Inuktitut and the concept of word-level prominence

Anja Arnhold, Emily Elfner and Richard Compton

February 19, 2020

Abstract

This chapter addresses the question of word prominence in Eastern Canadian Inuktitut, a part of the Inuit dialect continuum constituting a branch of the Eskimo-Aleut language family. Inuit is an extreme example of polysynthesis, with productive noun and verb incorporation that can be applied recursively, along with extensive word-internal modification. The main analysis is based on original data from South Baffin Island Inuktitut (Uqqurmiut), but the chapter also examines the literature on other varieties from across the language family. We present acoustic analyses of three potential correlates of stress or prominence: duration, fundamental frequency and intensity. Duration of syllables increased at the end of the word, while fundamental frequency and intensity dropped at the right word edge. Word-internally, no alternating or other regular patterns appeared. Comparing these results to hypotheses of what would be expected for metrical stress systems and other types of word prominence, we conclude that there is no indication that South Baffin Island Inuktitut has stress or another type of word-level prominence. Instead, in line with previous research on Inuit prosody, we find that the language regularly marks the borders of words and other prosodic constituents.

Keywords

Inuit, Eskimo-Aleut, prosody, prominence, stress, polysynthesis, syllable weight

1 Introduction

Eastern Canadian Inuktitut is a part of the Inuit dialect continuum constituting a branch of the Eskimo-Aleut language family. The Eastern Canadian Inuktitut dialect group examined here is spoken in central and eastern Nunavut, northern Quebec, and Labrador, Canada. In this chapter, we investigate the status of stress or word prominence for South Baffin Inuktitut (Uqqurmiut).

Section 2 lays out hypotheses that the rest of this chapter tests empirically. Since stress minimal pairs have not been observed for any Inuit variety, we concentrate on the possibility that Inuktitut may have fixed or rhythmic stress, with or without an additional influence of syllable weight (quantity-sensitivity). The chapter then empirically investigates acoustic measures to see if they indicate that syllables at a certain position or positions within the word are consistently prominent. Section 3 presents the data, which come from two corpora, dialogues and single-word utterances, from a website created by Inuit and targeted at beginning learners of Inuktitut (Pirurvik Centre 2015). Section 4 presents the analyses of 1) duration, 2) the acoustic correlate of pitch, fundamental frequency (f0), and 3) the acoustic correlate of loudness, intensity. All three measures showed clear effects of word boundaries, with duration increasing and f0 and intensity decreasing towards the end of the word, but not further regularities. Based on these results, we discuss in section 5 how various definitions of stress or other types of word prominence apply to Eastern Canadian Inuktitut—and, based on the available literature, Inuit languages and the Eskimo-Aleut language family more broadly. Section 6 concludes that, in contrast to Yupik, there is no reason to evoke the notions of stress or word-level prominence in the analysis of Inuit.

The remaining subsections of this introduction provide the necessary background on Inuktitut (and other Inuit varieties), establishing the morphological complexity of the language and introducing its phonology, particularly its prosody. To lay the foundation for a discussion of word-level prominence, particular attention is given to phonemic length, syllable structure and syllable weight to set up the investigation of the possibility of a quantity-sensitive stress system.

1.1 Morpho-syntax

Inuktitut has a complex and productive morphology. It is both head and dependant marking, with a rich system of agreement on verbs co-indexing both subjects and (absolutive) objects, as well as case marking on nouns,[1] see (1) (from Pirurvik Centre 2015, glosses added).[2] The patterning of case in the language includes both an ergative-absolutive alignment and an antipassive one in which objects are marked with oblique case, see (2) (from Johns 2006, 294, glossing slightly adjusted). Word order is variable, but SOV has been claimed to be the default order (Swift 2004). The language is almost exclusively suffixing, with only one prefix that is limited to the demonstrative system and is arguably frozen.

[1] In addition to verbs and their arguments, another example of simultaneous head and dependant marking is that possession is marked on both the possessor and its possessum, see (4) below.

[2] The orthography used in the examples here and below is phonemic and largely corresponds to IPA transcriptions, except for the following: *ng* = /ŋ/, *nng* = /ŋː/, *g* = /ɣ/, *r* = /ʁ/. Note also that double letters indicate long segments. See the next section for a segmental inventory.

(1) a. Uqaalaut namu-nnga-qqau-viuk?
phone(ABS) where-go-REC.PAST-INTER.2SG.3SG
'Where did you put the phone?'

b. Pulaarvim-mut.
living.room-ALL.SG
'In the living room.'

c. Atii, ai-guk.
okay get-IMP.2SG.3SG
'OK then, go get it.'

(2) a. Anguti-up nanuq kapi-jaa.
man-ERG polar.bear(ABS) stab-3SG.3SG
'The man stabbed the polar bear.'

b. Angut nanur-mik kapi-si-juq.
man(ABS) polar.bear-OBL stab-ANTIPASSIVE-3SG
'The man is stabbing the polar bear.'

The language makes a clear distinction between nouns and verbs, with dedicated inflectional paradigms for each. In addition to agreement with ergative and absolutive arguments, verbs are routinely marked for tense, aspect, negation, clause type, and changes to their argument structure, such as passivization and causativization. Nouns are marked for case, number, and possession, including the person and number of their possessor.

Inuktitut exhibits a cluster of phenomena frequently labelled as polysynthetic. A closed class of verbs triggers obligatory noun incorporation of their object (see Johns 2007), as illustrated in (3) (Pirurvik Centre from 2015, glosses added, except for (3-d), which is from Compton 2012, p. 10). These objects, and nominals generally, may themselves be morpho-syntactically complex, containing category-changing morphology and modifiers, see e.g. the nominalization combined with incorporation in (3-b) and incorporation with modification in (3-d) . A subset of incorporating verbs (those associated with location and motion) can incorporate nominals with oblique cases. Verbal complexes may also contain modifiers, which Compton (2012) argues to be adjectives and adverbs. There are no free articles, prepositions, complementizers, modals, or auxiliaries. Instead, case, clause-type marking, and suffixal verbs inside verbal complexes serve these functions. These various properties combine to create long, often holophrastic, words. Word structure in other polysynthetic languages of North America, for instance those of the Algonquian and Dene (Athabaskan) language families, has been described in terms of templates (see e.g. Oxford 2014, Rice 2000, for discussion). While words in Inuit generally begin with a root and end in inflection, possibly followed by clitics, the highly recursive nature and high degree of productivity of Inuit morphology make a templatic analysis untenable. For instance, Fortescue's (1980) analysis of West Greenlandic (Kalaallisut) word structure employs what are essentially a set of recursive phrase-structure rules, distinct from syntax, to explain the morpheme order in the language. Adopting Distributed Morphology (Halle & Marantz 1993, 1994), Compton & Pittman (2010) propose instead that syntactic phases (DP and CP) are realized as phonological words, with constituents smaller than these being realized word-internally.

(3) a. Kaapi-tu-ruma-vit?
coffee-consume-want-INTER.SG
'Would you like coffee?'

b. Sukali-suu-ngu-vit?
sugar-HAB.NOMLZ-COP-INTER.2SG
'Do you take sugar?'

c. Aupaq-tu-mik titirauti-qaq-qit?
red-DEC-OBL.SG pen-have-INTER.2SG
'Do you have a red pen?'

d. Uqalimaarvi-ralaa-qaq-tugut.
library-small-have-DEC.1PL
'We have a small library.'

While speakers generally agree on what constitutes a word in the language, and furthermore given that multiple writing systems across the Inuit dialect continuum have also converged on the same units, the status of words is not a trivial issue. Although Sadock (1980) lays out a number of diagnostics for identifying words in West Greenlandic that are also applicable to Inuktitut (for instance, that words constitute a domain for phonological operations and that they are inseparable and uninterruptible), this state of affairs appears compatible with these words being (i) complex heads, (ii) XP-sized constituents mapped to phonological words, or even (iii) phonological phrases that merely exhibit phonological properties typically associated with words in other languages.

1.2 Segmental Phonology

Inuit consonant inventories typically consist of a series of voiceless stops, a series of voiced continuants, nasals, a strident, a voiced lateral, and, in some dialects, its voiceless counterpart. Places of articulation include bilabial, coronal, velar, and uvular. Some dialects have debuccalized /s/ to /h/. Table 1 shows the consonant inventory for South Baffin Inuktitut.

Table 1: Consonant inventory of South Baffin Inuktitut (adapted from Dorais 2003, p. 98).

	bilabial	coronal	velar	uvular
stops	p	t	k	q
voiced fricatives	v		ɣ	ʁ
voiceless fricatives		s		
approximant		j		
lateral approximant		l		
nasals	m	n	ŋ	

Voiceless plosives and voiced continuants participate in a number of phonologically-conditioned allomorphic alternations at morpheme boundaries, with stops appearing after stops, and continuants appearing after vowels (see Compton 2009, for more details). For example, the intransitive second person singular interrogative morpheme appears as *-vit* in (3-a) and (3-b), where its initial consonant is in intervocalic position. By contrast, the morpheme appears as *-qit* in (3-c), with an initial plosive following the final plosive of the preceding morpheme.

The Inuit dialect continuum exhibits varying degrees of regressive place assimilation as well as manner assimilation. Moving from West to East, the potential place combinations in hetero-organic consonant clusters become more and more restricted, with Labrador and Greenlandic for the most part limiting clusters to geminates (dialectal overviews appear e.g. in Bobaljik 1996, Dorais 1986). For example, note that the first consonant of the above-mentioned interrogative morpheme in (3-c) is identical to the last consonant of the preceding root, i.e. the two consonants form a geminate. Yet a different form of the same morpheme, with its initial consonant likewise assimilated to the preceding one, appears in *qanuip**pit**?* 'How are you?'.

The vowel inventory consists of only three (surface) vowels, /a/, /i/, and /u/. A fourth vowel, schwa, has merged with /i/ in most contexts in all but one Inuit dialect, and yet in some dialects it continues to exhibit distinct phonological behaviour from etymological *i (see Compton & Dresher 2011, for details). Vowel length is contrastive. Any two vowels can occur in combination in most Inuit varieties, including South Baffin Inuktitut, whereas assimilation eliminates all combinations except word-final /ai/ in West Greenlandic. Most descriptions do not specify whether combinations of two vowels are homo- or heterosyllabic, but where the question is addressed, they are usually argued to always be homosyllabic (Massenet 1978, Rischel 1974, but see Kaplan 1981).[3] Three-moraic vowel sequences created by morphology are interrupted with an epenthetic consonant, typically /ŋ/.

1.3 Prosody

The notion of stress or accent had been used in early descriptions of Inuit (e.g. Pyle 1970, p. 131), but Rischel (1974, p. 91) concludes that it has no well-defined status in West Greenlandic. He further states that it is difficult to obtain agreement on stress placement, although both native speakers and phonetically trained Danish listeners tend to hear stress on either the antepenultimate or the last vowel. These two vowels are the locations of the two high pitch points the of the word-final falling-rising pitch contour appearing on all West Greenlandic words when spoken in isolation (Mase 1973) — in an autosegmental-metrical framework (Goldsmith 1976, Pierrehumbert 1980, Ladd 2008), this contour can be modelled as HLH tones associated with the last three vowel moras (see Arn-

[3]Thus, they form a diphthong according to the standard assumption that two vowels sharing a syllable nucleus are per definition diphthongs. Note, however, that there is no evidence that Inuit possesses separate phonemic diphthongs which would have to be added to the vowel inventory.

hold 2014, for a recent overview of research on West Greenlandic prosody). Finally, Rischel points to the potential importance of syllable weight, suggesting that heavy syllables may be perceived as stressed, especially if they carry a high tone. However, while Rischel discusses the great importance of moras for intonation and other aspects of West Greenlandic phonology, a classification of syllables into light and heavy is not completely straightforward.

An important prosodic characteristic of Inuit is that vowel and consonant length distinctions are phonemic and are used to mark lexical and morphological distinctions, e.g. [ataːt**a**] 'father' vs. [ataːt**aː**] 'his/her father' in West Greenlandic or [tu**t**uk] 'messy hair' vs. [tu**tː**uk] 'caribou' in Labrador Inuttut. Acoustic studies of West Greenlandic show that long consonants are generally about double the length of short ones (Mase & Rischel 1971, p. 235; Nagano-Madsen 1992, p. 61–63), while long vowels have double or triple the duration of short ones (Nagano-Madsen 1992, p. 61–63). Pigott (2012, p. 87) generally finds long consonants to be 2.7 times longer than short ones in Labrador Inuttut.

As complex onsets and codas are disallowed, Inuit syllable structure can be represented as (C)V(V)(C),[4] with onsetless syllables only occurring word-initially. Recognizing the importance of phonological length, Kleinschmidt (1985) proposed a quasi-moraic account of syllable weight for West Greenlandic in the 19th century (see especially his letters in Holtved 1964), assigning a value of 2 to (C)V syllables, a value of 3 to (C)VC syllables, a value of 4 to (C)VV syllables and a value of 5 to (C)VVC syllables. Note that as in modern moraic theory (with some exceptions, e.g. Ryan 2014), only vowel nuclei and coda consonants contribute to syllable weight, whereas onset consonants are not relevant in the distinction of the four different syllable types. A more atypical aspect of Kleinschmidt's suggestion is that while consonants and vowels both contribute to syllable weight, they do not do so equally, but the weight of a vowel (2) is double that of a consonant (1). Supporting a distinction between vowels and consonants, Mase (1973) finds that only vowel moras function as tone-bearing units in West Greenlandic.[5]

However, Jacobsen (2000) does not find significant duration differences between (C)VC and (C)VV syllables. Instead, one of her two speakers systematically distinguished three degrees of syllable weight in terms of duration—light / short: (C)V, heavy / long: (C)VC, (C)VV and extra-heavy / overlong: (C)VVC—while the other speaker only showed a bipartite system with the distinction between long and overlong not being significant. Jacobsen attributes this neutralization to the speaker's more extreme sub-phonemic variation in duration shortening the overlong (C)VVC (see section 5.3) and concludes that three syllable weights need to be distinguished, fitting a conventional moraic description that assigns the same weight to (C)VC and (C)VV.[6]

[4] We use the notation of VV representing long vowels and diphthongs purely for convenience here and below; that is we do not mean that long vowels are double vowels that should be represented as a sequence of two short vowels.

[5] He notes one possible exception: Nasals may form a tone-bearing unit together with a preceding vowel.

[6] Note that due to the difficulty of locating syllable boundaries within geminate consonants,

In contrast to Jacobsen's assertion of three categories of syllable weight for West Greenlandic, Pigott (2012, pp. 108–117) concludes that only two categories need to be distinguished for Labrador Inuttut when syllables in phrase-final position, which receive extra lengthening, are disregarded. The two categories are determined by vowel length, i.e. Pigott classifies V, CV and CVC syllables as short and VV, CVV and CVVC syllables as long. The mean durations he gives for long syllables are around double of those for short syllables with the same vowel quality, and he states that the durational ranges of the two categories do not overlap in most of his data, though some overlap appeared especially for closed syllables.

An acoustic investigation of the issue of stress, syllable weight and pitch with a production experiment appears in Jacobsen (2000). She finds that while the different syllable structures are reliably reflected in durational distinctions, durational and tonal patterns do not show significant co-variance in West Greenlandic. Instead she finds that the syllable containing the antepenultimate mora indeed had consistently higher f0 than the surrounding syllables, followed by an f0 drop making it sound prominent. However, the duration of this syllable was significantly shorter than that of the preceding syllable for one of the speakers and did not differ significantly for the other speaker. Compared to the following penultimate syllable, which had consistently low f0, the antepenultimate was significantly shorter for both speakers. The final syllable had consistently high f0 in one speaker's productions, but did not differ significantly from the penultimate in duration, whereas for the other speaker, it was significantly shorter and f0 stayed on a low level (presumably due to final lowering of the last H in HLH, also see Rischel 1974, Fortescue 1984). There were also no consistent tonal patterns associated with differences in syllable weight. Jacobsen (2000) therefore concludes that there is no autonomous category of stress in West Greenlandic phonology.[7]

Similarly, Pigott (2012) and Rose et al. (2012) investigate f0, intensity and duration in Labrador Inuttut and find no evidence for alternating or regular patterns that would point to metrical conditioning. Instead, Pigott (2012) observes marking of prosodic phrase boundaries that manifested in regular f0 patterns analyzable as boundary tones, though he does not specify tonal targets. Additionally, he observed durational lengthening and aspiration of phrase-final plosives in final position, in addition to systematic effects of phonological vowel length on syllable duration. By contrast, no systematic patterns appeared for in-

Jacobsen (2000) measured and compared units from vowel onset to vowel onset, i.e. including syllable onsets with the preceding syllable rhyme. However, all rhymes included in the comparison were measured together with the same following onset consonant, /k/, and were additionally controlled for position in the word.

[7]Recall that the use of the term 'syllable' is a simplification here, as Jacobsen (2000) measured units consisting of syllable rhymes plus the onset of the following syllable. Test words were chosen so that comparisons across positions in the word could be made for units with the same segmental content. Jacobsen (2000) additionally provides measurements for vowels only, which confirmed the significantly higher f0 for the antepenultimate compared to the penultimate. Regarding duration, one speaker had a significantly longer penultimate than final vowel, whereas all other differences were insignificant.

tensity, with intensity peaks appearing in varying locations and with differences between syllables often being small (also see Rose et al. 2012). These results resemble an earlier acoustic study by Massenet (1978, 1980) on the speech of speakers from Inukjuak, Quebec, living in Qausuittuq (Resolute), Nunavut. He finds regular phrase-final f0 patterns and durational lengthening marking utterance type, but no systematic variation in intensity, concluding that this variety has a 'musical accent', but no 'intensity accent'.

Based on auditory analysis of twelve Inuit dialects, with tape recordings of at least one speaker per dialect, Fortescue (1983) concludes that while there is considerable variation in intonation—notably in the location of phrase-final falling pitch, the presence vs. absence of a following rise and a contrast between mora-based and syllable-based systems—, stress does not seem to be part of the prosodic system of any Inuit variety. Instead, he assumes that the auditory impression of stress that may arise is based on a coincidence of a heavy syllable and a phrase-final pitch peak, though "a limited degree of extra stress may be utilised for discursive contrastive effect in some varieties" (Fortescue 1983, p. 115).

1.4 Prosody of South Baffin Inuktitut

Word- and phrase-level intonational (tonal) patterns of South Baffin Inuktitut were analyzed in Arnhold et al. (2018), based on the same corpus of dialogues used in the acoustic analysis presented in the current chapter. The principal findings of our previous study were the identification of two prosodic contours, one associated with 'prosodic word'-level domains (corresponding to orthographic words) and one with 'intonational phrase'-level domains (corresponding roughly to sentences or utterances composed of more than one word-level unit). Figure 1 illustrates this analysis with an extract consisting of the three intonational phrases in (4). The first (4-a) and third (4-c) are spoken by the male, the second (4-b) by the female speaker. Each phrase consists of two or three words, which are numbered for better readability in the figure due to their length.

(4) a. Pani-ga taku-qqau-viuk?
daughter-1SG.POSS see-REC.PAST-INTER.2SG.3SG
'Have you seen my daughter?'

b. Kisiani taku-qqau-jara irni-ra.
however see-REC.PAST-DEC.1SG.3SG son-1SG.POSS
'I have only seen my son.'

c. Asu, taku-qqau-nngit-tait pani-ga.
I.see see-REC.PAST-NEG-DEC.2SG.3SG daughter-1SG.POSS
'I see, so you haven't seen my daughter.'

