Disentangling word stress and phrasal prosody: a view from Georgian

Lena Borise[*]

Hungarian Research Centre for Linguistics – lena.borise@nytud.hu

This paper investigates the interaction of word stress and phrasal prosody in Georgian by studying the distribution of acoustic cues (duration, intensity, F0) in controlled data. The results show that initial syllables in Georgian words are marked by greater duration than all subsequent syllables, regardless of syllable count and phrasal context. After excluding domain-initial strengthening as an alternative explanation, this finding provides evidence in favor of fixed initial stress. Likewise, initial syllables are marked by greatest intensity, but the consistent gradual drop in intensity throughout the word suggests that this effect may not be stress-related. The F0 results align with the existing accounts: individual lexical words form Accentual Phrases, marked by a low pitch accent on the initial syllable and a high final boundary tone on the final syllable. Additionally, new evidence for a phrasal accent, aligned with the penult, is presented. F0 targets are shown to be completely absent in the context of post-focal deaccenting, which shows that F0-marking in Georgian is reserved for phrasal prosody and is not intrinsic to stress-marking. These results help account for the facts related to word stress, phrasal intonation, and their interplay in Georgian, the object of debate in the literature.

Keywords: Georgian; word stress; F0; duration; intensity

1 Introduction

When it comes to the issue of stress in an understudied language, many questions emerge: does it have word stress? How is its placement determined? What acoustic parameter (syllable/vowel duration,[1] F0, intensity) does the realization of word stress chiefly rely upon? Answering these may not be straightforward. To provide just a brief illustration, word stress in Hunzib, a small Tsezic language of Dagestan, has been described as free and cued by F0 (Bokarev 1967: 474), fixed on the initial syllable and not relying on F0 (Gamzatov 1975: 18), penultimate and often accompanied by high F0, but not for all speakers (van den Berg 1995: 28), and initial but with numerous exceptions, driven by morphological factors (Isakov & Khalilov 2012: 78). Finally, it has been suggested that Hunzib does not have word-level stress, and, instead, initial syllables carry phrasal accents (Kibrik & Kodzasov 1990: 332).

As the example of Hunzib demonstrates, another important question is about the interaction of word-level stress and the expression of phrasal intonation. This issue is a notoriously difficult one and has not

[1] Consonant duration can play a role in cuing stress as well (Everett 1998; Michael 2010).

been settled even in some better-studied languages. One reason for this is that in many languages individual words are isomorphic to the smallest prosodic phrases, which does not allow for a clean separation between the two. This issue has been shown to arise in languages as diverse as Korean (Jun 1993; 1998) and Chickasaw (Gordon 2003; 2004; 2005), among many others. If lexical words and prosodic phrases correspond, the methodological challenge is to determine what constitutes a word-level or a phrase-level property. It is often assumed that, in languages without lexical tone, changes in F0 reflect phrasal prosody, while other parameters, such as duration and intensity, have a closer connection with word-level stress. This is not necessarily the case, however, since both intensity and duration can interact with F0. Specifically, intensity values are dependent on F0 values, and duration can increase, e.g., to accommodate for several tonal targets in cases of tonal crowding. Furthermore, greater duration can also result from domain-initial strengthening, which is independent from word-level stress or intonation (Fletcher 2010).

To take a closer look at another example, consider Korean. There is no agreement as to whether word stress exists in Korean, and, if it does, where its location is; initial or second (Lee 1973), second (Huh 1985), and final syllables (Polivanov 1936, as cited in Lee 1990) have been argued to regularly carry stress. More fundamentally, though, it is unclear if the phenomenon that has been labelled 'word stress' is actually word-level stress or phrase-level prominence. Contrary to the previous accounts, Jun (1993) shows that stress placement in a word depends on its position in an Accentual Phrase, which suggests that the phenomenon at hand is phrasal in nature. Jun (1995) further shows, based on instrumental evidence, that whether the initial or second syllable is perceived as prominent is determined by a combination of factors, such as syllable count, syllable weight, and the position of the phrase in a larger utterance; she also shows that the main acoustic cue that stress/prominence relies on is F0. These factors, taken together, suggest that what is described as 'word stress' in Korean, in fact, fits the profile of a phrasal intonational F0 target.

The language that this paper focuses on, Georgian (Kartvelian), also poses interesting challenges with respect to the questions raised above.[2] Like Korean, Georgian has been variably analyzed as having either word-level stress, phrase-level stress, or both. There is no unanimity about the acoustic cues that mark stress in Georgian. Initial (Tschenkeli 1958; Tevdoradze 1978, a.o.), antepenultimate (Akhvlediani 1949; Gudava 1969, a.o.), and penultimate (e.g., Zhghenti 1958) syllables have been described as carrying stress by different authors, with more than one stress locus argued to be possible in longer words. Native speakers of Georgian have no consistent intuitions about stress placement, other than that stress never targets the ultima. There are no minimal pairs based on stress and no regular variation in stress placement in declensional or conjugational paradigms. Authors who advocate for the existence of word stress in Georgian acknowledge its acoustic weakness and often remark on the uncertainty of their observations (Robins & Waterson 1952; Zhghenti 1959; Tevdoradze 1978). Finally, lexical words in Georgian regularly correspond to Accentual Phrases (APs),[3] which means that the challenges that arise from the isomorphism of individual words and prosodic phrases affect Georgian as well.

Based on experimental evidence, this paper argues for stress in Georgian being fixed on the initial syllable and cued primarily by syllable duration, and, possibly, intensity (the latter also interacts with phrasal prosody). With respect to F0, the findings confirm the facts about Georgian that have been established in the literature: APs in Georgian carry phrasal intonational F0 targets on initial and final syllables. In neutral broad-focus declaratives, initial syllables host low pitch accents (L*), which may be manifested as dips in F0 or low plateaus, and right edges of APs carry high final boundary tones. Additionally, there is evidence for penultimate syllables hosting low phrase accents (L), which constitute the lowest F0 point per AP. The latter result aligns with a previously established fact of Georgian intonational phonology: penultimate syllables of predicates in narrow focus contexts and questions carry a

---

[2] Georgian is the biggest and the best studied language within the Kartvelian family, which includes another three languages (Megrelian, Laz, and Svan). The prosodic properties of these languages await detailed further study. For an overview of the existing work, see Borise (2020).

[3] The analytical terminology (e.g., names of prosodic constituents) are borrowed from Vicenik & Jun's (2014) analysis of Georgian prosody.

low F0 target, identified as a low phrase accent (Bush 1999; Vicenik & Jun 2014; Borise 2017). The penultimate F0 targets – the one described in the current paper and the low phrase accent described in the literature – likely constitute two subtypes of the same phenomenon.

This paper also addresses the challenge of isomorphism between lexical words and prosodic phrases and shows that there is a limited number of contexts in Georgian in which lexical words do not correspond to APs one-to-one. One such context is the parts of utterances that follow narrowly focused constituents. While post-focal deaccenting is optional in Georgian (Vicenik & Jun 2014), it is possible to find examples where it applies systematically, and post-focal constituents are stripped of all tonal targets. Another context is grouping of APs into Intermediate Phrases (ips). This is also an optional process that may apply to semantically/syntactically connected adjacent words, e.g., a noun and a modifying adjective. When combined into an ip, individual APs retain their pitch accents, but, phonologically, an extra layer of prosodic phrasing emerges between the levels of APs and full Intonational Phrases. The results presented in this paper show that initial syllables still carry word-level stress, cued by duration, even if the word is (i) found in the domain of post-focal deaccenting or (ii) embedded in an Intermediate Phrase. At the same time, the lack of tonal targets in the context of post-focal deaccenting convincingly demonstrates that F0 is only used in phrasal prosody in Georgian.

This paper is structured as follows. Section 2 discusses the existing research on word stress in Georgian, introspection-based (2.1) and experimental (2.2), presents the key properties of Georgian intonational phonology (2.3), and sums up the known facts (2.4). Section 3 provides the information about the experiment: the research questions (3.1), and the design and stimuli (3.2). Section 4 reports the results for syllable duration (4.1), intensity (4.2), and F0 patterns (4.3). Section 5 brings in additional data from a supplementary study (5.1), divided into Datasets 1a and 1b (5.2-5.3) and Datasets 2a and 2b (5.4-5.5). Section 6 offers a discussion of the results, and Section 7 concludes.

2   Georgian prosody: previous work

The prosodic properties of Georgian have received considerable attention in the literature, with the existing descriptions based both on introspection and instrumental analysis. Nevertheless, there is no agreement as to the existence of word stress in Georgian or rules governing its distribution. Initial, antepenultimate and/or penultimate syllables are most often postulated as possible stress loci, with potentially more than one of these carrying stress in longer words. The question about the size/type of the prosodic domain that 'stress' in question is assigned in – i.e., whether it is a lexical/prosodic word or a larger constituent, such as a prosodic phrase – has not been settled either.

2.1   Introspection-based accounts

According to Tschenkeli (1958: LX), stress in Georgian is found on the initial syllable in di- and trisyllabic words, and is harder to locate in longer words, though there, too, it is often initial. Tevdoradze (1978: 40) also argues for fixed initial stress, but notes that secondary stress may occur in longer words: on the penult in tetrasyllables, antepenult in pentasyllables, and antepenult or pre-antepenult in hexasyllables. Antepenultimate stress placement is advocated by Ioseliani (1840: 145), Gorgadze (1912: 3), Akhvlediani (1949: 135), and Gudava (1969: 106). Gorgadze notes that in longer words/phrases, the initial syllable receives secondary stress; exceptionally, if the antepenult consists of a vowel only, stress targets the pre-antepenult: *sá.i.dum.lo* 'mystery', *mí.i.rbi.na* '(he) came running'. Zhghenti (1958: 262) describes the Khevsuri and Mokheuri dialects of Georgian as regularly assigning stress to penults. Parting from most other works, Robins & Waterson (1952: 58) propose that stress assignment in Georgian follows a rhythmic pattern (while avoiding final syllables), with alternating non-adjacent syllables carrying stress (primary or secondary): disyllables have initial stress, trisyllables can be stressed on the first or second syllables, tetrasyllables either carry stress on the second syllable or on the first and third syllables, and pentasyllables either stress the first and third or the second and fourth syllables.

