Sonority and central vowels: A cross-linguistic phonetic study

Matthew Gordon
Edita Ghushchyan
Bradley McDonnell
Daisy Rosenblum
Patricia A. Shaw

Abstract

This paper reports results of a cross-linguistic study of four potential acoustic correlates of vowel sonority. Duration, maximum intensity, acoustic energy, and perceptual energy are measured in five languages (Hindi, Besemah, Armenian, Javanese, and Kʷak'ʷala) in order to determine whether there is an acoustic basis for the position of schwa at the bottom of vocalic sonority scales. The five targeted languages belong to two groups. In three languages (Armenian, Javanese, and Kʷak'ʷala), the reduced phonological sonority of schwa relative to peripheral vowels is manifested in the rejection of stress by schwa. In two languages (Hindi and Besemah), on the other hand, schwa is treated parallel to the peripheral vowels by the stress system. Results indicate that schwa is differentiated from most vowels along one or more of the examined phonetic dimensions in all of the languages surveyed regardless of the phonological patterning of schwa. Languages vary, however, in which parameter(s) is most effective in predicting the low sonority status of schwa. Furthermore, the emergence of isolated contradictions of the sonority scale whereby schwa is acoustically more intense than one or more high vowels suggests that phonological sonority in vowels may not be quantifiable along any single acoustic dimension.

1. Introduction

Sonority refers to the relative prominence of different sounds. Scales based on sonority have proven very useful in characterizing a wide range of phonological phenomena, including syllabification, phonotactic constraints, and stress (see Parker 2011 for an overview of sonority in phonological theory). One of the productive research programs belonging to the study of sonority is the examination of the physical properties defining sonority distinctions (Parker 2002, 2008, 2011). This paper contributes to this research agenda by examining the acoustic basis for one type of sonority distinction that is particularly important in the description of many stress systems: the sonority contrast between central and peripheral vowels. An acoustic study of five languages shows that a number of acoustic properties successfully predict sonority distinctions based on vowel quality, though there is no single parameter that correlates in all five languages with all of the sonority distinctions involving vowels.

Central non-low vowels such as /ə/ and /ɨ/ rank lower on many sonority scales than more peripheral vowel qualities (Parker 2002, 2008). Much of the evidence for their sonority profile is drawn from stress systems in which unstressed vowels reduce to schwa, e.g. in English, and languages in which central vowels reject stress in words containing at least one peripheral vowel (see Kenstowicz 1997 and de Lacy 2002, 2004 for overviews), e.g. Mari (Itkonen 1955, Kenstowicz 1997), Javanese (Herrfurth 1964, Horne 1974), Aljutor (Kodzasov and Muravyova 1978, Kenstowicz 1997). For example, stress in most varieties of Armenian (Vaux 1998) falls on the final syllable (1a) unless this syllable contains schwa (1b).

(1) Armenian stress (examples from Vaux 1998:132)

a. [moˈrukʰ] ‘beard’, [ɑrtɑˈsukʰ] ‘tears’, [jerkrɑkedronɑˈkɑn] ‘geocentric’

b. [ˈmɑnər] ‘small’, [jerˈpʰemən] ‘sometimes’

The sonority distinction between interior and peripheral vowels potentially presents a challenge to the quantification of sonority in terms of acoustic phonetic scales, because intensity, the most reliable acoustic correlate of sonority (Parker 2002, 2008) has been shown to be greater for mid-central vowels like schwa than for high peripheral vowels due to the increased degree of vocal tract aperture associated with central vowels relative to high vowels. Furthermore, the cross-linguistic tendency for lower vowel qualities to be longer than higher vowels might superficially appear to preclude another potential phonetic correlate of sonority, duration, from predicting the reduced phonological sonority of non-low central vowels.

An articulatory-based measure of sonority is potentially informative as a predictor of the phonological status of mid-central vowels like schwa since the tongue position associated with such vowels is closer to its default location in the center of the vocal tract. This holds true of all schwas regardless of whether they are underlying or the result of vowel reduction, epenthesis or excrescence triggered by a consonant (see Silverman 2011 for an overview of the various sources of schwa). Mid-central vowels thus require less movement of the tongue, and presumably less articulatory effort, than their more peripheral counterparts requiring vertical or horizontal movement of the tongue and jaw. Nevertheless, although an articulatory-driven account of the behavior of central vowels is intuitively appealing, it suffers from the notorious difficulty associated with quantifying physical effort. In contrast, measuring sonority along various acoustic dimensions is a far more tractable endeavor even if historically it has been a difficult task to pinpoint a single acoustic property that predicts all sonority distinctions (see Parker 2002, 2008 for discussion of this research program). For this reason, we believe it is worthwhile to exhaustively explore the potential acoustic basis for sonority before appealing to more elusive articulatory-based accounts. The merits of pursuing an acoustically-driven analysis of sonority are further justified by an extensive literature proposing numerous acoustic correlates of sonority (see Parker 2002, 2008 for an overview). Although most of these proposals are not accompanied by supporting acoustic evidence, Parker’s (2002) multi-dimensional phonetic examination of sonority in English and Spanish finds that a measurement of intensity correctly predicts the order of most classes of segments in cross-linguistic sonority scales assuming that other factors such as pitch are held constant. Building on his earlier work, Parker (2008) expands his study to include data from Quechua in addition to Spanish and English. He shows that a function based on a measure of intensity extremes, the peak intensity of vowels and the intensity nadir of consonants, fits closely to established sonority hierarchies. Jany et al (2007) find a similarly close fit between Parker’s intensity-based equation and data from consonants in four additional languages (Mongolian, Hindi, Arabic, and Malayalam), though they employ mean RMS amplitude rather than intensity extremes.

Given the demonstrated success of a measure of acoustic intensity as a correlate of sonority, we adopt the working hypothesis that the sonority of central vowels is predictable on acoustic grounds. Nevertheless, we will also explore the possibility that intensity is not the sole acoustic dimension on which sonority is projected.

2. The sonority of schwa

Vowel sonority adheres to a hierarchy predictable from height and centrality (Kenstowicz 1997, de Lacy 2002, 2004, Gordon 2006) as shown in figure 1.

Low V	Mid V	High V	Mid-central V	High-central V
æ, a	e, o	i, u	ə	ɨ
High sonority				Low sonority

Figure 1. Sonority scale for vowels

Low vowels such as /æ, a/ are the highest sonority vowels, followed by peripheral mid vowels /ɛ, ɔ/, followed by the peripheral high vowels /i, u/, the mid central vowel /ə/, and, at the bottom of the sonority scale, the high central vowel /ɨ/. This hierarchy is deducible from cross-linguistic observation of stress systems (see Kenstowicz 1997, de Lacy 2002, 2004), although most languages do not distinguish all levels of the hierarchy. The Kobon stress system (Davies 1981, Kenstowicz 1997) appears exceptional in exploiting all five distinctions in the hierarchy in figure 1.

Despite the intuitive basis of the vowel sonority scale in figure 1, its phonetic grounding is not entirely transparent and, like consonantal sonority, may not be reducible to a single phonetic parameter (Ohala 1992, Ohala et al 1997). A correlation between vowel height and both duration and intensity is well established (Lehiste 1970), such that lower vowels are longer and more intense than higher vowels. However, a purely height-based correlation between the phonetic properties of duration and intensity and the phonological feature of sonority is insufficient to account for the reduced sonority of the mid central vowel /ə/ relative to the high peripheral vowels /ɪ, ʊ/. This problem is apparent in the classic publications by Gordon Peterson and Ilse Lehiste (Lehiste and Peterson 1959, Peterson and Lehiste 1960), which show that stressed schwa in English, which is found in many British varieties of English including RP, is characterized by greater duration and/or average intensity (RMS amplitude) than many other more peripheral vowels. Intensity, the dimension shown by Parker (2002, 2008) to be the best predictor of sonority, is found by Lehiste and Peterson (1959) to be 5-8 decibels greater (depending on the experimental condition) for schwa than for the high vowels /ɪ, ʊ/.

On the other hand, it is clear that *unstressed* schwa in English is shorter and less intense than its stressed counterparts. Parker (2008) finds that peak decibel levels for schwa are 2.3 dB less than those averaged over the high vowels /i, ɪ, u, ʊ/ in American English. He also finds that barred-ɨ, e.g. in the second syllable of words like *roses*, is a further 3.7dB less intense than schwa in keeping with the lighter status of the high central vowel relative to the mid central vowel in Kobon. It is unclear, however, whether the lower intensity of the central vowels relative to their more peripheral counterparts in Parker's study is due to vowel quality or to stress since the two types of vowels are in virtual complementary distribution in American English.

Gordon (2002, 2006) presents phonetic data comparing schwa and the peripheral vowels /a, i/ of Javanese, a language that treats schwa as light in its stress system. He finds that schwa is much shorter and has less overall perceptual energy than the peripheral vowels. He does not, however, present intensity data independent of his measure of auditory energy, which is a temporal integration of intensity transformed to reflect auditory loudness. It is thus unclear how much of the energy difference in his data is due to the shorter duration of schwa as opposed to its reduced acoustic intensity.

Comparison of the Lehiste and Peterson (1959, 1960), Gordon (2002, 2006), and Parker (2008) studies suggests that schwa may be phonetically quite different between languages in keeping with differences in its phonological status (see Silverman 2011 on the phonological behavior of schwa cross-linguistically). On the one hand, in languages like British English in which it an underlying phoneme that may carry stress, schwa may display properties that are expected of its more peripheral mid vowel counterparts. On the other hand, in languages in which schwa resists stress whether underlying, as in Javanese, or the result of vowel reduction, as in American English, it appears to display phonetic characteristics that make it less prominent than not only the mid but also the high peripheral vowels. It is nevertheless unclear given the varied measurements taken in previous research (average intensity, duration, peak intensity, perceptual energy) exactly how central vowels phonetically differ on a language-specific basis.

The current work seeks to remedy this lacuna in our understanding of the phonetic basis for the sonority of the mid central vowel schwa by examining its characteristics in several languages. Of particular interest is the phonetic sonority of schwa relative to high peripheral vowels, which might be expected to be phonetically less prominent than schwa, but which nevertheless occupy a higher position on phonological sonority scales. The set of studied languages includes those in which schwa behaves as a lower sonority vowel than more peripheral vowels as well as those in which schwa lacks any phonological characteristics suggesting that it is less sonorous than other vowel qualities. We explore various potential phonetic correlates of sonority in both types of languages in order to determine whether there is any universal correlate of sonority that predicts its assumed position at the bottom of the sonority scale for vowels or whether its phonetic characteristics differ across languages according to its phonological sonority. More generally, the present work also provides an opportunity for cross-linguistic investigation of the phonetic correlates of sonority in all vowels, including peripheral vowels of different heights as well as pairs of vowels distinguished on the basis of the tense/lax feature. Cross-linguistic examination of the phonetic basis for the sonority of schwa and other vowels promises to shed light on broad issues related to the mapping between phonetic properties and phonological patterns and the role of language-specificity in the phonetic and phonological domains.

2. Methodology

Five languages were targeted for inclusion in the study. Three of these languages, Armenian, Javanese, and Kʷak'ʷala, possess phonemic schwa that asymmetrically rejects stress in contexts in which more peripheral vowels attract stress. The other two languages, Besemah and Hindi, have a schwa phoneme that does not display any propensity to reject stress. The targeted languages including genetic affiliation (according to the 16th edition of the *Ethnologue*, online at http://www.ethnologue.com) primary places where they are spoken, and sources of data on each are summarized in table 1. Further information concerning each language and the corpus of data analyzed for each is presented in the respective results sections (section 3) for individual languages.

Table 1. Languages targeted in the phonetic study

Language	Genetic affiliation	Primary location	Primary data source(s)
Armenian	Indo-European	Armenia	Vaux (1998)
Besemah	Austronesian	Sumatra, Indonesia	McDonnell (2008)
Hindi	Indo-European	India	Dixit (1963), Kelkar (1968), Ohala (1977, 1999)
Javanese	Austronesian	Java, Indonesia	Clynes and Rudyanto (1995), Horne (1974)
Kʷak'ʷala	Wakashan	British Columbia	Boas (1947), Bach (1975), Wilson (1986), Shaw (2009)

For all five languages, a series of potential phonetic correlates of sonority were measured for the phonemic vowel qualities of the language. All of the target vowels appeared in an interconsonantal environment and had the same level of stress in the target words, which were produced between two and five times (depending on the language) by each speaker in randomized order. In order to control for microprosodic effects, an attempt was made to control for surrounding consonants, in particular voicing, to the extent possible. Further details about the methodology employed for each language appear in the sections devoted to the results for the individual languages. The corpus for each language appears in the appendix.

Measured properties included the following: duration, maximum intensity, first formant values, total acoustic intensity (intensity integrated over time), and total perceptual energy (temporally integrated acoustic intensity submitted to a number of filters designed to model independently known properties of audition). Of these measurements, the first four (first formant, duration, maximum intensity, and average intensity) were calculated using Praat (Boersma and Weenink 2010), while the last two, the intensity and perceptual energy summations, were computed using Cricket, custom software developed at UCSB (Gordon and Nash 2007; downloadable at *http://www.linguistics.ucsb.edu/faculty/gordon/projects.htm*). This software is designed to perform an intensity summation as well as an auditory transform designed to capture the perceptual prominence of a sound rather than its physical intensity (see below for further discussion).

