The origin of click consonants

Thora Daneyko

May 31, 2017

Structure

1. What are clicks and where do they occur?
2. Do clicks correlate with large phoneme inventories?
3. What is the relationship between inventory size and clicks?

What are clicks?

- Consonants produced with an inward sucking movement of the tongue
- IPA distinguishes only five click symbols: [ʘ], [ǀ], [ǃ], [ǂ], [ǁ]
- But languages can have dozens of clicks!
  - Labial, dental, palatal, alveolar, lateral, retroflex, ...
  - Tenuis, aspirated, voiced, nasal, glottalized, aspirated nasal, ...
- Extremely rare phoneme: Only in southern Africa; one language in Australia

Known click languages (Güldemann 2007, Phoible)

- Inherited clicks:
  - All Khoisan languages (Northern, Southern, Central Khoisan)
- Borrowed clicks:
  - Some Bantu (Niger-Congo) languages
  - Dahalo (Afro-Asiatic)
  - Masalit (Nilo-Saharan)
  - Hadza (isolate)
  - Sandawe (isolate/Khoisan)
- Invented clicks:
  - Damin, speech register of Lardil (Tangkic, Australia)

Figure: African language families (Wikipedia).

Project: Inventory size

The effect of inventory size

- The languages with the largest phoneme inventories are click languages
- Click languages seem to often have large phoneme inventories

- My task: Do languages using clicks (as phonemes) have generally bigger sound inventories than languages not using clicks?

The Phoible database

- Comprises seven large phonetic databases and many source grammars
  - One inventory per source

  $\rightarrow$ Multiple inventories for some languages
- Contains 1672 languages, 2155 inventories and 2160 phonemes
- Transcription in IPA

Data collection

- Extracted data for all languages from Phoible
  - Inventory sizes (consonants, vowels, clicks, total phonemes)
  - Genus and family
  - Geographical data (latitude, longitude)
- Removed 5 languages with missing geographical data
- Took mean of inventory sizes for languages with multiple inventories

$\rightarrow$ 1649 non-click languages

$\rightarrow$ 18 click languages

Click languages in Phoible

- 6 Khoisan (Hai||om, Ju|'hoan, Khoekhoe, Khwe, Sandawe, !Xóõ)
- 9 Bantoid/Niger-Congo (Chichopi, Fwe, Kgalagadi, Ndebele, Ronga, Sesotho, Xhosa, Yeyi, Zulu)
- 1 Afro-Asiatic (Dahalo)
- 1 Nilo-Saharan (Masalit)
- 1 isolate (Hadza)

Results: Do clicks imply larger inventories?

