# Exploiting machine algorithms in vocalic quantification of African English corpora

**Adeiza Lasisi Isiaka**
Department of English Studies,
Adekunle Ajasin University,
adeiza.isiaka@yahoo.com

## Abstract

Towards procedural fidelity in the processing of African English speech corpora, this work demonstrates how the adaptation of machine-assisted segmentation of phonemes and automatic extraction of acoustic values can significantly speed up the processing of naturalistic data and make the vocalic analysis of the varieties less impressionistic. Research in African English phonology has, till date, been least data-driven – much less the use of comparative corpora for cross-varietal assessments. Using over 30 hours of naturalistic data (from 28 speakers in 5 Nigerian cities), the procedures for segmenting audio files into phonemic units via the Munich Automatic Segmentation System (MAUS), and the extraction of their spectral values in Praat are explained. Evidence from the speech corpora supports a more complex vocalic inventory than attested in previous auditory/manual-based accounts – thus reinforcing the resourcefulness of the algorithms for the current data and cognate varieties.

Keywords: machine algorithms; naturalistic data; African English phonology; vowel segmentation

## 1 Introduction

The basis of automatic segmentation through forced alignment is vital to the phonetic analysis of large naturalistic data. The processes are trained to execute the correlation between phonological categories and their acoustic properties. Depending on what to be analysed, linguistic categories are mapped unto the equivalent signal constituents, which are then broken into labelled segments. A manual alternative to this process has obvious limitations. First, it would be impressionistic - as the labelling of the phonemic units would rely on auditory evaluation rather than objective measurement. Second, it would take an awfully much longer time to complete a segmentation task for a truly naturalistic data. To walk around these, the Munich Automatic Segmentation System (MAUS) comes in handy. The MAUS web services make possible the uploading of audio and the corresponding txt.files into MAUS for processing. The results (usually in textgrids) are returned to the local computer. The MAUS algorithm combines simple forced alignment based on Hidden Markov Modelling (HMM) with ancillary mesh of statistical enhancement to accommodate specific variants of languages (Kisler, Reichel & Schiel 2017). It locates the most suitable correlation for the inputted speech signals based on pre-trained parameters for the language(s) in question. The machine's predictive potential however depends on the quality of inputs (clarity of audio recordings and purity of the txt.files) which it basically interprets against all embedded signals and categories. MAUS, as currently implemented, supports about 14 different languages which include English, and an independent SAM-PA that automatically makes use of equivalent HMMs to yield best relevant approximations for specific sounds.

## 2 Data segmentation

To adapt MAUS for the segmentation of vowel tokens, certain procedures are required. First, the transcription must be MAU-compatible. The txt.files are required as tab-delimited, and the audio recordings as wav.files. These done, MAUS force-aligns the inputted pair of audio signal and its corresponding txt.file and returns both as textgrids to an assigned folder on the computer. Since MAUS is pre-designed for phonemic segmentation based on prescribed symbols, an add-on Praat script was further run to 'cleans up' the textgrid and re-arrange the tiers as well as return the vowels based on Wells' Lexical Set (Wells 1982). For further editing, the audio files with their textgrids were re-opened in Praat (Boersma & Weenink, 2016). Measurements of spectral details including formants, pitch, fundamental frequency ($F_0$), bandwidths, etc were then extracted into spreadsheet for over 100,000 vowel tokens and sorted prior to statistical analysis in R (RStudio 2015).

### 2.1 Praat illustrations of MAUS-segmented files

**Figure 1:** A MAUS-segmented textgrid of a speaker saying *for one to be born again* in rapid speech (IFUO4_SI_1: Age =52, educated speaker, female, Nigerian).