Most of the targeted measurements are straightforward with the exception of auditory energy, which is described in detail below. The beginning and end points of the target vowels were marked using a waveform in conjunction with a time-aligned spectrogram with the second formant onset and offset serving as the demarcation point for the beginning and end of the vowel, respectively. Duration, maximum intensity, and first formant values were collected in Praat using a script. Values for the first formant were calculated over a 25 millisecond window centered on the midpoint of the vowel. Formant values were collected in order to assess whether any differences in the other measured properties might be predictable from phonetic differences in vowel quality that might not emerge in broad phonemic transcriptions used in phonological descriptions.

Total acoustic intensity was calculated in Cricket by sliding an 11.6 millisecond (256 points) window over the entire duration of each target vowel with a new window starting at the end point of the previous one. Within each window, intensity was averaged over the frequency range of 0-10 kHz with a resolution of 86Hz. The average intensity values for all the windows were summed together to yield a total intensity value integrated over time. In case the duration of the vowel was not a multiple of 11.6 ms, the last window factored into the summation was smaller than 11.6 ms.

We turn now to the measure of total auditory energy. The auditory energy values in the current work are based on power spectra calculated using the same sliding 11.6 millisecond (256 points) window used to perform the intensity integration. Within each window, intensity was calculated throughout the frequency range of 0-10 kHz with a resolution of 86Hz. Spectra were computed for successive windows stretching over the entire duration of each target rime. These spectra were submitted to a series of filters representing various processes that take place in the mapping of an acoustic signal to an auditory one. The first two stages in the auditory transform model the bandpass filtering properties of the outer and middle ear. The first filter is an outer ear filter capturing the bandpass filtering characteristics of the pinna and the outer auditory canal (the meatus). The natural resonating frequency of the outer ear is about 2.5 kHz with an approximately 10dB per octave attenuation on either side of 2.5 kHz (Shaw 1974). The lower skirt of this filter becomes flat at 1.25kHz, one octave below 2.5 kHz. The next filter represents the bandpass filter provided by the middle ear, where pressure fluctuations on the eardrum are converted to mechanical energy by the ossicles. The middle ear is a maximally effective transducer of energy at approximately 1.5 kHz, with a 15dB per octave attenuation at frequencies above and below 1.5 kHz (Nedzelnitsky 1980). Because a greater proportion of the measured frequency range falls above the center frequency of 1.5 kHz, the result is a greater relative diminution of energy at higher frequencies. The next step in the auditory transform models the bandpass filtering characteristics of the auditory system (Patterson et al. 1982, Moore and Glasberg 1983). Cricket uses a symmetric filter with a 60 dB per octave attenuation linearly interpolated from the center frequency to the base of the skirts, which increase in breadth as the center frequency increases. The filter was slid over the entire frequency range of each spectrum starting at the highest frequency and working downward calculating the attenuation of each intensity value in the range of frequencies affected by the filter (Bladon and Lindblom 1981). The net response at any frequency was the sum of the responses to the filter as it progresses through the frequency domain. The overall loudness of each spectrum was then calculated by summing the outputs of all the bandpass filters.

The next step involves the modeling of temporal effects in the auditory response as adaptation and recovery functions (e.g. Plomp 1964, Wilson 1970, Viemeister 1980). The adaptation function captures the gradual decline in sensation to a continued stimulus, while the recovery function reflects the boost in auditory response after a reduction in stimulus intensity. In the present model, adaptation and recovery were implemented as follows. First, the total loudness value for the second spectral slice in the rime is compared with the loudness values for the first spectral slice. If the loudness of the second frame exceeds that of the first frame, the difference in loudness between the two frames is multiplied by a recovery factor yielding a value that is added to the loudness value of the second frame to yield an output loudness value for the second frame. If, however, the loudness of the first frame is greater than that of the second frame, the difference in loudness between the two frames is multiplied by an adaptation factor that is subtracted from the loudness value of the second frame to yield an output loudness value for the second frame. The loudness of the third frame is compared with the output loudness value averaged over the previous two frames. This procedure proceeds from left to right throughout the entire duration of the rime by comparing the loudness of a given spectrum with a baseline loudness value reflecting the average of the output loudness values for all the previous spectra.

An adaptation factor of 2 decibels per frame was employed, while a recovery factor of 1 decibel per frame was assumed based on a synthesis of results from a variety of sources (Plomp 1964, Wilson 1970, Viemeister 1980). The final step in the auditory model is a simple summing of the auditory loudness values for each spectrum, thereby yielding a single measure of auditory energy.

Implementation of the auditory model potentially has a number of effects on the acoustic data feeding into it. We describe here some of these effects and their relationship to predictions about vowel sonority, focusing first on the frequency dependencies.

The outer and middle ear filter provide a boost in loudness to sounds characterized primarily by energy in the bottom half of the examined 0-10 kHz frequency range. In particular, frequencies falling between the peak of the middle ear filter at 1.5 kHz and the peak of the outer ear filter at 2.5 kHz receive the greatest boost. This frequency selectivity potentially accounts for the propensity of lower vowel qualities to attract stress in several languages. The first and the second formants for a low central vowel like the prototypical one found in most languages with a single low vowel lie close together near to the 1.5 kHz peak associated with the middle ear filter. For this reason, low vowels are perceived as louder than higher vowel qualities. In contrast, the first formant for higher vowel qualities is much lower than 1.5 kHz and would not benefit from the auditory boost. We also might expect high front vowels to have greater auditory energy than high back vowels due to the location of the second formant for high front vowels in the 2 kHz to 3 kHz range. The damping of acoustic energy associated with increases in frequency, however, potentially offsets any auditory advantage of the more forward articulation in the case of high vowels. The predictions relating vowel backness to auditory energy are thus less clear-cut.

Central vowels like schwa might also be predicted to receive a perceptual boost relative to high back vowels since their second formant values are closer to the 1.5 kHz center frequency of the middle ear filter. The auditory prominence of central vowels is potentially compromised, however, by two properties. First, they are often shorter than other vowels, at least in languages such as Javanese (Gordon 2002, 2006), in which they pattern as low sonority vowels in the stress system. Second, it is conceivable that the acoustic intensity of central vowels is low enough relative to other vowel qualities, again perhaps on a language-specific basis, to offset any perceptual boost they might receive due to their distribution of energy. Comparison of the acoustic measurements of duration and intensity with perceptual energy values in the present work will allow for assessment of the relative contribution of different acoustic parameters to the overall perceptual prominence of vowels. Furthermore, perceptual energy is another potential phonetic correlate of sonority whose efficacy in predicting phonological sonority can be compared with that of acoustic properties.

3. Results

Sections 3.1-3.5 present the results for the five languages targeted in our study, beginning in sections 3.1 and 3.2 with those for languages (Hindi and Besemah) in which schwa does not pattern differently from other vowels with respect to stress. In sections 3.3-3.5, we move on to languages in which schwa tends to reject stress (Armenian, Javanese, and Kʷak'ʷala).

3.1. Hindi

3.1.1. Background

Standard Hindi has a ten vowel system in the native vocabulary that is based partially on length and partially on vowel quality. Most of the peripheral vowels come in pairs, one of which is slightly more peripheral, i.e. tense, as well as longer than the other (lax) member of the pair (Ohala 1999). There are two central vowels, a schwa and a low vowel. The vowel phonemes of Hindi appear in table 2 following the conventions of Ohala (1999) with the exception of replacing /ɑ/ with /a/.

Table 2. Vowels of Hindi

	Front	Central	Back
High	i ɪ		u ʊ
Mid	e ɛ	ə	o ɔ
Low		a	

There is disagreement in the literature about the exact principles governing the location of stress in Hindi (see Ohala 1977 and Hayes 1995 for overview and analysis), though the weight criterion described is relatively consistent across accounts. Most scholars are in agreement that both closed syllables and those containing a tense vowel are treated as heavy, with certain scholars (e.g. Kelkar 1968) pointing to a third superheavy degree of weight assigned to syllables closed by two consonants and to closed syllables containing a tense vowel. The simplest characterization of the primary stress rule is the one adopted by Dixit (1963), as discussed in Ohala (1977). Stress falls on the rightmost non-final heavy syllable (2a) or on the final syllable if the only heavy syllable is final (2b). In words in which all syllables are light, stress fall on the penultimate syllable (2c). Examples are from Ohala (1977).

(2) Hindi stress
  a. [aˈvəʃjək] ‘necessary’, [ˈkimət] ‘price’, [paˈkɪstan] ‘Pakistan’, [ɪnˈsanɪjət] ‘humanity’
  b. [goˈbər] ‘cow dung’, [rəˈsoi] ‘kitchen’, [əmɪˈta] (proper name)
  c. [tɔˈlɪja] ‘towel’[1]

As the examples in (2) show, schwa patterns with other lax vowels in attracting stress when it is followed by a coda consonant and rejecting stress when it is in an open syllable. Other variants of the stress system reported by Dixit (1963) differ in the location of primary stress (see Hayes 1995 for discussion), but crucially for our purposes, schwa patterns together with all non-schwa lax vowels with respect to its ability to attract stress.

[1] In the appendix, Ohala (1977:336) transcribes certain all-light words with antepenultimate stress, although her text description (pg. 330) predicts penultimate stress for them.

3.1.2. Methodology

The target vowels for Hindi all appeared in the first and stressed syllable of a disyllabic word. Each measured vowel appeared in two words. Measured vowels were followed by a sonorant consonant (except for the vowel in the first syllable of ˈwɛʃja ‘prostitute’) and each vowel appeared in two words, one in which the target vowel occurred in a closed syllable and the other in which it was found in an open syllable. The Hindi words were embedded in the middle of the carrier phrase *ham ab* ________ *bolte hain* ‘We now say ________’. Each phrase and thus its embedded target word was read five times by a speaker of standard Hindi and recorded on a SONY DAT recorder in a soundproof booth at a sampling rate of 44.1kHz using a headworn microphone (Shure SM10) before being converted to .wav files in preparation for acoustic analysis. Data from three male speakers were analyzed.

3.1.3. Results

Figure 2 depicts first formant values for the measured Hindi vowels averaged across the three speakers. Bars are ordered from left to right in order of height (from lower to higher vowels) and then frontness (front to back) with schwa on the far right. Individual speaker means along with the number of tokens and standard deviations appear in table 3.

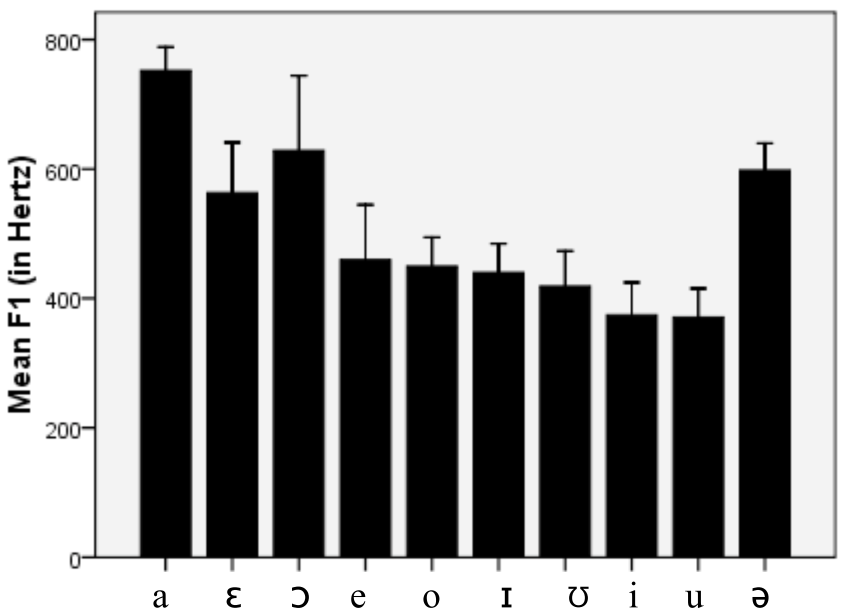


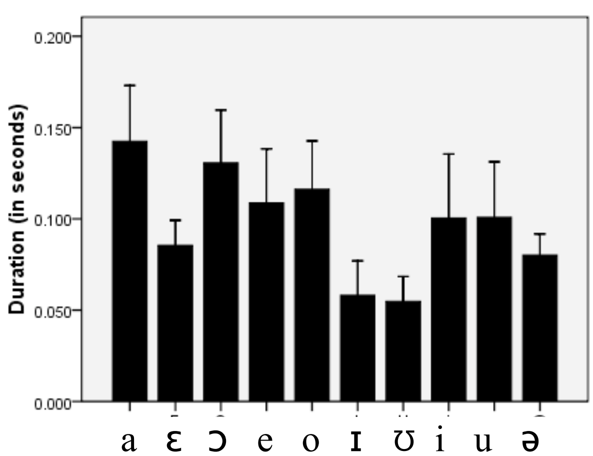
Figure 2. First formant values averaged across five tokens each produced by three Hindi speakers. Whiskers indicate one standard deviation from the mean.

