

12-2017

## Vowel-Length Contrasts and Phonetic Cues to Stress: An Investigation of their Relation

Anya Hogoboom  
*William & Mary*, ahogoboom@wm.edu

Jessica Campbell

Mark Hutchens

Nick Kalivoda

Follow this and additional works at: <https://scholarworks.wm.edu/aspubs>



Part of the [Linguistics Commons](#)

---

### Recommended Citation

Hogoboom, Anya; Campbell, Jessica; Hutchens, Mark; and Kalivoda, Nick, Vowel-Length Contrasts and Phonetic Cues to Stress: An Investigation of their Relation (2017). *Phonology*, 34(3), 565-580.  
<https://doi.org/10.1017/S0952675717000288>

This Article is brought to you for free and open access by the Arts and Sciences at W&M ScholarWorks. It has been accepted for inclusion in Arts & Sciences Articles by an authorized administrator of W&M ScholarWorks. For more information, please contact [scholarworks@wm.edu](mailto:scholarworks@wm.edu).

# Vowel-length contrasts and phonetic cues to stress: an investigation of their relation

Anya Lunden,<sup>\*</sup> Jessica Campbell,<sup>\*</sup> Mark Hutchens,<sup>\*</sup> and Nick Kalivoda<sup>†</sup>

## Abstract

The functional load hypothesis of Berinsein 1979 put forward the idea that languages which use a suprasegmental property (duration, F0) contrastively will not use it to realise stress. The functional load hypothesis is often cited when stress correlates are discussed, both when noting that the language under discussion follows the hypothesis and when noting that it fails to follow it. In the absence of a larger assessment of how frequently languages conform, or not, to the hypothesis, it is unknown whether the functional load hypothesis is an absolute, a strong tendency, a weak tendency, or unsupported. The results of a database of reported stress correlates and use of contrastive duration for 140 languages are presented and discussed. No support for the functional load hypothesis is found.

**Keywords:** stress correlates; functional load hypothesis; duration

## 1 Functional load hypothesis

Berinsein (1979) put forward and experimentally investigated the hypothesis, stated below, that a non-random relationship exists between correlates of stress and the phonological structure of languages.

### (1) Berinsein’s Functional Load Hypothesis

“Change in F0, increased duration, and increased intensity, in that order, constitute the unmarked universal hierarchy for perception of stress in languages with no phonetic contrasts in tone or vowel length; in languages with such contrasts the perceptual cue correlated with that contrast (i.e. F0 with tone and duration with length) will be superseded by the other cues in the hierarchy.” (p. 2)

We put aside the issue of a universal hierarchical ordering<sup>1</sup> and focus on the claim for which Berinsein is classically cited: That languages that use a suprasegmental property

---

<sup>\*</sup>College of William & Mary

<sup>†</sup>University of California, Santa Cruz

<sup>1</sup>The hierarchical ordering of F0 > duration > intensity in Berinsein is taken from Hyman (1977:40) who bases it on experiments such as those of Fry (1955, 1958). It is now understood that the primary importance of F0 is in fact due to intonational pitch accents aligning with the primary stress (e.g. Pierrehumbert 1980, Beckman & Edwards 1994, Sluijter & van Heuven 1996, Gordon 2014).

contrastively will not use that property, or will only minimally use it, in the realization of stress; a proposal typically referred to as the Functional Load Hypothesis (FLH).

Berinstein experimentally compared English, which lacks a length contrast and is known to use duration to cue stress, to K'ekchi, which has phonemic vowel length. Her key findings were that (i) increasing vowel length on syllables did not make K'ekchi speakers more likely to recognise those syllables as stressed, in contrast to the results for English speakers, and (ii) F0 ratio, peak F0, and peak vowel intensity all correlate with stress in K'ekchi, whereas acoustic measurements showed that duration “is poorly correlated with stress” (p. 37). To rule out a connection between the fact that K'ekchi does not use duration and that it has fixed final stress, Berinstein experimentally investigated the perceptual effect of increased vowel duration on speakers of Kaqchikel,<sup>2</sup> a language closely related to K'ekchi that also has final stress, but that lacks a phonemic length contrast. Berinstein found that, unlike speakers of K'ekchi, speakers of Kaqchikel do equate increased duration with stressed-syllable identification. Berinstein thus concluded that there is evidence for the hypothesis put forward, quoted in (1) above, concluding:<sup>3</sup>

(2) Berinstein’s generalised conclusion, based on K'ekchi

“Duration has little or no cue value to the stress/no stress distinction in a language in which there are long and short vowels distinguished solely by length.” (p. 46)

While Berinstein’s original hypothesis (in (1)) also makes predictions for use of F0, we restrict our investigation to uses of duration. Berinstein’s (1979) investigation focuses on uses of duration, and, further, she notes (fn 1) that F0 may be used in different ways within the same language, as a language could have, for example, a high tone, but still use F0 to signal stress with a rising or falling tone, without violating the FLH.

