Acoustic correlates of word stress: A cross-linguistic survey

Matthew K. Gordon, University of California, Santa Barbara
Timo B. Roettger, University of Cologne

1. Introduction

The study of the acoustic correlates of word stress has been a fruitful area of phonetic research since the seminal research on American English by Fry (1955, 1958) over 50 years ago. This paper presents results of a cross-linguistic survey designed to distill a clearer picture of the relative robustness of different acoustic exponents of word stress. The present paper will not attempt to address the complex issue of situating word stress within the broader taxonomy of prosodic systems (see Beckman 1986, Hyman 2006, 2014 inter alia). Rather, we assume word stress (or simply ‘stress’) to be the phonological marking of one or more prominent syllables within the phonological word. In practice, for many of the languages surveyed in this paper, the classification of the prosodic system is not conclusive. In order to be as inclusive as possible, studies of languages whose prosodic systems are open to alternative interpretations were included in the present study. Although future consensus might suggest that these languages are better classified as lacking stress, their inclusion in the present study at least allows for contextualizing their phonetic properties relative to the broader literature on acoustic correlates of prominence.

2. Methodology

Several different primary sources were consulted, including a number of phonetics and areal studies journals, working papers volumes and books and dissertations. The corpus (in form of a table) is publically available online at https://osf.io/9r2cd/ alongside a script to reproduce respective counts presented in this manuscript. To establish a reliable and informative corpus that can be used in the future, cited authors are encouraged to submit corrections, if we have interpreted respective aspects of their method and/or results incorrectly. Further, we would like to invite scholars that have published work on word stress that is not logged in the present corpus to share their results with us for inclusion in the database.
Although the database was intended to be as comprehensive as possible, many works that dealt with stress were excluded from the present study on various methodological grounds. First, papers in which methodological description was too sparse or vague to allow for replication were excluded. Likewise excluded were studies that did not present quantitative results. Also omitted were papers not explicitly focused on stress. Papers on stress were included, however, even if experimental design created confounds that could render definitive interpretation of results impossible. For example, several studies were based on words uttered in isolation where word-level stress is conflated with phrase-level prominence, while many others employed carrier phrases in which the target word was (either likely or explicitly) focused, thereby creating a potential confound between phrase-level prominence and word-level stress (see Roettger and Gordon this issue). Finally, we included only studies on populations consisting of adult speakers without reported speech impairments.

The corpus encompassed a total of 110 (sub-)studies on 75 languages or language varieties, e.g. Jordanian and Tunisian Arabic, American and British English. Languages in the survey are plotted in figure 1 and listed in table 1 along with their genetic affiliation according to the 19th edition of the Ethnologue (Lewis et al. 2016).

---

1 The survey conflates as a single variety, non-Javanese Indonesian, the results of the Adisasmito-Smith and Cohn (1996) study of Indonesian based on the speech of a non-Javanese substrate speaker and the results for the Toba Batak substrate speaker of Indonesian in the Goedemans and van Zanten (2007) study (which also includes results for a Javanese substrate speaker).
Figure 1: Geographical distribution of languages included in the survey of acoustic correlates of stress plotted via the “lingtypology” package (Moroz 2017) for R (2017).