At the word-level, Arnhold et al. (2018) found two tones, H and L, creating a pattern of falling pitch throughout the prosodic word. The H tone was typically realized close to the beginning of the word, aligned with the first or second

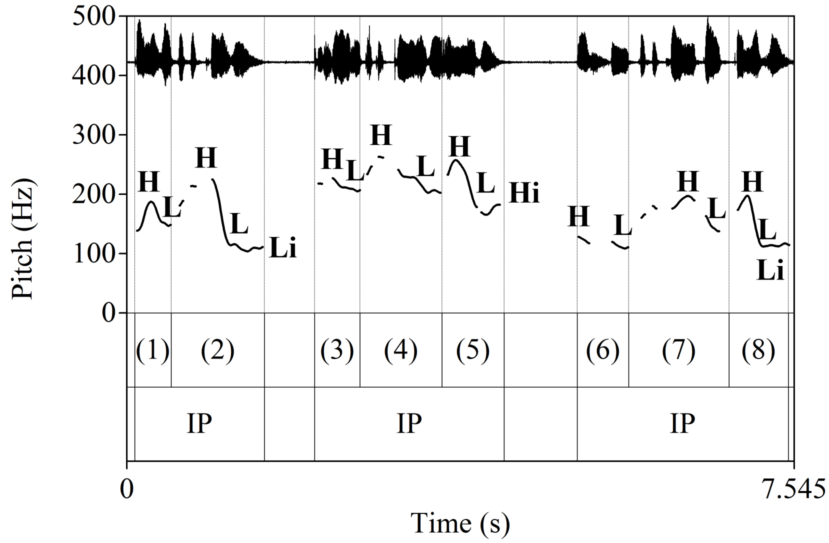

Figure 1: Annotated extract from dialogues corpus, consisting of three intonational phrases (marked IP): *(1) Paniga (2) takuqqauviuk?* 'Have you seen my daughter?', *(3) Kisiani (4) takuqqaujara (5) irnira.* 'I have only seen my son.', *(6) Asu, (7) takuqqaunngittait (8) paniga.* 'I see, so you haven't seen my daughter.', see (4) for glosses.

syllable, though in some cases, the H tone occurred later in the word, as late as the penultimate syllable. The L tone was much less variable; always following the H tone and almost always aligning with the final syllable. Noting only a single exception to this pattern (a word with a rising, rather than falling, f0 contour), Arnhold et al. concluded that this HL contour was indicative of a prosodic unit marking the prosodic word.

In addition, Arnhold et al. identify a larger prosodic domain, demarcated by prosodic pauses in the dialogue corpus, and which they identify as intonational phrase-level domains. They propose that the final (pre-pausal) words in such domains are marked at their right edge with an L boundary tone (marked Li in Figure 1). This boundary tone combines with the prosodic word HL contour with the effect that in such intonational phrase-final words, the L from the HL contour occurs slightly earlier, resulting in an extended stretch of low pitch, usually across the final two syllables, as seen in the first and last phrase in Figure 1. The middle phrase shows a small rise at the end instead, marked with a high boundary tone (Hi), which occurred much less frequently (the possible meaning associated with the difference between Li and Hi requires further study).

The identification of these two intonational patterns, pitch falls and following low stretches, lead Arnhold et al. to posit two levels of prosodic domains, which they identify as the prosodic word and the intonational phrase. They remark that while these domains are regularly marked in terms of intonation, they do not find any evidence of domains smaller than the prosodic word, nor of any domains intermediate between the prosodic word and intonational phrase (i.e. the phonological phrase).

Importantly, as in other Inuit varieties (see overview in Arnhold accepted), South Baffin Inuktitut intonation is highly regular, with a small set of pitch contours marking the edges of prosodic constituents. Note that work on Inuit prosody, including our own, has concentrated on the marking of the right edges of prosodic constituents, i.e. on word- or phrase-final phenomena (but see Arnhold 2014, for a tentative suggestion of a low tone associated with the left edge of the prosodic word in West Greenlandic). Therefore, when we refer to edge marking here and in the following, we generally intend reference right edges, though we certainly think marking of left edges, while underexplored, is possible.

2 Stress: Possible patterns and hypotheses

In this section, we provide an overview of various logically-possible hypotheses regarding the phonological patterns of stress placement in Inuktitut, which will be tested against the results of the phonetic study provided in section 4. The basic rationale of our phonetic study is this: If certain syllables reliably differ significantly from their neighboring syllables in terms of acoustics—prototypically, but not necessarily, by having higher fundamental frequency, longer duration and higher intensity—these syllables are consistently acoustically prominent. This would speak for these syllables bearing stress. The argument for stress could further be strengthened if emerging acoustic stress patterns were in line with patterns observed in the existing phonological and typological literature (e.g. Bakovic 1998, Gordon 2002, Hayes 1995, van der Hulst 2014).

Conversely, if no consistent patterns emerge there is no evidence for systematic word-level prominence. In this case, other types of phonological or phonetic evidence would be needed to argue for the relevance of stress in the grammar of Inuktitut. Should emerging patterns be reliable, but contradict previous descriptions of possible stress patterns, further discussion is required.

There is broad agreement that Inuit does not have contrastive stress and no minimal pairs have been reported. Therefore, we will concentrate on fixed and rhythmic metrical stress as the most likely possibility. In delineating possible stress patterns, we will take as a starting point the parameters for metrical stress languages, as outlined in Hayes (1995).

2.1 Culminativity: Placement of Primary Stress

Languages with metrical stress systems are expected to show one syllable in every prosodic word with a higher degree of prominence as compared to the other syllables in the word domain (Hayes 1995, p. 24–25). The notion of culminativity has received slightly different interpretations, with several authors interpreting it to mean that each word has maximally one most prominent syllable (e.g. Kager 2007) and distinguishing it from obligatoriness, the requirement that each word has at least one prominent syllable (e.g. Hyman 2006, 2009, who takes obligatoriness, but not culminativity to be the defining criterion of a stress language). In a prototypical stress language, each word has minimally and maximally, i.e. exactly, one syllable of highest prominence. This prominent syllable is the locus of primary word stress.

Phonetically, we may expect this highest prominence to be marked through some combination of the phonetic cues of pitch (f0), intensity and duration (see Gordon & Roettger 2017, for a recent cross-linguistic overview of stress cues).[8] Thus, we can hypothesize that if South Baffin Island Inuktitut has a metrical stress system, one syllable of each word will significantly differ from all other syllables with respect to one or more of these acoustic measures. While the most prototypical realization of prominence would be high pitch, high intensity and long duration, divergent cues to prominence are possible, for example reliably lower, rather than higher pitch, on stressed syllables compared to neighboring ones (see e.g. Chung 1983, Williams 1982, 1985).[9]

[8]Like most studies surveyed by Gordon & Roettger (2017), we concentrate on these three potential stress cues. Thus, we omit spectral measures, which correlate with voice quality (spectral tilt) and vowel quality (formants). There is no acoustic investigations of voice quality in Inuit and linguistic descriptions do not mention notable variations in voice quality. Similarly, the existing literature gives no hint that vowel quality could be a potential stress correlate in Inuit like it is e.g. in English or Russian: Inuit contrasts only three vowel qualities and there are no reports of a further reduction of this contrastive set in certain positions. Also, while there is expectably a lot of allophonic variation, it is mostly conditioned by the consonantal context (see Hagerup 2011 and overview in Arnhold accepted).

[9]We thank an anonymous reviewer for pointing us towards these references.

In line with Hayes's (1995) parametric setting of End Rule Left/Right and Optimality Theoretic alignment constraints, this primary stress is expected near the left or right word boundary.[10] Given that the existing literature on Inuit consistently reports word-final pitch contours, the right edge might be more likely.

2.2 Secondary stress: Bounded vs. unbounded systems

I

The primarily stressed syllable may be the only stress in a word or other, secondary, stresses may occur. The former is described as unbounded stress systems in a metrical account. Phonologically, unbounded stress systems include those where there is a single foot within the word, correlating with a single syllable demarcated for stress assignment.[11] Traditionally, the term "unbounded" derives from the idea that there is no limit on the size of the foot—the foot may include a single syllable, or as many syllables as are contained within the prosodic word (see e.g. Prince 1985, Hayes 1995). More recent work on unbounded stress systems analyses these as utilizing non-iterative binary feet (Bakovic 1998, McCarthy 2003).

The latter type of system, which has secondary stress in addition to primary (word-level) stress, is described as a bounded, iterative stress system in metrical theory. In a bounded metrical system, we predict that there will be multiple feet built across the word, where the upper size of the foot will be limited to two syllables or two moras.[12] Secondary stress is expected to fall on the head syllable of eachfoot; depending on the rhythmic type of the foot, i.e. trochee vs. iamb, this may be either the initial or final syllable of that foot.

Thus, we predict to see some evidence of secondary stress in alternating positions. Phonetically, we expect this to be marked similarly to the placement of primary stress, using suprasegmental features like pitch or intensity, although rhythmic adjustments to syllable structure, like vowel lengthening or consonant gemination, may also be interpreted as stress, as in other related languages (cf. discussion in section 5.3).

2.3 Quantity-sensitivity

Another factor which may come into play is quantity-sensitivity. Quantity-sensitive languages create a dichotomy between heavy and light syllables. This

[10] Both binary feet and those of unlimited size are designated as left- or right-headed, such that primary stress will occur either on the leftmost (word-initial) syllable or on the rightmost (word-final) syllable, resulting in regular initial stress or regular final stress. Other parameters could result in misalignment of stress from the edge of the word; for example, right-headed feet and final extrametricality would result in a language where stress regularly falls on the penultimate syllable of the word.

[11] Footless accounts of stress also exist, e.g. Gordon (2002). While we couch our predictions in a foot-based here, this decision is not crucial to our analysis.

[12] Hayes's foot inventory excludes quantity-insensitive bimoraic feet due to a requirement of syllable integrity (Hayes 1995, p. 121–124), but we consider the option here to maximize the chance of detecting stress in Inuktitut.

contrast can be modeled using Moraic Theory (Hayes 1989) by assigning moras (units of weight, see e.g. McCawley 1978, Yoshiba 1983) to segments occupying positions in the rhyme (nucleus and coda); onset consonants typically do not contribute to the weight of the syllable (though cf. Ryan 2014). Thus, in languages which contrast phonemically long and short vowels, syllables with long vowels are considered heavy (bimoraic) and short vowels are light (monomoraic). Languages differ with respect to whether or not coda consonants count toward the weight of the syllable. In languages where coda consonants are moraic, syllables with VC rhymes often pattern as heavy alongside syllables with VV rhymes. More rarely, languages may make a three-way contrast between light (V), heavy (VC/VV) and superheavy syllables (VVC) (Gordon 1999, Morén 1999, Rosenthall & van der Hulst 1999).

Because Inuktitut is a language which contrasts both long and short consonants and vowels, as well as a number of rhyme types, we may expect to find that weight plays a role in stress assignment, with heavy syllables attracting stress. There are several ways in which quantity-sensitivity may play a role in determining word-level prominence, depending on whether the language shows a bounded or unbounded stress system.

If Inuktitut shows a bounded, iterative stress system (i.e. multiple binary feet per word) as well as quantity-sensitivity, we would expect to find that the presence of heavy syllables would result in deviations from a strictly alternating pattern of stress assignment. Specifically, we may occasionally see evidence of stress clash (two adjacent stressed heavy syllables) or stress lapse (two adjacent unstressed light syllables), rather than regularly alternating patterns of prominence.[13].

Conversely, in an unbounded system, heavy syllables may attract stress away from the preferred word edge, resulting in stress that is sometimes distant from the position in which stress falls in words with only light syllables (where stress would predictably fall at a single word edge). In words with multiple heavy syllables, we expect to see a preference for such syllables to attract stress depending on their position in the word.

In order to determine whether Inuktitut possesses a quantity-sensitive unbounded system, we will compare words with only light syllables to words with one or more heavy syllables, in order to determine whether the presence of heavy syllables causes deviations from the patterns observed for words without heavy syllables. An additional factor which may come into play is the possibility of a three-way weight system distinguishing between light (V), heavy (VV, VC) and extra-heavy (VVC) rhymes (as apparent in West Greenlandic syllable durations, cf. Jacobsen 2000 and discussion above). In order to test this, we will need to conduct an additional comparison between words without any extra-heavy syllables and words with one or more extra-heavy syllables.

If Inuktitut possess a quantity-sensitive stress system, we can hypothesize that the distribution of acoustic stress cues differs between subsets of the data

[13]Provided that both primary and secondary stress are quantity-sensitive. If only primary stress is weight-sensitive, clashes or lapses are not necessarily expected, but heavy syllables should attract the primary stress as in an unbounded system.

containing only light syllables and subsets also containing heavy or superheavy syllables. Thereby, in the former subsets, heavy syllables should be significantly more likely to have acoustic characteristics pointing towards prominence than light syllables. In our statistical analyses, these differences should manifest themselves in a main effect of syllable weight or, more likely, in an interaction between syllable weight and position. In the latter type of subsets, where all syllables have the same weight, acoustically-distinguished syllables should reliably appear in the same positions, which should significantly differ from neighboring positions, i.e. a main effect of syllable position should appear.

2.4 Phrasal Prominence

As discussed in previous sections, it has previously been proposed that Inuit does not have any phonetic evidence of stress. In our previous work with this data set (Arnhold et al. 2018), we have observed that there are clear intonational markers of prosodic word edges. One additional hypothesis that can be tested is the idea that while Inuktitut does mark right domain edges at the prosodic word level, there may not be any evidence for word-internal metrical structure between the level of the syllable and the prosodic word—in other words, that Inuktitut lacks metrical foot structure.

This hypothesis is difficult to separate empirically from the possibility that Inuktitut shows an unbounded, quantity-insensitive system represented either by a single, multi-syllabic foot or a single foot which is aligned with the right edge of the prosodic word.[14] In this case, we may raise the question of prominence: Is there indeed a single syllable contained within the prosodic word which may be considered to be more prominent that all others, as outlined above?

3 Data and annotation

In order to examine the prosodic properties of Inuktitut, we obtained permission from the Pirurvik Centre to analyze recordings posted on their language-learning website, Tusaalanga (meaning ‘Let me hear!’, Pirurvik Centre 2015). The website contains both short dialogues between a female and a male speaker of South Baffin Inuktitut, as well as a glossary of words uttered in isolation.

The data considered in this paper consisted of 19 dialogues (containing 297 orthographic words) and 195 single words chosen from the corpus. Our selection procedure did not take into account the content of the dialogues or words. In this chapter, we will compare the acoustic analysis of data from both contexts with the goal of considering prominence from the perspective of words in isolation and words uttered within a phrasal context.

Note that our data was not specifically gathered for the purpose of the present analysis and therefore has some of the methodological shortcomings

[14] Evidence for feet without stress has been reported for several languages, see e.g. Bennett (2012). For Inuit, the only potential foot-based phenomena concern adjustments of syllable structure, which are however difficult to capture in metrical theory, see section 5.3.

that are common in studies of acoustic correlates of stress (see Roettger & Gordon 2017). Mitigating these shortcomings are the fact that the materials were created by native speakers, the inclusion of two corpora with different genres of speech and relative avoidance of researcher bias.

3.1 Dialogues

We analyzed 19 scripted dialogues between a female and a male native speaker of South Baffin Inuktitut. They consisted of 151 orthographic sentences, which in turn were made up of 297 orthographic word tokens, consisting of a total of 1121 syllable tokens of various types (see Tables 2 and 3 for distributions). The 297 words consisted of 216 unique word tokens, with most of them appearing only once.[15]

Table 2: Distribution of words in dialogue data by number of syllables.

Number of syllables	1	2	3	4	5	6	7	8	9
Number of word tokens	15	62	77	58	32	27	13	11	2

Table 3: Distribution of dialogue data by syllable types.

Syllable type	CV	CVC	CVV	CVVC	V	VC	VV	VVC
Number of syllables	527	216	155	91	46	50	22	14

Two research assistants segmented and annotated the dialogue recordings based primarily on the orthographic transcriptions (using a Roman alphabet) and English translations included on the website; additionally, we tagged each speaker for gender (male and female) and added morpheme glosses. Each orthographic word was segmented at the syllable level, as based on phonotactic constraints in the language (see section 1.3),[16] and syllables were then numbered beginning at the end of the word. Two of the authors of this paper annotated the recordings for the occurrence of phrase boundaries (based both on pausal and intonational cues), and marked the realization of high and low pitch targets.

[15] As our focus is on the influence of phonological structure and it is uncontroversial that there is no lexical stress in Inuit, we counted words consisting of the same root with different inflections, e.g. *ataatait* 'your father' and *ataataga* 'my father', as separate unique word tokens.

[16] Note in particular that we assumed that all adjacent vowels are homosyllabic, i.e. form diphthongs.

Table 4: Distribution of single-word utterance data by number of syllables.

Number of syllables	1	2	3	4	5	6	7	8	9
Number of words tokens	16	96	138	62	44	12	10	10	2

Table 5: Distribution of single-word utterance data by syllable types.

Syllable type	CV	CVC	CVV	CVVC	V	VC	VV	VVC
Number of syllables	498	371	184	169	12	21	40	35

3.2 Single-word utterances

We also analyzed a subset of the vocabulary list on the Pirurvik Centre (2015) website for learning Inuktitut, which provides audio-recordings together with transcriptions in syllabics and Roman orthography, as well as English translations. Each recording contains a single-word utterance or short phrase spoken twice by a female speaker of South Baffin Inuktitut (one of the speakers in the dialogue data), with a substantial pause between the two productions. Starting from the alphabetically sorted list, two research assistants first downloaded recordings of the first 20 to 22 words each starting with the letters a, i, k, m, n, p, q, s, t, and u. We chose these letters since words beginning with other letters are always loanwords. Restricting our data set to single-word utterances, we retained 195 unique words in two repetitions, i.e. 390 single-word utterances.[17] These words consisted of between 1 and 9 syllables, with trisyllabic words constituting the largest group by far (Table 4), making up a total of 1330 syllables.

4 Results

To test for acoustic correlates of stress or word-level prominence more generally, we analyzed the duration, fundamental frequency (f0; the acoustic correlate of pitch) and intensity of syllables in all words in the two annotated data sets. We were particularly interested in whether these measures were influenced by the syllable's position in the word, displaying a single prominent syllable pointing towards an unbounded stress system or rhythmical alternation of prominent and less prominent syllables that would establish the presence of a bounded, quantity-insensitive stress system. Moreover, looking for cues of a quantity-sensitive stress system, we separately inspected subsets of the data without

[17]Note that this includes two words which appeared twice in the vocabulary list, with different English translations: *aggaak* 'hands/gloves' and *airaapik* 'brother-in-law/sister-in-law'. Each of the two separate entries were associated with distinct sound files, so that our corpus contains a total of four repetitions of each of these two words (two sound files with two repetitions).

(extra-)heavy syllables. To the same end, we also investigated whether rhyme type influenced the acoustic measures in a way that would indicate heavy syllables attracting stress, and especially whether rhyme type interacted with syllable position in determining the acoustic prominence of a syllable.

The three sections on duration (section 4.2), f0 (section 4.3) and intensity (section 4.4) each start with a general description of results consistent for both data sets, before presenting detailed results of statistical analyses in two separate subsections. The following section outlines the procedure used for all statistical analyses.