Following Berinstein’s (1979) work, various languages have been reported to conform to the predictions of the FLH. Other languages with contrastive vowel length which eschew duration as a cue to stress include Anejom (Lynch 2000), Bhojpuri (Shukla 1981), Creek (Martin 2011), Fijian (Dixon 1988), Halkomelem/Musqueam (Suttles 2004), Hungarian (White & Mády 2008, Vogel *et al.* 2016), Konkow (Ultan 1967), Nisenan (Eatough 1999), and Palula (Liljegren 2008).

But while many languages provide support for the FLH, others run contrary to expectations. Languages with contrastive vowel length which *do* use duration as a stress correlate include Aleut (Rozelle 1997, Taff *et al.* 2001), Arabela (Rich 1963), Barok (Du 2010), Chickasaw (Gordon 2004), Diola-Fogny (Sapir 1969), Leti (Hume *et al.* 1997, Hume 1997), Masbatenyo (Rosero 2014), Nhandá (Blevins 2001), and Northern Paiute (Snapp *et al.* 1982, Thornes 2003).

As published work often reports only on a single language, most reflections on whether or not Berinstein’s FLH holds are made on a language-by-language basis. For example, Taff *et al.* (2001) specifically address the consequences of Aleut, which has both contrastive vowel length and uses duration to cue stress, for the FLH (see also Rozelle 1997). The authors

---

<sup>2</sup>Berinstein uses the alternative spelling *Cakchiquel* for this language.

<sup>3</sup>Berinstein at this point separates her proposal into two component parts, the Functional Load Hypothesis and the related hierarchical cue reordering.

propose that perhaps, while still correlating with stress, duration may be a weaker cue to stress in languages with contrastive vowel length. However, it is difficult to judge how robust the FLH is, or to what degree it holds on a case-by-case basis.

The exception to the language-specificity of many publications referencing the FLH is the ongoing work of the University of Delaware Stress Lab, whose website lists 15 languages that are being investigated with standardized methodology that separates word-level stress from sentence-level focus ([sites.google.com/site/udstresslab/home](http://sites.google.com/site/udstresslab/home)). The results of the acoustic investigation of four of these languages (Greek, Hungarian, Spanish, and Turkish) are presented in Vogel *et al.* (2016). The authors found results consistent with what they refer to as an extended version of the FLH: Hungarian, the language of the four with contrastive vowel length, was found not to use duration as a cue to stress (also see Vogel *et al.* 2015), and Greek and Spanish were found to use different cues for word-level stress (F0, in both cases) and sentence-level focus (duration, in both cases). This kind of careful experimental investigation is what is needed to ultimately assess the validity of the FLH across languages as well as understand more broadly what types of cues remain distinct.

It is worthwhile, however, to assess the correlation between vowel length and stress-related duration based on the current state of reported stress correlates across a wide variety of languages.<sup>4</sup> To this end, we have looked for evidence of Berinstitute’s FLH by compiling a database of languages for which we could find information on stress correlates. The Stress Correlate Database, which has 140 languages, is introduced in the following section, and the correlations between uses of duration are examined in §3.

## 2 Stress correlate database

The Stress Correlate Database (Lunden and Kalivoda 2016) is a compilation of reported stress correlates among the world’s languages. A main source for the database was grammars (books, PhD and MA theses), but it also reflects information from articles, manuscripts, and personal communications with researchers. The database includes information on 140 languages from 53 language families (with one creole, one unclassified, and four isolates counted as belonging to separate families). The parameters included are listed, grouped by type, in (3). We then discuss the coded levels for each and how they were determined.

### (3) Database parameters

- Language information
  - Language name
  - Primary family
  - Secondary family
  - ISO code
- Stress system information
  - Stress system type (single stress, binary stress, dual stress, ternary stress)

---

<sup>4</sup>Note that the term “correlate” is used here for phonetic cues, and in the further discussion of correlations, to mean “potential correlate.” On the other hand, we use the term “correlation” refers not to phonetic cues to stress, but to relations between the status of phonological vowel length and the phonetic stress correlates.

- Primary stress edge (left/right)
- Stress correlate information
  - Duration reported correlate
  - Pitch reported correlate
  - Intensity/loudness/energy/spectral tilt reported correlate
  - Primary correlate if known
  - Vowel reduction (phonological) in unstressed syllables<sup>5</sup>
- Phonological uses of suprasegmentals
  - Duration used contrastively outside of stress
  - Pitch used contrastively outside of stress<sup>6</sup>
- References
  - Sources consulted

The language name is generally the one given by the source article, although in some cases an updated one is used instead. The primary language family is the largest known family affiliation and the secondary family gives the relevant branch of larger families.

