Acoustic Correlates of Word Stress in Haisla

Grace Baleno
Canada Institute of Linguistics

Jonathan Janzen
Nicola Valley Institute of Technology

Brendon Yoder
Canada Institute of Linguistics and SIL International[1]

Abstract: Haisla typically assigns stress like other North Wakashan languages to the leftmost stress-bearing (moraic) syllable, or applies the default-to-opposite rule if the word contains no moraic syllable (Elfner 2008; Janzen 2015, 2023). Perceived exceptions exist in words containing long vowels and suffixes with lexical stress. Previous work on the Haisla stress systems either did not consider phonemic vowel length (Vink 1977; Lincoln & Rath 1986) or simply acknowledged that it may be a factor (Wilson 1987; Bach 1990). After finding that all long (bimoraic) vowels do attract stress, we conduct the first detailed phonetic study of the Haisla stress system. We investigate prosodic features pitch, duration, and intensity, as well as the first (F_1), second (F_2), and third (F_3) formants, of predictably stressed vowels in Haisla. We find that pitch, duration, and intensity are very highly significant ($p < 0.001$) correlates of stress and, interestingly, that the backness (F_2) of stressed /a/ is highly significant ($p < 0.01$). These results contribute to the understanding of the predictability of word stress in Haisla as well as additional considerations around orthographic revision for the Haisla speech community.

Keywords: Haisla, Wakashan, acoustic phonetics, duration, pitch, intensity

1 Introduction

Word stress is largely predictable in Haisla. Stress assignment follows the general tendency of North Wakashan languages to assign stress to the leftmost stress-bearing (moraic) syllable. Complicating this are the existence of two disruptive factors in word stress assignment: suffixes with lexical high pitch, and long vowels. In Haisla, certain suffixes always attract stress, regardless of the makeup of the rest of the word. In the absence of these suffixes, long vowels take stress, regardless of their syllable position. If a word does not have a suffix with a lexical high pitch, or any long vowels, then stress will fall on the leftmost moraic syllable. The general North Wakashan default-to-opposite rule also applies to Haisla, whereby if no syllable in the word is moraic, the final syllable is stressed (Elfner 2008; Janzen 2015, 2023).

The following four ordered generalizations summarize stress assignment in Haisla. The stressed syllable is marked with an acute accent and is in bold:

(1) Assign stress to a lexically stressed syllable.

*p̓ełdaudi***lás**	‘laundry room’ from -*ás* ‘place’
*qelqezu***wá***yu*	‘butter’ from -*áyu* ‘implement’
*λiku***má**	‘lend’ from -*má* ‘causative’

---

[1] Authors are listed alphabetically by last name.

Papers for the International Conference on Salish and Neighbouring Languages 59.
D. K. E. Reisinger, Laura Griffin, Ella Hannon, Gloria Mellesmoen, Sander Nederveen, Bruce Oliver, Julia Schillo, Lauren Schneider, Bailey Trotter (eds.). Vancouver, BC: UBCWPL, 2024.

(2) Assign stress to the leftmost heavy (bimoraic) syllable.

*p̓eł**dáu**dac̓i*	‘washing machine’
***x̄áa**x̄aax̄awais*	‘skin and bones’
*kikuxʷ**t̓áam***	‘feeling hot’

(3) If no syllable is bimoraic, assign stress to the leftmost moraic syllable.

***p̓él**ca*	‘moss-covered’
***c̓íks**du*	‘dirty’
***λ̓ú**x̄ʷilac̓i*	‘freezer’

(4) If no syllable contains a mora, assign stress to the rightmost syllable.

