becoming lyrics: how word prosody and musical meter negotiate the rhythmic terms of prominence

by

sally ransom, B.A., Linguistics

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF ARTS

THE UNIVERSITY OF TEXAS AT AUSTIN

December 2022

becoming lyrics: how word prosody and musical meter negotiate the rhythmic terms of prominence

APPROVED BY

SUPERVISING COMMITTEE:

Scott Myers, Supervisor

Katrin Erk, Supervisor

I think that when linguists discuss and dispute sound length and stress among themselves they would definitely benefit from inviting ethno-musicologists to join them. They could discuss the issues together, and not only based on written records but also sung records.

That what is written is fiction. Only that what is sung is truth.

Tormis (2007)

becoming lyrics: how word prosody and musical meter negotiate the rhythmic terms of prominence

sally ransom, M.A.
The University of Texas at Austin, 2022

Supervisors: Scott Myers
Katrin Erk

This paper analyses the acoustic realizations of sung vowels in Estonian *regilaul* lyrical folksong performances to examine the effects of musical metrical structure on the word-prosodic level of the language. Estonian is of particular interest as it has a fixed, predictable word stress pattern and three syllable weights. In light of these features, I compare documented acoustic correlates of both stress and syllable weight, and the effects of them falling "on" or "off" the beat within the trochaic tetrametric pattern of *regilaul*. Vowel space and duration of syllable nuclei is compared in stressed and unstressed syllables falling on and off the beat in song, and all three quantities in word-prosodic stressed position examined on and off the beat in the song.

Table of Contents

List of Tables

List of Figures

Chapter 1

Results

1.1 Lexical and Metrical Prominence Patterns

1.1.1 Distributions

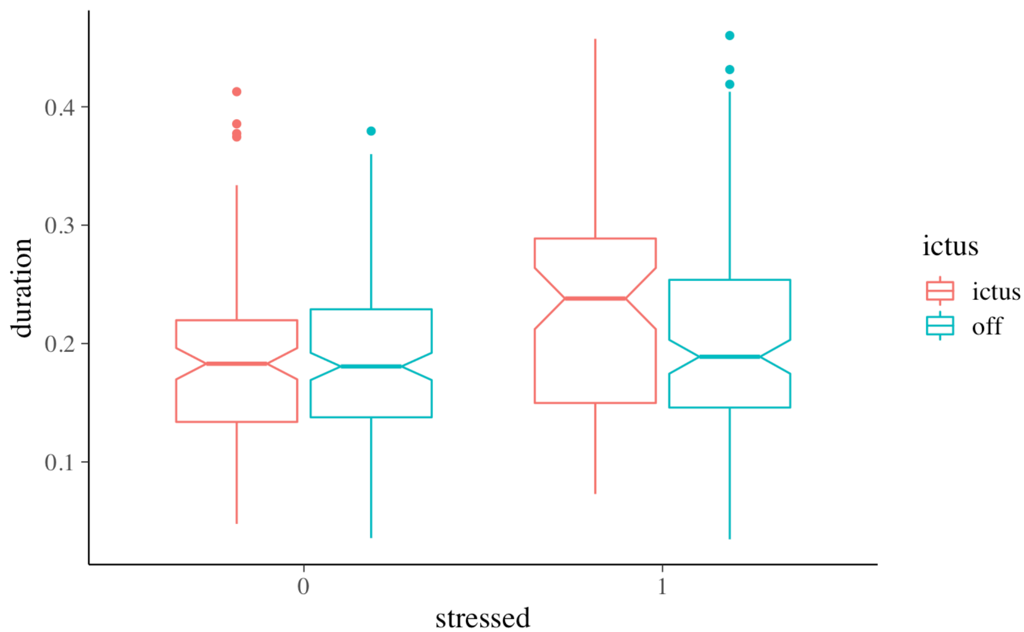

Figure 1.1: default

Before analyzing the measurements, I look at the overall occurrence patterns of conflicting and concordant lexical and lyrical metrical positions. That is, do stressed syllables have a requirement or tendency to fall "on the beat" in the song, or are they freely distributed? Likewise, do unstressed syllables tend to align with unaccented notes "off the beat?"