Table 1. Languages included in the survey of acoustic correlates of stress

<table>
<thead>
<tr>
<th>Language</th>
<th>Genetic Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aleut</td>
<td>Eskimo-Aleut</td>
</tr>
<tr>
<td>Apache, Jicarilla</td>
<td>Na Dene</td>
</tr>
<tr>
<td>Apache, San Carlos</td>
<td>Na Dene</td>
</tr>
<tr>
<td>Arabic, Jordanian</td>
<td>Afro-Asiatic</td>
</tr>
<tr>
<td>Arabic, Tunisian</td>
<td>Afro-Asiatic</td>
</tr>
<tr>
<td>Basque, Goizueta</td>
<td>Isolate</td>
</tr>
<tr>
<td>Belarusian</td>
<td>Indo-European</td>
</tr>
<tr>
<td>Besemah</td>
<td>Austronesian</td>
</tr>
<tr>
<td>Bininj Gun-wok</td>
<td>Australian</td>
</tr>
<tr>
<td>Bulgarian</td>
<td>Indo-European</td>
</tr>
<tr>
<td>Catalan</td>
<td>Indo-European</td>
</tr>
<tr>
<td>Chabacano, Cavite</td>
<td>Creole</td>
</tr>
<tr>
<td>Chickasaw</td>
<td>Muskogean</td>
</tr>
<tr>
<td>Language</td>
<td>Family</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Chuvash</td>
<td>Turkic</td>
</tr>
<tr>
<td>Czech</td>
<td>Indo-European</td>
</tr>
<tr>
<td>Dalabon</td>
<td>Australian</td>
</tr>
<tr>
<td>Dutch</td>
<td>Indo-European</td>
</tr>
<tr>
<td>Émérillon</td>
<td>Tupi-Guarani</td>
</tr>
<tr>
<td>English, American</td>
<td>Indo-European</td>
</tr>
<tr>
<td>English, British</td>
<td>Indo-European</td>
</tr>
<tr>
<td>Estonian</td>
<td>Uralic</td>
</tr>
<tr>
<td>Finnish</td>
<td>Uralic</td>
</tr>
<tr>
<td>Finnish, Ingrian</td>
<td>Uralic</td>
</tr>
<tr>
<td>German</td>
<td>Indo-European</td>
</tr>
<tr>
<td>Greek</td>
<td>Indo-European</td>
</tr>
<tr>
<td>Hebrew</td>
<td>Afro-Asiatic</td>
</tr>
<tr>
<td>Hungarian</td>
<td>Uralic</td>
</tr>
<tr>
<td>Indonesian, Javanese</td>
<td>Austronesian</td>
</tr>
<tr>
<td>Indonesian, non-Javanese</td>
<td>Austronesian</td>
</tr>
<tr>
<td>Italian</td>
<td>Indo-European</td>
</tr>
<tr>
<td>K’ekchi</td>
<td>Mayan</td>
</tr>
<tr>
<td>Kabardian</td>
<td>North Caucasian</td>
</tr>
<tr>
<td>Kuot</td>
<td>isolate</td>
</tr>
<tr>
<td>Lakhota</td>
<td>Siouan</td>
</tr>
<tr>
<td>Latvian</td>
<td>Indo-European</td>
</tr>
<tr>
<td>Lithuanian</td>
<td>Indo-European</td>
</tr>
<tr>
<td>Livonian</td>
<td>Uralic</td>
</tr>
<tr>
<td>Ma’ya</td>
<td>Austronesian</td>
</tr>
<tr>
<td>Macedonian</td>
<td>Indo-European</td>
</tr>
<tr>
<td>Meadow Mari</td>
<td>Uralic</td>
</tr>
<tr>
<td>Mongolian</td>
<td>Mongolic</td>
</tr>
<tr>
<td>Mordvin, Ezrya</td>
<td>Uralic</td>
</tr>
<tr>
<td>Mordvin, Moksha</td>
<td>Uralic</td>
</tr>
<tr>
<td>Nahuatl, Balsas</td>
<td>Uto-Aztecan</td>
</tr>
<tr>
<td>Paiwan</td>
<td>Austronesian</td>
</tr>
<tr>
<td>Papiamentu</td>
<td>Creole</td>
</tr>
<tr>
<td>Persian</td>
<td>Indo-European</td>
</tr>
<tr>
<td>Pirahã</td>
<td>Mura</td>
</tr>
<tr>
<td>Pitjantjatjara</td>
<td>Australian</td>
</tr>
<tr>
<td>Polish</td>
<td>Indo-European</td>
</tr>
<tr>
<td>Portuguese, Brazilian</td>
<td>Indo-European</td>
</tr>
<tr>
<td>Quechua, Conchucos</td>
<td>Quechua</td>
</tr>
<tr>
<td>Saisiyat</td>
<td>Austronesian</td>
</tr>
<tr>
<td>Language</td>
<td>Family</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>Savosavo</td>
<td>Central Solomons</td>
</tr>
<tr>
<td>Sekani</td>
<td>Na Dene</td>
</tr>
<tr>
<td>Sindhi</td>
<td>Indo-European</td>
</tr>
<tr>
<td>Spanish</td>
<td>Indo-European</td>
</tr>
<tr>
<td>Squamish</td>
<td>Salish</td>
</tr>
<tr>
<td>St'át'imcets</td>
<td>Salish</td>
</tr>
<tr>
<td>Swedish</td>
<td>Indo-European</td>
</tr>
<tr>
<td>Tagalog</td>
<td>Austronesian</td>
</tr>
<tr>
<td>Tamil</td>
<td>Dravidian</td>
</tr>
<tr>
<td>Tanana, Minto</td>
<td>Na Dene</td>
</tr>
<tr>
<td>Tanana, Salcha</td>
<td>Na Dene</td>
</tr>
<tr>
<td>Tarahumara</td>
<td>Uto-Aztecan</td>
</tr>
<tr>
<td>Tashlhiyt</td>
<td>Afro-Asiatic</td>
</tr>
<tr>
<td>Thai</td>
<td>Tai-Kadai</td>
</tr>
<tr>
<td>Tongan</td>
<td>Austronesian</td>
</tr>
<tr>
<td>Turkish</td>
<td>Turkic</td>
</tr>
<tr>
<td>Urdu</td>
<td>Indo-European</td>
</tr>
<tr>
<td>Ute, Southern</td>
<td>Uto-Aztecan</td>
</tr>
<tr>
<td>Uyghur</td>
<td>Turkic</td>
</tr>
<tr>
<td>Welsh</td>
<td>Indo-European</td>
</tr>
<tr>
<td>Witsuwi't'en, Babine</td>
<td>Na Dene</td>
</tr>
<tr>
<td>Yakima Sahaptin</td>
<td>Sahaptian</td>
</tr>
</tbody>
</table>