In numerous accounts, Georgian stress placement is described as dependent on syllable count: initial in disyllables and antepenultimate or penultimate in longer words (Marr 1925: 13; Rudenko 1940: 24; Vogt 1971: 15), or initial in di- and trisyllables and antepenultimate in longer words (Dirr 1904: 3; Janashvili 1906: 5; Akhvlediani 1949: 132); Dirr (1904: 3) also notes that these rules apply regardless of the morphological structure of a word. In words over four syllables long, a secondary stress on the initial syllable is possible; both are obligatory in words over six syllables long. A similar approach is adopted by Skopeteas & Féry (2016), who take disyllables to carry stress on the first syllable, and words of four syllables or longer to carry primary stress on the antepenult and secondary stress on the initial syllable. According to Aronson's grammar (1990: 18), in words up to four syllables long, stress falls on the antepenult or the initial syllable, while in longer words both are stressed. Finally, according to Hewitt (1995: 28), in trisyllables, the initial syllable carries stress; in longer words, stress is either antepenultimate or initial.

In contrast, some maintain that Georgian does not have word-level stress, and its prosodic organization only includes tonal targets that are assigned within prosodic phrases. This view goes back to Gorgadze's (1912: 13) notion of 'syntactic groups', Marr's (1925: 14) 'accentual complexes', and Zhghenti's (1953: 162; 1963: 144) 'rhythmic groups' as domains of 'stress' assignment in Georgian. Some evidence supporting this view comes from traditional Georgian poetry, which is based on syllable count and not alternation of stressed and unstressed syllables (Gachechiladze 1968). Finally, so-called mixed approaches advocate for there being both word stress and phrasal F0 targets in Georgian. This view is maintained by Chikobava (1942: 302) and Tschenkeli (1958: LXI), who point out that word stress in contemporary Georgian is considerably weaker than phrasal prosodic targets.

2.2   Experimental investigations

There have been several experimental studies of Georgian prosody, both word- and phrase-level. The conclusions that have been reached vary, similarly to how introspection-based reports do. In one of the earliest studies, Selmer (1935) reports on an instrumental investigation of word stress in Georgian, based on recordings of one speaker pronouncing 27 Georgian words (twenty disyllables, six trisyllables, and one tetrasyllable), some iterated twice, with the total stimuli count being 36. The experimental items were uttered in isolation. Measurements of F0 curves and vowel duration are reported. Selmer (1935) notes that the initial syllable invariably carries an F0 peak, with the average rise being 2.64 st (semitones). In disyllables, the two vowels are almost equal in duration, while in trisyllables the second vowel is the shortest, with the two others being comparable in duration.[4] With respect to vowel quality, Selmer notes that, in the initial syllable, vowel quality does not have a significant effect on vowel duration, which suggests that the durational effect on the initial syllable cannot be explained, e.g., by the intrinsically greater duration of low vowels. Overall, Selmer cautiously interprets his results as consistent with Vogt's initial assessment, later published as Vogt (1936; 1971), according to which di- and trisyllables are stressed on the initial syllable.

According to the results of Zhghenti's (1953; 1959) production experiment, all syllables in Georgian words other than the final two are high in prominence, which Zhghenti (1953; 1959) takes to be expressed in F0 values and intensity, while the final two syllables (or the final one in disyllables) are less so. Zhghenti's results are based on the analysis of F0 traces of individual words two to six syllables long, uttered in isolation. The total number of stimuli or speakers is not reported, but a number of F0 traces are discussed: disyllabic (n=6), trisyllabic (n=6), tetrasyllabic (n=7), pentasyllabic (n=4), and hexasyllabic (n=2). Zhghenti interprets his results as indicative of initial stress in di- and trisyllables; he takes both the initial and the second syllable to be stressed in tetrasyllables, and refrains from interpreting the results for penta- and hexasyllables. Zhghenti's (1953; 1959) results are summarized in Table 1.

---

[4] In words uttered in isolation, the extra duration on word-final syllables is likely due to final lengthening.

Table 1: Stress placement in Georgian according to syllable (σ) count (Zhghenti 1953; 1959)

σ count	Interpretation of measurements	Stressed σ
2σ	High F0 on the 1st σ	1st
3σ	High F0 on the 1st σ, high intensity on the 1st and 2nd σ	1st
4σ	High F0 on the 1st & 2nd σ, high intensity on the 1st, 2nd, and 3rd σ	1st & 2nd
5σ	High F0 and intensity on the 1st, 2nd, and 3rd σ	?
6σ	High F0 and intensity on the 1st, 2nd, 3rd, and 4th σ	?

Other experimental studies, such as Alkhazishvili (1959), argue that 'stress' placement in Georgian interacts with information structure. This suggests that the prosodic phenomenon in question is phrasal in nature, since word-stress placement typically does not depend on information-structural factors. Alkhazishvili (1959) argues that discussing 'stress' in Georgian is only possible with reference to particular information-structural contexts, and proposes that three such contexts determine 'stress' placement:

- Type I: broad-focus utterances, with neutral word order (e.g, with transitive predicates, SOV) (1);
- Type II: utterances with narrow focus on the preverbal constituent (2);
- Type III: verb-initial thetic utterances (3).

These utterance types, in Alkhazishvili's analysis, vary with respect to the distribution of "subject" and "predicate" prosodic phrases within them, which correspond to the more conventional notions of Topic and Comment, respectively. A Comment includes the verb and the immediately preverbal focused constituent (if present), while a Topic includes all other material in a clause. Stress placement within a phrase is determined by its type.

(1) [Topic Giorgi-m] [Comment pex-i   ar   ga-a-ndzr-i-a.]
         Giorgi-ERG   foot-NOM   NEG   PV-VER-move-SM-AOR.3SG[5]
      'Giorgi didn't move'

(2) [Comment Omaraʃvil-ma   da-i-xsn-a]   [Topic   gatʃ'irvebi-dan samartal-i.]
         Omarashvili-ERG   PV-VER-save-AOR.3SG   hardship-from   court-NOM
      'It was Omarashvili that led the court out of the difficult situation.'

(3) [Comment Ga-vid-nen   k'idev dɣeni da   tveni.]
         PV-go-IPFV.3PL   more   day.PL and   month.PL
      'More days and months went by.'   (Alkhazishvili 1959)

The results are based on the analysis of 21 recorded utterances (Type I = 12, Type II = 6, Type III = 3), pronounced by a male and a female native speaker. Two phoneticians, one of whom was a native speaker of Georgian, acted as analysts. In Topic phrases, which have an overall rising intonational pattern, initial stress was identified by the analysts, but it is not specified what acoustic cue this conclusion was based on. No acoustic evidence for antepenultimate stress was identified in Topic phrases. In Comment phrases, the picture is more complex. Here, in most cases, the analysts also identified initial stress, but about 20% of Comment phrases were identified as having antepenultimate stress. Acoustic evidence for this judgement is scarce, and small sample size did not allow Alkhazishvili to reach a conclusion about its nature. The differences in stress perception between native and non-native speaker analysts are not reported.

Finally, Jun et al. (2007) and Vicenik and Jun (2014: 156) report on a preliminary production study, which involved four speakers of Georgian and test words embedded in carrier phrases (disyllabic = 3, trisyllabic = 4, pentasyllabic = 1), each uttered twice. They found that the initial syllable in Georgian is characterized by higher intensity and longer duration. They also report a high-low tonal contour that spans

---

[5] The Leipzig Glossing Rules (https://www.eva.mpg.de/lingua/resources/glossing-rules.php) are followed for glossing language examples. These are supplemented by abbreviations listed in the Abbreviations section at the end.

the antepenult and penult, which they take to be a manifestation of phrase accent. Based on these results, they suggest that word stress in Georgian is fixed on the initial syllable, while the antepenult and penult are loci of phrasal intonational F0 targets. Borise & Zientarski (2018) arrive at the same conclusion – initial word stress and phrasal F0 targets anchored to the right edge of prosodic domains – based on a larger dataset (one speaker, 179 words of 1-6 syllables embedded in carrier phrases).

2.3 Intonational phonology of Georgian

Issues of Georgian intonational phonology have attracted considerable scholarly interest. In their work, Tevdoradze (1978; 2005), Bush (1999), Müller (2005), Skopeteas, Féry, and Asatiani (2009; 2018), Skopeteas and Fanselow (2010), Asatiani and Skopeteas (2012), Jun, Vicenik and Lofstedt (2007), Skopeteas and Féry (2010; 2014; 2016), and Vicenik and Jun (2014) cover a wide range of issues, including prosody of neutral statements, questions, and utterances containing narrow focus. Jun et al. (2007) and, especially, Vicenik & Jun (2014) offer a detailed Autosegmental-Metrical (AM) analysis of Georgian prosody, establishing the levels of prosodic phrasing and inventory of F0 targets. The key insights of Vicenik & Jun (2014) – the analysis that is adopted in this paper – are provided below.

