Stress in a Pitch Accent Language:
An Acoustic Analysis of Lithuanian Grave Syllables
Christopher Oakden
Rutgers University
Version: 4/2/18

1. Introduction

The goal of this paper is to answer the question: "What is stress in a tone/pitch accent language?" Metrical structure and tone/pitch accent have been argued to co-exist in many languages (e.g. Downing 1990, de Lacy 2002, and others). However, the question of how metrical heads are phonetically realized – i.e. perceptibly stressed – in such languages has not been extensively addressed. In many languages with stress, stress is realized by perturbing fundamental frequency (F0 – i.e. pitch) (e.g. Garellek & White 2012 on Tongan; see also Gordon 2014 and Gordon & Roettinger 2017). However, if a language uses F0 to convey pitch accent/tonal distinctions, which acoustic properties does stress utilize?

Lithuanian, a Balto-Slavic language, has been claimed to exhibit tone and stress, as well as both tone-driven stress and metrically-conditioned tone neutralization (Blevins 1993, de Lacy 2002). The basic generalization is that stress is attracted to the leftmost high-toned syllable (Leskien 1919, Senn 1966, Kiparsky & Halle 1977). Additionally, tonal contrasts are neutralized to ensure word-level high-tone culminativity – i.e. every word must surface with one and only one high tone. The following examples are from de Lacy (2002:21):

(1) *Tone driven stress and metrically-conditioned neutralization in Lithuanian*
a. /prà-dúrtè/ → [pràˈdúrtè] 'to pierce through'
b. /prà-nèʃù/ → [ˈprá nèʃù] 'I announce'

In (1a), word stress falls on the leftmost high-toned syllable, while in (1b), a word that has only low tones underlyingly surfaces with a high tone on its leftmost syllable, which is then stressed.

As stress and tone have been claimed to both co-exist and be clearly perceptible in Lithuanian, it provides an ideal opportunity to determine the acoustic correlates of stress in a tone/pitch-accent language.

This paper presents the results of a production experiment which analyzes the acoustic properties of so-called 'grave' syllables in Lithuanian nominals (see §2), showing that duration, intensity, and vowel quality (F1/F2) are not reliable acoustic correlates of stress, contra previous accounts. An increase in F0 does correlate with stressed position within a word as predicted in traditional descriptions of Lithuanian stress, but this increase is demonstrably categorical in nature, akin to the realization of a [H] tone target. In short, there is no clear perceptible realization of metrical structure—i.e. stress—on Lithuanian grave syllables. This finding contradicts previous descriptions of tone-driven stress in the language, and problematizes its status as a stress language.

This paper is organized as follows. Section 2 introduces relevant background information on Lithuanian and provides motivation for the acoustic study. The proposed acoustic experiment is outlined in Section 3. Methodology and results are presented in Section 4, while Section 5 discusses the results and their implications for phonological theory. Section 6 concludes.

2. Background

This section outlines the vowel inventory, pitch accent system, and stress assignment patterns of Lithuanian (particularly in the inflection of nominals), and discusses previous accounts of the phonetic realization of stress and pitch accent, both impressionistic and experimental. Theoretical motivation for the current study is also provided.

2.1 Phonemic Vowels and Syllable Structure

The phonemic vowel inventory of Lithuanian consists of monophthongs and diphthongs, with the former exhibiting a length contrast, as shown below (adapted from Ambrazas 1997:24):

Short Vowels

ɪ		ʊ
	ɛ	
	ɑ	

Long Vowels

i:			u:
	e:	o:	
	æ:		
	ɑ:		

Fig. 1: Lithuanian Monophthong Inventory; Short and Long Vowels

Sources differ on the nature of the diphthong inventory (Robinson 1968, Kenstowicz 1972, Ambrazas 1997), but agree on the distribution of Lithuanian monophthong vowels. Note the asymmetry in the short-long distinction; neither [æ:] nor [o:] have short equivalents, though Kenstowicz (1972:3) notes that the latter occurs in loanwords.

Lithuanian allows a variety of surface syllabic structures. Onsets are optional, and syllables containing both complex onsets and complex codas of up to three segments in length are also permitted (Ambrazas 1997). Lithuanian syllable structure can be schematized as (C)(C)(C)V(:/V)(C)(C)(C). The figure below outlines frequencies of certain syllable types, where ‘V^{w}’ indicates a diphthong (Girdenis 2003:324).

Syllable type	Syllable count	Relative frequency (%)	Syllable type	Syllable count	Relative frequency (%)
CV	22,813	54.66	V^{w}C	60	0.144
CVC	7,346	17.60	CCV^{w}C	59	0.141
CV^{w}	2,682	6.43	CCVCC	46	0.110
CCV	2,661	6.38	**CCCV**	34	0.081
VC	2,026	4.86	CCCVC	28	0.067
V	1,434	3.44	CCCV^{w}	20	0.048
CCVC	780	1.87	CV^{w}CC	6	0.014
CV^{w}C	573	1.37	CVCCC	4	0.010
CCV^{w}	494	1.18	CCCVCC	2	0.005
CVCC	405	0.97	CCV^{w}CC	2	0.005
VCC	138	0.33	CCVCCC	1	0.002
V^{w}	120	0.29			

Fig. 2: Frequency of syllable type in Lithuanian (Girdenis 2003:324)

2.2 Pitch Accent/Tone in Lithuanian

Lithuanian has been described a pitch accent/tonal language in the sense of Hyman (2006:229): i.e. a language "in which an indication of pitch enters into the lexical realization of at least some morphemes" (Mathiassen 1996, Ambrazas 1997, Girdenis 2003, Young 1991, Kenstowicz 1972, Dudas and O'Bryan 1972, Blevins 1993, Dogil and Möhler 1998). Examples of tonal minimal pairs in the language are presented in (2).

(2) *Tonal minimal pairs (Ambrazas 1997:55)*

Falling Contour Tone		Rising Contour Tone
a. [ʃâuk] 'shoot!'	vs.	[ʃǎuk] 'shout!'
b. [gînti] 'to defend'	vs.	[gǐnti] 'to drive (off)'
c. [tî:ræ] '(he) explored'	vs.	[tǐ:ræ] 'mush'

There is disagreement, however, regarding the pitch accent inventory of Lithuanian. Generative analyses of the language argue for a system with three distinct pitch accent types. The *acute* (sharp falling contour) and *circumflex* (gradual rising contour) pitch accents are realized on syllables with long monophthongs, diphthongs, and syllables with sonorant codas ('long' syllables), while *grave* (falling contour) pitch accents are realized on short monophthongs and syllables with obstruent codas ('short' syllables) (Kenstowicz 1972, Dudas and O'Bryan 1972, Blevins 1993, Dogil and Möhler 1998). This three-way distinction is summarized below, where 'R' indicates sonorant segments and 'O' obstruents.

	CV:	CVV	CVR	CVO	CV
Acute (falling contour – HL)	✓	✓	✓		
Circumflex (Rising contour – LH)	✓	✓	✓		
Grave (falling contour – HL)				✓	✓

Fig. 3: Lithuanian Pitch Accents

Descriptions of the language from the Structuralist tradition (Mathiassen 1996, Ambrazas 1997, Girdenis 2003, Young 1991) maintain that since grave syllables are not contrastive, they do not constitute proper pitch accents. Therefore, the pitch accent system consists of acute and circumflex only, while *grave* syllables merely constitute the realization of stress (Ambrazas 1997:13).

An analysis of Lithuanian nominals by Blevins (1993) analyzes the language's pitch accents in terms of tone. Long vowels, diphthongs and vowel+sonorant rimes are considered bimoraic in the account, with the acute/circumflex distinction formalized as a [H] tone associated to the first or second mora, respectively. Grave syllables are monomoraic, with a single [H] tone associated to the mora, while unaccented root syllables are toneless (Blevins 1993:243):

(3) *Blevins' (1993) tonal analysis of nominal roots*

a. unaccented:	b. grave:	c. acute	d. circumflex
$[\mu]\sigma\ [\mu\ \mu]\sigma$	$[\mu]\sigma$	$[\mu\ \mu]\sigma$	$[\mu\ \mu]\sigma$
	|	|	|
	H	H	H

Despite disagreements regarding the nature of the inventory, the culminativity of pitch accent (i.e. one and only one pitch accent per word) is consistent across previous accounts.

2.3 Stress Assignment in Nominals and de Saussure's Law

All Lithuanian nominals surface with one H tone and one stress – both on the same syllable. Previous analyses have argued that the leftmost underlying H tone survives and is stressed, a pattern known as the 'Basic Accentuation Principle' (Kiparsky & Halle 1977). As shown below, word stress on nominals can be realized on either the bound root or the suffix.

(4) *Basic Accentuation Principle in Lithuanian*

a. /víːr+ɑ́ms/ → [ˈvíː.rɑms] 'man DatPl'
b. /kɛlm+ɑ́ms/ → [kɛl.ˈmɑ́ms] 'tree-stump DatPl'

Stress assignment in Lithuanian is more complex than the BAP suggests. Traditional descriptions of the language conceptualize the variability in (4) as *stress mobility*, and understand its distribution in terms of *strong* and *weak* roots and suffixes, as well as stress-attracting vs. non-attracting suffixes. The present study does not attempt to offer an analysis of this distribution, but does make use of stress mobility effects in the experiment design (see Appendix 2 for discussion).

2.4 Acoustic Correlates of Stress and Phonological Evidence for Stress

Descriptions of the phonetic realization of stress in Lithuanian are largely impressionistic. Ambrazas (1996:54) describes stressed syllables as characterized by "[i]ncreases in loudness and pitch" with "longer duration and more precise timbre." Girdenis (2003) also asserts higher pitch is a correlate of stress, as well as greater articulatory effort. Unstressed vowels, on the other hand, are reduced both qualitatively and quantitatively (Kaukėnienė 2004). Additionally, tonal contrast is said to be neutralized in unstressed position (Ambrazas 1997, Young 1991, Girdenis 2003).

What is problematic about these descriptions is that tonal contrast is neutralized in unstressed position, so it is not possible to observe pitch accented syllables which are unstressed. Complicating the issue further is the interchangeable use of *stress/accent* and *pitch-accent/accent* across sources. Disentangling stress from tone in Lithuanian terminologically, phonologically, and acoustically, therefore, is non-trivial.

There is little evidence from phonological alternations for metrical structure in Lithuanian. Phonological lengthening of non-high short vowels /e/ and /a/ when in non-final stressed position has been reported (Kenstowicz 1972, Mathiassen 1996, Ambrazas 1997). For example, in (5), 'ice' has the underlying form /led/, and 'wheel' /rat/.

(5) *Non-high lengthening (Kenstowicz 1972:4)*

a. [le.ˈde]	'ice LocSg'	c. [ra.ˈte]	'wheel LocSg'
b. [ˈleː.das]	'ice NomSg'	d. [ˈraː.tas]	'wheel NomSg'

However, Kaukėnienė (2012) argues that this lengthening effect is imperceptible, even though there was a statistically significant distinction in the duration of lengthened and non-lengthened vowels.