4.1 Statistical analyses

For statistical analyses, we chose a subset of the data where a potential interaction between rhyme type and syllable position could be tested. For the dialogues data, we omitted the initial syllables of the two nine-syllable-words from the analysis to achieve this goal, resulting in a set of 1119 syllables. For the single-word utterances, testing the interaction was possible when evaluating the last six syllables in all words (1294 syllables, 97% of the complete data set).

Statistical results reported below are based on linear mixed-effects modelling (Baayen et al. 2008) as implemented in the package lme4 in R (Bates et al. 2015, R Core Team 2018). We fitted separate models for the three reported dependent measures: (1) syllable durations in milliseconds (ms), (2) f0 maximum measured during the syllable's vowel nucleus, measured in Hertz (Hz) and converted to semitones (st) relative to a reference value of 100 Hz and (3) mean intensity of the vowel in decibel (dB), measured over the mid 50% of the vowel's duration to exclude effects of neighboring consonants. We tested whether these dependent variables were influenced by the following predictor variables: (1) rhyme type (V, VC, VV or VVC), (2) position of the syllable in the word and (3) word length in number of syllables. We counted syllable positions from the end of the word, as previous research on Inuit describes prosodic patterns aligned to the right, but not the left edge of the word (cf. section 1.4). For the dialogue data, we additionally included the word's position in the intonational phrase as a predictor, distinguishing phrase-initial, phrase-medial and phrase-final words, as well as words that were the only word in the phrase.

The model fitting always proceeded the same way, starting with a model containing all predictors, as well as an interaction between syllable position and rhyme type. In addition to these predictors (the fixed-effects structure), linear-mixed effects models also have a random structure, which allows to take into account the relationship between measurements coming for example from the same speaker, thus modelling inherent (random) variation between speakers. For the initial model, we always specified random intercepts to account for inherent variation between the two speakers in the dialogues data and between vowel qualities (/a/, /i/, /u/)[18] for both data sets. We also tested a more

[18] For diphthongs, we specified vowel quality as that of the first part in accordance with cross-dialectal patterns reported by Dorais (1986, p. 32).

complex random effects structure with by-speaker or by-vowel quality effects of the predictors, e.g. assuming that word length affected different speakers or vowel qualities differently. However, such models rarely converged, meaning the size and composition of the data sets did not support such complexity. Where these models did converge, they did not show a significantly improved fit to the data, as determined by performing an ANOVA comparing a model with the initial random structure to a more complex model (see Matuschek et al. 2017, on principles of model comparison and determining the best linear mixed-effects structure). Thus, all models reported here contain only simple random effects. For models of the dialogue data, we tested whether the random structure could be further simplified by removing one of the two intercepts without decreasing model fit. Where this was possible, it is reported for the resulting model below.

Once the optimal random-effects structure was determined, we tested the fixed effects structure in the same way. First, we checked whether the interaction between syllable position and rhyme type could be removed without decreasing model fit. Next, we tested whether the predictors number of syllables and, for the dialogue data, the word's position in the intonation phrase significantly contributed to model fit and removed them where they did not. Finally, we removed data points with residuals of more than 2.5 standard deviations from 0 and refit the model to ensure that effects were not driven by outlier values. We report the number of removed data points for each model below. To assess significance, we used the package emmeans (Lenth 2018) to perform pairwise comparisons. Statements below regarding significant differences are based on these comparisons. Model summaries and the results of pairwise comparisons (significant differences only) appear in 6.

Note that our statistical models and data selection were designed for hypothesis testing, specifically for assessing potential main effects of all tested predictors and a potential interaction between syllable position and rhyme type. Thus, while there may be interactions between factors like a syllable's position in the word and the number of syllables in the word or between the syllable's position in the word and the word's position in the intonational phrase, such interactions were not tested here, as our goal was not to comprehensively model prosodic variation in the data.

4.2 Duration

Syllable durations were influenced by the position of the syllable in the word, the number of syllables in the word, and by rhyme type. As would be expected based on the phonological length distinction and in line with previous research on other Inuit varieties (cf. section 1.3), syllables containing long vowels or more segments had longer durations than those containing fewer segments and only short vowels. Thus, syllables with rhymes consisting of a single short vowel showed the shortest duration, while durations were longest for (C)VVC syllables in both the dialogues and the single-word utterances (see Table 6).[19]

[19]Note that while mean duration was longer for (C)VC than for (C)VV syllables in the single-word utterance data set, the opposite was true in the dialogues. However, this apparent

Table 6: Mean and standard deviation (SD) of syllable duration (in ms) by rhyme type for dialogue data set (dialogues) and single-word utterance data set (words).

Rhyme type	Mean, dialogues	SD, dialogues	Mean, words	SD, words
V	175	82	177	58
VC	217	67	324	78
VV	264	93	281	90
VVC	286	68	382	101

Table 7: Mean and standard deviation (SD) of syllable duration (in ms) by number of syllables in the word for dialogue data set (dialogues) and single-word utterance data set (words).

Number of syllables	Mean, dialogues	SD, dialogues	Mean, words	SD, words
1	437	89	533	61
2	260	123	343	118
3	215	92	284	105
4	204	80	258	96
5	196	69	240	97
6	191	68	224	83
7	185	59	199	65
8	189	63	200	61
9	179	37	185	64

Additionally, both data sets showed a tendency towards isochrony, with words with fewer syllables showing longer syllable durations than those with more syllables, which was especially noticeable for monosyllabic words (see Table 7).

Regarding syllable position, Figure 2 shows that syllable durations increased towards the end of the word, and especially for the last syllable. This word-final lengthening effect appeared clearly in both data sets, though its magnitude was larger for the single-word utterances, likely because all words were utterance-final (cf. the findings on phrase-final lengthening in the dialogues data in section 4.2.1). Notably, neither data set exhibited an alternation between longer and shorter syllables, as would be expected for a (quantity-insensitive) bounded stress system where duration is a cue to stress. Instead, except for the final

contrast is likely due to the lack of balance in the data set. Statistical analyses presented in section 4.2.1 indicated that when the influence of other factors was taken into account, durations of (C)VC syllables were consistently shorter than those of (C)VV syllables, though in the single-word utterance data set, the difference only reached significance in word-final position.

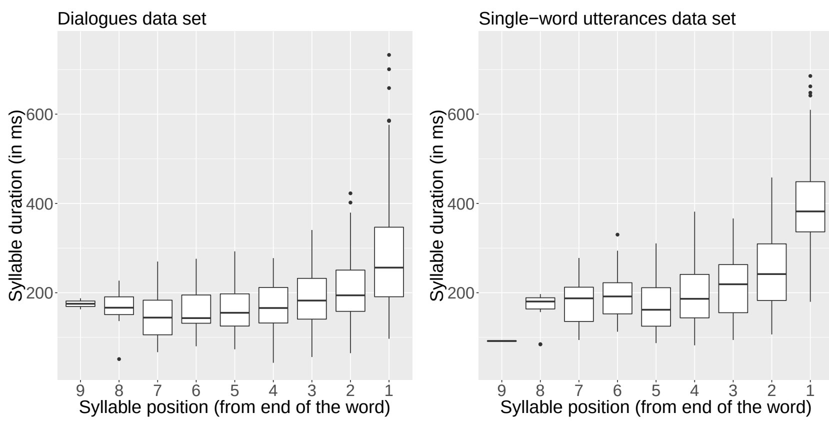

Figure 2: Syllable duration (in ms) by syllable position, i.e. syllable number as counted from the end of the word, for the whole dialogues data set (left panel) and the whole single-word utterances data set (right panel).

lengthening, syllable duration was similar across different positions.

To see whether the duration patterns matched a quantity-sensitive bounded stress system, we further checked whether differences between syllable positions were affected by syllable weight. In a quantity-sensitive system, heavy syllables, where present, would be expected to attract stress. In the absence of heavy syllables, regular alternations between lengthened and shortened syllables should emerge if stress is acoustically cued by duration. Figure 3 illustrates durations for words consisting only of (C)V syllables, which would doubtlessly count as light in a quantity-sensitive stress system. Note, however, that the figure represents an extremely small subset of the data, as only 34 words in the dialogues and only two single-word utterances fit this criterion. No regular alternation between lengthened and shortened syllables emerged. In particular, while the last four syllables of the single-word data may suggest a rhythmic pattern, with the antepenultimate and final syllables being longer than the syllables directly preceding them, this pattern was not present in the dialogue subset, where duration successively increased for the last three syllables. Figure 4 represents larger subsets of the data, which only excluded words containing syllables with VVC rhymes, which would clearly count as (extra-)heavy and thus stress-attracting in any quantity-sensitive stress system. Here, the distribution of syllable durations by position in the word looked very similar to the complete data sets (cf. Figure 2), showing final lengthening, but no other effects of syllable position. The next two subsections test statistically whether there was an interaction between syllable position and rhyme type in predicting syllable durations.

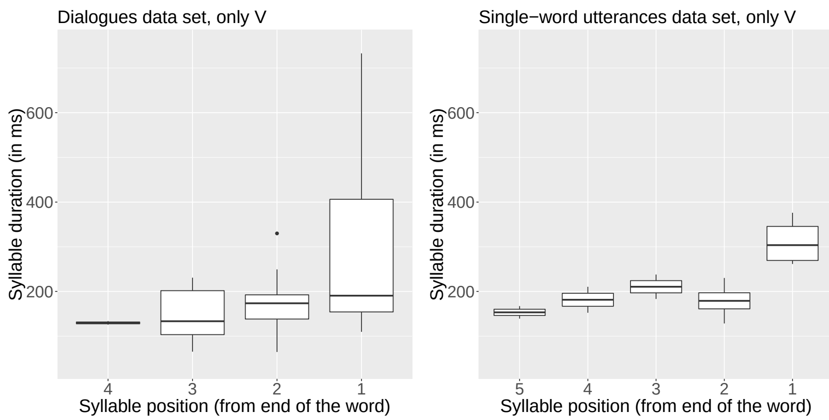

Figure 3: Syllable duration (in ms) by syllable position, i.e. syllable number as counted from the end of the word, for the 34 words in dialogues data set (left panel) and the two words in the single-word utterances data set (right panel) containing only (C)V syllables.

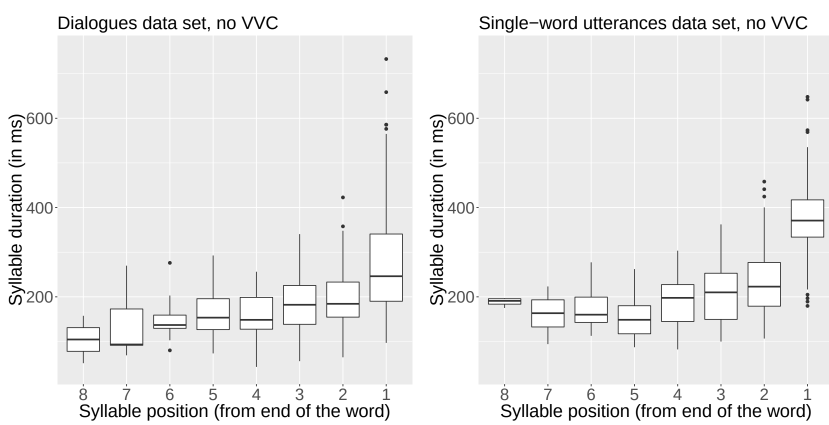

Figure 4: Syllable duration (in ms) by syllable position, i.e. syllable number as counted from the end of the word, for the 197 words in the dialogues data set (left panel) and the 99 words in the single-word utterances data set (right panel) not containing (C)VVC syllables.

4.2.1 Statistical analysis: Dialogues

The best linear mixed-effects model of syllable duration contained the predictors syllable position as counted from the end of the word, word length counted as number of syllables, the position of the word in the intonational phrase and rhyme type as significant predictors (24 outlier data points were trimmed, i.e. 2.1% of the data; see Table 8 in the Appendix). The interaction between syllable position and rhyme type was not significant, and was omitted from the final model as it did not improve model fit.

Results showed a clear effect of final lengthening, with syllables closer to the end of the word having longer durations. This effect was significant for the last three syllables, which had significantly longer durations than all preceding syllables, and additionally contrasted amongst each other with the ultima being significantly longer than both penultima and antepenultima, and the penultima being significantly longer than the antepenultima (see Table 9).

Regarding the effect of word length, syllable durations were significantly longer in monosyllabic words than in all other word lengths, while words consisting of two or more syllables did not differ significantly from each other (Table 10).

The effect of rhyme type was also persistent in spite of the uneven distribution of syllable types (recall Table 3). Pairwise comparisons showed significant differences between all different rhyme types, indicating a durational ranking of (C)V<(C)VC<(C)VV<(C)VVC (Table 11).

Finally, the model included a significant effect of the word's position in the intonational phrase: Syllable duration was significantly shorter for words in phrase-initial or -medial position than for phrase-final words and that constituted phrases on their own (i.e. that were the only words in single-word phrases), see Table 12. Though it seems likely that phrase-final lengthening especially affected word-final syllables, the interaction between position in the phrase and syllable number could not be tested in this data set due to gaps in the distribution.

4.2.2 Statistical analysis: Single-word utterances

The best statistical model (32 outliers removed, i.e. 2.5% of the data; see Table 13) contained the same effects as for the dialogues data: Syllables in monosyllabic words were again significantly longer than syllables in longer words, while words with two or more syllables did not differ significantly from each other (Table 14). The model also contained significant main effects of syllable position, pointing again towards longer durations later in the word, and of rhyme type, indicating longer durations for syllables containing more segments or long vowels.

Additionally, the model contained an interaction between syllable type and syllable position. This interaction meant that while the effects of syllable position and rhyme type were consistent, they did not reach significance in all pairwise comparisons, which could be due to the imbalance in the data set (cf. section 3.2). Thus, word-final syllables had significantly longer durations than all preceding syllables for all rhyme types, while the longer duration of the penultimate only reached significance compared to syllables 3, 4 and 5 as counted from the end of the word for V rhymes, compared to syllable 5 when comparing VC rhymes, compared to syllables 4, 5 and 6 for VV rhymes and compared to syllable 3 and 5 for VVC rhymes (Table 15). Comparisons between preceding syllables never reached significance.

Pairwise comparisons between all rhyme types reached significance in word-final position, with (C)V syllables being shorter than (C)VC, (C)VV and (C)VVC syllables, (C)VC being significantly longer than (C)V, but shorter than (C)VV and (C)VVC, and (C)VV syllables were shorter than (C)VVC ones, but significantly longer than all others (Table 16). These same contrasts, except for the one between syllables with VC and VV rhymes, were significant in penultimate position, in the antepenultimate and in the fourth syllable as counted from the end of the word. Finally, the contrasts between (C)V and (C)VVC and between (C)VC and (C)VVC were still significant in the fifth and sixth syllable from the end of the word, and for the fifth syllable, the contrast between (C)V and (C)VV additionally reached significance.

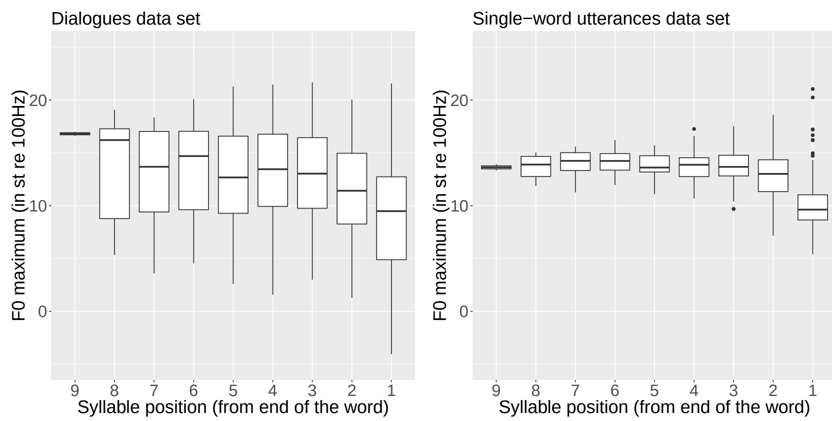

Figure 5: F0 maximum (in semitones relative to a reference value of 100 Hz) by syllable position, i.e. syllable number as counted from the end of the word, for the whole dialogues data set (left panel) and the whole single-word utterances data set (right panel).

4.3 Fundamental frequency

Figure 5 displays the highest f0 value measured for vowels in all syllable positions. The dialogues data, which included more varied utterances from two speakers, showed broader distributions than the single-word utterances, but the overall pattern was very similar in both: The f0 maximum stayed at about the same level for most of the word before decreasing for the penultimate and especially the last syllable. Neither set showed an alternation between syllables with higher and lower peaks that would indicate a rhythmic stress pattern or a syllable with a higher f0 peak than both neighbouring syllables as expected for a primary stress location marked by a pitch accent.

The size of the word-final f0 drop from the penultimate to the last syllable was larger for syllables with short vowels than for those with long ones: In the dialogues, the average f0 maximum in the antepenultimate syllable was14 semitones (st) for all rhyme types. For the penultimate syllable, the average was 11 st for V rhymes, and 12 st for VC and VVC rhymes and 13 st for VV, while in the last syllable, it was 9 st for V and 8 st for VC, compared to 12 st for both VV and VVC rhymes. In the single-word utterances, the average for the antepenultimate was 14 st and for the penultimate syllable it was 13 st for all rhyme types, whereas for the final syllable, it was 9 st with V and VC rhymes, but 12 st for VV and 11 st for VVC rhymes.[20] In spite of these small differences,

[20]This could potentially indicate that, contrary to the classification by Fortescue (1983), the mora and not the syllable is the tone-bearing unit in South Baffin Inuktitut, so that the final f0 fall takes place within the last syllable if it is bimoraic, but between the penultimate and the final if the last syllable carries only one mora. Further investigations beyond the scope of this chapter would be required to assess this possibility. While we thus leave this issue to be addressed future work, note that it is commonly assumed that stress is assigned to syllables

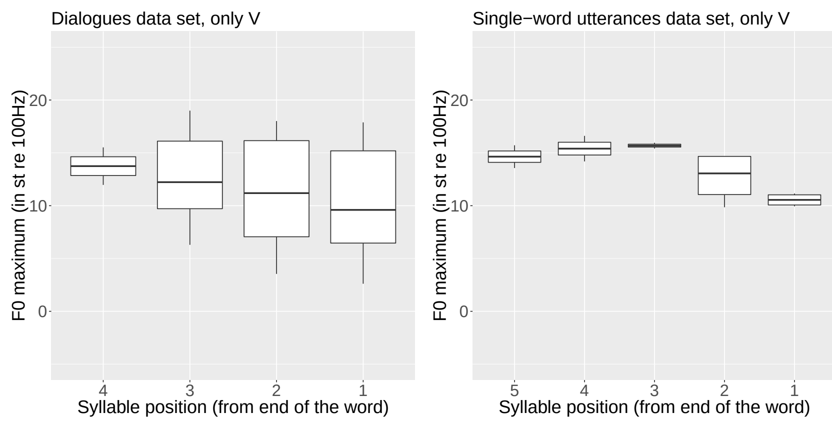

Figure 6: F0 maximum (in semitones relative to a reference value of 100 Hz) by syllable position, i.e. syllable number as counted from the end of the word, for the 34 words in dialogues data set (left panel) and the two words in the single-word utterances data set (right panel) containing only (C)V syllables.

the pattern of relatively level f0 values before a final f0 fall persisted in subsets excluding words with heavy syllables, cf. Figures 6 and 7.