The *stress system type* column indicates whether a language is reported to have only a single stress per word, a binary system, a dual stress system, or a ternary system. The single main stress in languages of the first type may be accompanied by secondary stress that is unpredictable in location and number (i.e. not rhythmic). If a primary stress and one secondary stress were both mentioned, with secondary stress occurring either two syllables after (if the primary stress was leftward) or before (if the primary stress was rightward), the language was assumed to be binary unless other information was provided. Stress systems are coded as dual if the primary and secondary stresses appear to come from opposite directions and no other (non-weight-sensitive) stresses were mentioned; e.g., if the primary stress is penultimate and the secondary stress initial. The cells in this column in the database each have comments that briefly describe the stress pattern. The stress pattern was taken from the stress correlate source where possible; otherwise additional sources were consulted. Following the usual practice for stress systems, quantity-sensitive systems (where weight affects the stress pattern) are described based on the stress pattern in words with all light syllables, unless otherwise stated. Whether stress is leftward or rightward is also coded. Lexical stress, if no stress window was listed, is not coded as either leftward or rightward.

The reported stress correlates were taken from descriptions of the realization of stress. Often, researchers only discussed the correlates used by the language, and so other suprasegmental correlates (duration, pitch, and/or intensity) are coded as not present even if not specifically mentioned. The exception to this practice is when a correlate was mentioned in passing in a different context, or when the status of a correlate is specifically mentioned as not known. In such cases the other correlates are left blank, rather than coded as not present.

---

<sup>5</sup>Vowel reduction is included for completeness of stress correlates. As it is not a suprasegmental property it is not considered further here.

<sup>6</sup>We looked for and coded whether F0 was used contrastively, although it is not referenced in the analysis presented here. Only 21 languages in the database do use F0 contrastively.

Sometimes correlates were stated with a hedge of some kind. Any statement that indicated the correlate was present more often than not was treated as a positive statement. Thus, correlates that were mentioned as occurring “frequently,” “often,” or “typically” are coded as being present, whereas those described as occurring “occasionally,” “to some degree,” or “for some speakers,” or were associated with the phrase “may occur,” are coded as being absent. The quote that each language’s stress correlates was based on is included as a cell comment in the duration correlate column (whether or not duration is used). Finally, because sources did not consistently clarify whether correlates were for primary or secondary stress, the database conflates the two.<sup>7</sup> If a correlate was mentioned as being present in the realization of any word-level stress, it was coded as positive. Sometimes stress, especially in older texts, was only referred to by a correlate (e.g. “the tone” Dillmann 2005/1907); however, if that feature clearly followed a culminative pattern that was present in every word,<sup>8</sup> we assumed that the author was referring to stress and coded the term the author used as a correlate.

Some further notes are relevant for each of the suprasegmental correlates. Regarding duration, the occurrence of long vowels only in stressed syllables was not taken as sufficient evidence for duration as a stress correlate, unless it was specifically mentioned that short vowels were longer when stressed (or, hypothetically, if all stressed vowels became long and long vowels only occurred under stress). The use of pitch as a word-level stress correlate can be difficult to untangle from pitch accents due to sentence-level intonation. We did not code pitch as a correlate if it was clear the author was discussing sentence-level stress. However, as Gordon (2014) and Vogel *et al.* (2016) discuss, many descriptions of stress correlates are based on words uttered in citation form and thus with sentence-level stress/intonation. Therefore, this confound was undoubtedly present in some data presented as information about word-level stress. Finally, we grouped statements referencing physical, acoustic, and auditory properties such as “force,” “energy,” “intensity,” “loudness,” and “spectral tilt” under the heading *intensity*, as the inconsistency in terminology does not, in most cases, appear to reflect a real difference in the cue being referenced.<sup>9</sup>

Although vowel reduction can serve as a correlate of stress, it is common for stress correlates to be discussed without mention of the segmental correlate of vowel reduction. Therefore, we did not code vowel reduction as not present when it was unmentioned in the discussion of stress correlates. In most cases we had to look more broadly (within a source or to a different source) in order to determine whether there was any evidence for or against vowel reduction. While vowel reduction occurs on a spectrum, we only considered phonological vowel reduction, categorised as significant reduction, to the point of different IPA symbols. (We assume that notable vowel reduction will reliably be reflected in IPA

---

<sup>7</sup>This was done because (i) the descriptions of only six languages listed correlates for primary and secondary stresses separately and (ii) Berinstitute’s FLH is not specific to a particular level of stress.