Each prosodic word in Georgian – defined as a lexical word, which may be accompanied by clitics, like postpositions or discourse particles – forms an Accentual Phrase (AP). This is based on the fact that prosodic words in Georgian carry final boundary tones, which means that they also form minimal prosodic phrases, such as APs. As part of the unmarked intonational pattern of all-new, broad-focus declarative utterances, each AP, except for the right-most one, carries a rising F0 contour. Vicenik & Jun (2014) analyze it as a low pitch accent L* on the initial syllable of the AP followed by a high final boundary tone on the final syllable, Ha (where 'a' indicates that the boundary is part of the AP). Typically, downstep applies to each successive Ha. This pattern is illustrated in Figure 1, with glosses provided in (4).[6] Importantly for our purposes, the final high boundary tone is phrasal, and not associated with word stress; the (steepest) rise in F0 and the F0 peak are typically contained within the final syllable. The full inventory of pitch accents and boundary tones can be found in Vicenik and Jun (2014).
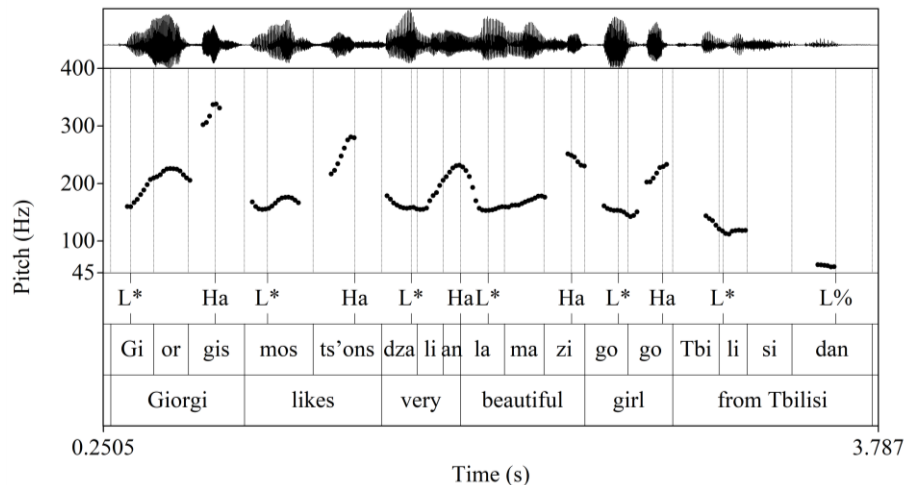

Figure 1. An example of declarative intonation

(4)   (Giorgi-s)_{AP}   (mosts'on-s)_{AP}   (dzalian)_{AP}   (lamaz-i)_{AP}   (gogo)_{AP}   (Tbilisi-dan)_{AP}.
Giorgi-DAT   like-PRS.3SG   very   beautiful-NOM   girl.NOM   Tbilisi-from
'Giorgi likes a very beautiful girl from Tbilisi.'

---

[6] The acoustic data used for illustrating the prosodic patterns of Georgian was collected during author's fieldwork in Georgia.

A prosodic constituent larger than an AP is an Intermediate Phrase (ip). Ip-formation in Georgian is optional and may apply to two or three syntactically/semantically related adjacent APs, such as a noun and an adjective/demonstrative. Each of the APs within an ip retains its own pitch accent, but the presence of an ip is signaled via a change in the F0 pattern: non-ip-final APs carry an H* La tonal pattern instead of L* Ha.[7] The ip-final AP carries an L* pitch accent and an H- final boundary tone, which overrides that of the AP. Ip-formation is illustrated in (5a) and Figure 2, where the ip corresponds to the nominal phrase *lamaz-ma kalbat'on-ma* 'beautiful-ERG lady-ERG' (boxed). Lack of ip-formation is illustrated in (5b) and Figure 3, where *lamaz-ma* 'beautiful-ERG' and *kalbat'on-ma* 'lady-ERG' form two independent APs; note that the OV vs. VO order in (5a) vs. (5b) has no bearing on the intonational contour of the subject. While the presence of an ip does not break the isomorphism between lexical words and prosodic phrases (APs), being embedded in another layer of prosodic phrasing may affect the phrasal prosodic make-up of the APs contained in an ip – e.g., because of it being adjacent to a higher-level prosodic boundary. Ip-containing contexts will be investigated in Datasets 1a-b of the supplementary study.

Finally, the largest prosodic constituent, an Intonational Phrase (IP), corresponds to a clause and typically carries a final low boundary tone L%, as shown in Figures 1-5 (other IP-boundary tones, like H% and HL%, are attested in other contexts; for a full tonal inventory proposed for Georgian, see Vicenik & Jun (2014)).

(5)   a.   ((Lamaz-ma)AP   (kalbat'on-ma)AP)ip   (k'aba)AP   (mo-i-zom-a)AP.
           beautiful-erg lady-erg   dress.nom   pv-ver-try-aor.3sg
           'A beautiful lady tried on a dress.'

      b.   (Lamaz-ma)AP   (kalbat'on-ma)AP   (mo-i-zom-a)AP   (kvedabolo)AP.
           beautiful-ERG   lady-ERG   PV-VER-try-AOR.3SG   skirt.NOM
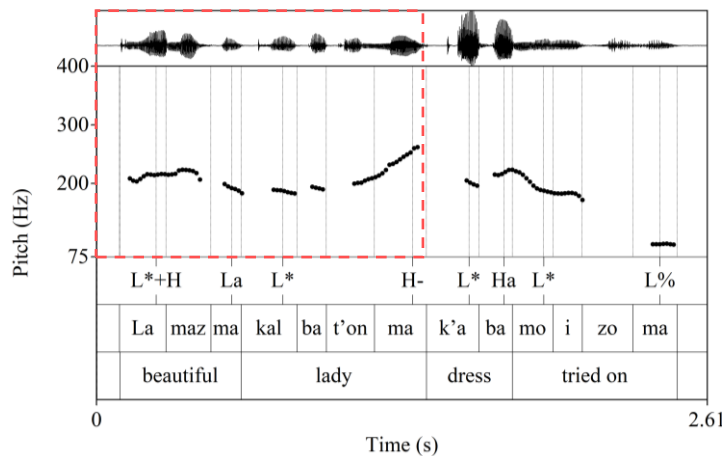   'A beautiful lady tried on a skirt.'

Figure 2. The intonational contour of an utterance containing *lamaz-ma kalbat'on-ma* 'beautiful-ERG lady-ERG' as an ip

---

[7] Alternatively, especially if the ip is utterance-initial, the leftmost pitch accent may be realized as L*+H instead of H*, with the F0 peak on the second syllable (Vicenik & Jun 2014: 185). The pitch accent on *lamaz-ma* 'beautiful-ERG' in Figure 2 is labelled as L*+H, given that this is the only kind of pitch accent, proposed for Georgian, where the location of the F0 peak is on the second syllable. Note, however, that the difference in F0 values between the first and second syllables in *lamaz-ma* 'beautiful-ERG' is very slight. More detailed research of the inventory of pitch accents in Georgian may be needed to label it more reliably.
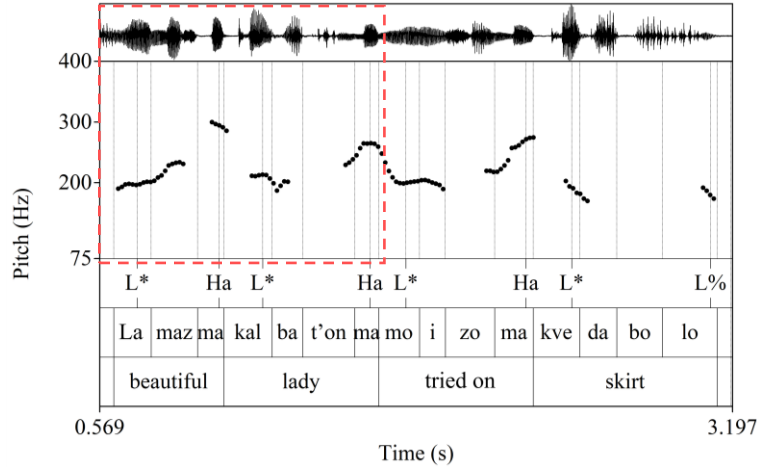
Figure 3. The intonational contour of an utterance containing *lamaz-ma kalbat'on-ma* 'beautiful-ERG lady-ERG' as two APs

The prosody of focus-marking in Georgian is not going to be discussed in detail here (see Skopeteas, Féry & Asatiani 2009; Skopeteas & Féry 2010; 2014; Borise 2019, a.o.), but one of the focus-related phenomena that has a bearing on breaking the isomorphism between lexical words and prosodic phrases is post-focal deaccenting/dephrasing. In many languages, the part of the utterance that follows a narrowly focused constituent exhibits no/little evidence for the presence of phrasal tonal targets (Ladd 1980; 1996; Jun 1993; Rahmani, Rietveld & Gussenhoven 2018, a.o.), or, if the tonal targets are present, the tonal range is dramatically reduced (Harnsberger & Judge 1996). If analyzed as eliminating tonal targets, including phrasal boundaries, post-focal deaccenting provides a context where the isomorphism between words and prosodic phrases does not hold. As such, it is an environment that allows to distinguish word- vs. phrase-level prosody. In Georgian, post-focal deaccenting is optional: the constituents that follow a narrowly focused one may either retain their tonal targets (often with a reduced tonal range), or be stripped of their tonal targets (Vicenik & Jun 2014: 179). The former is illustrated in (6a) and Figure 4, and the latter in (6b) and Figure 5; in Figure 4, creaky voice does not allow for pitch tracking on the final word *tve-ʃi* 'month-LOC'. The L pitch target on the verb in Figure 4 is a phrase accent, discussed in detail in the remainder of this section. Also, note that the focused subjects form ips with the verbs that follow them.[8] There is a strong preference for preverbal focus placement in Georgian, and focused constituents often form a prosodic constituent with the verb. In such cases, post-focal deaccenting applies to the material that follows the verb.

(6) a. ('Who drove a/the car last month?')
((*Mamuk'a_F*)_{AP}   (*a-t'ar-eb-d-a*)_{AP})_{ip}   (*mankana-s*)_{AP}   (*ts'ina*)_{AP}   (*tve-ʃi*)_{AP}.
Mamuka.NOM   PV-drive-SF-SM-IPFV.3SG   car-DAT   last   month-LOC
'Mamuka_F drove a/the car last month.'

b. ('Who made jam last week?')
((*Bebia_F*)_{AP}   (*a-ket-eb-d-a*)_{AP})_{ip}   *muraba-s*   *ts'ina*   *k'vira-s.*
grandma.NOM   PV-make-SF-SM-IPFV.3SG   jam-DAT   last   week-DAT
'Grandma_F made jam last week.'