Fundamental frequency was also influenced by word length: Words with fewer syllables, especially monosyllabic ones, had overall higher f0 maxima than longer words, even though the differences were small compared to the word-final f0 drop. In the dialogues data, monosyllabic words had an average f0 maximum of 13 st compared to averages between 10 st and 12 st for all longer words except words with nine syllables, of which there were only two tokens (mean: 14 st). For the single-word utterances, the average for monosyllabic words was 14 st compared to means between 12 st and 13 st for all other word lengths. This might indicate that the size of f0 movements is truncated rather than compressed in shorter words in this Inuit variety: Word-final pitch falls may have a smaller magnitude when they stretch over only one syllable (see Ladd 2008, p. 180–184, for a cross-linguistic overview on truncation vs. compression).

4.3.1 Statistical analysis: Dialogues

The best model of f0 maxima of vowel nuclei in the dialogues contained an interaction between syllable position and rhyme type, as well as significant main effects of word length and the position of the word in the phrase (20 outliers or 1.8% of data removed; random intercept for vowel quality removed). Pairwise comparisons to assess the interaction confirmed only four significant differences between rhyme types when syllable position was controlled: In word-final syllables, vowels in V rhymes had lower f0 maxima than those in VVC rhymes,

and not moras by definition (e.g. Hayes 1995, Hyman 2006, 2009).

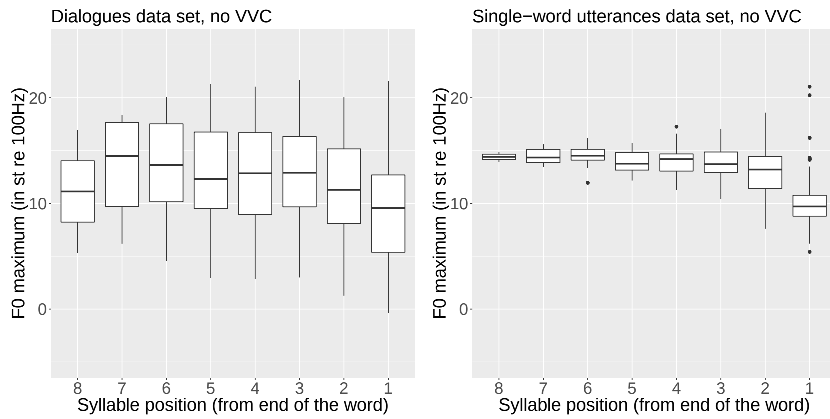

Figure 7: F0 maximum (in semitones relative to a reference value of 100 Hz) by syllable position, i.e. syllable number as counted from the end of the word, for the 197 words in the dialogues data set (left panel) and the 99 words in the single-word utterances data set (right panel) not containing (C)VVC syllables.

and VC rhymes had lower maxima than both VV and VVC (Table 18). In penultimate position, short vowels were lower than long ones in open syllables.

Conversely, pairwise comparisons between the same rhyme types in different positions showed that final syllable vowels were significantly lower than those in all preceding syllables when comparing VC and VV rhymes, and significantly lower than all except syllable 8 as counted from the end of the word for V rhymes (Table 19). Moreover, the penultimate had lower f0 maxima than syllables 3 to 7 as counted from the end of the word for V rhymes and lower maxima than syllable four when comparing VC rhymes. For VVC rhymes, differences in position were not significant. Thus, the model affirmed the word-final f0 drop illustrated in Figures 5 to 7, as well as showing slight differences in its realization depending on the rhymes of the word-final syllables.

Regarding the effect of word length, monosyllabic words had higher f0 maxima than longer words, but this was only significant compared to words with seven or eight syllables (Table 20).The same was true of words with four, five and six syllables, which likewise had higher maxima than those with seven and eight syllables, while disyllabic words had significantly higher values than those with six, seven and eight syllables. Trisyllabic words differed significantly from all words with five or more syllables. As to the last predictor in the model, the word's position in the phrase, medial words had higher f0 than initial and final words and those in one-word phrases (Table 21). Additionally, phrase-final words showed higher f0 when they were not the only word in the phrase.

4.3.2 Statistical analysis: Single-word utterances

For the single-word utterances, the best linear mixed-effects model of f0 maxima included an interaction between syllable position and rhyme type, as well as word length as a predictor (27 data points or 1.7% removed; see Table 22). As for the dialogues data, differences in rhyme types slightly modulated the word-final f0 drop: The final syllable had significantly lower f0 maxima than all preceding syllables in pairwise comparisons within all rhyme types, but comparisons between the penultimate and earlier syllables only consistently reached significance for V and VV rhymes, while VC rhymes were lower in penultimate position compared to positions 3 and 5 and penultimate VVC rhymes only differed from those in antepenultimate position (Table 23).

The additional effect of word length substantiated that mono- and disyllabic words had higher f0 maxima than all longer words, while trisyllabic words differed significantly from words with five, seven and eight syllables and words with four syllables had higher f0 than those with five and eight syllables (Table 24).

4.4 Intensity

As for the other two measures, intensity was clearly affected by the right word edge: It remained relatively stable within the word before dropping word-finally. This drop was more pronounced in the single-word data, where each word constituted its own intonational phrase, than in the dialogues data, which included phrase-initial and -medial words, as well (cf. Figure 8). Importantly, intensity measurements neither displayed alternations between louder and less loud syllables that would indicate iterative foot structure, nor a single syllable constituting a clear and consistent maximum pointing towards the location of primary stress. These patterns persisted also when investigation was restricted to words consisting of only light syllables (see Figure 9) or of light or heavy, but no (C)VVC syllables (Figure 10).

Similar to the findings for f0, the word-final intensity drop was again steeper for syllables with fewer segments and short vowels than for syllables with long vowels and heavier rhymes. In the dialogues data set, the difference in mean intensity between the last and the penultimate syllable position was 2 dB for V and VC rhymes and 1 dB for VV rhymes (VVC rhymes had 1 dB higher average intensity in final than in penultimate position). In the single-word utterance data set, the difference between the final and the penultimate position was 7 dB for V rhymes, 4 dB for VC and VV rhymes and 3 dB for VVC.[21] Even though we did not test this statistically, these numbers suggest that the final drop was generally larger in the single-word utterance data set, where all words were final, than in the dialogues data set. This is line with the significant effect of a word's position in an intonational phrase found within the dialogues data set, where phrase-final words showed lower intensity than initial and medial ones,

[21] Differences between the average intensity of the penultimate and the antepenultimate never exceeded 1 dB in either data set, but appeared in the statistical analyses of the single-word utterance data set (see section 4.4.2).

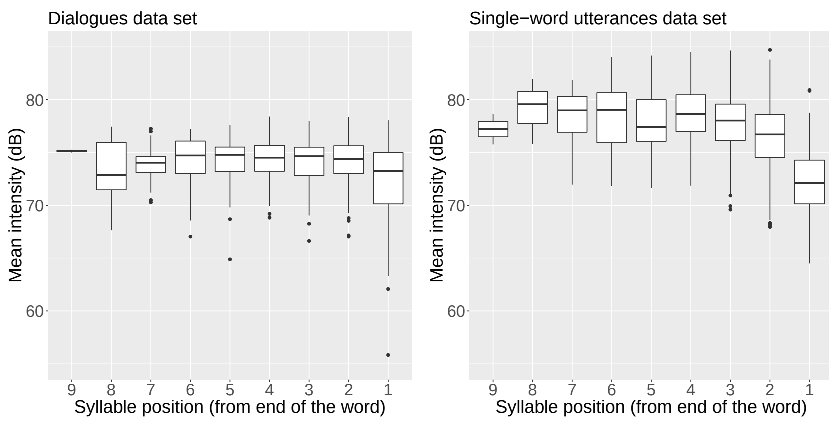

Figure 8: Mean intensity in the centre of the vowel (in dB) by syllable position, i.e. syllable number as counted from the end of the word, for the whole dialogues data set (left panel) and the whole single-word utterances data set (right panel).

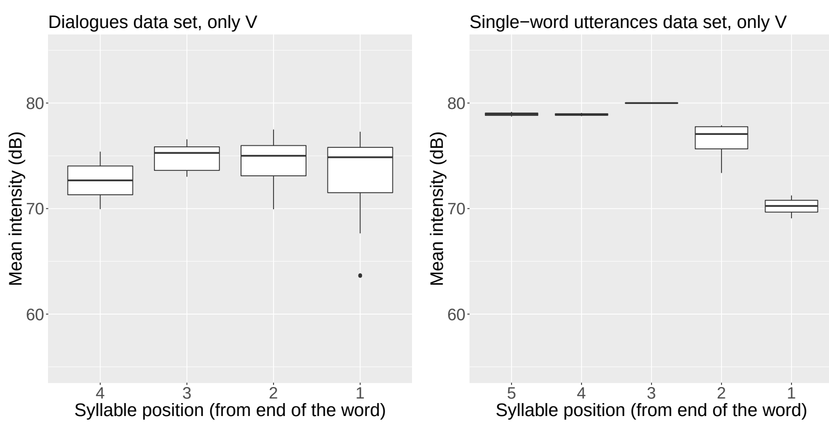

Figure 9: Mean intensity in the centre of the vowel (in dB) by syllable position, i.e. syllable number as counted from the end of the word, for the 34 words in dialogues data set (left panel) and the two words in the single-word utterances data set (right panel) containing only (C)V syllables.

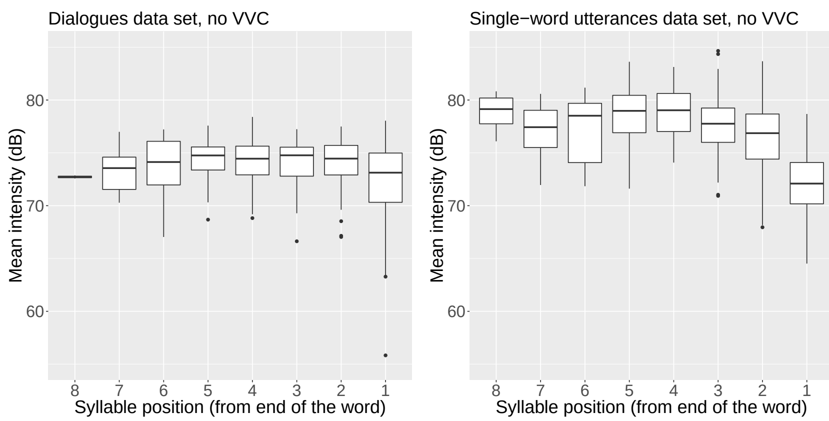

Figure 10: Mean intensity in the centre of the vowel (in dB) by syllable position, i.e. syllable number as counted from the end of the word, for the 197 words in the dialogues data set (left panel) and the 99 words in the single-word utterances data set (right panel) not containing (C)VVC syllables.

cf. section 4.4.1.

Additionally, intensity differed between words of different lengths, but these differences were not systematic across both data sets (see statistical analyses in the next two sections).

4.4.1 Statistical analysis: Dialogues

The best model of mean intensity in the centre of the vowel for the dialogue data contained a significant interaction between syllable position and rhyme type, as well as word length and position in the intonational phrase as additional predictors (21 data points or 1.9% excluded; vowel quality removed from the random-effects structure; see Table 25). The interaction suggested that the word-final intensity drop was not significant when comparing the vowel nuclei of VVC rhymes. By contrast, in comparisons among V rhymes, the word-final position showed significantly lower intensity than syllables 2 to 7 as counted from the end of the word (Table 26). For VC rhymes, intensity of the last syllable differed significantly from that of syllables 2, 4, 5, and 6, while for VV rhymes, intensity of the last syllable differed significantly from that of syllables 3, 4, 6, 7 and 8.

In line with this difference between VVC and other rhyme types regarding the final intensity drop, vowels in V and VC rhymes showed significantly lower intensity than those in VVC rhymes in final syllables (Table 27). Additionally, vowels in VC rhymes had lower intensity than those in VV rhymes in syllables 3and 8 as counted from the end of the word.

The significant effect of word length suggested that in words with two, three and four syllables, intensity was overall higher than in words with seven or eight syllables, and that it was higher in words with five or six syllables than in those with seven (Table 28).

Lastly, intensity was higher in words in phrase-initial and -medial position compared to phrase-final words and those were the only words in a phrase (Table 29). Additionally, single-word phrases had a lower intensity than other phrase-final words.

4.4.2 Statistical analysis: Single-word utterances

For the single-word utterances, the best linear mixed-effects model of intensity had an interaction between syllable position and rhyme type, as well as an effect of word length, as predictors (18 outliers, i.e. 1.4% of the data, removed; see Table 30). Similar to the dialogues data, the interaction revealed differences between rhyme types regarding the word-final intensity drop; however, these concerned the penultimate syllable in this data set.

Vowels in the final syllable had lower intensity than those in all preceding syllables in pairwise comparisons for all rhyme types (Table 31). The penultimate syllable had significantly lower intensity than all preceding syllables only when comparing vowels in VC rhymes. Comparing VV rhymes, intensity in the penultimate was significantly lower than in syllables 3, 4 and 5 as counted from the end of the word. For VVC rhymes, the penultimate had significantly lower intensity than syllables 3 and 4. The penultimate only differed significantly from syllable 4 as counted from the end of the word for V rhymes. In line with this, significant differences between rhyme types were mostly concentrated at the word edge (Table 31): Vowels in V and VC rhymes had significantly lower intensity than in VVC rhymes, with V rhymes additionally showing lower intensity than VC rhymes in final position. Vowels in VC rhymes also had significantly lower intensity than those in VV rhymes in penultimate position, and V rhymes had lower intensity than VV rhymes in third and fifths syllables as counted from the end of the word.

Regarding word length, pairwise comparisons showed significant differences, which differed from those found in the dialogues data: Monosyllabic words had higher intensity than all words with three or more syllables, though this difference was not significant compared to six-syllable words (Table 33). The same was true for disyllabic words. Intensity of three-syllable words was significantly higher than for words with five and nine syllables and words with four, five, seven and eight syllables had higher intensity than those with nine. Intensity in words with six syllables was also significantly higher than in all longer words. By contrast, words with four and five syllables had significantly lower intensity than six-syllable words.

5 Inuktitut and notions of word prominence

Acoustic investigations of two data sets of spoken South Baffin Inuktitut targeted at beginning language learners (Pirurvik Centre 2015), scripted dialogues and single-word utterances, displayed consistent marking of right word edges: Word-finally, syllable duration lengthened in both data sets, accompanied by a drop in fundamental frequency and intensity. Similarly, acoustic cues to a word's position in a prosodic phrase also appeared in the connected speech of the dialogue data: Phrase-final words—whether preceded by other words or forming intonational phrases on their own—had significantly longer durations and lower intensity than initial and medial ones.[22] These findings are in line with our previous analyses of the dialogues data set (see Arnhold et al. 2018), as well as with descriptions of the prosodic systems of other Inuit varieties, which show regular marking of the boundaries of prosodic constituents (see section 1.4).

Crucially for the question of word prominence, no other consistent prosodic patterns over and above this marking of finality appeared in the data analyzed here. Instead, syllable duration, f0 and intensity were relatively level before the word-final change. No single syllable appeard most prominent based on these measures, as would be expected by a stress system meeting the criterion of culminativity. Below, we relate these findings to the hypotheses laid out in section 2 and discuss whether South Baffin Inuktitut fits any of the metrical stress categories described in the literature (section 5.1) and whether the acoustic patterns should be described as phrasal prominence (section 5.2). We also discuss adjustments to syllable structure as a possible stress correlate (section 5.3).

5.1 Metrical stress and syllable weight

Of the metrical stress systems discussed in section 2, it is most straightforward to determine that the acoustic measurements of f0, duration and intensity in the present data set do not provide evidence for a bounded stress system, whether quantity sensitive or quantity insensitive. A bounded quantity-insensitive stress system would be expected to show alternations of more prominent syllables and less prominent syllables. For example, syllables two (penultimate), four, six and eight as counted from the end of the word should have prototypically shown higher f0, longer duration and/or higher intensity in a language with syllabic trochees or iambs with an extrametrical final syllable, while syllables one (final), three (antepenultimate), five, seven and nine should prototypically have had lower f0, shorter duration and/or lower intensity for iambs without extrametricality.[23] No such regularities appeared. We observed the same acoustic patterns, which lacked evidence for alternations and thus foot structure, in the complete data sets, as well as in subsets composed of only light syllables with V rhymes and in subsets excluding potentially extra-heavy VVC rhymes. Thus, the data neither support the hypothesis that South Baffin Island Inuktitut has a quantity-insensitive bounded stress system, nor the assumption that it has a quantity-sensitive bounded stress system, where heavy syllables would attract

[22] Effects of word position on fundamental frequency were not as easily explainable as boundary effects, with phrase-medial words showing higher f0 than all other positions and phrase-final words showing higher f0 when they were preceded by other words in the same phrase.

[23] We are aware that prominence could also be marked with low instead of high f0, but the data showed no regular f0 alternations at all.

stress.

This leaves the hypothesis of an unbounded stress system, which again might be quantity-sensitive or -insensitive. Before turning to these options, let us first address the question of syllable weight, as classification into light vs. heavy (and potentially extra-heavy) syllables is not established for South Baffin Island Inuktitut. We found clear evidence that the presence of additional segments or long segments was cued by longer durations in our data. Interestingly, durational distinctions appeared between all four rhyme types (V, VC, VV, VVC), in contrast to Pigott's (2012) finding that only two syllable types, long and short, have to be distinguished in Labrador Inuttut (a third type, overlong, appears exclusively in phrase-final position) and Jacobsen's (2000) conclusion that West Greenlandic has three categories, with VC and VV both counting as long. However, whereas the durational differences were overall significant in the dialogues data set, in the single-word utterances, the contrast between VC and VV syllables—which would both count as long in Jacobsen's, but not in Pigott's classification—was only significant in word-final position, where durations were overall larger. While the contrast between VC and VV, where it appeared, was in line with Kleinschmidt's claim that vowels contribute more towards syllable weight than coda consonants, as well as with Pigott's focus on vowel length as the determinant of syllable weight, more investigation is clearly required to assess whether the present findings truly point to a contrast between different Inuit varieties. Note also that statistically significant duration differences alone do not justify proclaiming a phonological contrast.

While duration thus clearly reflected the composition of the rhyme (in line with previous studies on segment quantity and mora-timing in Inuit, cf. section 1.3), f0 and intensity did not show persistent effects of rhyme type throughout the word. For both measures, word-final drops manifested slightly differently for syllables with different rhymes, so that contrasts between rhyme types were significant in final and sometimes penultimate syllables. In earlier syllables, no significant contrasts between rhyme types in f0 and intensity appeared for the dialogues data, and the few contrasts in the single-word utterances were not consistent for both measures. This lack of word-internal contrasts again contradicts the hypothesis of a quantity-sensitive bounded stress system.

In contrast to the hypothesis that South Baffin Island Inuktitut might have a quantity-sensitive unbounded stress system, patterns in the complete data sets corresponded to those in more restricted subsets excluding words with potentially heavy or extra-heavy syllables. Still, the differences between rhyme types in the realization of the word-final drops in f0 and intensity could be interpreted in terms of quantity-sensitive stress, as primary stress being localized at the word end and attracted to heavier syllables. In the dialogues data, lighter syllables had lower f0 than heavier syllables in final position and the contrast between the ultima and preceding syllables was significant for all rhyme types except VVC. This suggests that when the ultima had a VVC rhyme, the word-final f0 fall took place within this extra-heavy last syllable, whereas it took place between the penultimate and the final syllable when the final was light(er). Similar patterns appeared for intensity. The correlate of stress would then be the following drop in intensity and f0 (in autosegmental terms, the association with a H target preceding an L target associated with the end of the word). Accordingly, stress would be on the final syllable if it is heavy and otherwise on the penultimate.

