Open Library of Humanities

Acoustics of stress and weight in Central Alaskan Yup'ik

McKinley Alden*, Department of Linguistics, University of Alberta, Edmonton, AB, Canada, malden@ualberta.ca

Anja Arnhold, Department of Linguistics, University of Alberta, Edmonton, AB, Canada, arnhold@ualberta.ca

*Corresponding author.

In Central Alaskan Yup'ik, syllables with long vowels are always stressed, light syllables alternate stress, but only certain closed syllables are stressed. The acoustic correlates of stress, however, have only been the subject of one small-scale preliminary study so far. Moreover, there are divergent accounts of how phonological phenomena such as gemination and syllable closure affect weight. This article presents an acoustic investigation of gemination, stress, and phonemic length. Six Yup'ik recordings were annotated, resulting in a dataset of 2,602 syllable onsets and 2,282 vowels, which were then modelled using linear mixed-effects models. The first part of the study, examining the distribution of gemination as a metrical-adjacent phenomenon, revealed that singleton onsets were shorter than geminated onsets both within feet and across foot boundaries. The main study showed that stressed vowels were longer, louder, and, for short vowels only, higher in f0 than unstressed vowels, while long vowels were longer, louder, and featured greater f0 falls than short vowels. These results corroborate the literature that asserts that long, short unstressed, and short stressed vowels are all produced distinctively, and moreover, that vowel duration is affected by iambic lengthening and syllable closure. The identification of stress correlates and other metrical behaviors examined here set the stage for future prosodic work on Yup'ik.

1. Introduction

The Yupik languages are spoken across northeastern Siberia and western to southern Alaska. There are several varieties of Yupik; this project focuses on Central Alaskan Yup'ik, specifically the General Central dialect (*Yugtun*; henceforth *Yup'ik*). This article seeks to explore the acoustic correlates of stress and weight in Yup'ik by acoustically analyzing a series of recordings housed in the Alaska Native Language Archive (Alaskan Native Language Archive, n.d.). The goals of this paper are to investigate claims regarding the distribution of gemination, which is relevant to questions of stress as a process that affects syllable weight, and to examine the ways in which phonemic vowel length and stress are expressed.

Section 2 introduces Yup'ik stress patterns as described in the literature, as well as the metrical model used to assign stress in the recordings analyzed here. Section 3 summarizes the existing literature on correlates of stress in Yup'ik, which includes only one small-scale acoustic study. Section 4 introduces the materials and methods used in the present study before Section 5 and Section 6 present results regarding gemination and acoustic correlates of stress and vowel length, respectively. Section 7 compares the findings to previous descriptions and discusses their implications before Section 8 concludes.

2. Yup'ik stress patterns

This study utilizes extant metrical accounts of Yup'ik stress to categorically identify the syllables that receive stress in the acoustic dataset. The goal of this section is to provide an overview that informs and justifies the phonetic analysis of stress given in Section 6. To that end, this section will first present the basic facts of Yup'ik stress and the way they have been modelled in the literature, focusing on the metrical framework proposed in Hayes (1995). When predictions made by this model were applied to real-world data (see Section 4), some adjustments were made to the Hayes model; these are also discussed in this section (see supplementary materials in Alden and Arnhold (2022), for more details).

The literature on the Yup'ik stress system can be divided into two main groups: Foot-based metrical frameworks and apodal accounts. The metrical framework assumes the existence of feet as the primary vehicle for stress assignment and uses rules to constrain foot shape, while the apodal approach assigns stress based on a syllable's position within a word using classic phonological rules without making reference to feet. Authors who utilize a metrical framework include Halle (1990); Leer (1985); Miyaoka (1971); and Woodbury (1987), with the most comprehensive and theoretically-founded account presented by Hayes (1995). Authors who adopt the apodal approach to Yup'ik stress include Jacobson (1984, 1990) and Lipscomb (1992). Another account worth mentioning is van de Vijver (1998), which takes principles outlined by Jacobson and Woodbury and adapts them into an Optimality Theory framework (see also Bakovic, 1996).

These scholars agree that in Yup'ik, syllables with long vowels or diphthongs are always stressed, while open syllables with short vowels alternate stress from left to right (see **Table 1**). Intonational phrase-final syllables are destressed; however, unless otherwise stated, in this paper examples are assumed to not be IP-final for the purposes of demonstrating stress patterns.

Example	Transcription	Source: Jacobson & Jacobson, 1995
nuna 'land, village'	[nu.ˈna]	p. 4
patu 'cover'	[pa.ˈtu]	p. 4
quyana 'thank you'	[qu.ˈja.na]	p. 9
acaka 'my paternal aunt'	[a.ˈtʃa.ka]	p. 9
nalluyagucaqunaku 'don't forget it'	[na.ˈɬu.ja.ˈgu.tʃa.ˈqu.na.ˈku]	p. 10
qayaliqataraqama 'whenever I'm going to make a kayak'	[qa.ˈja.li.ˈqa.ta.ˈχa.qa.ˈma]	p. 10
ekumaqatalliniluni 'evidently being about to be in a conveyance'	[ə.ˈku.ma.ˈqa.ta.ˈɬi.ni.ˈlu.ni]	p. 10

Table 1: Examples of stress alternation in Yup'ik words made up of CV syllables.

Several factors complicate what may at first glance seem like a simple iambically alternating stress system. The first of them is the issue of syllable weight. The Yup'ik syllable is (C)V(V)(C), and intrasyllabic consonant clusters are disallowed (Jacobson, 1984, p. 9; Leer, 1985a, p. 164; Lipscomb, 1992, p. 65; Woodbury, 1987, p. 694). Syllables containing long vowels or diphthongs are always heavy and stressed, and open syllables with short vowels are always light, but whether or not the codas of closed syllables contribute to weight is contested. Closed syllables are always stressed word-initially, and while they can receive stress in other positions, they do not always (cf. examples in **Table 2**).

When determining which syllables receive stress in the acoustic dataset, it was important to be able to systematically identify which of these closed syllables receive stress. Most accounts across all frameworks, including Halle (1990), Jacobson (1984), Leer (1985b, 1985c, 1994), van de Vijver (1998), and Woodbury (1987), require complex solutions to resolve the closed syllable's intermediary status—specifying narrow environments as exceptions to general stress tendencies—in order to solve the CVC problem. Here, we follow Hayes (1995) in assuming that codas always contribute to syllable weight (also cf. the less categorical statement by Miyaoka, 2012, p. 222), except in positions of double clash (i.e., between two stressed syllables or between a stressed

Example	Transcription	Source
ayagtuq 'he leaves'	[a.ˈjax.tuq]	p. 18 (Jacobson & Jacobson, 1995)
up'nerkaq[1] 'spring'	[ˈup.nəχ.ˈkaq]	p. 226 (Jacobson & Jacobson, 1995)
akngirtaatnga 'they hurt me (indic.)'	[ˈɑk.ˈŋi̥χ.ˈtaːt.ŋa]	p. 254 (Hayes, 1995)
akngirtatnga 'they hurt me (interr.)'	[ˈɑk.ŋi̥χ.ˈtat.ŋa]	p. 254 (Hayes, 1995)

Table 2: Examples of stress alternation in Yup'ik words containing (C)VC syllables.

syllable and the right edge of the word). In double clash positions, the coda's mora is removed and the closed syllable destressed. For example, compare *ayagtuq* and *up'nerkaq* in **Table 2**: The middle CVC syllable is stressed in the former, but not in the latter, where it appears between two stressed syllables; the mora associated with /χ/ is lost in double-clash. While Hayes (1995) does not explicitly discuss CVVC syllables, he states that codas are always moraic except in double clash positions. We therefore assume that coda weight is consistent across all syllables, including closed, heavy CVVC syllables. Such syllables also lose a mora when in double-clash; however, as their nuclei are long, they retain their stress (see a more detailed discussion in Section 7).

A second complicating factor is that Yup'ik stress assignment is not only influenced by syllable weight, but frequently influences syllable structure in turn, at least according to Hayes' description (1995). Several phenomena are notable here. The first is that the onset of a heavy syllable may geminate to close a preceding open syllable, as illustrated in (1), which is called "automatic gemination" or "pre-long strengthening" (Bakovic, 1996, p. 5; Hayes, 1995, p. 243; Jacobson, 1984; Lipscomb, 1992, p. 70; Miyaoka, 1971, p. 225).[2]

(1)   *maqikaatggun* 'with their (other's) future steambath material'   /ma.qi.**k**aːt.xun/ -> [ma.ˈqi**k**.ˈ**k**aːt.xun] (Hayes 1995, p. 243)

---

[1] The apostrophe serves many functions in Yup'ik orthography. Here, it indicates that the rule that devoices nasals following voiceless plosives does not apply. An example where this rule does apply is akn*girtaatnga* [ˈɑk.ˈŋi̥χ.ˈtaːt.ŋa].

[2] Automatic gemination is not to be confused with the second type of consonant lengthening, lexical (or "marked") gemination, which refers to the presence of geminates pre-syllabification. Marked gemination is reflected orthographically and is phonemically distinctive: for example, compare taq'uq /taqːuq/ 'he quit' to taquq /taquq/ 'braid' (Reed, 1977). The sequence C'V, as in Yup'ik, indicates marked gemination in the orthography. Note that the apostrophe has many functions, and for it to express lexical gemination, the spelling must be in the sequence <C'V>, where C' is the onset of the syllable.

In (1), the second and third syllables /qi.kaːt/ constitute a gemination environment, in which the onset of the latter syllables spreads backwards to close the former. The result is [ma.ˈqik.ˈkaːt.xun]. Such gemination is prosodically impactful, insofar as it creates new CVC syllables, whose clash position must then be considered in the next cycle of stress derivation. In (1), the newly formed CVC syllable *qik* is not in a double-clash position, and so retains its stress.

Hayes (1995) introduces a constraint on the distribution of automatic gemination that is not consistent with the rest of the Yup'ik literature: While accounts such as Bakovic (1996), Jacobson (1984), Lipscomb (1992), and Miyaoka (1971) define the environment for gemination as any (C)V.CVV sequence, Hayes redefines gemination as a metrical process that only occurs within feet, in the environment ((C)V.CVV). The examples in (2) demonstrate gemination occurring within feet, as described by Hayes (2a), as well as across feet, as defined by Miyaoka and other authors (2b).

(2)   Gemination within vs. across feet (gemination marked based on Alden & Arnhold, 2022)
     a.   *ipuun* 'ladle' (i**p.p**uːn) (Jacobson & Jacobson, 1995, p. 493)
     b.   *tegumiaq* 'a thing one is taking along in his hands' (te.gu**m**).(**m**iaq) (Jacobson & Jacobson, 1995, p. 503)

As there were many instances in the acoustic dataset where gemination crossed foot boundaries, part of this study was an examination of the distribution of gemination (see Section 5). The results will show that gemination does occur in any (C)V.CVV environment and is not restricted to occurring within feet; this justifies the broader definition of the gemination environment used in the rest of the study.

The second weight-affecting phenomenon is iambic lengthening, also called "rhythmic lengthening", meaning that stressed short vowels in open syllables are lengthened (Bakovic, 1996; Hayes, 1995; Heinrich, 1979; Lipscomb, 1992; Miyaoka, 1970, 1971, 1985, 2012; Woodbury, 1987, 1995; for an overview of iambic lengthening in other languages, see Hyde, 2011; for specific languages, see Derbyshire, 1985, on Hixkaryana; Chung, 1983; Topping, 1973, on Chamorro; Árnason, 1985, on Icelandic; Michelson, 1988, on Kanien'kéha (Mohawk); Mithun & Basri, 1986, on Selayarese; Gordon & Munro, 2007, on Chickasaw; Nicklas, 1974, 1975, on Choctaw, among others).[3] In this paper, vowels that are lengthened in this way are marked with a half-long diacritic Vˑ, so as to distinguish them from vowels that are underlyingly long.

---

[3] There is also a relationship between iambic lengthening and higher-order prosody: while the Yup'ik destressing of IP-final syllables is a priori unexpected, given general cross-linguistic trends towards final lengthening, there is a tendency for final syllable suppression in languages that employ iambic lengthening (Hayes, 1995; see Buckley, 2019 for a discussion of cross-linguistic word-final vowel length and extrametricality). We thank an anonymous reviewer for pointing out this connection.

Iambic lengthening cannot apply to the schwa vowel /ə/, which, unlike the other Yup'ik vowel phonemes /a/, /i/, and /u/, never appears as long and cannot be lengthened by any phonological or prosodic process. If the nucleus of a stressed syllable is a schwa, additional steps must be taken to ensure that that stress is expressed. In the General Central dialect, the solution is schwa deletion (on different outcomes in other Yup'ik dialects, see Hayes, 1995; Jacobson, 1985; Miyaoka, 1985). To illustrate how schwa deletion works, take *qanruteqaka* /qan.χu.**tə**.qa.ka/ 'I speak about it' (Hayes 1995, p. 255): Iambic footing would result in *[(ˈqan).(χu.ˈ**tə·**).(qa.ˈka·)]. Instead the schwa is deleted, giving [(ˈqan).(ˈχut).(qa.ˈka·)]. Note that deletion of a stressed vowel will lead to the stress migrating to a neighboring syllable; since Yup'ik feet are iambic, stress will shift to the left as a result of the hanging onset attaching to the preceding syllable as its new coda. The closure of a formerly open syllable, as with gemination, triggers re-footing. While Hayes (1995) does not explicitly describe the behavior of stressed /ə/ vowels in closed syllables, occasional instances of stressed CəC syllables in our dataset showed that such instances, while rare, do occur, and follow the general ə-deletion tendency, with the orphaned coda becoming syllabic. For instance, /**məχ**.kaχ.tai.tə.łi.niuq/, where the initial syllable is closed, and thereby heavy and stressed, and contains a schwa nucleus, surfaces as [ˈ**m̩χ**.kaχ.ˈtait.łin.ˈniuq]. Importantly, optional schwa deletion in CəC syllables is non-metrical and thus does not lead to any differences in predicted stress locations.

