Phonology **34** (2017) 565–580. © Cambridge University Press 2017
doi:10.1017/S0952675717000288

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

Anya Lunden
Jessica Campbell
Mark Hutchens
College of William and Mary

Nick Kalivoda
University of California, Santa Cruz

The functional load hypothesis of Berinstein (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 it is observed that the language under discussion follows the hypothesis and when it fails to follow it. In the absence of a more wide-ranging assessment of how frequently languages do or do not conform to the functional load hypothesis, it is unknown whether it is an absolute, a strong tendency, a weak tendency or unsupported. The results from 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.

1 The functional load hypothesis

Berinstein (1979) put forward and experimentally investigated the hypothesis, stated in (1), that a non-random relationship exists between correlates of stress and the phonological structure of languages.

* E-mail: LUNDEN@WM.EDU, JACAMPBELL@EMAIL.WM.EDU, MHHUTCHENS@EMAIL.WM.EDU, NKALIVOD@UCSC.EDU.

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

(1) *Functional Load Hypothesis* (Berinstein 1979: 2)

> 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such contrasts the perceptual cue correlated with that contrast (i.e. F0 with tone and duration with length) will be superseded by the other cues in the hierarchy.

We put aside the issue of a universal hierarchical ordering,[1] and focus on the claim for which Berinstein is classically cited: languages that use a suprasegmental property contrastively will not use that property, or will use it only minimally, in the realisation of stress; this is the 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' (1979: 37). To rule out a connection between the facts 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 ('Cakchiquel' in Berinstein 1979), a language closely related to K'ekchi that also has final stress, but 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 FLH, observing:[2]

> 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 (1979: 46).

While Berinstein's hypothesis in (1) also makes predictions for the use of F0, we restrict our investigation to uses of duration. Her investigation focuses on uses of duration, and she further observes (note 1) that F0 may be used in different ways within the same language: 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 work, various languages have been reported to conform to the predictions of the FLH. Other languages with contrastive

[1] The hierarchical ordering of F0 > duration > intensity in Berinstein 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).

[2] Berinstein at this point separates her proposal into two component parts, the functional load hypothesis and the related hierarchical cue reordering.

vowel length which eschew duration as a cue to stress include Anejom̃ (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 1965), Leti (Hume *et al*. 1997, Hume 1997), Masbatenyo (Rosero 2014), Nhanda (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 have been made on a language-by-language basis. For example, Taff *et al*. (2001) specifically address the consequences of Aleut, which has contrastive vowel length and uses duration to cue stress, for the FLH (see also Rozelle 1997). The authors propose that, while still correlating with stress, duration may perhaps 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 Typology Lab, whose website lists 15 languages that are being investigated with standardised methodology that separates word-level stress from sentence-level focus.[3] 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 (see also 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 to understand more broadly what types of cues remain distinct.

It is worthwhile, however, to assess the correlation between vowel length and stress-related duration on the basis of the current state of reported stress correlates across a wide variety of languages.[4] To this end, we have looked for evidence for Berinstein's FLH by compiling a database of languages for which we could find information on stress correlates. This

[3] Available (August 2017) at sites.google.com/site/udstresslab/home.

[4] 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' not to refer to phonetic cues to stress, but to relations between the status of phonological vowel length and the phonetic stress correlates.

database, which contains 140 languages, is introduced in the following section, and the correlations between uses of duration are examined in §3.

2 The Stress Correlate Database

The Stress Correlate Database (Lunden & Kalivoda 2017) is a compilation of reported stress correlates among the world's languages.[5] A main source for the database was grammars (books, and 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 (2). We then discuss the coded levels for each, and how they were determined.

(2) *Database parameters*

a. *Language information*
Language name
Primary family
Secondary family
ISO 639-3 code

b. *Stress system information*
Stress system type
(single stress, binary stress, dual stress, ternary stress)
Primary stress edge (left/right)

c. *Stress correlate information*
Duration as reported correlate (yes/no)
Pitch as reported correlate (yes/no)
Intensity/loudness/energy/spectral tilt as reported correlate (yes/no)
Primary correlate, if known
Vowel reduction (phonological) in unstressed syllables (yes/no)[6]

d. *Phonological uses of suprasegmentals*
Duration used contrastively outside of stress (yes/no)
Pitch used contrastively outside of stress (yes/no)[7]

e. *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

[5] The database is available as online supplementary materials at https://doi.org/10.1017/S0952675717000288.