*p̓e**lép***	‘sister-in-law’
*le**ḡéxʷ***	‘dried berries’
*λe**xéx̄s***	‘crosspiece in a canoe’

Earlier scholars, such as Vink (1977), Lincoln and Rath (1986), and Wilson (1987) did not formally consider phonemic vowel length as an influence on the stress system. Instead, they analyzed vowel duration as a primary indicator of word stress. Wilson, however, correctly predicts the lexically-stressed suffixes as a disruptive element, stating, “Haisla stress/pitch, while largely predictable by the standard rules, is not nearly so regular as Nootkan [Nuu-chah-nulth] or Kwak’wala stress. Most of the irregularities appear to be due to paradigmatic leveling with commonly used suffixes” (1987: 326). Bach is the only previous scholar who takes note of vowel length as a factor in Haisla stress: “We know that the distinction between light and heavy syllables plays a role in the stress system of Kwak’wala and probably in Haisla as well” (1990:11). He does not, however, explore vowel length fully, nor come to any conclusions. The light and heavy syllables Bach refers to are non-stressed (non-moraic) and stressed (moraic), respectively.

Our own work shows that all long vowels are indeed stressed, and each of these scholars accurately transcribed some, but not all, long vowels in their work. We define long vowels as bimoraic (heavy), whereas regular vowels are simply moraic. Long vowels in Haisla are exhaustively as follows:

(5) /uu/, /ii/, /aa/, /ai/, /au/

Phonetically, /ai/ and /au/ are realized as [ɛ] and [ɔ], and surface sounding much like /e/ and /o/, respectively. The sequences /ia/ and /ua/ do not coalesce to form diphthongs, but are realized as nuclei of separate syllables, e.g., [əja ~ əy̓a] or [əwa ~ əw̓a].[2] There are some words that contain more than one long vowel. In such cases the leftmost of these receives word-stress (see above).

While previous scholars have analysed the phonology of Haisla stress, to date no detailed phonetic work has been done to examine the acoustic correlates of phonological stress. This paper seeks to fill that gap.

---

[2] Transcriptions in previous work (Vink 1977; Lincoln & Rath 1986; Bach 2006) sometimes differ from our own, particularly in the vowel length distinction.

2 Methodology

2.1 The data

Our data come from recordings made in 2020 in Kitimat, BC. The 393 mp3 files were recorded using The Combine[3] with semantic domain elicitation methodology (Rapid Word Collection). Sound files include two repetitions of each word. The dataset also includes at least six different speakers, all over the age of 65. All the speakers recorded are members of the Haisla Nation and one of them is speaking Haisla again after nearly three decades of not using it.

The quality of data for this acoustic study was limited due to the nature of the available data. The data were recorded in mp3 format, which is not ideal for acoustic analysis. However, we found that the files were good enough quality for our study due to a relatively low compression rate. The data are additionally variable in sound quality. Since the data were not recorded for the purpose of acoustic study, the recording environment conditions were not always ideal or consistent. Though they could not be completely avoided, factors like background noise and clarity of the speakers were taken into consideration when choosing which sound files to use. The number of speakers in the data may have also impacted our study and results. Male and female speakers are both represented in the data, as well as a mix of language fluency among the speakers. Moreover, the data were provided with existing transcriptions of the words in the community orthography as well as in phonological form. The stress marked in these transcriptions was used as the basis for which words follow the expected Haisla stress assignment (predictable stress pattern words) and which do not (unpredictable stress pattern words). It is possible that not all these transcriptions are accurate, primarily the location of stress. A few of the transcriptions were revised as the acoustic study progressed. Each of these factors invariably contributed to high standard deviation of some of our measurements, but we expect that the key patterns and major trends of stress in the language have come through the data nevertheless.

2.2 Measurement conventions

For this study, 277 monomoraic vowel segments — /a,i,u/ — were targeted from 231 unique word recordings. From the 393 usable sound files, vowels in multi-word entries were automatically excluded as well as any vowels within VV sequences, like diphthongs, long vowels, and other ambiguous sequences. Additionally, at least one recording included major variation in pitch pattern between the two repetitions, most likely list intonation, and was excluded as well. Data were pulled from the first of the two tokens recorded of each word unless the recording quality or other acoustic factors noticeably hindered the ability to make accurate measurements in the first repetition. Recordings with bad background noise were also avoided, though a few were included when there would be less than 30 tokens of the target vowel otherwise.