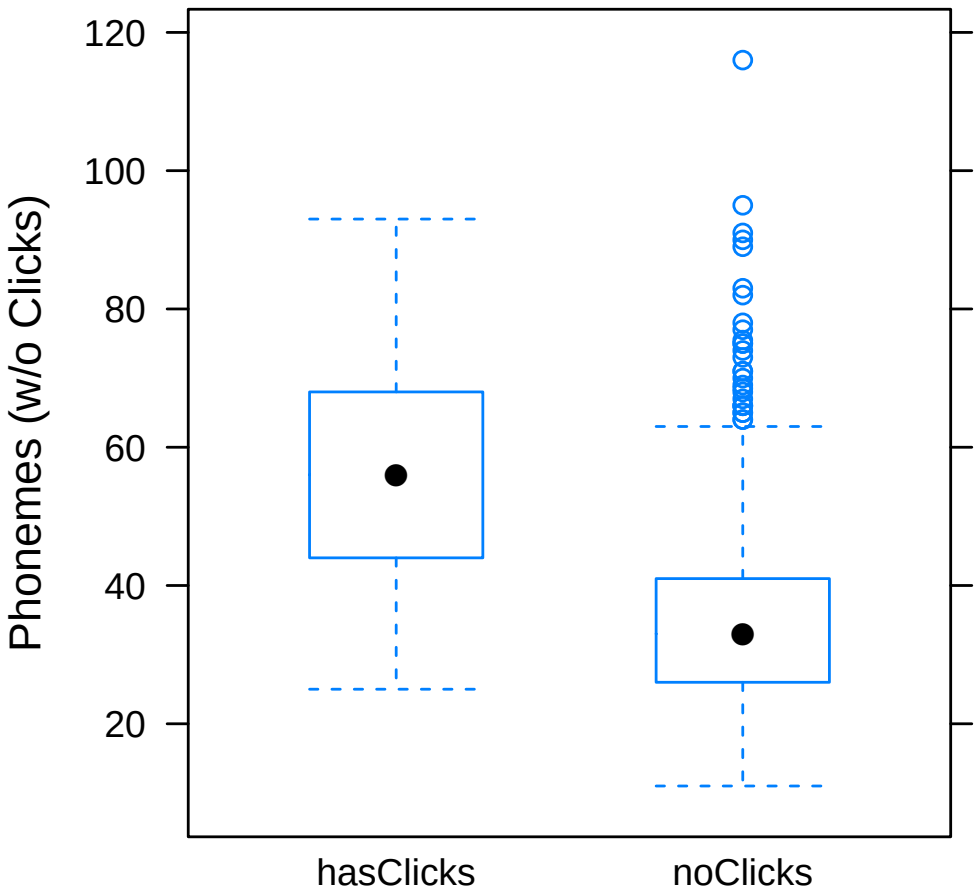

Figure: The number of phonemes in languages with and without clicks. The difference is significant with $p = 6.361 \times 10^{-7}$ according to the Wilcoxon rank sum test.

Results: Do clicks imply larger inventories?

- Having clicks seems to be correlated with consonant inventory size, not vowel inventory size

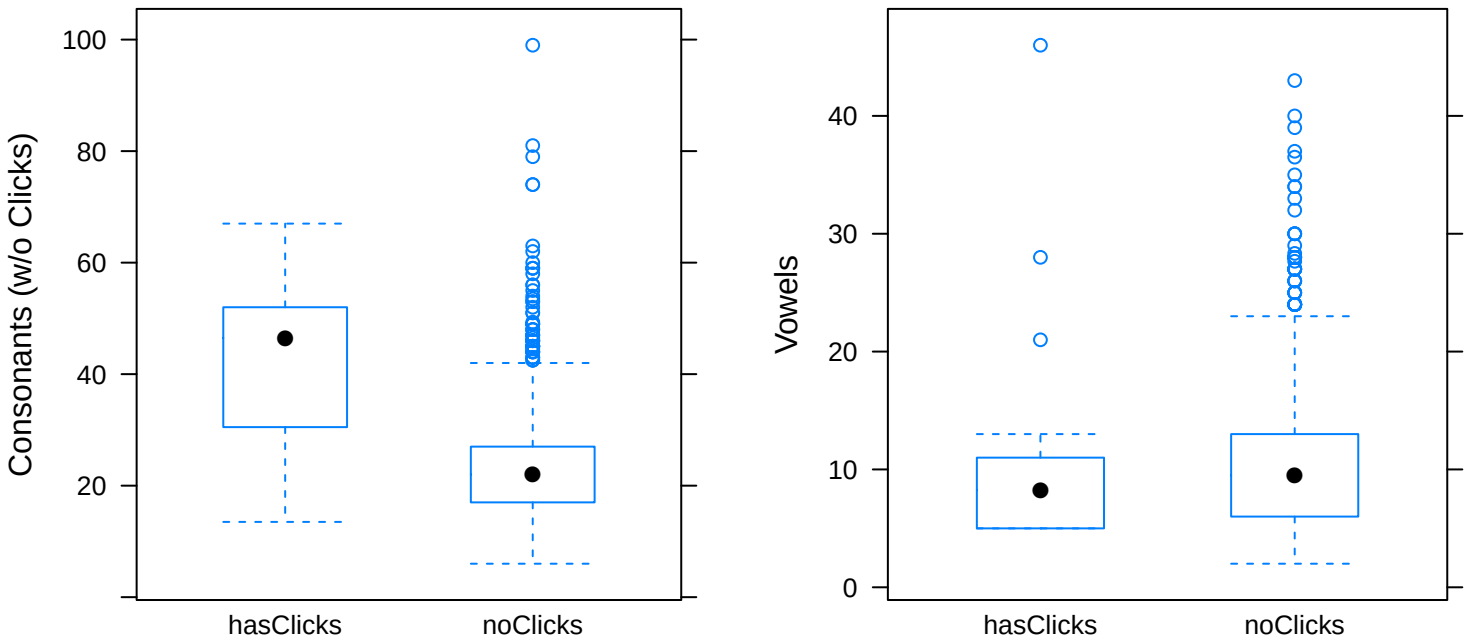

Figure: The number of consonants and vowels in languages with and without clicks. The difference is significant for consonants with $p = 5.591 \times 10^{-8}$, but not for vowels with $p = 0.358$ (Wilcoxon).

