An Acoustic Study of Georgian Vowels as a Function of Syllable Prominence
Submitted by Roy Becker

1. Introduction

The purpose of this study is to provide a preliminary yet well-controlled acoustic description of the vowel inventory of Georgian as a function of various prosodic positions. This study goes beyond a mere description of the inventory in establishing acoustic correlates of prosodic prominence.

Generally speaking, the Georgian 5-vowel inventory is the ‘canonical’ $i$-$e$-$a$-$o$-$u$, which is the most common inventory in the world’s languages. Like most other languages with this inventory type, all vowels are permitted and potentially contrastive in any syllable in the word. Syllable prominence (stress) is not contrastive in Georgian, and it seems that native speakers are not aware of relative prominence between syllables. Nonetheless, non-native listeners tend to have a strong sensation of relative syllable prominence in Georgian, and typically judge the first syllable of the word to carry primary stress, and also report sensation of secondary stress in odd-ordered non-final syllables. Depending on language background, non-native listeners (in particular English-speaking listeners) sometimes report sensation of primary stress elsewhere in the Georgian word if the word is sufficiently long.

Clearly, if it turns out that cross-linguistic correlates of stress, such as duration, magnitude of articulatory displacement (as reflected indirectly in the acoustic correlates of articulatory positions), intensity and pitch vary systematically, for all vowels, between different syllabic positions in Georgian, then Georgian speakers have in their grammar a syllable-prominence-dependent control mechanism, even if they are unaware of the notion of syllable stress and its presence in their language.

Among these correlates, pitch is the least reliable one, both cross-linguistically (the type of default stress-anchored melody varies substantially across languages) and methodologically (in order to avoid unreliable acoustic signal resulting from creaky-voice at utterance edges, it is essential to embed vowels in carrier sentences, with the inevitable result of sentential intonation patterns overriding any reliable within-word melody pattern). Hence this study focuses on the other correlates of stress, namely duration, displacement magnitude and intensity.

In order to obtain vowels in various contexts, this study used nonsense words embedded in carrier sentences read aloud from printed text (luckily, in Georgian, grapheme-to-phoneme relation is a straightforward injective function, leaving very little room to text misinterpretation). Use of nonsense...
words is of course somewhat problematic, for it primes the speaker to pay special attention to any sound, thus flattening attention hierarchies embedded in the grammar normally applied on familiar words, with the inevitable result of flattening prominence differences, which is precisely the object of research in the current study. However, just like in any other language, the Georgian lexicon (and even more so the commonly-used lexicon) does not contain a comprehensive set of entries that matches the needs of the current study, and clearly, using a mixed set of target words, some real and some nonsensical, would introduce a contaminating effect of partial lexical familiarity. It was therefore decided to neutralize lexical familiarity altogether and use nonsense words throughout.

It should be noted that, given the clear speech mode typically used by a speaker when pronouncing nonsense words, any significant prominence differences that do emerge are most likely reflecting even stronger differences in the ‘normal’ grammar. Moreover, the combination of the exclusively segmental-phonemic nature of the Georgian orthography with native-speaker’s unawareness of syllable prominence relations, perhaps primes the speaker towards excessive attention to phoneme realization, but definitely not towards syllable prominence, which is the object of inspection here, that is, the plausible flattening of syllable prominence hierarchies is an artifact of excessive attention to the phonemic contents of syllables, rather than to the syllables themselves.

The current study is based on elicitation with one female speaker, Manana Batashvili. I am indebted to Manana for her great patience and consistency, in the somewhat tedious and monotonous task of nonsense-word elicitation. As will become clear, interpretable and consistent tendencies did emerge despite the limited amount of data. This would not have been achieved without Manana’s cooperative attitude, of which she is so unique.