Finally, some Yup'ik morphemes bear lexical stress (Jacobson, 1984). Lexical stress differs from metrical stress as it is phonemically distinctive and cannot be predicted. Lexically stressed syllables are identified by the sequence (C)V'(C) in the orthography, as in *qavartu'rtuq* [qa.ˈvaχ.ˈ**tuχ**.ˈtuq] 'he keeps on sleeping' (Jacobson, 1995, p. 25). Although not explicitly stated by Hayes, this lexical stress must be assigned before initial footing, meaning that syllables marked as lexically stressed are assigned their own feet before all other syllables, in a sort of pre-initial footing. Lastly, this stress cannot be removed at any point in the derivation, even in clash.

It is worth stating here that, while words can contain multiple stressed syllables, it is uncontroversially claimed by most authors that there is no need to distinguish between primary and secondary stress levels, as Yup'ik stress is non-culminative (Jacobson, 1985; Woodbury, 1987). While rare, this is not unheard of in highly polysynthetic languages (e.g., Blackfoot, Stacy, 2004; Arapaho, Bogomolets, 2011; Mapudungun, Molineaux, 2018; other Yupik languages, Woodbury, 1987). There are claims that the rightmost stress in a Yup'ik word is the most prominent (see Miyaoka, 1970, 1971, 1985). However, these claims are unsubstantiated by acoustic data and are not maintained across the literature. Since there is no basis in the literature by which the distinction between primary and secondary stresses can be made, this paper will employ only primary stress diacritics in its examples, following Yup'ik transcription traditions.

To sum up, according to Hayes (1995), Yup'ik stress is binary, quantity sensitive, iambic, constructed left-to-right, and iterative, with the foot's head being obligatorily heavy. When applying Hayes' metrical model to our data for the main acoustic study presented here, some adjustments were needed to assign labels of stressed vs. unstressed to all syllables. These include definitively assuming that codas always contribute a mora, including in syllables with long nuclei; explicating the consequences of schwa deletion in closed syllables; expanding the automatic gemination environment from (C)V.CVV sequences within a foot to all (C)V.CVV sequences, to more closely align with the rest of the Yup'ik literature (see Section 5); and the explicit inclusion of lexical stress and gemination. For more information about the stress derivation model employed in this study, see Alden and Arnhold (2022).

3. Literature on Yup'ik stress correlates and hypotheses

In acoustic terms, a syllable's nucleus can be made more prominent by increasing its duration, increasing its amplitude, raising its pitch, or any combination thereof (Gordon & Roettger, 2017). Little of the literature discusses which of these strategies Yup'ik employs. Miyaoka (1971) claims that a strong syllable is phonetically accompanied by a higher pitch (p. 220). He goes on to state that the highest pitch in the word, which is associated with main stress, is on the last stressed vowel of the word. Woodbury (1985) concurs, stating that "syllable stress consists of pitch movement (usually upward), represented by Pierrehumbert (1980) as pitch accent, with an increase in duration and amplitude" (p. 695).

There is one acoustic study of Yup'ik stress, presented in Gabas (1996). Gabas used the Computerized Extraction of Components of Intonation in Language (CECIL) software to analyze the acoustic correlates of stress. The study used a very small sample size (135 words, total number of syllables not specified), all of which were spoken in isolation. Nevertheless, Gabas demonstrates a cue hierarchy in Yup'ik stress wherein f0 is primary, duration is secondary, and intensity is tertiary. Furthermore, he proposes that "short" words (six syllables or less) are prosodically different from "long" words (seven syllables or more). Short words begin high in intensity, which slowly falls, and exhibit the highest pitch on the right-most stressed syllable, as posited by Miyaoka (1971). Long words, on the other hand, begin with high intensity and pitch, both of which decrease across the word, while the right-most stressed syllable is slightly higher in pitch but considerably louder than its neighboring syllables. Whether or not these are intonational effects is unclear. The present research seeks to corroborate Gabas' claims and account for intonational finality effects.

Finally, the acoustic differentiation of phonemic length (short vs. long vowels) from metrical length (stressed vs. unstressed vowels) has yet to be addressed in Central Alaskan Yup'ik. It is unclear what the phonetic result of iambic lengthening is, to what degree lengthened vowels are

longer, and whether the resultant length is the same as for phonemically long vowels (but see Koo & Badten, 1974, on Central Siberian Yupik).

Based on extant literature, we predict that stressed syllables are made distinct from unstressed syllables via increased duration and higher f0; furthermore, intensity may behave as a secondary cue. To examine the effects of a higher f0, we investigated f0 maxima as well as f0 falls as a potential indicator of stress. The realization of stress is particularly interesting due to a three-way distinction of syllable types: Short, open syllables (CV); short, closed syllables (CVC); and long syllables (CV:/CVV). For short, open syllables to carry stress, they must be either lengthened or closed via gemination of the following syllable onset, resulting in iambic lengthening on CV syllables. What, then, distinguishes a stressed CV syllable from a CV:/CVV syllable? How do intonational phrase-final effects interact with word-level stress? An acoustic study can clarify these questions and set the stage for future work in Yup'ik prosody.

Our hypotheses regarding the acoustic correlates of stress in Yup'ik are as follows:

H_{DUR}: The presence of stress significantly increases the duration of a syllable's nucleus.

H_{INT}: The presence of stress significantly increases the intensity of a syllable's nucleus.

H_{f0MAX}: The presence of stress significantly increases the maximum f0 of a syllable's nucleus.

H_{f0FALL}: The presence of stress significantly increases the post-peak f0 fall of a syllable's nucleus.

$H_{VowelLength}$: There is an acoustic distinction between long, stressed-short, and unstressed-short vowels, such that underlyingly long vowels are distinct from underlyingly short vowels, and the acoustic values of short vowels are increased by the presence of stress.

4. Materials and methods

The acoustic study in this project analyzed six recordings of Central Alaskan Yup'ik, spoken by four different speakers. All of the recordings are housed in the Alaskan Native Language Archive (Alaskan Native Language Center, n.d.). In this section, we will detail the recordings, as well as our transcription practices and analysis methods.

Of the six recordings, four were educational, meant to be listened to as a supplement to a textbook (Reed, 1977), and two were narratives. Two of the speakers were male and two were female. The advantage of using educational recordings (ANLA identifiers: ANLC3111a, ANLC3111b, ANLC3112a, and ANLC3113a) is that the words are clearly articulated and repeated several times. The narratives, meanwhile, provide many more words in connected discourse; the narratives used in this study include Paschal Afcan's *Napam Cuyaa* (Afcan & Hofseth, 1972) and Annie Blue's *Cikmiumalria Tan'gaurluq Yaqulegpiik-llu* from the book *Cungauyaraam Qulirai:*

Annie Blue's Stories (Blue, 2007) (ANLA identifiers: CY(SCH)967A1972g and CY970B2007, respectively). These six recordings were chosen specifically because they were publicly available for download from the ANLA website and all have written transcriptions.

The first author manually annotated each recording in Praat (Boersma & Weenink, 2020). Tiers were added for the entire intonational phrase (IP), the word level, the syllable level, and the segment level. The most important criteria for transcription were the onset of frication in consonants surrounding the vowels and movement in the higher vowel formants (Turk, Nakai, & Sugahara, 2012). **Figure 1** shows two annotation samples, the words *qayacuar* 'little kayak' and *atu'urkaq* 'article of clothing'. These examples demonstrate several notable transcription practices employed in this study. Firstly, the examples given in **Figure 1** both contain vowels in open and closed syllables (note that syllable boundaries were marked on a separate tier that, for the sake of simplicity, has been excluded from these figures). It also demonstrates how boundaries were placed considering fluctuations in higher formants in neighboring segments when a vowel was beside a voiced continuant: Note the rise in F2 between the first /a/ and /j/ in **Figure 1a**. **Figure 1b**, meanwhile, further shows how boundaries were placed around fricated segments and plosive closures. Both panels also show that the onset of frication and the offset of higher formants outranked F1 and voicing as segmentation criteria, as illustrated by the boundaries between /χ/ and the preceding vowels. Given that this study examines intensity, F0, and duration, we used these criteria systematically so all three measurements could be obtained from the same intervals.

IP boundaries were determined by pauses between words, combined with punctuation cues (commas and periods) in the written narratives. This means that a word in isolation constituted its own IP. Stress correlates were measured only on vowels for consistency and ease of measurement (see Section 7 for details).

At this point, the preliminary study of gemination was performed (see Section 6) to confirm the distribution of geminates in the dataset. The Yup'ik literature presents divergent descriptions of gemination. According to models that rely on syllable shape and environment to predict gemination (Halle, 1990; Jacobson, 1984; Miyaoka, 1971; Woodbury, 1987), the onset consonant preceding a long vowel geminates leftwards when preceded by an open, light syllable. In Hayes' model, this gemination is restricted to occurring only within a foot. As gemination results in the closure of open syllables, and syllable closure can in turn affect the stress environments of surrounding syllables, it was important to identify where gemination occurs prior to performing the full metrical analyses of the tokens for the acoustic study.

The dataset contained a total of 458 IPs, 440 words, and 2,282 vowels. Analysis was done via linear mixed-effect models using the package lme4 in R (Baayen, Davidson, & Bates, 2008; Bates, Mächler, Bolker, & Walker, 2015b, 2015a; R Core Team, 2014). The best models reported below
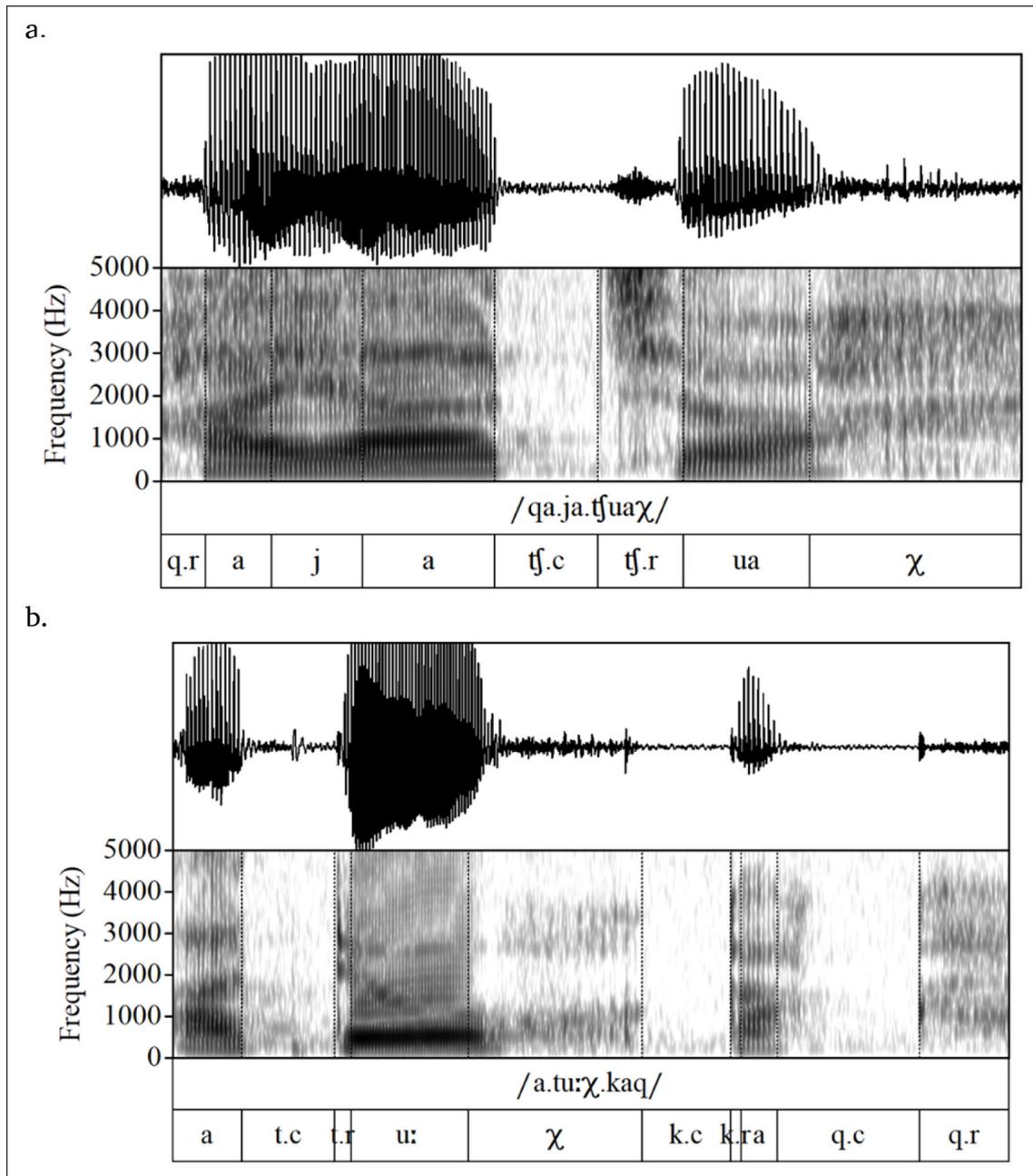
Figure 1: Annotated waveform and spectrogram for the words qayacuar 'little kayak' and atu'urkaq 'article of clothing'. Note that 'C.c' marks the closure phase of a plosive, and 'C.r' marks the release phase.

for each dependent variable were chosen to be only as complex as justified by an improved fit to the data (Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017). Improved model fit was determined by ANOVA comparison between models with or without each added variable. If

the ANOVA did not reveal significant differences between two models, the model without the added variable was preferred. Where ANOVA comparison indicated a significant difference, the model with a higher log likelihood and a lower Akaike Information Criterion (AIC) was chosen. Model selection began with a model containing only the most relevant predictor as a fixed effect, and as random intercepts speaker and genre (educational or narrative). Models of potential stress correlates as reported in Section 6 additionally included vowel quality (/a, i, u/ or /ə/; diphthongs were coded for the quality of their first segment (e.g., /au/ was simplified to /a/, but retained its long specification) as a random intercept. Starting with this model, the fixed variables were forward fitted, so long as each addition improved model fit; each results section for the gemination and stress studies will describe the tested fixed effects variables for each model. The random-effects structure was then forward fitted by stepwise adding by-subjects random slopes for each of the fixed effects variables. However, none of the models with random slopes converged, so all models below contain only random intercepts. Finally, we tested whether all random intercepts significantly contributed to model fit and simplified models where appropriate. The dependent and independent variables for each resulting best model are specified in the results (see Section 5 and Section 6).