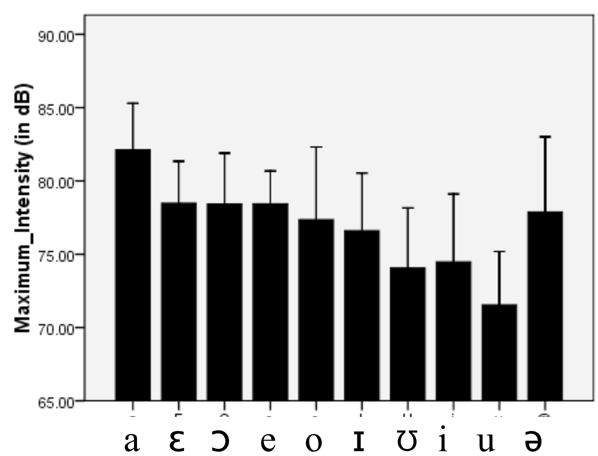
Table 3. Mean first formant values (in Hz) for three Hindi speakers

	Speaker								
	M1			M2			M3		
Vowel	N	Mean	Std.Dev.	N	Mean	Std.Dev.	N	Mean	Std.Dev.
a	10	729	25	10	789	31	10	739	21
ɛ	10	550	106	10	544	39	5	629	28
ɔ	10	547	64	10	620	142	10	718	51
e	10	516	113	9	400	16	10	457	49
o	10	476	44	9	403	28	10	466	15
ɪ	10	412	15	10	431	38	8	462	88
ʊ	10	402	32	9	381	25	10	469	56
i	10	352	25	9	416	66	10	358	28
u	10	352	30	8	383	62	9	381	37
ə	10	576	31	9	604	33	10	614	52

As expected, given its low tongue body position, first formant values are highest for the low vowel /a/. The mid vowels, in turn, have higher first formant values than the high vowels, although this relationship only holds for vowels of equivalent tenseness/laxness. The lax high vowels /ɪ/ and /ʊ/ interestingly do not have reliably lower first formant values than the tense mid vowels /e/ and /o/, suggesting that the contrast between these two sets of vowels resides at least partially in the second formant. Schwa occupies a height equivalent to that of the lax mid vowels /ɛ, ɔ/, lower than the tense mid vowels but considerably higher than the low vowel.

Graphs depicting results for the four measured correlates of sonority appear in figure 3, duration in the top left, maximum intensity in the top right, acoustic energy in the bottom left, and perceptual energy in the bottom right. Individual speaker values for each dimension are given in tables 4-7.





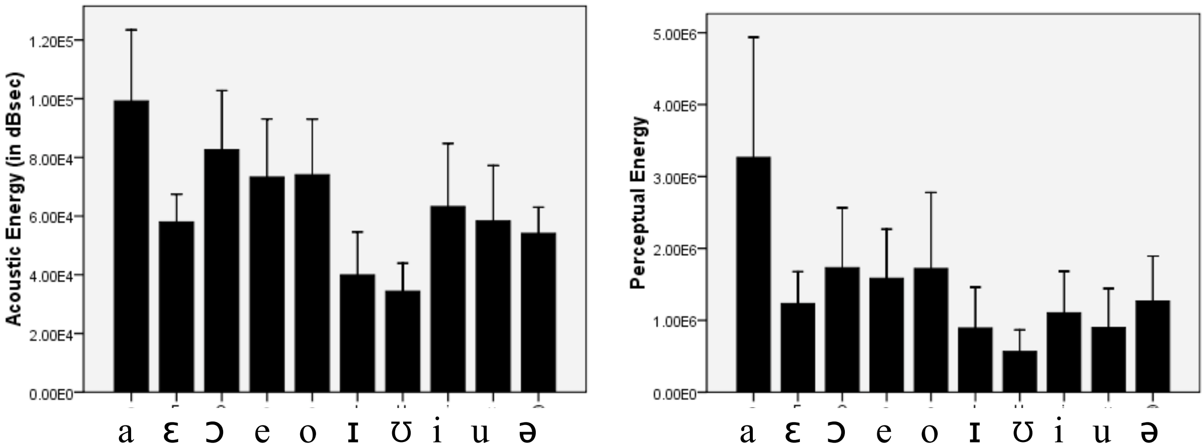


Figure 3. Duration (top left), maximum intensity (top right), acoustic energy (bottom left), and perceptual energy (bottom right) values averaged across three Hindi speakers. Whiskers indicate one standard deviation from the mean.

Table 4. Mean duration values (in seconds) for three Hindi speakers

	Speaker								
	M1			M2			M3		
Vowel	N	Mean	Std.Dev.	N	Mean	Std.Dev.	N	Mean	Std.Dev.
a	10	0.116	0.011	10	0.134	0.013	10	0.178	0.022
e	10	0.089	0.005	9	0.093	0.023	10	0.142	0.017
o	10	0.103	0.015	9	0.101	0.019	10	0.144	0.019
ɛ	10	0.091	0.013	10	0.082	0.014	5	0.081	0.014
ɔ	10	0.119	0.012	10	0.109	0.008	10	0.164	0.024
i	10	0.077	0.014	9	0.079	0.022	10	0.143	0.010
u	10	0.083	0.020	8	0.083	0.019	9	0.137	0.009
ɪ	10	0.055	0.011	10	0.053	0.015	8	0.069	0.027
ʊ	10	0.050	0.014	9	0.049	0.010	10	0.064	0.011
ə	10	0.078	0.010	9	0.082	0.010	10	0.081	0.015

Table 5. Mean maximum intensity values (in decibels) for three Hindi speakers

	Speaker								
	M1			M2			M3		
Vowel	N	Mean	Std.Dev.	N	Mean	Std.Dev.	N	Mean	Std.Dev.
a	10	78.5	1.0	10	83.7	0.9	10	84.2	2.9
e	10	77.3	2.8	9	79.1	1.6	10	79.0	1.8
o	10	71.5	1.8	9	80.1	3.1	10	80.8	2.4
ɛ	10	78.2	2.7	10	77.1	1.1	5	82.0	3.1
ɔ	10	75.5	2.9	10	78.2	2.2	10	81.6	2.2
i	10	70.2	3.2	9	77.5	4.8	10	76.1	1.8
u	10	67.9	1.5	8	73.7	1.1	9	73.7	3.6
ɪ	10	74.9	3.2	10	75.8	2.5	8	79.8	4.6
ʊ	10	70.0	2.3	9	75.9	2.4	10	76.5	3.7
ə	10	75.8	6.2	9	76.5	4.4	10	81.1	2.8

Table 6. Mean total acoustic energy values (in dB sec) for three Hindi speakers

	Speaker								
	M1			M2			M3		
Vowel	N	Mean	Std.Dev.	N	Mean	Std.Dev.	N	Mean	Std.Dev.
a	10	76777	7806	10	94819	10505	10	126039	18538
e	10	60047	5742	9	64274	17181	10	94660	11761
o	10	61761	9364	9	66534	12431	10	93346	15738
ɛ	10	59910	10184	10	55069	8914	5	60115	9216
ɔ	10	73482	10606	10	68342	7213	10	106016	14825
i	10	46503	8780	9	53215	12261	10	88999	8872
u	10	44983	11856	8	50882	11229	9	79843	9784
ɪ	10	37118	8911	10	34964	10131	8	49650	20924
ʊ	10	29537	8774	9	32052	7280	10	41486	8327
ə	10	50633	10075	9	54157	5489	10	57627	9426

Table 7. Mean perceptual energy values (in arbitrary units) for three Hindi speakers

	Speaker								
	M1			M2			M3		
Vowel	N	Mean	Std.Dev.	N	Mean	Std.Dev.	N	Mean	Std.Dev.
a	10	1.70E+6	3.87E+5	9	3.16E+6	5.30E+5	10	4.94E+6	1.65E+6
e	10	1.06E+6	3.77E+5	9	1.36E+6	3.80E+5	10	2.31E+6	4.88E+5
o	10	8.87E+5	2.54E+5	10	1.76E+6	8.11E+5	10	2.52E+6	1.16E+6
ɛ	10	1.23E+6	4.71E+5	10	1.01E+6	2.59E+5	5	1.68E+6	3.99E+5
ɔ	10	1.09E+6	2.56E+5	9	1.39E+6	3.28E+5	10	2.72E+6	6.27E+5
i	10	5.85E+5	1.75E+5	8	1.06E+6	2.74E+5	10	1.67E+6	5.36E+5
u	10	5.08E+5	1.35E+5	10	7.73E+5	2.66E+5	9	1.45E+6	5.73E+5
ɪ	10	6.67E+5	2.44E+5	9	6.93E+5	3.46E+5	8	1.43E+6	7.43E+5
ʊ	10	3.85E+5	1.54E+5	9	4.92E+5	2.09E+5	10	8.13E+5	3.34E+5
ə	10	1.07E+6	6.27E+5	10	9.54E+5	2.46E+5	10	1.75E+6	6.03E+5

One-factor analyses of variance (ANOVA) conducted for each of the four phonetic parameters indicated a significant effect of vowel quality on all parameters: for duration, $F(9, 275) = 37.312$, $p<.001$; for maximum intensity, $F(9, 275) = 16.171$, $p<.001$; for acoustic energy, $F(9, 175) = 35.793$, $p<.001$; for perceptual energy, $F(9, 275) = 24.086$, $p<.001$.

Table 8 summarizes the vowels differentiated at $p<.05$ or less according to Scheffe posthoc tests along each of the four phonetic parameters measured. “D” stands for duration, “I” for maximum intensity, “A” for acoustic energy, and “P” for perceptual energy. Sonority reversals, cases in which a phonetic parameter contradicts the ranking of two vowel qualities along phonological sonority scales, are indicated by “!” after the relevant parameter. Thus, for example, schwa has greater intensity than /u/ even though /u/ ranks higher on sonority scales.

Table 8. Summary of vowels distinguished by different phonetic parameters in Hindi

	ɛ	ɔ	e	o	ɪ	ʊ	i	u	ə
a	DAP	IP	DAP	IAP	DIAP	DIAP	DIAP	DIAP	DIAP
ɛ		DA		D		DA		I	
ɔ					DA	DIAP	DA	DIA	DA
e					DA	DIAP		I	DA
o					DA	DAP		I	DIA
ɪ						DA	DA	DI	
ʊ								DA	A!
i									
u									I!

The most reliable sonority distinction among the vowels is between the low vowel /a/ and all other vowels with /a/ having greater perceptual energy and, in most cases, greater duration, maximum intensity and acoustic energy, than other vowel qualities. There is also a somewhat weaker tendency for the mid vowels to be differentiated from the high vowels, in particular, the lax high vowels, although the relevant differentiating parameter(s) varies depending on the vowels involved.

Although schwa is distinguished from most of the mid vowels (with the exception of /ɛ/), the difference between schwa and the high vowels is not robustly manifested along the measured phonetic dimensions. In fact, there are two sonority reversals in which schwa is less prominent than a high vowel occupying a position higher on the sonority scale. Schwa thus has greater maximum intensity than /u/ and schwa has greater acoustic energy than /ʊ/. On the other hand, these reversals are only reversals when considered from a phonological standpoint. Phonetically, the fact that schwa may be more prominent than high vowels along certain dimensions is not surprising given schwa's lower tongue body position.

Of the measured correlates of sonority, duration and acoustic energy make the most sonority distinctions, both distinguishing 25 of the 45 possible pairwise comparisons. Both fare well in differentiating the low vowel from most other vowels and distinguishing the mid vowels (with the exception of /ɛ/) from the high vowels. Neither duration nor acoustic energy, however, differentiates the high lax vowels from schwa in the direction predicted by sonority scales.

Maximum intensity and perceptual energy predict 16 and 12, respectively, of the pairwise comparisons. Of all the measures, perceptual energy is the best at distinguishing the low vowel from other vowels, although it does not distinguish any of the high vowels from schwa. Maximum intensity is the most successful differentiator of mid and high vowels, although it too fails to capture the lesser phonological sonority of schwa relative to the high vowels.

3.2. Besemah

3.2.1. Background

Besemah (McDonnell 2008) is an Austronesian language of Sumatra with a conservative vowel system consisting of four vowel phonemes (see table 9) including schwa.

Table 9. Vowels of Besemah

	Front	Central	Back
High	i		u
Mid		ə	
Low		a	

Schwa contrasts with the other vowel phonemes in the penultimate syllable in both open and closed syllables, but is in complementary distribution with /a/ in word-final position with /a/ occurring in closed syllables and schwa in open syllables.

Although the properties of word-level stress are not entirely clear, it is clear that in words in isolation stress falls on the final syllable, even when schwa is in the final syllable (3).

(3) Besemah stress

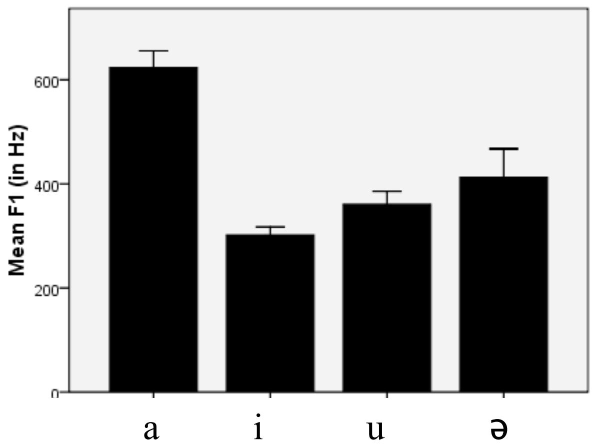
tiˈpu 'spy', kaˈtə 'word', tiˈtu 'that', aˈpi 'fire'

3.2.2. Methodology

The Besemah data was recorded as part of a study of vowel quality in Besemah so the recording conditions differ slightly between Besemah and the other examined languages. The measured vowels for Besemah appeared in the penultimate (unstressed) syllable of a disyllabic word, a context in which all four vowel phonemes are found. In stressed (final) syllables, /a/ and schwa are in complementary distribution, with schwa occurring in open syllables and /a/ in closed syllables. The analyzed Besemah words were uttered twice in isolation. Each word was recorded on a Marantz PMD670 solidstate recorder with an Audio-Technica AT825 stereo microphone at a sampling frequency of 48 kHz. The sound files were stored as .wav audio files in preparation for analysis. Data from two male speakers and two female speakers were collected.

3.2.3. Results

Figures 7 and 8 show mean first formant values averaged across male and female speakers, respectively. Individual speaker results for all speakers appear in table 14.