---

[8] In Figure 4, the verb does not seem to carry its own pitch accent, and it may be possible to analyze the focused constituent and the verb as forming a single AP, not ip. I label it as ip due to a stronger break between the verb and postverbal material than what would be expected between two APs. Vicenik & Jun (2014: 180; fn.3) mention and briefly discuss such cases but do not provide a definitive analysis. I leave examples like this for further research.
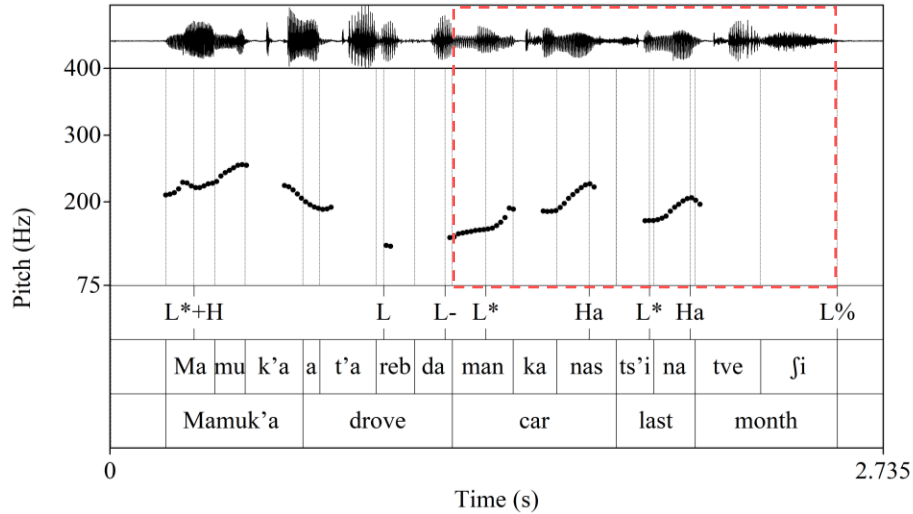
Figure 4. The intonational contour of an utterance with narrow focus on *Mamuka* (personal name), with no post-focal deaccenting
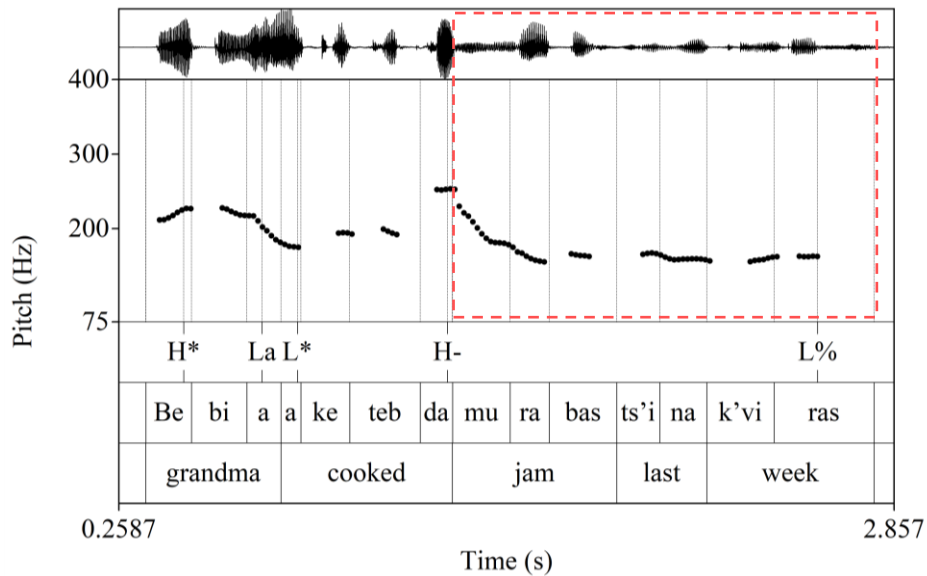

Figure 5. The intonational contour of an utterance with narrow focus on *bebia* 'grandma', with post-focal deaccenting

In addition to pitch accents and boundary tones, Vicenik and Jun (2014) argue for there being another F0 target in Georgian – a phrase accent. In the AM literature, phrase accents have been variably analyzed as boundary tones for mid-level prosodic phrases (e.g., phonological phrases), or as F0 targets found between the rightmost pitch accent and a final boundary tone (Bruce 1977; Pierrehumbert 1980; Ladd 1983; Grice, Ladd & Arvaniti 2000).[9] The distribution and properties of phrase accents in Georgian have been addressed in several studies, and it has been established that they anchor to penultimate syllables, in certain contexts (Bush 1999; Müller 2005; Vicenik & Jun 2014; Borise 2017). Penultimate placement makes phrase accents reminiscent of antepenultimate/penultimate word stress, as postulated by some analyses.

---

[9] Phrase accents are known to have more complex distribution than other F0 targets, sometimes taking the form of pitch stretches or 'elbows', rather than being associated with a particular syllable (Del Giudice et al. 2007; Reichel & Salveste 2015; Flemming 2018).

Let us briefly review the properties of Georgian phrase accents. In a study of prosody of yes/no-questions in Georgian, Bush (1999) notes that penultimate syllables of verbs in yes/no-questions are always marked by a low tone, before a sharp rise on the final syllable. He takes the low F0 target on the penult to be (part of) a phrase accent, though notes that the precise anchoring of the low tone to the penultimate syllable is atypical of a phrase accent. Müller (2005) in her study of the prosodic right periphery in yes/no-questions in Georgian also notes the low F0 target on the penult of the yes/no-question on the whole and takes it to be a phrase accent. Note that, since the questions in Bush's study were verb-final, Müller's findings also replicate those of Bush. Similarly, Vicenik and Jun (2014) analyze the penultimate syllable in yes-no questions, wh-questions, and narrow focus contexts as carrying a phrase accent. They specify further that the low target on the penult is preceded by a high target on the antepenult, and analyze the phrase accent as H+L. Finally, the issue of the phrase accent is also taken up in Borise (2017), where two conclusions are reached. First, it is shown that the phrase accent is associated with the verb, as opposed to the right edge of the question. This is demonstrated with the help of questions in which the verb is clause-initial or -medial, but the low F0 target is still associated with the verb. Second, it is shown that the high F0 target on the antepenult does not always accompany the low tone on the penult, which confirms the intuition that the low tone constitutes the main tonal element of the phrase accent. In line with these conclusions, phrase accents are labelled L in this paper. The realization of the phrase accent is illustrated in (7) and Figure 6 (as well as in Figure 4 above).

(7) **ʃe-tʃ'am-a**   Manana-m   alubal-i?
     PV-eat-AOR.3SG Manana-ERG   cherry-NOM
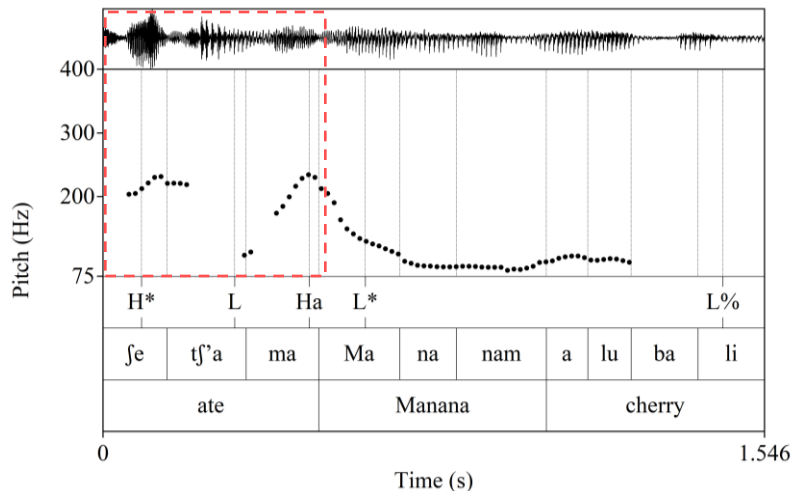   'Did Manana eat the/a cherry?'

Figure 6. The realization of the low phrase accent L on the penultimate syllable of the verb
ʃe-tʃ'am-a 'PV-eat-AOR.3SG'

Most importantly for our purposes, the existence of a phrase accent in Georgian is another factor to consider when interpreting the instrumental results.

2.4   Summary

There is no unanimity on the nature and distribution of stress in Georgian, and stress interacts with phrasal prosody in a complex way. It seems uncontroversial that shorter words (di- and trisyllables) carry stress on the initial syllable, but the picture is less clear in longer words, where many authors note the presence of another stress locus on the antepenult or penult, with no agreement as to which of the loci carries main stress. Additionally, factors like information structure may systematically affect the F0 contours, which also influences the perceived location of word stress. Lexical words are commonly isomorphic to APs,

which makes it difficult to distinguish word- and phrase-level prominence, but there are two contexts in Georgian where the isomorphism may not hold: (i) the domain of post-focal deaccenting and (ii) ip-formation. Several types of tonal targets are distinguished, including pitch accents, boundary tones, and phrase accents.