After model selection, the residuals were plotted and datapoints with residuals more than 2.5 standard deviations from zero were trimmed. The models were then refit to the trimmed dataset. Information about the number of trimmed data points for each model can be found in the table captions for each model below. Additionally, p-values for fixed effects were obtained with the package lmerTest (Kuznetsova, Brockhoff, & Christensen, 2017). For models containing significant interactions (see Section 5 and Section 6.1), pairwise comparisons were performed with the lsmeans function from the package emmeans (Lenth, 2022), to further examine the interactions, using the Tukey method for p-value adjustments and the (default) Kenward-Roger method for calculating degrees of freedom.

5. First study: Gemination and onset length

In order to examine the distribution of gemination, prior to stress derivation all onsets in the dataset were isolated, resulting in a total of 2,602 consonants;[4] see **Table 3** for distribution. Each onset consonant was then assigned one of three categorical labels: One group in which our predictions matched Hayes', in which geminates occurred within a foot (N = 179, labelled 'agreed'); one group in which we predicted gemination where Hayes would not, across foot boundaries (N = 71, 'predicted'); and one group of uncontroversial non-geminates (N = 2352,

---

[4] The disparity between the number of onsets and vowels in the dataset is due to the fact that Yup'ik vowels, especially schwas, are often deleted, and that syllables with onsets are more common than syllables without them, which are illegal word-medially. Syllabic consonants, which are also frequent as a result, were excluded in the data for this article.

'none'). Secondly, each onset was assigned a categorical label based on manner of articulation in order to examine the degree to which duration is affected by consonant identity. Onsets were either plosive closures (N = 492), fricatives (N = 1379), nasals (N = 459), or approximants (N = 272). **Figure 2** shows the distribution of onset duration for the different gemination and manner categories.

Gemination	Manner of Articulation			
	Approximant	Fricative	Nasal	Plosive Closure
None	247	1252	410	443
Predicted	5	34	25	7
Agreed	20	93	24	42

Table 3: Distribution of onsets by manner and gemination.
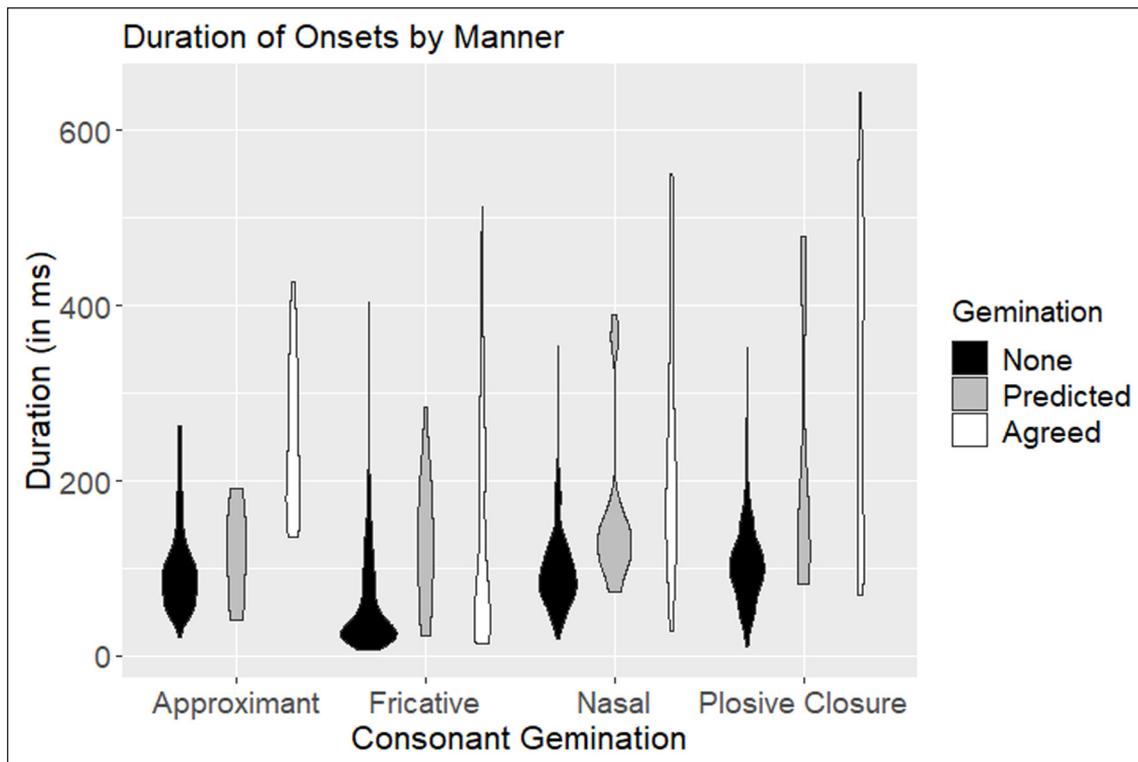
Figure 2: Durations of onsets by manner and predicted gemination.

These onsets were then compared via a mixed-effects linear regression model in order to examine the degree to which duration is affected by geminate status, and the degree to which manner of articulation plays a role in this effect. The base model included gemination (agreed, predicted, none), manner (fricative, approximant, nasal, plosive closure), and syllable stress as fixed effects; model comparison showed that all predictors contributed significantly to model fit. While both speaker and genre as random effects were tested, only the inclusion of the speaker as a random effect improved model fit. The model was then releveled for both gemination and manner of articulation, such that the intercept in **Table 4** is an unstressed fricative geminate within a foot (in the 'agreed' category of geminates). Finally, the model was further improved by including an interaction between gemination and manner.

Gemination Model	Estimate	Std. Error	Df	t value	Pr($>$|t|)
Intercept (Agreed Geminated Fricative)	102.4313	14.2775	4.0535	7.1743	0.0019
Gemination (None)	–37.7734	5.5226	2504.017	–6.8398	9.93E-12
Gemination (Predicted)	23.1689	9.6902	2503.993	2.391	0.0169
Manner (Approximant)	146.4164	11.5735	2504.222	12.651	$<$2E-16
Manner (Nasal)	97.8518	12.047	2504.004	8.1225	7.09E-16
Manner (Plosive Closure)	237.7714	11.5671	2504.363	20.5559	$<$2E-16
Gemination (None): Manner (Approximant)	–104.852	12.0068	2504.197	–8.7327	$<$2E-16
Gemination (Predicted): Manner (Approximant)	–141.028	24.8306	2504.033	–5.6796	1.51E-08
Gemination (None): Manner (Nasal)	–56.6288	12.3329	2503.997	–4.5917	4.61E-06
Gemination (Predicted): Manner (Nasal)	–84.7549	17.4622	2503.981	–4.8536	1.29E-06
Gemination (None): Manner (Plosive Closure)	–190.803	11.8459	2504.404	–16.1071	$<$2E-16
Gemination (Predicted): Manner (Plosive Closure)	–229.733	24.8575	2504.062	–9.242	$<$2E-16
Syllable Stressed	4.9129	1.8832	2503.953	2.6087	0.0091

Table 4: Fixed effects summary of best model of onset duration. Trimming removed 82 data points (3.15%).

The significant interaction between manner and gemination, as given in **Table 4**, shows that differences between gemination categories were not equally pronounced within the individual manners of articulation. This is further reflected in **Table 5**, which summarizes the gemination results by manner of articulation. Interestingly, geminated onsets in the 'agreed' category have much broader ranges than the other two categories, and there is a notable overlap across categories of gemination for all manners (cf. **Figure 2**).

Approximant					
Contrast	**Estimate**	**SE**	**Df**	**t ratio**	**p value**
None vs. Predicted	–26.0	21.10	2504	–1.230	0.4356
None vs. Agreed	–143.2	10.90	2504	–13.138	<.0001
Predicted vs. Agreed	–117.2	23.31	2504	–5.030	<.0001
Fricative					
Contrast	**Estimate**	**SE**	**Df**	**t ratio**	**p value**
None vs. Predicted	–61.3	8.36	2504	–7.334	<.0001
None vs. Agreed	–70.8	5.79	2504	–12.239	<.0001
Predicted vs. Agreed	–9.5	9.96	2504	–0.953	0.6066
Nasal					
Contrast	**Estimate**	**SE**	**Df**	**t ratio**	**p value**
None vs. Predicted	–32.7	10.27	2504	–3.187	0.0042
None vs. Agreed	–94.1	11.26	2504	–8.360	<.0001
Predicted vs. Agreed	–61.4	14.81	2504	–4.149	0.0001
Plosive Closure					
Contrast	**Estimate**	**SE**	**df**	**t ratio**	**p value**
None vs. Predicted	–22.9	20.98	2504	–1.089	0.5210
None vs. Agreed	–104.8	10.70	2504	–9.792	<.0001
Predicted vs. Agreed	–81.9	23.30	2504	–3.517	0.0013

Table 5: Pairwise comparison results for gemination. Negative estimates correspond to smaller values for the left one of the two compared factor combinations.

Table 5 presents the results of the pairwise comparisons, which tested the significance of all gemination contrasts within the individual manners of articulation. These results are most reliable for fricatives, as they were the most frequent manner of articulation in the dataset (c.f. **Table 3**). Within the category of fricative, non-geminated onsets were significantly shorter than predicted geminates, as well as being shorter than agreed geminates. Crucially, there was no significant difference between predicted and agreed geminates. Together, these results indicate that gemination occurs both within and across foot boundaries, with no significant difference between the two. Within nasals, there were significant differences between all three gemination categories. Non-geminates were significantly shorter than both predicted and agreed gemination, and predicted geminates were significantly shorter than agreed ones. Finally, within approximants and plosive closures, predicted geminates differed significantly from agreed geminates, but their difference from non-geminates did not reach significance. This would imply that for approximants and plosive closures, geminates do not cross foot boundaries; however, those results are unreliable due to the critically low number of predicted geminates within these categories, with only five and seven predicted geminates, respectively. For this reason, we take the fricative (and nasal) results as the only reliable outcomes of the pairwise comparisons.

The results of this gemination study show that Hayes was not wrong to attest that geminated onsets are longer within feet; however, consonants in any (C)V.CVV sequence, indicated by the 'predicted' portion of **Figure 2**, are longer than other onsets, justifying their analysis as geminates regardless of their relationship to foot boundaries and stress. Because of Yup'ik's interactions between vowel length, syllable closure, and stress (recall Section 2), broadening the environment for automatic gemination has significant effects on stress derivation. More opportunities for the creation of new closed syllables create more clash environments in which closed syllables are destressed. Together, the observations that automatic gemination is not limited by foot boundaries and that it affects syllable shapes lead to the conclusion that automatic gemination is a phonological process with metrical consequences. This diverges from Hayes' account, which asserts that gemination is already beholden to phonological structures, namely the foot, before it applies. However, these findings are in line with the rest of the literature on automatic gemination in Yup'ik.

One question, then, concerns the motivating factors of gemination. There are three options. First, it could be an entirely phonetic process, with no mind paid to the phonology (similarly to sub-phonemic duration adjustments which shorten the overall duration of one of two adjacent CVVC syllables in Kalaallisut; see Jacobsen, 2000). Second, it could be triggered by specific phonological environments without reference to foot structure (similar to most scholars' analyses of the Law of Double Consonants, or Schneider's Law, in Nunavik Inuktitut and Labrador Inuttut, cf. (e.g., Dresher & Johns, 1995; Rose, Pigott, & Wharram, 2012). Third, gemination could indeed be contingent on the presence of a foot, as assumed by Hayes, in which case the lengthening across

foot boundaries would need to be explained by appealing to another mechanism. Of these, the third option can probably be discarded as unnecessarily complex in the absence of firm acoustic evidence for a difference between gemination within and across foot boundaries. Therefore, the motivating factor of gemination seems to be the presence of a V.CVV environment, as is assumed in the literature on Yup'ik.[5] Regarding the first two options, the evidence that gemination is phonological (and involves the addition of a mora) can be seen insofar as syllables that are closed via gemination and are *not* in double-clash environments behave as heavy for the purpose of stress assignment, as described by Hayes. The mora is always added to syllables preceding geminated onsets, but it is not subsequently removed unless in double-clash. This implies a rule order wherein gemination must occur before double-clash resolution.

To conclude, while we leave the precise analysis of Yup'ik gemination across foot boundary as a subject for future research, based on the preliminary findings presented here, we treated any consonant in a V.CVV environment as geminated for the purpose of stress assignment as evaluated in Section 7. Note, however, that due to the small number of geminates straddling foot boundaries in the present dataset, the impact of this decision should be relatively minor.

6. Acoustic correlates of stress

The goal of the study in this section is to examine the acoustic correlates of stress in spoken Yup'ik. All 440 words in the dataset were taken through the stress derivation process as described in the User's Guide to Central Alaskan Yup'ik Stress Derivation (Alden & Arnhold, 2022), which is based on Hayes' (1995) description of the Yup'ik stress system (cf. Section 3). Each syllable in the output was thereby assigned a label of 'stressed' or 'unstressed'. While it would have been ideal for these stress judgments to be corroborated by native speakers, this was not possible. This is a limitation in this study; however, the use of a guide that was developed, tested, and adapted to be as accurate as possible to the spoken data, based on established descriptions in the literature, does ensure consistency in the stress labels.

Once stressed syllables were identified, a Praat script (Arnhold, 2018) took measurements of duration, f0, and intensity for the sum total 2,282 vowels in the data corpus. Words included in the dataset were between one and nine syllables long; outlier words ten syllables or longer

---

[5] Gemination is but one form of onset fortition attested in the Yupik language family. Alutiiq has been claimed to feature onset fortition in order to demarcate foot boundaries (Leer, 1985a, 1985b, 1994). Leer distinguishes between three levels of consonant length: short, fortified, and geminated. Gemination in Alutiiq is limited to the word-initial environment #V.CVV, but fortition affects all foot-initial consonants. Unlike gemination, Leer does not connect onset fortition with moraic weight, i.e. he does not assume that fortified onset consonants are truly geminated or moraic. The correlates of Alutiiq fortition are not clear, though it may be related to voicing and preclosure (non-moraic onset lengthening), nor is it consistent across speakers. We are planning a follow-up investigation into Alutiiq onset fortition, which would shed light onto other forms of onset fortition in the language family and their relationship to gemination, and the degree to which gemination in this study is similar to onset fortification and gemination processes in Alutiiq.

were removed (four words, 44 vowels total). As a result, there were 436 words and 2,238 vowels in the final dataset, including 854 /a/ vowels, 560 /i/ vowels, 648 /u/ vowels, and 176 schwas. Within the 2,238 vowels, 506 were long (and thereby obligatorily stressed), 785 were short and stressed, and 947 were short and unstressed, cf. **Table 6**. Altogether around 80% of evaluated vowels appeared in first, second or third syllables. In spite of the use of educational materials, in which words were produced in isolation, only 7.2% of analyzed vowels appeared in IP-final syllables.