However...

- All click languages from single geographical area
- Most click languages from Khoisan and Bantu family
- Is inventory size rather correlated with this area and these families?
  - $\rightarrow$ Test correlation for click families?
    - Clicks only evolved independently in Khoisan, but all Khoisan languages have clicks $\rightarrow$ pointless to test
    - Only single click language for Afro-Asiatic, Nilo-Saharan and Hadza languages $\rightarrow$ pointless to test
    - Bantu?

Inventory size in Bantu languages

- New hypothesis: Bantu languages with clicks have larger consonant inventories than Bantu languages without clicks
- Repeat tests for Phoible's Bantu languages
  - Phoible only has Bantoid as a genus $\rightarrow$ Bantoid = Bantu + related languages exclusively from Nigeria and Cameroon

  $\rightarrow$ Remove languages from these countries from the sample
- 9 click vs. 93 non-click languages

Inventory size in Bantu languages

- Clear tendency for Bantu click languages to have larger consonant inventories than their non-click relatives

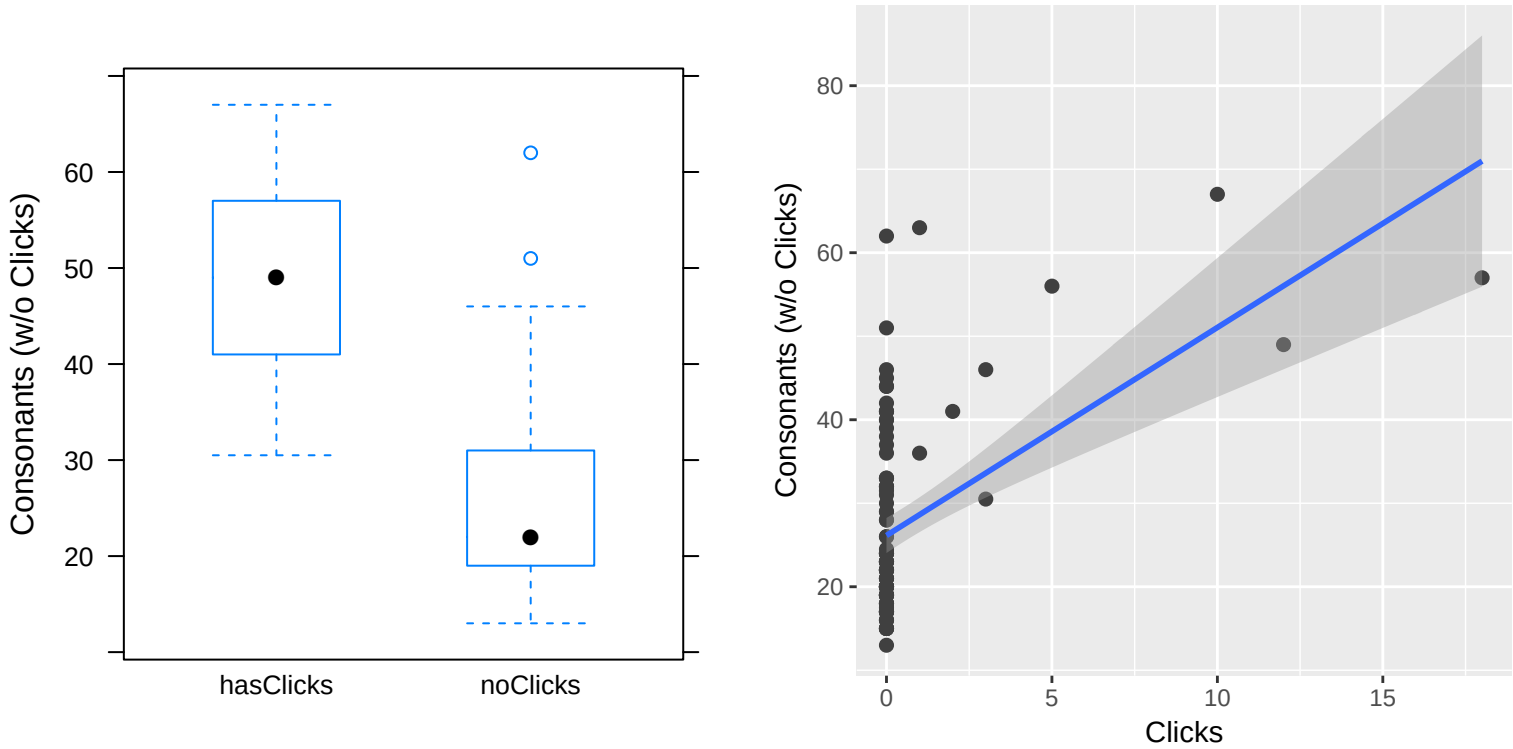

Figure: The number of consonants in Bantu languages with and without clicks. The difference is significant with $p = 1.751 \times 10^{-5}$ (Wilcoxon). The number of consonants and clicks is significantly correlated with $p = 8.095 \times 10^{-8}$ according to Pearson's correlation test.

Distance to Khoisan languages

- Is the inventory size really correlated with clicks or rather another effect of proximity to Khoisan?
- Calculated geographical distance to closest Khoisan language for each Bantu language

Distance to Khoisan languages

- Proximity to Khoisan does not seem to have an effect on both inventory size and click borrowing