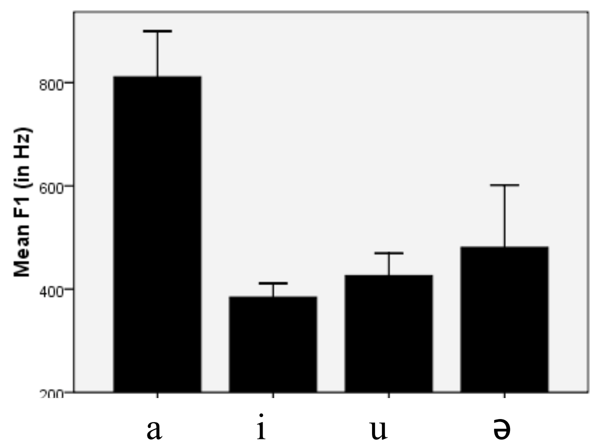


Figure 4. Mean first formant values averaged across two male (left) and two female (right) Besemah speakers. Whiskers indicate one standard deviation from the mean.

Table 10. Mean first formant values (in Hz) for four Besemah speakers

	Speaker											
	F1			F2			M1			M2		
Vowel	N	Mean	Std.Dev	N	Mean	Std.Dev	N	Mean	Std.Dev	N	Mean	Std.Dev
a	4	886	55	4	736	13	4	637	34	4	610	27
i	4	400	25	4	369	20	4	307	15	4	297	15
u	4	459	37	4	393	8	4	372	22	4	350	25
ə	4	527	159	4	435	55	4	367	25	4	458	30

Formant values suggest a four-way height distinction with /i/ being highest followed in turn by /u/, /ə/ and /a/. The higher F1 values associated with /u/ is consistent with a common cross-linguistic pattern (de Boer 2011).[2] This four-way distinction is most clearly evinced in the data from speakers F1 and M2. The distinction between schwa and /u/ is least robust across speakers as one of the male speakers (M1) has virtually identical first formant values for the two vowels, and one of the female speakers (F2) has a relatively small 42Hz difference in first formant values between these two vowels. Speaker F2 also displays the smallest difference between the two phonemic high vowels /i, u/.

Graphs depicting results for the four measured correlates of sonority appear in figure 5, duration in the top left, maximum intensity in the top right, acoustic energy in the bottom left, and perceptual energy in the bottom right. Individual speaker values for each dimension are given in tables 11-14.

[2] Thanks to Steve Parker for pointing out this cross-linguistic tendency.

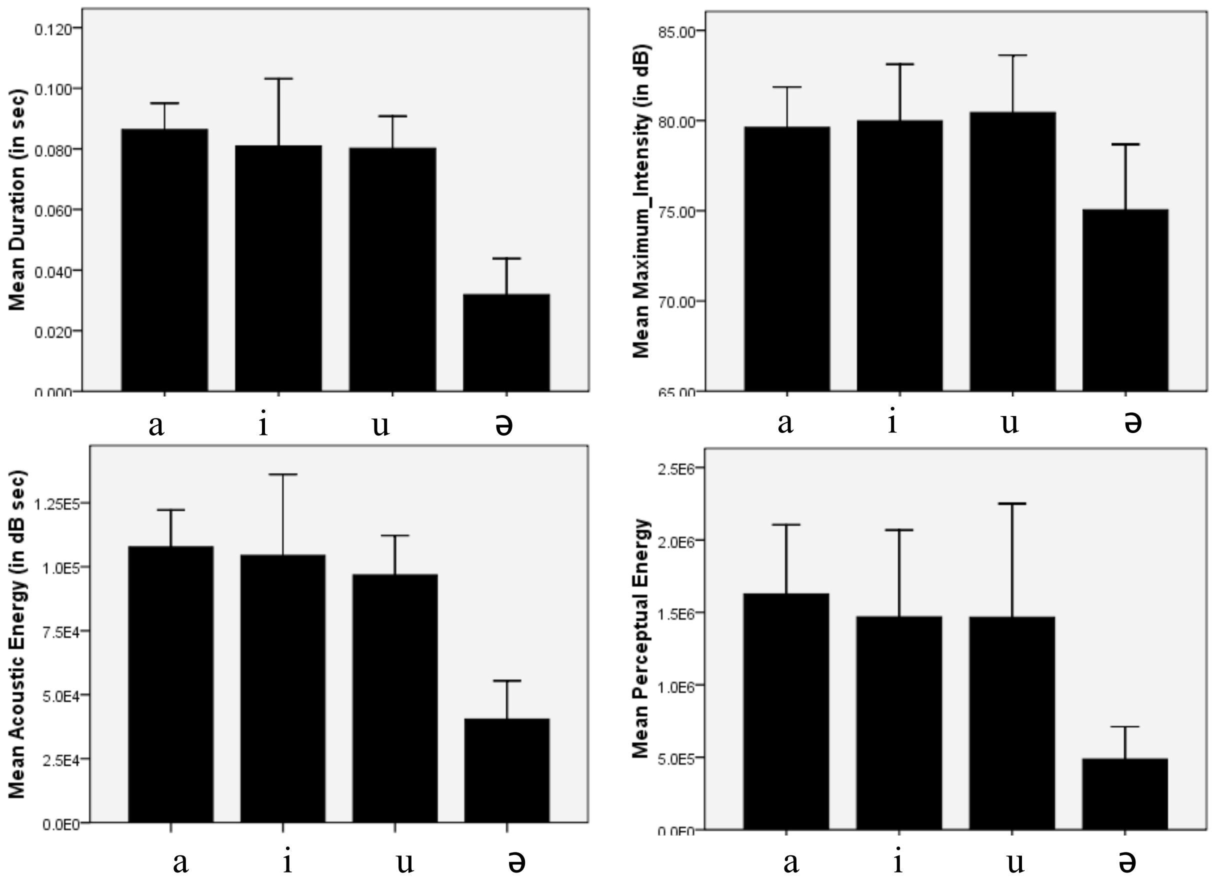


Figure 5. Duration (top left), maximum intensity (top right), acoustic energy (bottom left), and perceptual energy (bottom right) values averaged across four Besemah speakers. Whiskers indicate one standard deviation from the mean.

Table 11. Mean duration values for four Besemah speakers

	Speaker											
Vowel	F1			F2			M1			M2		
	N	Mean	Std.Dev	N	Mean	Std.Dev	N	Mean	Std.Dev	N	Mean	Std.Dev
a	4	0.082	0.012	4	0.090	0.011	4	0.088	0.006	4	0.086	0.005
i	4	0.062	0.016	4	0.088	0.025	4	0.078	0.006	4	0.097	0.025
u	4	0.072	0.011	4	0.089	0.006	4	0.081	0.009	4	0.079	0.012
ə	4	0.032	0.016	4	0.026	0.007	4	0.034	0.016	4	0.035	0.009

Table 12. Mean maximum intensity values for four Besemah speakers

	Speaker											
Vowel	F1			F2			M1			M2		
	N	Mean	Std.Dev	N	Mean	Std.Dev	N	Mean	Std.Dev	N	Mean	Std.Dev
a	4	79.06	2.30	4	78.90	1.97	4	79.63	2.61	4	80.90	2.41
i	4	79.14	3.57	4	78.74	3.45	4	80.47	2.13	4	81.59	3.64
u	4	78.88	2.06	4	81.26	4.06	4	79.18	2.29	4	82.45	3.61
ə	4	73.15	2.42	4	72.43	2.63	4	74.56	1.21	4	80.08	1.83

Table 13. Mean acoustic energy values for four Besemah speakers

Vowel	F1 N	F1 Mean	F1 Std.Dev	F2 N	F2 Mean	F2 Std.Dev	M1 N	M1 Mean	M1 Std.Dev	M2 N	M2 Mean	M2 Std.Dev
a	4	10121	18872	4	106510	12677	4	1458371	494461	4	115498	10496
i	4	80351	18679	4	99784	27526	4	726867	731605	4	133150	25653
u	4	87783	11180	4	94670	11277	4	1323827	320330	4	107883	18360
ə	4	37841	18229	4	30574	79561	4	491690	259154	4	52961	8724

Table 14. Mean perceptual energy values for four Besemah speakers

V	F1 N	F1 Mean	F1 Std.Dev	F2 N	F2 Mean	F2 Std.Dev	M1 N	M1 Mean	M1 Std.Dev	M2 N	M2 Mean	M2 Std.Dev
a	4	1.61E+6	4.13E+5	4	1.36E+6	3.30E+5	4	1.46E+6	4.94E+5	4	2.09E+6	4.41E+5
i	4	1.19E+6	4.92E+5	4	1.22E+6	6.57E+5	4	1.32E+6	1.42E+5	4	2.14E+6	5.35E+5
u	4	9.25E+5	1.73E+5	4	1.32E+6	4.07E+5	4	1.32E+6	3.20E+5	4	2.30E+6	1.18E+6
ə	4	4.44E+5	2.11E+5	4	2.89E+5	6.41E+4	4	4.92E+5	2.59E+5	4	7.26E+5	4.83E+4

One-factor analyses of variance (ANOVA) conducted for each of the four phonetic parameters indicated a significant effect of vowel quality on all parameters except acoustic energy: for duration, $F(3, 60) = 50.227$, $p<.001$; for maximum intensity, $F(3, 60) = 10.531$, $p<.001$; for perceptual energy, $F(3, 44) = 20.432$, $p<.001$. Note that data from the second male speaker was excluded from the two ANOVAs involving energy since his acoustic and perceptual energy values were sharply divergent from those of the other speakers.

Table 15 summarizes the various phonetic parameters distinguishing (at $p<.05$ or less according to Scheffe posthoc tests) the vowels of Besemah.

Table 15. Summary of vowels distinguished by different phonetic parameters in Besemah

	i	u	ə
a			DIP
i			DIP
u			DIP

Schwa is differentiated from the other three vowels in duration, maximum intensity, and perceptual energy, a result that accords with the position of schwa lower on the phonological sonority scale than other vowels cross-linguistically, even though schwa is not distinguished from other vowels in terms of its ability to attract stress in Besemah itself. None of the three non-schwa vowels, however, are distinguished along any of the measured dimensions.

3.3. Armenian

3.3.1. Background

Standard Western Armenian possesses six vowel phonemes, one of which is schwa (Vaux 1998).

Table 16. The vowels of Armenian

	Front	Central	Back
High	i		u
Mid	ɛ	ə	ɔ
Low		a	

Primary stress in most varieties of Armenian falls on the final syllable unless the final syllable contains schwa (4a), in which case stress retracts onto the penult (4b). There are no content words whose only vowels are schwa. Vaux (1998) reports that secondary stress characteristically falls on the initial syllable (Vaux 1998). Examples of Armenian stress, repeated from (1), with secondary stress also marked, appear in (4).

(4) Armenian stress (Vaux 1998:132)

a. ˌmɔˈrukʰ 'beard', ˌɑrtɑˈsukʰ 'tears', ˌjɛrkrɑkdrɔnɑˈkɑn 'geocentric'

b. ˈmɑnər 'small', ˌjɛrˈpʰɛmən 'sometimes'

3.3.2. Methodology

The measured vowels for Armenian appeared in the penultimate (secondary stressed) syllable of a disyllabic word. The Armenian words were uttered five times in the carrier phrase *ɑˈsɑ* ___ *nɔˈrits* 'Say ___ again'. Each word was recorded on a Marantz PMD660 solidstate recorder as a .wav file using a unidirectional microphone (Shure SM10) at a sampling frequency of 44.1 kHz. Data from two male speakers and two female speakers of Eastern Armenian were collected. Three of the four speakers were born in Yerevan, Armenia before emigrating to the United States as children, while the one of the speakers was born in the United States of parents who speak the Eastern Armenian dialect. All of the speakers spoke Armenian as their first language.

3.3.3. Results

3.3.3.1. First Formant

Mean first formant values averaged across the male speakers and female speakers are shown in figures 13 and 14, respectively. Individual speaker results appear in table 17.

Both the male and female speakers make a clear three-way height distinction between high vowels, mid vowels, including schwa, and a low vowel. For the male speakers, the high back and mid vowels are associated with slightly higher first formant values, i.e. a lower tongue body position, than their front mid counterparts, i.e. /u/ is slightly lower than /i/, and /ɔ/ is somewhat lower than /ɛ/, in keeping with a common cross-linguistic tendency (de Boer 2011) found earlier in the Besemah data.

Graphs depicting results for the four measured correlates of sonority appear in figure 7, duration in the top left, maximum intensity in the top right, acoustic energy in the bottom left, and perceptual energy in the bottom right. Individual speaker values for each dimension are given in tables 18-21.