	un	stressed
on	107	129
off	164	210

Table 1.1: counts of conflict(*) and concord strong-weak combinations

The contingency table in **??** illustrates the counts of stressed and unstressed Estonian syllables as they fall on (ictus) and off the musical beats in the song corpus. On the beat, in ictus position, both stressed and unstressed syllables occur, and Chi-squared test finds no significant difference. Thus, both stressed and unstressed syllables may occupy the ictus "on beat" note positions in a measure. A similar pattern is seen in note positions that are off the beat, with both stressed an unstressed syllables given equal opportunity to fall in note positions that are off the beat.

This suggests that word-level and song-level rhythmic prominence are able to act independently of each other: that is, one cannot predict from beat position in the song the stressed status of the syllable, or vice a versa. a chi-squared test on the contingency table using the scipy library in python (**??**)[1]

In 1.2, we see that there is no significant ($p = 0.78$) tendency for stress to align with ictus or not, and the expected values are very close to the actual value counts in 1.3.

[1](**?**) can be found at `https://github.com/sally-ran-some` (**?**)

Table 1.2: chi-squared of stress-ictus contingency table

chi-stat	0.077
p-value	0.78

htb

Table 1.3: expected values predicted from chi-squared

	un	stressed
on	166	207
off	104	131

1.1.2 Vowel Duration: song & word stress

To examine the overall effects of stress and ictus on vowel duration, I exclude Q3 syllables, as they always fall in stressed position in native Estonian words.[2]

Using the (Goodrich et al., 2020) Stan extension package for R, I first construct a design-based Bayesian heirarchical regression model following **?**.

$$duration\ stress + ictus + stress * ictus + (1|singer) \tag{1.1}$$

However, upon looking at the posterior distributions, it was clear there was a pattern not being explained by the model. The random intercept for singer had been chosen based on the assumption that singers would vary from each other. However, as you can see in **??**, there are distinct groupings

[2]Vowel durations for Q3 falling on and off the beat are included in appendix C

that may instead correspond to each song. The singer as random intercept notion had been based on phonetics research in natural speech, however, in this case the song was the best random intercept to include in the model for vowel duration. Performers tend to be consistent compared with each other in songs, while songs with differing speeds (beats per minute) are better predictors for vowel duration. Thus, the (ad hoc) model I will focus on for this analysis is as follows:

$$duration\ stress + ictus + stress * ictus + (1|song) \tag{1.2}$$

In 1.4 is the summary of posterior statistics.

Parameter	Rhat	n_eff	mean	sd	se_mean	2.5%
Intercept	1.004	963	0.190	0.024	0.001	0.142
ictusoff	1.000	3073	0.004	0.010	0.000	-0.014
stressed1	1.000	2791	0.041	0.012	0.000	0.016
quantity2	1.000	3076	-0.038	0.012	0.000	-0.061
ictusoff:stressed1	1.000	3494	-0.018	0.013	0.000	-0.042
stressed1:quantity2	1.000	3166	0.029	0.013	0.000	0.005
ictusoff:quantity2	1.000	2818	-0.002	0.013	0.000	-0.027
b[(Intercept) song:9]	1.003	1022	0.088	0.025	0.001	0.039
b[(Intercept) song:18]	1.004	917	-0.012	0.024	0.001	-0.057
b[(Intercept) song:41]	1.003	982	0.031	0.025	0.001	-0.018
b[(Intercept) song:55]	1.004	1085	-0.027	0.025	0.001	-0.077
b[(Intercept) song:65]	1.005	909	0.067	0.024	0.001	0.020
b[(Intercept) song:69]	1.002	1303	-0.044	0.029	0.001	-0.101
b[(Intercept) song:77]	1.003	1111	-0.104	0.026	0.001	-0.156
b[(Intercept) song:92]	1.003	1086	0.027	0.026	0.001	-0.026
b[(Intercept) song:94]	1.004	896	-0.036	0.025	0.001	-0.084
sigma	0.999	4020	0.065	0.002	0.000	0.061
Sigma[song:(Intercept),(Intercept)]	1.000	1375	0.005	0.003	0.000	0.002

mean_PPD	1.000	4102	0.201	0.004	0.000	0.192
log-posterior	1.002	1049	632.780	3.684	0.114	624.523

Table 1.4: Posterior Summary Statistics, Q1 and Q2 sylla

Bibliography

Boersna, P., & Weenink, D. (2022). Praat: Doing Phonetics by Computer.