2. Method

2.1 Segmental Context of the vowels

Since the primary articulator of vowels is the tongue dorsum, establishing dorsal behaviour correlates of syllable prominence requires limiting contextual dorsal co-articulation to the minimum. Therefore it was decided that the vowels be flanked by labial consonants. In addition, in order to best interpret the acoustic signal for articulatory magnitude, duration and intensity, it is essential that the vowels have (a) constant and consistent laryngeal quality, (b) constant and consistent number of resonating cavities and points of aperture, and (c) straightforward boundaries with adjacent consonants. The first requirement disfavours spread- or constricted-glottis consonants (ruling out the aspirated and ejective labial stops), while the second one disfavours nasalization (ruling out the labial nasal) and the third one disfavours voiced fricatives and approximants (ruling out the labio-dental
voiced fricative). By elimination, this leaves us with /b/, which was chosen as both pre- and post-vocalic consonant throughout the experiment, yielding an immediate sequence \(bVb\).

As co-articulatory effects may be caused by vowels at adjacent syllables as well, it was essential to limit such effects by confining the vocalic contents of these syllables to the vowel which is least co-articulatorily aggressive. Based on previous exposure to many elicitations, it was assumed that the low vowel \(a\) was the most suitable, as this vowel seems to involve only jaw lowering, which is counter-balanced by the labial stops intervening between the target vowel and the context vowels, yielding the sequence \(abVba\). Moreover, as it is well-known that regressive co-articulation is significantly more robust than progressive co-articulation, it was decided to augment the post-vocalic context by another labial (nasal) stop, yielding the sequence \(abVbma\) (\(bm\) is a grammatical consonantal cluster in Georgian). Finally, as it was required that target vowels appear in word initial syllables in certain tokens, the preceding syllable \(a\) had to appear word-finally in such elicitations. It was decided that the syllable preceding the target syllable would always be \(ma\), enabling the use of the ergative suffix –\(ma\) as the ending of the word preceding the target word in the carrier sentence. Thus the target vowel was always embedded in the sequence \(ma(#)bVbma\). It was assumed that such a segmental context would incur relatively minimal co-articulation to the target vowels.

2.2 Syllable-prominence context of the vowels

The experiment was designed to find acoustic correlates of two hypotheses regarding relative syllable prominence in Georgian, namely (a) that Georgian has word-initial primary stress, and (b) that Georgian has secondary stress on oddly-ordered non-initial syllables. In practice, this was translated to the following hypotheses: (a) that comparing first-syllable to second-syllable target vowels would yield significant acoustic differences, and (b) that third-syllable target vowels would pattern consistently between these two extremes.

Thus, the target vowels in this experiment were always embedded in either the first, second or third syllable in three-, four- or five-syllable words, respectively. The decision to have two syllables after the target vowel’s syllable was made in order to test yet another hypothesis made elsewhere, namely that in 5-syllable words it is the third syllable, rather than the first, that carries primary stress.

It was assumed that at distances beyond one syllable from the target vowel the segmental and prosodic contents have only negligible effects, if any, on the target vowel. In order to deviate from the labial-stop – low-vowel routine, it was decided to use \(de\) as the last syllable in all target words, and \(ge\) as the initial syllable in 5-syllable target words, yielding the sequence \((ge)mabVbma\).

2.3 Carrier sentence and higher prosodic context
As the target vowel was evidentially embedded in nonsense words made by a sequence of
alternating stops and vowels, it seemed reasonable to embed these nonsense words in the semantic
context of reporting baby babbling, where such sentences would actually make sense. Thus, the
carrier sentence was always: bavhsvma X tkva (‘The baby said X’), where X is the target nonsense
word. Furthermore, in order to ensure flowing and natural pronunciation of the target word in this
context, it seemed reasonable to familiarize Manana with the word by pronouncing it once before the
carrier sentence, yielding the form: X, bavshvma X tkva (‘X, the baby said X’). However, it turned
out that Manana pronounced the target nonsense words naturally and flawlessly to begin with, which
enabled comparison of the two instances of the target vowel as a function of higher prosodic position,
as the first instance of the nonsense word spanned a whole phonological phrase, whereas the second
instance of the nonsense word was only one component of a multi-component phonological phrase.