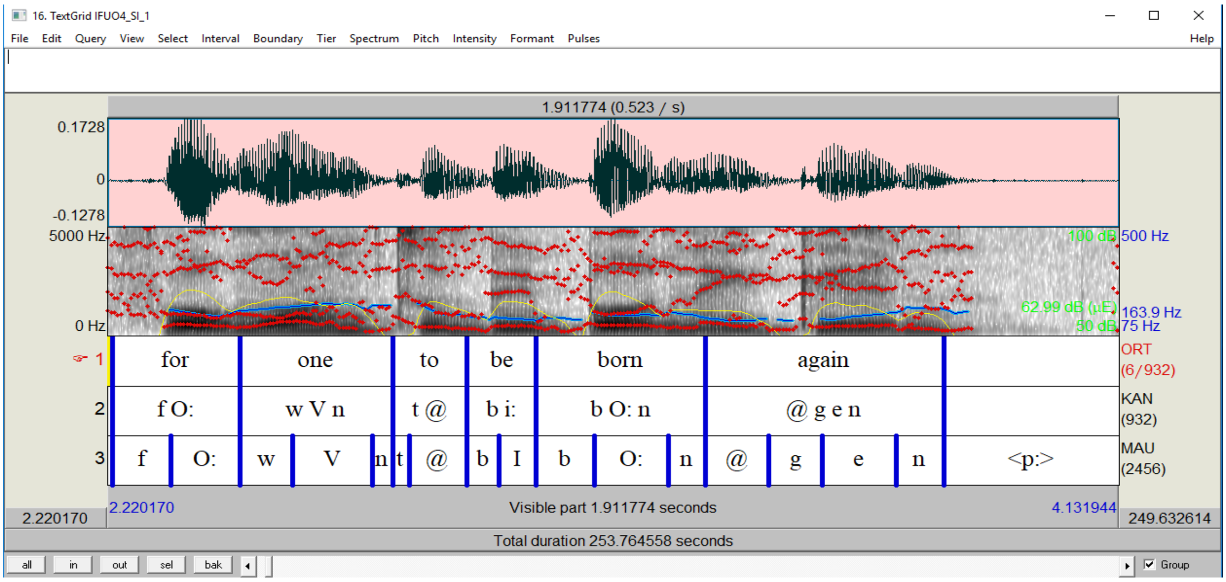
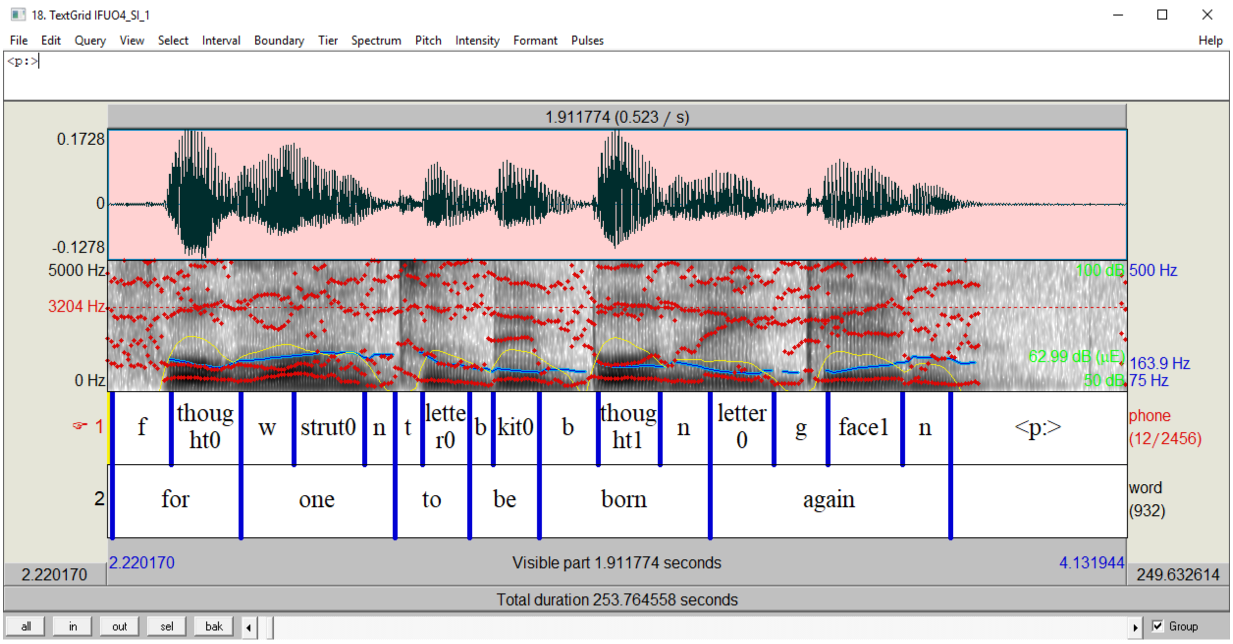
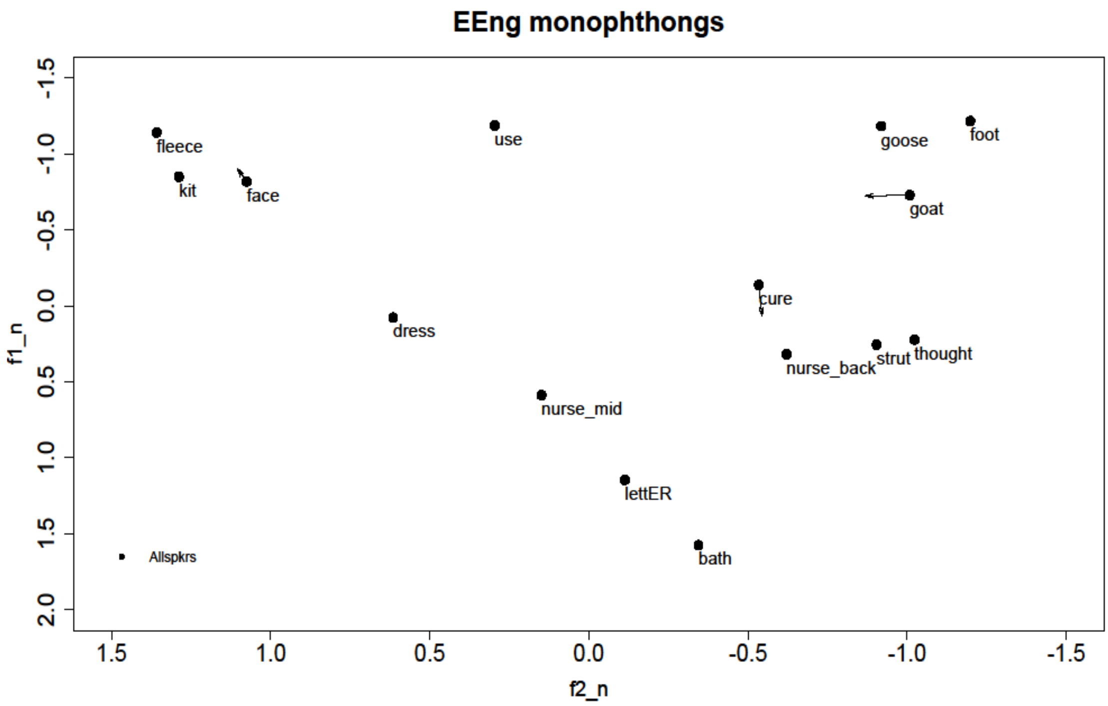