When possible, 30 tokens of each vowel were used, in stressed and unstressed syllables. There were not enough words with unpredictable stress to do quantitative analysis. When there were over 30 tokens of a vowel to choose from, data were chosen across as many different words as possible. However, due to the size of the dataset and distribution of the target vowels, many words were used for multiple data points. For example, the recording of the /ɢˈałdəma/ ‘pajamas’ was used for two

---

[3] The Combine (https://thecombine.app/) is a tool developed by SIL International to support language documentation, specifically designed to support Rapid Word Collection (RWC, https://rapidwords.net/).

different predictable stress word tokens — stressed /a/ and unstressed /a/. A summary of the data is given in Table 1.

Table 1: Number of tokens used for each category of data

	Target vowel	**Predictable stress**	**Unpredictable stress**
Stressed	/a/	37	8
	/i/	36	3
	/u/	34	3
Unstressed	/a/	50	3
	/i/	50	3
	/u/	44	4

For each word used, the target vowels were annotated for their duration (ms), pitch (Hz), and intensity (dB). Boundaries were marked with the word waveform, spectrogram, and intensity in consideration. Where boundaries were unclear or transitions were very gradual, the boundary was marked at the midpoint of the transition. When target vowels were word final, Praat often read acoustic properties for longer than the vowel was impressionistically produced (e.g., reading an echo in the recording location). In these cases, the end boundary was marked within the final intensity transition, typically toward the end, also using the waveform for cues. Duration of the target vowels was measured and recorded from these segments.

The maximum pitch (F_0) and intensity were also measured for each vowel. For most vowels, the time that maximum pitch and intensity were measured over was the same as the segmented duration of the vowel. Sometimes, however, there was a clear transition or effect from an adjacent sound, like a jump in pitch caused by a preceding voiceless stop. In these types of instances, segmentation for pitch and intensity measurements were adjusted to exclude transitions or spikes not characteristic of the target vowel. Maximum pitch and intensity over the whole word containing each vowel were also measured. The word maxima were taken manually since the sound files were not segmented for the two repetitions recorded. This also helped catch, and adjust for, misreadings due to recording quality or voice quality effects; e.g., spikes in the pitch trace due to glottalization. The first three formants (F_1, F_2, F_3) of each target vowel were also measured. Each formant was measured at three points throughout the vowel: at the beginning, middle, and end. Working from the vowel's duration segmentation, these points were also taken manually, avoiding visible transitions in the formant tracks for the beginning points and end points.

From these data, we calculated the difference between the maximum pitch in the vowel and maximum pitch in the word from each vowel (Df0). The same was calculated from the intensity measurements for each vowel (Dint). For vowels from predictable stress pattern words, the average and standard deviation of duration, Df0, and Dint were found for stressed and unstressed vowels.

Following, we ran a Welch's two-sample t-test comparing the acoustic properties for the stressed and the unstressed vowels. Results can be found in Table 2 in the appendix.

With the formant measurements, we computed the average and standard deviation of F_1, F_2, and F_3 for each individual vowel. From there, we again applied the t-test to each formant, comparing stressed vowels and unstressed vowels; i.e., F_1 of stressed /a/ and unstressed /a/, F_2 of stressed /a/ and unstressed /a/, and F_3 of stressed /a/ and unstressed /a/; and the same for /u/ and /i/.

We ran a Welch's two-sample t-test on these data as well. Results can be found in Table 3 in the appendix.

3 Results

We found that three acoustic features correlate with the vowel of the stressed syllable to a very highly significant degree: intensity (Section 3.1), pitch (Section 3.2), and duration (Section 3.3). Additionally, formant structure is marginally relevant to word stress (Section 3.4).

3.1 Intensity

Stressed syllables are louder than unstressed syllables. This is realized in the acoustic signal with greater intensity of vowels on stressed syllables, measured in decibels. The results are shown in Figure 1, which shows the difference between two measurements: the peak intensity in the word minus the peak intensity within the target vowel.

