiences at **Fonologi i Norden** and the **PhilSoc ECR Forum** for feedback.

Contact

💻 https://tiny.cc/sjn

✉ stephen.nichols@manchester.ac.uk

References I

Beckman, Jill N. 1997. Positional faithfulness, positional neutralisation and Shona vowel harmony. *Phonology* 14(1). 1–46.

Clements, George N. 1991. Vowel Height Assimilation in Bantu Languages. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society: Special Session on African Language Structures* 17. 25–64.

Downing, Laura J. 1998a. On the prosodic misalignment of onsetless syllables. *Natural Language & Linguistic Theory Linguistic Theory* 16. 1–52.

Downing, Laura J. 1998b. Prosodic misalignment and reduplication. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1997*, 83–120. Dordrecht: Kluwer.

Downing, Laura J. 1999. Prosodic Stem ≠ Prosodic Word in Bantu. In T. Alan Hall & Ursula Kleinhenz (eds.), *Studies on the phonological word*, 73–98. Amsterdam: John Benjamins.

Downing, Laura J. 2010. Opacity is a matter of representation: Shimakonde vowel harmony and vowel reduction. *ZAS Papers in Linguistics* 52. 159–91.

Downing, Laura J. & Al D. Mtenje. 2017. *The Phonology of Chichewa*. New York: Oxford University Press.

Eberhard, David M., Gary F. Simons & Charles D. Fennig. 2019. Lozi. In *Ethnologue: Languages of the World*, Dallas, TX: SIL International. 22nd edition. URL: http://www.ethnologue.com/language/loz.

Fortune, George. 2001. *An Outline of siLozi Grammar*. Lusaka: Bookworld Publishers. First published in 1977 by the Institute for African Studies University of Zambia in *Language in Zambia; Grammatical Sketches*.

Givón, Talmy. 1970. *The Si-Luyana Language: A Preliminary Linguistic Description*. Lusaka: University of Zambia, Institute for Social Research.

Gorman, W. A. R. 1950. *Simple Silozi: A Guide for Beginners*. London: Longmans, Green & Co.

Gowlett, Derek. 2003. Zone S. In Derek Nurse & Gérard Philippson (eds.), *The Bantu Languages*, 609–38. London: Routledge.

Gowlett, Derek F. 1967. Morphology of the Verb in Lozi. MA dissertation, University of the Witwatersrand.

Gowlett, Derek F. 1989. The Parentage and Development of Lozi. *Journal of African Languages and Linguistics* 11(2). 127–50.

Harris, John. 1994. Monovalency and opacity: Chichewa height harmony. *UCL Working Papers in Linguistics* 6. 509–47.

Harris, John. 1997. Licensing inheritance: An integrated theory of neutralization. *Phonology* 14. 315–70.

Hoch, Rev. E. 1998. *Hippocrene Concise Dictionary: Bemba-English/English-Bemba*. New York: Hippocrene Books. First published in 1960.

References II

Hyman, Larry M. 1991. Cyclic phonology and morphology in Cibemba. University of California, Berkeley, ms.

Hyman, Larry M. 1999. The Historical Interpretation of Vowel Harmony in Bantu. In Jean-Marie Hombert & Larry M. Hyman (eds.), *Bantu Historical Linguistics: Theoretical and Empirical Perspectives*, 235–95. Stanford, CA: CSLI Publications.

Hyman, Larry M. 2003. Segmental phonology. In Derek Nurse & Gérard Philippson (eds.), *The Bantu Languages*, 42–58. London: Routledge.

Jacottet, Édouard. 1896. *Études sur les langues du Haut-Zambèze. Textes originaux, recueillis et traduits en français et précédés d'une esquisse grammaticale. Première partie : grammaires soubiya et louyi* (Bulletin de Correspondence Africaine XVI). Paris: L'École des Lettres d'Alger.

Jacottet, Édouard. 1901. *Études sur les langues du Haut-Zambèze. Textes originaux, recueillis et traduits en français et précédés d'une esquisse grammaticale. Troisième partie : textes louyi : contes et légendes, supersitions, etc. et vocabulaires* (Bulletin de Correspondence Africaine XVI). Paris: L'École des Lettres d'Alger.

Jalla, Adolphe. 1982. Database of 'Dictionary of the Lozi Language, Vol. 1: Lozi–English (3rd edition)'. Available to download from the Comparative Bantu Online Dictionary. Contributed by John Lowe. URL: http://www.cbold.ish-lyon.cnrs.fr/Dico.asp?Langue=Lozi.

Katamba, Francis. 1984. A nonlinear analysis of vowel harmony in Luganda. *Journal of Linguistics* 20(2). 257–75.

Maho, Jouni Filip. 2009. NUGL Online: The online version of the New Updated Guthrie List, a referential classification of the Bantu languages, ms. URL: http://goto.glocalnet.net/mahopapers/nuglonline.pdf.

Moto, Francis. 1989. Phonology of the Bantu lexicon. PhD thesis, University College London.

Mtenje, Al D. 1985. Arguments for an autosegmental analysis of Chichewa vowel harmony. *Lingua* 66(1). 21–52.

Mwisiya, M. W. 1977. *Introduction to Silozi Grammar*. Lusaka: National Educational Company of Zambia Ltd.

Odden, David. 2015. Bantu Phonology. In *Oxford Handbooks Online*, Oxford: Oxford University Press. URL: http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199935345.001.0001/oxfordhb-9780199935345-e-59.

Parkinson, Frederick. 1996. The representation of vowel height in phonology. PhD thesis, Ohio State University.

Scullen, Mary Ellen. 1992. Chichewa vowel harmony and Underspecification Theory. *Linguistic Analysis* 22. 218–45.