[6] Vowel reduction is included for completeness of stress correlates. As it is not a suprasegmental property, it is not considered further here.

[7] 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 use F0 contrastively.

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 in the source, with secondary stress occurring either two syllables after (if the primary stress was leftward) or before (if the primary stress was rightward), the language is assumed to be binary unless other information was provided. Stress systems are coded as dual if the primary and secondary stresses appeared to come from opposite directions and no other (non-weight-sensitive) stresses were mentioned; e.g. if the primary stress was 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 on the basis of the stress pattern in words with all light syllables, unless otherwise stated. Whether stress is leftward or rightward is also coded. If no stress window was listed, lexical stress is not coded as either leftward or rightward.

The reported stress correlates were taken from descriptions of the realisation of stress. Researchers often only discussed the correlates used by the language; in such cases, 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 was specifically mentioned as not known. In these cases the other correlates are left blank, rather than coded as not present.

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 which 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* column. Finally, because sources did not consistently clarify whether correlates were for primary or secondary stress, the database conflates the two.[8] If a correlate was mentioned as being present in the realisation 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

[8] This was done because (i) the descriptions of only six languages listed correlates for primary and secondary stresses separately and (ii) the FLH is not specific to a particular level of stress.

1907); however, if that feature clearly followed a culminative pattern that was present in every word, we assumed that the author was referring to stress and coded the term the author used as a correlate.[9]

Some further points 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 were 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.[10]

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 not mentioned in the discussion of stress correlates. In most cases we had to look further (within a source or at a different source), in order to determine whether there was any evidence for or against vowel reduction. While vowel reduction occurs on a continuum, we only considered phonological vowel reduction, which we assumed would reliably be reflected in IPA transcriptions, whereas vowel undershoot would not be. Phonological vowel reduction typically involves a reduction of the number of vowel contrasts, although not necessarily; e.g. Bhojpuri phonologically reduces all vowels, although they remain distinct (Shukla 1981). Phonetic vowel undershoot (in the sense of Lindblom 1963 and Stevens & House 1963) and other phonetic reduction were not considered to be 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,

[9] Culminativity is classically considered a property of stress systems which is not present in tonal systems (e.g. Hyman 1977, Hayes 1995).

[10] The exception may be spectral tilt, which is an acoustic measurement of the intensity distribution across frequencies. Sluijter & 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'.

mention of vowel reduction was absent. The observations in the sources that prompted the determinations are given as quotations in comments in the cells of the *vowel reduction* column.

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 than the importance of the others.

Whether length or pitch was 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.), an additional source was consulted in order to ascertain whether vowel length was 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; quotations 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 to stress that could be utilised by listeners, who would then not need to attend to the actual vowel length.[11]

2.1 Testing the database

Some of the reported correlates are the result of acoustic investigation on a 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 realisation 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 with 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

[11] There were four such languages, including English, which Berinstein considers not to have a length contrast. See note 13 for the adjustment to Table III if these languages were coded as having contrastive vowel length.

language sets of the same size taken 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 compared this to the average for each correlate in 18 languages randomly pulled from the database. On average, over 10,000 random sets of 18 languages, 59% used duration, 66% pitch and 61% intensity. Thus we can see that the Austronesian languages are more likely than any average group of 18 languages to use duration and pitch, but just as likely to use intensity.

Each value (from each of 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 by tending either to use or eschew a particular stress correlate. The 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.[12] 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 observes that increased duration, on the other hand, cannot be seen as a 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 as well, but that this should not hold for the co-occurrence of either with

[12] 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).

duration. Table Ia 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 Ib is restricted to only those languages reported to have exactly two of the three suprasegmental correlates.

(a)

pitch and intensity	59
pitch but not intensity	34
intensity but not pitch	25

(b)

pitch and intensity	33
pitch and duration	20
intensity and duration	15

Table I

Correlation of use of pitch and intensity: (a) for all languages; (b) for languages with exactly two correlates.

We can see from these 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 χ^2 tests of equal proportions for the data in Table I 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 FLH.

3 Correlations

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

		stress correlate		
		yes	*no*	*total*
contrastive vowel length	*yes*	45	37	82
	no	38	20	58
	total	83	57	140

Table II

Correlations of stress-based duration and contrastive vowel length.