An open issue regarding this lengthening process is its effect on the realization of pitch accent. According to Kenstowicz's (1972:4) notation, the roots in (5b) and (5a) are realized with a circumflex accent. This annotation may suggest that short (and therefore *grave*) syllables surface as circumflex post-lengthening uniformly, or it could be that these vowels are long underlyingly, and are shortened in unstressed position. Examining these effects is not within the scope of the current study, and will be left for future work (see the discussion of target selection in §3.1)

2.5 Previous Phonetic Studies

Acoustic analyses of Lithuanian pitch accent using modern instrumentation go back to Robinson (1968), whose analysis of diphthongs in the language is among the first to cast doubt on the generalizations described in impressionistic accounts; i.e. that pitch enters into the phonetic distinction between acute and circumflex syllables. The results of the acoustic analysis indicate that fundamental frequency (i.e. F0, pitch) does not reliably distinguish acute and circumflex diphthongs. Rather, the study found that duration and vowel quality are more robust indicators of the acute/circumflex distinction. Subsequent work has produced similar results for monophthongs (Dogil and Möhler 1998, Dogil and Williams 1999). A more recent study by Kudirka (2005) examines a number of acoustic features of stressed acute and circumflex monophthongs, including F0, intensity, duration, and formant structure. Only duration yielded significant effects. These effects have also been reproduced in work by Kelly (2013) which echoes the conclusions from earlier studies that pitch does not provide a robust cue for the distinction between acute and circumflex pitch accents, contra traditional descriptions.

The studies mentioned above focus primarily on the two-way pitch accent distinction between acute and circumflex, neglecting grave syllables. Additionally, data are limited to stressed syllables, precluding any comparison with unstressed syllables. Other phonetic work (Kaukėnienė 2003, 2012; Grigorjevs & Jaroslavienė 2015) that does examine qualitative and quantitative differences between stressed and unstressed syllables neglects the effect of tone. Additionally, these studies largely compare stressed syllables with pre- and post-tonic unstressed syllables within a word, apart from being mostly comparative in nature.

In short, a comprehensive understanding of the acoustic realization of stress in Lithuanian (and how it relates to tone) requires reference to grave syllables, and must include an acoustic profile of stress that controls for the effect of word position. In other words, the comparison should be made between stressed and unstressed syllables in the same position. Such is the goal of this study.

2.6 Theoretical Interest

The tone-bearing status of grave syllables is also of interest as it has implications for prosodic theories which attempt to formalize the interaction of tone and stress. de Lacy (2002)'s theory of tone-driven stress, for example, cites Lithuanian as a case of the convergence of canonical tone-driven stress and metrically-conditioned neutralization in a single language. Formally, this interaction is mediated through the relative ranking of constraints which govern the realization of tone on metrical (non)-heads, stress placement, and faithfulness to underlying tonal specification information. The tableau in (6) illustrates the general argument for (1a) (de Lacy 2002: 21).

(6) *Tableau for /prà-dúrtè/*

/prà-dúrtè/	*HD/L	IDENT(T)	ALLFTL
a. (ˈpràdúr)tè	*!		
b. (ˈprádúr)tè		*!	
☞ c. prà(ˈdúrtè)			*

Here, the high ranking of the markedness constraint *HD/L ("Incur a violation for a low-toned foot head") ensures that an optimal candidate will not realize stress on a L-toned syllable, but rather it will seek out the leftmost (ALLFTL) lexically-H-toned (IDENT(T)) syllable. Metrically-conditioned neutralization is achieved through the ranking *HD/L » IDENT(T) included in (6). That is, stress must not be realized on a L-toned syllable, even if violation of tonal faithfulness is necessary. This characterization, however, hinges on the assumptions implicit in Blevins (1993)'s tonal account of the language, namely that grave syllables are tone-bearing. Whether or not these syllables can in fact bear tone is of theoretical interest.

Furthermore, a system like Lithuanian is puzzling from an acoustic perspective, apart from the general question of how tonal pitch/accent languages realize stress. When tone-driven stress and metrically-conditioned tonal neutralization converge, tone and stress only surface together. How, then, is it possible to disambiguate the two phonetically?

The current study incorporates grave syllables into consideration as potentially tone-bearing units in the Lithuanian pitch accent system, and offers an analysis of their acoustic properties as they relate to tone and stress, namely duration, intensity, vowel quality, and fundamental frequency. Specifically, the data for this study come from recordings of Lithuanian nominals collected in a laboratory experiment, discussed in the next section.

3. Experiment

This section outlines details of the experiment from which acoustic data for the study were gathered. The primary purpose of the experiment is to establish an acoustic profile for stressed and unstressed grave syllables on nominal forms.

3.1 Targets

Multiple controls were implemented during target selection. The primary control limited targets to a specific syllabic structure and default stress placement: trisyllabic words of the shape [CVCVCV(C)] and tetrasyllabic words of the shape [CVCVCVCV(C)], with default stress on the penult. Potential targets with complex onsets in any position (pre-tonic, tonic, post-tonic) were excluded. Additionally, target selection minimized root syllables containing codas [CVC] or lacking onsets [V]. Due to lexical gaps, however, two target roots were selected which contained these suboptimal structures: *apetit-* 'apetite' and *mygtuk-* 'button'.

Grave syllables containing the vowels [u] and [i] were given preference in selection (n = 32 and 28, respectively), as these vowels are not reported to undergo phonological lengthening in stressed position (see discussion of "non-high lengthening" in the previous section). Nominal forms with a stressed [e] were also included for the purpose of comparison (n = 4), as well as to determine the phonetic reality of non-high lengthening for that vowel.

Segments surrounding the vowel on the stressed syllable were also controlled. To aid in segmentation and offer reliable duration measures, a portion of the selected nominal forms contained a tonic vowel flanked by stops (n = 36). A smaller proportion of targets contained tonic vowels flanked by sonorants [l], [m] and [n] (n = 24); vowels on these targets were reserved for extracting fundamental frequency data. The one exception to this control was the root *palikim-* 'inheritance'.

Targets conforming to the above restrictions were further limited to roots of the same declension class (Ia) and accent class (2) to ensure uniformity of accentual properties across stimuli (see Ambrazas 1997 and Appendix 2 for more discussion). Four suffixes from this declension class were selected to combine with nominal roots based on predicted stress placement: two suffixes (NomSg [-as] and GenSg [-o]) for which stress surfaces on the root, and two suffixes (InstSg [-u] and AccPl [-us]) for which stress is predicted to surface on the suffix. A sample of targets and predicted stress patterns is provided below.

NomSg	GenSg	InstSg	AccPl
katúkas	katúko	katukú	katukús
tupíkas	tupíko	tupikú	tupikús
pakétas	pakéto	paketú	paketús
rėmímas	rėmímo	rėmimú	rėmimús
tolúmas	tolúmo	tolumú	tolumús

Fig. 4: Sample of Targets and Predicted Stress Patterns

This paradigm allows direct comparison of stressed and unstressed (or destressed) vowels on grave syllables.

Ten nominal roots with obstruent-flanked tonic syllable vowels were selected, as well as six nominal roots with sonorant-flanked tonic syllable vowels. Each root was inflected for the four cases described above, totaling 64 target words (see Appendix 1 for full list of roots and suffixes). A native speaker consultant examined the target list and confirmed that the words were appropriate for colloquial conversation, and that none of the forms were archaic, loanwords or specialized vocabulary.

3.2 Scenario/Frame Sentences

Target words were placed within two frame sentences, designed to illicit vernacular speech production. These sentences describe an informal conversation scenario between two people in a bar or restaurant. The subject notices that someone has scrawled a word on the wall next to them, and observes that the word is written in a strange manner (perhaps because the writer of the word was in an advanced stage of inebriation at the time of writing). Subjects read the target word in both frame sentences:

(7) *Target frame sentences*

a.	Žiurek,	čia	parašytas	žodis	_______
	ʒurɛk	t͡ʃa	p^{h}aɾaʃiːtas	ʒodis	
	look	there	write.pass	word	

‘Look, the word ______ is written over there.’

b.	Žodis	_____	paraštas	keistai
	ʒodis		p^{h}aɾaʃiːtas	k^{h}eistai
	word		write.pass	strange

‘The word _____ is written strangely.’

The rationale behind selecting a written word in the ambient environment is that the constructed frame sentences equally accommodate any word regardless of case marking.

Note that the position of the target word in the first frame sentence is sentence-final. This was unavoidable given the information structure of Lithuanian; new information (‘rheme’ in the functional sense) is introduced either sentence-finally or sentence-medially, but the latter placement triggers a sentence-level falling intonation pattern (Ambrazas 1997:690-692, Anderson p.c.). To mitigate any confounding effects of this stress pattern, the choice was made to place the target at the end of the sentence in (7a).

3.3 Participants

Participants in the study were adult native speakers of Lithuanian (as determined by a demographic questionnaire) recruited from the central New Jersey area. Data from four speakers were collected, one of which is used in the current study. This speaker (female, early 30s) was born and spent her childhood and adolescent years in Lithuania, moving to the United States after reaching adulthood. Though residing in the United States, her primary home language continues to be Lithuanian, which she speaks on a daily basis with her husband (an L1 Lithuanian speaker) and children (heritage speakers), suggesting a low possibility of L1 attrition. The subject was naïve to the purpose of the study, was not trained in linguistics, and had no history of speech impairment.

3.4 Procedure

Recording took place inside a sound-attenuated booth at the Phonology and Field Research Laboratory at Rutgers University. The subject was fitted with an AKG C420 head-worn microphone with behind-the-neck headband to maintain a constant distance from the mouth. The recording was made using *GoldWave* v6.10 software at 44.1k Hz sampling rate and 16-bit quantizing resolution in mono. A laptop screen presented a series of images. Each image contained a scene with characters in a bar/restaurant, with one pointing to a strangely-written word. The same word in standard orthography appeared at the top left-hand corner of the screen.

vaikas

Fig. 5: Example prompt

The frame sentence did not appear on the screen, but was printed on a separate sheet of paper for reference. The subject read the word in both frame sentences, then pressed a key to advance to the next picture.

Three repetitions of the targets within the frame sentences were recorded. The order of targets was randomized and counterbalanced in each repetition, with filler words evenly spaced among stimuli. Five additional fillers were placed at the beginning of each repetition to accommodate any initial hyper- or mis-articulation in case the participant felt anxious about the task. All three repetitions were completed during a single session with a 10-to-15 minute break between each repetition.

The speaker produced each of the 64 target words in two frame sentences across three repetitions. The total number of tokens was 384.

3.5 Hypotheses

Traditional descriptions of stress in Lithuanian assert the presence of certain acoustic correlates of stress. Following those accounts (see §2.4), the following hypotheses for grave syllables in the current study are proposed:

(8) *Experiment hypotheses*

a. Duration: Stressed grave syllables are longer in duration than unstressed syllables
b. Intensity: Stressed syllables have greater intensity than unstressed syllables
c. F1 and F2: Unstressed syllables are centralized compared to unstressed syllables
d. F0: Stressed syllables exhibit a higher F0 than unstressed syllables

3.6 Data Processing and Statistical Analysis

After data collection concluded, sound files were annotated using TextGrids in Praat (Boersma & Weenink 2015). Boundaries were marked for all vowels within target words as well as fixed intervals in both frame sentences for the purpose of duration normalization. Left edges of target word vowels were identified by placing a boundary at the zero crossing of the first periodic, non-deformed waveform of the vowel (Francis et al. 2002). The right edges of vowels were determined by inserting a boundary at the end of the second formant of the vowel (Turk et al. 2006:7).

