

# Asymmetries in vowel-pair frequencies and height harmony in Bantu

Stephen Nichols  
University of Manchester

50<sup>th</sup> Annual Conference on African Linguistics  
University of British Columbia

Vancouver • 23 May 2019



# Introduction: This study

- ▶ I examine vowel-pair frequencies in the **nouns** of six **five-vowel** Bantu languages and the relationship with **height harmony** in **verbs**.
- ▶ Two vowel pairs are of particular interest, namely **/e.i/** and **/o.u/**.
- ▶ The results suggest a bias against both vowel pairs but an relatively **stronger bias against /o.u/ than /e.i/** in these languages
- ▶ I then focus on **Lozi** and argue that, there, the only phonotactic vowel co-occurrence restriction is a **part-of-speech-blind ban on /o.u/**.



# Introduction: The sample

Language	Guthrie code	Harmony system	Data source
Chewa	M.31b	Canonical	Mtenje (2001)
Kalanga	S.16	Canonical	Mathangwane (1994)
Lozi	K.21	Back height harmony only	Jalla (1982)
Makhuwa	P.31	Back height harmony only	Kisseberth (1996)
Pende	L.11	"Quasi-canonical"	Gusimana (1972)
Yao	P.21	Canonical	Ngunga (2001)

**Table 1:** Six-language sample

# Introduction: Bantu height harmony

- ▶ Height harmony is **widespread** in Bantu.
  - 📄 See e.g. Clements (1991), Hyman (1999:§2, 2003), Odden (2015:§1).
- ▶ In almost all cases, it is described as being **confined to verbs**.
- ▶ By far the commonest variety is the **“canonical”** asymmetric pattern.
  - Found in, for example: Bemba (M.42), Luganda (E.15), Shona (S.11), Swahili (G.42).
- ▶ This has been the focus of the **majority** of 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 front height harmony in Bemba

(1) **Unsuffixed:**

- a. -bila  
    'to sew'
- b. -tunga  
    'to thread'
- c. -lemba  
    'to write'
- d. -longa  
    'to pack'
- e. -kaka  
    'to tie'

(2) **Applicative suffix:**

- a. -bilila  
    'to sew for'
- b. -tungila  
    'to thread for'
- c. -lembela  
    'to write to'
- d. -longela  
    'to pack for'
- e. -kakila  
    'to tie for'

(Hoch 1998: sub vocibus; own fieldwork)

# Canonical back height harmony in Bemba

## (3) **Unaffixed:**

- a. -bila  
    'to sew'
- b. -tunga  
    'to thread'
- c. -lemba  
    'to write'
- d. -longa  
    'to pack'
- e. -kaka  
    'to tie'

## (4) **Reversive suffix:**

- a. -bilulula  
    'to unsew'
- b. -tungulula  
    'to unthread'
- c. -lembulula  
    'to rewrite/erase'
- d. -longololo  
    'to unpack'
- e. -kakulula  
    'to untie'

(Hoch 1998: sub vocibus; own fieldwork)



# Front height harmony in Pende

(5) **Unaffixed:**

- a. -shita  
'fermer'
- b. -tunga  
'construire'
- c. -suega  
'cacher'
- d. -sola  
'défricher'
- e. -tala  
'examiner'

(6) **Applicative suffix:**

- a. -shitila  
'fermer pour'
- b. -tungila  
'construire pour'
- c. -suegela  
'cacher pour'
- d. -solela  
'défricher pour'
- e. -talela  
'surveiller pour'

(Gusimana 1972: sub vocibus)



# Back height harmony in Pende

(7) **Unaffixed:**

- a. -shita  
'fermer'
- b. -tunga  
'construire'
- c. -suega  
'cacher'
- d. -soga  
'se démancher'
- e. -tala  
'examiner'

(8) **Repetitive suffix:**

- a. -shitulula  
'fermer de nouveau'
- b. -tungulula  
'construire de nouveau'
- c. -suegulula  
'cacher encore de nouveau'
- d. -sogolola  
'démancher de nouveau'
- e. -talulula  
'regarder de nouveau'

(Gusimana 1972: sub vocibus)





# No front 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) **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'

(11) **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



(12) **Unaffixed:**

- a. -kiya  
    'to lock'
- b. -luka  
    'to weave'
- c. -leka  
    'to buy'
- d. -longa  
    'to pack'
- e. -tama  
    'to fold'