For each of the studies (and sub-studies within a single work) that satisfied the criteria for inclusion in the survey, a series of pieces of information was logged, including the name of the language, whether the language is tonal (which includes languages often regarded as having lexical “pitch accent” rather than canonical tone), the word stress levels examined (primary stress (1S), secondary stress (2S) and unstressed (US)), the acoustic parameters used to express word stress, as well as other methodological aspects (see Roettger & Gordon, this issue, for discussion).
3. Acoustic correlates of stress

Studies in the database differed in the acoustic dimension(s) investigated. These can be coarsely broken down into four categories: duration, fundamental frequency, intensity, and spectral characteristics.

In most studies, duration values were taken of only the vowel. Also attested were measures of the syllable rime (labeled “R” in the database), the nucleus (labeled “N”) if the nucleus could be a consonant, the entire syllable (labeled “syll”), and consonant durations, typically of the constriction for syllable onsets (labeled “O”) and, more rarely, of voice-onset-time (labeled “VOT”) values for onsets or duration values for syllable coda consonants (labeled “C”). Most duration measurements were absolute measures calculated over a given domain, although some studies employed measures relative to another adjacent segment.

The most common fundamental frequency measurement (unlabeled in the corpus) was the mean for the vowel. Other fundamental frequency measurements (typically from the vowel) included peak F0, F0 at vowel midpoint or at the intensity peak, variability of F0 (calculated as F0 standard deviation) as well as time varying characteristics such as F0 slope, or values taken at regular intervals of either a fixed length or relative portions of the entire span of an interval.

The most frequent measure of intensity in the database was also the mean (usually calculated over the vowel), sometimes taken as a relative measure between stressed and unstressed syllables, which helps to mitigate fluctuations in intensity attributed to differences in the distance between the mouth and the microphone (if not worn on the head). Less common were measurements of peak intensity and intensity at the midpoint of the vowel and the intensity integral, the overall intensity aggregated over the entire duration of the target. This integration of intensity over time captures the increased perceptual loudness of a longer stimulus relative to a shorter one, at least over relatively short durations characteristic of vowels (see Moore 2013).