<sup>8</sup>Culminativity is classically considered a property of stress systems and not present in tonal systems (e.g. Hyman 1977, Hayes 1995).

<sup>9</sup>The exception may be spectral tilt, which is an acoustic measurement of the intensity distribution across frequencies. Sluijter and van Heuven (1996) found that while this intensity distribution (higher intensity specifically in frequencies above 500 Hz) is a reliable correlate for stress in Dutch, overall intensity is not. However, since spectral tilt is not discussed frequently enough to be separated out in the database we coded it as “intensity.”

transcriptions, whereas vowel undershoot would not be.) This typically involves a reduction of the number of vowel contrasts, although not necessarily; i.e. Bhojpuri phonologically reduces all vowels although they remain distinct (Shukla 1981). Thus, phonetic vowel undershoot (the sense of undershoot in Lindblom 1963 and Stevens and House 1963) and other phonetic reduction were not considered vowel reduction. Vowel reduction was also coded as not present if we could clearly see that every vowel could be present in both stressed and unstressed positions, or if, in a section devoted to phonological quality changes of vowels or vowel distribution, mention of vowel reduction was absent. The quotes that prompted the determinations are given as comments in the cells.

In cases where multiple stress correlates were given, one or two were often mentioned as being stronger, more reliable, or more important. The database therefore codes for the primary correlate when either duration, pitch, or intensity was singled out. No primary correlate is coded in cases where two correlates were said to be more important than a third because, in addition to not identifying a unique primary correlate, this type of statement was often used to show the weakness of one correlate rather the importance of the others.

Whether length or pitch are used contrastively outside each language's stress correlates was also determined. If the stress correlate source did not include the relevant information (phonemic inventory, etc), then an additional source was consulted in order to ascertain whether vowel length is contrastive in the language. We also searched for any mention of tone or contrastive pitch accent in the language. We assumed that a contrast was absent when it was not mentioned; quotes are therefore provided whenever the feature is listed as present, but not if the feature is not present unless the author explicitly mentioned its absence. When languages were described as having a vowel contrast based on both length and quality, we coded these as not having contrastive length. This was done because the vowel reduction gives a cue stress that could be utilized by listeners, who would then not need to attend to the actual vowel length.<sup>10</sup>

## 2.1 Testing the database

Some of the reported correlates are the result of acoustic investigation on the language; however, most are the researchers' auditory impressions of how stress is realised. While care was taken to apply a consistent set of criteria to the coding of the presence/absence of various correlates, the fact that the information came from a wide variety of researchers across times and locations means that the auditory criteria applied by the researchers were not in any way standardised.

In order to test the database for evidence of consistency, we worked from the assumption that languages within a family are more likely to resemble each other in their realization of stress than a randomly-selected subgroup from the database. Looking for robust-enough language families to compare to random samplings, we found there to be seven language families that have at least four members.

A Monte Carlo simulation was used to determine where each of these seven language families fell, for each parameter, in the distribution of language sets of the same size taken

---

<sup>10</sup>There were four such languages, including English, which Berinstein (1979) considers not to have a length contrast. See fn. 12 for the adjustment to Table 4 if these languages were coded as having contrastive vowel length.

from the database (based on 10,000 random pulls from the database). We compared the number of actual languages within each family that use each of duration, pitch, and intensity to the average number found for sets of language groupings of the same number of languages from the overall database. For example, of the 18 Austronesian languages, 72% have duration as a reported stress correlate, 83% have pitch, and 61% have intensity. We then compare this to the average number to use each correlate for 18 languages randomly pulled from the database. The average over 10,000 random sets of 18 languages was that 59% used duration, 66% used pitch, and 61% used intensity. Thus we can see that the actual Austronesian languages are more likely than any average group of 18 languages to use duration and pitch, but are just as likely to use intensity.

Each value (from each the three parameters in each of the seven language families) was calculated as a z-score (based on distributional curves for each language family size). The absolute values of these z-scores were taken, as a language family may show internal consistency either by tending to use a particular stress correlate or by tending to eschew a particular stress correlate. The actual mean of the absolute z-scores was 0.7361. In comparison, the mean absolute z-scores across sampled same-sized sets of languages was 0.5107, with a standard deviation of 0.0905.

Thus, the value from our actual families is 2.4908 standard deviations away from the expected value, which is a significant difference ( $p = 0.0127$ , two-tailed). This shows that the actual languages within families are more similar to each other than we would typically find with random groupings of languages. Specifically, we would have to pull random language sets of the size of those seven families an average of 87 times (based on 500 iterations) in order to find one with a z-score of at least 0.7361.<sup>11</sup> Thus we have evidence that the database shows something of the internal consistency that we would expect to find.

