Vowel-pair frequencies & phonotactic restrictions in Lozi

Stephen Nichols
University of Manchester

Philological Society Early Career Researcher Forum
University of Oxford

9 March 2019

► A case study of vowel-pair frequencies in Lozi.
  ▫ Using data from a large digitised dictionary, namely Jalla (1982).

► The results of which have implications for formal analyses.
  ☞ The only synchronic phonotactic vowel co-occurrence restriction in the language is a part-of-speech-blind ban on the vowel pair /o.u/ (in ~the non-prefixal domain).

Bantu height harmony

- ▶ Vowel height harmony is extremely common in the Bantu languages (see e.g. Clements 1991, Hyman 1999:§2, 2003, Odden 2015:§1).
  - — In the vast majority of cases, harmony is confined to verbs.

- ▶ 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).

▶ Thus, it can, descriptively at least, be split into front and back height harmony.

CANONICAL FIVE-VOWEL BANTU HEIGHT HARMONY

(1) a. Front height harmony: $\textbf{i} \rightarrow \textbf{e}$ / {**e o**} (C) _

    b. Back height harmony: $\textbf{u} \rightarrow \textbf{o}$ / **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. -peta
      'to fold'

   d. -longa
      'to pack'

   e. -kaka
      'to tie'

(3) **Applicative suffix:**

   a. -bilila
      'to sew for'

   b. -tungila
      'to thread for'

   c. -p<u>ete</u>la
      'to fold for'

   d. -l<u>o</u>ng<u>e</u>la
      'to pack for'

   e. -kakila
      'to tie for'

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

(4) **Unsuffixed:**

 a. -bila
  'to sew'

 b. -tunga
  'to thread'

 c. -peta
  'to fold'

 d. -longa
  'to pack'

 e. -kaka
  'to tie'

(5) **Reversive suffix:**

 a. -bilulula
  'to unsew'

 b. -tungulula
  'to unthread'

 c. -petulula
  'to unfold'

 d. -l<u>o</u>ng<u>olo</u>la
  'to unpack'

 e. -kakulula
  'to untie'

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

▶ Lozi is a Bantu language spoken mainly in Zambia.
  — Around 730,100 speakers (Eberhard et al. 2019).
  — Guthrie code: K.21 (Maho 2009).

▶ Most closely related to Sotho (S.33), Pedi (S.32),
  Tswana (S.31) and Kgalagadi (S.311).
  — But has also been heavily influenced by Luyana (K.31).
    📖 See Gowlett (1989) for a discussion of Lozi's history.

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)

(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'
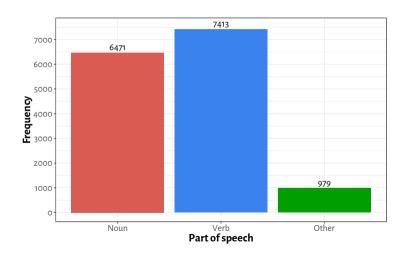
b. -lukulula
'to unweave'

c. -lekulula
'to resell'

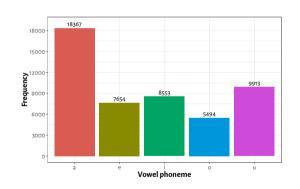d. -l<u>o</u>ng<u>olo</u>la
'to unpack'

e. -tamulula
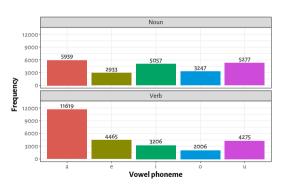'to unfold'

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

METHODOLOGY I

▶ The data come from Jalla (1982).
  — This is a Lozi–English dictionary database.
  — 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 tagged for part of speech.

▶ Perfective verb forms were then removed, which left a final total was 14,863.
  — *-tamile* ← *-tama* 'to tie';
  — *-lekezi* ← *-lekela* 'to buy for';
  — *-lutuluzi* ← *-lutulula* 'to unthatch';
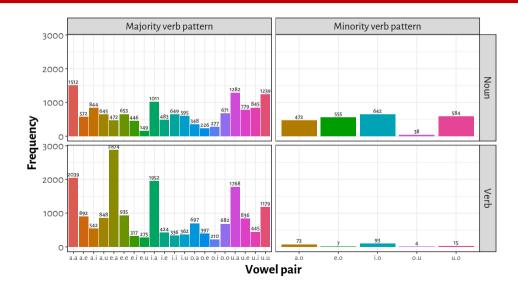 — *-mizize* ← *-miza* 'to swallow'.

Methodology III

▶ Orthographic long vowels were reduced to short vowels and intervening consonants were ignored.

— This means that VCV sequences were treated the same as VV sequences.