For fundamental frequency, there were occasional cases where the measurement script could not return values (excessive creak, noise, poor recording quality, et cetera). With these errors additionally removed, the resultant dataset used for the fundamental frequency models contained 2,166 vowels.

Data Metric	Number of Tokens	Percentage
Total vowels	2,238 (2,166)	–
Long vowels	506 (504)	22.6%
Short stressed vowels	785 (757)	35.08%
Short unstressed vowels	947 (905)	42.3%
Word-initial syllables	733 (707)	32.8%
Syllables in position 2	666 (649)	29.7%
Syllables in position 3	392 (375)	17.5%
Syllables in position 4	213 (207)	9.5%
Syllables in position 5	128 (124)	5.7%
Syllables in position 6	54 (53)	2.4%
Syllables in position 7	31 (30)	1.4%
Syllables in position 8	17 (17)	0.8%
Syllables in position 9	4 (4)	0.2%
IP-final syllables	162 (160)	7.2%
Non-IP-final syllables	2076 (2006)	92.8%

Table 6: Data metrics for the acoustic investigation. The number of tokens in parenthesis are the amount considered in the analysis of f0 measurements.

Analysis was done via linear mixed-effect models, as described in Section 4. The following dependent variables were modeled individually: Vowel duration (in ms), mean intensity (in dB, scaled to a reference level of 70 dB), f0 maximum (in semitones, relative to a reference frequency of 100 Hz), and f0 fall (in st) from maximum f0 to the f0 at 80% of the vowel's total duration. The 80% mark was chosen over the end of the vowel to reduce segment-final creak or effects from neighboring consonants. The mean intensity was measured over the central 50% of the vowel duration; the first and last 25% of the vowel were excluded in the analysis to avoid confounds from transitions to and from segments preceding and following the vowel. This was not necessary for fundamental frequency, as the chosen measure was the pitch maximum, not mean pitch. The models employed for this study were meant to assess the relationship between a vowel's acoustic characteristics and stress. To test the hypotheses shown in Section 4, the independent variables included whether a syllable was stressed and the underlying length of the vowel.

Model selection always began with a model containing the ternary stress-length distinction (long, stressed-short, unstressed-short) as a fixed effect, releveled with stressed-short as the intercept, and as random intercepts speaker, genre, and vowel quality. The fixed variables were forward fitted, so long as each addition improved model fit. The tested fixed variables included the syllable's position within the IP (final or non-final), the syllable number counted from the left edge of the word (as a factor), the number of syllables within the word (also as a factor), whether the syllable contained an onset, and whether the syllable contained a coda. We also tested the contribution of random intercepts to the fit and found that genre only significantly improved model fit for intensity; genre is not included as an intercept in any of the other models for this reason. Once the best random effects were accounted for in the model, backward-fitting for the stress-length effects was done to ensure that the distinction was critical to the models.

The following subsections present results of the statistical modelling for duration, intensity, maximum f0 and f0 fall, respectively; the intercept is always the vowel of a short, stressed, non-IP-final syllable with both an onset and a coda. In the following sections, each line of a table represents a comparison to the intercept: In **Table 7**, for example, the first two lines indicate the estimated duration of a short, stressed vowel (125.7905 ms) and the duration of a long vowel (estimated by the model as 86.6617 ms longer than the intercept, or 212.4522 ms).

6.1. Duration

For the linear mixed-effects models of vowel duration, tested predictors included whether the syllable was stressed and phonemic length (combined into a single factor as described above); IP finality; syllable position relative to the left edge; total number of syllables in the word; presence of an onset; presence of a coda. For duration, all factors improved model fit. Including genre of the audio as a random intercept did not improve model fit, and so this term was excluded in the final model reported in **Table 7**.

Duration Model	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept (Short Vowel, Stressed)	125.7905	11.4606	11.7826	10.976	1.54E-07
Vowel Length (Long Vowel, Stressed)	86.6617	24.0265	3.0165	3.607	0.03626
Vowel Length (Short Vowel, Unstressed)	–32.969	2.5207	2147.1971	–13.079	<2.00E-16
Syllable Number (2)	3.8006	2.7959	2148.6001	1.359	0.17418
Syllable Number (3)	4.9795	3.4772	2147.5454	1.432	0.15227
Syllable Number (4)	11.6831	4.2485	2147.8389	2.75	0.00601
Syllable Number (5)	7.3923	5.233	2145.985	1.413	0.15791
Syllable Number (6)	2.8677	7.4937	2145.2682	0.383	0.702
Syllable Number (7)	5.8321	9.704	2144.6718	0.601	0.5479
Syllable Number (8)	–7.8135	13.1426	2144.4838	–0.595	0.55223
Syllable Number (9)	–14.9254	25.3569	2143.8695	–0.589	0.55618
Word length (2 syllables)	–22.9186	7.6455	2147.3254	–2.998	0.00275
Word length (3 syllables)	–32.7853	7.815	2147.5667	–4.195	2.84E-05
Word length (4 syllables)	–46.7042	8.0584	2146.7237	–5.796	7.81E-09
Word length (5 syllables)	–51.0858	8.1381	2147.0861	–6.277	4.16E-10
Word length (6 syllables)	–47.0694	8.6077	2146.1206	–5.468	5.07E-08
Word length (7 syllables)	–54.2454	9.2366	2146.8828	–5.873	4.95E-09
Word length (8 syllables)	–37.9638	9.5978	2147.1316	–3.955	7.89E-05

(Contd.)

Duration Model	Estimate	Std. Error	df	t value	Pr(>|t|)
Word length (9 syllables)	–61.6575	12.2941	2147.2907	–5.015	5.73E-07
Onset (No onset)	0.2667	3.0478	2146.4982	0.088	0.93027
Coda (No coda)	44.1644	2.3908	2139.2113	18.472	<2.00E-16
IP Position (Final)	49.003	4.4385	2148.564	11.041	<2.00E-16

Table 7: Fixed effects summary of best linear mixed-effects model of duration. Trimming removed 61 data points (2.7%).

Table 7 shows that long vowels were significantly longer than stressed short vowels, while unstressed short vowels were significantly shorter than stressed short vowels; however, the boxplots in Figure 3 below demonstrate that the ranges of all three categories overlap considerably. Relevelling also confirmed that unstressed short vowels had shorter durations than stressed long vowels (estimate = –86.6618; std. error = 24.026; df = 3.0166; t = –3.607; Pr(>|t|) = 0.03625). While the result that long vowels are longer than short vowels is not surprising, it does demonstrate that the result of iambic lengthening is not the same as phonemic length. Another notable distinction is between stressed-short and unstressed-short vowels. The fact that the only difference between the two is stress, and that that difference is statistically significant, demonstrates that a change in duration is in fact correlated with stress in Yup'ik.

The rest of the results in Table 7 suggest that vowels got longer the further they got from the left edge of the word, but this effect was only significant for the fourth syllable when compared to the first syllable. Word length also affected vowel duration, with all other word lengths having significantly shorter vowel durations than monosyllabic words. The estimates indicate that the shortening compared monosyllabic words increases with each extra syllable, although six-syllable and eight-syllable words form an exception to this generalization, likely due to a paucity of data for exceptionally long words (cf. Table 6). Regarding syllable structure, the intercept was a vowel with an onset and a coda: While the absence of an onset did not significantly affect length, the vowel's duration was significantly longer when a coda was absent. Finally, IP-final vowels were significantly longer than non-final vowels, in line with cross-linguistic tendencies towards final lengthening (see overview in Fletcher, 2010).

Following up on these results, in order to specifically assess iambic lengthening, a linear mixed-effect model examined the extent to which syllable closure has a measurable effect on the duration of vowels in different phonemic length and stress categories. If iambic lengthening applies only to Yup'ik stressed short vowels, as described in the literature, then we expect that

syllable closure will have a noticeable effect on stressed short vowels that is distinct from syllable closure's effect on long or unstressed short vowels.

The model tested the duration of vowels by phonemic length and whether a coda was present in the syllable, with speaker and vowel phoneme as random effects. The fixed effects additionally included an interaction between the stress-length distinction and presence/absence of a coda, which significantly improved model fit. The intercept of this model was a stressed short vowel in a closed syllable. **Table 8** provides the distribution of the data, and the model summary is given in **Table 9**.

Stress-Length Distinction	**Syllable Closure**	
	Closed	**Open**
Stressed Long	169	337
Stressed Short	497	288
Unstressed Short	243	704

Table 8: Distribution of syllable closure and length for pairwise comparisons.

Duration Model	**Estimate**	**Std. Error**	**df**	**t value**	**Pr(>|t|)**
Intercept (Stressed Short Vowel, Coda Present)	90.843	13.848	5.793	6.560	0.000693
Stress-Length (Stressed Long)	83.514	4.101	2165.222	20.363	<2.00E-16
Stressed-Length (Unstressed Short)	−10.541	3.582	2164.978	−2.942	0.003291
Coda (No Coda)	64.396	3.416	2165.808	18.851	<2.00E-16
Stress-Length (Stressed Long): Coda (No Coda)	−24.811	5.564	2165.071	−4.459	8.65E-06
Stress-Length (Unstressed Short): Coda (No Coda)	−38.311	4.821	2165.076	−7.946	3.07E-15

Table 9: Fixed effects summary of linear mixed-effects model of syllable closure and duration. Trimming removed 61 data points (2.7%).

As the interaction between the stress-length distinction and closure was significant, pairwise comparisons were run in order to examine the effect of closure across the stress-length categories,

see **Table 10**. For the purposes of examining iambic lengthening, the crucial result in **Table 10** is that stressed short vowels in open syllables are longer than those in closed syllables, as expected. However, this holds true for all three stress-length categories. As reflected in **Figure 3**, the closure of a syllable significantly affected the duration of a vowel, regardless of that vowel's underlying length or stress: Vowels in closed syllables consistently had shorter durations than those in open syllables.

Stressed Long Vowel					
Contrast	**Estimate**	**SE**	**Df**	**t ratio**	**p value**
Closed vs. Open	–39.6	4.40	2165	–8.994	<.0001
Stressed Short Vowel					
Contrast	**Estimate**	**SE**	**Df**	**t ratio**	**p value**
Closed vs. Open	–64.4	3.42	2166	–18.844	<.0001
Unstressed Short Vowel					
Contrast	**Estimate**	**SE**	**Df**	**t ratio**	**p value**
Closed vs. Open	–26.1	3.45	2166	–7.554	<.0001

Table 10: Pairwise comparisons for syllable closure by phonemic length and stress. Negative estimates correspond to smaller values for the left one of the two compared factor combinations.

A second set of pairwise comparisons, given in **Table 11**, shows the differences in duration between stress-length categories within closed and open syllables, respectively. Note that the difference in duration between stressed and unstressed short vowels is significant also in closed syllables, where iambic lengthening does not apply. This shows that the distinction between the two categories is due to more than iambic lengthening alone. Specifically, it shows that stress significantly distinguishes short vowels even when iambic lengthening is not involved. However, the effect of iambic lengthening is observable insofar as the estimate for the difference between the two categories in open syllables (48.9 ms) is much larger than the difference in closed syllables (10.5 ms). Here, the former seems to be a combination of the effects of both iambic lengthening and stress: The larger duration effect may be the result of two lengthening processes affecting the vowel. The latter, meanwhile, is presumably the lengthening effect of stress alone— or, alternatively, it may be the case that in closed syllables, some of the stress effect appears on the coda. The critical result is that the ternary stress-length distinction is preserved even when iambic lengthening does not apply (i.e., regardless of syllable closure).

Figure 3: Boxplot of the vowel durations for each of three different stress-length combinations by syllable closure.

Closed Syllable					
Contrast	**Estimate**	**SE**	**df**	**t ratio**	**p value**
Stressed Long vs. Stressed Short	83.5	4.10	2165	20.359	<.0001
Stressed Long vs. Unstressed Short	94.1	4.63	2166	20.308	<.0001
Stressed Short vs. Unstressed Short	10.5	3.58	2165	2.942	0.0092
Open Syllable					
Contrast	**Estimate**	**SE**	**df**	**t ratio**	**p value**
Stressed Long vs. Stressed Short	58.7	3.73	2164	15.718	<.0001
Stressed Long vs. Unstressed Short	107.6	3.14	2166	34.213	<.0001
Stressed Short vs. Unstressed Short	48.9	3.23	2166	15.118	<.0001

Table 11: Pairwise comparisons for phonemic length and stress by syllable closure. Negative estimates correspond to smaller values for the left one of the two compared factor combinations.

Altogether, the results in this section indicate that vowel duration serves multiple functions: The expression of syllable structure, such that the presence of a coda shortens the vowel, and the expression of a syllable's stress, such that stressed syllable vowels are longer than unstressed ones. In this way, duration is an important acoustic correlate for communicating syllable shape information as well as stress in Yup'ik. Moreover, vowel duration reliably cues phonemic vowel length at the same time.

6.2. Intensity

Modelling of intensity tested the same fixed and random effects as for duration. The only predictor that did not improve model fit, and was therefore excluded from the final linear mixed-effects model (**Table 12**), was the sum total number of syllables in the word. Note also that the intensity model is the only one to incorporate genre as a random effect. While this effect was tested in every model, it only contributed to model fit for intensity. This is likely due to differences in microphone sensitivity and audio quality among the recordings, where the educational materials were very loud and clear, while the narrative materials were often much quieter, which also affected the scaled intensity values evaluated here.