3 Instrumental investigation: methods

3.1 Research questions

Cross-linguistically, the expression of stress typically relies on duration, F0, and intensity-related measures (Beckman 1986; Gordon & Roettger 2017). Stressed vowels/syllables commonly have greater duration than unstressed ones (De Jong & Zawaydeh 1999; Eriksson & Heldner 2015; Garellek & White 2015 a.o.). Stressed syllables may also carry an F0 target or a particular tonal contour, though it is not always clear whether this F0 specification is lexical (word-level) or post-lexical (phrasal) in nature. Stressed syllables may be identified by intensity-based measurements, such as overall intensity (Remijsen & Van Heuven 2005; Vogel, Athanasopoulou & Pincus 2016), or frequency-sensitive intensity (Sluijter & van Heuven 1996). In the existing studies, duration is most often mentioned as an acoustic cue for stress, followed by F0, intensity, and formant and spectral qualities (Gordon & Roettger 2017). In turn, the expression of phrasal prosody commonly relies on F0 targets, aligned with stressed syllables or edges of prosodic domains, though it may also be cued by other effects: most commonly, final lengthening (Edwards, Beckman & Fletcher 1991), initial strengthening (Hock 1988), and/or glottalization (Dilley, Shattuck-Hufnagel & Ostendorf 1996). The questions that this paper aims to answer, therefore, are the following:

(i)   Is there consistent evidence that certain acoustic cues (F0 targets, durational and intensity effects) are realized on the syllables that have been identified as possible stress loci in Georgian (initial, antepenultimate, penultimate)?

(ii)   How is the distribution of these cues best interpreted phonologically (i.e., as word- or phrase level)?

(iii)   How does this evidence relate to the existing descriptions and earlier experimental results available for Georgian?

The experiment addresses (i) and allows for formulating preliminary responses to (ii). The evidence brought forward by the supplementary study provides further evidence relevant for answering (ii). The data from both investigations is considered to address (iii).

3.2 Stimuli and design

The experiment builds on previous instrumental work, especially Vicenik and Jun (2014) and Borise & Zientarski (2018). The stimuli consisted of Georgian words (n=172), two to five syllables long, of CV syllable structure, where onsets contain a single voiced consonant (a nasal, liquid, or a voiced stop or fricative), and the vowel is any of the five vowels of Georgian /i, u, ɛ, ɔ, a/. CV syllable structure was selected for two reasons. First, to avoid complex onsets, which are very common in Georgian and often include both voiced and voiceless consonants, which would disrupt pitch tracking. Second, to unify syllable weight across stimuli, since syllable weight has not been mentioned as relevant for stress assignment in the existing literature on Georgian. The test words included nouns, adjectives, and participles.[10] Given the

---

[10] As part of the 172 stimuli, 15 non-words of three to five syllables were used; in the dataset, they were randomly interspersed with real words, to determine whether the word vs. non-word status of a stimulus may affect its prosodic realization. Prior to the analysis, disfluent non-word tokens, like disfluent real word ones (due to pauses, errors, list intonation, etc.) were discarded. No statistically significant differences in the behavior of the remaining non-word

observation, made in the existing grammars, that morphological structure is not a relevant factor in stress placement in Georgian, both mono- and polymorphemic stimuli were used. A representative sample of the stimuli is provided in Table 12; a full list of stimuli is available in the supplementary materials.[11]

Table 2. Sample stimuli

σ count	Sample stimuli	n
2 σ	*mama* 'father.NOM'	30
	bude 'nest.NOM'	
3 σ	*ʒiʒin-i* 'bird twitter-NOM'	54
	malamo 'balm.NOM'	
4 σ	*bagabug-i* 'thumping-NOM'	55
	monazon-i 'monk-NOM'	
5 σ	*ga-gor-eb-ul-i* (PV-roll-SF-PTCP-NOM) 'rolled'	33
	mineral-ur-i (mineral-ADJ-NOM) 'mineral (adj.)'	
Total		**172**

The stimuli were embedded in one of three carrier phrases. In them, the test word acted as the object in an SOV clause, and corresponded to an AP:

(8)   a. *Me   sit'q'va*   ([stimulus])_{AP}   *v-i-mɣer-e.*
        1SG   word.NOM   1SG-VER-sing-AOR.1SG
        'I sang the word "[stimulus]".'

    b. *Me   sit'q'va*   ([stimulus])_{AP}   *v-i-xmar-e.*
        1SG   word.NOM   1SG-VER-use-AOR.1SG
        'I used the word "[stimulus]".'

    c. *Me   sit'q'va*   ([stimulus])_{AP}   *da-v-c'er-e.*
        1SG   word.NOM   PV-1SG-write-AOR.1SG
        'I wrote the word "[stimulus]".'

Seven native speakers of Georgian participated in the study: two males (M1, M2) and five females (F1-F5).[12] All participants were natives of Tbilisi, with the age range 22-35 y.o., mean age 26.8 y.o. Speaker M2 was recorded in Tbilisi, Georgia, speaker F2 in College Park, Maryland, and the other five speakers in Cambridge, Massachusetts. Of the speakers recorded in the US, two (F1 and F2) have lived there for eight and seven years, respectively, but still use Georgian on a daily basis with their families. Speaker F5 arrived in the US a year prior to the recording time, and speakers M1, F3 and F4 arrived in the US less than a year prior to the recording time. The data from speakers F2 and M2 was collected using a lavalier microphone in a quiet classroom; the other five speakers were recorded in a sound-proof booth using a head-worn microphone. All data was recorded at the sampling rate of 44.100 Hz and 16 bits per sample; both microphones had the frequency response range 20Hz – 20kHz.

---

tokens, as compared to that of real words, were detected. Consequently, they were included into the dataset. Exclusion of non-words from the dataset or their inclusion does not affect the results.

[11] The variation in the total number of words with different syllable counts stems from several factors, such as the CV syllable structure requirement, the requirement for voiced consonants, and avoiding consonant clusters. Mono- and hexasyllabic stimuli (n=9 and 4, respectively) were also recorded, but the low numbers meant that the data could not be used in a conclusive way. They were excluded from the final dataset.

[12] One of the speakers, F2, was recorded by Xavier Zientarski, with the results reported in Borise & Zientarski (2018).

The same set of stimuli was used for all speakers. Each speaker was provided with a randomized list of all stimuli and, as a separate list, the three carrier phrases. They were instructed to pick each word from the list, one by one, insert it into one of the carrier phrases, also picked consecutively, and pronounce the resulting combination. The two lists were kept separate to avoid the effect of reading. Each stimulus was iterated three times in a row – i.e., each type contributed three tokens.[13] These two factors – different verbs accompanying different experimental stimuli, and the fact that each stimulus was iterated three times – were intended to discourage a contrastive reading on the consecutive stimuli. The stimuli that were pronounced with list intonation – the first and/or the second iteration of the carrier phrase with the same stimulus – were discarded; rising F0 at the end of the carrier phrase was taken to be the main indicator of list intonation. Since no additional context was provided for the stimuli embedded in the carrier phrases, their information-structural status was taken to be that of neutral/broad focus declaratives.[14]

After eliminating disfluent tokens (due to pauses, speech errors, throat clearing, etc.), the final dataset consisted of 1,146 word types, which equals to 3,252 word tokens and 11,388 syllables. A breakdown of the dataset by speaker is provided in Table 3[15], and by syllable count in Table 4.

Table 3. The final dataset broken down by speaker

speaker	words		syllables
	types	tokens	
F1	165	500	1,758
F2	167	502	1,770
F3	167	505	1,768
F4	169	529	1,851
F5	168	511	1,794
M1	142	177	597
M2	168	528	1,850
Total	**1,146**	**3,252**	**11,388**

Table 4. The final dataset broken down by syllable count

σ count	2 σ	3 σ	4 σ	5 σ	**Total**
n(tokens)	597	1,038	1,012	605	**3,252**

The data was manually annotated in Praat (Boersma & Weenink 2021) by trained research assistants, who labelled words and syllables using the segmentation criteria in Machač & Skarnitzl (2009). Duration, maximum intensity, and mean F0 of each syllable, as well as F0 at four fixed points throughout a syllable (25%, 50%, 75%, right edge) were measured using a Praat script based on Elvira-García (2014); all F0 measurements were performed in Hz. Note that the intensity data excluded data from speakers F2 and M2 because they were recorded with a lavalier microphone with gain normalization; the total item counts used

---

[13] Occasionally, a speaker would repeat the stimulus more than three times; for these speakers, the total number of tokens may exceed the total number of types multiplied by three.

[14] Despite the effort to eliminate the possibility of contrastive reading on the stimuli, speakers may have interpreted the stimuli as narrowly focused, which could have affected their prosodic realization. Since Sluijter & van Heuven's (1996) seminal work, it is considered best practice to also record stimuli that are explicitly out-of-focus within the carrier phrase (i.e., in the presence of another narrowly focused constituent). This was done as part of the supplementary study (Datasets 2a and 2b).

[15] Speaker M1 contributed considerably fewer tokens than the other speakers because he used list intonation much more frequently than other participants. His data is included into the final dataset to make it more gender-balanced. Its exclusion or inclusion does not affect the results.

for the investigation of intensity are provided in Table 10 in Section 4.2. Statistical analysis of the data was carried out using the `glmer` function in the lme4 package (Bates et al. 2015) for R (R Core Team 2020).