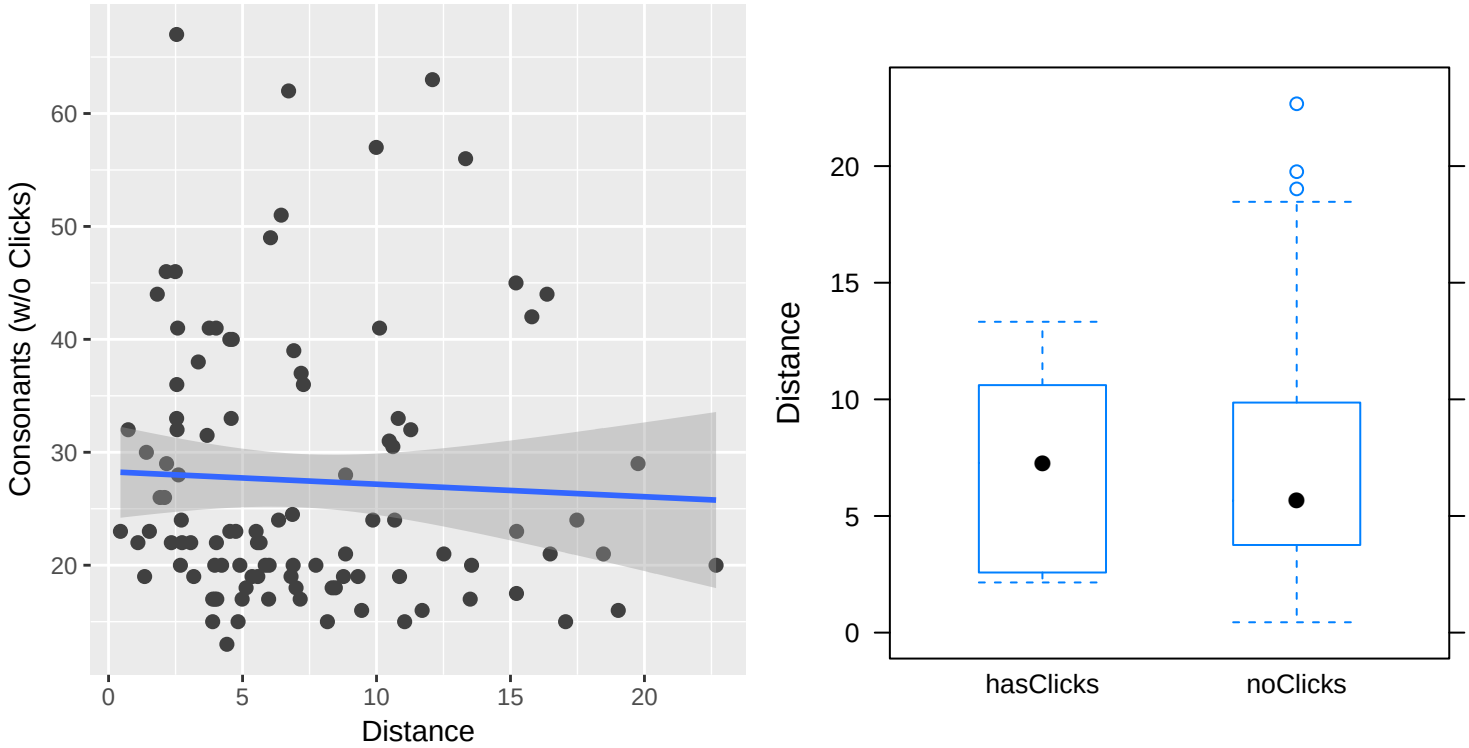

Figure: The size of the consonant inventories of Bantu languages and whether they contain clicks with respect to the distance to the nearest Khoisan language. The number of consonants and distance to Khoisan are not correlated with $p = 0.649$ (Pearson), neither is the effect of distance on having clicks significant with $p = 0.75$ (Wilcoxon).