Intensity Model	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept (Short Vowel, Stressed)	78.8326	4.0547	1.0497	19.442	0.028523
Vowel Length (Long Vowel, Stressed)	1.2004	0.221	2170.4585	5.431	6.21E-08
Vowel Length (Short Vowel, Unstressed)	–1.3019	0.1949	2171.8781	–6.68	3.03E-11
Syllable Number (2)	–0.3716	0.2114	2172.9126	–1.758	0.078886
Syllable Number (3)	–1.2176	0.2493	2172.46	–4.885	1.11E-06
Syllable Number (4)	–2.3477	0.3015	2172.9078	–7.786	1.06E-14
Syllable Number (5)	–3.3722	0.3693	2172.1623	–9.132	<2.00E-16
Syllable Number (6)	–3.6842	0.5327	2171.5264	–6.916	6.07E-12
Syllable Number (7)	–4.2178	0.6789	2170.6194	–6.213	6.23E-10
Syllable Number (8)	–5.2103	0.9147	2170.6231	–5.696	1.39E-08

(Contd.)

Intensity Model	Estimate	Std. Error	df	t value	Pr(> |t|)
Syllable Number (9)	−2.0675	1.8365	2169.9757	−1.126	0.260387
Onset (No onset)	−0.9202	0.2348	2168.074	−3.919	9.15E-05
Coda (No coda)	0.6399	0.1819	2170.7299	3.519	0.000443
IP Position (Final)	−3.9498	0.3345	2172.5633	−11.809	<2.00E-16

Table 12: Fixed effects summary of best linear mixed-effects model of intensity. Trimming removed 40 data points (2.11%).
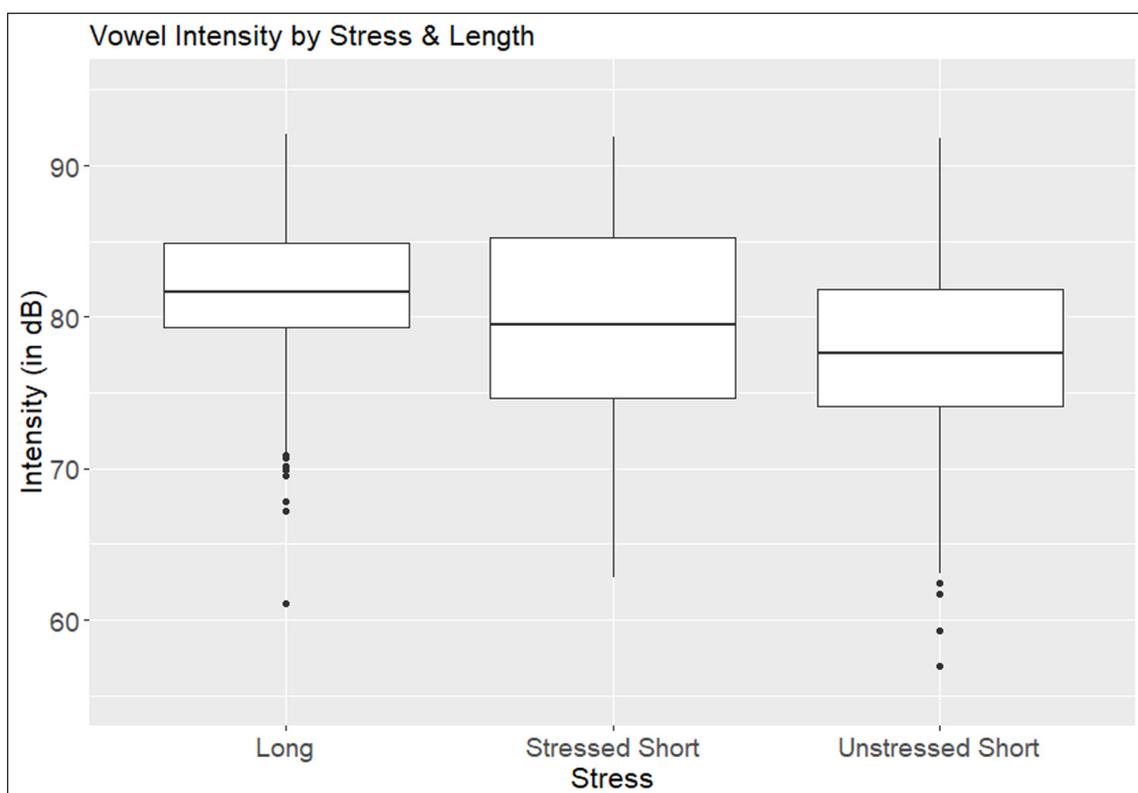
Figure 4: Boxplot of the vowel intensity for each of three different stress-length combinations.

The final model for intensity in **Table 12** showed that long vowels were significantly louder than stressed short vowels, while unstressed short vowels were significantly less loud than stressed short vowels, as also illustrated in **Figure 4**. Relevelling also confirmed that unstressed short vowels had significantly lower intensities than stressed long vowels (estimate = −1.2004;

std. error = 0.2210; df = 2170.46; t = –5.431; Pr(>|t|) = 2.61e-08). Like duration, then, intensity is associated with the production of both long vowels and stressed vowels.

In the rest of **Table 12**, decreasing intensity estimates for syllable number mean that vowels are quieter the later they come in the word, relative to an initial syllable. This difference was significant except for syllables in second position, which only differed marginally from initial ones, and for the ninth syllable from the left edge. Syllables in second position come early in the word and are likely affected by the intensity of the initial syllable. That the effect was not significant for ninth syllables is likely due to the low number of nine-syllable words in the data. While vowels without onsets had significantly lower intensity than those with onsets, vowels in an open syllable were louder than in a closed one. The last vowels in IP-final words were significantly quieter than non-IP-final vowels.

6.3. Maximum f0

The tested factors that did not improve model fit for maximum f0 include the presence of an onset and total number of syllables; as with the duration model, recording genre as a random effect did not improve model fit, and so it was excluded in the final maximum f0 model.

Maximum f0 Model	Estimate	Std. Error	Df	t value	Pr(>|t|)
Intercept (Short Vowel, Stressed)	8.35351	1.21844	3.42104	6.856	0.004115
Vowel Length (Long Vowel, Stressed)	–0.65611	0.51476	3.02931	–1.275	0.291437
Vowel Length (Short Vowel, Unstressed)	–1.51746	0.157	2077.69558	–9.665	<2.00E-16
Syllable Number (2)	–0.04885	0.16507	2075.45388	–0.296	0.767298
Syllable Number (3)	0.05506	0.19338	2078.1831	0.285	0.775893
Syllable Number (4)	–0.58626	0.23973	2078.35679	–2.446	0.014547
Syllable Number (5)	–1.10458	0.29392	2077.37198	–3.758	0.000176
Syllable Number (6)	–0.44352	0.42786	2076.82697	–1.037	0.300039
Syllable Number (7)	–0.94626	0.55937	2075.36566	–1.692	0.090863

(Contd.)

Maximum f0 Model	Estimate	Std. Error	Df	t value	Pr(>|t|)
Syllable Number (8)	−0.91483	0.72007	2076.48214	−1.27	0.204054
Syllable Number (9)	0.10948	1.45021	2075.13676	0.075	0.939832
Coda (No coda)	0.76608	0.14612	2076.42625	5.243	1.74E-07
IP Position (Final)	−2.22453	0.26372	2078.9518	−8.435	<2.00E-16

Table 13: Fixed effects summary of best linear mixed-effects model of maximum f0. Trimming removed 68 data points (3.14%).
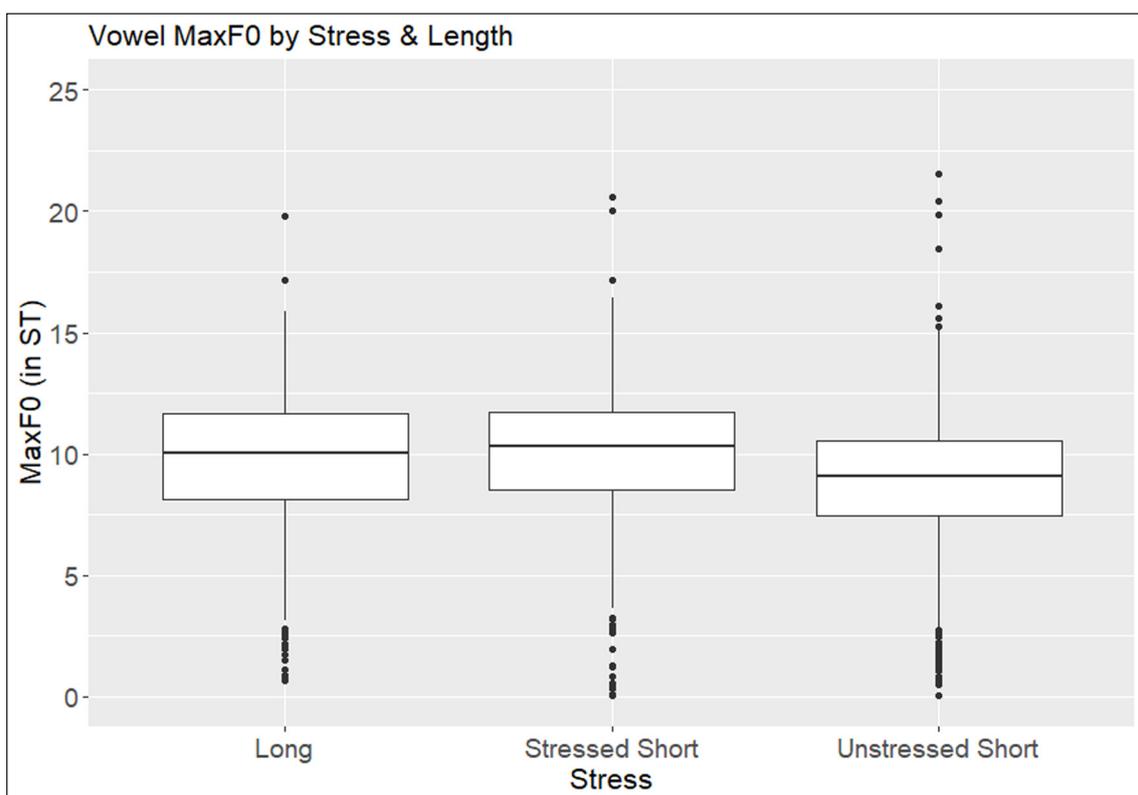
Figure 5: Boxplot of the vowel maximum f0 for each of three different stress-length combinations.

Table 13 presents the fixed effects of the best linear mixed-effect model for maximum f0. Recall that the f0 models were based on a smaller dataset than the other analyses due to the removal of any f0 measurement errors. The results given in the first three rows of **Table 13** show that while stressed short vowels were significantly higher in f0 than unstressed short vowels,

they were not significantly different from long vowels. Relevelling revealed that the max f0s of unstressed short vowels were not significantly different than those of stressed long vowels (estimate = 0.86135; std. error = 0.51264; df = 2.97738; t = 1.680; Pr(>|t|) = 0.192). This result indicates that max f0 is not associated with length, but is associated with stress on short vowels, although, as **Figure 5** shows, the max f0 ranges for stressed and unstressed short vowels overlap.

The most striking feature of the rest of **Table 13** is the general tendency for f0 to fall across the word: The later in the word a syllable comes, the lower its pitch. Compared to the initial syllable, the difference was significant for the fourth, fifth, and seventh syllable (again, the effect was probably not significant for later syllables due to their small number). Lastly, **Table 13** shows that vowels in open syllables were higher in max f0 than those in closed syllables, while IP-final vowels were lower in max f0 than non-IP-final ones.

6.4. f0 fall

In the modelling of f0 fall, including syllable number and the presence of an onset as a predictor did not improve model fit and these factors were therefore excluded from the final model, presented here in **Table 14**. As with duration and maximum f0, genre as a random intercept did not improve model fit.

f0 Fall Model	Estimate	Std. Error	Df	t value	Pr(>|t|)
Intercept (Short Vowel, Stressed)	10.4893	2.2808	207.4453	4.599	7.38E-06
Vowel Length (Long Vowel, Stressed)	4.5056	0.8888	1947.7928	5.069	4.38E-07
Vowel Length (Short Vowel, Unstressed)	–1.538	0.7492	1946.0471	–2.053	0.040221
Word length (2 syllables)	–4.0142	2.1928	1947.7834	–1.831	0.067315
Word length (3 syllables)	–5.7668	2.2194	1947.5491	–2.598	0.009437
Word length (4 syllables)	–5.4179	2.2801	1946.6136	–2.376	0.017589
Word length (5 syllables)	–5.9493	2.2885	1945.8254	–2.6	0.009404
Word length (6 syllables)	–6.9504	2.4289	1946.4903	–2.862	0.00426

(Contd.)

f0 Fall Model	Estimate	Std. Error	Df	t value	Pr(> |t|)
Word length (7 syllables)	–8.3744	2.6039	1947.1291	–3.216	0.001321
Word length (8 syllables)	–5.6231	2.7087	1942.0919	–2.076	0.038034
Word length (9 syllables)	–7.4338	3.3558	1947.7951	–2.215	0.026862
Coda (No coda)	2.3855	0.7036	1919.4521	3.39	0.000713
IP Position (Final)	12.1406	1.3416	1947.6247	9.05	2.00E-16

Table 14: Fixed effects summary of best linear mixed-effects model of f0 fall. Trimming removed 58 data points (2.95%).
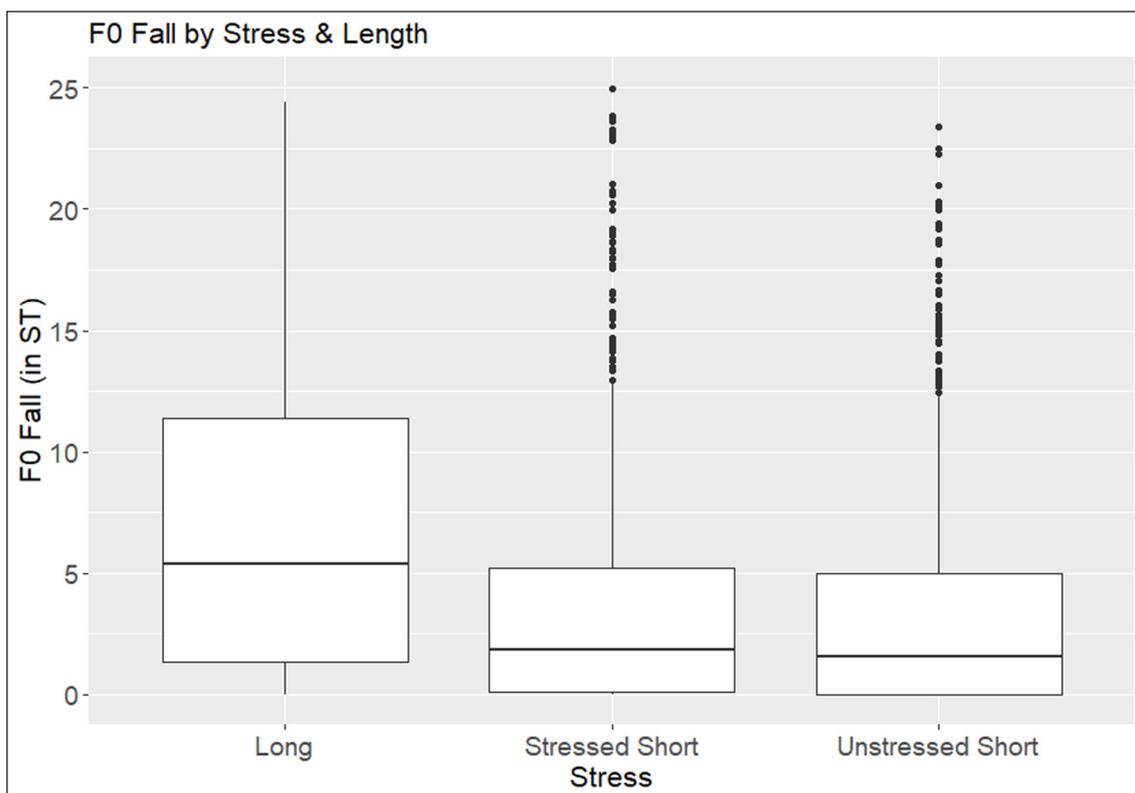
Figure 6: Boxplot of the f0 fall for each of three different stress-length combinations.