2.4 The tokens and the elicitation procedure

The following table lists all the nonsense words used in the experiment:

<table>
<thead>
<tr>
<th></th>
<th>1st syllable</th>
<th>2nd syllable</th>
<th>3rd syllable</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>bibmade</td>
<td>mabibmade</td>
<td>gemabibmade</td>
</tr>
<tr>
<td>e</td>
<td>bebmade</td>
<td>mabebebmade</td>
<td>gemabebmade</td>
</tr>
<tr>
<td>a</td>
<td>babmade</td>
<td>mababmade</td>
<td>gemababmade</td>
</tr>
<tr>
<td>o</td>
<td>bobmade</td>
<td>mabobmade</td>
<td>gemabobmade</td>
</tr>
<tr>
<td>u</td>
<td>bubmade</td>
<td>mabubmade</td>
<td>gemabubmade</td>
</tr>
</tbody>
</table>

Each of the 15 token-embedding sentences was elicited twice. In addition there were 10
dummy tokens. Thus the total number of sentences elicited was 40. The first two sentences and
the last sentence were dummy tokens, and the rest of the sentences were randomly ordered.

The 40 sentences were printed on small sheets of paper, each sentence on a separate page,
using the Georgian (Mkhedruli) alphabet in large fonts (the target words were printed in boldface
fonts). Latinized-script transcription, in smaller fonts, was printed at the bottom of each page in
order to facilitate the experimenter’s control. Following is an example of a token sheet:
Manana was told that the purpose of the elicitation was phonetic research, and was briefly familiarized with the notion of the use of nonsense words in phonetic research in general and the form of the sentence tokens used here in particular. The ultimate goal of the elicitation, i.e. the pronunciation of vowels in various syllable prominence context, was not disclosed to her.

Seated in a sound-proof booth in front of a Shure SM10A microphone, Manana read the sentences aloud, one after the other, in four blocks of ten sentences. Sentences were immediately re-elicited whenever Manana’s pronunciation was hesitant, interrupted or otherwise mistaken.

The sentences were recorded digitally using PRAAT at a sampling frequency of 22,050Hz. Annotation and analysis were performed using PRAAT as well.

2.5 Acoustic Analysis

With very few exceptions, the recorded material proved very easy for acoustic analysis. Except for a handful of instances where the post-vocalic /b/ was rendered as a narrow bilabial approximant, requiring a more sophisticated segmentation algorithm, segmentation of the target vowel from the flanking bilabial stops was always straightforward using abrupt changes in the aligned waveform, spectrogram and formant trackers. Segmentation between the preceding stop and the vowel was made at the burst, which might be theoretically inadequate but nevertheless legitimate for the current purposes. For each instance of a target vowel the following measurements were recorded:

**Duration:**
- a) Duration of the vowel.
- b) Duration of the post-vocalic *made* sequence.

**Spectral data:**
- c) The aperture formant (F1) at the 20th, 50th and 80th percentile of the vowel’s duration.
- d) The formant corresponding to the front cavity (F3 for *i*, F2 elsewhere) at the 20th, 50th and 80th percentile of the vowel’s duration.

**Intensity:**
- e) Intensity at the 20th, 50th and 80th percentile of the vowel’s duration.
- f) Reference intensity of the stressed vowel *a* at the point 80msec after the burst of the bilabial stop in the word *bavšvma*.

---

1 It can be easily claimed that the burst itself as well as the initial portion of the transition, in particular when this portion includes some frication, is in fact part of the stop and not of the vowel. This would require that segmentation be performed elsewhere had the consonantal context been parameterized. However, as the prevocalic consonant is constant and is not part of the analysis, the burst is a legitimate temporal anchoring event.