TextGrid intervals were marked to indicate a) the vowel, b) predicted stress pattern based on traditional accounts, c) position within a word or frame sentence, and d) number of syllables per target word containing the given interval. These factors form the basis of the fixed effects structure of statistical analysis. The labeling schema is summarized below.

(9) *TextGrid labeling schema*

a. Vowel: [a], [e], [i], [o], [u]
b. Stress: upper case indicates stressed, lower case indicates unstressed
c. Position: 1/2/3 for trisyllabic words in 1st frame sentence; 4/5/6 for trisyllabic words in 2nd frame sentence; 0/10/20/30, for quadrisyllabic words in 1st frame sentence; 00/40/50/60 for quadrisyllabic words in 2nd frame sentence
d. Examples: U2 = stressed [u], 2nd syllable of target word in 1st frame sentence (3σ)
a6 = unstressed [a], 3rd syllable of target word in 2nd frame sentence (3σ)
I50 = stressed [i], 3rd syllable of target word in 2nd frame sentence (4σ)

Duration data were extracted using boundary marks on TextGrids for each audio file. A normalization multiplier was created by dividing the raw duration of the marked fixed interval in each frame sentence by the mean of all fixed intervals. This yielded 128 target repetition- and frame sentence-specific multipliers. TextGrid intervals for vowels in each target word in a given frame sentence were multiplied by the corresponding multiplier. Intensity, F1 and F2 data were extracted using customized Praat scripts. Values were extracted from the midpoint of the vowel for each, to ensure that measurements were taken from the most stable portion of the vowel. F0 and spectral tilt measurements were extracted using the Praat script Prosody Pro (Xu 2013). To test the hypothesis that stressed vowels are higher in pitch than unstressed vowels, maximum F0 for each vowel interval was calculated.

A series of linear mixed effects models were fit to the data using the *lmer()* function in the lme4 package (Bates et al. 2015) in R (R Development Core Team 2017). For the vowels [i] and [u], five separate models were fit to the data, with acoustic measures (duration, intensity, F1, F2, max F0) as dependent variables, while predicted stress (stressed/unstressed), frame sentence, segmental environment (grave syllables flanked by obstruents or sonorants), and number of syllables per target word (3 or 4) were set as fixed effects. The random effects structure for each model specified by-word random slopes for frame sentence, segmental environment, and number of syllables per target word. For each model, a null model excluding the fixed effect of stress was also created. Corresponding models were then compared with a Likelihood Ratio Test using R's *anova*() function to test for significant differences between the likelihood of the models. To evaluate the appropriateness of LME models on this data and to ensure the reliability of model comparison, visual inspection of residual plots was performed to check model diagnostics (Schroeder et al 1986). No indication of model assumption violations—normal distribution and equal variation (homoscedasticity) of the residuals—was observed for the duration, intensity, F1 or F2 models (but see discussion in §4.5).

The following subsections present the experiment results by acoustic measure: duration, intensity, F1/F2, and F0. General discussion includes figures from the vowels [e], [i], and [u]. As mentioned in the previous section, preference was given to [i] and [u] in target selection, resulting in a low number of measurements for [e] (n = 24) by comparison ([i]: n = 168, [u]: n = 192). Therefore, statistical analysis in this paper is limited to the acoustic measures for vowels [i] and [u].

4. Results

This section presents the results of the experiment. Since preference was given to grave syllables containing the vowels [i] and [u] in target selection, discussion of results will focus primarily on these vowels (and to a lesser extent [e]).

4.1 Duration

This subsection presents (normalized) duration results from the acoustic experiment to test the hypothesis developed from impressionistic accounts that increased duration is a correlate of stress.

Figure 6 presents mean duration values for stressed and unstressed vowels, categorized by vowel. For the vowels [i] and [u], the mean duration of stressed syllables is shorter than for unstressed syllables, though the difference is slight, and overlap between values for the two categories is substantial. In the case of [e], stressed syllables are longer (mean = 67.99ms (s.d. = 10.81ms)) than unstressed syllables (mean = 54.75ms (s.d. = 6.69ms)). As noted in the previous section, results from the acoustic analysis of [e] will not undergo statistical analysis in the current study due to this potentially confounding factor. Additionally, the small sample size of grave syllables with [e] precludes determination of significant effects.

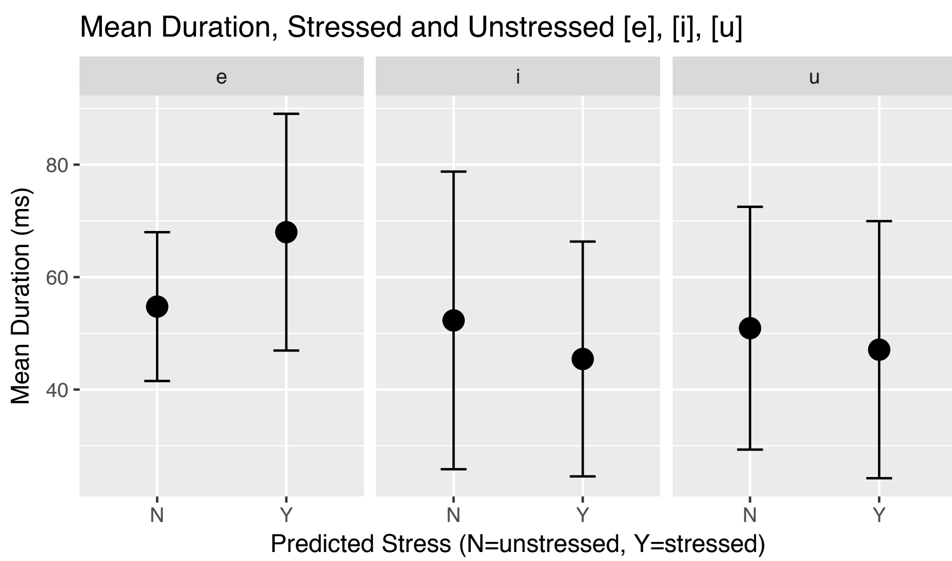

Fig. 6: Mean Duration

4.1.1 [i]

Mean duration results for [i] vowels on grave syllables in trisyllabic and tetrasyllabic words are presented in the following table, separated by predicted stress pattern (stressed 'Y' or unstressed 'N'), frame sentence, and number of syllables in target word.

Frame/#	Stressed	Unstressed
$1^{st}/3\sigma$	48.63 (7.82)	55.73 (11.96)
$2^{nd}/3\sigma$	51.00 (7.90)	59.76 (11.46)
$1^{st}/4\sigma$	39.23 (10.79)	45.72 (9.73)
$2^{nd}/4\sigma$	40.00 (10.88)	44.32 (13.44)

Fig. 7: [i] vowels in grave syllables:
Mean duration values in milliseconds (with standard deviation)

For both three- and four-syllable words and between both frame sentences, stressed [i] grave syllables exhibit mean duration shorter than that of unstressed syllables. This generalization is the opposite of the hypothesis generated from traditional accounts. The linear mixed effects model for duration on [i] vowels failed to converge, so number of syllables per word was removed from the random effects structure. Comparison of the resulting model with the null model determined that the difference between stressed and unstressed grave syllables with [i] is significant ($\chi^2(1)=8.69$, $p=0.0032$), such that stress on [i] in grave syllables decreases duration by about 7 ms ± 2.3 ms (standard error).

4.1.2 [u]

Results for the vowel [u] are similar that of [i], as illustrated in the following figure.

Frame/#	Stressed	Unstressed
$1^{st}/3\sigma$	49.42 (13.47)	51.46 (11.03)
$2^{nd}/3\sigma$	47.06 (10.45)	51.36 (11.75)
$1^{st}/4\sigma$	43.95 (8.36)	48.30 (9.81)
$2^{nd}/4\sigma$	43.35 (9.51)	50.51 (8.62)

Fig. 8: [u] vowels in grave syllables:
Mean duration values in milliseconds (with standard deviation)

Mean duration values for stressed [u] on grave syllables are lower than unstressed counterparts, though there is a substantial degree of overlap between the two considering their respective standard deviations. In spite of this fact, the difference was found to be (marginally) significant ($\chi^2(1)=5.45$, $p=0.01959$), with stress corresponding to a decrease in duration by approximately 4.16 ms ± 1.7 ms (standard error). Again, these results do not confirm the hypothesis that longer duration is an acoustic correlate of stress for Lithuanian grave syllables.

4.2 Intensity

Traditional accounts describe Lithuanian stressed syllables as louder than unstressed syllables. Therefore, higher intensity is expected on stressed syllables. Intensity measurements on vowels indicate a correlation between stress and increased mean intensity for vowels [e], [i], and [u], as shown in Figure 9. A slight increase in intensity is indicated for both [i] and [u] (approximately 1 dB), with the latter showing a slightly larger difference.

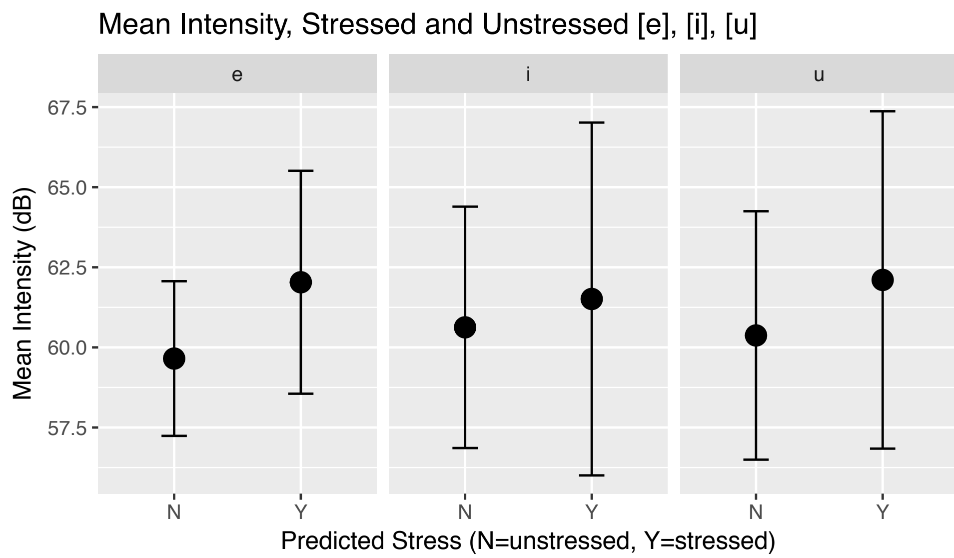

Fig. 9: Mean intensity values

4.2.1 [i]

Intensity values (in dB) for stressed and unstressed grave syllables with the vowel [i] demonstrate little variation regardless of frame sentence or number of syllables per target word.