Figure 6 illustrates that stressed long vowels had the greatest f0 fall. As shown in **Table 14**, linear mixed-effects modelling indicated that their falls were significantly larger than those of stressed short vowels. Relevelling also confirmed that unstressed short vowels had smaller falls

than stressed long vowels (estimate = –4.5056; std. error = 0.89; df = 1947.7928; t = –5.069; Pr(>|t|) = 4.38e-07). The difference between stressed and unstressed short vowels, however, was only marginally significant. F0 fall was further significantly impacted by the total number of syllables in the word, such that the longer the word, the less the f0 falls. An open syllable had a larger f0 fall than a closed syllable. Finally, an IP-final syllable had a significantly larger f0 fall than a non-IP-final syllable.

Since these results suggest that f0 falls are affected by phonemic vowel length, but not by stress, it stands to reason that the size of the fall is mainly affected by the "space" available for it (i.e., by vowel duration). In order to examine the degree to which f0 fall and duration are correlated, a Pearson correlation test was run between the two variables. The result was significant (t = 17.816, df = 2017, p-value < 2.2E-16), indicating that there is indeed a positive correlation between a vowel's duration and its f0 fall (r = 0.44). This result is reflected in **Figure 7**. In this graph, values with low f0 fall are clearly associated with lower vowel durations. Together with the model in **Table 14**, these results imply that f0 fall is more of a characteristic of long vowels than it is a cue for stress, though its marginal relationship with stress may be associated with the role of duration as a stress correlate (see section 6.1).
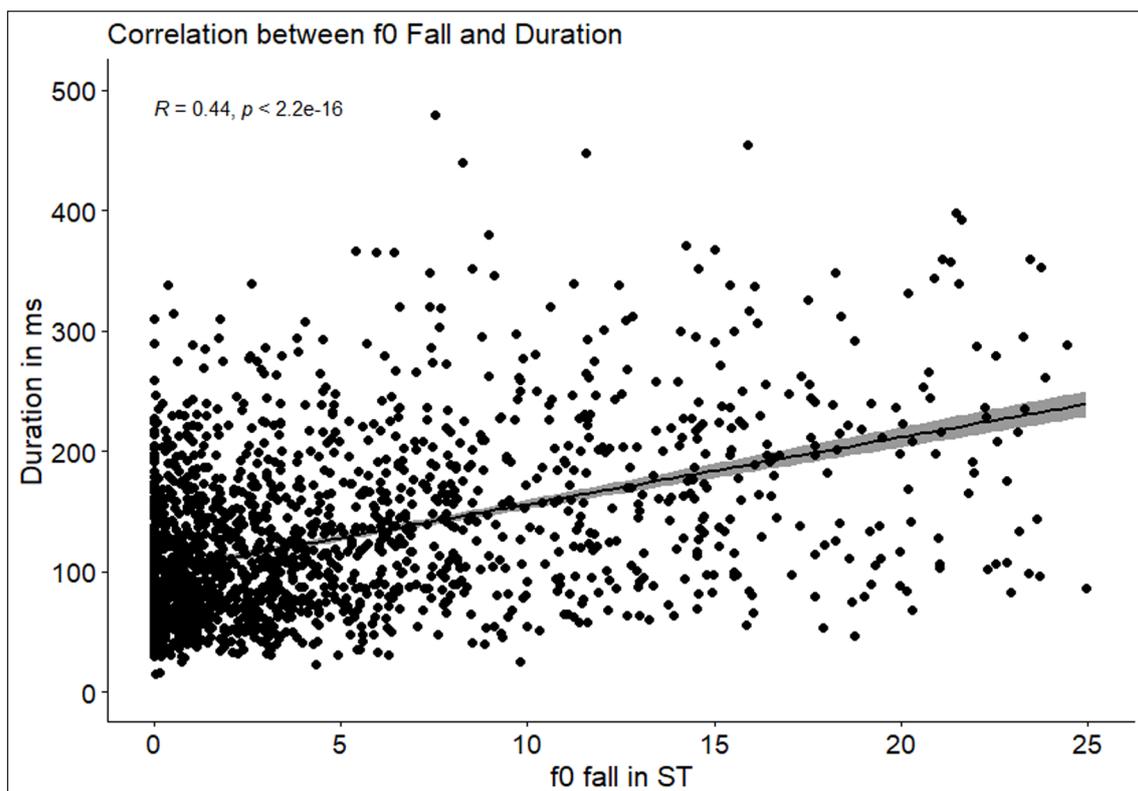
Figure 7: Scatter plot of f0 fall and duration, demonstrating the correlation.

7. Discussion

Table 15 briefly summarizes the results of the statistical models.

	Duration	**Intensity**	**Max f0**	**F0 Fall**
Stress	Significant	Significant	Significant	Marginal
Phonemic Length	Significant	Significant	Not Significant	Significant

Table 15: Summary of the effects of stress and length on each examined acoustic correlate.

In order to examine characteristics of stress alone, we can consider the differences between stressed and unstressed short vowels. Duration, intensity, and maximum f0 were significantly different between the two. This demonstrates that duration, intensity, and f0 maximum are all correlated with stress. These results affirm the hypotheses H_{DUR}, H_{INT}, and H_{f0MAX}, all of which predicted that values would be higher in stressed vowels than in unstressed vowels. F0 fall was shown to be only marginally affected by stress, leading to a rejection of H_{f0FALL}. In general, however, there was a correlation between f0 fall and duration, such that longer vowels had more time for their f0 to fall. In this way, the marginal relationship of f0 to stress may be more associated with the durational effect of stress than f0 fall behaving as a stress correlate itself.

These findings corroborate and notably expand upon observations by Miyaoka (1971) and Woodbury (1987), both of whom describe pitch movement on strong (stressed/heavy) syllables, without distinguishing between stress and length effects. The results of the present study also affirm Gabas' observations about duration, intensity, and f0 affecting stress production. However, there is no evidence that any of the acoustic correlates of stress—duration, intensity, and max f0—are ordered relative to one another, as suggested by Gabas. A perception study that isolates these correlates would serve to test Gabas' claims regarding the cue hierarchy. While our dataset suffers from a dearth of exceptionally long words, comprising 2.4% of the data, results indicate that syllables that come later in the word (syllable number as counted from the left edge) tend to be shorter and quieter. Furthermore, word length was only significant for duration and f0 fall, both of which decrease as word length increases—that is to say, syllables in longer words tend to have nuclei that are shorter and feature less f0 fall. Our results in this regard do not corroborate Gabas' claim about two categories of word length affecting prosodic performance (i.e. stress being marked differently in long vs. short words). Instead, the present study found consistent correlates of stress across the whole data set. Finally, this study limits its scope to exploring the relationship between unstressed short, stressed short, and long vowels. A follow-up study to compare stressed vowels to one another and relative to the right edge will clarify Gabas' claims regarding penultimate prominence.

In addition to stress, the present study also investigated the acoustic correlates of phonemic vowel length. Our results showed a three-way distinction in the acoustic characteristics of the stress-length categories wherein stressed long vowels were longer and louder than stressed short vowels, which were in turn longer and louder than unstressed short vowels. Each of the three categories is therefore made audibly distinct by the speaker, and critically, these distinctions were preserved even in cases of confounding phonology (e.g., iambic lengthening and effects of syllable closure). The result is a durational scale of vowel durations, illustrated in **Table 16**.

Short, Closed, Unstressed	<	Short, Open, Unstressed	<	Short, Closed, Stressed	<	Short, Open, Stressed (Iambically Lengthened)	<	Long, Closed	<	Long, Open
(C)VC		(C)V		ˈ(C)VC		ˈ(C)V·		ˈ(C)VːC		ˈ(C)Vː

Table 16: Observed scale of vowel duration by syllable closure, underlying vowel length, and stress.

Similar to the findings in Koo and Badten (1974) regarding a ternary stress-length distinction in the durations of Central Siberian Yupik vowels, the acoustic portion of this study revealed a distinction in the expression of phonemic and metrical length, such that they are audibly distinct from one another. While Koo and Badten (1974) only investigated duration, the present results revealed differences between long and short vowels also regarding other acoustic correlates, in line with what has been observed for a range of other languages such as Estonian (Lippus, 2011; Lippus, Asu, Teras, & Tuisk, 2013), Finnish (Järvikivi, Vainio, & Aalto, 2010; Vainio, Järvikivi, Aalto, & Suni, 2010), Japanese (Isei-Jaakkola, 2004; Kubozono, Takeyasu, Giriko, & Hirayama, 2011; Yoshida, de Jong, Kruschke, & Päiviö, 2015), and Sakha (Vasilyeva, Järvikivi, & Arnhold, 2016; also see Yu, 2010, on general effects of f0 on perceived duration). Acoustic measures that were significantly different between stressed short and stressed long vowels highlight the acoustic effects of phonemic length: Duration, intensity, and f0 fall, but not maximum f0, were associated with length. This result affirms the hypothesis $H_{VowelLength}$, which predicted that this phonological distinction would be observable in the acoustics. Indeed, phonemic vowel distinctions in length are preserved when stressed. Stress surfaces mostly the same way on long and short vowels and is marked by a higher duration and intensity, as well as higher f0 on stressed short vowels only. However, the function of f0 falls appears to be to differentiate (stressed) short vowels from long vowels, as long vowels have a greater f0 fall than both stressed and unstressed short vowels. In short, the phonemic length contrast and metrical lengthening phenomenon of Yup'ik are acoustically distinct.

Finally, the results also showed acoustic correlates of prosodic units both above the foot (the word and the IP) and below the foot (the syllable). The results for duration (Section 6.1) showed generally shorter vowel durations in longer words than in shorter ones and shortening of vowels in syllables with onsets or codas compared to those without (i.e., tendencies towards isochrony at both the word and the syllable level). Furthermore, the duration results also showed that IP-final syllables are significantly longer than non-IP final syllables. In terms of intensity (Section 6.2), syllables towards the end of a word are quieter than those towards the beginning. Lastly, f0 maxima for later syllables tend to be lower than those for early syllables (Section 6.3), and longer words tend to exhibit overall more f0 fall than shorter words (Section 6.4). Although this study did not explicitly seek to corroborate claims regarding prosodic units above the foot, all of these observations are in line with trends towards isochrony and finality effects that are common cross-linguistically (see overview in Fletcher, 2010; also see Arnhold, accepted, for an overview of phonetic and phonological marking of prosodic domains and adjustments to syllable structure in Inuit and Yupik languages).

One particularly interesting theoretical question that arises out of the acoustic results addresses the relationship between the mora and its measurable value as a unit of phonemic length (also see Broselow, Chen, & Huffman, 1997; Cohn, 2003; Duanmu, 1994; Gordon, 2002, 2004, 2007; Ham, 2013; Khattab & Al-Tamimi, 2014). Our study suggests that Yup'ik may not simply have binary system of syllable weight, but instead is reminiscent of what Gordon (2002) calls scalar syllable weight systems. The acoustic results show a large number of distinctions in vowel durations, as summarized in **Table 16**, which cannot straightforwardly be explained in terms of moraic constituency as laid out by Hayes (1995). Recall that, according to Hayes, Yup'ik unstressed open syllables with short vowels are monomoraic, while closed syllables (except in double clash environments), stressed open syllables with short vowels, and syllables containing long vowels are bimoraic. A particular challenge to this account is the fact that stressed short vowels had shorter durations than long vowels, both of which are assumed to be bimoraic in Hayes' (1995) conception of Yup'ik syllable weight. That the two are distinctively produced points to the extent to which Yup'ik prioritizes maintaining the ternary stress-length distinction. However, if the addition of a mora is the vehicle of iambic lengthening, then it is surprising that short vowels that are lengthened in this way are not as long as vowels with underlying length.

In a scalar system, as found, for example, in Klamath (Barker, 1964), Kashmiri (Kenstowicz, 1994), Chickasaw (Munro & Willmond, 1994), and Yapese (Jensen, Pugram, Iou, & Defeg, 2019), syllables have intermediary weight categories, more than a simple light-heavy dichotomy. Scalar weight systems are still fundamentally binary: Gordon interprets the scale of weight between, for instance, CVV, CVC, and CV syllables as a series of binaries, such that {CVV > CVC, CV} and

{CVV, CVC > CV}. This scalar framing mirrors Yup'ik vowel distinctions, in which long syllables are longer than both stressed short and unstressed short syllables, and all stressed syllables are longer than unstressed short syllables.