2 Most similar descriptive studies use F1 and F2 as co-ordinates for all vowels. However, it is well known that, at the high front periphery, F2 and F3 cavity affiliations are reversed, and F2 is not informative about the size of the front cavity (and is therefore unreliable for estimation of the magnitude of articulatory displacement).
3. Results

By and large, Manana’s pronunciation was very consistent, in the sense that the two repetitions of each token were very similar to one another. Hence, despite the very limited amount of data, averages of acoustic readings from the two repetitions are reliable. Surprisingly, most results are very conclusive regarding the presence or absence of acoustic correlates of syllable prominence, either reaching very high statistical significance ($p<0.01$) or none at all ($p>0.3$).

3.1 Duration

3.1.1 Results

The following table summarizes the average duration readings (in ms.) for the each of the vowels and the post-vocalic residue in the target word (bmade):

<table>
<thead>
<tr>
<th></th>
<th>1st syl.</th>
<th>2nd syl.</th>
<th>3rd syl.</th>
<th>1st syl.</th>
<th>2nd syl.</th>
<th>3rd syl.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>vowel</td>
<td>bmade</td>
<td>vowel</td>
<td>vowel</td>
<td>bmade</td>
<td>vowel</td>
</tr>
<tr>
<td>i</td>
<td>0.141</td>
<td>0.464</td>
<td>0.082</td>
<td>0.474</td>
<td>0.101</td>
<td>0.471</td>
</tr>
<tr>
<td>e</td>
<td>0.151</td>
<td>0.472</td>
<td>0.091</td>
<td>0.471</td>
<td>0.117</td>
<td>0.518</td>
</tr>
<tr>
<td>a</td>
<td>0.147</td>
<td>0.428</td>
<td>0.100</td>
<td>0.473</td>
<td>0.126</td>
<td>0.456</td>
</tr>
<tr>
<td>o</td>
<td>0.152</td>
<td>0.499</td>
<td>0.079</td>
<td>0.428</td>
<td>0.122</td>
<td>0.458</td>
</tr>
<tr>
<td>u</td>
<td>0.175</td>
<td>0.464</td>
<td>0.086</td>
<td>0.492</td>
<td>0.105</td>
<td>0.439</td>
</tr>
</tbody>
</table>

Remarkably, the duration of the bmade post-vocalic residue shows no effect of any syllabic or prosodic contrast (despite, for example, a difference in the total number of syllables in the target word between different syllable conditions). Without exception, all matched-pair t-test results reached a non-significant probability score higher than 0.3. That the absence of any such effect is a result of highly consistent pronunciation, rather than an artifact of ‘messy’ and inconsistent pronunciation, is evident from the fact that the standard deviation of bmade duration, calculated over all the 60 instances, was 44ms, less than 10% of the average duration of 461ms.

Conversely, syllabic and prosodic contrasts show unambiguous effects on the duration of the target vowel: The target vowel was significantly longer in the first syllable than in the second syllable, both in the full P-phrase condition ($p<0.001$) and in the P-phrase component condition ($p<0.001$). Similarly, the target vowel was significantly longer in the first syllable than in the third syllable, both in the full P-phrase condition ($p<0.01$) and in the P-phrase component condition ($p<0.002$). Interestingly, the target vowel was significantly longer in the third syllable than in the second syllable in the full P-phrase condition ($p<0.004$) but not in the P-phrase component condition ($p>0.7$). The effect of higher prosodic condition seemed overwhelming, with full-P-phrase-embedded target vowels being significantly longer than the corresponding P-phrase-component-embedded target vowels ($p<0.001$). However, further analysis discovered that, while this is indeed...
true for first-syllable and third-syllable target vowels ($p<0.002$ for both), there was no significant duration difference for second-syllable target vowels between the two prosodic conditions ($p>0.65$).

It is worth mentioning that the prosodic condition also had an unequivocal effect on the duration of the phonetic material preceding the target vowel: comparing the full P-phrase condition vs. the P-phrase component condition vis-à-vis the duration of pre-vocalic $ab$ (4-syllable target words) and $emab$ (5-syllable target words) showed significant difference ($p<0.02$ for $ab$ in 4-syllable target words, $p<0.0004$ for $emab$ in 5-syllable target words, $p<0.002$ for the combination of the two together)$^3$. It is beyond the scope of the current paper to explain why varying the higher-level prosodic position had duration effect on the phonetic material preceding the target vowel and on the vowel itself but not on the post-vocalic residue.