Another reasonable expectation we can test the database for is the correlation of intensity and pitch. As noted by Lehiste (1970:125), and supported by Titze (1989), the subglottal pressure that results in increased intensity also causes an increase in the rate of vocal fold vibration. Lehiste notes that increased duration, on the other hand, cannot be seen as physiological consequence of increased respiratory effort. Given this, we would expect to find that languages reported to have either pitch or intensity as a stress correlate would have the other one as well, but that this should not hold for the co-occurrence of either one with duration. Table 1 shows that the occurrence of pitch and intensity both being reported as stress correlates is more frequent than either being reported without the other. Table 2 is restricted to only those languages reported to have exactly two of the three suprasegmental correlates.

---

<sup>11</sup>If we increase the number of members necessary for a language family to be included to seven, then only three families meet the criterion. While the mean absolute z-score for the parameters of the families is similar (0.9639) as is the mean across all language sets pulled from the database (0.6447), the standard deviation of those random sets is much higher (0.1751) because of the small number of sets involved, meaning the value from our actual language families is only 1.8229 standard deviations away from the expected value, which is not significant ( $p = 0.0683$ , two-tailed).



Table 1: Correlation of use of Pitch/Intensity for all languages

Pitch and Intensity	59
Pitch but not Intensity	34
Intensity but not Pitch	25

Table 2: Correlation of use of Pitch/Intensity for languages with exactly 2 correlates

Pitch and Intensity	33
Pitch and Duration	20
Intensity and Duration	15

We can see from both tables that if either intensity or pitch is reported to be a stress correlate, it is more likely than not that the other is also a reported correlate, and that the likelihood of either occurring with duration is much lower. One-sample Chi-square tests of equal proportions for the data in Tables 1 and 2 found that the three categories are not distributed equally ( $p < 0.001$ ,  $p = 0.022$  respectively).

Both the Monte Carlo simulation and the pitch/intensity correlation indicate that the database is robust enough to be informative regarding the Functional Load Hypothesis.

### 3 Correlations

Table 3 shows the correlations between contrastive vowel length and duration (of any sort) as a reported correlate of stress.

Table 3: Correlations of stress-based duration and contrastive vowel length

		Stress correlate		
		Yes	No	Total
Contrastive V length	Yes	45	37	82
	No	38	20	58
Total		83	57	140

We see that of the 82 languages that have contrastive vowel length, 54.9% use duration as a stress correlate, which is not very different from the 65.5% of the 58 languages that do not have contrastive vowel length that use duration as a stress correlate. Looking at it the other way around, of the 83 languages which reportedly use duration as a stress correlate, 54.2% nevertheless have contrastive vowel length, compared to 64.9% of the 57 languages

which do not use duration as a stress correlate and have contrastive vowel length. A Chi-Square statistical test shows there is not a significant relationship between duration as a stress correlate and contrastive vowel length ( $\phi = -0.107$ ;  $\chi^2(1) = 1.593$ ,  $p = 0.207$ ).

A few of the languages that use duration as a correlate of stress are in fact described as lengthening a (preceding or following) consonant in at least some cases. For example, Łukaszewicz (2015) found that secondary stress in Polish is realised through additional duration of the onset consonant compared to that of the preceding vowel, and the following consonant is reported to lengthen in Nambikuara as the durational cue to stress in closed syllables (Kroeker 2001). Languages were coded for whether they always used only consonant duration as a correlate (N=5), sometimes used consonant duration without simultaneously using vowel duration as a correlate (N=4), or sometimes or always used consonant duration as a correlate in addition to vowel duration (N=1). Because of the greater frequency of vowel duration as a correlate in comparison to consonant duration, we assumed that mention of duration referred to the length of the vowel unless the author clearly specified otherwise. Table 4 is a reexamination of Table 3 which limits the durational correlate to consistent vowel duration. Languages in which consonant lengthening, whether consistent as in Polish, or structure-dependent as in Nambikuara, is the durational stress correlate (without concurrent vowel duration) have been reassigned to the second column.

Table 4: Correlations of stress-based vowel duration and contrastive vowel length

		Stress correlate		
		Yes	No	Total
Contrastive V length	Yes	40	42	82
	No	34	24	58
Total		74	66	140

Even with the cases involving consonantal duration counted as not using duration as a stress correlate, 48.8% of languages with a vowel length contrast use (specifically vowel) duration as a stress correlate, which is not very different from the 58.6% of the languages without a vowel contrast that use (vowel) duration as a cue to stress. A Chi-square test finds no relationship between the use of vowel duration to cue stress and the presence of contrastive vowel length ( $\phi = -0.097$ ;  $\chi^2(1) = 1.320$ ,  $p = 0.251$ ).<sup>12</sup>

We could imagine that increased duration might be present to some degree in stressed syllables, and therefore be reported as a stress correlate, but not carry the primary load of cueing stress. While Berinstein’s conclusion (quoted in (2)) states that duration will have “little or no [stress] cue value” in languages that use duration contrastively, her original hypothesis (quoted in (1)) suggests that languages might still use stress cues that also have a contrastive function in the language, but to a lesser degree than other stress cues.