However, several observations contradict this interpretation. First, while the dialogues data set showed differences between the rhyme types in the final syllable, the single-word utterances uniformly showed lower f0 and intensity on the ultimate compared to all preceding syllables for all rhyme types, with differences emerging for the penultimate. Thus, if the pitch and intensity falls were interpreted as stress correlates, the two data sets would point towards different stress locations. Second, even in the dialogues data, some significant differences in between rhyme types appeared in the penultimate syllable for f0, though not for intensity, complicating the determination of a potential weight-dependent position of primary stress further. Third, establishing categories of syllable weight in South Baffin Island Inuktitut is not trivial, as remarked above. Instead of clear contrasts between light and heavy or between light, heavy and extra-heavy, acoustic differences in duration, f0 and intensity revealed a continuum from lighter to heavier depending on the number and type of segments in the syllable rhyme: V<VC<VV<VVC. This would at least complicate an analysis as a quantity-sensitive stress system. Fourth, as Figures 5 to 7 illustrate, the last high syllable before the f0 fall, and thus the most likely syllable to be stressed and associated with a tone, is the antepenultimate, for which no significant differences between rhyme types appeared. Altogether, the present results do not lend support to an analysis of South Baffin Island Inuktitut as having a quantity-sensitive unbounded stress-system.

The final metrical stress hypothesis to consider, then, is the presence of a quantity-insensitive unbounded stress system. Given the localization of the word-final f0 fall, the most viable variant of this hypothesis would posit the antepenultimate syllable as a fixed stress location, meaning that each word contains a single syllabic trochee aligned with the right edge of the word and that final syllables are extrametrical. The main counter-argument to this hypothesis is that acoustic cues do not converge to make the antepenultimate, or any other position, clearly prominent (Jacobsen 2000, employs the same argument against stress for West Greenlandic). Whereas in terms of f0, the antepenultimate would seem most prominent as a high-pitched syllable followed by a fall, the word-final fall in intensity took place between the penultimate and the final syllable in the dialogues, thus making the penultimate, if any, syllable prominent. In the single-word utterances, the intensity fall stretched from the fourth-last syllable to the end of the word (cf. Figure 8). Moreover, while f0 and intensity decreased towards the end of the word, duration increased with the final syllable having the longest duration.

It could perhaps be argued that the last syllable always carries stress (i.e. the single foot is an iamb aligned to the right word edge) if one considered low pitch instead high pitch, i.e. the end of the f0 fall instead of its beginning, as a marker of prominence. However, there are again several arguments against this position. First, the most prominent syllable would then be the one with the lowest intensity. Second, South Baffin Inuktitut words are clearly associated with two tonal targets, H and L, and the choice to interpret one of them as an exponent of stress would be arbitrary. Since the location of the H varies, it cannot easily be interpreted as a trailing tone for an L accent — nor is it suitable to interpret the L, which remains stable near the right word edge, as a trailing tone for the variable H. Third, increases in duration towards the end of the word are gradual, and neither is there categorical change in pitch or duration from the penultimate to the last syllable. Thus, it would be hard to argue that the final syllable is categorically different from all other syllables based on its acoustic characteristics. All in all, the edge-marking of prosodic constituents, especially of right edges, is extremely pervasive in Inuit (recall section 1.4) and interpreting falling f0 and intensity and increasing duration as finality marking is in line with an overwhelming corpus of typological research. Interpreting the low f0, low intensity and long duration of the final syllable as exponents of stress, by contrast, is an extra stipulation for which additional evidence would be desirable.

In conclusion, no syllable position emerged as a clear candidate for a prominent syllable and thereby the location of primary stress, i.e. there was no acoustic evidence of culminativity. Thus, the data failed to lend support to the hypothesis that South Baffin Island Inuktitut has a metrical stress system.

5.2 Phrasal prominence

As remarked in section 2.4, an unbounded quantity-insensitive stress system with a single prominent syllable can be difficult to distinguish from prosodic marking of word boundaries. With respect to the present data, however, the absence of a consistent location of maximum acoustic prominence does favour an interpretation in terms of boundary-marking over an account in terms of demarcative stress. Our previous, more qualitative, analysis of the dialogues data (Arnhold et al. 2018) described words as associated with two tones, H and L. While the L tone showed stable association with the end of the word, the location of pitch peaks, which we interpreted as realizations of the H, varied, as is typical for what Féry (2010, 2016) calls phrase languages, where intonation is primarily the product of tones associated with prosodic domains rather than with prominent syllables inside them. Our present analyses, showing no significant differences in f0 before the word-final f0 drop, are in line with this previous description of the H tone as having no fixed association.

Whether the term '(phrasal) prominence' is appropriate to describe the regular prosodic marking of prosodic domains in South Baffin Island Inuktitut and other Inuit varieties is somewhat questionable. As detailed above, duration, f0 and intensity did not consistently point towards a specific syllable as most prominent in the present data, similar to Jacobsen's (2000) findings for West Greenlandic. In the absence of acoustic correlates, the most persuasive evidence for prominence would be native speaker judgements, which we have not tried to obtain here, though future perception studies are desirable. Recall, however, that Rischel (1974) reports that native speakers of West Greenlandic do not consistently identify one syllable as the most prominent in their language.

5.3 Syllable structure adjustments

While evidence for consistent acoustic prominence has not been found for any Inuit variety investigated so far, another phenomenon has been described under the label stress in some Inuit and related Yupik varieties, namely systematic adjustments of syllable structure (see, e.g., Petersen 1970 on South Greenlandic, Dorais & Lowe 1982 on Siglitun, Kaplan 1985 on Seward Peninsula Inupiaq, Leer 1985*a* on Sugpiaq / Alutiiq / Pacific Yupik, Miyaoka 2019 on Central Alaskan Yupik, Jacobson 2019 on Central Alaskan and Siberian Yupik, Lipscomb 1992 for an overview and comparison between Inuit and Yupik). The most relevant such phenomenon in the present context is Schneider's Law, also called the Law of Double Consonants, which is active in Nunavik and Nunatsiavut varieties of Inuktitut (e.g. Smith 1977, Dorais 1976, 1986, Dresher & Johns 1995, Rose et al. 2012). This law prohibits consonant clusters or geminates to occur across two consecutive syllable boundaries (*. . . CCV(V)CC. . .), i.e. it disallows sequences of closed syllables except when the second one is word-final. For example in Labrador Inuttut, when the vialis affix /-kːut/ is attached to /nunak/ 'land', the result is /nuna**kː**ut/ 'through / across the land', but for /ilːuk/ 'house', which already contains a geminate, vialis form is /il**ː**u**k**ut/ 'through the house' (example from Rose et al. 2012, p. 2). Interestingly, vowel length is not relevant, i.e. Schneider's Law is purely concerned with coda consonants. This suggests that it is not an adjustment of syllable weight and means the process is impossible to describe in terms of moras (thus contrasting, e.g. with iambic vowel lengthening as described for Central Alaskan Yupik, see e.g. Miyaoka 1985, 2012, Hayes 1995). As pointed out by Dresher & Johns (1995) and Rose et al. (2012), this speaks against Schneider's Law being metrically conditioned. Moreover, Pigott (2012) investigated acoustic correlates of prominence, i.e. f0, intensity and duration, and concluded that there is no evidence of metrical stress in Labrador Inuttut and thus no support for a metrical motivation of Schneider's Law. However, Rose et al. (2012) suggest that there is nonetheless prosodic conditioning and that Schneider's Law and similar processes in other Eskimo-Aleut languages are based on a metrical system present in the ancestor language (contra Fortescue's 1983 suggestion that "the stress system of Yupik is probably an innovation since common Eskimo time", p. 120).

In this context, it is interesting that Kaplan (1985, p. 193) states that stress and what he calls 'syllable strength' are independent in Seward Peninsula Inupiaq: Whereas all non-final closed syllables and all syllables containing long vowels or diphthongs are stressed, 'syllable strength' alternates from left to right: If the first syllable is 'strong', containing a long vowel or diphthong or a coda consonant, it will be followed by an open syllable with a short vowel, then by another 'strong' syllable and so on, with the reverse pattern appearing for words starting with a 'weak' syllable, i.e. (C)V. To achieve this pattern, consonants are strengthened or weakened in consonant gradation processes similar to those captured by Schneider's Law, whereas long vowels and diphthongs are not adjusted and may thus interrupt the alternating pattern with a strong syllable. Stress, according to Kaplan, is assigned based on the output of consonant gradation, i.e. a non-final syllable that has a coda due to gemination is strong.

A further relevant observation comes from Jacobsen (2000). In addition to consistent significant differences between short and long segments, she reports consistent small duration adjustments depending on the length of neighbouring sounds in West Greenlandic, e.g. the difference between a long and a short vowel is larger if followed by a short consonant than when followed by a long consonant. In other words, long vowels are slightly shortened preceding long consonants, though they are still significantly longer than short vowels in the same context (see, e.g. Lehtonen 1970, for similar findings in Finnish, another quantity language). Jacobsen (2000) refers to this sub-phonemic variation as 'rhythmic adjustments' of the syllable (it could also be described as a tendency towards syllable isochrony) and points out the similarity between this sub-phonemic variation and more systematic patterns adjusting syllable structure in other Inuit varieties, such as Schneider's Law.

Jacobsen (2000) further found that in her test word containing two adjacent CVVC-syllables (/taːmaː**lː**aː**lː**i-aːsiːt/ 'as usual'), both speakers adjusted the duration of the geminates: The first geminate was shorter than the second geminate for one of her two speakers, whereas it was the other way around for the other speaker. While the difference was significant in both cases, it was smaller than the difference between phonemically long and short consonants. No such adjustment of geminates across consecutive syllable boundaries appeared when vowels were short. While she concludes that stress is not a relevant category in West Greenlandic, Jacobsen refers to the adjustment of consecutive extra-heavy syllables as "rhythmicization proper" (2000, p. 64), in contrast to the general tendency for shorter segment durations in CVVC syllables compared to other syllable types.

Neither Schneider's Law nor similar syllable structure alternations have been described for South Baffin Island Inuktitut to our knowledge (see Dorais 1976, on the absence of Schneider's Law differentiating Baffin Island varieties from those immediately to the south). Indeed, 43 words or about 15% of the words in our dialogues corpus and 42 single-word utterances (11%) violate Schneider's Law, i.e. they contain a CCV(V)CC sequence, with some of these words containing more than one violation, e.g. /iqːanaija**jːaːŋːitː**uŋa/ 'I am not working'. Due to their relatively small absolute number and the fact that our corpora were not controlled for syllable structure, we did not directly measure the duration of the individual syllables in these words. However, as there was a very systematic correlation between syllable duration and syllable structure as indicated by the quasi-phonemic orthographic transcriptions provided by Pirurvik Centre (2015) (cf. section 4.2), they suggest that Schneider's Law is indeed not active in South Baffin Island Inuktitut. Cursory inspection also does not reveal any other tendencies towards alternations of phonologically heavy and light syllables (though, as discussed in section 5.1, whether and how to classify syllable weight in South Baffin Island Inuktitut remains somewhat open at this point). For example, among quadrisyllabic words, the two most common combinations of syllable rhymes were V-V-V-VC and V-VV-V-V ($N = 4$ or 7% each) in the dialogues and VV-VV-VC-VC and VV-V-VVC-VC ($N = 8$ or 13% and $N = 6$ or 10%, respectively) in the single-word utterances. However, future investigations of potential phonological alternations or sub-phonemic adjustments of duration using a carefully constructed corpus are desirable.

The question remains whether adjustments to syllable structure on their own should be considered stress from a synchronic point of view. If one considers rhythmicity as a or the defining criterion of stress, it is natural to model at least the Yupik phenomenon of alternating vowel lengthening in underlyingly CV-syllables (so-called 'iambic lengthening') in terms of foot structure, as e.g. Hayes (1995) does. However, if cumulativity is considered the defining criterion, one syllable would need to be more prominent than the others. At present, there is no indication that this is the case for most Inuit varieties. Moreover, cases like Seward Peninsula Inupiaq (Kaplan 1985) and Labrador Inuttut (Pigott 2012, Rose et al. 2012), where syllable structure adjustments do not correlate with stress or phonetic cues to prominence, may suggest that at least for Inuit, syllable structure adjustments do not constitute word prominence proper (in contrast to Yupik, where syllable structure adjustments are described as related to stress / feet and stress is acoustically cued by f0, duration and intensity, see e.g. Miyaoka 1985, 2012, Woodbury 1987, Krauss 1985*b*, Leer 1985*b*, Martínez-Paricio & Kager 2017, although these analyses typically also treat all stresses within a word as of equal prominence). Supporting this position, acoustic analyses of West Greenlandic show that similar adjustments to syllable weight occur also in an Inuit variety where they do not hold phonological status and do not enforce the creation of a strict alternation of light and heavy syllables (Jacobsen 2000).

6 Conclusion

In line with previous research on other Inuit varieties, our acoustic investigation of potential acoustic cues to stress did not support the hypothesis that South Baffin Island Inuktitut has a metrical stress system—neither bounded nor unbounded, neither quantity-sensitive nor quantity-insensitive. We also did not find clear evidence that a single syllable within each word carried acoustic characteristics marking it as the most prominent, as would be expected of a primary stress.

While adjustments of syllable structure that create rhythmic alternations have been reported for many varieties in the language family, it is doubtful whether they should be regarded as a type of word prominence on their own. For South Baffin Island Inuktitut, we found no indication that such syllable structure adjustments play a role, although further research is warranted.

We therefore conclude that the notion of stress, and indeed of word-level prominence, does not have a well-defined status in the phonology of South Baffin Island Inuktitut as represented in our data. Based on the existing literature, we assume that this conclusion extends to other varieties of Inuktitut and, likely, Inuit more generally. As previously observed (e.g. Fortescue 1983, Krauss 1985*a*, Lipscomb 1992), this seems to constitute a major contrast between Inuit and the related Yupik languages.

Appendix: Statistical modelling

This section presents the best linear-mixed effects models, as well as results of pairwise comparisons referenced in the text above. It is divided into sections for the three acoustic measures evaluated, i.e. duration, fundamental frequency and intensity, which are in turn each divided into sections presenting model summaries and pairwise comparisons for the dialogues data set and the single-word utterance data. Tables with pairwise comparisons only show significant differences, listed in the order in which they are discussed in the text.

Duration: Dialogues

Table 8: Model summary of fixed effects for best linear mixed-effects model of syllable duration (in ms) in dialogues data set.

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	325.67	14.51	31.07	22.44	0.00
VC rhyme	53.41	3.53	1060.98	15.13	0.00
VV rhyme	96.64	4.22	1071.64	22.90	0.00
VVC rhyme	117.69	5.02	1070.06	23.47	0.00
Syl. position 2	-54.96	4.04	1036.82	-13.61	0.00
Syl. position 3	-79.18	4.42	1069.00	-17.93	0.00
Syl. position 4	-97.79	5.13	1071.64	-19.05	0.00
Syl. position 5	-99.62	6.23	1071.22	-16.00	0.00
Syl. position 6	-108.91	7.57	1070.28	-14.39	0.00
Syl. position 7	-126.44	10.32	1070.74	-12.26	0.00
Syl. position 8	-133.53	14.22	1071.49	-9.39	0.00
Word length 2 syl.	-119.29	13.82	1045.23	-8.63	0.00
Word length 3 syl.	-127.69	13.59	1067.24	-9.40	0.00
Word length 4 syl.	-130.81	13.68	1062.95	-9.56	0.00
Word length 5 syl.	-129.58	14.01	1068.42	-9.25	0.00
Word length 6 syl.	-130.71	14.05	1067.36	-9.31	0.00
Word length 7 syl.	-129.57	14.48	1066.70	-8.95	0.00
Word length 8 syl.	-127.86	14.66	1062.42	-8.72	0.00
Word length 9 syl.	-117.61	17.76	1068.75	-6.62	0.00
IP-medial	6.20	5.03	1071.88	1.23	0.22
IP-final	37.85	4.04	1070.76	9.37	0.00
IP only word	38.24	4.20	1072.46	9.10	0.00

Table 9: Pairwise comparisons of syllable duration (in ms) between syllable positions (counted from the end of the word) in dialogues data set.

	Estimate	Std. Error	df	t-ratio	p-value
Syl. position 1 vs. 2	54.962	4.076	1036	13.483	<.0001
Syl. position 1 vs. 3	79.1801	4.435	1069	17.853	<.0001
Syl. position 1 vs. 4	97.787	5.139	1072	19.030	<.0001
Syl. position 1 vs. 5	99.624	6.232	1071	15.985	<.0001
Syl. position 1 vs. 6	108.908	7.568	1070	14.390	<.0001
Syl. position 1 vs. 7	126.439	10.32	1071	12.252	<.0001
Syl. position 1 vs. 8	133.525	14.234	1072	9.381	<.0001
Syl. position 2 vs. 3	24.218	4.324	1072	5.601	<.0001
Syl. position 2 vs. 4	42.825	5.091	1072	8.412	<.0001
Syl. position 2 vs. 5	44.662	6.211	1072	7.191	<.0001
Syl. position 2 vs. 6	53.945	7.586	1071	7.111	<.0001
Syl. position 2 vs. 7	71.477	10.325	1072	6.923	<.0001
Syl. position 2 vs. 8	78.563	14.241	1072	5.517	<.0001
Syl. position 3 vs. 4	18.606	5.151	1071	3.612	0.0076
Syl. position 3 vs. 5	20.443	6.264	1071	3.264	0.0250
Syl. position 3 vs. 6	29.727	7.607	1072	3.908	0.0025
Syl. position 3 vs. 7	47.258	10.323	1071	4.578	0.0001
Syl. position 3 vs. 8	54.344	14.237	1072	3.817	0.0036

Table 10: Pairwise comparisons syllable duration between word lengths (counted in number of syllables) in dialogues data set.

	Estimate	Std. Error	df	t-ratio	p-value
1 vs. 2 syllables	119.286	13.950	1045	8.551	<.0001
1 vs. 3 syllables	127.694	13.663	1067	9.346	<.0001
1 vs. 4 syllables	130.814	13.768	1063	9.502	<.0001
1 vs. 5 syllables	129.582	14.089	1068	9.197	<.0001
1 vs. 6 syllables	130.712	14.118	1067	9.258	<.0001
1 vs. 7 syllables	129.570	14.556	1067	8.901	<.0001
1 vs. 8 syllables	127.86	14.752	1062	8.667	<.0001
1 vs. 9 syllables	117.611	17.845	1069	6.591	<.0001

Table 11: Pairwise comparisons syllable duration between syllable rhyme types in dialogues data set.

	Estimate	Std. Error	df	t-ratio	p-value
V vs. VC	-53.408	3.554	1061	-15.028	<.0001
V vs. VV	-96.642	4.232	1072	-22.838	<.0001
V vs. VVC	-117.695	5.034	1070	-23.381	<.0001
VC vs. VV	-43.234	4.793	1037	-9.021	<.0001
VC vs. VVC	-64.287	5.544	1042	-11.596	<.0001
VV vs. VVC	-21.053	5.842	1071	-3.604	0.0019

Table 12: Pairwise comparisons syllable duration between different positions of a word in an intonational phrase in dialogues data set.