The final macro-category of measurements comprised various spectral measures. The most common frequency-sensitive measure consisted of formant values, typically for the first two formants. The other type of spectral measure observed in the database reflects the tendency for
stressed vowels to display relatively less attenuation of energy at higher frequencies relative to unstressed values. Measurements of spectral tilt were quantified in various ways depending on the study, including frequency-bounded intensity bands, relative amplitude of the first and second harmonic (H1-H2), amplitude values of harmonics proximal to formants, slope of intensity declination as a function of frequency, and frequency-adjusted loudness scales such as the phon.

The stacked bar plot in figure 2 graphically depicts both the number of (sub-) studies (out of a total of 110) for which at least one study identified each acoustic parameter as a marker of stress (dark bars) vs. the number of studies for which a given parameter was examined but found not to signal stress (grey bars). A parameter is identified as a successful marker of stress if it distinguishes at least two levels of stress, i.e. primary stressed vs. unstressed, primary stress vs. secondary stress, or secondary stress vs. unstressed. The two frequency-sensitive measures, formant frequencies and spectral tilt, are separated due to the inclusion of data on both in many studies.

Figure 2. Number of (sub-) studies for which various acoustic correlates of stress were successful (black bars), unsuccessful (grey bars), and not measured (white bars) in differentiating stress level.
3.1. Duration

Duration was by far the most frequently measured property in the database (72 of 75 languages), and also the most successful marker of stress, distinguishing stress in 90% (65 of 72 languages) of the languages studied, all except Hungarian, Tunisian Arabic, Javanese, K’ekchi, Mordvin and the Javanese-substrate variety of Indonesian discussed in the Goedemans and van Zanten (2007) study.

Interestingly, for several languages, only consonant and not vowel duration successfully distinguished stress level. Thus, in Estonian (Lehiste 1966, Gordon 1995) and Peninsular Spanish Ortega-Llebaria (2006) onsets were lengthened in stressed syllables, while in Welsh (Williams 1998) stressed codas were durationally enhanced. Vowels in Lakhota (Cho 2006) were also not lengthened under stress, although VOT values for aspirated stops were greater in onset position of stressed syllables relative to their unstressed counterparts. Vowels in Yakima Sahaptin (Hargus 2005) were also not durationally distinct as a function of stress. However, contrary to the Lakhota results, VOT values in Yakima Sahaptin were shorter for stops in the onset of stressed syllables. The divergence between Lakhota and Yakima Sahaptin can be understood in terms of contrast enhancement: the phonemically aspirated stops of Lakhota are enhanced by lengthening VOT, whereas the unaspirated stops of Yakima Sahaptin are enhanced by shortening VOT values.

Finally, there are a few further studies in which an overall lengthening effect on either the syllable rime (Bond 1991 on Latvia, Chian and Chiang 2005 on Saisiyat) or the entire syllable (Lehiste et al. 2005 on Meadow Mari, Sadeghi 2011 on Persian) emerged under stress, leaving the segmental source of the additional length unclear.

3.2. Fundamental Frequency

Looking at F0 to examine word stress is notoriously difficult due to the common co-occurrence of word level prominence and post-lexical tonal events such as pitch accents (e.g. Bolinger 1958, 1961, Huss 1978, Beckman 1986, Ladd 2008 inter alia). As discussed in Roettger & Gordon (this issue), many studies in our corpus do not allow for teasing these levels apart. The following discussion about F0 as a marker of word stress should thus be interpreted with caution.
F0 measures marked alleged word stress in 73% (46 of 63) of the languages for which it was targeted for investigation in at least one study. The success rate of F0 as a correlate of stress becomes even higher if one excludes the five tone languages that fail to use F0 to mark stress.