<sup>12</sup>The results do not change if the four languages in the database described as having a length contrast with accompanying tenseness difference (see discussion at the end of §2.0) are coded as having contrastive vowel length ( $\phi = -0.072$ ;  $\chi^2(1) = 0.730$ ,  $p = 0.393$ ).

Many sources, where multiple stress correlates were listed, singled out one which was said to be the most important or the key component. We coded cases where a single correlate was identified as the primary correlate so that the correlations of this subset with the use of contrastive vowel length could be examined. Languages for which only one correlate was reported were coded as having that correlate as the primary correlate. Table 5 shows, in the first column, the primary correlate for the 63 languages for which a primary correlate was identified, and, in the subsequent columns, the number of those that use vowel duration, pitch, or intensity as a correlate (of any standing).<sup>13</sup>

Table 5: Correlate use for languages with proposed primary correlates

		N with primary stress correlate	also use V duration	also use pitch	also use intensity
Primary Correlate	V Duration	18	NA	4	1
	Pitch	31	8	NA	12
	Intensity	14	3	3	NA

For those languages that are reported to use duration as a correlate of stress, Table 6 shows the correlations between using vowel duration (i.e. the more restrictive “duration as a stress correlate” criterion reflected in Table 4) as the primary correlate and having a vowel length contrast.

Table 6: Correlations of duration as primary stress cue and contrastive vowel length among languages that use stress-based duration

		Primary stress correlate		
		Yes	No	Total
Contrastive V length	Yes	10	8	18
	No	8	3	11
Total		18	11	29

As seen in Table 6, even when the dataset is restricted to those languages that are identified as having a primary stress correlate, no aversion to using duration as the primary cue to stress is seen in languages with contrastive vowel length. Of the 18 languages that have contrastive vowel length and have a primary correlate, 55.6% use duration as that primary

<sup>13</sup>Of the nine languages for which consonant duration is sometimes or consistently the durational cue to stress, six languages are reported to have a primary correlate of duration. Those six are entirely excluded here in order to focus on uses of vowel duration. Of the three remaining consonant-duration languages, one was not reported to have a primary correlate, and the other two were reported to have different primary correlates. These two languages are included in Table 5 but are not included in the tally of “also use V duration.”

stress correlate, compared to 72.7% of the 11 languages lacking a vowel length contrast that use duration as the primary stress correlate. A Fisher’s exact test (used because the cell counts are too small for a Chi-square test) finds no significant relationship between the primary stress correlate and the use of contrastive vowel length among languages that use duration to cue stress ( $\phi = -0.172$ ,  $p = 0.449$ ).

A reviewer suggested examining the correlations for the subset of languages for which the reported correlates were based on acoustic studies. There are 46 such languages,<sup>14</sup> and the correlation within this subset of vowel-based duration as a cue to stress with contrastive vowel length is given in Table 7.

Table 7: Correlations of stress-based vowel duration and contrastive vowel length in acoustically-investigated languages

		Stress correlate		
		Yes	No	Total
Contrastive V length	Yes	17	8	25
	No	16	5	21
Total		33	13	46

We see a very similar distribution to that found in Table 4 for the full dataset. Of the 25 languages with contrastive vowel length, 68.0% use duration as a cue to stress, which is similar to the 71.7% of languages without contrastive vowel length which use duration as a cue to stress. A Chi-square test finds no relationship between the use of vowel duration to cue stress and the presence of contrastive vowel length in this subset of the data ( $\phi = -0.091$ ;  $\chi^2(1) = 0.378$ ,  $p = 0.539$ ).

## 4 Conclusion

The Functional Load Hypothesis is an intuitively appealing hypothesis, and if correct would provide a natural explanation for, e.g., the stress-related suprasegmental differences between K’ekchi and Kaqchikel. Is not unusual to find Berinsein’s work referenced as demonstrating that languages with a contrastive vowel length do not use duration in the realization of stress. However, many languages must be considered in order to truly test the hypothesis. The stress correlate database presented here is designed to do this, and indicates, perhaps surprisingly, that the FLH is not borne out. Among the languages in the database, there is no correlation between a language’s use of vowel duration in the realization of stress and that language’s use of contrastive vowel length, and this does not change when the relative importance of the stress correlates in a given language is taken into account.

<sup>14</sup>Since this division was not originally coded, we identified this subset based on whether the stress cue quotes referenced measurements, and, in cases that possibly, but not clearly, included acoustic evidence, we went back to the source. We did not, however, revisit every source.