	Estimate	Std. Error	df	t-ratio	p-value
Initial vs. final	-37.853	4.04	1071	-9.370	<.0001
Initial vs. only	-38.24	4.212	1073	-9.078	<.0001
Medial vs. final	-31.653	4.823	1071	-6.564	<.0001
Medial vs. only	-32.04	4.982	1073	-6.431	<.0001

Duration: Single-word utterances

Table 13: Model summary of fixed effects for best best linear mixed-effects model of syllable duration (in ms) in single-word data set.

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	350.89	12.90	657.99	27.19	0.00
VC rhyme	78.14	5.95	1066.12	13.14	0.00
VV rhyme	148.26	9.76	1226.34	15.19	0.00
VVC rhyme	194.50	6.99	1226.53	27.83	0.00
Syl. position 2	-113.03	5.99	1185.18	-18.87	0.00
Syl. position 3	-138.61	6.24	1229.18	-22.22	0.00
Syl. position 4	-136.97	6.82	1229.11	-20.10	0.00
Syl. position 5	-148.93	7.61	1229.44	-19.57	0.00
Syl. position 6	-117.22	10.40	1229.57	-11.27	0.00
Word length 2 syl.	-52.72	11.69	1228.33	-4.51	0.00
Word length 3 syl.	-55.26	11.68	1228.66	-4.73	0.00
Word length 4 syl.	-56.76	11.92	1229.50	-4.76	0.00
Word length 5 syl.	-60.12	12.02	1229.62	-5.00	0.00
Word length 6 syl.	-59.78	12.68	1229.69	-4.72	0.00
Word length 7 syl.	-64.63	12.89	1229.76	-5.02	0.00
Word length 8 syl.	-65.13	13.01	1229.01	-5.01	0.00
Word length 9 syl.	-83.22	16.51	1229.98	-5.04	0.00
VC rhyme : Syl. position 2	10.35	7.83	1003.36	1.32	0.19
VV rhyme : Syl. position 2	-44.50	11.88	1213.95	-3.75	0.00
VVC rhyme : Syl. position 2	-31.14	8.97	1223.25	-3.47	0.00
VC rhyme : Syl. position 3	10.79	8.97	1205.21	1.20	0.23
VV rhyme : Syl. position 3	-46.52	11.27	1229.92	-4.13	0.00
VVC rhyme : Syl. position 3	-60.72	10.17	1229.84	-5.97	0.00
VC rhyme : Syl. position 4	-1.02	16.09	1212.67	-0.06	0.95
VV rhyme : Syl. position 4	-60.02	12.15	1228.46	-4.94	0.00
VVC rhyme : Syl. position 4	-41.22	15.91	1229.93	-2.59	0.01
VC rhyme : Syl. position 5	-27.34	15.98	1225.01	-1.71	0.09
VV rhyme : Syl. position 5	-68.73	16.60	1229.99	-4.14	0.00
VVC rhyme : Syl. position 5	-76.69	15.17	1229.83	-5.05	0.00
VC rhyme : Syl. position 6	-70.23	29.45	1228.73	-2.38	0.02
VV rhyme : Syl. position 6	-98.91	19.29	1223.15	-5.13	0.00
VVC rhyme : Syl. position 6	-61.82	22.92	1229.79	-2.70	0.01

Table 14: Pairwise comparisons syllable duration between word lengths (counted in number of syllables) in single-word data set.

	Estimate	Std. Error	df	t-ratio	p-value
1 vs. 2 syllables	52.719	11.696	1228	4.508	0.0002
1 vs. 3 syllables	55.256	11.679	1229	4.731	0.0001
1 vs. 4 syllables	56.758	11.93	1230	4.758	0.0001
1 vs. 5 syllables	60.118	12.034	1230	4.996	<.0001
1 vs. 6 syllables	59.782	12.691	1230	4.710	0.0001
1 vs. 7 syllables	64.633	12.915	1230	5.004	<.0001
1 vs. 8 syllables	65.126	13.049	1229	4.991	<.0001
1 vs. 9 syllables	83.215	16.539	1230	5.031	<.0001

Table 15: Pairwise comparisons syllable duration between syllable positions (counted from the end of the word) by rhyme types in single-word data set.

	Estimate	Std. Error	df	t-ratio	p-value
Syl. position 1 vs. 2 for V	113.026	6.046	1182	18.693	<.0001
Syl. position 1 vs. 3 for V	138.612	6.244	1229	22.204	<.0001
Syl. position 1 vs. 4 for V	136.967	6.834	1229	20.042	<.0001
Syl. position 1 vs. 5 for V	148.928	7.614	1229	19.561	<.0001
Syl. position 1 vs. 6 for V	117.221	10.407	1230	11.263	<.0001
Syl. position 1 vs. 2 for VC	102.676	4.887	1128	21.011	<.0001
Syl. position 1 vs. 3 for VC	127.824	6.48	1181	19.726	<.0001
Syl. position 1 vs. 4 for VC	137.988	14.573	1217	9.469	<.0001
Syl. position 1 vs. 5 for VC	176.271	14.277	1226	12.347	<.0001
Syl. position 1 vs. 6 for VC	187.451	27.793	1229	6.745	<.0001
Syl. position 1 vs. 2 for VV	157.529	10.231	1230	15.397	<.0001
Syl. position 1 vs. 3 for VV	185.134	9.524	1230	19.439	<.0001
Syl. position 1 vs. 4 for VV	196.986	10.422	1230	18.901	<.0001
Syl. position 1 vs. 5 for VV	217.654	15.058	1230	14.455	<.0001
Syl. position 1 vs. 6 for VV	216.130	16.676	1226	12.961	<.0001
Syl. position 1 vs. 2 for VVC	144.164	6.63	1229	21.743	<.0001
Syl. position 1 vs. 3 for VVC	199.334	8.146	1230	24.471	<.0001
Syl. position 1 vs. 4 for VVC	178.185	14.564	1229	12.234	<.0001
Syl. position 1 vs. 5 for VVC	225.615	13.382	1230	16.86	<.0001
Syl. position 1 vs. 6 for VVC	179.042	20.651	1229	8.67	<.0001
Syl. position 2 vs. 3 for V	25.586	4.779	1192	5.354	<.0001
Syl. position 2 vs. 4 for V	23.941	5.599	1207	4.276	0.0047
Syl. position 2 vs. 5 for V	35.902	6.632	1226	5.413	<.0001
Syl. position 2 vs. 5 for VC	73.595	14.488	1230	5.08	0.0001
Syl. position 2 vs. 4 for VV	39.457	8.682	1230	4.544	0.0015
Syl. position 2 vs. 5 for VV	60.125	13.958	1229	4.307	0.0041
Syl. position 2 vs. 6 for VV	58.602	15.631	1230	3.749	0.0353
Syl. position 2 vs. 3 for VVC	55.171	8.212	1230	6.718	<.0001
Syl. position 2 vs. 5 for VVC	81.451	13.359	1229	6.097	<.0001

Table 16: Pairwise comparisons syllable duration between rhyme types by syllable positions (counted from the end of the word) in single-word data set.

	Estimate	Std. Error	df	t-ratio	p-value
V vs. VC in syl. position 1	-78.145	6.052	1057	-12.912	<.0001
V vs. VV in syl. position 1	-148.263	9.799	1226	-15.13	<.0001
V vs. VVC in syl. position 1	-194.504	7.015	1226	-27.728	<.0001
VC vs. VV in syl. position 1	-70.118	8.687	1229	-8.072	<.0001
VC vs. VVC in syl. position 1	-116.359	5.458	1223	-21.318	<.0001
VV vs. VVC in syl. position 1	-46.241	9.083	1230	-5.091	0.0001
V vs. VC in syl. position 2	-88.495	5.066	1206	-17.467	<.0001
V vs. VV in syl. position 2	-103.76	6.907	1230	-15.022	<.0001
V vs. VVC in syl. position 2	-163.366	5.682	1230	-28.753	<.0001
VC vs. VVC in syl. position 2	-74.871	6.161	1230	-12.152	<.0001
VV vs. VVC in syl. position 2	-59.607	7.698	1229	-7.743	<.0001
V vs. VC in syl. position 3	-88.933	6.727	1230	-13.221	<.0001
V vs. VV in syl. position 3	-101.74	5.699	1230	-17.853	<.0001
V vs. VVC in syl. position 3	-133.781	7.451	1229	-17.955	<.0001
VC vs. VVC in syl. position 3	-44.849	8.706	1230	-5.152	0.0001
VV vs. VVC in syl. position 3	-32.041	7.956	1229	-4.027	0.0128
V vs. VC in syl. position 4	-77.124	14.725	1230	-5.238	0.0001
V vs. VV in syl. position 4	-88.244	7.253	1229	-12.167	<.0001
V vs. VVC in syl. position 4	-153.286	14.346	1230	-10.685	<.0001
VC vs. VVC in syl. position 4	-76.162	19.625	1229	-3.881	0.0221
VV vs. VVC in syl. position 4	-65.042	14.854	1229	-4.379	0.0031
V vs. VVC in syl. position 5	-117.817	13.46	1228	-8.753	<.0001
VC vs. VVC in syl. position 5	-67.016	18.288	1228	-3.664	0.047
V vs. VV in syl. position 5	-79.537	13.462	1228	-5.908	<.0001
V vs. VVC in syl. position 6	-132.683	21.715	1229	-6.11	<.0001
VC vs. VVC in syl. position 6	-124.768	34.155	1230	-3.653	0.0488

Fundamental frequency: Dialogues

Table 17: Model summary of fixed effects for best linear mixed-effects model of fundamental frequency (in st re 100 Hz) in dialogues data set.

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	9.43	3.49	1.11	2.70	0.21
VC rhyme	-0.74	0.36	1055.00	-2.02	0.04
VV rhyme	1.68	0.54	1055.00	3.13	0.00
VVC rhyme	3.94	0.55	1055.00	7.20	0.00
Syl. position 2	2.45	0.27	1055.00	9.01	0.00
Syl. position 3	4.60	0.31	1055.00	14.76	0.00
Syl. position 4	4.60	0.35	1055.00	13.12	0.00
Syl. position 5	5.95	0.42	1055.00	14.33	0.00
Syl. position 6	6.21	0.53	1055.00	11.63	0.00
Syl. position 7	6.40	0.74	1055.00	8.64	0.00
Syl. position 8	4.68	1.72	1055.00	2.72	0.01
Word length 2 syl.	-0.47	0.78	1055.00	-0.60	0.55
Word length 3 syl.	-0.32	0.78	1055.00	-0.41	0.68
Word length 4 syl.	-0.91	0.78	1055.00	-1.16	0.25
Word length 5 syl.	-1.45	0.80	1055.00	-1.81	0.07
Word length 6 syl.	-1.54	0.81	1055.00	-1.92	0.06
Word length 7 syl.	-3.05	0.83	1055.00	-3.66	0.00
Word length 8 syl.	-2.79	0.83	1055.00	-3.35	0.00
Word length 9 syl.	-2.53	1.01	1055.00	-2.49	0.01
IP-medial	0.76	0.26	1055.00	2.93	0.00
IP-final	0.08	0.21	1055.00	0.37	0.71
IP only word	-0.46	0.22	1055.01	-2.12	0.03
VC rhyme : Syl. position 2	1.48	0.51	1055.00	2.92	0.00
VV rhyme : Syl. position 2	0.73	0.71	1055.00	1.02	0.31
VVC rhyme : Syl. position 2	-3.26	0.69	1055.00	-4.69	0.00
VC rhyme : Syl. position 3	0.28	0.54	1055.00	0.52	0.60
VV rhyme : Syl. position 3	-0.93	0.69	1055.00	-1.35	0.18
VVC rhyme : Syl. position 3	-3.87	0.82	1055.00	-4.71	0.00
VC rhyme : Syl. position 4	1.78	0.65	1055.00	2.74	0.01
VV rhyme : Syl. position 4	-0.93	0.75	1055.00	-1.23	0.22
VVC rhyme : Syl. position 4	-3.03	0.88	1055.00	-3.43	0.00
VC rhyme : Syl. position 5	0.19	0.73	1055.00	0.25	0.80
VV rhyme : Syl. position 5	-1.84	0.95	1055.00	-1.94	0.05
VVC rhyme : Syl. position 5	-4.89	1.38	1055.00	-3.56	0.00
VC rhyme : Syl. position 6	0.18	0.85	1055.00	0.21	0.84
VV rhyme : Syl. position 6	-2.37	1.02	1055.00	-2.32	0.02
VVC rhyme : Syl. position 6	-4.90	2.50	1055.00	-1.96	0.05
VC rhyme : Syl. position 7	0.78	1.32	1055.00	0.59	0.56
VV rhyme : Syl. position 7	-2.21	1.25	1055.00	-1.76	0.08
VVC rhyme : Syl. position 7	-4.63	2.55	1055.00	-1.82	0.07
VC rhyme : Syl. position 8	1.98	2.03	1055.00	0.98	0.33
VV rhyme : Syl. position 8	0.89	2.16	1055.00	0.41	0.68
VVC rhyme : Syl. position 8	-1.69	2.44	1055.00	-0.69	0.49

Table 18: Pairwise comparisons of fundamental frequency (in st re 100 Hz) between rhyme types by syllable positions (counted from the end of the word) in dialogues data set.

	Estimate	Std. Error	df	t-ratio	p-value
V vs. VVC in syl. position 1	-3.94	0.548	1055	-7.197	<.0001
VC vs. VVC in syl. position 1	-4.677	0.606	1055	-7.712	<.0001
VC vs. VV in syl. position 1	-2.416	0.595	1055	-4.059	0.0191
V vs. VV in syl. position 2	-2.407	0.472	1055	-5.098	0.0002

Table 19: Pairwise comparisons of fundamental frequency (in st re 100 Hz) between syllable positions (counted from the end of the word) by rhyme types in dialogue data set.

	Estimate	Std. Error	df	t-ratio	p-value
Syl. position 1 vs. 2 for V	-2.447	0.272	1055	-9.006	<.0001
Syl. position 1 vs. 3 for V	-4.6	0.312	1055	-14.762	<.0001
Syl. position 1 vs. 4 for V	-4.6	0.351	1055	-13.121	<.0001
Syl. position 1 vs. 5 for V	-5.948	0.415	1055	-14.325	<.0001
Syl. position 1 vs. 6 for V	-6.211	0.534	1055	-11.628	<.0001
Syl. position 1 vs. 7 for V	-6.405	0.741	1055	-8.643	<.0001
Syl. position 1 vs. 2 for VC	-3.93	0.429	1055	-9.166	<.0001
Syl. position 1 vs. 3 for VC	-4.884	0.444	1055	-11.008	<.0001
Syl. position 1 vs. 4 for VC	-6.383	0.559	1055	-11.434	<.0001
Syl. position 1 vs. 5 for VC	-6.134	0.623	1055	-9.85	<.0001
Syl. position 1 vs. 6 for VC	-6.387	0.693	1055	-9.219	<.0001
Syl. position 1 vs. 7 for VC	-7.181	1.139	1055	-6.305	<.0001
Syl. position 1 vs. 8 for VC	-6.66	1.153	1055	-5.776	<.0001
Syl. position 1 vs. 2 for VV	-3.175	0.657	1055	-4.833	0.0007
Syl. position 1 vs. 3 for VV	-3.667	0.623	1055	-5.887	<.0001
Syl. position 1 vs. 4 for VV	-3.67	0.683	1055	-5.37	<.0001
Syl. position 1 vs. 5 for VV	-4.109	0.875	1055	-4.698	0.0013
Syl. position 1 vs. 6 for VV	-3.842	0.902	1055	-4.262	0.0086
Syl. position 1 vs. 7 for VV	-4.197	1.06	1055	-3.961	0.0275
Syl. position 1 vs. 8 for VV	-5.567	1.349	1055	-4.127	0.0147
Syl. position 2 vs. 3 for V	-2.153	0.323	1055	-6.673	<.0001
Syl. position 2 vs. 4 for V	-2.153	0.359	1055	-6	<.0001
Syl. position 2 vs. 5 for V	-3.502	0.42	1055	-8.333	<.0001
Syl. position 2 vs. 6 for V	-3.765	0.538	1055	-7.005	<.0001
Syl. position 2 vs. 7 for V	-3.958	0.743	1055	-5.33	0.0001
Syl. position 2 vs. 4 for VC	-2.453	0.552	1055	-4.447	0.004

Table 20: Pairwise comparisons of fundamental frequency (in st re 100 Hz) between word lengths (counted in number of syllables) in dialogues data set.

	Estimate	Std. Error	df	t-ratio	p-value
1 vs. 7 syllables	3.0453	0.832	1055	3.661	0.0081
1 vs. 8 syllables	2.792	0.834	1055	3.346	0.0239
4 vs. 7 syllables	2.14	0.336	1055	6.363	<.0001
4 vs. 8 syllables	1.887	0.351	1055	5.384	<.0001
5 vs. 7 syllables	1.594	0.346	1055	4.605	0.0002
5 vs. 8 syllables	1.340	0.358	1055	3.744	0.0059
6 vs. 7 syllables	1.501	0.336	1055	4.470	0.0003
6 vs. 8 syllables	1.248	0.35	1055	3.565	0.0114
2 vs. 6 syllables	1.074	0.333	1055	3.223	0.0353
2 vs. 7 syllables	2.574	0.389	1055	6.624	<.0001
2 vs. 8 syllables	2.321	0.399	1055	5.823	<.0001
3 vs. 5 syllables	1.131	0.268	1055	4.214	0.0009
3 vs. 6 syllables	1.224	0.283	1055	4.332	0.0005
3 vs. 7 syllables	2.724	0.345	1055	7.891	<.0001
3 vs. 8 syllables	2.471	0.358	1055	6.899	<.0001
3 vs. 9 syllables	2.209	0.689	1055	3.206	0.0372

Table 21: Pairwise comparisons of fundamental frequency (in st re 100 Hz) between different positions of a word in an intonational phrase in dialogues data set.

	Estimate	Std. Error	df	t-ratio	p-value
Initial vs. medial	-0.762	0.26	1055	-2.928	0.0183
Medial vs. final	0.684	0.25	1055	2.734	0.0323
Medial vs. only	1.222	0.259	1055	4.724	<.0001
final vs. only	0.538	0.188	1055	2.865	0.0221

Fundamental frequency: Single-word utterances

Table 22: Model summary of fixed effects for best best linear mixed-effects model of fundamental frequency (in st re 100 Hz) in single-word data set.