Distance to Khoisan languages

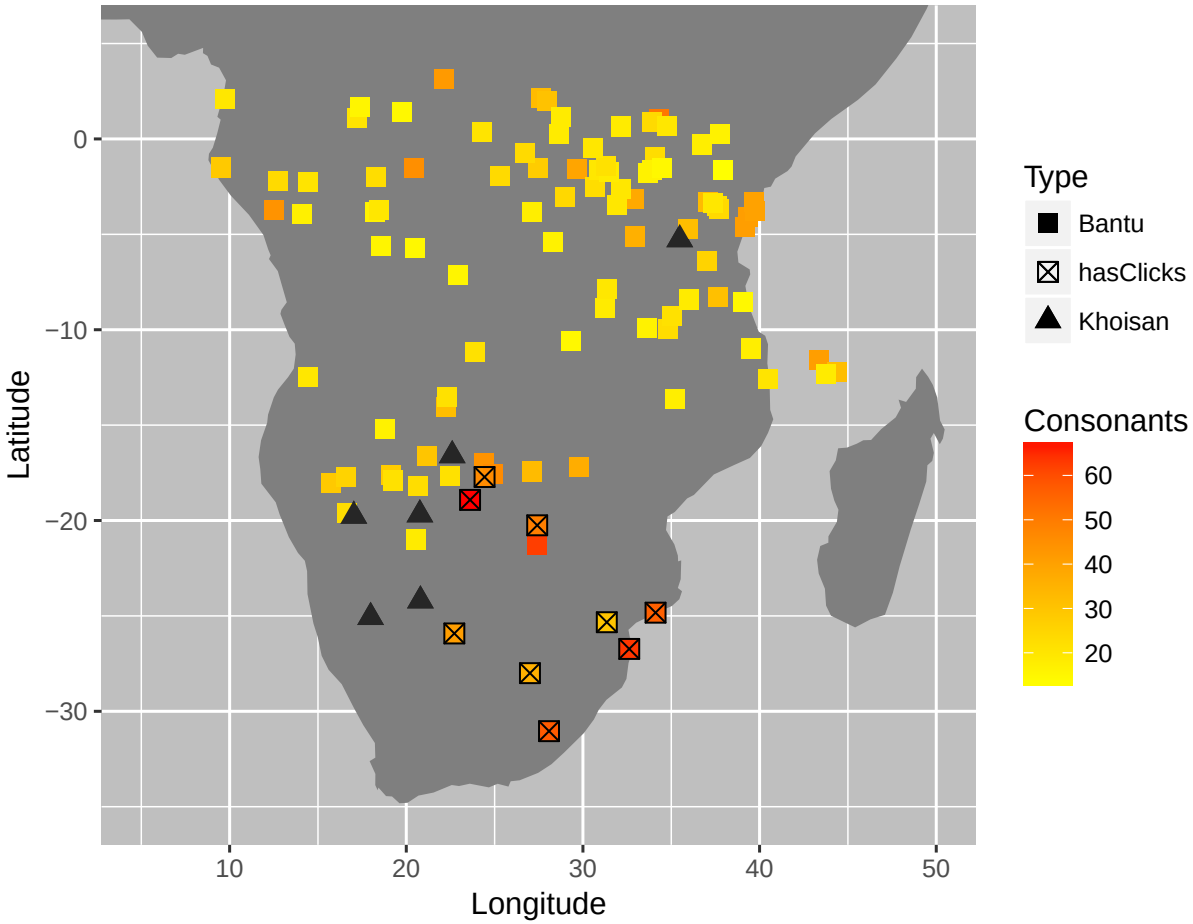

Figure: The location of Bantu (squares) and Khoisan (triangles) languages in southern Africa. The consonant inventory size of the Bantu languages is indicated by color (yellow = small, red = large). Click Bantu languages are marked by a black box.