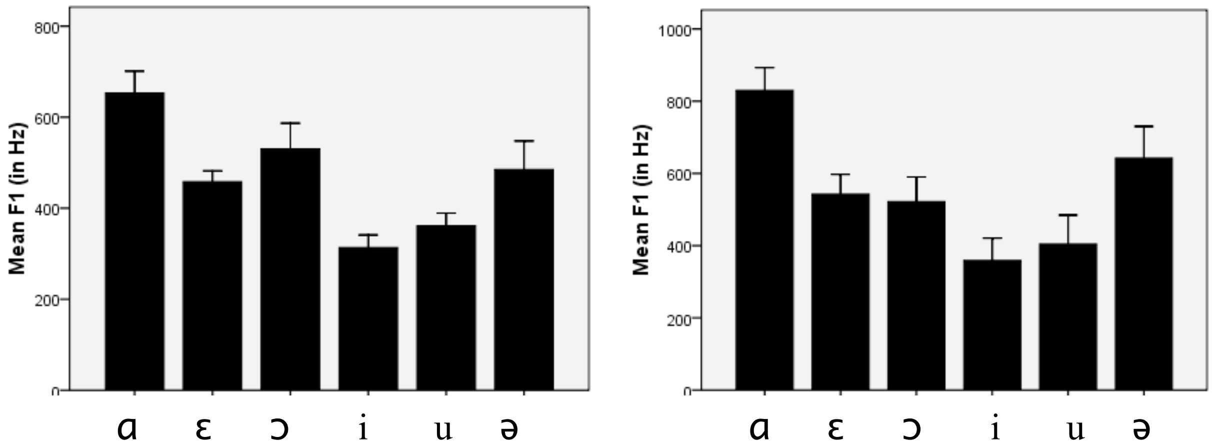
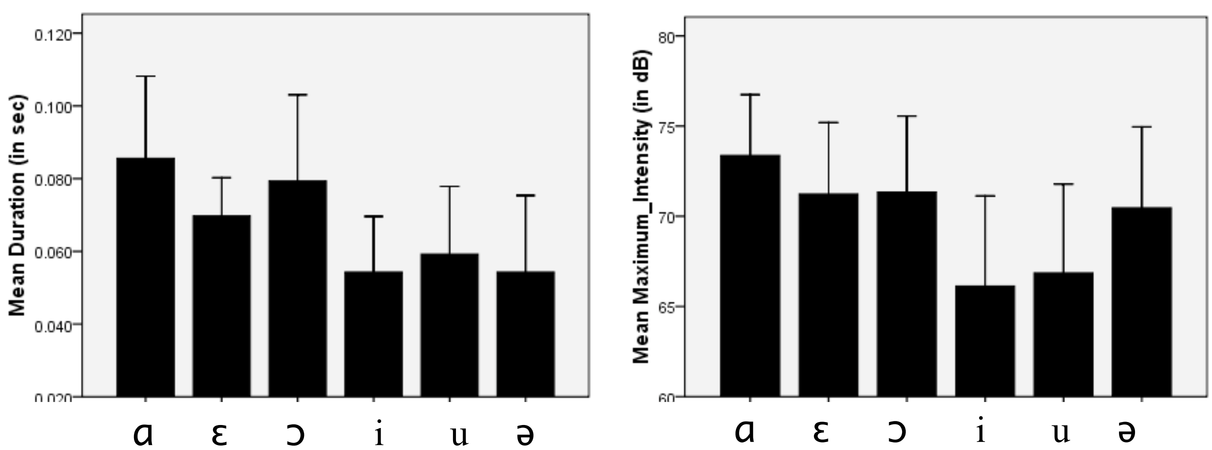


Figure 6. Mean first formant values averaged across two male (left) and two female (right) Armenian speakers. Whiskers indicate one standard deviation from the mean.

Table 17. Mean first formant values (in Hz) for four Armenian speakers

	Speaker											
V	F1			F2			M1			M2		
	N	Mean	Std.Dev	N	Mean	Std.Dev	N	Mean	Std.Dev	N	Mean	Std.Dev
a	10	790	21	9	870	66	10	608	24	11	690	25
ɛ	6	601	50	8	511	19	11	451	29	10	465	17
ɔ	9	575	39	10	474	51	10	488	34	10	573	40
i	10	417	16	10	302	20	10	314	30	10	313	26
u	10	455	71	10	347	44	9	358	34	9	365	18
ə	19	678	80	24	607	83	19	471	45	29	497	73



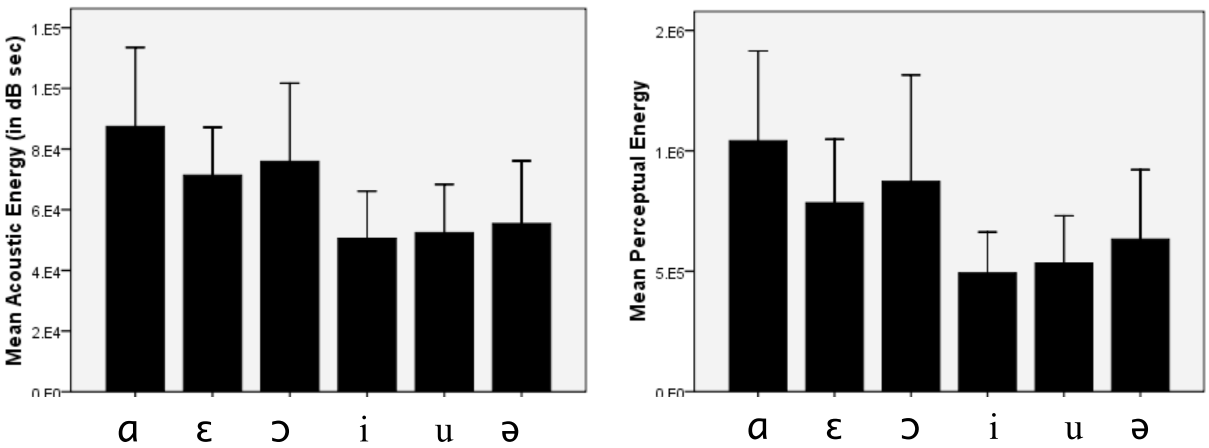


Figure 7. Duration (top left), maximum intensity (top right), acoustic energy (bottom left), and perceptual energy (bottom right) values averaged across four Armenian speakers. Whiskers indicate one standard deviation from the mean.

Table 18. Mean duration values (in seconds) for four Armenian speakers

	Speaker											
Vowel	F1			F2			M1			M2		
	N	Mean	StdDev	N	Mean	StdDev	N	Mean	StdDev	N	Mean	StdDev
a	10	0.099	0.016	9	0.090	0.021	10	0.077	0.017	11	0.076	0.028
ɛ	6	0.072	0.009	8	0.078	0.011	11	0.064	0.009	10	0.064	0.006
ɔ	9	0.090	0.021	10	0.092	0.029	10	0.070	0.018	10	0.067	0.017
i	10	0.064	0.013	10	0.055	0.007	10	0.055	0.019	10	0.043	0.014
u	10	0.068	0.015	10	0.061	0.027	9	0.062	0.010	9	0.045	0.014
ə	19	0.068	0.026	24	0.067	0.015	19	0.044	0.013	29	0.045	0.017

Table 19. Mean maximum intensity values for four Armenian speakers

	Speaker											
Vowel	F1			F2			M1			M2		
	N	Mean	StdDev	N	Mean	StdDev	N	Mean	StdDev	N	Mean	StdDev
a	10	74.36	2.32	9	73.86	3.11	10	70.54	2.34	11	74.35	4.11
ɛ	6	69.48	2.99	8	74.39	2.45	11	67.16	2.71	10	72.08	3.49
ɔ	9	70.42	2.72	10	76.51	2.65	10	67.40	2.85	10	70.95	2.24
i	10	66.97	2.14	10	71.34	2.24	10	61.28	4.00	10	64.95	4.90
u	10	68.17	3.41	10	73.16	1.33	9	63.17	1.02	9	63.21	4.17
ə	19	71.85	2.14	24	74.77	2.79	19	66.93	3.20	29	69.67	4.82

Table 20. Mean acoustic energy values (in dB seconds) for four Armenian speakers

	Speaker											
Vowel	F1			F2			M1			M2		
	N	Mean	StdDev	N	Mean	StdDev	N	Mean	StdDev	N	Mean	StdDev
a	10	99881	19327	9	96047	22997	10	74659	21545	11	78864	31529
ɛ	6	70810	18160	8	83159	13331	11	59432	9865	10	68402	13132
ɔ	9	78161	17671	10	95687	35052	10	65726	14752	10	64273	19634
i	10	61045	10325	10	54820	6621	10	45961	16927	10	40352	18176
u	10	60238	10671	10	58676	22705	9	51905	9761	9	38037	8187
ə	19	70084	21754	24	68958	14328	19	43151	11438	29	47163	18995

Table 21. Mean perceptual energy values (in arbitrary units) for four Armenian speakers

	Speaker											
V	F1			F2			M1			M2		
	N	Mean	StdDev	N	Mean	StdDev	N	Mean	StdDev	N	Mean	StdDev
a	10	1.09E+6	2.96E+5	9	1.17E+6	4.22E+5	10	8.04E+5	1.96E+5	11	1.08E+6	4.48E+5
ɛ	6	6.65E+5	1.78E+5	8	1.04E+6	2.37E+5	11	5.92E+5	1.35E+5	10	7.31E+5	2.13E+5
ɔ	9	8.18E+5	2.60E+5	10	1.30E+6	6.24E+5	10	6.13E+5	1.49E+5	10	7.52E+5	2.28E+5
i	10	5.35E+5	9.03E+4	10	5.86E+5	1.30E+5	10	4.13E+5	1.47E+5	10	4.38E+5	2.36E+5
u	10	6.04E+5	1.43E+5	10	7.05E+5	2.36E+5	9	4.60E+5	6.09E+4	9	3.66E+5	1.35E+5
ə	19	7.63E+5	2.19E+5	24	9.08E+5	2.68E+5	19	4.22E+5	1.21E+5	29	5.44E+5	2.62E+5

One-factor analyses of variance (ANOVA) conducted for each of the four phonetic parameters indicated a significant effect of vowel quality on all parameters: for duration, $F(5, 277) = 21.775$, $p<.001$; for maximum intensity, $F(5, 277) = 16.309$, $p<.001$; for acoustic energy, $F(5, 277) = 22.865$, $p<.001$; for perceptual energy, $F(5, 277) = 20.296$, $p<.001$.

Table 22 summarizes the various phonetic parameters distinguishing (at $p<.05$ or less according to Scheffe posthoc tests) the vowels of Armenian.

Table 22. Summary of vowels distinguished by different phonetic parameters in Armenian

	ɛ	ɔ	i	u	ə
a	DAP		DIAP	DIAP	DIAP
ɛ			DIAP	IAP	DA
ɔ			DIAP	DIAP	DAP
i					I!
u					I!

The clearest distinction emerging overall is the bifurcation between the high vowels and schwa, on the one hand, and the peripheral mid vowels and /a/, on the other hand. The low vowel is also distinguished from the front mid vowel /ɛ/ along three of the four measured dimensions. Maximum intensity draws a further distinction between the high vowels and schwa with schwa displaying greater peak intensity values than the two high vowels, a contradiction of phonological sonority scales placing schwa below high vowels.

3.4. Javanese

3.4.1. Background

The standard Javanese of central Java is typically characterized as having six or eight vowel phonemes (Clynes and Rudyanto 1995, Horne 1974). In the six vowel system, which appears to characterized the speech of our consultants, [ɛ] and [ɔ] are in complementary distribution with [e] and [o], respectively. The lower allophone occurs when the following vowel is schwa or another mid vowel, or /i, u/ in an open syllable. The schwa occurs in pre-final syllables, both open and closed, but only in closed syllables word-finally.

Table 23. Vowels of Javanese

	Front	Central	Back
High	i		u
Mid	e (ɛ)	ə	o (ɔ)
Low		a	

Stress in Javanese falls on the penultimate syllable (5a) unless the penult has a schwa (Herrfurth 1964, Horne 1974), in which case, stress shifts to the final syllable (5b).

(5) Javanese stress
   a. ˈpantun ‘rice plant’, ˈkates ‘papaya’
   b. kəˈtes ‘slap’, kəˈtan ‘sticky rice’, jənˈtən ‘cumin, caraway seed’

3.4.2. Methodology

The target vowels for Javanese appeared in the final (stressed) syllable of a disyllabic word containing a schwa in the penult. Each vowel appeared in a closed syllable in three words in the corpus. Each Javanese word was uttered four times in isolation and recorded using a high quality unidirectional microphone connected to a solidstate recorder. Data from two male speakers were collected; one speaker from Bojonegoro in East Java (at a sampling rate of 44.1kHz) and one speaker from Semarang in Central Java (recorded as part of a larger phonetic study of Javanese at a sampling rate of 22.05Hz). The sound files were stored as .wav audio files in preparation for analysis.

3.4.3. Results

Figure 8 shows first formant values for the measured Javanese vowels averaged across the two speakers. Individual speaker means appear in table 24.

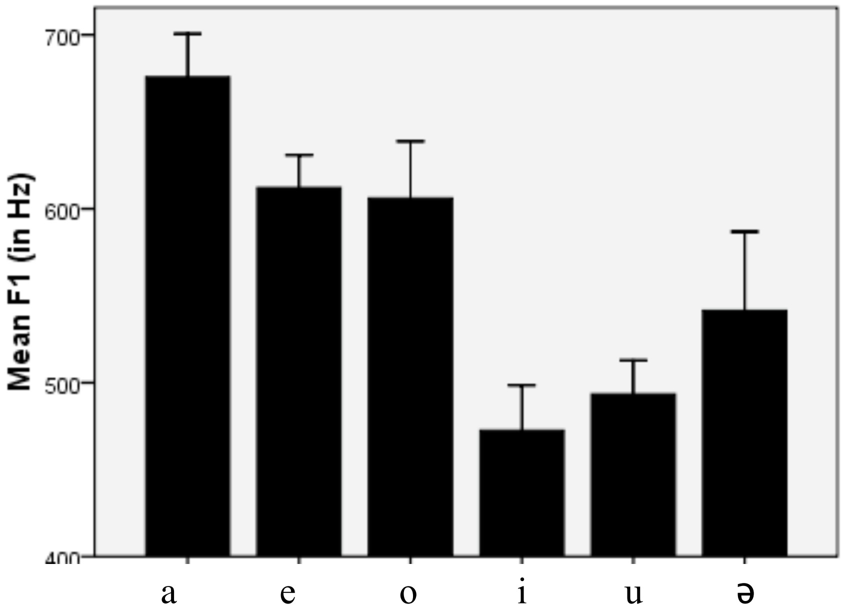


Figure 8. First formant values averaged across two Javanese speakers. Whiskers indicate one standard deviation from the mean.