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 χ^2 statistical test shows there is not a significant relationship between duration as a stress correlate and contrastive vowel length ($\varphi = -0.107$; $\chi^2(1) = 1.593$, $p = 0.207$).

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, 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 source clearly specified otherwise. Table III is a revised version of Table II, with the durational correlate limited 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.

		stress correlate		
		yes	*no*	*total*
contrastive vowel length	*yes*	40	42	82
	no	34	24	58
	total	74	66	140

Table III
Correlations of stress-based vowel duration and contrastive vowel length.

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, not very different from the 58.6% of the languages without a vowel contrast that use (vowel) duration as a cue to stress. A χ^2 test finds no relationship

between the use of vowel duration to cue stress and the presence of contrastive vowel length ($\varphi = -0.097$; $\chi^2(1) = 1.320$, $p = 0.251$).[13]

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 above states that duration will have 'little or no [stress] cue value' in languages that use duration contrastively, her original hypothesis 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.

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 IV shows, in the first column, the primary correlate in the 63 languages for which one was identified, and, in the subsequent columns, the number of those that also use vowel duration, pitch or intensity as a correlate.[14]

		primary stress correlate	also vowel duration	also pitch	also intensity
primary correlate	vowel duration	18	*n/a*	4	1
	pitch	31	8	*n/a*	12
	intensity	14	3	3	*n/a*

Table IV

Languages for which a primary correlate is specified: primary and other correlates.

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

[13] The results would 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) were coded as having contrastive vowel length ($\varphi = -0.072$; $\chi^2(1) = 0.730$, $p = 0.393$).

[14] Of the nine languages for which consonant duration is sometimes or consistently the durational cue to stress, six are reported to have a primary correlate of duration. Those six are 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 IV, but are not included in the column of those which also use vowel duration.

		primary stress correlate		
		yes	*no*	*total*
contrastive vowel length	*yes*	10	8	18
	no	8	3	11
	total	18	11	29

Table V

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

As seen in Table V, 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 the primary stress correlate, compared to 72.7% of the eleven 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 χ^2 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 ($\varphi = -0.172$, $p = 0.449$).

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, and the correlation within this subset of vowel-based duration as a cue to stress with contrastive vowel length is given in Table VI.[15]

		stress correlate		
		yes	*no*	*total*
contrastive vowel length	*yes*	17	8	25
	no	16	5	21
	total	33	13	46

Table VI

Correlations of stress-based vowel duration and contrastive vowel length in acoustically investigated languages.

[15] Since this division was not originally coded, we identified this subset based on whether the stress cue quotes referenced measurements. In cases that possibly, but not clearly, included acoustic evidence, we went back to the source. We did not, however, revisit every source.

We see a very similar distribution to that found in Table III for the full dataset. Of the 25 languages with contrastive vowel length, 68% 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 χ^2 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 ($\varphi = -0.091$; $\chi^2(1) = 0.378$, $p = 0.539$).

4 Conclusion

The functional load hypothesis is intuitively appealing, and, if correct, would provide a natural explanation for the stress-related suprasegmental differences between K'ekchi and Kaqchikel. It is not unusual to find Berinstein's work referenced as demonstrating that languages with contrastive vowel length do not use duration in the realisation 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 realisation of stress and its 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.

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 long vowels in Aleut 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 coexist alongside contrastive length, without obscuring these contrasts. Further, as a reviewer notes, there are multiple additional sources of durational differences, from final lengthening, syllable shape, effects from neighbouring consonants, etc., and so it perhaps should not be surprising to find that duration can be used both for contrastive length in vowels and to cue stress.

Others have also noted that Berinstein'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 emphasise 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. Furthermore, Vogel *et al*. 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 Berinstein's work actually investigated sentence accent. It may therefore be that there is a level at which Berinstein'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 will be helpful not only in showing that languages may use duration both in vocalic contrasts and in the domain of stress, but also in serving as a useful reference for identifying languages for which more detailed acoustic investigations should be carried out to clarify the use of duration at different levels.

REFERENCES

Beckman, Mary E. & Jan Edwards (1994). Articulatory evidence for differentiating stress categories. In Patricia A. Keating (ed.) *Phonological structure and phonetic form: papers in laboratory phonology III*. Cambridge: Cambridge University Press. 7–33.