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	13.49	0.47	423.61	28.61	0.00
VC rhyme	-0.16	0.22	1136.54	-0.75	0.45
VV rhyme	0.65	0.37	1235.24	1.74	0.08
VVC rhyme	0.77	0.25	1234.84	3.02	0.00
Syl. position 2	3.44	0.22	1216.87	15.50	0.00
Syl. position 3	4.48	0.23	1234.81	19.46	0.00
Syl. position 4	4.79	0.25	1235.95	19.06	0.00
Syl. position 5	5.24	0.28	1235.16	18.60	0.00
Syl. position 6	5.52	0.39	1235.26	14.33	0.00
Word length 2 syl.	-2.70	0.42	1234.11	-6.40	0.00
Word length 3 syl.	-4.09	0.42	1234.83	-9.72	0.00
Word length 4 syl.	-4.28	0.43	1235.57	-9.93	0.00
Word length 5 syl.	-4.85	0.43	1235.68	-11.16	0.00
Word length 6 syl.	-4.69	0.46	1235.74	-10.21	0.00
Word length 7 syl.	-4.77	0.47	1235.85	-10.23	0.00
Word length 8 syl.	-5.02	0.47	1235.73	-10.67	0.00
Word length 9 syl.	-5.34	0.60	1236.00	-8.84	0.00
VC rhyme : Syl. position 2	-0.09	0.29	1107.52	-0.30	0.76
VV rhyme : Syl. position 2	-1.10	0.45	1229.09	-2.46	0.01
VVC rhyme : Syl. position 2	-1.07	0.33	1233.23	-3.25	0.00
VC rhyme : Syl. position 3	0.76	0.33	1222.29	2.29	0.02
VV rhyme : Syl. position 3	-0.61	0.43	1235.67	-1.42	0.16
VVC rhyme : Syl. position 3	-0.92	0.37	1236.00	-2.46	0.01
VC rhyme : Syl. position 4	0.07	0.60	1228.97	0.12	0.90
VV rhyme : Syl. position 4	-0.07	0.46	1235.86	-0.16	0.87
VVC rhyme : Syl. position 4	-1.83	0.59	1235.90	-3.11	0.00
VC rhyme : Syl. position 5	0.47	0.59	1234.11	0.80	0.43
VV rhyme : Syl. position 5	-0.81	0.62	1235.77	-1.30	0.19
VVC rhyme : Syl. position 5	-1.43	0.56	1235.51	-2.55	0.01
VC rhyme : Syl. position 6	0.32	1.09	1234.43	0.29	0.77
VV rhyme : Syl. position 6	-0.40	0.72	1233.88	-0.55	0.58
VVC rhyme : Syl. position 6	-1.55	0.85	1235.45	-1.83	0.07

Table 23: Pairwise comparisons of fundamental frequency (in st re 100 Hz) between syllable positions (counted from the end of the word) by rhyme types in single-word data set.

	Estimate	Std. Error	df	t-ratio	p-value
Syl. position 1 vs. 2 for V	-3.437	0.223	1217	-15.401	<.0001
Syl. position 1 vs. 3 for V	-4.484	0.231	1235	-19.448	<.0001
Syl. position 1 vs. 4 for V	-4.786	0.252	1236	-19.02	<.0001
Syl. position 1 vs. 5 for V	-5.236	0.282	1235	-18.587	<.0001
Syl. position 1 vs. 6 for V	-5.522	0.386	1235	-14.322	<.0001
Syl. position 1 vs. 2 for VC	-3.349	0.186	1183	-18.004	<.0001
Syl. position 1 vs. 3 for VC	-5.243	0.239	1207	-21.986	<.0001
Syl. position 1 vs. 4 for VC	-4.859	0.54	1232	-8.996	<.0001
Syl. position 1 vs. 5 for VC	-5.707	0.529	1235	-10.783	<.0001
Syl. position 1 vs. 6 for VC	-5.844	1.032	1235	-5.665	<.0001
Syl. position 1 vs. 2 for VV	-2.338	0.387	1236	-6.047	<.0001
Syl. position 1 vs. 3 for VV	-3.877	0.364	1235	-10.655	<.0001
Syl. position 1 vs. 4 for VV	-4.712	0.396	1236	-11.913	<.0001
Syl. position 1 vs. 5 for VV	-4.427	0.565	1235	-7.837	<.0001
Syl. position 1 vs. 6 for VV	-5.126	0.624	1235	-8.214	<.0001
Syl. position 1 vs. 2 for VVC	-2.369	0.242	1235	-9.796	<.0001
Syl. position 1 vs. 3 for VVC	-3.561	0.3	1236	-11.874	<.0001
Syl. position 1 vs. 4 for VVC	-2.954	0.539	1235	-5.482	<.0001
Syl. position 1 vs. 5 for VVC	-3.808	0.495	1235	-7.699	<.0001
Syl. position 1 vs. 6 for VVC	-3.97	0.764	1235	-5.196	0.0001
Syl. position 2 vs. 3 for V	-1.046	0.178	122	-5.893	<.0001
Syl. position 2 vs. 4 for V	-1.349	0.207	1228	-6.508	<.0001
Syl. position 2 vs. 5 for V	-1.799	0.246	1235	-7.306	<.0001
Syl. position 2 vs. 6 for V	-2.084	0.361	1235	-5.776	<.0001
Syl. position 2 vs. 3 for VV	-1.54	0.281	1235	-5.475	<.0001
Syl. position 2 vs. 4 for VV	-2.375	0.317	1235	-7.484	<.0001
Syl. position 2 vs. 5 for VV	-2.089	0.514	1234	-4.063	0.0111
Syl. position 2 vs. 6 for VV	-2.788	0.577	1236	-4.832	0.0004
Syl. position 2 vs. 3 for VC	-1.894	0.261	1234	-7.247	<.0001
Syl. position 2 vs. 5 for VC	-2.358	0.539	1235	-4.375	0.0031
Syl. position 2 vs. 3 for VVC	-1.192	0.304	1235	-3.921	0.019

Table 24: Pairwise comparisons of fundamental frequency (in st re 100 Hz) between word lengths (counted in number of syllables) in single-word data set.

	Estimate	Std. Error	df	t-ratio	p-value
1 vs. 2 syllables	2.697	0.422	1234	6.395	<.0001
1 vs. 3 syllables	4.088	0.421	1235	9.713	<.0001
1 vs. 4 syllables	4.276	0.431	1236	9.923	<.0001
1 vs. 5 syllables	4.846	0.435	1236	11.151	<.0001
1 vs. 6 syllables	4.69	0.46	1236	10.198	<.0001
1 vs. 7 syllables	4.774	0.468	1236	10.209	<.0001
1 vs. 8 syllables	5.02	0.472	1236	10.643	<.0001
1 vs. 9 syllables	5.342	0.605	1236	8.826	<.0001
2 vs. 3 syllables	1.392	0.135	1236	10.321	<.0001
2 vs. 4 syllables	1.579	0.153	1229	10.326	<.0001
2 vs. 5 syllables	2.149	0.161	1227	13.313	<.0001
2 vs. 6 syllables	1.993	0.222	1233	8.987	<.0001
2 vs. 7 syllables	2.077	0.238	1216	8.737	<.0001
2 vs. 8 syllables	2.323	0.246	1209	9.445	<.0001
2 vs. 9 syllables	2.645	0.451	1235	5.862	<.0001
3 vs. 5 syllables	0.758	0.133	1236	5.697	<.0001
3 vs. 7 syllables	0.685	0.218	1233	3.150	0.0440
3 vs. 8 syllables	0.931	0.227	1229	4.102	0.0014
4 vs. 5 syllables	0.57	0.14	1234	4.060	0.0017
4 vs. 8 syllables	0.744	0.231	1233	3.221	0.0354

Intensity: Dialogues

Table 25: Model summary of fixed effects for best linear mixed-effects model of mean vowel intensity (in dB) in dialogues data set.

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	73.67	0.71	36.92	104.07	0.00
VC rhyme	-0.15	0.30	1054.02	-0.49	0.62
VV rhyme	0.76	0.45	1054.10	1.69	0.09
VVC rhyme	2.41	0.46	1054.19	5.28	0.00
Syl. position 2	1.80	0.23	1054.13	7.89	0.00
Syl. position 3	2.04	0.26	1054.00	7.82	0.00
Syl. position 4	2.00	0.29	1054.00	6.84	0.00
Syl. position 5	2.97	0.35	1054.04	8.54	0.00
Syl. position 6	3.31	0.45	1054.05	7.43	0.00
Syl. position 7	2.52	0.62	1054.00	4.10	0.00
Syl. position 8	1.83	1.43	1054.00	1.28	0.20
Word length 2 syl.	-0.23	0.65	1054.28	-0.36	0.72
Word length 3 syl.	-0.52	0.64	1054.11	-0.80	0.42
Word length 4 syl.	-0.29	0.65	1054.12	-0.44	0.66
Word length 5 syl.	-0.88	0.66	1054.37	-1.32	0.19
Word length 6 syl.	-0.78	0.67	1054.06	-1.17	0.24
Word length 7 syl.	-2.27	0.69	1054.02	-3.29	0.00
Word length 8 syl.	-1.46	0.69	1054.15	-2.12	0.03
Word length 9 syl.	-1.98	0.83	1054.14	-2.39	0.02
IP-medial	-0.13	0.22	1054.12	-0.60	0.55
IP-final	-1.55	0.17	1054.05	-8.89	0.00
IP only word	-1.08	0.18	1054.51	-6.02	0.00
VC rhyme : Syl. position 2	0.02	0.42	1054.05	0.05	0.96
VV rhyme : Syl. position 2	-0.06	0.59	1054.02	-0.11	0.91
VVC rhyme : Syl. position 2	-2.24	0.58	1054.44	-3.87	0.00
VC rhyme : Syl. position 3	-0.79	0.44	1054.01	-1.78	0.07
VV rhyme : Syl. position 3	0.21	0.57	1054.19	0.36	0.72
VVC rhyme : Syl. position 3	-1.56	0.68	1054.09	-2.29	0.02
VC rhyme : Syl. position 4	0.60	0.54	1054.02	1.12	0.26
VV rhyme : Syl. position 4	0.32	0.63	1054.01	0.52	0.60
VVC rhyme : Syl. position 4	-0.51	0.73	1054.01	-0.70	0.48
VC rhyme : Syl. position 5	-0.04	0.61	1054.02	-0.06	0.95
VV rhyme : Syl. position 5	-0.43	0.77	1054.05	-0.56	0.58
VVC rhyme : Syl. position 5	-1.54	1.14	1054.16	-1.35	0.18
VC rhyme : Syl. position 6	-1.02	0.71	1054.02	-1.42	0.15
VV rhyme : Syl. position 6	-0.05	0.85	1054.33	-0.05	0.96
VVC rhyme : Syl. position 6	-0.56	2.07	1054.10	-0.27	0.79
VC rhyme : Syl. position 7	0.14	1.04	1054.02	0.13	0.89
VV rhyme : Syl. position 7	1.41	1.04	1054.04	1.36	0.18
VVC rhyme : Syl. position 7	0.65	2.11	1054.08	0.31	0.76
VC rhyme : Syl. position 8	-1.87	1.68	1054.07	-1.12	0.26
VV rhyme : Syl. position 8	2.70	1.79	1054.01	1.51	0.13
VVC rhyme : Syl. position 8	0.19	2.03	1054.16	0.09	0.93

Table 26: Pairwise comparisons of mean vowel intensity (in dB) between syllable positions (counted from the end of the word) by rhyme types in dialogues set.

	Estimate	Std. Error	df	t-ratio	p-value
Syl. position 1 vs. 2 for V	-1.802	0.228	1054	-7.89	<.0001
Syl. position 1 vs. 3 for V	-2.044	0.261	1054	-7.822	<.0001
Syl. position 1 vs. 4 for V	-2	0.293	1054	-6.837	<.0001
Syl. position 1 vs. 5 for V	-2.97	0.348	1054	-8.542	<.0001
Syl. position 1 vs. 6 for V	-3.306	0.445	1054	-7.428	<.0001
Syl. position 1 vs. 7 for V	-2.522	0.615	1054	-4.1	0.0164
Syl. position 1 vs. 2 for VC	-1.822	0.35	1054	-5.211	0.0001
Syl. position 1 vs. 4 for VC	-2.602	0.459	1054	-5.672	<.0001
Syl. position 1 vs. 5 for VC	-2.933	0.512	1054	-5.731	<.0001
Syl. position 1 vs. 6 for VC	-2.288	0.585	1054	-3.913	0.0328
Syl. position 1 vs. 3 for VV	-2.253	0.517	1054	-4.358	0.0058
Syl. position 1 vs. 4 for VV	-2.324	0.567	1054	-4.098	0.0165
Syl. position 1 vs. 6 for VV	-3.261	0.748	1054	-4.361	0.0057
Syl. position 1 vs. 7 for VV	-3.933	0.879	1054	-4.475	0.0035
Syl. position 1 vs. 8 for VV	-4.524	1.116	1054	-4.053	0.0195

Table 27: Pairwise comparisons of mean vowel intensity (in dB) between syllable positions (counted from the end of the word) by rhyme types in dialogues data set.

	Estimate	Std. Error	df	t-ratio	p-value
V vs. VVC in syl. position 1	-2.405	0.455	1054	-5.283	0.0001
VC vs. VVC in syl. position 1	-2.553	0.5	1054	-5.11	0.0002
VC vs. VV in syl. position 3	-1.905	0.391	1054	-4.879	0.0006
VC vs. VV in syl. position 8	-5.475	1.355	1054	-4.041	0.0204

Table 28: Pairwise comparisons of mean vowel intensity (in dB) between word lengths (counted in number of syllables) in dialogues data set.

	Estimate	Std. Error	df	t-ratio	p-value
2 vs. 7 syllables	2.033	0.323	1055	6.293	<.0001
2 vs. 8 syllables	1.23	0.33	1054	3.730	0.0063
3 vs. 7 syllables	1.75	0.285	1054	6.130	<.0001
3 vs. 8 syllables	0.947	0.295	1054	3.213	0.0364
4 vs. 7 syllables	1.979	0.279	1054	7.103	<.0001
4 vs. 8 syllables	1.175	0.29	1054	4.067	0.0017
5 vs. 7 syllables	1.389	0.287	1055	4.840	0.0001
6 vs. 7 syllables	1.485	0.279	1054	5.318	<.0001

Table 29: Pairwise comparisons of mean vowel intensity (in dB) between different positions of a word in an intonational phrase in dialogues data set.

	Estimate	Std. Error	df	t-ratio	p-value
Initial vs. final	1.55	0.174	1054	8.886	<.0001
Initial vs. only	1.084	0.18	1055	6.016	<.0001
Medial vs. final	1.422	0.207	1054	6.858	<.0001
Medial vs. only	0.955	0.213	1055	4.473	0.0001
final vs. only	-0.467	0.156	1055	-2.993	0.0150

Intensity: Single-word utterances

Table 30: Model summary of fixed effects for best linear mixed-effects model of mean vowel intensity (in dB) in single-word data set.

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	73.25	0.78	491.83	93.44	0.00
VC rhyme	1.39	0.37	1106.79	3.79	0.00
VV rhyme	1.55	0.61	1243.54	2.54	0.01
VVC rhyme	3.34	0.43	1242.04	7.85	0.00
Syl. position 2	6.32	0.37	1211.66	17.00	0.00
Syl. position 3	7.29	0.39	1242.94	18.80	0.00
Syl. position 4	8.23	0.42	1243.88	19.48	0.00
Syl. position 5	7.61	0.47	1243.19	16.09	0.00
Syl. position 6	7.41	0.65	1243.33	11.46	0.00
Word length 2 syl.	-1.22	0.70	1242.12	-1.73	0.08
Word length 3 syl.	-3.04	0.70	1242.69	-4.32	0.00
Word length 4 syl.	-3.56	0.72	1243.53	-4.95	0.00
Word length 5 syl.	-3.76	0.72	1243.55	-5.18	0.00
Word length 6 syl.	-2.28	0.77	1243.77	-2.96	0.00
Word length 7 syl.	-3.84	0.78	1243.82	-4.93	0.00
Word length 8 syl.	-3.94	0.78	1243.49	-5.05	0.00
Word length 9 syl.	-6.41	1.01	1244.00	-6.33	0.00
VC rhyme : Syl. position 2	-2.41	0.49	1055.39	-4.93	0.00
VV rhyme : Syl. position 2	-0.52	0.74	1236.28	-0.70	0.48
VVC rhyme : Syl. position 2	-3.16	0.55	1238.68	-5.71	0.00
VC rhyme : Syl. position 3	-0.96	0.56	1224.32	-1.73	0.08
VV rhyme : Syl. position 3	0.21	0.71	1243.34	0.30	0.77
VVC rhyme : Syl. position 3	-2.14	0.63	1243.99	-3.41	0.00
VC rhyme : Syl. position 4	0.20	1.00	1233.51	0.20	0.85
VV rhyme : Syl. position 4	0.08	0.76	1243.97	0.10	0.92
VVC rhyme : Syl. position 4	-0.54	0.99	1243.97	-0.54	0.59
VC rhyme : Syl. position 5	-0.14	0.99	1240.93	-0.14	0.89
VV rhyme : Syl. position 5	2.41	1.04	1243.62	2.33	0.02
VVC rhyme : Syl. position 5	-2.84	0.94	1243.58	-3.02	0.00
VC rhyme : Syl. position 6	3.03	1.84	1242.48	1.65	0.10
VV rhyme : Syl. position 6	0.15	1.21	1241.24	0.12	0.90
VVC rhyme : Syl. position 6	0.15	1.42	1243.58	0.10	0.92

Table 31: Pairwise comparisons of mean vowel intensity (in dB) between syllable positions (counted from the end of the word) by rhyme types in single-word data set.

	Estimate	Std. Error	df	t-ratio	p-value
Syl. position 1 vs. 2 for V	-6.321	0.375	1212	-16.864	<.0001
Syl. position 1 vs. 3 for V	-7.295	0.388	1243	-18.795	<.0001
Syl. position 1 vs. 4 for V	-8.232	0.423	1244	-19.442	<.0001
Syl. position 1 vs. 5 for V	-7.609	0.473	1243	-16.08	<.0001
Syl. position 1 vs. 6 for V	-7.413	0.647	1243	-11.451	<.0001
Syl. position 1 vs. 2 for VC	-3.915	0.306	1162	-12.811	<.0001
Syl. position 1 vs. 3 for VC	-6.331	0.404	1206	-15.688	<.0001
Syl. position 1 vs. 4 for VC	-8.428	0.907	1236	-9.293	<.0001
Syl. position 1 vs. 5 for VC	-7.471	0.889	1242	-8.406	<.0001
Syl. position 1 vs. 6 for VC	-10.447	1.734	1243	-6.027	<.0001
Syl. position 1 vs. 2 for VV	-5.804	0.633	1243	-9.164	<.0001
Syl. position 1 vs. 3 for VV	-7.505	0.599	1243	-12.537	<.0001
Syl. position 1 vs. 4 for VV	-8.31	0.654	1244	-12.715	<.0001
Syl. position 1 vs. 5 for VV	-10.02	0.941	1243	-10.648	<.0001
Syl. position 1 vs. 6 for VV	-7.561	1.042	1243	-7.254	<.0001
Syl. position 1 vs. 2 for VVC	-3.165	0.406	1243	-7.792	<.0001
Syl. position 1 vs. 3 for VVC	-5.155	0.502	1244	-10.266	<.0001
Syl. position 1 vs. 4 for VVC	-7.697	0.904	1243	-8.513	<.0001
Syl. position 1 vs. 5 for VVC	-4.765	0.829	1243	-5.748	<.0001
Syl. position 1 vs. 6 for VVC	-7.562	1.28	1243	-5.91	<.0001
Syl. position 2 vs. 3 for VC	-2.416	0.437	1242	-5.533	<.0001
Syl. position 2 vs. 4 for VC	-4.513	0.918	1244	-4.918	0.0003
Syl. position 2 vs. 5 for VC	-3.556	0.903	1244	-3.937	0.0179
Syl. position 2 vs. 6 for VC	-6.532	1.747	1244	-3.739	0.0365
Syl. position 2 vs. 3 for VV	-1.701	0.464	1243	-3.67	0.0462
Syl. position 2 vs. 4 for VV	-2.506	0.527	1243	-4.752	0.0006
Syl. position 2 vs. 5 for VV	-4.216	0.86	1243	-4.902	0.0003
Syl. position 2 vs. 3 for VVC	-1.99	0.512	1243	-3.887	0.0216
Syl. position 2 vs. 4 for VVC	-4.532	0.908	1243	-4.993	0.0002
Syl. position 2 vs. 4 for V	-1.912	0.349	1228	-5.472	<.0001

Table 32: Pairwise comparisons of mean vowel intensity (in dB) between rhyme types by syllable positions (counted from the end of the word) in single-word data set.