Table 24. Mean first formant values (in Hz) for two Javanese speakers

	Speaker					
Vowel	M1			M2		
	N	Mean	Std.Dev.	N	Mean	Std.Dev.
a	12	681	21	12	671	28
e	6	607	25	9	616	14
o	11	628	19	12	586	30
i	12	465	33	12	480	14
u	12	497	14	12	490	24
ə	12	525	44	12	558	42

First formant values distinguish five vowel heights with the two peripheral mid vowels /e, o/ being slightly lower (consistent with a transcription as /ɛ, ɔ/ rather than /e, o/) than schwa, which occupies the middle of the vowel space.

Graphs depicting results for the four measured correlates of sonority appear in figure 9, duration in the top left, maximum intensity in the top right, acoustic energy in the bottom left, and perceptual energy in the bottom right. Individual speaker values for each dimension are given in tables 25-28.

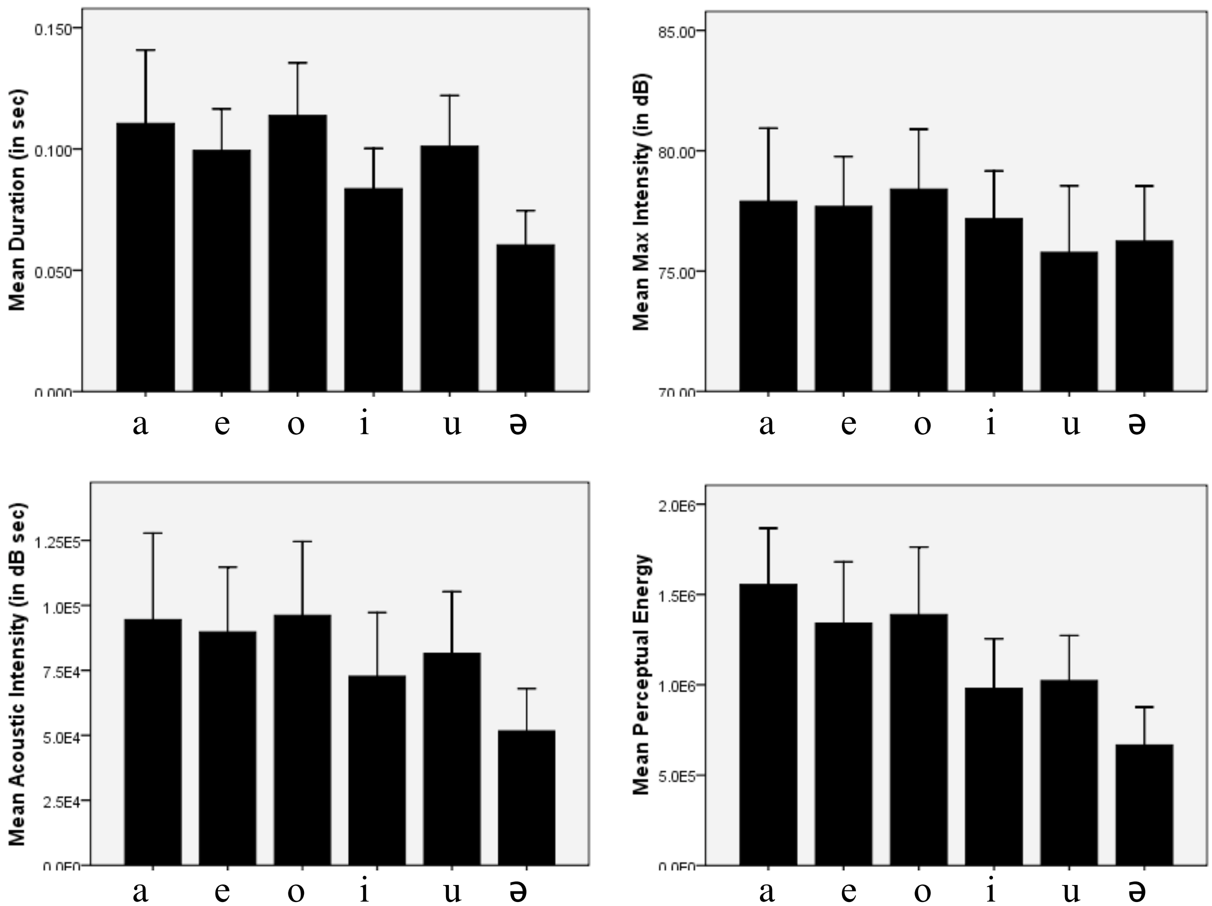


Figure 9. Duration (top left), maximum intensity (top right), acoustic energy (bottom left), and perceptual energy (bottom right) values averaged across two Javanese speakers. Whiskers indicate one standard deviation from the mean.

Table 25. Mean duration values (in seconds) for two Javanese speakers

	Speaker					
Vowel	M1			M2		
	N	Mean	Std.Dev.	N	Mean	Std.Dev.
a	12	0.108	0.038	12	0.113	0.021
e	6	0.102	0.018	9	0.098	0.018
o	11	0.114	0.028	12	0.113	0.016
i	12	0.079	0.018	12	0.089	0.014
u	12	0.105	0.025	12	0.098	0.016
ə	12	0.061	0.018	12	0.060	0.010

Table 26. Mean maximum intensity values (in dB) for two Javanese speakers

Vowel	Speaker					
	M1			M2		
	N	Mean	Std.Dev.	N	Mean	Std.Dev.
a	12	78.97	2.52	12	76.84	3.23
e	6	78.71	1.81	9	77.00	2.04
o	11	79.43	2.09	12	77.47	2.53
i	12	77.87	1.93	12	76.48	1.85
u	12	76.18	1.73	12	75.38	3.54
ə	12	76.80	1.90	12	75.69	2.58

Table 27. Mean acoustic energy values (in dB seconds) for two Javanese speakers

Vowel	Speaker					
	M1			M2		
	N	Mean	Std.Dev.	N	Mean	Std.Dev.
a	12	6.77E+4	2.11E+4	12	1.21E+5	1.71E+4
e	6	6.76E+4	1.39E+4	9	1.05E+5	1.88E+4
o	11	7.25E+4	1.75E+4	12	1.18E+5	1.60E+4
i	12	5.24E+4	1.20E+4	12	9.30E+4	1.47E+4
u	12	6.41E+4	1.56E+4	12	9.90E+4	1.66E+4
ə	12	3.96E+4	1.17E+4	12	6.39E+4	9.79E+3

Table 28. Mean perceptual energy values (in arbitrary units) for two Javanese speakers

Vowel	Speaker					
	M1			M2		
	N	Mean	Std.Dev.	N	Mean	Std.Dev.
a	12	1.50E+6	3.76E+5	12	1.61E+6	2.36E+5
e	6	1.18E+6	3.90E+5	9	1.45E+6	2.70E+5
o	11	1.35E+6	4.89E+5	12	1.43E+6	2.40E+5
i	12	7.94E+5	2.04E+5	12	1.17E+6	1.98E+5
u	12	9.24E+5	2.91E+5	12	1.12E+6	1.55E+5
ə	12	5.84E+5	2.27E+5	12	7.51E+5	1.60E+5

One-factor analyses of variance (ANOVA) conducted for each of the four phonetic parameters indicated a significant effect of vowel quality on all parameters: for duration, $F(5, 128) = 21.195$, $p<.001$; for maximum intensity, $F(5, 128) = 3.826$, $p=.003$; for acoustic energy, $F(5, 128) = 9.972$, $p<.001$; for perceptual energy, $F(5, 128) = 28.750$, $p<.001$.

Table 29 summarizes the phonetic dimensions that distinguish (at $p<.05$ or less according to Scheffe posthoc tests) the vowels of Javanese.

Table 29. Summary of vowels distinguished by different phonetic parameters in Javanese.

	e	o	i	u	ə
a			DAP	P	DAP
e			P		DAP
o			DAP	IP	DAP
i					DP
u					DAP

The reduced prominence of schwa relative to other vowels is consistent with phonological scales placing schwa at the bottom of the sonority hierarchy and also accords with the light status of schwa in the Javanese stress system. All pairwise comparisons involving schwa are distinguished through duration, acoustic energy, and perceptual energy, with the exception of the comparison of schwa with /i/, which is not differentiated through acoustic energy.

The low and mid vowels are distinguished from the high vowels with the exception of /e/ and /u/. These distinctions, however, are manifested along different acoustic dimensions. Perceptual energy is the most reliable differentiator of high vowels from both low and mid vowels, failing only to distinguish /e/ and /u/. Duration and acoustic energy differentiate /i/ from both /a/ and /o/, while maximum intensity serves to separate /o/ from /u/.

3.5. Kʷak'ʷala

3.5.1. Background

The Kʷak'ʷala vowel system can be characterized in terms of three "full" vowel phonemes distributed around the periphery of the vowel space: /i, a, u/, plus a shorter central vowel, schwa /ə/ (Boas 1947, Grubb 1977). The relative height and/or backness of particular realizations within each of these phonemic vowel categories varies across dialects and speakers, generally reflecting a broad spectrum of co-articulatory effects with adjacent consonants.

The distinction between schwa and the other full vowels plays a fundamental role in the locus of stress (Boas 1947, Bach 1975, Wilson 1986, Zec 1988, Shaw 2009). The basic generalization is that primary stress falls on the leftmost full vowel of a word (6a) and, in the absence of a full vowel anywhere in the word, on the rightmost schwa (6b). The dichotomous stress behavior of the "full" vowels vs. schwa is attenuated by the relative sonority of coda consonants. A schwa followed by a sonorant coda thus follows the same stress generalizations as syllables with a full vowel nucleus, and whichever is leftmost will receive primary stress (6c). Laryngealization in a coda (represented by an apostrophe in (6)), on the other hand, has a prominence-reduction effect. Syllables with a glottalized resonant or a glottal stop in the coda (Shaw 2009), regardless of vowel quality, are thus skipped in the scan for the leftmost stressable syllable (6d).

(6) a. ˈkʷak'ʷala 'Kwakwala', səˈbaju 'searchlight', bəq'ʷəɬəˈla 'sleepy, drowsy'
b. tsəˈGəɬ 'thimbleberry', tsəGəɬˈm'əs 'thimbleberry plant', dzəˈGʷəd 'coal'
c. ˈt'əmxʷ.m'əs 'wild gooseberry plant', ɬə.ˈnəm.di 'red elderberry plant'
d. gəl'.ˈdzud 'to crawl onto a flat thing', gʷaʔ.sə.ˈla 'people of Smith's Inlet'

3.5.2. Methodology

The four vowels in Kʷak'ʷala all appeared in a stressed final syllable in isolation. One male speaker of Kʷak'ʷala was recorded repeating each word five times. Recordings were made at a sampling rate of 48 kHz using a Marantz PMD670 solidstate recorder via a desktop Audio-Technica AT-831b cardioid condenser microphone.

3.5.3. Results

A one-way ANOVA indicates a significant effect of vowel quality on first formant values: F (3, 87) = 50.068, p<.001. Figure 10 shows first formant values for the measured Kʷak'ʷala vowel, followed by mean values in table 30.

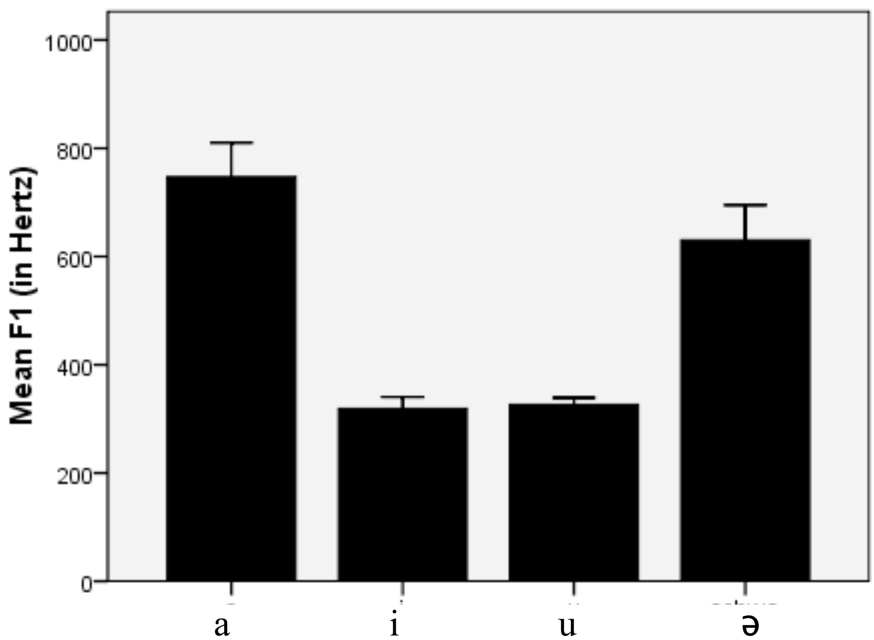


Figure 10. First formant values averaged across one male Kʷak'ʷala speaker. Whiskers indicate one standard deviation from the mean.

Table 30. Mean first formant values (in Hz) for one Kʷak'ʷala speaker

Vowel	F1		
	N	Mean	Std.Dev.
a	14	743	63
i	24	319	21
u	15	327	12
ə	10	631	64

As the phonemic transcription of the four vowels suggests, first formant values confirm that there is a three-way height distinction with /a/ lowest in quality, /i/ and /u/ highest, and schwa intermediate in height.

