Aitken‘s Law Revised

The Scottish Vowel Length Rule in 21st century SSE

Introduction

One of the most notable features of Scottish English is the Scottish Vowel Length Rule (SVLR) / Aitken‘s Law which states that vowels are „long in stressed open syllables, before voiced fricatives and /r/, and before morpheme boundaries and short elsewhere; with the two exceptions of /ɪ/ and /ʌ/ which are invariably short.“ (McClure 1977). The SVLR partially contradicts with the standard quantity alternation pattern present in most varieties of English and many other languages, namely the Voicing Effect (VE) (Chen 1970). Whereas the VE generally triggers longer vowel realizations before all voiced consonants, the SVLR conditions a shortening of tense vowels before following laterals, nasals and voiced oral stops. In this respect, the vowel length patterns in Scottish English differ from all other varieties of English (Chevalier 2019) (see Tab. 1).

Vowel length patterns according to SVLR and VE in different contexts

Vowel length pattern	voiceless consonants	voiced plosives, nasals and laterals	voiced fricatives, /r/ and morpheme boundaries
SVLR	short	short	long
VE	short	long	long

*(for a more detailed overview, see Rathcke & Stuart-Smith 2016: 406)

Even though several studies were carried out on the SVLR, the datasets & methods of most previous investigations are clearly outdated and
- analysed vowel duration in word list/carrier sentence readings
- investigated a limited number of vowels and contexts
- had a limited geographical scope (mostly central belt)
- had a limited & unbalanced (age,sex) number of informants

Only the most recent studies (Rathcke & Stuart Smith 2016, Chevalier 2019, Stuart-Smith et al. 2019) investigated Scottish vowel length patterns in naturally occuring language with more advanced statistical methods. Yet, these studies predominantly use Scots data and lay the focus on a particular region. Thus, the unresolved questions of the **geographical scope of Aitken‘s Law** as well as the influence of **age and gender-related variation** have not yet been satisfactorily answered, especially not for Scottish Standard English (SSE).

Aims of this study

The present study addresses these questions. It aims to:
- investigate SVLR and VE patterns in contemporary 21st century SSE
- achieve representativeness in terms of region, age and sex
- account for the influence of suprasegmental factors
- use advanced statistical methods and large & up-to-date datasets with manually corrected segmentations & detailed social background info

Data

The dataset incorporates spoken SSE from a variety of sources. Core criteria of the dataset include:

- up-to-dateness (only 21st century speech)
- authentic language use from real life contexts
- only speakers with sociolinguistic background info
- countrywide representativeness in terms of speakers‘ age, gender and regional background
- differentiation between scripted (political speeches, broadcasts etc.) & unscripted language (private conversations, discussions etc.)

The dataset includes speech from the following corpora:

ICE Scotland

1 million word corpus of SSE with manually corrected phonemic transcriptions (Schützler, Gut u. Fuchs 2017). Special thanks to Prof Ulrike Gut and Zeyu Li at the University of Münster.

SCOTS Corpus

Spoken component of the Scottish Corpus of Texts & Speech (SCOTS)[1]. Special thanks to Prof Jane Stuart-Smith, Dr Rachel Macdonald, Brian Aitken & the GULP at the University of Glasgow.

SCOSYA

Parts of the Scots Syntax Atlas (SCOSYA) (Smith et al. 2019). Special thanks to Prof Jennifer Smith and Frankie MacLeod at the University of Glasgow.

Recordings Alba

Own data collection of private conversations recorded in different local libaries in Scotland in February 2020. The recordings include detailed sociolinguistic background information.

Home Recordings

Own data collection from online sources (radio broadcasts, podcasts etc.). The recordings include detailed social background information.

Pol&Lang Corpus

Own data collection of another research project investigating the influence of political identitiy on pronunciation. Special thanks to the Scottish Parliament.