(13) **Reversive suffix:**

- a. -kiyulula  
    'to unlock'
- b. -lukulula  
    'to unweave'
- c. -lekulula  
    'to resell'
- d. -longolola  
    'to unpack'
- e. -tamulula  
    'to unfold'

(Jalla 1982: sub vocibus; own fieldwork)

- ▶ The data come from the Comparative Bantu Online Dictionary:
  -  <http://www.cbold.ish-lyon.cnrs.fr/>
    -  Gusimana (1972), Jalla (1982), Mathangwane (1994), Kisseberth (1996), Mtenje (2001), Ngunga (2001)
- ▶ Each file was hand-corrected for machine-readability and then processed in R.
  - Long and short vowels were collapsed and intervening consonants ignored.
  - **Observed** and **expected frequencies** of all 25 possible vowel pairs were calculated.
  - As were the corresponding **observed-expected ratios**.
- ▶ Each individual entry was (already) tagged for part of speech.



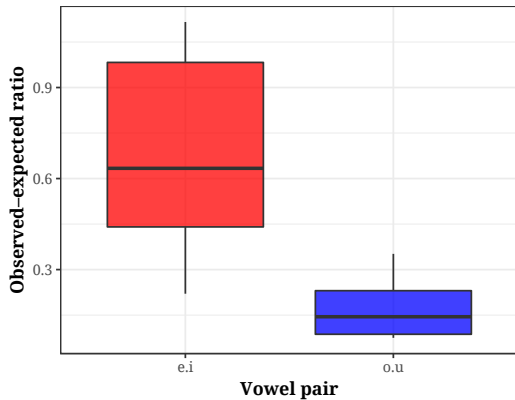
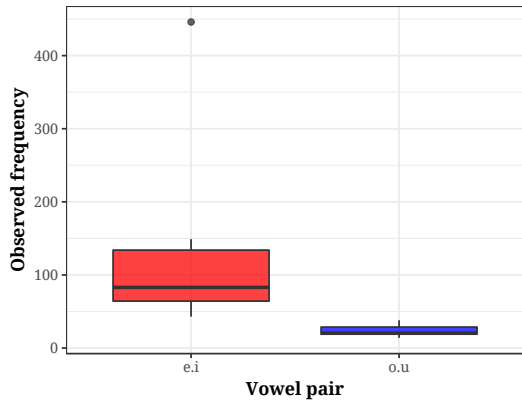
# Data set at large: Results I

	<i>/e.i/</i>			<i>/o.u/</i>		
	O. freq.	E. freq.	Ratio	O. freq.	E. freq.	Ratio
<b>Chewa</b>	60	146	0.41	22	89	0.25
<b>Kalanga</b>	77	69	1.12	19	54	0.35
<b>Yao</b>	43	195	0.22	14	176	0.08
<b>Pende</b>	149	202	0.74	19	253	0.08
<b>Lozi</b>	446	419	1.06	38	349	0.11
<b>Makhuwa</b>	89	168	0.53	31	172	0.18

**Table 2:** Observed/expected frequencies and ratios for */e.i/* and */o.u/*



# Data set at large: Results II





# Data set at large: Discussion I

- ▶ The under-representation of both /e.i/ and /o.u/ might be interpreted as a **gradient** version in nouns of the **categorical** rule in verbs.
  - Cf. the under-representation heteromorphemic geminates in English or compounds in Navajo disagreeing in anteriority Martin (see 2011).
- ▶ It might be argued also that this is simply predictable on **inductive** grounds.
  - Cf. Archangeli et al. (2012a,b).