### 3.1.2 Discussion

If vowel duration is a universal correlate of syllable prominence, or stress, which seems to be cross-linguistically true, then the results imply the following:

- The first syllable is always more prominent than the second or third syllable, implying word-initial stress.
- The third syllable is more prominent than the second syllable in the full P-phrase condition but not in the P-phrase component condition, that is, at high prosodic positions (special focus), the third syllable is assigned secondary stress, but at lower prosodic positions the third syllable is virtually identical to the second syllable, both being unstressed.
- At high prosodic positions, the stressed syllable is even further strengthened.

Notice that these results are counter-evidence to the claim that, in 5-syllable words, it is the third, rather than the first syllable, that is stressed. This is particularly evident from the fact that, at the lower prosodic position, there was practically no duration difference between second-syllable and third-syllable vowels. As it is probably non-controversial that the second syllable is never stressed, the only interpretation of the duration identity is that the third syllable is not stressed either.

Whereas most of the claims above should be qualified substantially if we assume that duration differences are merely an artifact of domain-initial strengthening, rather than stress, thus rendering the word-initial syllable longer and the P-phrase-initial syllable even longer (higher domain = greater lengthening), the fact that, in the full P-phrase condition, there was a consistent

---

$^3$ Notice that the word initial consonant was never subject to duration measurement. This is because (a) it is impossible to determine the beginning of an utterance-initial stop, which appears in 3-syllable and 5-syllable target words, and (b) utterance-initial consonants are sporadically excessively lengthened, which was indeed the case for quite a few of the utterance initial $m$ in 4-syllable target words.
difference between the second and the third syllable, none of which is domain initial, must imply that duration is indeed a correlate of syllable prominence, i.e. stress.

3.2 Magnitude of articulatory displacement

3.2.1 Results

The following table summarizes the average frequency readings (in Hz) at the 50th percentile of the duration of the vowel for the aperture formant AF (=F₁) and the front cavity formant FCF (=F₃ for i, =F₂ otherwise) for the each of the vowels:

<table>
<thead>
<tr>
<th></th>
<th>target word as full P-phrase</th>
<th>target word as component of P-phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st syl.</td>
<td>2nd syl.</td>
</tr>
<tr>
<td></td>
<td>AF</td>
<td>FCF</td>
</tr>
<tr>
<td>i</td>
<td>377</td>
<td>3177</td>
</tr>
<tr>
<td>e</td>
<td>580</td>
<td>2148</td>
</tr>
<tr>
<td>a</td>
<td>835</td>
<td>1453</td>
</tr>
<tr>
<td>o</td>
<td>568</td>
<td>1083</td>
</tr>
<tr>
<td>u</td>
<td>404</td>
<td>1017</td>
</tr>
</tbody>
</table>

At the absence of a fixed reference point and the non-trivial relationship between articulatory displacement and the numeral differences in the Hz scale (or any other frequency scale), it is impossible to run meaningful significance tests on these data. However, the following qualitative points should provide a clear picture of the magnitude of articulatory displacement in the various conditions:

- **How low is the aperture formant for the high vowels i and u** (the lower the formant, the greater the constriction and the greater the displacement)? Evidently, AF is always lowest, for both i and u in both prosodic conditions, when these vowels appear in the first syllable. Likewise, AF is lower for high vowels in a third syllable than in a second syllable, except for the case of u in the P-phrase component condition, where AF was lower in the second syllable than in the third syllable. These results imply a displacement hierarchies 1st>3rd>2nd (full P-phrase condition) and 1st>3rd=2nd (P-phrase component condition).

- **How high is the aperture formant for the low vowel a** (the higher the formant, the greater the aperture and the greater the displacement)? The displacement hierarchies emerging given AF values for a are 3rd>1st>2nd (full P-phrase condition) and 1st>3rd=2nd (P-phrase component condition).