There are several analyses that may explain the scalar nature of Yup'ik vowel durations. First, moras may contribute different amounts of length in different positions. Second, the observed length distinctions may be exclusively stress effects compounding on underlying length distinctions. In this view, iambic lengthening does not apply as described in the literature. Third, if stress always adds a mora to affected syllables, in closed syllables, that mora may be shared between the nucleus and the coda, resulting in stressed vowels in closed syllables appearing shorter, but having equal weight, to vowels in open syllables. Fourth, there may be a disparate number of moras between stressed short and long vowels, such as in Nilotic languages, while further length differences are non-moraic. Finally, iambically lengthened (short stressed) vowels and long vowels may have the same number of moras, namely two, per Hayes (1995). Here, the observed vowel length disparities must be either non-moraic or the result of incomplete neutralization.

The first possible analysis would assert that the mora that is added to a short vowel when it is stressed contributes less duration than either mora of a long vowel. This implies a quality distinction among moras: Though all moras contribute to weight, such that the presence of two moras constitutes a heavy syllable, some contribute more acoustically than others, resulting in some heavy syllables surfacing as longer than others. In this case, moras assigned to underlyingly long vowels would contribute more to duration than moras assigned by the metrical component. While theoretically interesting, there is little evidence for this line of argumentation, which would need to be backed up by strong cross-linguistic data due to the far-reaching implications of assuming non-uniformity among moras. Moreover, even if such an analysis were chosen to account for the difference between stressed short and long vowels, it would either need to acknowledge that there are additional durational differences that cannot be accounted for by moras (cf. **Table 16**) or would need to be expanded into a system where moras were able to have not only two different values or degrees of contributing to duration, but several.

The second possible analysis reconsiders the assumption that the observed lengthening of stressed short vowels is iambic lengthening as it is described in the literature. In this view, rather than iambic lengthening resulting from an added mora, the trigger for the observed lengthening may be the presence of stress—that is, it may not be the case that stress requires that a light syllable be made heavy via iambic lengthening, but rather that the durational effects of stress itself, compounded onto the underlying length of a short or long vowel, result in the observed lengthening.

There is some evidence against this option, however. If the observed lengthening on stressed short vowels in open syllables was just a stress effect, without any durational contributions from iambic lengthening, then, presumably, this single effect would apply evenly to stressed

short vowels regardless of syllable closure. We would not expect the significant difference between stressed short vowels in open and closed syllables that was observed in the acoustic measurements of duration (cf. **Table 10**). While one could explain this difference as an effect of syllable isochrony (vowels are shorter in syllables with codas than in those without, as seen for all vowel categories), it is noteworthy that the difference between vowels in open and closed syllables was much larger for stressed short vowels (estimated by pairwise comparisons as 64 ms) than for both unstressed short vowels and stressed long vowels (40 ms and 26 ms, respectively, cf. **Table 10**). This strongly suggests that iambic lengthening is in fact distinct from regular stress lengthening, and that only stressed short vowels in open syllables are affected by both. Since iambic lengthening seems to have a distinct effect on vowel duration, the addition of a mora can be assumed to explain why iambic lengthening occurs in the first place: To ensure that all foot heads are heavy.

Thirdly, it may be the case that the codas of stressed closed syllables with short nuclei share the burden of stress lengthening with the vowel. This would allow stress to lengthen stressed open and closed syllables equally, but in closed syllables, some of that lengthening would be realized on the coda, and their nuclei would appear shorter. This would be similar to Levantine Arabic, where, outside of the word-final position, codas dominate moras when the preceding vowel is long. In word-final position, however, the mora is shared between the nucleus and the coda (Broselow et al., 1997). Broselow et al. show that, in Levantine Arabic, because a non-final coda shares its mora with the preceding (long) nucleus, only long vowels are shortened in closed syllables. Short vowels have the same duration in both open and closed syllables, as they do not share moras with codas. This, however, is not the case in Yup'ik, where both long and short vowels in open syllables are longer than those in closed syllables, as indicated by the results of the present study (cf. Section 6.1). Though the gap between open and closed syllables is much larger for short stressed vowels than it is for long vowels, if the discrepancy was due to mora sharing alone, there would be no difference between long open and long closed syllables. Thus, while mora sharing could contribute to the observed special behavior of stressed short vowels, it cannot explain it completely.

The fourth analysis posits that the difference is seen between vowels with ostensibly the same number of moras because that they do not, in fact, have the same number of moras. That is, stressed short/iambically lengthened vowels have more moras than unstressed short vowels, but fewer than long vowels. In this view, durational differences are explained as a combination of moraic and non-moraic factors. **Table 17** visualizes such an analysis, following the assumptions that a) iambic lengthening is caused by the addition of a mora to short vowels targeted for stress in open syllables (following Hayes, 1995), b) codas contribute to weight outside of clash environments (again following Hayes, 1995, and supported by the fact that these closed syllables receive stress, but extending Hayes' assumption that moras contribute

weight also to syllables with long vowels), and c) the consistent durational distinctions between the three stress-length categories as observed in the acoustic measurements can be attributed to different mora counts. The resulting system would posit that Yup'ik syllables can have between one and four moras. This is notably more complex than Hayes's account, but is not unheard of in the language family (Arnhold, accepted; Holtved, 1964; Jacobsen, 2000; Kleinschmidt, 1851, pp. 7–8; Nagano-Madsen, 1990). There is also typological precedence for a phonetic long-overlong vowel distinction in languages that permit trimoraic vowels. In Nilotic languages, such as Dinka and Nuer, for instance, the ternary vowel length distinction between short, long, and overlong vowels is phonemically contrastive, and expressed via three degrees of duration (see Monich, 2017; Remijsen & Gilley, 2008).

Syllable	Mora Count
Short, Open, Unstressed	1 mora (from the nucleus)
Short, Closed, Unstressed	1 mora (1 from the nucleus, 1 from the coda; however, the second mora is removed due to double clash, resulting in an unstressed syllable)
Short, Open, Stressed	2 moras (1 from the nucleus underlyingly, 1 from iambic lengthening)
Short, Closed, Stressed	2 moras (1 from the nucleus, 1 from the coda)
Long, Open, Stressed	3 moras (3 from the nucleus, in order to differentiate it from stressed short vowels)
Long, Closed, Stressed	4 moras (3 from the nucleus, 1 from the coda)

Table 17: Possible distribution of moras accounting for observed durational effects of phonemic length, iambic lengthening, and the contribution of codas to syllable weight, but not for effects of syllable closure on vowel duration.

While this analysis does account for the difference between long and short stressed vowels, it leads to another discrepancy: If **Table 17** is a true account of moraic constituency in Yup'ik and moras are directly reflected in vowel duration, then we expect unstressed vowels and short stressed vowels in closed syllables to have roughly the same duration, as they have the same number of moras, namely one (since the second mora of stressed CVC syllables comes from the coda). However, as seen in **Table 10** and **Table 11**, this is not the case. In other words, this hypothesis cannot account the observed vowel duration differences beyond iambic lengthening and phonemic length, nor does it account for the observed effects of codas on vowel durations (i.e., isochronic lengthening/shortening). Thus, even expanding the inventory of possible mora counts to syllables with up to four moras does not allow for an exhaustive mapping between

moras and duration effects (i.e., there are durational effects that cannot be accounted for with a strictly moraic analysis).

The fifth and final analysis of the durational discrepancy between stressed short and long vowels strictly follows the description of moraic distribution in Hayes (1995): Syllables in Yup'ik are maximally bimoraic, with all heads necessarily containing two moras. This account justifies iambic lengthening as moraic, giving light, monomoraic syllables an extra mora in order to function as a foot head. However, it also implies that all bimoraic syllables are made equal, for instance, that iambically-lengthened short vowels are equal in length to long vowels, which is not supported by the evidence in this paper.

There are two ways to maintain this analysis. Firstly, we could assume that moras are more or less completely divorced from acoustic duration. The only correspondence between the two would be that monomoraic short unstressed vowels have shorter durations than short stressed and long vowels, both of which are bimoraic, at least in open syllables. All other observed durational differences summarized in **Table 16** would not be reflected in the moraic distribution. Moras would thus be central to the assignment of stress, but untethered from the rest of the phonological system and its phonetic expression. Such a strong dissociation between duration and moras is probably not theoretically desirable. Furthermore, the acoustic results clearly distinguish three vowel categories, and demonstrate that Yup'ik prioritizes preserving length distinctions wherever possible. This suggests that these vowel categories play a central role in the phonology of the language, as in other Inuit and Yupik languages (cf. overview in Arnhold, accepted), and that vowel length distinctions should be captured in moras.

The second alternative would be that, while long vowels and stressed short vowels are both bimoraic—long vowels by nature of their underlying length, and stressed short vowels by way of iambic lengthening—the discrepancy between the two is the result of incomplete neutralization. This echoes monomoraic lengthening in Japanese (Braver, 2019). Japanese has a bimoraic minimality requirement: Each prosodic word must contain one foot, and each foot must contain two moras (McCarthy & Prince, 1986, 1993). In order to fulfill this requirement, monomoraic nouns with short vowels are lengthened. This is parallel to Yup'ik's iambic lengthening fulfilling the requirement that foot heads be heavy. While in Japanese, the mora counts of lengthened monomoraic vowels and bimoraic vowels are ostensibly identical, Braver and Kawahara (2016) found that the vowel durations of lengthened nouns were shorter than for underlyingly long nouns. Lengthened vowels were produced with an intermediate duration, between unlengthened short vowels and long vowels. This very strongly resembles the observed stress-length distinction in Yup'ik.

Braver (2019) interprets the results of Braver and Kawahara (2016) to mean that monomoraic lengthening is attempting to neutralize the length distinction between monomoraic and bimoraic vowels. However, this neutralization is incomplete, resulting in the intermediary duration of

lengthened monomoraic vowels. In Braver's Optimality Theoretical account, constraints that govern the surface vowel's faithfulness to its phonemic length (underlyingly short vowels surfacing as short) compete with those that require that monomoraic short vowels be lengthened. The consequence is a trichotomy of length which distinguishes the durations of unlengthened short, lengthened short, and long vowels.

The connections between Japanese monomoraic lengthening and Yup'ik iambic lengthening are quite clear. Both use the addition of a mora to induce lengthening on an otherwise short vowel (which is underlyingly monomoraic), so that it can fill a bimoraicity requirement (binary feet in Japanese, heavy iambic heads in Yup'ik). However, this lengthening competes with the language's desire to maintain its phonemic length distinction. As a result, the moraic lengthening process does not fully neutralize the phonetic differences, even if the targeted vowel is now treated as heavy.

In sum, considering typological parallels, the most plausible analyses for the Yup'ik findings presented here are the possibility of polymoraic syllables, as has been reported elsewhere in the language family (option four above), or that iambic lengthening is an incomplete neutralization process (option five above). In the latter case, in closed syllables, the added mora may be shared between the nucleus and coda (as mentioned under option three above). In open syllables, meanwhile the vowels' desire to remain short conflicts with their newfound weight, resulting in a phonetic distinction between lengthened short vowels and long vowels.

8. Conclusion

In this study, six recordings of spoken Central Alaskan Yup'ik were analyzed in order to examine the acoustic correlates of stress in Yup'ik syllabic nuclei. The Yup'ik stress system can be summarized as a left-to-right iambic system wherein codas contribute to weight and stress is sensitive to weight. The results of this study demonstrate that in Central Alaskan Yup'ik, a vowel's stress affects that vowel's duration, intensity, and f0. The significant effects observed for f0 fall as a stress cue for long vowels are likely more related to length than they are to stress, while a significant effect of stress on the height of the f0 maximum appeared only for short stressed vowels. The results further corroborate the foundational tenet of Yup'ik metrics: that there is a ternary relationship between stress and length, and that stressed long vowels, stressed short vowels, and unstressed short vowels are all different. This project establishes baseline patterns of stress distribution, behavior, and expression that address the outstanding questions surrounding stress in Central Alaskan Yup'ik and open the door for further investigations in the future.

Competing interests

The authors have no competing interests to declare.

Author contributions

This project arose out of research done as a requirement for McKinley Alden's (henceforth MA) Ph.D. in Linguistics at the University of Alberta, supervised by Dr. Anja Arnhold (henceforth AA). MA and AA co-designed the study. MA selected the data from the Alaskan Native Language Archive Central Alaskan Yup'ik collections, annotated the data, and performed statistical and phonological analyses under the supervision of AA. MA wrote the first draft of the article, and both authors revised the article and have approved the submitted version.

References

Afcan, P., & Hofseth, E. (1972). *Napam cuyaa*. Eskimo Language Workshop, University of Alaska.

Alaskan Native Language Center. (n.d.). Retrieved January 13, 2022, from https://www.uaf.edu/anlc/

Alden, M., & Arnhold, A. (2022). *User's guide to Central Alaskan Yup'ik stress derivation [Report]*. https://doi.org/10.7939/r3-4dd8-0b37

Árnason, K. (1985). Icelandic word stress and metrical phonology. *Studia Linguistica, 39*(2), 93–129. DOI: https://doi.org/10.1111/j.1467-9582.1985.tb00747.x

Arnhold, A. (2018). *Duration, intensity, F0 measurement script*. Https://Sites.Ualberta.ca/~arnhold/Scripts/OnlyMeasurements/MeasureIntensityDurationF0minF0maxF0contourpoints.Praat.

Arnhold, A. (n.d.). Phonology and phonetics of Inuit, Yupik and Unangan languages. In A. Berge, A. Arnhold, & N. B. Trondhjem (Eds.), *The Inuit, Yupik and Unangan languages*. Oxford University Press.

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. DOI: https://doi.org/10.1016/j.jml.2007.12.005

Bakovic, E. (1996). Foot harmony and quantitative adjustments [Qualifying paper/unpublished manuscript, Rutgers University]. In *Rutgers Optimality Archive*. DOI: https://doi.org/10.7282/T33B61TX

Barker, M. A.-R. (1964). Klamath grammar. In *University of California Publications in Linguistics*. University of California Press.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015a). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. DOI: https://doi.org/10.18637/jss.v067.i01

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015b). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1–48. DOI: https://doi.org/10.18637/jss.v067.i01

Blue, A. (2007). *Cungauyaraam qulirai: Annie Blue's stories*. Alaskan Native Language Center.

Boersma, P., & Weenink, D. (2020). Praat: doing phonetics by computer. In *retrieved 23 March 2020* (6.1.10).