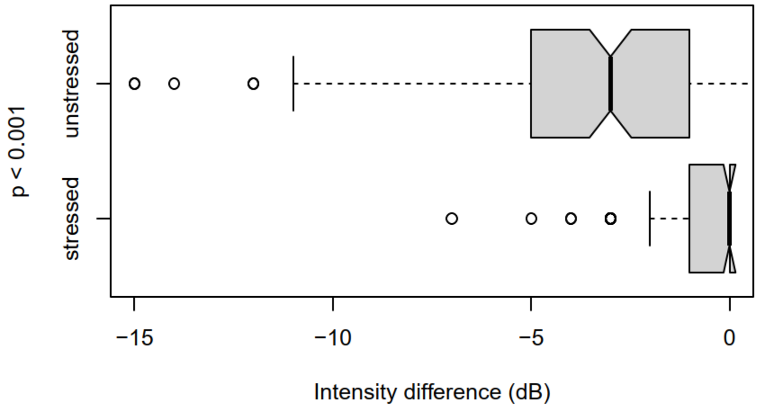

Figure 1: Intensity differences in the vowels of stressed and unstressed syllables

If intensity is a correlate of word stress, then the highest intensity in any given word should be within the stressed syllable. Thus, peak intensity in the word minus peak intensity within the target vowel should equal zero. As seen in Figure 1, the distribution for vowels in stressed syllables is close to zero. By contrast, the intensity of vowels in unstressed syllables should be lower than the peak intensity of the word, which should be in the stressed syllable elsewhere. Thus, we expect a negative value for vowels in unstressed syllables. As seen in the figure, this is indeed the case. A t-test run on these two groups shows the difference is very highly significant ($p < 0.001$).

3.2 Pitch

Stressed syllables have higher pitch than unstressed syllables. This is realized in the acoustic signal with a higher fundamental frequency on vowels in stressed syllables than those in unstressed syllables. The results are shown in Figure 2, where the numbers represent the difference between two measurements: the highest pitch in the word minus the highest pitch within the target vowel.

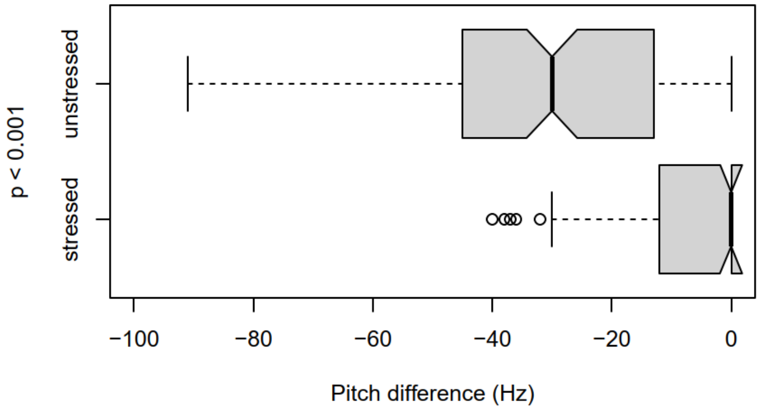

Figure 2: Pitch differences in the vowels of stressed and unstressed syllables

If pitch is an acoustic correlate of stress, the highest pitch in any given word should be within the stressed syllable. Thus, the highest pitch in a word minus the highest pitch in the stressed syllable should be zero. As seen in Figure 2, this figure is indeed near zero. By contrast, the highest pitch found in vowels in unstressed syllables is consistently lower than the highest pitch in the word. A t-test run on these two groups shows the difference is very highly significant ($p < 0.001$).

3.3 Duration

Stressed syllables have longer vowels than unstressed syllables. This is shown in Figure 3, where the duration of vowels in unstressed syllables is shown on the top and the duration of vowels in stressed syllables is shown on the bottom.