We conjecture that these issues for the FLH are due at least in part to languages' flexibility in putting suprasegmentals to work. For example, Taff *et al.* (2001) document that in Aleut, long vowels are statistically significantly longer than short vowels, and stressed vowels are statistically significantly longer than unstressed vowels of the same phonological length. This results in four levels of duration, showing that increased duration under stress does not conflate the vowel length contrast. The phonetic details of duration in Aleut give an insight into how languages can use duration both for a phonological contrast and as a phonetic correlate of stress. Thus we see that duration as a correlate of stress may co-exist alongside contrastive length without obscuring these contrasts. Further, as a reviewer of this paper notes, there are multiple additional sources of durational differences, from final lengthening, syllable shape, effects from neighboring consonants, etc., and so it perhaps should not be surprising to find that duration can be used for both contrastive length in vowels and to cue stress.

Others have also noted that Berinsein's FLH may be too strong. For example, Turk (2012:251) suggests that "prosodic use of duration is constrained, but not precluded, by its use for segmental contrasts," and also suggests that other uses of duration may be the reason that three-way quantity distinctions are rare; as two-way systems leave room for other influences on duration.

It is important to emphasize that while our results raise a question of whether there is any correlation between language's use of contrastive vowel length and use of duration as a stress correlate, we do not expect the results of our typological study to conclusively decide the fate of the FLH. As Vogel *et al.* (2016) point out, inconsistent methodologies cast doubt on even the validity of acoustic investigations. Further, Vogel *et al.* (2016) find support for a pressure for languages to distinguish the cues to word-level stress and focus accent. Suomi *et al.* (2003:114) suggest that Berinsein's work actually investigated sentence accent. And so it may be that there is a level at which Berinsein's FLH does hold, even if it is not between segmental contrasts and word-level stress.

We hope that the survey of languages presented here is helpful not only in showing that languages may use duration both in vocalic contrasts and in the domain of stress, but that it serves as a useful reference for identifying languages for which more detailed acoustic investigations should be done to clarify the use of duration at different levels.

## Acknowledgements

The Stress Correlate Database has benefited from many additional researchers/assistants from 2011-2016 that has allowed it to reach its current form. Listed chronologically, they are: Sara Carter, Rachel Nabulsi, Lindsay Baldwin, Kylie Ehlers, Kelly Luo, and Diana Worthen. The authors also thank Anna Henshaw. We thank Kim Love of K. R. Love Quantitative Consulting and Collaboration for consultation regarding aspects of the statistical tests reported here.

## References

- Beckman, Mary E. & Jan Edwards (1994). Articulatory evidence for differentiating stress categories. In Patricia A. Keating (ed.) *Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form*. Cambridge: Cambridge University Press. 7–33.
- Berinstein, Ava (1979). *A Cross-Linguistic Study on the Perception and Production of Stress*. Master's thesis, University of California, Los Angeles.
- Blevins, Juliette (2001). Nhanda: An aboriginal language of Western Australia. *Oceanic Linguistics Special Publications No. 30*. i–170.
- Dillmann, August & Carl Bezold (2005). *Ethiopic Grammar*. Eugene, Oregon: Wipf and Stock Publishers. Revised version of Dillmann 1907.
- Dixon, Robert M.W. (1988). *A Grammar of Boumaa Fijian*. Chicago: University of Chicago Press.
- Du, Jingyi (2010). *Towards a Grammar of the Usen Dialect of the Barok Language New Ireland, Papua New Guinea*. PhD dissertation, La Trobe University.
- Eatough, Andrew (1999). *Central Hill Nisenan Texts with Grammatical Sketch*. University of California Publications in Linguistics 132. Berkeley: University of California Press.
- Fry, D.B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America* **27**. 765–768.
- Fry, D.B. (1958). Experiments in the perception of stress. *Language and Speech* **1**. 126–152.
- Gordon, Matthew (2004). A phonological and phonetic study of word-level stress in Chickasaw. *International Journal of American Linguistics* **70**. 1–32.
- Gordon, Matthew (2014). Disentangling stress and pitch-accent: a typology of prominence at different prosodic levels. In Harry van der Hulst (ed.) *Word Stress: Theoretical and Typological Issues*. Cambridge University Press. 83–118.
- Hayes, Bruce (1995). *Metrical Stress Theory: Principles and Case Studies*. Chicago Illinois: University of Chicago Press.
- Hume, Elizabeth (1997). Vowel preservation in Leti. *Oceanic Linguistics* **36**. 65–101.
- Hume, Elizabeth, Jennifer Muller & Aone van Engelenhoven (1997). Non-moraic geminates in Leti. *Phonology* **14**. 371–402.
- Hyman, Larry (1977). On the nature of linguistic stress. In Larry Hyman (ed.) *Studies in Stress and Accent*. Vol. 4. Southern California Occasional Papers in Linguistics. 37–82.
- Kroeker, Menno (2001). A descriptive grammar of Nambikuara. *International Journal of American Linguistics* **67**. 1–87.