	Estimate	Std. Error	df	t-ratio	p-value
V vs. VC in syl. position 1	-1.393	0.373	1109	-3.731	0.0376
V vs. VVC in syl. position 1	-3.337	0.427	1242	-7.821	<.0001
VC vs. VVC in syl. position 1	-1.944	0.33	1235	-5.885	<.0001
VC vs. VV in syl. position 2	-2.049	0.439	1244	-4.664	0.0009
V vs. VV in syl. position 3	-1.764	0.356	1244	-4.95	0.0002
V vs. VV in syl. position 5	-3.964	0.839	1242	-4.723	0.0006

Table 33: Pairwise comparisons of mean vowel intensity (in dB) between word lengths (counted in number of syllables) in single-word data set.

	Estimate	Std. Error	df	t-ratio	p-value
1 vs. 2 syllables	1.218	0.704	1242	1.731	0.7272
1 vs. 3 syllables	3.035	0.703	1243	4.318	0.0006
1 vs. 4 syllables	3.56	0.719	1244	4.950	<.0001
1 vs. 5 syllables	3.755	0.725	1244	5.179	<.0001
1 vs. 7 syllables	3.841	0.781	1244	4.918	<.0001
1 vs. 8 syllables	3.945	0.784	1244	5.034	<.0001
1 vs. 9 syllables	6.406	1.013	1244	6.325	<.0001
2 vs. 3 syllables	1.817	0.227	1243	8.014	<.0001
2 vs. 4 syllables	2.342	0.257	1229	9.102	<.0001
2 vs. 5 syllables	2.537	0.272	1230	9.321	<.0001
2 vs. 7 syllables	2.623	0.401	1210	6.550	<.0001
2 vs. 8 syllables	2.727	0.407	1188	6.706	<.0001
2 vs. 9 syllables	5.188	0.758	1242	6.842	<.0001
3 vs. 5 syllables	0.72	0.223	1244	3.231	0.0344
3 vs. 9 syllables	3.371	0.74	1244	4.553	0.0002
4 vs. 9 syllables	2.847	0.747	1244	3.812	0.0046
5 vs. 9 syllables	2.651	0.74	1244	3.580	0.0107
7 vs. 9 syllables	2.565	0.779	1243	3.292	0.0283
8 vs. 9 syllables	2.461	0.777	1242	3.166	0.0419
6 vs. 7 syllables	1.557	0.441	1243	3.531	0.0128
6 vs. 8 syllables	1.661	0.445	1244	3.733	0.0061
6 vs. 9 syllables	4.122	0.781	1244	5.280	<.0001
4 vs. 6 syllables	-1.276	0.353	1243	-3.618	0.0094
5 vs. 6 syllables	-1.471	0.355	1243	-4.144	0.0012

References

Arnhold, A. (2014), Prosodic structure and focus realization in West Greenlandic, *in* S.-A. Jun, ed., 'Prosodic Typology II', Oxford University Press, New York/Oxford, pp. 216–251.

Arnhold, A. (accepted), Phonology and phonetics, *in* A. Berge, A. Arnhold & N. Trondhjem, eds, 'The Eskimo-Aleut languages', Oxford University Press, Oxford.

Arnhold, A., Compton, R. & Elfner, E. (2018), Prosody and wordhood in South Baffin Inuktitut, *in* M. Keough, N. Weber, A. Anghelescu, S. Chen, E. Guntly, K. Johnson, D. Reisinger & O. Tkachman, eds, 'Proceedings of the Workshop on Structure and Constituency in the Languages of the Americas 21', University of British Columbia Working Papers in Linguistics 46, pp. 30–39.

Baayen, R. H., Davidson, D. J. & Bates, D. M. (2008), 'Mixed-effects modeling with crossed random effects for subjects and items', *Journal of Memory and Language* **59**(4), 390–412.
URL: *http://linkinghub.elsevier.com/retrieve/pii/S0749596X07001398*

Bakovic, E. (1998), *Unbounded stress and factorial typology*, RuLing Papers 1: Working Papers from Rutgers University, Department of Linguistics, Rutgers University, New Brunswick, NJ.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015), 'Fitting linear mixed-effects models using lme4', *Journal of Statistical Software* **67**, 1–48.

Bennett, R. (2012), Foot-conditioned phonotactis and prosodic constituency, Phd thesis, University of California, Santa Cruz.

Bobaljik, J. D. (1996), 'Assimilation in the Inuit languages and the place of the uvular nasal', *International Journal of American Linguistics* **62**(4), 323–350.
URL: *http://www.jstor.org/stable/1265705*

Chung, S. (1983), 'Transderivational relationships in Chamorro phonology', *Language* **59**(1), 35–66.

Compton, R. (2009), 'Contrast in Inuit Consonant Inventories'.
URL: *https://twpl.library.utoronto.ca/index.php/twpl/article/view/6528*

Compton, R. (2012), The syntax and semantics of modification in Inuktitut: Adjectives and adverbs in a polysynthetic language, Phd thesis, University of Toronto.

Compton, R. & Dresher, B. E. (2011), 'Palatalization and "strong i" across Inuit dialects', *The Canadian Journal of Linguistics / La revue canadienne de linguistique* **56**(2), 203–228.
URL: *http://muse.jhu.edu/content/crossref/journals/canadian_journal_of_linguistics/v056/56.2.compton.ht*

Compton, R. & Pittman, C. (2010), 'Word-formation by phase in Inuit', *Lingua* **120**(9), 2167–2192.

Dorais, L.-J. (1976), 'Some phonological characteristics of Eastern Canadian Inuit', *Anthropological Linguistics* **18**(8), 387–392.
URL: *http://www.jstor.org/stable/30027454*

Dorais, L.-J. (1986), 'Inuktitut surface phonology: A trans-dialectal survey', *Internation Journal of American Linguistics* **52**(1), 20–53.
URL: *https://www.jstor.org/stable/1265501*

Dorais, L.-J. (2003), *Inuit uqausiqatigiit: Inuit languages and dialects*, 2nd revised edn, Nunavut Arctic College, Iqaluit.

Dorais, L.-J. & Lowe, R. (1982), 'Les dialectes de l'Arctique de l'ouest', *Études/Inuit/Studies* **6**(2), 127–133.
URL: *https://www.jstor.org/stable/42869357*

Dresher, B. E. & Johns, A. (1995), 'The law of double consonants in Inuktitut', *Linguistica Atlantica* **17**(October 1992), 79–95.

Féry, C. (2010), Indian languages as intonational 'phrase languages', *in* I. Hasnain & S. Chaudhury, eds, 'Problematizing language studies. Festschrift for Rama Agnihotri', Aakar Books, Delhi, pp. 288–312.

Féry, C. (2016), *Intonation and prosodic structure*, Key topics in phonology, Oxford University Press, Oxford.

Fortescue, M. (1980), 'Affix ordering in West Greenlandic derivational processes', *Internation Journal of American Linguistics* **46**(4), 259–278.

Fortescue, M. (1983), 'Intonation across Inuit dialects', *Études/Inuit/Studies* **7**(2), 113–124.

Fortescue, M. (1984), *West Greenlandic*, Croom Helm, London/Sydney/Dover, New Hampshire.

Goldsmith, J. (1976), Autosegmental Phonology, Phd thesis, MIT.

Gordon, M. (1999), Syllable Weight: Phonetics, Phonology, and Typology, Phd thesis, UCLA.

Gordon, M. (2002), 'A factorial typology of quantity-insensitive stress', *Natural Language & Linguistic Theory* **20**(3), 491–552.
URL: *http://www.jstor.org/stable/4048055*

Gordon, M. & Roettger, T. (2017), 'Acoustic correlates of word stress: A cross-linguistic survey', *Linguistics Vanguard* .

Hagerup, A. (2011), A phonological analysis of vowel allophony in West Greenlandic, Ma thesis, NTNU.
URL: *http://hdl.handle.net/11250/244004*

Halle, M. & Marantz, A. (1993), Distributed morphology and the pieces of inflection, *in* K. Hale & S. J. Keyser, eds, 'The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger', Current studies in linguistics 24, MIT Press, Cambridge, MA, pp. 111–176.

Halle, M. & Marantz, A. (1994), 'Some key features of distributed morphology', *MIT working papers in linguistics* **21**, 275–288.

Hayes, B. (1989), 'Compensatory Lengthening in moraic phonology', *Linguistic Inquiry* **20**, 253–306.

Hayes, B. (1995), *Metrical stress theory: Principles and case studies*, University of Chicago Press, Chicago.

Holtved, E. (1964), *Kleinschmidts Briefe an Theodor Bourquin*, Meddelelser om Grønland140(3), Reitzel, Copenhagen.

Hyman, L. M. (2006), 'Word-prosodic typology', *Phonology* **23**(02), 225–257.
URL: *http://www.journals.cambridge.org/abstract_S0952675706000893*

Hyman, L. M. (2009), 'How (not) to do phonological typology: the case of pitch-accent', *Language Sciences* **31**(2-3), 213–238.
URL: *http://dx.doi.org/10.1016/j.langsci.2008.12.007*

Jacobsen, B. (2000), 'The Question of "Stress" in West Greenlandic: An Acoustic Investigation of Rhythmicization, Intonation, and Syllable Weight', *Phonetica* **57**, 40–67.

Jacobson, S. A. (2019), 'The stress conspiracy and stress-repelling bases in the Central Yup'ik and Siberian Yupik Eskimo Languages', **50**(3), 312–324.
URL: *https://www.jstor.org/stable/1265552*

Johns, A. (2006), Ergativity and change in Inuktitut, *in* A. Johns & D. Massam, eds, 'Ergativity: Emerging issues', Studies in natural language and linguistic theory 65, Springer, Dordrecht, pp. 293–311.

Johns, A. (2007), 'Restricting noun incorporation: Root movement', *Natural Language & Linguistic Theory* **25**(3), 535–576.

Kager, R. (2007), 'Feet and metrical stress'.
URL: *http://www.let.uu.nl/ Rene.Kager/personal/Cambridge-Handbook-2007.pdf*

Kaplan, L. D. (1981), *Phonological issues in North Alaskan Inupiaq*, Alaska Native Language Center Research Papers 6, Alaska Native Language Center, Alaska University.

Kaplan, L. D. (1985), Seward Peninsula Inupiaq consonant gradation and its relationship to prosody, *in* M. E. Krauss, ed., 'Yupik Eskimo prosodic systems: Descriptive and comparative studies', Alaska Native Language Center Research Papers 7, Alaska Native Language Center, University of Alaska, Fairbanks, pp. 191–210.

Kleinschmidt, S. (1985), *Grammatik der grönländischen sprache mit theilweisem einschluss des Labradordialects*, G. Reimer, Berlin.

Krauss, M. E. (1985*a*), Introduction, *in* M. E. Krauss, ed., 'Yupik Eskimo prosodic systems: Descriptive and comparative studies', Alaska Native Language Center Research Papers 7, Alaska Native Language Center, University of Alaska, Fairbanks, pp. 1–6.

Krauss, M. E. (1985*b*), Supplementary notes on Central Siberian Yupik prosody, *in* M. Krauss, ed., 'Yupik Eskimo prosodic systems: Descriptive and comparative studies', Alaska Native Language Center Research Papers 7, Alaska Native Language Center, University of Alaska, Fairbanks, pp. 47–50.

Ladd, D. R. (2008), *Intonational phonology*, 2nd ed edn, Cambridge University Press, Cambridge.

Leer, J. (1985*a*), Prosody in Alutiiq (The Koniaq and Chugach dialects of Alaskan Yupik), *in* M. E. Krauss, ed., 'Yupik Eskimo prosodic systems: Descriptive and comparative studies', Alaska Native Language Center Research Papers 7, Alaska Native Language Center, University of Alaska, Fairbanks, pp. 77–134.

Leer, J. (1985*b*), Toward a metrical interpretation of Yupik prosody, *in* M. E. Krauss, ed., 'Yupik Eskimo prosodic systems: Descriptive and comparative studies', Alaska Native Language Center Research Papers 7, Alaska Native Language Center, University of Alaska, Fairbanks, pp. 159–173.

Lehtonen, J. (1970), *Aspects of Quantity in Standard Finnish*, Studia Philologica Jyväskyläensia VI, Gummerus, Jyväskylä.

Lenth, R. (2018), *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.2.3.
URL: *https://CRAN.R-project.org/package=emmeans*

Lipscomb, D. R. (1992), 'Differences and similarities in the prosodic systems of Western and Eastern Eskimo', *Acta Linguistica Hafniensia: International Journal of Linguistics* **25**(1), 63–81.

Martínez-Paricio, V. & Kager, R. (2017), 'Metrically conditioned pitch and layered feet in Chugach Alutiiq', *Loquens* **3**(2), e030.

Mase, H. (1973), A Study of the role of syllable and mora for the tonal manifestation in West Greenlandic, *in* 'Annual Report of the Institute of Phonetics, University of Copenhagen', Vol. 7, University of Copenhagen, Copenhagen, pp. 1–98.

Mase, H. & Rischel, J. (1971), A study of consonant quantity in West Greenlandic, *in* 'Annual Report of the Institute of Phonetics, University of Copenhagen', Vol. 5, University of Copenhagen, Copenhagen, pp. 175–247.

Massenet, J.-M. (1978), Quelques aspects de la quantité vocalique et consonantique en eskimo, PhD thesis, University of Toronto.

Massenet, J.-M. (1980), 'Notes sur l'intonation finale en inuktitut', *Cahier de linguistique* **10**, 195–214.
URL: *http://id.erudit.org/iderudit/800091ar*

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H. & Bates, D. (2017), 'Balancing type I error and power in linear mixed models', *Journal of Memory and Language* **94**, 305–315.
URL: *http://arxiv.org/abs/1511.01864*

McCarthy, J. J. (2003), 'OT constraints are categorical', *Phonology* **20**, 75–138.

McCawley, J. D. (1978), What is a tone language?, *in* V. A. Fromkin, ed., 'Tone: A linguistic survey', Academic Press, Inc., New York, pp. 113–131.
URL: *http://dx.doi.org/10.1016/B978-0-12-267350-4.50009-1*

Miyaoka, O. (1985), Accentuation in Central Alaskan Yupik, *in* M. E. Krauss, ed., 'Yupik Eskimo prosodic systems: Descriptive and comparative studies', Alaska Native Language Center Research Papers 7, Alaska Native Language Center, University of Alaska, Fairbanks, pp. 51–76.

Miyaoka, O. (2012), *A grammar of Central Alaskan Yupik (CAY)*, Mouton Grammar Library 58, De Gruyter Mouton, Berlin.

Miyaoka, O. (2019), 'On Syllable Modification and Quantity in Yuk Phonology', **37**(4), 219–226.
URL: *https://www.jstor.org/stable/1264513*

Morén, B. (1999), Distinctiveness, Coercion and Sonority: A Unified Theory of Weight, Phd thesis, University of Maryland.

Nagano-Madsen, Y. (1992), *Mora and prosodic coordination. A phonetic study of Japanese, Eskimo and Yoruba*, Travaux de l'institut de linguistique de Lund 27, Lund University Press, Lund.

Oxford, W. (2014), Microparameters of agreement: A diachronic perspective on Algonquian verb inflection, Phd thesis, University of Toronto.
URL: *https://tspace.library.utoronto.ca/bitstream/1807/68125/1/Oxford_William_R_201406_PhD_thesis.pdf*

Petersen, R. (1970), 'On phonological length in the eastern Eskimo dialects', *Folk* **11/12**, 329–344.

Pierrehumbert, J. B. (1980), The phonology and phonetics of English intonation, PhD thesis, MIT.

Pigott, P. (2012), Degemination and prosody in Labrador Inuttut: An acoustic study, MA thesis, Memorial University of Newfoundland.

Pirurvik Centre (2015), 'www.tusaalanga.ca'.

Prince, A. (1985), Improving tree theory, *in* 'Annual Meeting of the Berkeley Linguistics Society 11', Berkeley Linguistic Society, Berkeley, pp. 471–490.

Pyle, C. (1970), '{W}est {G}reenlandic {E}skimo and the Representation of Vowel Length', *Papers in Linguistics* **3**(1), 115–146.

R Core Team (2018), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. Version 3.5.1.
URL: *https://www.R-project.org/*

Rice, K. (2000), *Morpheme order and semantic scope*, Cambridge University Press, Cambridge.

Rischel, J. (1974), *Topics in West Greenlandic phonology: regularities underlying the phonetic appearance of wordforms in a polysynthetic language*, Akademisk Forlag, Copenhagen.

Roettger, T. & Gordon, M. (2017), 'Methodological issues in the study of word stress correlates', *Linguistics Vanguard* **3**(1), 1–11.
URL: *http://www.degruyter.com/view/j/lingvan.2017.3.issue-1/lingvan-2017-0006/lingvan-2017-0006.xml*

Rose, Y., Pigott, P. & Wharram, D. (2012), Schneider's law revisited: The syllable-level remnant of an older metrical rule, *in* 'McGill Working Papers in Linguistics', Vol. 22.1, McGill University, Montreal, pp. 1–12.

Rosenthall, S. & van der Hulst, H. (1999), 'Weight-by-position by position', *Natural Language and Linguistic Theory* **17**, 499–540.

Ryan, K. M. (2014), 'Onsets contribute to syllable weight: Statistical evidence from stress and meter', *Language* **90**(2), 309–341.
URL: *http://muse.jhu.edu/content/crossref/journals/language/v090/90.2.ryan.html*

Sadock, J. M. (1980), 'Noun incorporation in Greenlandic: A case of syntactic word formation', *Language* **56**(2), 300–319.
URL: *http://www.jstor.org/stable/413758*

Smith, L. R. (1977), 'Some morphophonemic processes of Labrador Inuttut affixation', *Internation Journal of American Linguistics* **43**(2), 77–84.
URL: *https://www.jstor.org/stable/1264925*

Swift, M. D. (2004), *Time in child Inuktitut: A developmental study of an Eskimo-Aleut language*, Studies On Language Acquisition 24, Mouton de Gruyter, Berlin, New York.

van der Hulst, H. (2014), The study of word accent and stress: past, present, and future, *in* H. van der Hulst, ed., 'Word stress: Theoretical and typological issues', Cambridge University Press, Cambridge, chapter 1, pp. 3–55.

Williams, B. (1982), Some problems in the description of stress in modern Welsh, *in* 'Cambridge papers in phonetics and experimental linguistics, 1', Cambridge University.

Williams, B. (1985), 'Pitch and duration in Welsh stress perception: The implications for intonation', *Journal of Phonetics* **13**(4), 381–406.

Woodbury, A. C. (1987), 'Meaningful phonological processes: A consideration of Central Alaskan Yupik Eskimo prosody', *Language* **63**(4), 685–740.

Yoshiba, H. (1983), Moric phonology: Toward the establishment of a new phonological unit, PhD thesis, University of Washington.