Braver, A. (2019). Modelling incomplete neutralisation with weighted phonetic constraints. *Phonology*, *36*(1), 1–36. DOI: https://doi.org/10.1017/S0952675719000022

Broselow, E., Chen, S. I., & Huffman, M. (1997). Syllable weight: convergence of phonology and phonetics. *Phonology*, *14*(1), 47–82. DOI: https://doi.org/10.1017/S095267579700331X

Buckley, E. (2019). Stress, tone, and pitch accent. In D. Siddiqi, M. Barrie, C. Gillon, J. Haugen, & E. Mathieu (Eds.), *The Routledge Handbook of North American Languages* (1st edition, pp. 66–89). Routledge. DOI: https://doi.org/10.4324/9781315210636-3

Chung, S. (1983). Transderivational relationships in Chamorro phonology. *Language*, *59*(1), 35–66. DOI: https://doi.org/10.2307/414060

Cohn, A. C. (2003). Phonological structure and phonetic duration: the role of the mora. *Working Papers of the Cornell Phonetics Library*, *15*, 69–100.

Derbyshire, D. (1985). *Hixkaryana and linguistic typology*. Summer Institute of Linguistics; University of Texas at Arlington.

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

Duanmu, S. (1994). Syllabic weight and syllabic duration: a correlation between phonology and phonetics. *Phonology*, *11*(1), 1–24. DOI: https://doi.org/10.1017/S0952675700001822

Fletcher, J. (2010). The prosody of speech: timing and rhythm. In W. Hardcastle, J. Laver, & F. Gibbon (Eds.), *The Handbook of Phonetic Sciences* (2nd ed., pp. 521–602). Blackwell Publishing Ltd. DOI: https://doi.org/10.1002/9781444317251.ch15

Gabas, N. (1996). Phonetic correlates of stress in Yup'ik. *Prosody, Grammar, and Discourse in Central Alaskan Yup'ik, Santa Barbara Papers in Linguistics, Vol. 7*, 17–36.

Gordon, M. (2002). A phonetically driven account of syllable weight. *Language*, *78*(1), 51–80. DOI: https://doi.org/10.1353/lan.2002.0020

Gordon, M. (2004). Syllable weight BT – Phonetically based phonology. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically based phonology* (pp. 277–312). Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511486401.009

Gordon, M. (2007). Syllable weight: Phonetics, phonology, typology. In *Syllable Weight: Phonetics, Phonology, Typology* (1st ed.). Routledge. DOI: https://doi.org/10.4324/9780203944028

Gordon, M., & Munro, P. (2007). A phonetic study of final vowel lengthening in Chickasaw. *International Journal of American Linguistics*, *73*(3), 293–330. DOI: https://doi.org/10.1086/521729

Gordon, M., & Roettger, T. (2017). Acoustic correlates of word stress: a cross-linguistic survey. *Linguistics Vanguard*, *3*(1), 20170007. DOI: https://doi.org/10.1515/lingvan-2017-0007

Halle, M. (1990). Toward a metrical interpretation of Yupik prosody. *Alaska Native Language Center Research Papers*, *7*, 159–173. https://www.jstor.org/stable/4047697

Ham, W. (2013). Phonetic and phonological aspects of geminate timing [Doctoral dissertation, Cornell University]. In *Phonetic and Phonological Aspects of Geminate Timing.* DOI: https://doi.org/10.4324/9781315023755

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

Heinrich, A. C. (1979). Yup'ik Eskimo Grammar; Irene Reed, Osahito Miyaoka, Steven Jacobson, Paschal Afcan, Michael Krauss. *ARCTIC, 32*(1). DOI: https://doi.org/10.14430/arctic2757

Holtved, E. (1964). *Kleinschmidts Briefe an Theodor Bourquin, Meddelelser on Gronland*. Reitzel.

Hyde, B. (2011). The iambic-trochaic law. In *The Blackwell Companion to Phonology*. DOI: https://doi.org/10.1002/9781444335262.wbctp0044

Isei-Jaakkola, T. (2004). *Lexical quantity in Japanese and Finnish* [Doctoral dissertation]. University of Helsinki.

Jacobsen, B. (2000). The question of "stress" in West Greenlandic: an acoustic investigation of rhythmicization, intonation, and syllable weight. *Phonetica, 57*, 40–67. DOI: https://doi.org/10.1159/000028458

Jacobson, S. (1984). The stress conspiracy and stress-repelling bases in the Central Yup'ik and Siberian Yupik Eskimo languages. *International Journal of American Linguistics, 50*(3), 312–324. https://www.jstor.org/stable/1265552. DOI: https://doi.org/10.1086/465838

Jacobson, S. (1985). Siberian Yupik and Central Yupik prosody. In *Yupik Eskimo prosodic systems: descriptive and comparative studies* (pp. 25–45).

Jacobson, S. (1990). Comparison of Central Alaskan Yup'ik Eskimo and Central Siberian Yupik Eskimo. *International Journal of American Linguistics, 56*(2), 264–286. DOI: https://doi.org/10.1086/466153

Jacobson, S., & Jacobson, A. (1995). *A practical grammar of the Central Alaskan Yup'ik Eskimo language: with Yup'ik readings*. Alaska Native Language Center.

Järvikivi, J., Vainio, M., & Aalto, D. (2010). Real-time correlates of phonological quantity reveal unity of tonal and non-tonal languages. *PLoS ONE, 5*(9), e12603. DOI: https://doi.org/10.1371/journal.pone.0012603

Jensen, J. T., Pugram, L. D., Iou, J. B., & Defeg, R. (2019). Yapese reference grammar. In *Yapese Reference Grammar*. University of Hawai'i Press. DOI: https://doi.org/10.2307/j.ctv9zcks7

Kenstowicz, M. (1994). *Phonology in generative grammar* (1st ed.). Blackwell: Oxford.

Khattab, G., & Al-Tamimi, J. (2014). Geminate timing in Lebanese Arabic: the relationship between phonetic timing and phonological structure. *Laboratory Phonology, 5*(2), 231–269. DOI: https://doi.org/10.1515/lp-2014-0009

Kleinschmidt, S. (1851). *Grammatik der grönländischen sprache, mit theilweisem einschluss des Labradordialects [grammar of the Greenlandic language under partial inclusion of the Labrador dialect]*. G. Reimer. DOI: https://doi.org/10.1515/9783111698830

Koo, J. H., & Badten, L. (1974). Acoustic measurements of the "fake" vowel length and degrees of vowel length in St. Lawrence Island Eskimo. *Phonetica, 40*(4), 213–220. DOI: https://doi.org/10.1159/000259490

Kubozono, H., Takeyasu, H., Giriko, M., & Hirayama, M. (2011). Pitch cues to the perception of consonant length in Japanese. In W. S. Lee & E. Zee (Eds.), *17th International Congress of Phonetic Sciences (ICPhS XVII)* (pp. 1150–1153). City University of Hong Kong.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. DOI: https://doi.org/10.18637/jss.v082.i13

Leer, J. (1985a). Evolution of prosody in the Yupik languages. In M. Krauss (Ed.), *Yupik Eskimo prosodic systems: descriptive and comparative studies* (pp. 135–158). Alaskan Native Language Center.

Leer, J. (1985b). Prosody in Alutiiq. In M. Krauss (Ed.), *Yupik Eskimo prosodic systems: descriptive and comparative studies* (pp. 77–134). Alaskan Native Language Center.

Leer, J. (1985c). Toward a metrical interpretation of Yupik prosody. In M. Krauss (Ed.), *Yupik Eskimo Prosodic Systems: Descriptive and Comparative Studies* (pp. 159–173). Alaska Native Language Center.

Leer, J. (1994). The phonology of the Kenai Peninsula dialect of Chugach Alutiiq. *Languages of the North Pacific Rim, Hokkaido University Publications in Linguistics, 7*, 45–148.

Lenth, R. (2022). Package 'emmeans.' *CRAN*.

Lippus, P. (2011). *The acoustic features and perception of the Estonian quantity system* [Doctoral dissertation]. University of Tartu.

Lippus, P., Asu, E. L., Teras, P., & Tuisk, T. (2013). Quantity-related variation of duration, pitch and vowel quality in spontaneous Estonian. *Journal of Phonetics, 41*(1), 17–28. DOI: https://doi.org/10.1016/j.wocn.2012.09.005

Lipscomb, D. R. (1992). Differences and similarities in the prosodic systems of western and eastern Eskimo. *Acta Linguistica Hafniensia, 25*(1), 64–81. DOI: https://doi.org/10.1080/03740463.1992.10412278

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. DOI: https://doi.org/10.1016/j.jml.2017.01.001

McCarthy, J., & Prince, A. (1986). Prosodic morphology. *Linguistics Department Faculty Publication Series, 13*.

McCarthy, J., & Prince, A. (1993). Prosodic morphology: constraint interaction and satisfaction. *Linguistics Department Faculty Publication Series, 14*. https://scholarworks.umass.edu/linguist_faculty_pubs/14

Michelson, K. (1988). *A comparative study of Lake-Iroquoian accent.* Springer Dordrecht. DOI: https://doi.org/10.1007/978-94-009-2709-4

Mithun, M., & Basri, H. (1986). The Phonology of Selayarese. *Oceanic Linguistics, 25*(1), 210–254. DOI: https://doi.org/10.2307/3623212

Miyaoka, O. (1970). Vowel lengthening in Western Eskimo (Yuk). *Hoppo Bunka Kenkyu (Bulletin of the Institute for the Study of North Eurasian Cultures), 4*, 157–168.

Miyaoka, O. (1971). On syllable modification and quantity in Yuk phonology. *International Journal of American Linguistics, 37*(4), 219–226. https://www.jstor.org/stable/1264513. DOI: https://doi.org/10.1086/465169

Miyaoka, O. (1985). Accentuation in Central Alaskan Yupik. In M. Krauss (Ed.), *Yupik Eskimo prosodic systems: descriptive and comparative studies* (pp. 51–76). Alaskan Native Language Center.

Miyaoka, O. (2012). *A grammar of Central Alaskan Yupik*. De Gruyter Mouton. DOI: https://doi.org/10.1515/9783110278576

Molineaux, B. (2018). Pertinacity and change in Mapudungun stress assignment. *International Journal of American Linguistics*, *84*(4), 513–558. DOI: https://doi.org/10.1086/698855

Monich, I. V. (2017). Vowel length in Nuer. In K. Jesney, C. O'Hara, C. Smith, & R. Walker (Eds.), *Proceedings of the Annual Meetings on Phonology* (Vol. 4). DOI: https://doi.org/10.3765/amp.v4i0.4006

Munro, P., & Willmond, C. (1994). Chickasaw: an analytical dictionary. In *Language* (Issue 2). University of Oklahoma Press. DOI: https://doi.org/10.2307/416681

Nagano-Madsen, Y. (1990). Quantity manifestation and mora in West Greenlandic Eskimo: Preliminary analysis. *Lund University Department of Linguistics Working Papers*, *36*, 123–132.

Nicklas, T. (1975). Choctaw morphophonemics. In J. M. Crawford (Ed.), *Studies in Southeastern Indian languages* (pp. 237–250). University of Georgia Press.

Nicklas, T. D. (1974). *The elements of Choctaw* [Doctoral dissertation]. University of Michigan.

Pierrehumbert, J. (1980). The phonology and phonetics of English intonation [Doctoral dissertation, MIT]. In *RELC Journal*. DOI: https://doi.org/10.1177/003368828401500113

R Core Team. (2014). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*. URL Http://Www.R-Project.Org/.

Reed, I. (1977). *Yupik Eskimo grammar*. Alaskan Native Language Center.

Remijsen, B., & Gilley, L. (2008). Why are three-level vowel length systems rare? Insights from Dinka (Luanyjang dialect). *Journal of Phonetics*, *36*(2), 318–344. DOI: https://doi.org/10.1016/j.wocn.2007.09.002

Rose, Y., Pigott, P., & Wharram, D. (2012). Schneider's Law revisited: the syllable-level remnant of an older metrical rule. *McGill Working Papers in Linguistics*, *22*, 1–12.

Stacy, E. (2004). *Phonological aspects of Blackfoot prominence* [Master's thesis]. University of Calgary.

Topping, D. M. (1973). Chamorro reference grammar. In *Chamorro Reference Grammar*. University of Hawai'i. DOI: https://doi.org/10.1515/9780824841263

Turk, A., Nakai, S., & Sugahara, M. (2012). Acoustic segment durations in prosodic research: a practical guide. In S. Sudhoff, D. Lenertova, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter, & J. Schleisser (Eds.), *Methods in Empirical Prosody Research* (pp. 1–28). De Gruyter Mouton. DOI: https://doi.org/10.1515/9783110914641.1

Vainio, M., Järvikivi, J., Aalto, D., & Suni, A. (2010). Phonetic tone signals phonological quantity and word structure. *The Journal of the Acoustical Society of America*, *128*(3), 1313–1321. DOI: https://doi.org/10.1121/1.3467767

van de Vijver, R. (1998). *The iambic issue: iambs as a result of constraint interaction*. Holland Institute of Generative Linguistics.

Vasilyeva, L., Järvikivi, J., & Arnhold, A. (2016). Phonetic correlates of phonological quantity of Yakut read and spontaneous speech. *The Journal of the Acoustical Society of America*, *139*(5), 2541–2550. DOI: https://doi.org/10.1121/1.4948448

Woodbury, A. (1985). The functions of rhetorical structure: a study of Central Alaskan Yupik Eskimo discourse. *Language in Society*, *14*(2), 153–190. https://www.jstor.org/stable/4167628. DOI: https://doi.org/10.1017/S0047404500011118

Woodbury, A. (1987). Meaningful phonological processes: a consideration of Central Alaskan Yupik Eskimo prosody. *Language*, *63*(4), 685–740. https://www.jstor.org/stable/415716. DOI: https://doi.org/10.2307/415716

Woodbury, A. (1995). *The postlexical prosody of Central Alaskan Yup'ik [Report]*. https://www.uaf.edu/anla/record.php?identifier=CY978W1995

Yoshida, K., de Jong, K. J., Kruschke, J. K., & Päiviö, P. M. (2015). Cross-language similarity and difference in quantity categorization of Finnish and Japanese. *Journal of Phonetics*, *50*, 81–98. DOI: https://doi.org/10.1016/j.wocn.2014.12.006

Yu, A. C. L. (2010). Tonal effects on perceived vowel duration. *Laboratory Phonology*, *10*, 151–168. DOI: https://doi.org/10.1515/9783110224917.2.151