- Lehiste, Ilse (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.
- Liljegren, Henrik (2008). *Towards a Grammatical Description of Palula: An Indo-Aryan Language of the Hindu Kush*. PhD dissertation, University of Stockholm.
- Lindblom, Björn (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* **35**. 1773–1781.
- Lukaszewicz, Beata (2015). Polish rhythmic stress revisited: phonetic evidence of an iterative system. Poster given at Phonetics and Phonology in Europe, Cambridge, June 2015.
- Lunden, Anya & Nick Kalivoda (2017). Stress correlate database. Online database of stress correlates: <http://wmpeople.wm.edu/sllund/stresscorrelatedatabase>.
- Lynch, John (2000). *A Grammar of Anejom*. Pacific linguistics. The Australian National University, Canberra.
- Martin, Jack (2011). *A Grammar of Creek (Muskogee)*. University of Nebraska Press.
- Pierrehumbert, Janet (1980). *The Phonology and Phonetics of English Intonation*. PhD dissertation, Massachusetts Institute of Technology.
- Rich, Furne (1963). Arabela phonemes and high-level phonology. *Studies in Peruvian Indian Languages* **1**. 193–206.
- Rosero, Michael Wilson I. (2014). *A Grammatical Sketch of Masbatenyo*. Master's thesis, University of the Philippines.
- Rozelle, Lorna (1997). The effect of stress on vowel length in Aleut. *UCLA Working Papers in Phonetics* **95**. 91–101.
- Sapir, J. David (1969). *A Grammar of Diola-Fogny: A Language Spoken in the Basse-Casamance Region of Senegal*. Cambridge University Press.
- Shukla, Shaligram (1981). *Bhojpuri Grammar*. Washington D.C.: Georgetown University Press.
- Sluijter, Agaath & Vincent van Heuven (1996). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America* **100**. 2471–2485.
- Snapp, Allen, John Anderson & Joy Anderson (1982). Northern Paiute. In Ronald Langacker (ed.) *Studies in Uto-Aztecan Grammar III*. Dallas, TX: Summer Institute of Linguistics and the University of Texas. 1–92.
- Stevens, Kenneth & Arthur House (1963). Perturbation of vowel articulations by consonantal context: An acoustical study. *Journal of Speech and Hearing Research* **6**. 111–128.
- Suomi, Kari, Juhani Toivanen & Riikka Ylitalo (2003). Durational and tonal correlates of accent in Finnish. *Journal of Phonetics* **31**. 113–138.

- Suttles, Wayne (2004). *Musqueam Reference Grammar*. Vol. 2. University of British Columbia Press.
- Taff, Alice, Lorna Rozelle, Taehong Cho, Peter Ladefoged, Moses Dirks & Jacob Wegelin (2001). Phonetic structures of Aleut. *Journal of Phonetics* **29**. 231–271.
- Thornes, Timothy (2003). *A Northern Paiute Grammar with Texts*. PhD dissertation, University of Oregon, Eugene.
- Titze, Ingo R. (1989). On the relation between subglottal pressure and fundamental frequency in phonation. *The Journal of the Acoustical Society of America* **85**. 901–906.
- Turk, Alice (2012). Temporal implementation of prosodic structure. In Abigail Cohn, Cécile Fougeron & Marie Huffman (eds.) *Oxford Handbook of Laboratory Phonology*. Oxford University Press. 242–253.
- Ultan, Russell (1967). *Konkow Grammar*. PhD dissertation, University of California, Berkeley.
- Vogel, Irene, Angeliki Athanasopoulou & Nadya Pincus (2015). Acoustic properties of prominence in Hungarian and the Functional Load Hypothesis. In Katalin É. Kiss, Balázs Surányi & Éva Dékány (eds.) *Approaches to Hungarian 14*. Amsterdam: John Benjamins. 267–292.
- Vogel, Irene, Angeliki Athanasopoulou & Nadya Pincus (2016). Prominence, contrast and the functional load hypothesis: An acoustic investigation. In Jeffrey Heinz, Rob Goedemans & Harry van der Hulst (eds.) *Dimensions of Phonological Stress*. Cambridge University Press. 123–167.
- White, Laurence & Katalin Mády (2008). The long and the short and the final: Phonological vowel length and prosodic timing in Hungarian. In *4th Speech Prosody Conference, Campinas, Brasil*. 363–366.