Even if the languages for which studies demonstrated F0 to be a signal of stress but which relied on isolation forms are excluded, this still leaves a strong majority of languages in the database that used F0 to distinguish stress level. Most studies in which F0 was used to differentiate stress employed a static measure, typically the mean, but in some languages, only a dynamic and not a static measure of F0 was diagnostic of stress, e.g. Estonian (Liiv 1985, Gordon 1995), Thai (Potisuk et al. 1996), and Italian (Eriksson et al. 2016).

Of the nine tone languages in the database in which F0 was examined, it only was reliably used to cue stress in two, both of which lack the canonical profile of a tone language. In Goizueta Basque (Hualde et al. 2008), a language in which tone is limited to certain lexical items (a property characteristic of traditional “pitch accent” languages), F0 distinguishes stress level only in words lacking lexically-specified tone. In Balsas Nahuatl (Guion et al. 2010), F0 has been retained as a diagnostic of stress even in dialects that have developed incipient tone distinctions while still retaining vestiges of the original penultimate stress system.

In most languages in which fundamental frequency diagnosed stress, F0 values were greater in stressed than unstressed syllables, although there were a pair of studies in which lowered F0 was symptomatic of stress: the speakers from Lahore (but not the one from Karachi) in Hussain’s (1997) study of Urdu and the isolation words (but not those in context) in Eriksson’s et al.’s (2016) research on Italian.

Although many of the F0 effects observed in the database could be attributed to post-lexical prominence (see Roettger & Gordon, this issue, for a discussion), certain languages in the database still display an effect of stress on F0 when these factors are apparently controlled for by placing the target word in an utterance in which another word is explicitly focused, e.g. Finnish (Suomi et al. 2001), Greek, Hungarian, and Peninsular Spanish (Vogel et al. 2016). On the other hand, in a third tone language, Minto Tanana (Tuttle 1998), F0 has a marginal status as a stress correlate, only used to differentiate stress for short but not long vowels.
hand, the possibility that target words are still associated with a phrasal accent cannot be
definitely excluded even in cases where another constituent is explicitly focused. Vogel et al.
(2016:134) allude to this possibility any time a metalinguistic carrier phrase in which the target
word is systematically varied while the rest of the phrase is held constant, as in their study

3.3. Overall intensity

Non-frequency-dependent measures of intensity (e.g. mean, peak, midpoint) had similar success
to F0 in their capacity to diagnose stress, functioning as a marker of stress in 75% (39 of 52) of
languages. In three studies encompassing two languages, Dobrovolsky (1999) on Chuvash,
Lieberman (1960) and Beckman (1986) on American English, the relevant intensity measure was
the intensity integral, which incorporates duration.

A finding that casts doubt on the efficacy of overall measures of intensity is the observation that
none of the studies that controlled for phrase-level prominence found intensity to be a robust
exponent of stress. Of the four languages (Greek, Hungarian, Spanish, and Turkish) in the Vogel
et al. (2016) study, only Hungarian used mean intensity to distinguish stress in non-focused
target words. Otherwise, only in Papiamentu (Remijsen & van Heuven 2002) was mean intensity
reliably associated with stress in a clearly defocused condition. Notably, though, Papiamentu is a
tone language, in which, as mentioned earlier, F0 is less readily available for conveying stress
distinctions. It is likely no coincidence that six of the seven tone languages in the database for
which an overall measure of intensity was taken (all except Thai) employed intensity as a marker
of stress.

3.4. Frequency-sensitive intensity

Of the 19 languages for which at least one study targeted a frequency-dependent intensity
measure, 16 (84%) used such a measure to differentiate stress levels, where the intensity of
stressed vowels was weighted (in virtually all cases) toward higher frequencies in comparison to
unstressed vowels. The exceptional languages in which spectral tilt was not an exponent of stress
were Pitjantjatjara (Tabain et al. 2014), Peninsular Spanish (Ortega-Llebaria & Prieto 2010), and
Brazilian Portuguese (Barbosa et al. 2013).
Studies differ considerably in how they quantify spectral tilt. Most studies (the unmarked case in the corpus) compare the relative intensity of different frequency bands in the spectrum as an index of stress, where the frequency of these bands varies across studies potentially contributing to differences between studies in results for the same language, e.g. Prieto and Ortega-Llebaria (2006) vs. Ortega-Llebaria and Prieto (2010) on Peninsular Spanish. On the other hand, Hussain (1997), Guion et al. (2010)\(^3\), Garellek and White (2015) and Caballero and Carroll (2015)\(^4\) examine the relative intensity of the first two harmonics (H1-H2), which is typically analyzed as an index of voice quality (Gordon and Ladefoged 2001). Synthesizing the H1-H2 results across these studies suggests a pattern of increased breathiness in unstressed vowels relative to their stressed counterparts.