Frame/#	Stressed	Unstressed
$1^{st}/3\sigma$	62.15 (2.14)	61.14 (1.41)
$2^{nd}/3\sigma$	61.80 (2.09)	61.06 (1.90)
$1^{st}/4\sigma$	61.17 (3.39)	60.09 (1.94)
$2^{nd}/4\sigma$	60.624 (3.42)	59.89 (2.1)

Fig. 10: [i] vowels in grave syllables:
Mean intensity values in decibels (with standard deviation)

Stressed syllables exhibit higher intensity than unstressed syllables, though this difference is seldom more than one decibel. Intensity values for both types of syllables concentrate around 60 dB. Model comparison reveals that the difference did not reach significance ($\chi^2(1)=2.796$, $p=0.09449$). The hypothesis is therefore not confirmed for syllables with [i]; stress on grave syllables does not realize as an increase in intensity.

4.2.2 [u]

Recall that, for [u], the stressed/unstressed gap in intensity appeared larger than for [i]. Mean intensity values separated by frame sentence and Indeed, stressed grave syllables with [u] vowels show an increase in intensity on unstressed equivalents by roughly 2 dB. This is uniform across frame sentence and syllable number of target word.

Frame/#	Stressed	Unstressed
$1^{st}/3\sigma$	62.17 (3.67)	60.54 (1.88)
$2^{nd}/3\sigma$	62.29 (1.71)	60.10 (2.02)
$1^{st}/4\sigma$	62.05 (1.76)	60.57 (1.83)
$2^{nd}/4\sigma$	61.41 (1.99)	60.47 (2.14)

Fig. 11: [u] vowels in grave syllables:
Mean intensity values in decibels (with standard deviation)

Unlike [i], model comparison determined that the intensity difference between stressed and unstressed grave syllables with [u] does reach significance ($\chi^2(1)=14.408$, $p=0.000147$). Stressed [u] indicates an estimated intensity increase of 1.63 dB ± 0.38 dB (standard error).

4.3 F1 and F2

Comparison of the formant structure of stressed and unstressed grave syllables tests the traditional accounts' predictions about qualitative differences as a reflex of stress. A reasonable interpretation of the impressionistic descriptions of stress as correlating with greater "articulatory effort" or "clarity" is that stressed vowels are more peripheralized than unstressed counterparts (Gordon & Roettger 2017). Under this assumption, high front vowels in stressed syllables should be realized as higher (lower F1) and more fronted (higher F2) than unstressed vowels. High back vowels should also exhibit a lower F1 (higher in the vowel space), but have a lower F2, indicating a greater degree of backness.

4.3.1 [i]

Mean F1 and F2 values for the vowel [i] are presented below. Formants 1 and 2 are mostly consistent for unstressed vowels across both frame sentences and target word length, while stressed vowels display more variation, especially in F2.

Frame/#		Stressed	Unstressed
$1^{st}/3\sigma$	F1	350.7 (39.4)	323.3 (38.1)
	F2	1660.8 (381.1)	1576.5 (204.4)
$2^{nd}/3\sigma$	F1	351.3 (47.2)	323.8 (36.5)
	F2	1762.3 (341.4)	1605.8 (233.0)
$1^{st}/4\sigma$	F1	372.4 (55.1)	356.7 (46.3)
	F2	1839.3 (188.3)	1596.8 (173.4)
$2^{nd}/4\sigma$	F1	352.6 (53.1)	339.4 (47.0)
	F2	1940.3 (163.1)	1614.0 (122.6)

Fig. 12: [i] vowels in grave syllables:
Mean F1 and F2 values in Hertz (with standard deviation)

F1 and F2 values are higher for stressed [i], suggesting slight fronting and lowering. The plots in the following figure map the vowel space of stressed and unstressed [i], separated by syllable length in the target word (recall the TextGrid schema in (16); capital letter indicates predicted stress, number indicates position in frame sentence, etc.).

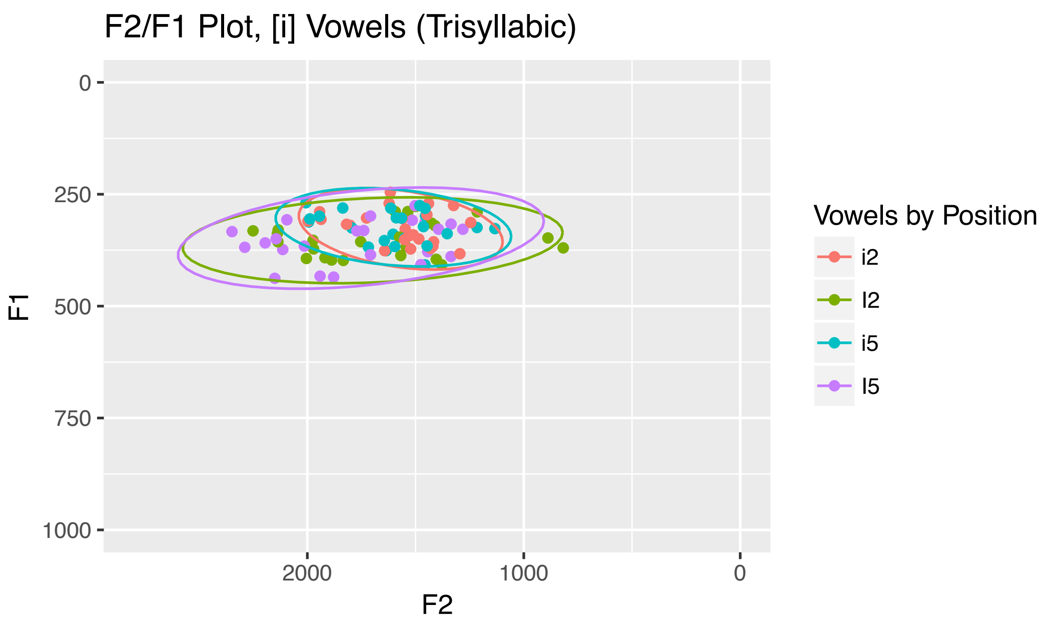

Fig. 13a: F1/F2 plots, [i] vowels on grave syllables (three-syllable words)

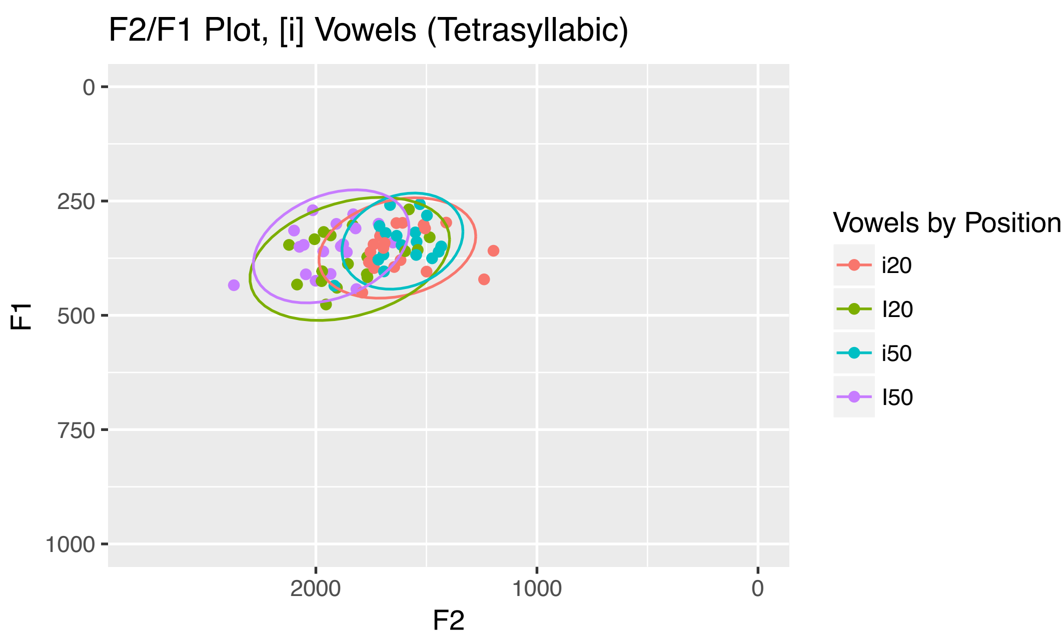

Fig. 13b: F1/F2 plots, [i] vowels on grave syllables (four-syllable words)

Plotting the F1/F2 figures also reveals noticeable intersection in the vowel space of stressed and unstressed vowels. The fronting trend suggested in the mean values is also apparent.

Comparison between models for both F1 and F2 showed significant effects. For F1, a marginally-significant effect was observed comparing the model integrating stress as a factor and the null model ($\chi^2(1)=6.3608$, $p=0.01167$); stressed grave syllables with the vowel [i] were lower than unstressed counterparts; F1 increased by 23.61 Hz ± 8.26 Hz (standard error) as a function of stress. For F2, the null model failed to converge, so number of syllables was removed from the set of random slopes. A likelihood ratio test between the resulting convergent models showed significance ($\chi^2(1)=9.7781$, $p=0.001766$). Stressed grave syllables containing [i] were more fronted than unstressed syllables containing the same syllable; the estimated increase in F2 was 222.02 Hz ± 60.44 Hz (standard error).

The vowel quality results do not uniformly support the hypothesis developed from traditional accounts of stress in Lithuanian. While the higher degree of fronting (higher F2) on stressed syllables supports the hypothesis, the F1 results—that is, stressed syllable vowels realize lower in the vowel space than unstressed syllables vowels—contradict impressionistic accounts.

4.3.2 [u]

Mean F1 and F2 values for [u] vowels on stressed and unstressed syllables are presented in Figure 14.

Frame/#		Stressed	Unstressed
$1^{st}/3\sigma$	F1	360.4 (36.5)	329.5 (38.0)
	F2	976.13 (213.9)	971.0 (121.4)
$2^{nd}/3\sigma$	F1	353.4 (42.2)	318.6 (42.7)
	F2	948.2 (131.2)	1024.4 (161.7)
$1^{st}/4\sigma$	F1	349.0 (40.8)	320.1 (35.1)
	F2	1138.1 (184.3)	1053.5 (90.7)
$2^{nd}/4\sigma$	F1	359.9 (53.4)	308.4 (35.4)
	F2	1089.8 (146.1)	1110.7 (88.5)

Fig. 14: [i] vowels in grave syllables:
Mean F1 and F2 values in Hertz (with standard deviation)

Like [i], F1 values for [u] vowels on stressed syllables tend to be higher than unstressed syllables. The F2 values, however, are inconsistent. Mean F2 values for stressed syllable vowels are either higher than unstressed syllables (in the case of tetrasyllabic target words in the first frame sentence), lower (as in trisyllabic and tetrasyllabic target words in the second frame sentence), or roughly the same (trisyllabic target words in the first frame sentence). There is less variation in these values than for the [i] values, as shown by the standard deviations, but overlap is substantial. This overlap is also evident in vowel plots for the vowels, where no clear distinction between vowels in stressed and unstressed vowels is discernible.