- **How low is the front cavity formant for the back-rounded vowels u and o** (the lower the formant, the backer the dorsum and the greater the displacement)? The displacement hierarchies emerging given FCF values for u and o are 3rd>1st>2nd in (full P-phrase condition) and 1st>3rd>2nd (P-phrase component condition).
- **How high is the front cavity formant for the front-spread vowels *i* and *e* (the higher the formant, the fronter the dorsum and the greater the displacement)?** The displacement hierarchy emerging given FCF values for the front vowels is strictly 1st>3rd>2nd (both vowels, both prosodic conditions). The third-syllable front vowel is closer to the extremely displaced first-syllable front vowel in the full P-phrase condition, and closer to the more centralized second-syllable front vowel in the P-phrase component condition.

- **What is the total area (in Hz²) of the pentangle created by the points representing the vowels in a AF~FCF plot (the greater the area, the greater the overall displacement for the whole inventory)?** The following figure provides area sizes and visual plots for the pentangles in the six conditions (3 syllable conditions x 2 P-phrase conditions):

![Image of AF~FCF plot]

The precise quantitative values mean very little, if at all, but qualitatively speaking, there are clearly 3 area-size types: Large (first syllable in both prosodic conditions and third syllable in the full P-phrase condition), medium (second syllable, full P-phrase condition) and small (second and third syllable in P-phrase component condition – notice that both pentangles are virtually identical).

### 3.2.2 Discussion

To a great extent, the spectral data results are in agreement with the duration results: Articulatory displacement is always much greater in the first syllable than in the second syllable, and is the same for the second syllable and the third syllable in the P-phrase component
condition. Two major differences from the duration results are (a) in the full P-phrase condition, articulatory displacement does not seem to differ consistently between first and third syllable, and (b) the effect of higher prosodic condition is not always clear – the magnitude of articulatory displacement is indeed smaller in the P-phrase component condition than in the full P-phrase condition for the second syllable, and a similar difference at a much larger scale holds also for the third syllable, but no such difference exists for the first syllable.

Overall, however, given the assumption that articulatory displacement is a cross-linguistic correlate of syllable prominence, or stress, then again it seems that the first syllable is always stressed, the second syllable is always unstressed, and the third syllable exhibits dual behaviour: like the second syllable (i.e. unstressed) in the P-phrase component condition, and like the first syllable (i.e. stressed) in the full P-phrase condition.

3.3 Intensity

3.3.1 Results

The following table summarizes the average target-to-reference intensity ratios (in dB) at the 50th percentile of the duration of the vowel relative to the reference point (80msec into the stressed a after the stop burst in šavšvma):

<table>
<thead>
<tr>
<th>target word as full P-phrase</th>
<th>target word as component of P-phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st syl.</td>
<td>2nd syl.</td>
</tr>
<tr>
<td>i</td>
<td>-2.79</td>
</tr>
<tr>
<td>e</td>
<td>8.86</td>
</tr>
<tr>
<td>a</td>
<td>6.43</td>
</tr>
<tr>
<td>o</td>
<td>11</td>
</tr>
<tr>
<td>u</td>
<td>-5.07</td>
</tr>
</tbody>
</table>

Somewhat surprisingly, it turns out that there are no significant differences between the intensity ratios across syllable conditions, and the emerging picture is rather ‘messy’. Thus, while the intensity hierarchy for e in the full P-phrase condition matches the duration and displacement hierarchies (1st>3rd>2nd), precisely the reverse hierarchy exists for u in the same prosodic condition (2nd>3rd>1st). Only the difference between the louder first syllables and softer second syllables in the P-phrase component condition seems to be consistent and is marginally statistically significant (p<0.05), but given the overall unclear picture, the significance here might be coincidental.