▶ The observed and expected frequencies of all 25 possible vowel pairs were calculated.
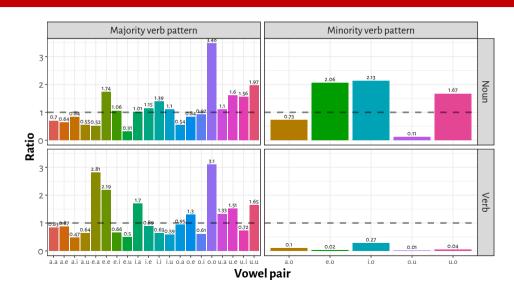
▶ As were the corresponding observed–expected ratios.

▶ The near-total absence of /o.u/ in both verbs and noun
  — Suggests there is an active phonotactic vowel co-occurrence restriction against /o.u/;
  — And that 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-] when preceded by /**o**/;
  — But elsewhere surface faithfully as [-**u**l**u**l-].

DISCUSSION II

- Places where /o.u/ this restriction does not apply:

  (11)   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)

  (12)   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

- Exceptions in the data set:
    - 4 instances occur across boundaries in compounds or with reduplication (cf. 12).
    - 4 are loan words from English:

  (13) a. bish**op̲u̲**   'bishop'
        b. sit**of̲u̲**   'stove'
        c. ingil**op̲u̲**   'envelope'
        d. w**ol̲u̲**pulete   'wall plate'

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

  (14) a. nj**op̲u̲**   'damp, dewy place'
        b. mal**op̲u̲**   'beer'
        c. nd**op̲u̲**   'elephant'
        d. s**op̲u̲**   'fine grass growing in fertile soil'

▶ In addition, the intervening segments are not random.

— 63% have an intervening labial (see also 13 and 14):

(15) a. bub**o**f**u** 'blindness'

b. siy**o**p**u** 'hut used for ritual confinements'

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

(16) a. lub**o**l**u** 'double chin'

b. muh**o**l**u** 'stomach, tripe'

▶ 60% occur word-finally.

— In an ongoing (indirectly-related) production study of Bemba, Nyanja and Lozi, I find that, for some Lozi speakers, final /u/ is deleted/devoiced word-finally and, when retained, it often appears to be phonetically lowered after /o/.

DISCUSSION V

▶ 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.:

(17)   a. lubot**olo**   'bottle'

       b. kop**olo**   'corporal'

       c. mabasik**olo**   'bicycle'

       d. -p**olo**fita   'to prophesy'

       e. sin**odo**   'synod'

       f. sitol**opo**   'strap (for inspanning oxen)'

       g. c**oko**   'chalk'

       h. d**oko**ta   'doctor'

       i. n**oto**   'musical note'

▶ One hypothesis as to the origin of the disparity of nouns and verbs might be:

 — /a.o, e.o, i.o, u.o/ were once also phonotactically disallowed throughout the language;

 — Previous examples of these pairs were removed and new ones prevented from arising;

 — When these restrictions were lifted, innovative word forms with these pairs arose;

 — They reoccurred more in nouns because of a higher rate of lexical innovation.

▶ This claim though requires more thorough investigation.

▶ However, on the face of it, examining the data it seems unlikely:

 — Only 44 of the 2,456 instances in nouns are in English loans;

 — And 21 marked as being from Luyana.

► To a certain extent, this may also be an artefact of the data set:

— Whereas verbs are prefixless in their citation forms, nouns are usually included with their relevant noun class prefix, which may include /a, i, u/;

— Taking this into account, although /a.o, i.o, u.o/ are less frequent, they are still not very infrequent (and certainly not as strikingly infrequent as /o.u/).

▶ Taking inspiration from work by Martin (2011) on Navajo sibilant harmony and English geminates, an alternative view to this could be:

— A ban on /a.o, e.o, i.o, o.u, u.o/ was once was active only in verbs;
— /o.u/ is the most marked or phonetically natural of these gaps;
— The phonotactic restrictions in verbs on /a.o, e.o, i.o, u.o/ were lifted;
— A small number of innovative verb forms containing these pairs arose;
— The remaining phonotactic ban on /o.u/ in verbs exerted a gradient effect on nouns;
— Until this reached a tipping point and the ban became blind to part of speech.

▶ Contra, e.g. Archangeli et al. (2012a,b), who claim that similar effects are due to inductive support rather than naturalness or markedness.

☞ In which case, they do not predict the pattern seen in Lozi /a.o, e.o, i.o, u.o/ also find inductive support but show no comparable effect.

- I have presented you with vowel-pair frequency data from Lozi.

- And argued that these suggest a part-of-speech-blind ban on the vowel pair /o.u/ but not on /a.o, e.o, i.o, u.o/.
  - ☞ Something that any formal account of height harmony in Lozi must reflect.

- I have also provided some discussion of the potential origin of this pattern.

CONTACT

🖥 https://tiny.cc/sjn

✉ 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. 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*.

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: 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.

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.

Martin, Andrew. 2011. Grammars leak: Modeling how phonotactic generalizations interact within the grammar. *Language* 87(4). 751–70.

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.

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.

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