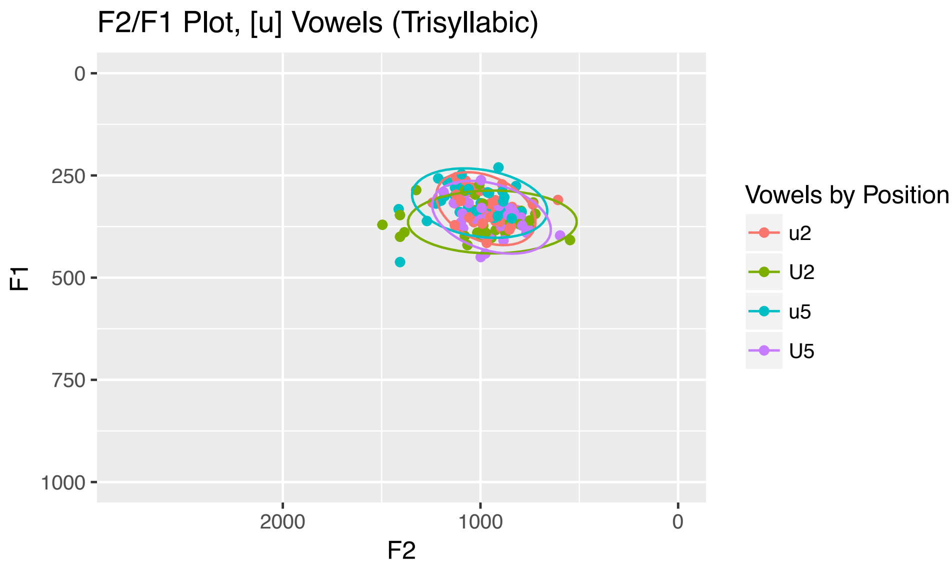

Fig. 15a: F1/F2 plots, [u] vowels on grave syllables (three-syllable words)

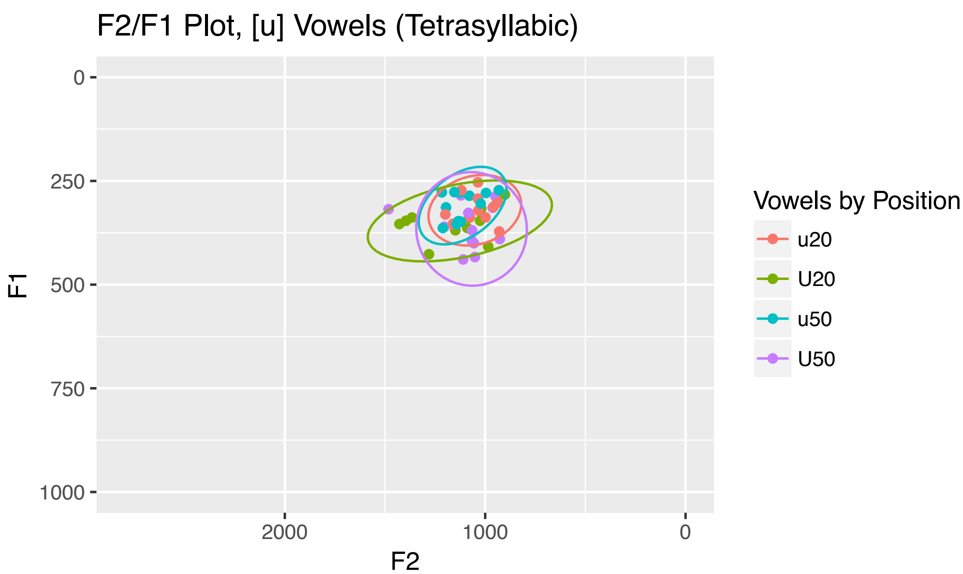

Figure 15b: F1/F2 plots, [u] vowels on grave syllables (four-syllable words)

Model comparisons on F2 confirm these generalizations. As was also the case with [i], the null model with a maximally-specified random effects structure failed to converge. Removing number of syllables from the random effects structure resulted in convergent models. A likelihood ratio test conducted on these models did not produce significant effects ($\chi^2(1)=0.7662$, $p=0.3814$), indicating that no qualitative difference on [u] vowels in terms of backness was observed between stressed and unstressed syllables.

The fully-specified F1 models yielded the same result as was observed with [i]. A significant lowering effect ($\chi^2(1)=26.668$, $p<0.001$) was evident, such that stress on syllables correlated with an increase of F1 by 36.71 Hz ± 5.55 Hz (standard error).

The observed results do not support the vowel quality hypothesis. Peripheralization of vowels on stressed syllables was not evident from either the F1 or F2 figures.

4.4 F0

This subsection presents the results of maximum F0 measurements to test the hypothesis developed from traditional descriptions of Lithuanian which assert that an increase in pitch is a correlate of stress. The prediction is that stressed syllables should exhibit a higher maximum fundamental frequency on vowels than unstressed syllables.

Mean values for vowels [e], [i], and [u] on stressed and unstressed syllables suggest that the former are realized with a higher maximum F0.

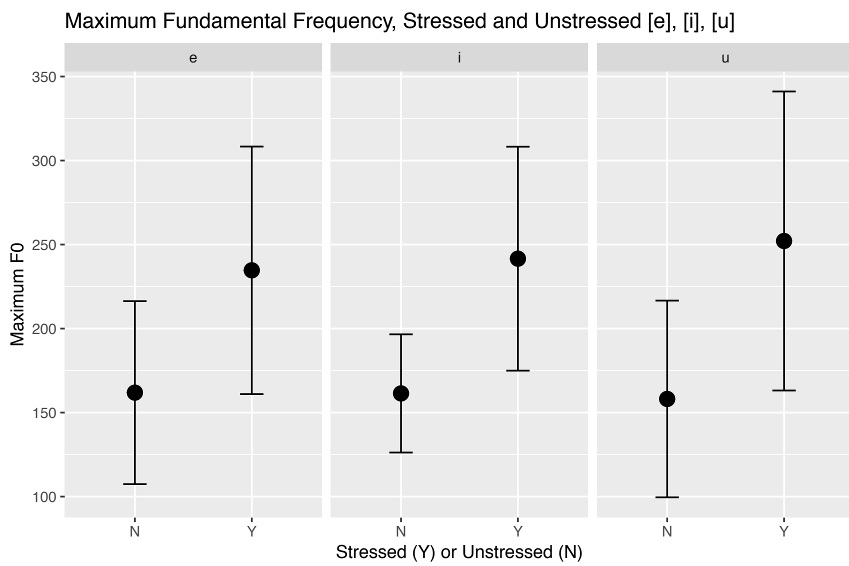

Fig. 16: Mean maximum F0 Values

Across all three vowels, F0 maxima are higher on stressed syllables. In the case of [u], this mean difference approaches 100 Hz, though the spread of maximum F0 values is large. The other vowels vary less by comparison. Detailed results from individual vowels are presented separately below.

4.4.1 [i]

Separated by frame sentence and number of syllables in target word, a uniform difference between maximum F0 values for [i] vowels and stressed and unstressed syllables is apparent. Stressed syllables reach a higher maximum fundamental frequency; the approximate range of this difference is 70 to 100 Hz.

Frame/#	Stressed	Unstressed
1st/3σ	229.54 (31.28)	158.42 (21.05)
2nd/3σ	239.26 (35.72)	160.67 (17.02)
1st/4σ	252.59 (27.14)	158.84 (27.78)
2nd/4σ	249.79 (34.94)	158.50 (14.61)

Fig. 17: [i] vowels in grave syllables:
Mean maximum F0 values in Hertz (with standard deviation)

While there is some variation in the mean figures for stressed syllables, unstressed syllables containing [i] are consistent in maximum F0 across frame sentences and between words of varying syllable length.

A linear mixed effects model with maximal random effects specification was compared to a null model (omitting Predicted Stress as a fixed effect) using a likelihood ratio test. The resulting comparison reached significance ($\chi^2(1)=72.808$, $p<0.001$); stress on syllables with [i] correlated to an increase in maximum F0 by approximately 80.62 Hz ± 4.699 Hz (standard error). This result supports the hypothesis and traditional descriptions of Lithuanian, that is, that increase in pitch is a correlate of stress.

Close inspection of the model residuals, however, suggests that the maximum F0 data do not constitute a continuous variable, but rather a categorical one. A violation of homoscedasticity is evident in the residual plot.

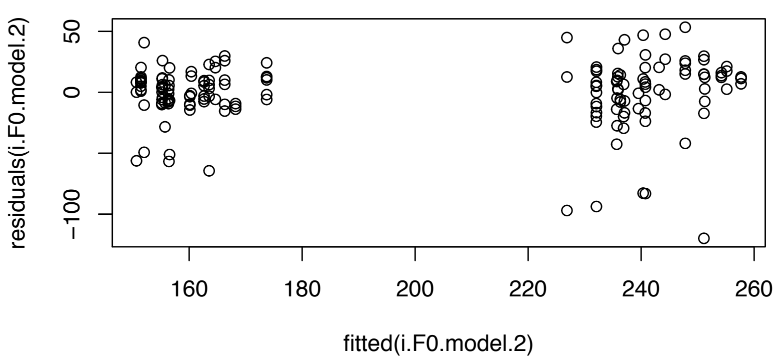

Fig. 18: Heteroscedastic residuals for maximum F0, [i] vowels

The residuals form two distinct clusters at either end of the plot, with no residuals in the middle. This result suggests a non-normal distribution in the data itself. The results of a Shapiro-Wilk test indicate that the maximum F0 data for [i] vowels are not normally distributed ($W=0.93132$, $p<0.001$). A histogram confirms a bimodal distribution of the data, with values for unstressed syllables congregating around the lower local maximum and stressed syllables congregating around the higher local maximum.

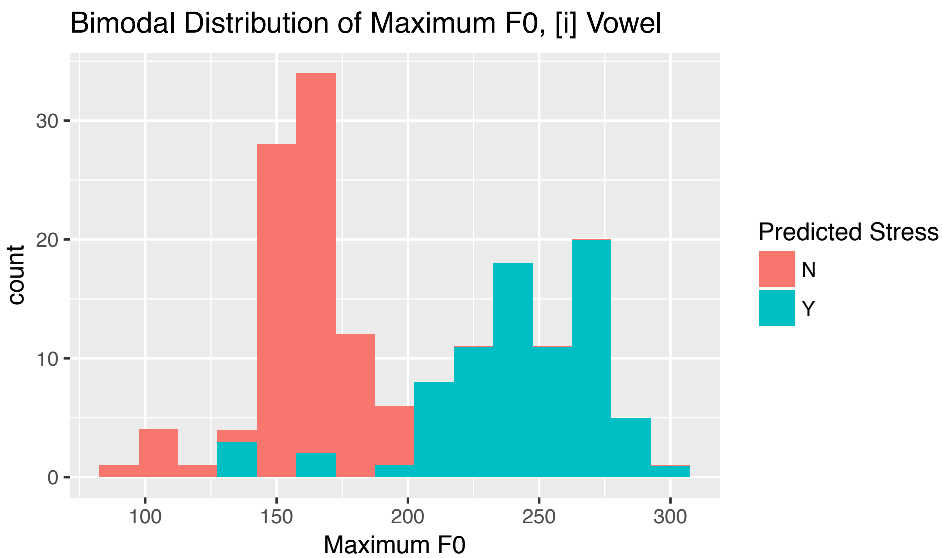

Figure 19: Distribution of Maximum F0, [i]

4.4.2 [u]

Mean values of fundamental frequency figures for [u] vowels corroborate the tendency identified in Figure 19. Mean maximum F0 on [u] vowels in stressed syllables range between 70-100 Hz higher than vowels on unstressed syllables, as the figure below illustrates.