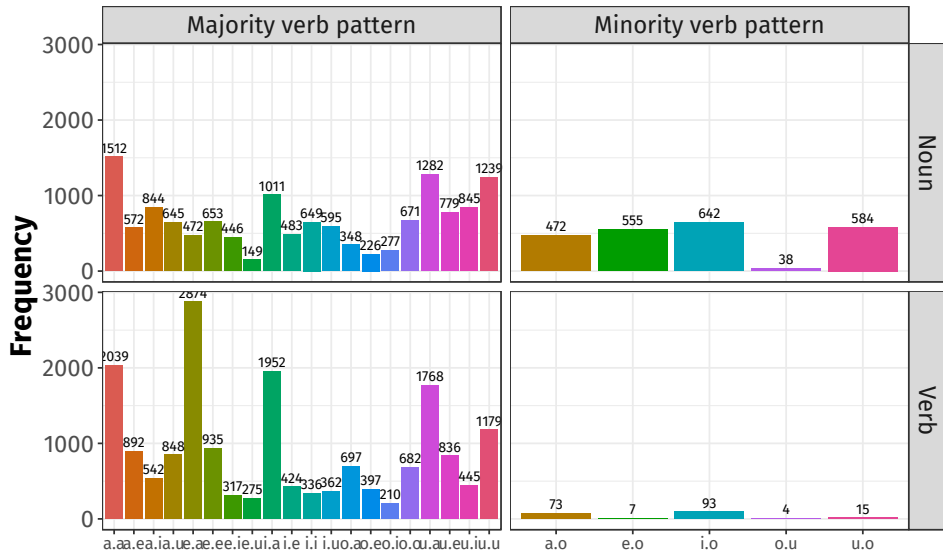


## Data set at large: Discussion II

- ▶ However, as /e.i/ and /o.u/ behave **similarly in verbs**, such explanations, would not expect a **significant difference** between their behaviour in **nouns**.
- ▶ This might instead suggest that the **under-representation** of these vowel pairs is simply due to the **well-groundedness** of their avoidance.
- ▶ And that there is a **stronger bias against /o.u/** than against /e.i/.
  - In fact, /e.i/ is over-represented in Kalanga and (very marginally) in Lozi.