Discussion

Sample size

- 18 click vs. 1649 non-click languages → statistical significance?
- Disagreement on inventory sizes within Phoible database:
  - Incorporation of seven overlapping databases with different counting criteria
  - e.g. Telugu (Dravidian): 43 phonemes in UPSID, 68 in SPA (SPA counts geminate consonants, UPSID does not)
  - e.g. Khoekhoe (Khoisan): 4 clicks in UPSID, 20 in SPA (same basic clicks, more distinctions in manner and phonation in SPA)

Open questions

- Yes, languages with clicks have larger sound inventories than languages without clicks
- This trend is also apparent in the Bantu subfamily

- What is the relation between having clicks and inventory size?

Clicks: Sounds of the proto-language?

- Clicks = archaic sounds of proto-human?
- Inventory sizes decrease with distance from Africa
- Were they the first sounds to be lost and are only retained in southern African languages?
  - Popular theory
  - But no linguistic evidence for that!

Many consonants → clicks?

- Do languages develop clicks when they already have many other consonants?
- Clicks are 'marked' and only adopted when there are only few other possibilities left
  - What is a 'marked' sound?
  - Clicks are only rare as phonemes, not as sounds (→ paralinguistic usage)
  - Clicks are easy to produce and distinguish

??? $\rightarrow$ clicks and other consonants?

- Why do Bantu click languages have so many more consonants?
- Did they borrow them together with the clicks?
- How and why did Bantu borrow its clicks?

The Bantu click inventory (Herbert 1990)

- Bantu has some Khoisan loanwords
- Bantu borrowed some Khoisan consonants, and almost exclusively clicks
- Bantu did not borrow Khoisan vowels, phonotactics, productive morphology or syntax

- Clicks do also occur in Bantu cognates:
  - *-consa* [|onsa] 'fall, drip, leak' vs. *ilithonsi* [ilit^{h}onsi] 'drop of liquid'
- Often, the original Bantu word is retained with a different meaning:
  - *-chela* [|^{h}ela] 'sprinkle' vs. *-thela* [t^{h}ela] 'pour (out)'

→ How did the Bantu speakers adopt the Khoisan clicks?!

Bantu: Clicks as a means of respect? (Herbert 1990)

- Hlonipha [ɬonipʰa] ‘respect’
- Avoidance custom of Nguni (Bantu subgroup) people including linguistic taboo
  - Married women may not pronounce the names of their father-in-law and other senior males
  - Also forbidden: Any syllable contained in these names
- Strategies to follow taboo: Lexical or phonetic substitution

Bantu: Clicks as a means of respect? (Herbert 1990)

- Clicks as substitution for offensive consonants?
- Very different from (small) native inventory
- Will not create homophonous words
- Explains strange nature of click borrowings:
  - Tendency to borrow most 'alien' sound in large quantities
  - Integration into native vocabulary
  - Coexistence of same native words with and without clicks

!! Hlonipha is most pervasive in the Bantu languages with the most clicks (Zulu and Xhosa)

Hlonipha → clicks and other consonants?

- Were the clicks just some of many consonants that were borrowed for Hlonipha?
- Hlonipha → many consonants, including clicks?

Conclusion

Conclusion

- Clear correlation between inventory size and having clicks
- Unclear relation between these features
  - Loss of clicks in other languages?
  - Causal? (Many consonants $\rightarrow$ clicks)
  - Two sides of the same coin? (Hlonipha $\rightarrow$ many consonants + clicks)
- All imply that clicks are highly 'marked' sounds that are either lost early or adopted late or consciously

References

Güldemann, Tom (2007). *Clicks, Genetics, and "proto-world" from a Linguistic Perspective*. Inst. für Afrikanistik.

Herbert, Robert K. (1990). "The sociohistory of clicks in Southern Bantu". In: *Anthropological linguistics*, pp. 295–315.

Moran, Steven, Daniel McCloy, and Richard Wright, eds. (2014). *PHOIBLE Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URL: http://phoible.org/ (visited on 05/06/2017).

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org.