In summary, although spectral tilt is certainly a promising correlate of stress, the diversity of implementations makes it difficult to definitively establish its reliability relative to other potential markers of stress.

3.5. Formant frequency

The final measure assessed in several studies was formant frequency, most commonly the first (F1) and second (F2) formant, which can be interpreted as indices of centrality along the height dimension in the case of F1, reflecting degree of jaw opening (Erickson and Kawahara 2016) and backness in the case of F2, reflecting tongue dorsum advancement/retraction (Erickson 2002) dimensions. Typically, stressed vowels tend to be more peripheral than unstressed vowels, although there is a contrary effect observed in some languages whereby stressed vowels may be lower in the acoustic space (reflecting a lowered jaw position) than their unstressed counterparts even if this entails a more central articulation, e.g. in the case of high vowels (see Crosswhite 2004 on the typology of stress-related effects on vowel quality).

In interpreting the database results, formant frequency was classed as a reliable correlate of stress in a language if either the first or second formant reliably distinguished any phonemic vowels as a function of stress in one or more studies of the language. In 86% (25 of 29) of languages for

\(^3\) Guion et al. (2010) also analyze H1-A2 (the intensity of the harmonic closest to the second formant).

\(^4\) Garellek and White (2015) also take a measure of cepstral peak prominence (CPP) to assess the degree of periodicity in the signal.
which formant data appeared did vowel quality differ as a function of stress, though it should be mentioned that the database did not include studies of certain languages demonstrated in other work to have stress-induced vowel reduction, e.g. English (Lindblom 1963), Russian (Padgett & Tabain 2005), and Finnish (Wiik 1965).

In many of those languages in which vowel quality differed as a function of stress, the effect was limited to certain vowels and/or only one formant. The formant(s) and vowel qualities differentiated by stress varied from language to language making it difficult to draw any salient cross-linguistic generalizations about the phonetic nature of reduction other than the well-known tendency for stressed vowels to occupy a more peripheral vowel space than their unstressed counterparts. It is also noteworthy that none of the studies that controlled for phrase-level prominence found reliable differences in the first and second formant between stressed and unstressed syllables.

3.6. Relative efficacy of different cues

It is possible in principle to evaluate the relative effectiveness of different acoustic cues not only in aggregate across languages but also in languages that use multiple properties to distinguish stress levels. There are eight studies representing eleven different languages in the database that use statistical analyses, either logistic regression or linear discriminant analysis, to assess the relative capacity of different acoustic dimensions to predict stress level. Crucially, the estimations are based purely on production data and do not imply any perceptual weighting. In six of the eleven languages, an F0 property (either mean or change) was the most reliable predictor of stress level: Berinstein (1979) on K’ekchi, Garellek and White (2015) on Tongan, and Vogel et al. (2016) on Greek, Hungarian, Spanish, and Turkish. In the remaining five languages, duration emerged as the most predictive of stress: Potisuk et al. (1996) on Thai, Sluijter and van Heuven (1996) on Dutch, Remijsen (2001) on Ma’ya, Remijsen and van Heuven (2002) on Papiamentu, and Silber-Varod et al. (2016) on Hebrew. The edge in favor of F0 becomes even greater if one excludes the three tone languages among the eleven, Ma’ya, Thai and Papiamentu, in all of which duration is a better predictor of stress.