Berinstein, Ava E. (1979). A cross-linguistic study on the perception and production of stress. *UCLA Working Papers in Phonetics* **47**. 1–59.

Blevins, Juliette (2001). *Nhanda: an aboriginal language of Western Australia*. Honolulu: University of Hawai'i Press.

Dillmann, August (1907). *Ethiopic grammar*. 2nd edn, enlarged and improved by Carl Bezold. Translated by James A. Crichton. London: Williams & Norgate.

Dixon, R. M. W. (1988). *A grammar of Boumaa Fijian*. Chicago & London: 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*. Berkeley: University of California Press.

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *JASA* **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. *IJAL* **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: Cambridge University Press. 83–118.

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

Hume, Elizabeth (1997). Vowel preservation in Leti. *Oceanic Linguistics* **31**. 65–101.

Hume, Elizabeth, Jennifer Muller & Aone van Engelhoven (1997). Non-moraic geminates in Leti. *Phonology* **14**. 371–402.

Hyman, Larry M. (1977). On the nature of linguistic stress. In Larry M. Hyman (ed.) *Studies in stress and accent*. Los Angeles: Department of Linguistics, University of Southern California. 37–82.

Kroeker, Menno (2001). A descriptive grammar of Nambikuara. *IJAL* **67**. 1–87.
Lehiste, Ilse (1970). *Suprasegmentals*. Cambridge, Mass.: 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. *JASA* **35**. 1773–1781.
Łukaszewicz, Beata (2015). Polish rhythmic stress revisited: phonetic evidence of an iterative system. Poster presented at *Phonetics and Phonology in Europe*, Cambridge.
Lunden, Anya & Nick Kalivoda (2017). Stress correlate database. Available (August 2017) at http://wmpeople.wm.edu/site/page/sllund/stresscorrelatedatabase.
Lynch, John (2000). *A grammar of Anejom̃*. Canberra: Australian National University.
Martin, Jack B. (2011). *A grammar of Creek* (*Muskogee*). Lincoln & London: University of Nebraska Press.
Pierrehumbert, Janet B. (1980). *The phonology and phonetics of English intonation*. PhD dissertation, MIT.
Rich, F. (1963). Arabela phonemes and high-level phonology. In *Studies in Peruvian Indian languages*. Vol. 1. Norman: Summer Institute of Linguistics. 193–206.
Rosero, Michael Wilson I. (2014). *A grammatical sketch of Masbatenyo*. BA 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 (1965). *A grammar of Diola-Fogny: a language spoken in the Basse-Casamance region of Senegal*. London: Cambridge University Press.
Shukla, Shaligram (1981). *Bhojpuri grammar*. Washington, D.C.: Georgetown University Press.
Sluijter, Agaath M. C. & Vincent J. van Heuven (1996). Spectral balance as an acoustic correlate of linguistic stress. *JASA* **100**. 2471–2485.
Snapp, Allen, John Anderson & Joy Anderson (1982). Northern Paiute. In Ronald W. Langacker (ed.) *Studies in Uto-Aztecan grammar*. Vol. 3: *Uto-Aztecan grammatical sketches*. Dallas: Summer Institute of Linguistics & University of Texas at Arlington. 1–92.
Stevens, Kenneth N. & Arthur S. 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. *JPh* **31**. 113–138.
Suttles, Wayne (2004). *Musqueam reference grammar*. Vancouver: University of British Columbia Press.
Taff, Alice, Lorna Rozelle, Taehong Cho, Peter Ladefoged, Moses Dirks & Jacob Wegelin (2001). Phonetic structures of Aleut. *JPh* **29**. 231–271.
Thornes, Timothy J. (2003). *A Northern Paiute grammar with texts*. PhD dissertation, University of Oregon.
Titze, Ingo R. (1989). On the relation between subglottal pressure and fundamental frequency in phonation. *JASA* **85**. 901–906.
Turk, Alice (2012). The temporal implementation of prosodic structure. In Abigail C. Cohn, Cécile Fougeron & Marie K. Huffman (eds.) *The Oxford handbook of laboratory phonology*. Oxford: 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*. Vol. 14. Amsterdam & Philadelphia: 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: 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 *Proceedings of Speech Prosody 2008*. 363–366. Available (August 2017) at http://isle.illinois.edu/sprosig/sp2008/papers/id120.pdf.