4   Results

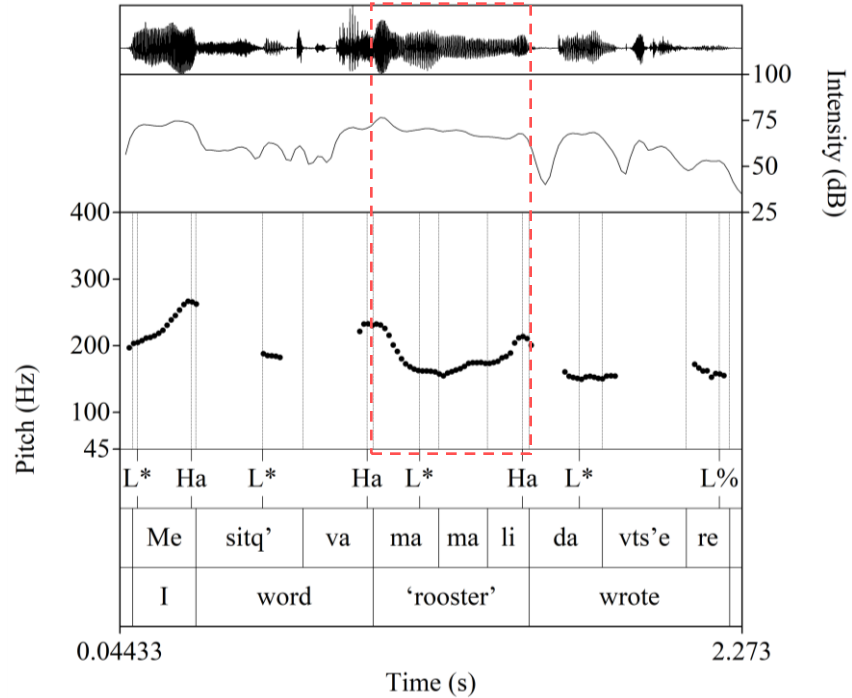Typical responses, for the sake of illustration, are provided in Figure 7 and Figure 8.

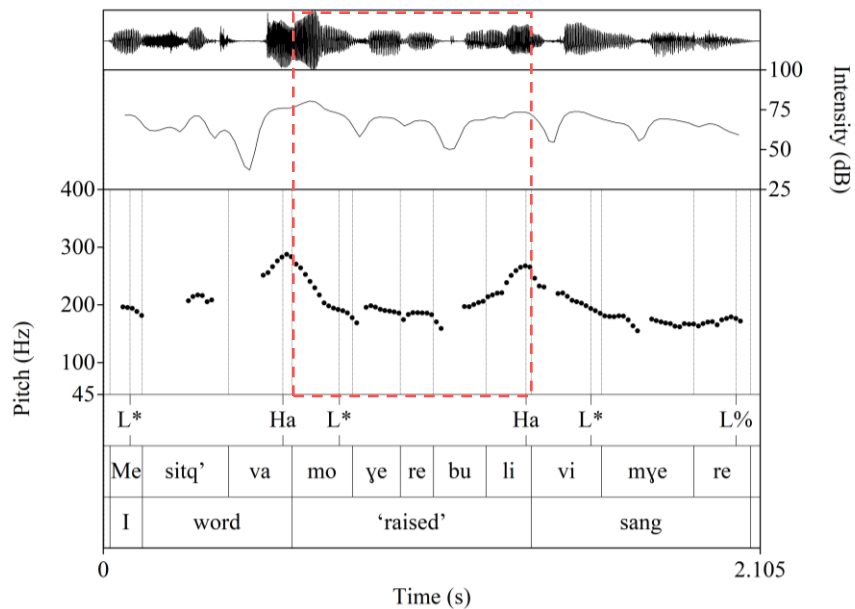Figure 7. A trisyllabic test-word *mamali* 'rooster' in a carrier phrase

Figure 8. A pentayllabic test-word *moɣerebuli* 'raised.PTCP' in a carrier phrase

14

4.1 Duration

The duration results are visualized in Figure 9, and mean values per syllable are provided in Table 5. They show that the initial syllable has greater duration than all subsequent syllables in words of all syllable counts (two to five syllables).[16]
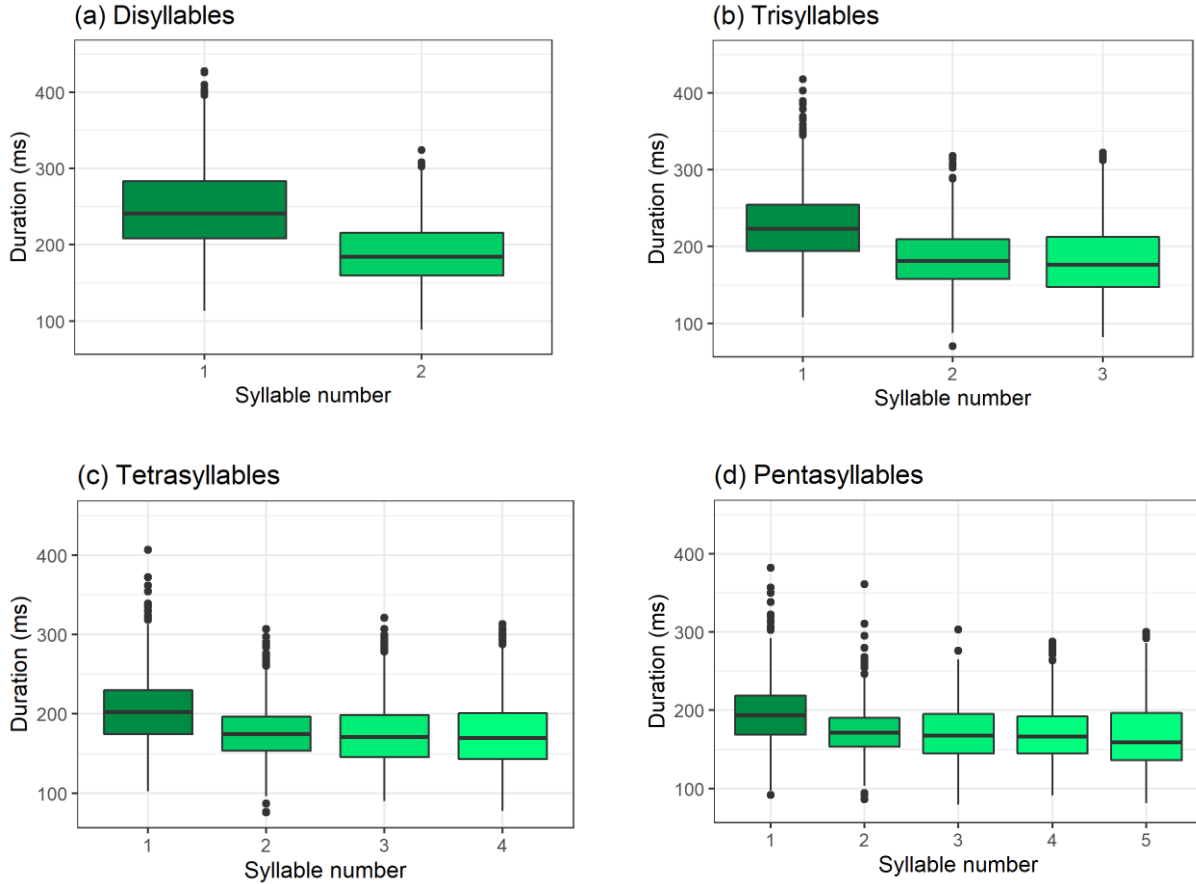

Figure 9. Syllable duration in words 2-5 syllables long

Table 5. Mean syllable duration (ms) and standard deviation (in brackets, ms) in words 2-5 syllables long

σ no. → / σ count ↓	1st σ	2nd σ	3rd σ	4th σ	5th σ
2 σ	250 (58)	189 (44)			
3 σ	226 (46)	185 (40)	181 (45)		
4 σ	205 (41)	177 (35)	174 (39)	174 (43)	
5 σ	196 (39)	173 (32)	170 (37)	171 (37)	168 (45)
Mean (SD)	**218 (50)**	**181 (38)**	**176 (41)**	**173 (40)**	**168 (45)**

---

[16] Syllable duration rather than vowel duration alone was measured here because the durational effect of stress may affect the consonant(s) in the stressed syllable, either in addition to or instead of the vowel: e.g., stressed syllables are marked by lengthened onsets in Estonian (Gordon 1995; Lehiste 1966) and codas in Welsh (Williams 1999); for more examples and discussion, see Gordon and Roettger (2017). The inherent durational differences between onset consonants of different types (stops, liquids, and nasals) are captured by the statistical model with the random intercept ITEM (which also captures any inherent differences between vowel types).

Before carrying out the statistical analysis, a mixed-effects model was fit to the data to test for the effect of polysyllabic shortening (Lehiste 1972). SYLLABLE DURATION was taken as the dependent variable, SYLLABLE NUMBER (1st, 2nd, etc; categorical factor) and SYLLABLE COUNT (per word) as fixed effects and SPEAKER and ITEM as random intercepts.[17] A main effect of SYLLABLE COUNT was obtained ($\chi^2(3)$=88.7, p<0.0001). Therefore, to control for the effect of SYLLABLE COUNT, manifested as polysyllabic shortening, the model testing for durational differences, introduced below, was run separately for words of each syllable count, in order to have a group-specific intercept for each group.

For a mixed-effects model analysis, SYLLABLE DURATION was taken as the dependent variable, SYLLABLE NUMBER as a fixed effect and SPEAKER and ITEM as random intercepts. The model was run three times, with the initial syllable, penult, and antepenult (the positions described as locations of stress in the previous literature) as intercepts. Including random slopes improved the fit for some of the models: SYLLABLE NUMBER by SPEAKER and by ITEM for disyllables, SYLLABLE NUMBER by SPEAKER for trisyllables, and SYLLABLE NUMBER by ITEM for tetrasyllables. As the results in Table 6, Table 7, and Table 8 show, only initial syllables are significantly different in their duration from all other syllables in words of all syllable counts.[18] This corresponds to the picture presented in Figure 9 and Table 5, where the main durational difference is shown to be between the initial syllable and all subsequent syllables. Note that in Table 7 and Table 8 the syllable counts where (ante)penults correspond to initial syllables are omitted because these results are presented in Table 6.

Table 6. Results of statistical analysis of syllable duration for all syllables, with the initial syllable as the intercept. Asterisks mark levels of significance: p<0.05 (*), p<0.01 (**), and p<0.001(***).