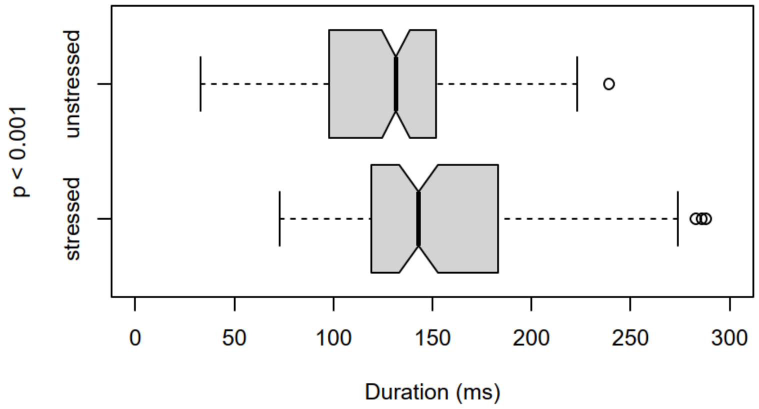

Figure 3: Duration of vowels in stressed and unstressed syllables

The duration of vowels varies considerably: 73–357 ms in stressed syllables and 33–239 ms in unstressed syllables. This is expected since we did not control for syllable type: the data include measurements from vowels in both open and closed syllables, in all word positions. For example, vowels in open syllables at the end of words are of course longer than vowels in closed syllables early in the word. Despite this confounding variable, the difference between the two sets should still show as statistically significant with enough data. A t-test run on these two groups shows the difference is very highly significant ($p < 0.001$).

3.4 Vowel formants

Vowels in stressed syllables are not generally articulated in more extreme positions in the oral tract in comparison to vowels in unstressed syllables. In most cases, measurements of F_1, F_2, and F_3 were not significantly different between the vowels of stressed and unstressed syllables.

A surprising exception to this generalization is the second formant of the vowel /a/. F_2 of /a/ is significantly lower in stressed syllables than in unstressed syllables, meaning that stressed /a/ is articulated farther back in the vocal tract than unstressed /a/. This is shown in Figure 4.

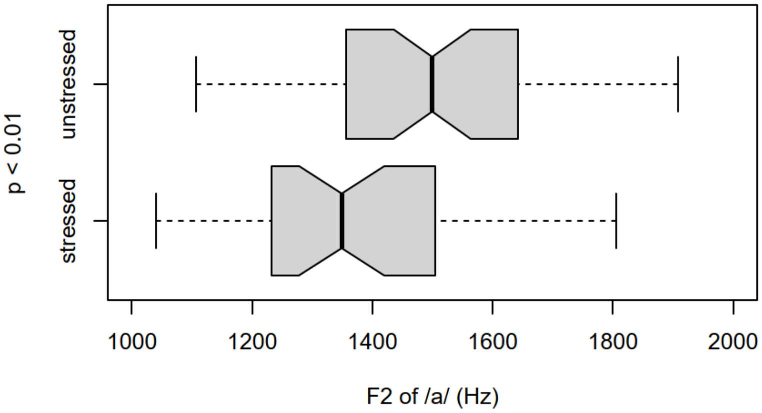

Figure 4: The F_2 of the vowel /a/ in stressed and unstressed syllables

The measurements shown in Figure 4 were taken at the midpoint of the vowel. A t-test run on these two groups shows the difference is highly significant ($p < 0.01$). We also took the average of three F_2 readings taken at the beginning, midpoint, and end of the vowel. Results are nearly identical.

It is not yet clear why stressed /a/ is articulated farther back than unstressed /a/. In the study by Gordon and Roettger (2017), formants correlated with word stress in 86% of the languages examined. This included all languages in which at least one formant correlated with at least one vowel. According to Gordon and Roettger's classification, Haisla is a language in which formant structure correlates with word stress; however, formants are certainly not a primary means of indicating stress in Haisla.

4 Implications

This section briefly addresses the implications of this phonetic research, both for scholars and for the Haisla language learning community. The phonetic analysis of word stress presented here has revealed that stress is signaled by higher pitch, greater intensity, and moderately greater duration of the vowel.