Frame/#	Stressed	Unstressed
$1^{st}/3\sigma$	239.22 (44.62)	167.17 (24.91)
$2^{nd}/3\sigma$	256.58 (47.52)	159.51 (27.72)
$1^{st}/4\sigma$	235.30 (21.33)	151.40 (34.73)
$2^{nd}/4\sigma$	224.76 (56.12)	135.16 (35.99)

Fig. 20: [u] vowels in grave syllables:
Mean maximum F0 values in Hertz (with standard deviation)

These rough estimates are compatible with the hypothesis that increase in pitch is an acoustic correlate of stress in Lithuanian.

Though the comparison of linear mixed effects models reached significance for the data ($\chi^2(1)=62.073$, $p<0.001$), indicating that stress correlated to an average increase in maximum F0 of 83.4 Hz ± 6.233 Hz (standard error), the model's residuals were heteroscedastic, showing a bimodal distribution much like those from the models specified for [i].

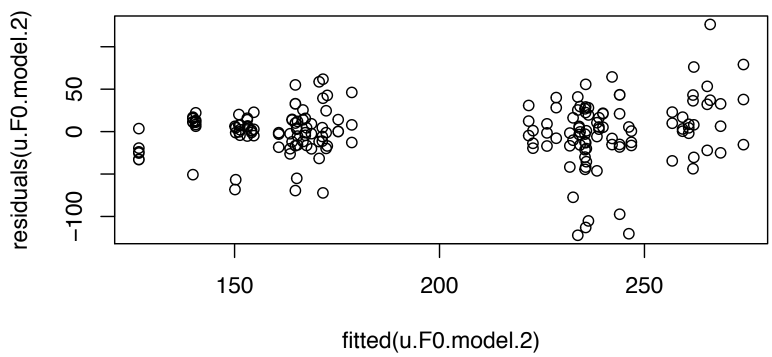

Fig. 21: Heteroscedastic residuals for maximum F0, [u] Vowels

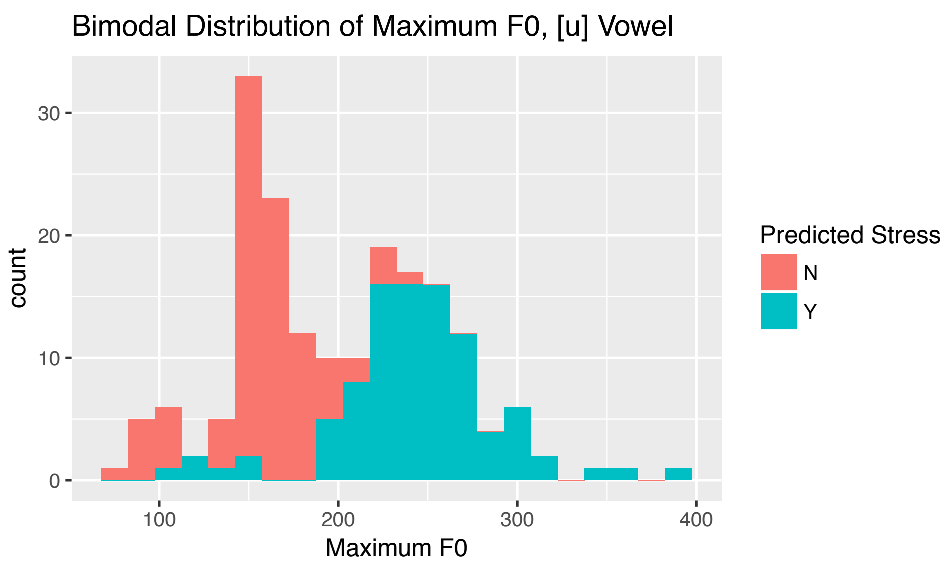

Fig. 22: Distribution of maximum F0, [u]

A Shapiro-Wilk test for normality on the data itself determined that the data are not normally distributed (W=0.97348, p=0.001699), with a histogram (Figure 22) suggesting bimodality.

The following section discusses the results of the acoustic experiment as they relate to the hypotheses developed in §3.5.

5. Discussion

The results of the production experiment do not uniformly confirm the hypotheses in (8). The acoustic correlates identified in traditional accounts of Lithuanian stress were not observed in stressed grave syllables in the current study of Lithuanian nominal forms. This section summarizes these results, and considers the current study's implications for phonological theory.

5.1 Summary of Results

The current study examined five acoustic measures which traditional sources claim to signal word stress in Lithuanian: duration, intensity, F1, F2, and F0.

Increased duration was not a reliable cue for stress on syllables with vowels [i] and [u]. Rather, stress correlated with a *decrease* in duration for the two vowels. These differences were found to be statistically significant. The respective decreases of 7 ms and 4 ms, however, are below the threshold of the Just-Noticeable-Difference for duration at 20% (Klatt 1976:1219). It is unlikely, then, that these differences are reliably perceptible. The limited data for syllables with the vowel [e] was suggestive of an increase in duration as a function of predicted stress, though these results are inconclusive given the size of the data set. Additionally, since vowel lengthening has been reported for [e] in stressed position, it is unclear whether the 13ms increase—though above the JND for duration—is a reflex of stress or some other phonological process.[1]

For vowels [e], [i], and [u], stressed syllables were realized with greater intensity than unstressed vowels. For [u], the resulting estimated increase of 1.63 decibels was found to be statistically significant. While this difference is likely to be perceptible (Harris 1963), the results for [i] did not yield significant effects. The salience of intensity as a correlate of stress is inconsistent among vowels, and therefore cannot be considered a reliable cue for stress.

Qualitative (F1/F2) differences between vowels in stressed and unstressed grave syllables did not confirm the hypothesis based on impressionistic accounts of qualitative reduction of vowels in unstressed position. Stress did not correlate reliably with vowel peripheralization for either [i] or [u]. Both vowels exhibited lowering (higher F1), an effect which was shown to be statistically significant. This effect on high vowels is the opposite of what is expected under the reduction hypothesis. While fronting (higher F2) effects for [i] were shown to be significant, no symmetric backing effect for [u] was observed.

Increase in pitch (as measured by maximum fundamental frequency) was the only acoustic measure included in the study which correlated consistently with stress. Analysis using linear mixed effects models yielded significant effects, indicating that stressed syllables realize an maximum F0 increase of approximately 80 Hz across both vowels. These results must be carefully interpreted, though, since the residuals produced from these models clearly violate the homoscedasticity assumption. This residual spread indicates a non-continuous dependent variable, rendering linear mixed effects models insufficient in modeling the data. Logarithmic transformation of the maximum F0 figures, in addition, was unsuccessful in stabilizing the variance of the residual points. Fitting a generalized linear mixed effects model to the data (using the *glmer()* function), which can accommodate categorical dependent variables, yielded significant effects for maximum F0 as a function of stress ($p<0.001$) for both [u] and [i].

Transformation of the data and model selection, however, should not obscure the critical result of the production experiment: there is no clear perceptible realization of metrical structure on grave syllables other than pitch modulation. F0 modulation, furthermore, is categorical in nature – more akin to the realization of tonal targets than simply a gradient phonetic increase in pitch (Keating 1996, Cohn 1998). In other words, the results of the acoustic experiment are consistent with an analysis in which what have been described as 'stressed' grave syllables are realized with a [H] tone, and 'unstressed' grave syllables with a (perhaps default) [L] tone.

---

[1] If [e] vowels in this position are found to lengthen significantly/perceptibly, then an account would need to be made for this lengthening without reference to stress (assuming the strong hypothesis from the acoustic results that there is no metrical structure in Lithuanian). Since the duration results are inconclusive—given the small sampling size—this question will be left to future research.

The current study failed to find acoustic evidence for metrical structure (i.e. stress) in Lithuanian, casting doubt on whether its categorization as a stress language is indeed correct. This has ramifications for phonological theory, as discussed below.

5.2 Implications for Phonological Theory

If Lithuanian is not a stress language, it cannot be considered evidence for phonological theories which formally characterize tone-stress interaction. This creates a gap in the empirical motivation for such theories. For example, de Lacy (2002)'s theory formalizes generalizations that metrical heads have an affinity for high tone (and non-heads for low tone), and claims that languages can manifest these tendencies in three different ways, described in (10) (de Lacy 2002:22 (35d)).

(10) *Manifestations of tone-stress interaction*
a. attraction of tone to (non-)heads (i.e. stress-driven tone)
b. attraction of (non-)heads to tone (i.e. tone-driven stress)
c. neutralization of tone on (non-)heads (i.e. stress-conditioned neutralization)

These processes are argued to be independent of one another, but can occur together in the same language. Lithuanian is presented as an example of a language with both tone-driven stress (10b) and metrically-conditioned neutralization (10c). However, the present study has failed to find acoustic evidence of metrical structure on Lithuanian grave syllables, undermining the claim that it has tone-driven stress.

The combination of two such processes in a single language like Lithuanian, in conjunction with [H]-tone culminativity, should arouse some concern. This is because it results in a system for which tone and stress are surface indistinguishable (recall discussion in §2.6). Word stress and word-level pitch accent (tone) are restricted to the same environment on surface forms, raising the question of whether they are two separate—albeit correlated—phenomena or just a single phenomenon. Intersection of their purported acoustic profiles (F0 as a reflex of both stress and pitch accent/tone) complicates the question further.[2]

• *Does tone-driven stress exist?*

The much broader question broached by this study is whether tone-driven stress exists at all. While several languages have been reported to have tone-driven stress, none have been subjected to acoustic analysis. It is therefore disturbing that Lithuanian, which impressionistic descriptions agree is a relatively uncontroversial case of tone-driven stress, turns out to not have reliable indicators of stress for grave syllables. Future work is required to examine other putative cases of tone-driven stress.

• *Tone-only analysis of Lithuanian*

Analysis of Lithuanian nominals need not make reference to stress. Blevins' (1993) analysis of Lithuanian nominals, for example, refers only to [H] tone and morae, conceptualizing 'strong' and 'weak' roots and suffixes in terms of underlying specification of [H] tone (see example (3)). Expanding on that rule-based account into an Optimality Theoretic framework, the

---

[2] One may reasonably ask: is such a language learnable? In other words, how could learners disambiguate stress and tone without some output-oriented indication that they are indeed separate? Though relevant to the case of Lithuanian, the question of learnability is left for future research.

distribution of nominal roots and suffixes according to the Basic Accentual Principle requires only a simple set of constraints. An align constraint ALIGNL(H,ω) (McCarthy & Prince 1993; see Alderete 2001 for application of align specifically to pitch accent) compels [H] alignment to the left edge of a prosodic word, while CULMINATIVITY (Alderete 2001) requires H-tone culminativity within a word. Additionally, a tonal faithfulness constraint IDENT(T) (McCarthy & Prince 1995), is violated whenever output tonal specification does not match that specified in the input. The ranking CULMINATIVITY » IDENT(T) » ALIGNL(H,ω) yields observed surface forms. Consider again the forms in (4), this time evaluated in Blevins' (1993) analysis with the proposed constraints. An additional form *kelm+as* 'tree-stump Nom.Sg' shows how the metrically-conditioned neutralization examples work without reference to metrical structure.