Cousins, M., & Hepworth-Sawyer, R. (2014). Logic pro X.

De Jong, K. J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *The journal of the acoustical society of America*, *97*(1), 491–504.

Duddington, J., Avison, M., Dunn, R., & Vitolins, V. (1995). eSpeak: Speech Synthesizer.

Goodrich, B., Gabry, J., Ali, I., & Brilleman, S. (2020). Rstanarm: Bayesian applied regression modeling via Stan. R package version 2.21.1.

Jadoul, Y., Thompson, B., & de Boer, B. (2018). Introducing Parselmouth: A Python interface to Praat. *Journal of Phonetics*, *71*, 1–15.

Laur, S., Orasmaa, S., Särg, D., & Tammo, P. (2020). EstNLTK 1.6: Remastered estonian NLP pipeline. In *Proceedings of the 12th Language Resources and Evaluation Conference*, (pp. 7154–7162). Marseille, France: European Language Resources Association.

Lehiste, I. (1992). The Phonetics of Metrics. *Empirical Studies of the Arts*, *10*(2), 95–120.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle, & A. Marchal (Eds.) *Speech Production and Speech Modelling*, (pp. 403–439). Dordrecht: Springer Netherlands.

Palmer, C., & Kelly, M. H. (1992). Linguistic Prosody and Musical Meter in Song. *Journal of Memory and Language*, *31*(4), 525–542.

Robertson, A., & Plumbley, M. (2007). B-Keeper: A beat-tracker for live performance. In *Proceedings of the 7th International Conference on New Interfaces for Musical Expression - NIME '07*, (p. 234). New York, New York: ACM Press.

Ross, J. (1989). A study of timing in an Estonian runic song. *The Journal of the Acoustical Society of America*, *86*(5), 1671–1677.

Ross, J. (1990). Formant frequencies in estonian folk singing.

Ross, J., & Lehiste, I. (1994). Lost Prosodic Oppositions: A Study of Contrastive Duration in Estonian Funeral Laments. *Language and Speech*, *37*(4), 407–424.

Ross, J., & Lehiste, I. (1996). Trade-off between quantity and stress in Estonian folksong performance? *Folklore: Electronic Journal of Folklore*, *02*, 116–123.

Ross, J., & Lehiste, I. (1998). Timing in Estonian Folk Songs as Interaction between Speech Prosody, Meter, and Musical Rhythm. *Music Perception*, *15*(4), 319–333.

Särg, T. (2005). The Early History of Estonian Speech Prosody Studies. *Linguistica Uralica*, *41*(2), 81.

Smiljanić, R., & Bradlow, A. R. (2005). Production and perception of clear speech in Croatian and English. *The Journal of the Acoustical Society of America*, *118*(3 Pt 1), 1677–1688.

Tampere, H. (2016). Anthology of Estonian Traditional Music.

Tormis, V. (2007). Some_problems_with_that_regilaul.pdf. In *RING*.

Van Rossum, G., & Drake Jr, F. L. (1995). *Python Reference Manual*. Centrum voor Wiskunde en Informatica Amsterdam.

Vita

Sarah Marie Ransom-Laud was born in Sarasota, Florida on 11 July 1990. Her early career was as a songwriter and recording artist, releasing several full-length albums combining digital and analog recording methods over the last decade.

In 2018 she received the Bachelor of Arts degree in Linguistics from the University of Wisconsin, Milwaukee. Upon graduation, she moved to Austin with her partner, Kavi, and taught English as a Second Language (ESL) to adults until her admission into the PhD program in the department of Linguistics at the University of Texas at Austin, where she matriculated in Fall 2019.

Permanent address: 305 E. 23rd Street STOP B5100
Austin, Texas 78712

This thesis was typeset with LaTeX[†] by the author.

[†] LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.