4.1 Implications for the phonology of Haisla stress

A majority of words in Haisla do not have long vowels, yet all have stressed syllables. These stressed syllables are significantly greater in intensity, higher in pitch, and longer in duration than other syllables within the same prosodic word. A stressed syllable, therefore, does not necessarily have an underlyingly long (or bimoraic) nucleus, but its phonetic duration is indeed measurably longer than non-stressed vowels.[4]

Earlier scholars assumed that relatively greater vowel duration in one syllable over others within a prosodic word correlated to word stress in all cases, as it does in other North Wakashan languages like Kwak’wala and likely ’Wùik̓ala.[5] They, therefore, assumed stress was not predictable, as it did not conform to the standard first described by Boas (1911). From Boas to as late as Grubb (1974) and Bach (1975), stress was assumed to correlate to phonemic vowel length, and was redundantly marked with an acute accent in transcription. Such an analysis has been proven unnecessary, assuming instead that acoustic correlates of stress, including vowel duration, are applied post-lexically to signal the stressed syllable of a word. For all other North Wakashan languages, positing a phonemic vowel length contrast is not necessary to accurately describe the stress systems in those languages. Only Haisla has retained contrastive vowel length in modern speech.[6] Earlier scholars may have, by analogy, assumed that phonemic vowel length was either always present in stressed syllables, or never present at all. Either assumption would shroud lexical vowel length as a factor in the predictability of Haisla word stress.

Our analysis of contrastive vowel length is based on moracity, similar to the analysis of the Kwak’wala stress system by Janzen (2015). This also follows the analysis of Grubb (1974), who states that schwa is non-moraic in Kwak’wala, and we assume all North Wakashan languages by extension. He states that schwa can be morphophonemic, or the result of either epenthesis at the morphophonemic level, or vowel reduction (Grubb 1974:42). Since schwa is non-moraic, it should not matter at what level that vowel enters the phonology, in regard to stress placement. It does, however, have some bearing on another prosodic constituent, namely prosodic feet. Until now, no analysis has been presented for any North Wakashan language that uses metrical feet to explain some unresolved aspects of the phonology. It may be that a close look at the interplay of underlying schwa against epenthetic schwa might yield some results that would help better explain the remaining inconsistencies in the stress placement of Haisla and its sister languages.

4.2 Implications for the Haisla speech community

A practical implication to our findings on vowel length is orthographic revision. The orthographies used throughout the literary history of Haisla have obscured some of the vowel length distinctions that have been uncovered through this phonetic analysis. Because stress was seen as largely unpredictable in this language, an acute accent on the stressed syllable was orthographically

---

[4] A study on the relative length of long vowels against short-but-stressed vowels is needed.

[5] More research on word stress in Oowikela is needed. Through superficial observation of data on the ’Wùik̓ala First Voices page and personal communication with John Rath and David Stevenson, we assume here that ’Wùik̓ala stress patterns as Kwak’wala does.

[6] Earlier accounts of the language (e.g., Boas 1911) posited phonemic vowel length of a version of the language that is now some 200 years old. South Wakashan languages do have contrastive vowel length either, so either Haisla is the one conservative North Wakashan language in this regard, or Haisla innovated a vowel length contrast, perhaps by analogy through intense contact with the surrounding Tsimshianic languages.

necessary for any non-monosyllabic word, including the orthography in use by the Haisla Nation today.

The task of integrating this research into the community orthography would mean a fairly comprehensive overhaul of the lexical corpus. Nearly every word entry in the modern dictionary would have to be evaluated for its stress predictability. When the stress is not correctly predicted, a further analysis would have to be carried out to ascertain if the discrepancy could be explained by the presence of a mis-transcribed long vowel. Part of this work has already been done through the research in this paper. The list of exceptions to predictable stress fell from 88 to 10 when a closer analysis of vowel length was applied. It is very likely that many more Haisla words have long vowels, mistaken for simple stressed syllables of greater duration.

The burden of going through an update of Haisla spelling to reflect predictable stress patterns might indeed be worthwhile. Remembering to write in the stress marking is by far the most neglected aspect of writing Haisla by learners and speakers alike. It is also an extra aspect to learn and remember when students take in more vocabulary. Learning the predictable stress system outlined above might well outweigh the burden for students of having to remember an extra diacritic.