Method

The transcription work is mainly carried out via ELAN and LaBB-CAT and it is partially supported by different ASR systems. The broad transcriptions are then converted into textgrid formats and forced aligned using WebMAUS (Kisler, Reichel & Schiel 2017) and HTK. Phonemic annotations are manually corrected and further intralinguistic annotations are added semi-automatically (utterance & intonation phrase structure, utterance position, prominece, speaker mean speech rate, local speech rate, word type, word frequency, word stress, syllable number, syllable structure, SVLR context classification, VE context classification, segment number, morpheme boundaries). The approach takes into account the pause definition of 150 milliseconds by Tsao and Weismer (1997) and analyses the following vowels in different contexts while accounting for suprasegmental features and the problem of relative segmentability (Lehiste and Peterson 1960: 694):

/ i / / e / / ɔ / / o / / ʉ / / a / / ʌi / / ʌu / / ɔɪ /

The project makes use of the following software, tools and coding languages:

Transcription software

ELAN | Praat | LaBB-CAT

Data management

sciebo

Forced Alignment

WebMAUS | HTK

Coding languages & data wrangling

Praat | R | Python

The vocalic interval measurement is based on standard phonetic criteria (Lehiste & Peterson 1960) (Start = beginning of stable F1 | Close: End of stable F2) (Gut 2009: 70). The data is then sorted for the different vowels. Function words and hits with the target vowel in unstressed syllable position are sorted out while extralinguistic background information (regional background, age, gender etc.) are added.

Statistical analyses are conducted in R applying mixed effects modelling as well as random forrests (Tagliamonte & Baayen 2012) to control the variation of vowel duration depending on extralinguistic as well as intralinguistic factors. *Word* and *Speaker* are treated as random factors. To avoid the influence of collinearity, each vowel set will be modelled independently for SVLR and VE contexts effects. Post-hoc tests (TukeyHSD) are implemented as well in the analyses.

References

Chen, M. (1970). “Vowel length variation as a function of the voicing of the consonant environment.” *Phonetica* 22,129-159.

Chevalier, F. (2019). “On sound change and gender: the case of vowel length in Scottish English.” *Anglophonia* 27 [Online].

ELAN (Version 6.0) [Software]. (2020). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from https://archive.mpi.nl/tla/elan

Gut, U. (2009). *Non-native speech. A Corpus-based Analysis of Phonological and Phonetic Properties of L2 English and German.* Frankfurt: Peter Lang.

HTK (Version 3.4.1) [Software]. (2009). University of Cambridge. Retrieved from http://htk.eng.cam.ac.uk

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

LaBB-CAT (Version 20161003.1931) [Software]. (2016). NZILBB, University of Canterbury, New Zealand. Retrieved from http://labbcat.sourceforge.net/

Lehiste, I. & Peterson G.E.(1960). “Duration of syllable nuclei in English.” *Journal of the Acoustical Society of America,* 32(6), 693-703.

McClure, J. D. (1977). “Vowel duration in a Scottish accent.” *Journal of the International Phonetic Association*, 7 (1), 10-16.

PRAAT (Version 6.1.08) [Software]. (2019). Boersma P. and D. Weenink, University of Amsterdam. Retrieved from http://www.praat.org/

Rathcke, T. V. & Stuart-Smith, J. H. (2016). On the Tail of the Scottish Vowel Length Rule in Glasgow. *Language and Speech* 59 (3), 404-430.

Schützler, O., Gut, U. & Fuchs, R. (2017). New perspectives on Scottish Standard English: Introducing the Scottish component of the International Corpus of English. In: S. Hancil & J. C. Beal (Eds.), *Perspectives on Northern Englishes*. Berlin: de Gruyter Publishing.

Stuart-Smith, J. H., Macdonald, R. & the SPADE consortium (2019). “The fate of the SVLR in contemporary Scottish English.” [Online].

Tagliamonte, S. A. & R. H. Baayen (2012). “Models, forests, and trees of York English: Was/were variation as a case study for statistical practice.” *Language Variation and Change*, 24, 135-178.

Tsao, Y.-C. & Weismer, G. (1997). ”Interspeaker variation in habitual speaking rate: Evidence for a neuromuscular component.” *Journal of Speech, Language, and Hearing Research*, 40, 858–866.

Smith, J., Adger, D., Aitken, B., Heycock, C., Jamieson, E. & Thoms, G. (2019). *The Scots Syntax Atlas*. University of Glasgow. Retrieved from https://scotssyntaxatlas.ac.uk

Andreas Weilinghoff, M.Ed.
Research Assistant & PhD student

TU Dortmund University, Germany
Chair of English Linguistics (Prof. Dr. Sarah Buschfeld)
Emil-Figge-Str. 50 | 44227 Dortmund | Room 3.313

Version:
12/2020