Graphs depicting results for the four measured correlates of sonority appear in figure 11, duration in the top left, maximum intensity in the top right, acoustic energy in the bottom left, and perceptual energy in the bottom right. Individual speaker values for each dimension are given in tables 31-34.

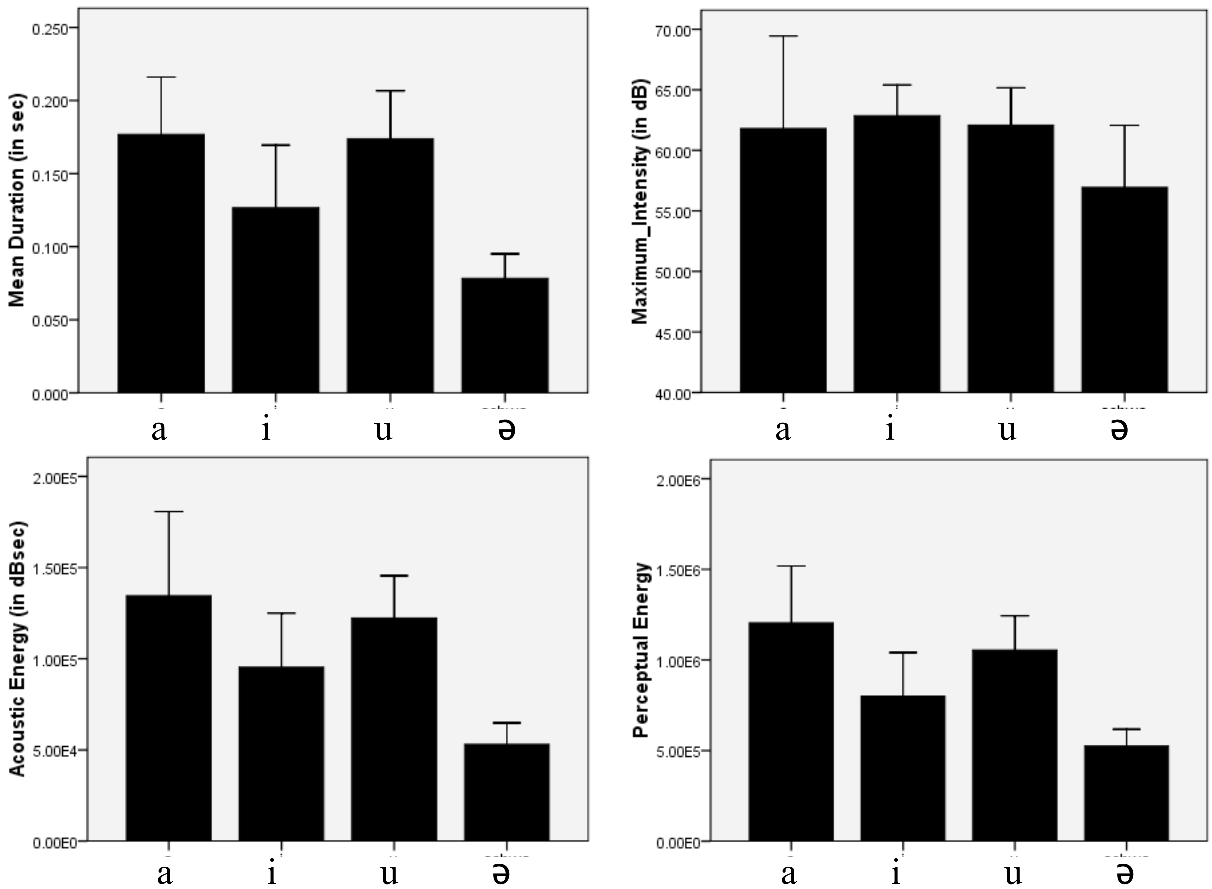


Figure 11. Duration (top left), maximum intensity (top right), acoustic energy (bottom left), and perceptual energy (bottom right) values averaged across one male K^{w}ak'^{w}ala speaker. Whiskers indicate one standard deviation from the mean.

Table 31. Mean duration values (in seconds) for one K^{w}ak'^{w}ala speaker

Vowel		F1	
	N	Mean	Std.Dev.
a	14	.177	.039
i	24	.127	.043
u	15	.174	.033
ə	10	.078	.017

Table 32. Mean maximum intensity values (in dB) for one Kʷak'ʷala speaker

Vowel	F1		
	N	Mean	Std.Dev.
a	14	61.81	7.63
i	24	62.86	2.54
u	15	62.08	3.09
ə	10	56.94	5.12

Table 33. Mean acoustic energy values (in decibel seconds) for one Kʷak'ʷala speaker

Vowel	F1		
	N	Mean	Std.Dev.
a	14	134559	46159
i	24	95414	29547
u	15	122239	23230
ə	10	53109	11717

Table 34. Mean perceptual energy values (in arbitrary units) for one Kʷak'ʷala speaker

Vowel	F1		
	N	Mean	Std.Dev.
a	14	1204508	313873
i	24	799776	240869
u	15	1054565	189172
ə	10	906745	323689

One-factor analyses of variance (ANOVA) conducted for each of the four phonetic parameters indicated a significant effect of vowel quality on all parameters: for duration, $F(3, 59) = 19.376$, $p<.001$; for maximum intensity, $F(3, 59) = 3.981$, $p=.012$; for acoustic energy, $F(3, 59) = 15.908$, $p<.001$; for perceptual energy, $F(3, 59) = 20.331$, $p<.001$.

Table 35 summarizes the phonetic dimensions that distinguish (at $p<.05$ or less according to Scheffe posthoc tests) the vowels of Kʷak'ʷala.

Table 35. Summary of vowels distinguished by different phonetic parameters in Kʷak'ʷala.

	i	u	ə
a	DAP		DAP
i		DP	DIAP
u			DAP

The clearest phonetic distinction is between schwa and the three full vowels, all of which are more prominent than schwa along at least three dimensions. Only maximum intensity fails to differentiate all three full vowels from schwa. Interestingly, the low vowel differs from /i/ (in all measures except maximum intensity) but not from /u/. On the other hand, /u/ is differentiated from its front high counterpart /i/ in both duration and perceptual energy, suggesting a sonority distinction, at least phonetically, between higher sonority /u/ and lower sonority /i/.

4. Discussion

Comparison of the results across the five examined languages indicates a number of similarities as well as certain differences in the relative prominence of different vowel qualities along the various studied phonetic dimensions as well as in the particular phonetic parameters used to differentiate vowels. Table 36 encapsulates the phonetic distinctions between different vowel qualities in the examined languages and the properties used to distinguish the vowels. To facilitate comparison across the five languages, all of which except Hindi do not make tense vs. lax distinctions, the tense vowels but not the lax vowels of Hindi are included in the table and the mid vowels of all other languages are represented as /e, o/ regardless of their phonetic height within the mid vowel subspace, i.e. whether they are phonetically /e/ or /ɛ/ and /o/ or /ɔ/. Sonority reversals in which a lower sonority vowel according to phonological scales has greater prominence along a given dimension are represented with "!" after the relevant phonetic parameter. Phonemic vowel pairs that are not differentiated phonetically along the measured dimensions in a given language are indicated by Ø. Light shaded cells occur at the intersection of contrasts that do not occur in a given language.

Table 36. Summary of the phonetic distinctions between vowels in five languages

		e	o	i	u	ə
	Hindi	DAP	IAP	DIAP	DIAP	DIAP
	Besemah	Ø	Ø	Ø	Ø	DIP
a	Armenian	DAP	Ø	DIAP	DIAP	DIAP
	Javanese	Ø	Ø	DAP	P	DAP
	Kʷak’ʷala			DAP	D	DAP
	Hindi		Ø	Ø	I	DA
	Besemah		Ø			
e	Armenian		Ø	DIAP	DIAP	DA
	Javanese		Ø	P	Ø	DAP
	Kʷak’ʷala					
	Hindi			Ø	I	DIA
	Besemah					
o	Armenian			DIAP	DIAP	DAP
	Javanese			DAP	IP	DAP
	Kʷak’ʷala					
	Hindi				Ø	Ø
	Besemah				Ø	DIP
i	Armenian				Ø	I!
	Javanese				Ø	DP
	Kʷak’ʷala				DP	DIAP
	Hindi					I!
	Besemah					DIP
u	Armenian					I!
	Javanese					DAP
	Kʷak’ʷala					DAP

4.1. The universality of the link between phonetic prominence and phonological sonority

As the table shows, the vowel that is most consistently distinguished phonetically from all other vowels along at least one of the measured dimensions is schwa. The only pairwise comparison involving schwa that is not manifested phonetically is the distinction between schwa and /i/ in Hindi. Nevertheless, although schwa is nearly universally differentiated from other vowels, the phonetic dimension(s) along which these distinctions are expressed differ between languages and even within languages between vowels paired with schwa.

The low vowel is differentiated in 11 of 18 pairwise comparisons involving vowels other than schwa. Of these 11 successful comparisons, 8 involve /a/ and the high vowels, which differ along at least one of the measured dimensions in four of the five languages (the exception being Besemah). The low vowel is less reliably distinguished (3 of 8 pairwise comparisons) from the mid vowels in four of the five languages (the exception being Kʷak'ʷala) with mid vowels.

Mid and high vowels are distinguished in 9 of 12 pairwise comparisons. Of the 9 distinctions, three entail a single distinguishing phonetic parameter. By comparison, 9 of the 11 pairwise distinctions involving a low vowel are conveyed by at least three phonetic properties.

Results point to uniformity among classes of vowels sharing height features. The two mid vowels are thus not phonetically differentiated along the measured dimensions in any of the languages. Furthermore, the two high vowels are only distinguished in prominence in one of the five languages, Kʷak'ʷala.

4.2. Language specificity in the phonological sonority of schwa and its phonetic properties

In the current data, differences between languages in the phonological status of schwa in the stress system do not correlate with interlanguage variation in the phonetic prominence of schwa relative to other vowels. Rather the dominant pattern is for schwa to be phonetically less prominent than other vowels regardless of its phonological behavior. Of the three languages with mid vowels, both the one in which schwa behaves parallel to other vowels with respect to stress placement, Hindi, and the two in which schwa rejects stress, Javanese and Armenian, have a schwa that is phonetically weaker than both mid vowels along at least two phonetic dimensions. The phonetic strength of the high vowels relative to schwa is not as consistent across languages, but this variation is not predictable from the phonological behavior of schwa. In one language with a schwa that attracts stress, Besemah, and in two languages with a schwa that rejects stress, Javanese and Kʷak'ʷala, high vowels are more prominent than schwa according to at least two phonetic parameters. In Hindi, which treats schwa like other vowels for stress, /i/ is not more prominent than schwa and /u/ is actually less intense than schwa. Perhaps more surprisingly, in Armenian, which avoids stress on schwa, there is no phonetic property among those measured that predicts the light status of schwa relative to the two high vowels. In fact, schwa has greater maximum intensity than both high vowels in contradiction of the sonority hierarchy. The failure of the measured parameters to distinguish schwa from the high vowels in Armenian in the correct direction raises questions about Gordon's (2002, 2006) hypothesis that phonological weight distinctions are predictable from acoustic properties or from perceptual properties ultimately derived from the acoustic signal via auditory transforms. The present work suggests that it might be necessary to explore an alternative hypothesis that syllable weight, and perhaps more generally, sonority, is at least partially grounded in speech production. Under this view, the reduced sonority of schwa could be to some degree attributed to the proximity of schwa to the tongue's rest position, an articulatory setting that would require less physical effort to achieve than more peripheral vowel qualities. A potential complication for this effort-based approach to sonority is the fact that high vowels may require greater tongue displacement from the rest position than low vowels, even though high vowels rank lower in phonological sonority than low vowels. It is conceivable that a measure of effort that penalizes jaw movement more than tongue movement due to the greater mass of the jaw could be invoked to account for the greater sonority of low vowels relative to high vowels (see Mooshammer et al 2007 for an overview of articulatory studies of jaw height). Appealing to effort-based considerations, of course, predicts that sonority could be context-sensitive. For example, one might expect high vowels to be less sonorous than low vowels in coronal and velar contexts where less movement is required to produce a high vowel than in bilabial contexts, where a greater articulatory excursion is necessary to produce a low vowel than a high vowel.

4.3. Assessing the robustness of different phonetic correlates as a predictor of sonority

Excluding cells representing the intersection of vowels of equivalent phonological height, there are 51 possible pairwise comparisons of vowels in the five examined languages. Of these 51 pairs, duration and perceptual energy each distinguish 32, acoustic energy distinguishes 27, and maximum intensity differentiates 22. Interestingly, of the 22 distinctions made by maximum intensity, three are cases in which schwa, a vowel of low phonological sonority, has greater intensity than a vowel ranking higher in phonological sonority, /i/ and /u/ in Hindi and /i/ in Armenian. (It may be noted that an additional reversal involving schwa and the high back lax vowel based on acoustic energy was found in Hindi but is not included in table 36, which excludes the lax vowels of Hindi.) It thus appears that the close link between maximum intensity and vocal tract aperture is adept at predicting phonological sonority distinctions based on vowel height, but is less successful in predicting sonority differences between peripheral and central vowels.

There is, however, no other single parameter that adequately predicts the sonority distinction between peripheral vowels and schwa. Duration is the most consistent property differentiating schwa from other vowels, being used in 18 of 21 total vowel distinctions involving schwa across the five languages. Yet, as mentioned above, duration fails to distinguish schwa from either /i/ or /u/ in Armenian, even though schwa is demonstrably lower in sonority than the high vowels in Armenian on the basis of its stress system. Perceptual energy and acoustic energy differentiate only 14 and 13, respectively, of the 21 total vowel pairs involving schwa and their success is likely attributed in large part to the fact that both measures are integrated over time and thus receive a boost in longer vowels.