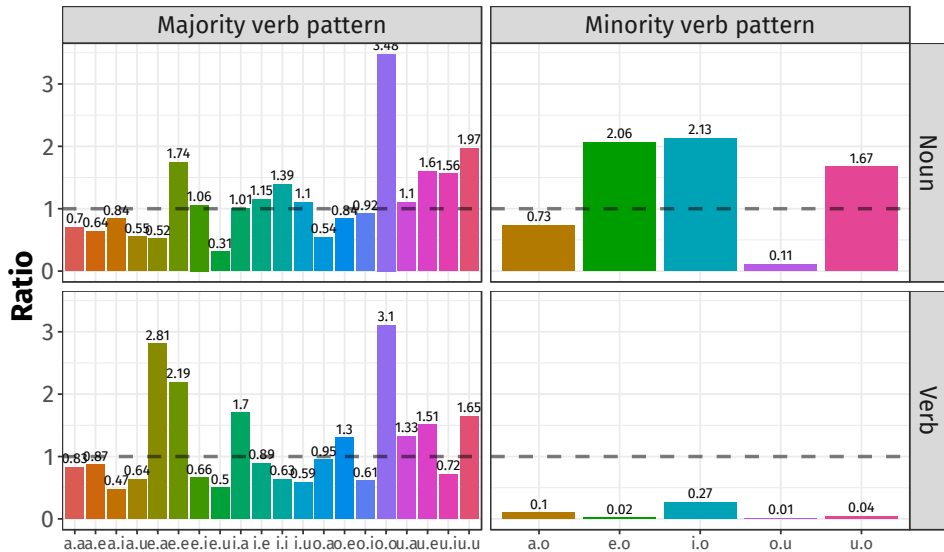


# Lozi: Results I

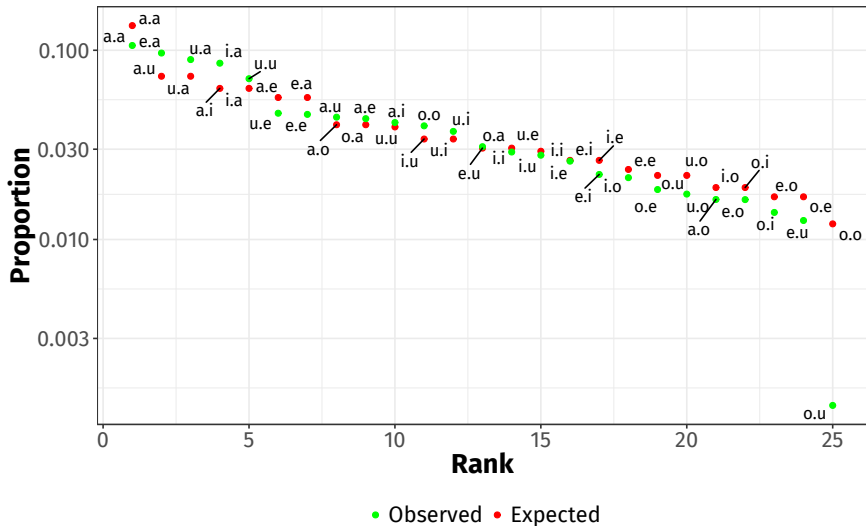




# Lozi: Results II



# Lozi: Results III





# Lozi: 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 /-ulul-/.
  - Undergoes a phonotactically-governed change to [-olol-] following /o/;
  - But elsewhere surface faithfully as [-ulul-].



# Lozi: Discussion (synchrony) II

► Exceptions in the data set:

- 4 instances occur across boundaries in compounds or with reduplication (see (19)).
- 4 are loan words from English:

- (14)
- |    |                                     |              |
|----|-------------------------------------|--------------|
| a. | bish <u>o</u> <u>p</u> <u>u</u>     | 'bishop'     |
| b. | sit <u>o</u> <u>f</u> <u>u</u>      | 'stove'      |
| c. | ingil <u>o</u> <u>p</u> <u>u</u>    | 'envelope'   |
| d. | w <u>o</u> <u>l</u> <u>u</u> pulete | 'wall plate' |

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

- (15)
- |    |                                |                                      |
|----|--------------------------------|--------------------------------------|
| a. | nj <u>o</u> <u>p</u> <u>u</u>  | 'damp, dewy place'                   |
| b. | mal <u>o</u> <u>p</u> <u>u</u> | 'beer'                               |
| c. | nd <u>o</u> <u>p</u> <u>u</u>  | 'elephant'                           |
| d. | s <u>o</u> <u>p</u> <u>u</u>   | 'fine grass growing in fertile soil' |



# Lozi: Discussion (synchrony) III

- ▶ In addition, the intervening segments are not entirely random.

- 63% have an intervening labial (see also (14) and (15)):

- (16) a. bubofu            'blindness'  
      b. siyopu            'hut used for ritual confinements'

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

- (17) a. lubolu            'double chin'  
      b. muholu            'stomach, tripe'

- ▶ In addition, 60% of these exceptions occur word-finally.



# Lozi: Discussion (synchrony) IV

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

(18)	a.	lubot <u>o</u> <u>o</u>	'bottle'
	b.	kop <u>o</u> <u>o</u>	'corporal'
	c.	mabasik <u>o</u> <u>o</u>	'bicycle'
	d.	-p <u>o</u> <u>o</u> fita	'to prophesy'
	e.	sin <u>o</u> <u>o</u>	'synod'
	f.	sitol <u>o</u> <u>o</u>	'strap (for inspanning oxen)'
	g.	c <u>o</u> <u>o</u>	'chalk'
	h.	d <u>o</u> <u>o</u> kota	'doctor'
	i.	n <u>o</u> <u>o</u>	'musical note'



# Lozi: Discussion (synchrony) V

► Places where the /o.u/ restriction does not apply:

(19) Across boundaries in compounds and with reduplication:

- a. kutweloo-buutuku 'pity, compassion' (Jalla 1982: sub voce)
- b. mafuloo-fuulo 'eagerness, zeal' (Jalla 1982: sub voce)
- c. mulyooluumbo 'senior person' (Mwisiya 1977:7)
- ↗ mulya u lumbo

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

- a. ne-ni-ta-too-kuuta 'I was going to have my hair cut' (Gowlett 1967:249)
- b. aba-too-luu-tusa 'they are not coming to help us' (Gowlett 1967:272)
- c. Bo-Muuwae 'Honourable Princess' (Fortune 2001:12)
- d. koo-kuu-mezi 'at a wet place' (Fortune 2001:33)



# Lozi: Discussion (synchrony) VI

- ▶ An explanation though may be found with reference to prosodic structure.
- ▶ It is not uncommon for **compounds** to be made up of **different prosodic words**.
- ▶ Similarly, **base** and **reduplicant** are often **separate prosodic words**.
- ▶ Some speakers prefer to write **prefixes** as **separate orthographic 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).
- ▶ 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).