Of course, this decision, as for all concerning how the language is used and presented, is up to the community and its language authority.

References

Bach, Emmon. 1990. Stem extensions in Haisla. *Papers for ICSNL* 25:1–15.

Bach, Emmon W. 1975. Long Vowels and Stress in Kwakiutl. *Texas Linguistic Forum* 2. 9–19.

Bach, Emmon W. 2006. Paradigm Regained: Deixis in Northern Wakashan. *SOAS Working Papers in* Linguistics 14:267–281

Boas, Franz. 1911. Kwakiutl. In Franz Boas (ed.), *Handbook of American Indian Languages* (Smithsonian Institution Bureau of American Ethnology 40), vol. 1. Washington, DC: Government Printing Office.

Elfner, Emily. 2008. *Focus in Kwak'wala*. University of Massachusetts Amherst.

Gordon, Matthew & Timo Roettger. 2017. Acoustic correlates of word stress: A cross-linguistic survey. *Linguistics Vanguard* 3(1).

Grubb, David. 1974. The position of /ə/ in the phonology of Kʷak̓ʷala and its use in the orthography together with a linguistic key to the alphabet. *Papers for ICSNL* 9:39–48.

Janzen, Jonathan. 2015. *Wałda̱mas? An exploration into the phonological-word in Kwak'wala*. Langley, BC: Trinity Western University MA thesis.

Janzen, Jonathan. 2023. North Wakashan stress systems. *Papers for ICSNL* 58:188–198.

Lincoln, Neville J. & John C. Rath. 1986. *Phonology, dictionary and listing of roots and lexical derivatives of the Haisla language of Kitlope and Kitimaat, B.C.* (Canadian Ethnology Service 103). Vol. 1. Ottawa, ON: Canadian Museum of Civilization.

Vink, Hein. 1977. A Haisla phonology. *Papers for ICSNL* 12:111–131.

Wilson, Stephen A. 1987. The development of tone in Heiltsuq. In *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*, 321–329. Berkeley: Linguistic Society of

America. https://journals.linguisticsociety.org/proceedings/index.php/BLS/article/view/1823. (22 May, 2024).

Appendix: Statistics

Table 2: Statistical summary of intensity, pitch, and duration

Correlate		Stressed	Unstressed	Welch's 2-sample t-test
Difference in intensity	Mean	-0.68 dB	-2.75 dB	$p < 0.001$
	SD	1.30 dB	7.05 dB	
Difference in F_0	Mean	-7 hZ	-33 hZ	$p < 0.001$
	SD	10 hZ	26 hZ	
Duration	Mean	154 ms	127 ms	$p < 0.001$

Table 3: Statistical summary of formants

Correlate		Stressed	Unstressed	Welch's 2-sample t-test
F_2 of /i/	Mean	377 Hz	356 Hz	$p > 0.05$
	SD	47 Hz	57 Hz	
F_2 of /u/	Mean	387 Hz	395 Hz	$p > 0.05$
	SD	68 Hz	88 Hz	
F_2 of /a/	Mean	704 Hz	706 Hz	$p > 0.05$
	SD	175 Hz	155 Hz	
F_2 of /i/	Mean	2171 Hz	2091 Hz	$p > 0.05$
	SD	307 Hz	439 Hz	
F_2 of /u/	Mean	1241 Hz	1224 Hz	$p > 0.05$
	SD	343 Hz	309 Hz	
F_2 of /a/	Mean	1362 Hz	1509 Hz	$p < 0.01$
	SD	174 Hz	227 Hz	
F_3 of /i/	Mean	2713 Hz	2726 Hz	$p > 0.05$
	SD	340 Hz	336 Hz	
F_3 of /u/	Mean	2554 Hz	2558 Hz	$p > 0.05$
	SD	259 Hz	227 Hz	
F_3 of /a/	Mean	2715 Hz	2636 Hz	$p > 0.05$
	SD	361 Hz	278 Hz	