In summary, the failure of any single phonetic property to accurately predict all aspects of sonority scales for vowels suggests that phonological sonority may not be quantifiable along any single dimension, at least in the case of vowels, but rather may reflect some weighted aggregate of multiple phonetic factors, a view espoused by Ohala and colleagues in earlier work (Ohala 1992, Ohala & Kawasaki-Fukumori 1997). The relevant parameters for predicting sonority may thus be multidimensional in nature, encompassing some, as yet undiscovered, combination of acoustic, perceptual, and articulatory properties.

5. Conclusions

The primary goal of this paper was to explore the hypothesis that the phonological status of schwa as a lower sonority vowel than peripheral vowels is predictable on phonetic grounds. Results from five languages with schwa indicate that, although sonority distinctions in individual languages are typically predicted by at least one phonetic parameter, there is no single parameter that predicted sonority distinctions across all languages. Although schwa is characteristically less prominent than other vowel qualities along multiple phonetic dimensions, there emerged several instances in which schwa not only failed to display reduced prominence relative to other vowels, but actually was characterized by greater prominence. Thus, in two languages (Armenian and Hindi), schwa was associated with greater peak intensity values than at least one of the peripheral vowels. Most strikingly, one of these sonority reversals even occurred in a language (Armenian) that treats schwa as phonologically lighter than peripheral vowels in its stress system. Although appealing to a property other than maximum intensity as a correlate of sonority eliminates instances of reversals in the data, there is still no single phonetic property that correctly predicts the lower sonority of schwa in all languages. The dimensions that are most successful cross-linguistically in our data set, duration and perceptual energy, do not distinguish schwa from high vowels in Armenian. Furthermore, these two properties are only partially successful in making sonority distinctions between the peripheral vowels. The present results thus underscore the challenges confronted by any model of the phonetics-phonology interface that posits a single phonetic dimension underlying phonological sonority.

Acknowledgments

The authors gratefully acknowledge the many speakers providing the recordings discussed in this paper. We thank the several K^{w}ak'^{w}ala consultants (in particular, the late Lorraine Hunt, Beverly Lagis, Margaret Pu'tsa Hunt, Chief Robert Joseph, Daisy Sewid-Smith, and Pauline Alfred) contributing the K^{w}ak'^{w}ala forms and generalizations about stress, many of which appear for the first time in print here. We would also like to thank Universitas Sriwijaya for their support in collecting the Besemah data as well as many Besemah speakers from the village of Karang Tanding for volunteering to take part in this research. We also extend our gratitude to Marc Garellek for assisting in recording one of the Javanese consultants. Furthermore, we acknowledge the generous financial support of NSF grant BCS0343981 to Matthew Gordon, SSHRC grant *Kan's kwak'wale' xan's yak'anda's! Let's keep our language alive!* to Patricia A. Shaw (in partnership with U'mista Cultural Center (the late Andrea Sanborn, Director), Lilawagila School, Kingcome Inlet (Mike Willie), 'Namgis First Nation (Chief William T. Cranmer), and T'lisalagi'lakw School, Alert Bay), as well as the financial, intellectual and institutional support of the 2008 InField Institute held at University of California at Santa Barbara and directed by Carol Genetti. Finally, thanks to Steve Parker and two anonymous reviewers for their many constructive and useful comments on an earlier draft of this paper.

References

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

Bladon, Anthony & Björn Lindblom. 1981. Modeling the judgment of vowel quality differences. *Journal of the Acoustical Society of America* 69, 1414-1422.

Boas, Franz. 1947. *Kwakiutl Grammar with a Glossary of the Suffixes.* Edited by Helene Boas Yampolsky with the collaboration of Zellig S. Harris. Transactions of the American Philosophical

Society 37:3. 199/202-377.
Boersma, Paul & David Weenink. 2010. *Praat: Doing phonetics by computer* (version 5.1.42) (www.praat.org).
Clynes, Adrian & C. Rudyanto. 1995. Javanese. In Darrel T. Tryon (ed.), *Comparative Austronesian Dictionary: An Introduction to Austronesian Studies*. Berlin: Mouton de Gruyter.
Davies, John. 1981. *Kobon*. [Lingua Descriptive Studies, Vol. 3]. Amsterdam: North-Holland Publishing Company.
de Boer, Bart. 2011. First formant difference for /i/ and /u/: A cross-linguistic study and an explanation. *Journal of Phonetics* 39, 110-114.
de Lacy, Paul. 2002. The Formal Expression of Markedness. Doctoral dissertation. University of Massachusetts Amherst.
de Lacy, Paul 2004. Markedness conflation in optimality theory. *Phonology* 21, 145–199.
Dixit, R. P. 1963. *The Segmental Phonemes of Contemporary Hindi*. MA thesis, University of Texas at Austin.
Gordon, Matthew & Carlos Nash. 2007. *Cricket* (www.linguistics.ucsb.edu/faculty/gordon/projects.html)
Gordon, Matthew. 2002. A phonetically-driven account of syllable weight. *Language* 78, 51-80.
Gordon, Matthew. 2006. *Syllable Weight: Phonetics, Phonology, Typology*. New York: Routledge.
Grubb, David. 1977. *A Practical Writing System and Short Dictionary of Kwakw'ala (Kwakiutl).* Mercury Series, Canadian Ethnology Service Paper no. 34. Ottawa: National Museums of Canada.
Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.
Herrfurth, Hans. 1964. *Lehrbuch des modernen Djawanisch*. Leipzig: VEB Verlag.
Horne, Elinor. 1974. *Javanese-English Dictionary*, New Haven and London: Yale University Press.
Itkonen, Erkki. 1955. Über die Betonungsverhältnisse in den finnisch-ugrischen Sprachen. *Acta Linguistica Academiae Scientiarum Hungaricae* 5, 21–34.
Jany, Carmen, Matthew Gordon, Carlos Nash & Nobutaka Takara. 2007. How universal is the sonority hierarchy?: A cross-linguistic acoustic study. *Proceedings of The XVIth International Congress of Phonetic Sciences*, pp. 1401-1404.
Kelkar, Ashok R. 1968. *Studies in Hindi-Urdu I: Introduction and Word Phonology*. Poona: Deccan College.
Kenstowicz, Michael. 1997. Quality-sensitive stress. *Rivista di Linguistica* 9, 157–188.
Kodzasov, S. & I. Muravyova. 1978. Stress in the Alutor language. *Estonian Papers in Phonetics* 1978, 47–8.
Lehiste, Ilse. & Gordon E. Peterson. 1959. Vowel amplitude and phonemic stress in American English. *Journal of the Acoustical Society of America* 31, 428-435
Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, Mass.: MIT Press.
McDonnell, Bradley. 2008. A Conservative Vowel Phoneme Inventory of Sumatra: The Case of Besemah. *Oceanic Linguistics* 47(2), 409–432.
Moore, Brian & Glasberg, Brian. 1983. Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *Journal of the Acoustical Society of America* 74, 750-753.
Mooshammer, Christine, Hoole, Philip, & Anja Geumann. 2007. Jaw and Order. *Language and Speech* 50, 145-176.

Nedzelnitsky, V. 1980. Sound pressures in the basal turn of the cat cochlea. *Journal of the Acoustical Society of America* 68, 1676-1689.

Ohala, John J. 1992. Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In M. Ziolkowski, M. Noske, K. Deaton (es.), *CLS 26: Papers from the 26th Regional Meeting of the Chicago Linguistic Society*, *Vol. 2: The Parasession on the Syllable in Phonetics and Phonology*, pp. 319–338. Chicago: Chicago Linguistic Society.

Ohala, John J. & Haruko Kawasaki-Fukumori. 1997 Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In S. Eliasson & E. H. Jahr (eds.), *Language And Its Ecology: Essays In Memory Of Einar Haugen* [*Trends in Linguistics. Studies and Monographs, Vol. 100*], pp. 343-365. Berlin: Mouton de Gruyter.

Ohala, Manjari. 1977. Stress in Hindi. In Larry Hyman (ed.), *Studies in stress and accent*. [*Southern California Occasional Papers in Linguistics* 4], pp. 327-338. Los Angeles: USC Department of Linguistics.

Ohala, Manjari. 1999. *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press, pp. 100–103.

Parker, Stephen. 2002. *Quantifying the Sonority Hierarchy*. Doctoral Dissertation. University of Massachusetts Amherst.

Parker, Stephen. 2008. Sound level protrusions as physical correlates of sonority. *Journal of Phonetics* 36, 55-90.

Patterson, R., I. Nimmo-Smith, D. Weber & R. Milroy. 1982. The deterioration of hearing with age: frequency selectivity, the critical ratio, the audiogram, and speech threshold. *Journal of the Acoustical Society of America* 72, 1788-1803.

Peterson, Gordon E. & Ilse Lehiste. 1960. Duration of syllable nuclei in English. *Journal of the Acoustical Society of America* 32, 693–703.

Plomp, Reinier. 1964. Rate of decay of auditory sensation. *Journal of the Acoustical Society of America* 36, 277-282.

Shaw, E. A. G. 1974. The external ear. In W. D. Keidel and W. D. Neff (eds.), *Handbook of Sensory Physiology*, vol. 5, pp. 455-490. Berlin: Springer.

Shaw, Patricia A. 2009. Default-to-Opposite Stress in Kʷak'ʷala: Quantity Sensitivities in a Default-to-Right System. Paper presented at SSILA, University of California at Berkeley.

Silverman, Daniel. 2011. Schwa. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell Companion to Phonology*. Wiley-Blackwell.

Vaux, Bert. 1998, *The Phonology of Armenian*, Clarendon Press, Oxford.

Viemeister, Neal. 1980. Adaptation of masking. In G. van den Brink and F. A. Bilsen (eds.), *Psychophysical, Physiological and Behavioural Studies in Hearing, Proceedings of the 5th International Symposium on Hearing*, pp. 190-198. Delft, Netherlands: Delft University Press.

Wilson, J. P. 1970. An auditory afterimage. In R. Plomp and G. F. Smoorenberg (eds.), *Frequency Analysis and Psychophysics of Hearing*, pp. 303-315. Leiden: Sijthoff.

Wilson, Stephen A. 1986. Metrical structure in Wakashan phonology. *Proceedings of Berkeley Linguistics Society* 12, 283-291.

Zec, Draga. 1998. *Sonority Constraints on Prosodic Structure*. PhD. Dissertation Stanford University, Palo Alto. (Published by Garland, New York)

Appendix: Corpora (target vowels in bold)

Hindi

ˈm**ə**nɪ	jewel
ˈg**ə**rdən	neck
ˈw**ɛ**ʃja	prostitute
ˈp**ɛ**nʃən	pension
ˈm**ɪ**lə	obtain
ˈb**ɪ**rla	rare
ˈb**ɔ**nə	dwarf
ˈk**ɔ**mtʃa	skimmer
ˈs**ɔ**nɔ	listen
ˈk**u**rsi	chair
ˈl**a**nə	bring
ˈs**a**rda	goddess
ˈl**e**nə	take
ˈk**e**lnə	play
ˈn**i**lə	blue
ˈk**i**rtən	mentioning, praising
ˈs**o**nə	sleep, gold
ˈʃ**o**rbə	broth
ˈs**u**nə	lonely
ˈm**u**rti	statue

Javanese

kəˈt**a**n	sticky rice
kəˈt**a**t	tight, constricting
ənˈt**a**s	bring in out of the rain
jənˈt**ə**n	cumin seed, caraway seed
bənˈt**ə**t	exact amount, completely filled
mənˈt**ə**s	well filled out
əmˈp**e**t	crowded
kəˈt**e**n	agile
gəˈt**e**n	hardworking, industrious
gənˈt**o**s	change, replacement
bəˈt**o**n	pit of a jackfruit
jəˈp**o**t	go away to avoid conversation
gəˈp**i**t	squeeze
kəˈt**i**s	(of voice) sweet and clear
səˈt**i**n	satin
kənˈt**u**t	fart
gəˈt**u**n	remorseful
bənˈt**u**s	bump

Armenian

ˌg**ɑ**ˈtɑ	kind of pastry
ˌh**ɑ**ˈsɑk	age
ˌg**ɛ**ˈtɑk	small river
ˌh**ɛ**ˈsɑn	grindstone
ˌg**o**ˈti	belt
ˌh**ɔ**ˈsɑnk	stream
ˌg**i**ˈtɑk	connoisseur
ˌh**i**ˈsun	fifty
ˌg**u**ˈtʰan	plough
ˌh**u**ˈsɑl	to hope
ˌg**ə**ˈtɑkʰ	you found
ˌh**ə**ˈʁi	pregnant

Kʷak'ʷala

səˈb**a**s	reply, echo
bəsˈb**a**s	to eat biscuits
dəˈn**a**s	inner bark of red-cedar
gəˈp**u**d	to unbutton, unwedge something
gəpˈst**u**d	to tuck in, to stuff up a hole; to plug it up
əpˈs**u**t	the opposite side
gəˈp'**i**d	to button up, to tip, slip money (etc.) to someone
səˈb**i**ɬ	sun rays striking floor
bəχˈs**i**s	to lance the foot
məˈɬ**i**k	salmon, sockeye

əˈn**i**s	aunt
nəpˈbəs	always throwing rocks
nəˈp’**ə**p	hair on chest

Besemah

p**a**ˈtah	snap
t**a**ˈtap	touch
p**i**ˈtuŋ	hold
t**i**ˈtu	this
p**u**ˈtih	white
t**u**ˈtus	strike
p**ə**ˈtaŋ	evening
t**ə**ˈtak	cut