(11) */ví:r+ɑ́ms/ 'man DatPl.', /kɛlm+ɑ́ms/ 'tree-stump DatPl.', and /kɛlm+ɑs/ 'tree-stump NomSg.'*

a.

/ví:r+ɑ́ms/	CULM	IDENT(T)	ALIGNL(H,ω)
a. ví:r+ɑ́ms	*!		
b. vi:r+ɑ́ms		*	*!
☞ c. ví:r+ɑms		*	

b.

/kɛlm+ɑ́ms/	CULM	IDENT(T)	ALIGNL(H,ω)
☞ a. kɛlm+ɑ́ms			*
b. kɛ́lm+ɑms		*!	
c. kɛ́lm+ɑ́ms	*!	*	

c.

/kɛlm+ɑs/	CULM	IDENT(T)	ALIGNL(H,ω)
a. kɛlm+ɑs	*!		
b. kɛlm+ɑ́s		*	*!
☞ c. kɛ́lm+ɑs		*	

The above constraints—with Blevins' (1993) tonal specification paradigm—account for both the BAP as well as the stress movement effects described in traditional accounts, thus obviating the need for reference to strong and weak roots and suffixes (see discussion in Appendix 2). The root *kelm* 'tree stump', described in the traditional literature as a weak stem, is in the current account tonally-underspecified. Combining with the nominative singular suffix (also underspecified), the combined effect of CULMINATIVITY and ALIGNL(H,ω) force realization of [H] tone on the root (12c). Suffixation with [H]-toned *–ams* 'DatPl.' (12b) yields the so-called stress shift effect, which is in fact a reflex of the ranking IDENT(T) » ALIGNL(H,ω).

• *What were descriptions reporting as 'stress' in Lithuanian?*

A relevant issue to consider is why so many sources report stress in Lithuanian if metrical structure is not realized phonetically.

The existing literature varies widely in its characterization of Lithuanian prosody; scholars from both generative and functional traditions have offered treatments of this topic (Mathiassen 1996, Ambrazas 1997, Girdenis 2003, Young 1991, Kenstowicz 1972, Dudas and O'Bryan 1972).

Stress and pitch accent are understood as separate in some accounts and are conflated in others, and there is little consistency in the analyses between (and even within) traditions. Generative accounts of Lithuanian pitch accents, for example, include Blevins' (1993) [H]-tone and mora representation, but also Dudas & O'Bryan's (1972) analysis in which acute pitch accent is represented as stress on a high tone and circumflex as stress on a *low* tone. Kenstowicz (1972) proposes a system in which stress assignment precedes tone assignment within a rule-based framework, with the grammar switching from syllable to mora representation halfway through the derivation. These analyses, furthermore, vary in their compatibility with de Lacy's (2002) theory of tone-stress interaction; the latter two accounts cannot be incorporated into such a theory, either through contradicting the generalization that metrical heads seek out high tone, or through conflating tone and stress in such a way that they would fail to interact as the theory predicts. Even assuming representational consistency in a pitch accent/tone system, disambiguating tone and stress in systems which purportedly contain both is a nontrivial task, and conflicting accounts among sources is difficult to avoid (see, for example, Downing 2006 and the case of Chitumbuka).

Inconsistency across linguistic schools of thought is exacerbated by the ambiguous use of terminology, especially the use of the terms 'stress' and 'accent' and their conflation. This is particularly true in functional work on the language. Girdenis (2003:267), for example, offers the following atheoretical characterization of prominence in Lithuanian: "[the] contrast between clearer and less-clear pronunciations of syllables is accentuation, and the highlighting of one syllable against the background of others is stress or accent." It is unclear how the two differ from one another, or how the generative notion of stress as the realization of metrical structure fits into such a generalization. Incorporating pitch accent/tone and the ambiguity of its terms results in more confusion. Sources differ on how to classify the prosodic system of Lithuanian, such as Young (1991:5), who argues that the language is a "pitch-accent language" and not a "stress-accent language." While most functional accounts agree that Lithuanian exhibits a system of pitch accents, Young's rejection of the "stress-accent" characterization is unorthodox (though, if adopting "stress-accent" in the sense of Beckman (1986), is also compatible with the results of the current study).

There is, however, an acoustic reality to pitch modulation on grave syllables, one that is demonstrably perceptible. This is true whether it is understood as stress, pitch accent, or the interaction of the two. Traditional functional descriptions of the language (Ambrazas 1996, Mathiassen 1997, Girdenis 2003) resist analyses in which grave syllables bear tone, since these syllables are not tonally contrastive like acute and circumflex pitch accents. The increase in pitch is salient (approximately 80 Hz for both [u] and [i]), and there is impetus to distinguish the prosodic characteristics of grave syllables from other short (light) syllables. Stress—in Girdenis' (2003) sense of 'highlighting' a syllable against the backdrop of other syllables—allows the functional tradition to characterize this acoustic reality. Caution in importing the terms 'stress' and 'accent' from functional literature is crucial, as the theoretical assumptions behind this terminology varies, even within a single tradition. Additionally, disentangling stress from pitch accent, both in terms of the theoretical assumptions of relevant literature and in experiment design, is important to avoid potential confounds (see Roettger & Gordon 2017 for discussion of methodological implications).

5.3 Future Work

The experimental design used in this study could be extended to examine the acoustic properties of acute and circumflex syllables. Indeed, a comprehensive analysis of the phonetic realization of what has been called *stress* in Lithuanian would integrate data from all three pitch accents in the language. Given that previous acoustic work has been largely devoted to the acute/circumflex distinction and not the acute/toneless and circumflex/toneless distinctions (see §2.5), such a comparison would offer novel data to the extant body of phonetic studies of Lithuanian.

Results from the current study make predictions about these distinctions which can form hypotheses for further iterations of the experiment. The lack of acoustic evidence for metrical structure in grave syllables would predict that acute and circumflex syllables do not differ from tonally-neutralized syllables in terms of duration, intensity, or vowel quality (F1/F2). The presence of an H tonal segment on these syllables would predict higher maximum fundamental frequency, mirroring the distributions in Figures 19 and 22. Despite the fact that previous acoustic analyses of acute and circumflex have not found pitch to enter into their distinction, this does not preclude pitch from distinguishing acute and circumflex from syllables for which tone has been neutralized. This question is beyond the scope of the current study, and will be left for future research.

6. Conclusion

This paper has provided an acoustic analysis of controversial grave syllables in nominal forms in Lithuanian (a language which has been claimed to exhibit both pitch accent and word stress), in an attempt to offer a phonetic characterization of stress in a pitch accent/tone language. Results of a production experiment indicate no clear perceptible realization of metrical structure on these syllables, contra previous impressionistic accounts, and including duration, intensity, and formant structure. An increase in pitch (maximum F0) was found to correlate consistently with stress position as predicted in traditional accounts. The distribution of this acoustic characteristic across the experiment data was found to be categorical rather than gradient, consistent with the realization of a [H] tone target. This finding is problematic for the characterization of Lithuanian as a stress language, and therefore its status as evidence for theories which formalize tone-stress interaction.

In other words, the present study fails to answer the question “what is stress in a tone/pitch accent language?” as it did not find clear phonetic correlates of metrical structure in Lithuanian grave syllables. So, it shows that Lithuanian does not constitute valid evidence for theories which formalize such an interaction. A language which could provide insight into the above questions would exhibit distinct evidence for lexical tone and metrical structure, with metrical structure realized with acoustic properties other than a difference in fundamental frequency.

To this point, Lithuanian has been presented as an uncontroversial case of tone-driven stress. However, as no other tone-driven stress languages have been subject to acoustic analysis, it is not an unreasonable stance at this time to question the existence of tone-driven stress.

Appendix 1: Target Roots and Suffixes

Root/Suffix	Orthography	IPA	Gloss
Root	katuk	[kɑtʊk]	'scythe'
Root	ratuk	[rɑtʊk]	'wheel'
Root	degut	[dɛgʊt]	'tar'
Root	lupik	[lʊpɪk]	'loan shark'
Root	ridik	[rɪdɪk]	'radish'
Root	paket	[pɑkɛt]	'package'
Root	mygtuk	[miːgtʊk]	'button'
Root	apetit	[ɑpɛtɪt]	'appetite/stomach'
Root	kalakut	[kɑlɑkʊt]	'turkey'
Root	palikim	[pɑlɪkɪm]	'legacy'
Root	gynim	[giːnɪm]	'defense'
Root	rėmim	[ræːmɪm]	'support'
Root	cepelin	[t͡sɛpɛlɪn]	(Lit. dish)
Root	sēnum	[seːnʊm]	'age'
Root	malonum	[mɑloːnʊm]	'pleasure'
Root	tylum	[tiːlʊm]	'silence'
Suffix	as	[ɑs]	NomSg.
Suffix	o	[o]	GenSg.
Suffix	u	[ʊ]	InstSg.
Suffix	us	[ʊs]	AccPl.

Appendix 2: Functional Characterizations of Nominal Stress and de Saussure's Law

This Appendix describes functional characterizations of stress and stress movement in Lithuanian nominals, expanding on discussion in §2.3.

Word stress in Lithuanian is culminative, and can be realized on both 'short' and 'long' syllables. Assignment of stress is reportedly sensitive to pitch accent: the leftmost lexically-specified pitch accented syllable is stressed, otherwise the leftmost syllable. This pattern is known as the 'Basic Accentuation Principle' (Kiparsky & Halle 1977). The BAP is, however, an oversimplification of stress and accentual assignment in the language (Ambrazas 1997, Mathiassen 1996, Girdenis 2003).

A Lithuanian nominal consists of a root and inflectional suffixes. Nominal stems do not occur in isolation – they obligatorily appear with a case-marking suffix.

(12) Case-marking on the nominal root vyr- 'man'

a. [ˈvíːr-ɑs]	'man NomSg.'	d. [kɛ́lm-ɑs]	'tree-stump NomSg.'
b. [ˈvíːr-o]	'man GenSg.'	e. [ˈkɛ́lm-ɑː]	'tree-stump AccSg.'
c. [ˈvíːr-ʊs]	'man AccPl.'	f. [ˈkɛ́lm-eː]	'tree-stump LocSg.'

Nominal roots are further specified for pitch accent. For example, the stems *vyr* 'man' and *kelm* 'tree stump' in (13) are both *acute* stems.

When a root is stressed, the lexically-specified pitch accent is pronounced. When certain roots combine with certain suffixes, however, stress is realized on the syllable containing the suffix instead, as in the dative plural suffix in (14) below.

(13) *Stress pattern with the dative plural suffix*
a. /ví:r+ɑ́ms/ → [ˈví:.rɑms] 'man DatPl.'
b. /kɛlm+ɑ́ms/ → [kɛl.ˈmɑ́ms] 'tree-stump DatPl.'

When the root *vyr* 'man' takes the dative plural suffix *–ams*, stress is realized on the root, and *vyr* is pronounced with the canonical acute intonation. When the suffix combines with *kelm*, stress is realized on the second syllable. Pitch accent on the first syllable is neutralized. This generalization holds for acute and circumflex roots equally.

The functional literature therefore divides Lithuanian nominals into four accent classes, based on their behavior in combination with different suffixes (Ambrazas 1997, Matthiassen 1996). Roots which retain stress are classified as *strong*, while those which do not are classified as *weak*.