The results of the linear discriminant analyses in the Vogel et al. (2016) study of Greek, Hungarian, Spanish, and Turkish demonstrate overall intensity and vowel quality to be relatively
unreliable predictors of stress. It should be noted, however, that only two of the five studies that directly compared cues, Sluijter and van Heuven (1996) on Dutch and Remijnse (2002) on Ma’ya, incorporated a measure of spectral tilt, making it difficult to assess the efficacy of spectral tilt as a marker of stress relative to other acoustic properties.

4. Acoustic evidence for secondary stress

Although most studies in the survey evaluated only the acoustic distinction between primary stressed and unstressed syllables, there were 21 papers that also considered the acoustic evidence for secondary stress, a contentious issue in the stress literature for many languages in which either the presence or absence of stress or its location is contested, e.g. Polish, Estonian (see Hayes 1995 for these and other cases). Perhaps not surprisingly, evidence for secondary stress as distinct from both primary stress and lack of stress was less compelling in the survey than evidence for a distinction between primary stressed syllables and unstressed ones. In most studies, secondary stress was distinguished from other levels using only a subset of properties that were used to distinguish primary stress from lack of stress. Only two studies, Gordon (2004) on Chickasaw and Rietveld et al. (2004) on Dutch, distinguished secondary stressed syllables from both their primary stressed and unstressed counterparts along all the dimensions that differentiated primary stressed and unstressed syllables. Otherwise, secondary stressed syllables were neutralized with either primary stressed or unstressed syllables for at least one parameter that marked the contrast between primary stress and lack of stress. The most tenuous distinction in most cases was between secondary stress and lack of stress. Vowels claimed in the phonological literature to carry secondary stress were not different from unstressed vowels along any dimension in Erzya Mordvin (Lehiste et al. 2003), Pitjantjatjara (Tabain et al. 2014), Polish (Dogil 1999, Newlin-Łukowicz 2012), and Brazilian Portuguese (Barbosa et al. 2013). Similarly, the distinction between secondary stress and lack of stress in German (Kleber & Klipphahn 2006) was only evident for duration for only two (of six) vowel qualities and only for one or two (of six) speakers. In Ingrian Finnish (Gordon 2010), only slight lengthening of voiced onsets emerged as a potential cue to secondary stress as distinct from lack of stress, while F0, intensity, and lengthening of all onsets differentiated primary stressed syllables from unstressed syllables. Garallek and White (2015) find a similar pattern of stronger acoustic evidence for primary stress relative to secondary stress in their study of Tongan: in a linear discriminant analysis, they
observe much higher classification rates for the primary stress vs. unstressed distinction than the secondary stress vs. unstressed difference (89.1% vs. 64.5%).

In summary, the search for secondary stress as a distinct level of prominence proved generally more elusive in the database than the diagnosis of primary stress, a finding that is consistent with the abundance of disputed cases of secondary stress in the phonological literature.

5. Acoustic correlates of stress and prosodic taxonomy

The database provides acoustic evidence for stress in a prosodically diverse set of languages. Evidence for stress emerged for languages with predictable phonological stress, both weight-sensitive stress, e.g. Chickasaw, Squamish, and strictly deliminative stress, e.g. Polish, Finnish, as well as those with robust phonemic stress distinctions, e.g. Russian, Hebrew, and with mixtures of phonemic and predictable stress, e.g. English, Spanish. For a few of these languages generally accepted to have stress, the consulted studies were too preliminary to offer compelling acoustic evidence of stress. For example, small studies of stress in Czech (Dubêda 2006) and Lakhota (Cho 2006) failed to provide definitive corroboration of stress either due to their confinement to a single potential correlate of stress (e.g. in Czech) or their small sample size, e.g. the single speaker examined in the Lakhota study. Presumably, future studies of these languages will provide more convincing evidence of stress.

Evidence for stress also emerged for tone languages, ranging from those with more canonical one-to-one mappings between syllables and tones, e.g. Thai and Pirahã, to those with more limited tone, i.e. pitch accentual, systems, e.g. Basque and Swedish. Not surprisingly, in languages with lexical tone contrasts, F0 typically played a subservient role in signaling stress.