σ no. → σ count ↓	1st σ	2nd σ	3rd σ	4th σ	5th σ
2 σ		*p<0.0001**** *β=-0.29* *t=-5.38*			
3 σ		*p<0.0001**** *β=-0.21* *t=-6.91*	*p<0.0001**** *β=-0.24* *t=-7.4*		
4 σ		*p<0.0001**** *β=-0.15* *t=-4.82*	*p<0.0001**** *β=-0.18* *t=-5.7*	*p<0.0001**** *β=-0.16* *t=-6.14*	
5 σ		*p<0.0001**** *β=-0.13* *t=-13.09*	*p<0.0001**** *β=-0.15* *t=-15.04*	*p<0.0001**** *β=-0.14* *t=-14.71*	*p<0.0001**** *β=-0.16* *t=-17.07*

[17] Including random slopes led to non-convergence of the model.

[18] The duration of stay in the US did not give rise to significant inter-speaker differences. A mixed-effects model with DURATION (of the initial syllable, antepenult, or penult) as a dependent variable, SPEAKER LOCATION (US LONG-TERM vs. OTHER) as a fixed effect, and SPEAKER and ITEM as random effects did not detect any significant differences between speakers from different locations (initial: p=0.54, $β$=0.08, t=0.61; antepenult: p=0.65, $β$=-0.06, t=-0.46; penult: p=0.55, $β$=-0.08, t=-0.6). Similarly, a model with F0 as a dependent variable (other parameters held constant) did not detect significant differences between the two populations either (initial: p=0.42, $β$=-0.21, t=-0.8; antepenult: p=0.42, $β$=-0.21, t=-0.82; penult: p=0.56, $β$=-0.15, t=-0.59). No significance was obtained for intensity, either (other parameters held constant; initial: p=0.17, $β$=-0.078, t=-1.38; antepenult: p=0.17, $β$=-0.09, t=-1.37; penult: p=0.34, $β$=-0.07, t=-0.96).

Table 7. Results of statistical analysis of syllable duration with the penultimate syllable as the intercept

σ no. → σ count ↓	1st σ	2nd σ	3rd σ	4th σ	5th σ
3 σ	*p<0.0001*** *β=0.21* *t=6.91*		*p=0.25* *β=-0.03* *t=-1.15*		
4 σ	*p<0.0001*** *β=0.18* *t=6.87*	*p=0.17* *β=0.02* *t=1.36*		*p=0.63* *β=-0.01* *t=-0.48*	
5 σ	*p<0.0001*** *β=0.14* *t=14.71*	*p=0.1* *β=0.02* *t=1.62*	*p=0.74* *β=-0.003* *t=-0.33*		*p<0.05* *β=-0.02* *t=-2.35*

Table 8. Results of statistical analysis of syllable duration with the antepenultimate syllable as the intercept

σ no. → σ count ↓	1st σ	2nd σ	3rd σ	4th σ	5th σ
4 σ	*p<0.0001*** *β=0.15* *t=5.04*		*p=0.17* *β=-0.02* *t=-1.37*	*p=0.35* *β=-0.03* *t=-0.93*	
5 σ	*p<0.0001*** *β=0.15* *t=15.04*	*p=0.05* *β=0.02* *t=1.95*		*p=0.74* *β=0.003* *t=0.33*	*p<0.05* *β=-0.02* *t=-2.02*

Anticipating the discussion of the duration results in Section 6, two follow-up tests were carried out to exclude alternative accounts of the results. The first one checked whether the durational effect on the initial syllable is driven by vowel quality; the second one verified whether this effect results from an increase in the duration of the consonant, vowel, or both segments.

The first test checked whether non-high vowels occur more frequently in initial syllables than in subsequent syllables, which would affect the results, since non-high vowels have inherently greater durations. For this, the initial syllables in the dataset were coded for vowel height: NON-HIGH ([a, o, e]) and HIGH ([i, u]). The mean duration of NON-HIGH initial syllables equals 220 ms, and that of HIGH initial syllables is 210 ms. A linear mixed-effects model, with DURATION as a dependent variable, VOWEL HEIGHT as a fixed effect, and ITEM and SPEAKER as random intercepts, revealed no significant difference between the two vowel heights (p=0.12, $β$=0.04, t=1.55). Accordingly, the durational effect on the initial syllable cannot be explained away as stemming from vowel quality.[19]

The second test checked the relative contribution of the consonants and vowels in initial syllables to their duration. To this end, relative duration of the segments in the initial syllable was compared to that of the segments in the second syllable. A subset of stimuli from the experiment (n of tokens = 1,228), representative of the full dataset, was selected, and the initial and second syllables were annotated for

---

[19] There is some evidence that a minimal durational difference between vowels of different heights is expected in Georgian: Shosted & Chikovani (2006: 262) found that Georgian /u/ and /ɔ/ largely overlap, and, likewise, /i/ and /ɛ/ have similar F1 values. Given that vowel height is a predictor of inherent vowel duration, the fact that (some of) the high and non-high vowels have comparable F1 values is likely to contribute to the results.

segment duration.[20] Since all syllables in the study are of CV shape, the ratio of the two segments in different syllables can be compared. The segment duration results are provided in Table 9:

Table 9. Duration of individual segments in the initial and second syllables (ms) and standard deviation (in brackets, ms)

Syllable	1st		2nd	
Segment	C_1	V_1	C_2	V_2
Mean duration (SD)	103 (30)	105 (27)	83 (20)	96 (24)

As Table 9 shows, the increase in duration that the initial syllables receive, as compared to the second syllables, does not result from an increase in the duration of just one segment – instead, both receive an increase in duration. A linear mixed effects model, with SEGMENT DURATION as a dependent variable, SEGMENT POSITION as a fixed effect, and ITEM and SPEAKER as random intercepts, shows that the duration of consonants in the initial syllables is significantly greater than that of consonants in the second syllables ($p<0.0001$***, β=-0.22, t=-22.15). Vowels in the initial syllables are also significantly greater in duration than vowels in the second syllables ($p<0.0001$***, β=-0.1, t=-10.98). These results confirm that the greater duration of the initial syllable does not stem from greater duration of the initial consonant alone. The significance of this fact is further taken up in Section 6.

4.2 Intensity

Table 10 provides the numerical breakdown of the subset of the data used for the study of intensity (with the data from speakers M2 and F2 excluded).

Table 10: The dataset used for intensity measurements, by syllable count

σ count	2 σ	3 σ	4 σ	5 σ	**Total**
n(tokens)	412	716	684	412	**2,224**

The intensity results (maximum intensity per syllable) are shown in Figure 10, and averaged values per syllable are provided in Table 11. They show that the initial syllable has greater intensity than all subsequent syllables in words two to five syllables long (note that while the magnitude of differences is small, it is likely to be perceptually substantial).

Table 11. Maximum syllable intensity, averaged per syllable number (dB), and standard deviation (in brackets, dB) in words 2-5 syllables long

σ no. → σ count ↓	1st σ	2nd σ	3rd σ	4th σ	5th σ
2 σ	67.52 (7.11)	66.04 (7.7)			
3 σ	67.22 (6.32)	65.28 (6.44)	64.86 (7.5)		
4 σ	67.06 (6.32)	66.22 (6.61)	63.98 (6.83)	63.79 (7.8)	
5 σ	67.64 (6.49)	66.00 (6.75)	65.3 (6.76)	63.04 (7.12)	63.31 (8.3)
Mean (SD)	**67.3 (6.5)**	**66.03 (6.82)**	**64.63 (7.1)**	**63.5 (7.55)**	**63.31 (8.3)**

---

[20] A representative subset of the full dataset was created in the following way. 15 words of each syllable count (di-, tri-, tetra- and pentasyllables; total=60/172 tokens) were selected for by-segment annotation. Since the first two syllables were the target of the additional analysis, their parameters were the crucial selection criteria. With respect to the $C_1V_1C_2V_2$ template, words with both identical and non-identical C_1 and C_2 and V_1 and V_2 were selected (in roughly equal ratios, i.e., per sample of stimuli of a given syllable count, ca. ¼ = $C_xV_aC_xV_a$, ¼ = $C_xV_aC_yV_a$, ¼ = $C_xV_aC_xV_b$, ¼ = $C_xV_aC_yV_b$).
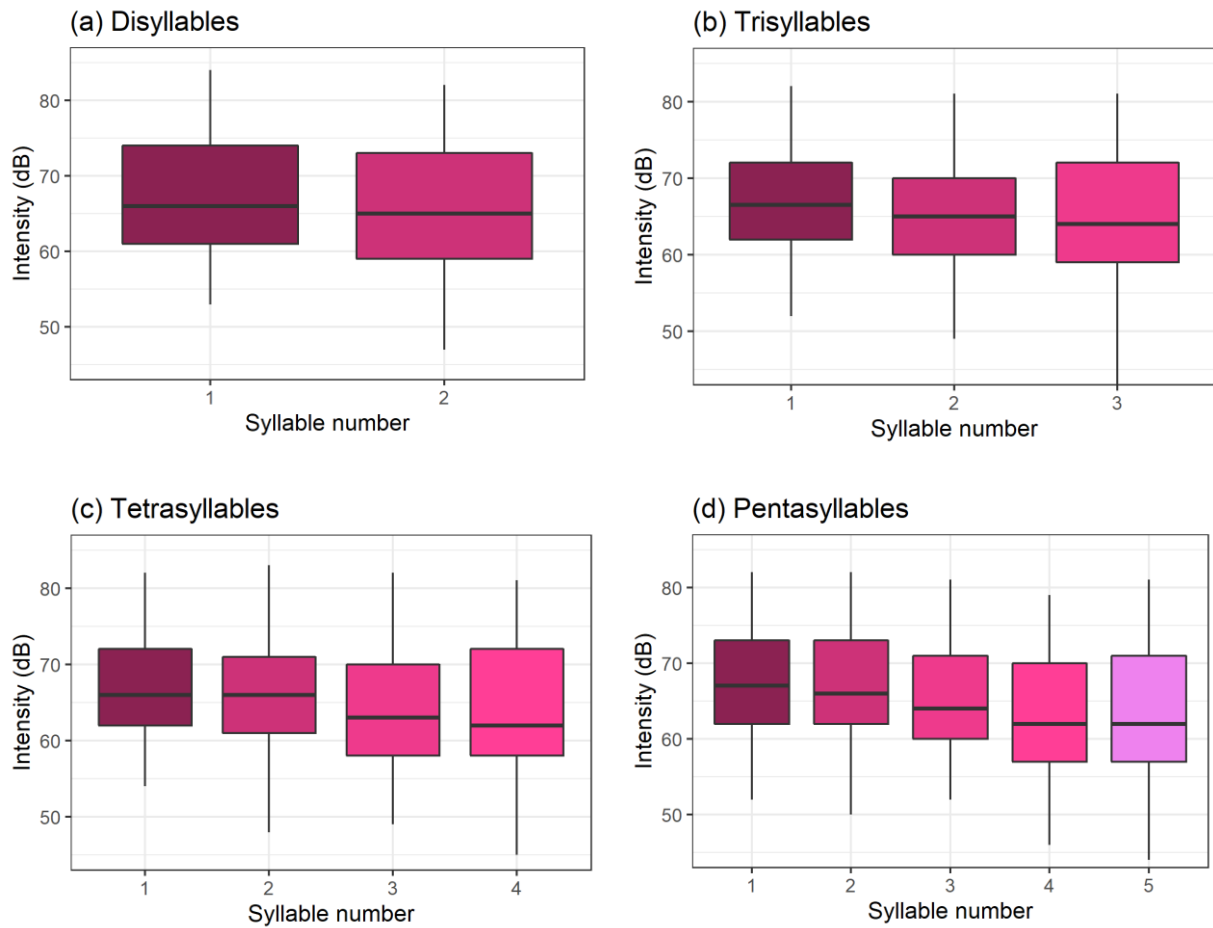
Figure 10. Syllable intensity in words 2-5 syllables long