**Figure 2**: A Praat-cleaned version of the textgrid above – showing the alignment of MAUS coding with the Lexical Set paradigm. All analysable tokens were coded with ‘1’ and those to be excluded with ‘0’.



## 3 Summary and Conclusion

**Figure 3**: F1/F2 Lobanov-normalised visual summary of Ebira English pure vowels (including monophthongal diphthongs)




This work proposes the adaptation of machine algorithms such as MAUS and Praat for the analysis of L2 English corpora. Mid-point formant measurements of over 100,000 naturalistic vowel tokens of a Nigerian English variety were extracted via a Praat script and analysed based on the Lexical Sets paradigm. A qualitative variance between the cluster of KIT and FLEECE was supported in F1 (p= 0.00655), as well as between FOOT and GOOSE in F2 (p=0.0144). The distinction between LOT and THOUGHT based on formant values was however insignificant (p=0.826). Also, STRUT was significantly distinct from the low back classes of LOT/THOUGHT vowels (p=0.000531). Other instances of differentiation between the vowels were however cued by durational differences. With regard to the size of pure vowels, the analysis supports a 13-vowel system for the variety. This is thus indicative of a more complex monophthongal inventory than previously attested for Nigerian English vowels; as well as the methodological validity of MAUS, Praat and R's algorithms over manual and auditory-based assessments.

## References

J.C. Wells, Accents of English I: An Introduction, University Press, Cambridge, 1982.

R Studio Team, Integrated Development for R. RStudio Inc. http://www.rstudio.com/, 2015 (accessed 03 March 2015).

P. Boersma, D. Weenink, Praat: Doing Phonetics by Computer [Computer program], http://www.praat.org/, (accessed 02 June 2016).

T. Kisler, U.D. Reichel, F. Schiel, Multilingual processing of speech via web services, Computer Speech & Language 45 (2017) 326-347.