It stands to reason that intensity, like pitch, is more correlated with sentential intonation patterns, which mask any word-internal patterns. This possibility is remarkably corroborated by the highly significant difference between the intensity ratios of the two prosodic conditions: target vowels at the full P-phrase condition were always much louder than their corresponding target...
vowels in the P-phrase component condition, for each of the syllable positions \((p=0.009, 0.011, 0.007\) respectively for first, second and third syllables) and for all positions together \((p<0.000001)\).

3.3.2 Discussion

The non-significant intensity results imply either that the experimental format, the type of measurement or the amount of data, or some of these, are not sufficient/suitable for testing the manifestation of cross-linguistic correlation between intensity and syllable prominence, or that this correlation does not hold for Georgian. It is very likely that the first implication is in fact true, since intensity seems to pattern more consistently with sentence-level prosody than with word-level prosody, and the consistent intensity differences between the two higher-level prosodic conditions demonstrates that Georgian is no exception to this.

4. General discussion and conclusions

4.1 Vowel qualities

There is hardly anything particular or ‘striking’ about vowel qualities in Georgian as emerging from the current set of data. Perhaps the most interesting finding is that, unlike most other ‘canonical’ 5-vowel inventories (but in a manner similar to e.g. Czech), the front mid vowel has a somewhat higher first formant (i.e. greater aperture) than the back mid vowel\(^4\). While it is difficult to interpret the front cavity formant (F2) of the low vowel \(a\), which is around 1450Hz, it is worth noting that the average difference in F2 between the low vowel \(a\) and the mid back vowel \(o\) is only about 400Hz, substantially smaller than the corresponding difference in inter-labial \(a\) and \(o\) of e.g. Hebrew female speakers (more than 550Hz). This fact, together with at least one phonological pattern in Georgian grouping \(a\) with back vowels, namely that the lateral consonant surfaces as velarized before \(a\), as it does before the rounded back vowels \(u\) and \(o\), might indicate that the low vowel in Georgian is indeed back and not central. In other words, the accurate phonetic description (and transcription) of the inventory is probably /i ɛ a ɔ u/, rather than the universally most common /i ɛ a ɔ u/. It is likely that precisely this retraction of the low vowel leaves more space at the lower-fronter periphery of the vowel space, allowing a lower rendition of the mid front vowel.

4.2 Phonetic correlates of stress

Three cross-linguistic phonetic correlates of syllable prominence, or stress, were examined in this experiment: duration, magnitude of articulatory displacement and intensity. Intensity provided evidence for nothing other than a higher prosodic domain effect. Duration and

\(^4\)In the vast majority of languages it is the back vowel that has a slightly higher first formant, both in the mid and high vowel pairs. This universal generalization definitely holds for the high vowels in Georgian as apparent in the data here.
articulatory displacement, on the other hand, provided unequivocal evidence for a prominence hierarchy, according to which:

- The first syllable is always the most prominent, and hence probably carries primary stress.
- The second syllable is always the least prominent, and hence is probably always unstressed.
- The third syllable fluctuates between the two: it is virtually identical to the second syllable, therefore unstressed, in words that are part of greater phonological phrases, but it is clearly more prominent than the second syllable, and in certain aspects rather similar to the first syllable, in words that span a whole phonological phrase on their own (e.g. topicalized or focalized words, words in citation form, one-word utterances etc.). In the latter condition, it can be safely claimed that the third syllable has secondary stress in Georgian (at least in words with 5 syllables).

It is quite likely that the initial impression of non-native listeners that Georgian has regular secondary stress on odd non-initial syllables, regardless of the word’s position in the higher prosodic hierarchy, resulted from unbalanced exposure to citation forms, one-word utterances and sentences pronounced with pauses between words, either for clearer elicitation or due to an online translation process from English to Georgian. However, one cannot rule out that secondary stress is indeed regular in Georgian and is manifested in other phonetic aspects such as pitch. The data, however, do not suggest that, in 5-syllable words, primary stress is on the third syllable, and support the claim that primary stress is always word-initial, at least in this type of declarative sentences.