# Lozi: Discussion (diachrony) I

- ▶ Modern Sotho does not exhibit 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/ [bɔfʊlʊla] → Lozi [bɔfɔlɔla] 'to outspan', not \*[bofulula]
  - But: Sotho [-tsʊma] → Lozi [-zuma] 'to hunt', not \*[-zoma]



## Lozi: 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.



## Lozi: 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)



# Lozi: 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:
  - longa 'être plein' — longola 'sortir (de la nourriture d'un pot)'
  - pumba 'mettre de la terre' — pumbula 'déterrer'
- ▶ It is not clear whether back height harmony is found in nouns.
  - 👉 Though based on the Luyana loans in Lozi in (15), it is possible it is not.

# Summary & conclusions

- ▶ I have presented you with vowel-pair frequency data from six Bantu languages.
- ▶ And argued on this basis that avoiding /o.u/ is better ground than avoiding /e.i/ (though both are grounded).
- ▶ I then focused on Lozi and argued that the data from this language actually suggest a ban on /o.u/ only and that this is blind to part of speech.
- ▶ This is something that any formal account of height harmony in Lozi must reflect.

## Acknowledgements

I would like to thank 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**, the **PhilSoc ECR Forum** and the **Manchester Forum in Linguistics** for their feedback.

## Contact

 <https://tiny.cc/sjn>

 [stephen.nichols@manchester.ac.uk](mailto:stephen.nichols@manchester.ac.uk)

- Archangeli, Diana, Jeff Mielke & Douglas Pulleyblank. 2012a. From Sequence Frequencies to Conditions in Bantu Vowel Harmony: Building a grammar from the ground up. *McGill Working Papers in Linguistics* 22(1).
- Archangeli, Diana, Jeff Mielke & Douglas Pulleyblank. 2012b. Greater than noise: frequency effects in Bantu height harmony. In Bert Botma & Roland Noske (eds.), *Phonological Explorations: Empirical, Theoretical and Diachronic Issues*, 191–222. Berlin: De Gruyter.
- 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* 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.
- 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.
- Gusimana, Barthelemy. 1972. *Dictionnaire pende-français*. Available to download from the *Comparative Bantu Online Dictionary*. Contributed by John Lowe. URL: <http://www.cbold.ish-lyon.cnrs.fr/Dico.asp?Langue=Gusimana>.
- Harris, John. 1994. Monovalency and opacity: Chicheŵa 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.
- 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, superstitions, 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 (3<sup>rd</sup> 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.
- Kisseberth, Charles W. 1996. Makhua–English lexicon. Available to download from the Comparative Bantu Online Dictionary. Contributed by Charles M. Kisseberth. URL: <http://www.cbold.ish-lyon.cnrs.fr/Dico.asp?Langue=Makhua>.
- Martin, Andrew. 2011. Grammars leak: Modeling how phonotactic generalizations interact within the grammar. *Language* 87(4). 751–70.
- Mathangwane, Joyce. 1994. Kalanga lexical database. Available to download from the Comparative Bantu Online Dictionary. Contributed by Joyce Mathangwane. URL: <http://www.cbold.ish-lyon.cnrs.fr/Dico.asp?Langue=Kalanga>.
- Moto, Francis. 1989. Phonology of the Bantu lexicon. PhD thesis, University College London.
- Mtenje, Al D. 1985. Arguments for an autosegmental analysis of Chicheŵa vowel harmony. *Lingua* 66(1). 21–52.
- Mtenje, Al D. 2001. Database of 'Chewa dictionary'. Available to download from the Comparative Bantu Online Dictionary. Contributed by Al D. Mtenje. URL: <http://www.cbold.ish-lyon.cnrs.fr/Dico.asp?Langue=Chewa>.
- Mwisiya, M. W. 1977. Introduction to Silozi Grammar. Lusaka: National Educational Company of Zambia Ltd.
- Ngunga, Armindo. 2001. Yao lexical database. Available to download from the Comparative Bantu Online Dictionary. Contributed by Armindo Ngunga. URL: <http://www.cbold.ish-lyon.cnrs.fr/Dico.asp?Langue=Yao>.
- 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. Chicheŵa vowel harmony and Underspecification Theory. *Linguistic Analysis* 22. 218–45.