(14) *Four accent classes*

Class Type	**Example**	**Gloss**
Class 1: 'strong' acute stems	*vyr*	'man'
Class 2: 'strong' circumflex (and grave) stems	*ind*	'dish'
Class 3: 'weak' acute stems	*kelm*	'tree stump'
Class 4: 'weak' circumflex (and grave) stems	*vaik*	'child'

Similarly, the inventory of inflectional suffixes is partitioned into *strong* and *weak* classes. Strong suffixes can attract stress off of weak stems, but not strong stems. When a weak root and weak suffix combine, stress defaults to the stem. Four classes of roots combine with two types of suffixes, totaling eight logically-possible combinations. Their predicted stress patterns are outlined below.

Root	Strong Suffix	Weak Suffix
Class 1 (strong)	Root	Root
Class 2 (strong)	Root	Root
Class 3 (weak)	Suffix	Root
Class 4 (weak)	Suffix	Root

Fig. 23: Predicted stress position in Lithuanian nominals

The stress patterns in (15) are thus predicted for strong suffixes like dative plural *–ams* (stress movement cases indicated in **bold**) and weak stems like nominative singular *–as*.

(15) *Strong and weak roots and suffixes and resulting surface stress*

a.	/vi:r+ɑms/	→	[ˈvi:.rɑms]	'man DatPl.'
b.	/ɪnd+ɑms/	→	[ˈɪn.dɑms]	'dish DatPl.'
c.	/kɛlm+ɑms/	→	[kɛl.ˈ**mɑms**]	'tree stump DatPl.'
d.	/vaik+ɑms/	→	[vɑi.ˈ**kɑms**]	'child DatPl'
e.	/vi:r+ɑs/	→	[ˈvi:.rɑs]	'man NomSg.'
f.	/ɪnd+ɑs/	→	[ˈɪn.dɑs]	'dish NomSg.'
g.	/kɛlm+ɑs/	→	[ˈkɛl.mɑs]	'tree stump NomSg.'

h. /vɑik+ɑs/ → [ˈvɑi.kɑs] ‘child NomSg.’

Admittedly, the full stress paradigm of Lithuanian nominals is more complicated than the above generalization. In addition to the *strong*/*weak* dichotomy, suffixes can be either *attracting* or *non-attracting*. Attracting suffixes attract stress away from roots of accent classes 2 and 4—that is, circumflex or grave—but never classes 1 or 3 (the acute roots). This generalization goes back to de Saussure (1876); these cases of stress movement are thus termed **de Saussure’s Law** effects. As a result, suffixes can be categorized into four types, based on whether they are strong or weak, attracting or non-attracting.

(16) *Four-way distinction of suffixes based on de Saussure’s Law*
a. Weak, Non-Attracting: will never have stress
b. Weak, Attracting: stressed after roots of accent classes 2 and 4
c. Strong, Non-Attracting: stressed after roots of accent classes 3 and 4
e. Strong, Attracting: stressed after roots of accent classes 2, 3, and 4

Figure 24 displays the full range of possible accent patterns on four roots and four suffixes. Accent class on roots is indicated in parentheses to the right of the gloss, while the accentual properties of the suffix follows the order in (17) (A = weak, non-attracting, B = weak, attracting, etc.). Bold indicates stress.

	/vɑrn/ ‘crow’ (1)	/rɑnk/ ‘hand’ (2)	/ʒmo:n/ ‘wife’ (3)	/dɑin/ ‘song’ (4)
/-ɑ:/ AccSg. (A)	**vɑrn**-ɑ:	**rɑnk**-ɑ:	**ʒmo:n**-ɑ:	**dɑin**-ɑ:
/-ɑ/ InstSg. (B)	**vɑrn**-ɑ	rɑnk-**ɑ**	**ʒmo:n**-ɑ	dɑin-**ɑ**
/-o:ms/ DatPl. (C)	**vɑrn**-o:ms	**rɑnk**-o:ms	ʒmo:n-**o:ms**	dɑin-**o:ms**
/-ɑ/ NomSg. (D)	**vɑrn**-ɑ	rɑnk-**ɑ**	ʒmo:n-**ɑ**	dɑin-**ɑ**

Fig. 24: Four-way distinction among roots and suffixes

Much work has been devoted to explaining these effects; some posit phonological explanations (Blevins 1993, Dogil 1999, Yamazaki 2016) and others morphological analyses (Ambrazas 1997, Girdenis 2003, Young 1991). The purpose of this study is not to provide an account of de Saussure’s Law patterns, though these effects do figure into the experiment design (see §3).

References

Alderete, John D. (2001). Dominance effects as transderivational anti-faithfulness. *Phonology* **18**. 201-253.

Ambrazas, Vytautas, ed. (1997). *Lithuanian grammar.* Vilnius: Baltos Lankos.

Anderson, Cori (2017). Personal communication. Rutgers University.

Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**. 1–48.

Beckman, Jill (1986). *Stress and non-stress accent*. Foris, Dordrecht.

Blevins, Juliette (1993). A tonal analysis of Lithuanian nominal accent. *Language* **69**. 237-273.

Boersma, Paul & David Weenink (2016). *Praat: doing phonetics by computer* (version 6.0.25). http://praat.org.

Cohn, Abigail (1998). The phonology-phonetics interface revisited: Where's phonetics? *Texas Linguistic Forum* **41**. 25-40.

de Lacy, Paul (2002). The interaction of tone and stress in Optimality Theory. *Phonology* **19**. 1-32.

de Saussure, Ferdinand (1876). Accentuation lituaniene. *IF* **6**. 157-166.

Dogil, Gregorz (1999). Baltic languages. In H. van der Hulst (ed.) *Word Prosodic Systems in the Languages of Europe*. Berlin: Mouton de Gruyter. 877-897.

Dogil, Gregorz & Briony Williams (1999). The phonetic manifestation of word stress. In H. van der Hulst (ed.) *Word Prosodic Systems in the Languages of Europe*. Berlin: Mouton de Gruyter. 273-334.

Dogil, Gregorz & Gregor Möhler (1998). Phonetic invariance and phonological stability: Lithuanian pitch accents. In *Proceedings of the 5th International Conference on Spoken Language* (ICSLP). 75-86.

Downing, Laura (1990). Local and metrical tone shift in Nguni. *Studies in African Linguistics* **21**(3). 261-317.

Downing, Laura (2006). The prosody and syntax of focus in Chitumbuka. *ZAS Papers in Linguistics* **43**. 55-79.

Dudas, Karen & Margie O'Bryan (1972). Lithuanian verbal accentuation. *Studies in the Linguistic Sciences* **2**. 86-107.

Francis, Alexander, Valter Ciocca & Jojo Man Ching Yu (2002). Accuracy and variability of acoustic measures of voicing onset. *JASA* **113**. 1025-1032.

Garellek, Marc & James White (2012). Stress correlates and vowel targets in Tongan. *UCLA Working Papers in Phonetics* **110**. 65-85.

Girdenis, Aleksas (2003). *Teroinai Lietuviu fonologijos pagrindai* [Theoretical foundations of Lithuanian phonology]. Vilnius: Mokslo Ir Enciklopediju Leidybos Institutas.

Gordon, Matthew (2014). Disentangling stress and pitch-accent: a typology of prominence at different prosodic levels. *Word Stress: Theoretical and Typological Issues* **2018**. 83-118.

Gordon, Matthew & Timo Roettger (2017). Acoustic correlates of word stress: a cross-linguistic survey. *Linguistics Vanguard* **3**. Article # 20170007.

Grigorjevs, Juris & Jurgita Jaroslavienė (2015). Comparative study of the qualitative features of the Lithuanian and Latvian monophthongs. *Baltistica* **50**. 57-89.

Harris, Donald J. (1963). Loudness discrimination. *Journal of Speech and Hearing Disorders*. (Monograph Supplement No. 11).

Kaukėnienė, Lidija (2003). Lietuviu kalbos nekirčiuotu balsiu keikybė [The quantity of unstressed vowels in Standard Lithuanian]. *Acta Linguistica Lithuanica* **48**. 35-47.

Kaukėnienė, Lidija (2012). Dabartiniu baltu bendriniu kalbu kirčiuotu ir nekirčiuotu žemutiniu balsiu spektrai bei trukmė [The spectra and duration of low vowles in Standard Lithuanian and Latvian]. Respectus Philologicus **21**. 216-225.

Keating, Patricia (1996). The phonology-phonetics interface. *UCLA Working Papers in Phonetics* **92**. 45-60.

Kelly, Niamh (2013). An acoustic description of Lithuanian pitch accent. Talk given at Texas Linguistic Society, University of Texas at Austin, 1-3 November 2013.

Kenstowicz, Michael (1972). Lithuanian phonology. *Studies in the Linguistic Sciences* **2**. 1-85.

Kiparsky, Paul & Morris Halle (1977). Towards a reconstruction of the Indo-European accent. In Larry Hyuman (ed.) *Studies in Stress and Accent* (Southern California Occasional Papers in Linguistics 4). Los Angeles: University of Southern California, 209-238.

Klatt, Dennis H. (1976). Linguistic uses of segmental duration in English: acoustic and perceptual evidence. *JASA* **59**. 1208-1221.

Leskien, August (1919). *Litauisches Lesebuch*. Heidelberg: Carl Winter.

Mathiassen, Terje (1996). *A short grammar of Lithuanian*. Columbus, OH: Slavica Publishers.

McCarthy, John & Alan Prince (1993). Generalized alignment. In van Marle, G. B. J. (ed.) *Yearbook of Morphology*. Dordrecht: Kluwer, 79-153.

McCarthy, John & Alan Prince (1995). Faithfulness and reduplicative identity. In Beckman, J., Dickey, I., & Urbanczyk, S. (eds.) *Papers in Optimality Theory: University of Massachusetts Occasional Papers in Linguistics* **18**. 249-384.

R Development Core Team (2017). *R: a language and environment for statistical computing* (version 3.4.3). Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Robinson, David F. (1968). Some acoustic correlates of tone in Standard Lithuanian. *The Slavic and East European Journal* **12**. 206-212.

Roettger, Timo & Matthew Gordon (2017). Methodological issues in the study of word stress correlates. *Linguistics Vanguard* **3**. Article # 20170006.

Schroeder, Larry D., David L. Sjoquist and Paula E. Stephan. (1986). *Understanding regression analysis: an introductory guide*. Newbury Park, CA: Sage Publications.

Senn, Alfred (1966). *Handbuch der litauischen Sprache*. Heidelberg: Carl Winter.

Turk, Alice, Satsuki Nakai & Mariko Sugahara (2006). Acoustic segment durations in prosodic research: a practical guide. In Stefan Sudhoff, Denisa Lenertova, Roland Meyer, Sandra Pappert, Petra Augurzky, Ina Mleinek, Nicole Richter & Johannes Schließer (eds.) *Methods in Empirical Prosody Research*. Berlin: Mouton de Gruyter. 1–28.

Yamazaki, Yoko (2016). Monosyllabic circumflexion in Lithuanian. PhD Dissertation, Stockholm University.

Young, S.R. (1991). *The prosodic structure of Lithuanian*. University Press of America.