Also included in the database were studies of some languages whose relationship to the tone-stress continuum is less clear. Recent literature has revealed the existence of some languages lacking evidence for either lexical tone or word-level stress. In these “intonation-only” systems, the most salient prosodic events are attributed to the intonation system in the form of phrasal tones realized at or near edges of prosodic phrases potentially consisting of multiple words. Languages fitting this profile of having phrasal prosody rather than word-level stress include both some not appearing in the database, e.g. Korean (Jun 1993) and French (Jun and Fougeron
as well as a few examined in studies considered here. One relevant case is Indonesian, which has traditionally been regarded as a language with word-level stress but whose membership in this prosodic category has more recently been questioned (see Goedemans and van Zanten 2007 for discussion). Goedemans and van Zanten (2007) show that the acoustic correlates of stress in Indonesian, which functions as a lingua franca for speakers with diverse native language backgrounds, diverge sharply based on the substrate language of the speaker. Thus, their speaker of Toba Batak, a language with clearly discernible stress distinctions in the acoustic domain, marks stress in Indonesian along multiple dimensions (duration, F0, and intensity), whereas their speaker of Javanese, another language lacking robust word-level stress, fails to signal stress through any of these acoustic properties. The results for their Toba Batak speaker parallel those for the non-Javanese speakers of Indonesian in the earlier Adisasmito-Smith and Cohn (1996) study, suggesting that Indonesian potentially lacks acoustic evidence for word-level stress independent of transfer effects associated with speakers form other languages with word stress.

Another language in the database that plausibly lacks both tone and word-level stress is Tashlhiyt. When controlling for phrase-level confounds, Roettger et al. (2015, see also Roettger accepted for a detailed analysis) find no evidence for consistent stress on the final syllable contra earlier results from Gordon and Nafi (2012). Yet another language in the survey lacking compelling evidence for word-level stress is Tamil, in which none of the potential acoustic correlates (duration, intensity, and F0) of stress emerged as reliable in Keane’s (2006) study.

In summary, although stress appears to be an acoustically manifested phonological property in both stress languages as well as in tone languages, its universal status remains to be corroborated. Rather, existing evidence suggests the need for a third category of language whose prosody is defined neither in terms of tone nor stress but rather by phrase-level rather than word-level properties.

6. Summary

Results of a survey of 110 studies of 75 languages indicate that a large number of parameters potentially signal stress, including duration (not just of the vowel but also the onset consonant),
various F0 features, overall intensity, assorted frequency-weighted measures of intensity, and vowel formant frequencies. Studies vary considerably in which subset of these potential stress correlates are examined, making it difficult to establish which ones are most consistently cues to stress. Statistically, duration was the most reliable exponent of stress across languages, although all of the measured parameters succeeded in differentiating stress in the majority of languages for which they were assessed. In most studies that investigated secondary stress, it was distinguished from primary stress and/or lack of stress through only a subset of parameters differentiating primary stress from no stress.

This study thus offers a first cross-linguistic assessment of the relative robustness of different potential acoustic exponents of word stress. However, as remarked throughout the manuscript, the findings need to be considered in light of the methodology employed in the studies comprising the survey. Carefully evaluating experimental design choices and statistical analyses of the discussed studies (see Roettger & Gordon, this issue) leads to a more conservative view of what the results can genuinely tell us about the phonetic manifestation of word stress.

References


Pycha, Anne. 2006. A duration-based solution to the problem of stress realization in Turkish. UC Berkeley Phonology Lab Annual Reports.


Roettger, Timo B. and Matthew K. Gordon. This issue. Methodological issues in the study of word stress correlates.


Tuomainen, Jyrki, Stefan Werner, Jean Vroomen, and Beatrice de Gelder. 1999. Fundamental frequency is an important acoustic cue to word boundaries in spoken Finnish. *Proceedings of the 14th International Congress of Phonetic Sciences, San Francisco*, 921-923.