For a mixed-effects model analysis, MAX INTENSITY (per syllable) was taken as the dependent variable, SYLLABLE NUMBER (1st, 2nd, etc.) as a fixed effect and SPEAKER and ITEM as random intercepts. The model was run separately for words of each syllable count, with the initial syllable, penult, and antepenult as intercepts. Including random slopes improved the fit for some of the models: SYLLABLE NUMBER by SPEAKER and by ITEM for disyllables, and SYLLABLE NUMBER by SPEAKER for trisyllables.

As the results in Table 12, Table 13, and Table 14 show, most differences in intensity are significant or borderline significant. Note, though, that Table 13 demonstrates lack of significance between the penult and ultima, which aligns well with what we see in Figure 10 and Table 11: the penult and ultima consistently have comparable intensity values. Significant differences between other adjacent syllables correspond to the gradual drop in intensity between the syllables other than penult and ultima, also shown in Figure 10 and Table 11.

Table 12. Results of statistical analysis of intensity for all syllables, with the initial syllable as the intercept. Asterisks mark levels of significance: *p*<0.05 (*), *p*<0.01 (**), and *p*<0.001(***).

σ no. → / σ count ↓	1st σ	2nd σ	3rd σ	4th σ	5th σ
2 σ		*p=0.08* *β=-0.02* *t=-1.74*			
3 σ		*p=0.05* *β=-0.03* *t=-1.96*	*p=0.08* *β=-0.04* *t=-1.76*		
4 σ		***p<0.0001**** *β=-0.01* *t=-3.99*	***p<0.0001**** *β=-0.05* *t=-14.71*	***p<0.0001**** *β=-0.05* *t=-15.98*	
5 σ		***p<0.05**** *β=-0.01* *t=-2.21*	***p<0.0001**** *β=-0.04* *t=-7.81*	***p<0.0001**** *β=-0.07* *t=-15.73*	***p<0.0001**** *β=-0.07* *t=-15.12*

Table 13. Results of statistical analysis of intensity with the penultimate syllable as the intercept

σ no. → / σ count ↓	1st σ	2nd σ	3rd σ	4th σ	5th σ
3 σ	*p=0.05* *β=0.03* *t=1.96*		*p=0.53* *β=-0.01* *t=-0.63*		
4 σ	***p<0.0001**** *β=0.05* *t=14.71*	***p<0.0001**** *β=0.04* *t=10.71*		*p=0.16* *β=-0.01* *t=-1.39*	
5 σ	***p<0.0001**** *β=0.07* *t=15.73*	***p<0.0001**** *β=0.06* *t=13.53*	***p<0.0001**** *β=0.04* *t=-7.93*		*p=0.54* *β=0.002* *t=0.61*

Table 14. Results of statistical analysis of intensity with the antepenultimate syllable as the intercept

σ no. → / σ count ↓	1st σ	2nd σ	3rd σ	4th σ	5th σ
4 σ	***p<0.0001**** *β=0.01* *t=3.37*		***p<0.0001**** *β=-0.04* *t=-10.6*	***p<0.0001**** *β=-0.06* *t=-14.29*	
5 σ	***p<0.0001**** *β=0.04* *t=6.81*	***p<0.0001**** *β=0.03* *t=5.02*		***p<0.0001**** *β=-0.04* *t=-7.47*	***p<0.0001**** *β=-0.05* *t=-8.81*

4.3 F0 values

Figure 11 visualizes the F0 contours that span the test words of different syllable counts, and Table 15 provides the mean F0 values per syllable.
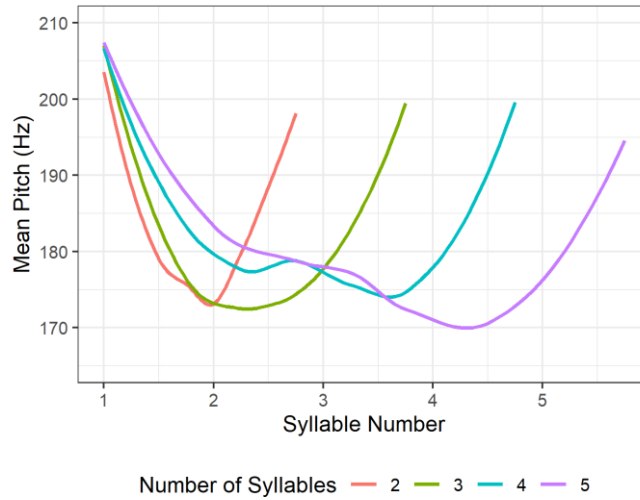

Figure 11. Mean F0 values at four points per syllable in stimuli of all syllable counts; smoothed at 0.6; tick marks correspond to the left edges of respective syllables (e.g., '1' marks the left edge of the first syllable).

Table 15. Mean F0 values (Hz) and standard deviations (in brackets, Hz) per syllable in stimuli of all syllable counts

σ no. → σ count ↓	1st σ	2nd σ	3rd σ	4th σ	5th σ
2 σ	175.75	188.76			
	(47.64)	(50.14)			
3 σ	178.01	173.33	190.18		
	(47.37)	(48.86)	(53.15)		
4 σ	180.68	178.49	172.54	191.05	
	(48.25)	(47.50)	(45.85)	(51.88)	
5 σ	182.39	179.32	172.68	169.74	187.54
	(48.28)	(46.81)	(42.86)	(42.65)	(50.71)
Mean	**179.25**	**178.87**	**179.41**	**182.91**	**187.54**
(SD)	**(47.9)**	**(48.55)**	**(48.91)**	**(49.64)**	**(50.71)**

As Figure 11 shows, words of all syllable counts have an overall falling-rising F0 contour. The F0 target on the initial syllable (L*) is realized as a sharp dip from the high target (Ha) on the preceding word, and the high target on the ultima (Ha) is realized as a steep rise. In words of all syllable counts, the lowest F0 values are found on the penult. The penultimate syllable, therefore, acts as a turning point between the falling and rising subparts of the F0 contour. The L* Ha tonal contour is expected to span APs in broad-focus declaratives, as discussed in Section 2.3, but the low target on the penultimate syllable has not been discussed in this context before. Its distribution, though, fits well with that of a low phrase accent, which targets penultimate syllables of predicates in focus contexts, as discussed in Section 2.3.

For a mixed-effects model analysis, MEAN F0 (per syllable) was taken as the dependent variable, SYLLABLE NUMBER as a fixed effect and SPEAKER and ITEM as random intercepts. As was the case with the

21

duration and intensity measurements, including random slopes improved the fit for some of the models: SYLLABLE NUMBER by SPEAKER and by ITEM for disyllables, and SYLLABLE NUMBER by SPEAKER for trisyllables. The model was run separately for words of each syllable count. In turns, the initial syllable, penult, and antepenult were taken as intercepts.

As the results in Table 16, Table 17, and Table 18 show, most comparisons turned out to be significant, except those between the initial and second syllables in tri-, tetra- and pentasyllabic words, second and third syllables in disyllabic words, and the third and fourth syllables in pentasyllabic words.

Table 16. Results of statistical analysis of F0 for all syllables, with the initial syllable as the intercept. Asterisks mark levels of significance: $p<0.05$ (*), $p<0.01$ (**), and $p<0.001$(***).

σ no. → σ count ↓	1st σ	2nd σ	3rd σ	4th σ	5th σ
2 σ		*p<0.01** β=0.08 t=2.67*			
3 σ		*p=0.64 β=-0.02 t=-0.46*	*p=0.18 β=0.07 t=1.33*		
4 σ		*p=0.08 β=-0.01 t=-1.76*	*p<0.001*** β=-0.04 t=-6.76*	*p<0.001*** β=0.06 t=9.23*	
5 σ		*p=0.06 β=-0.01 t=-1.88*	*p<0.001*** β=-0.05 t=-6.49*	*p<0.001*** β=-0.06 t=-8.34*	*p<0.001*** β=0.04 t=4.95*

Table 17. Results of statistical analysis of F0 for all syllables, with the penult as the intercept

σ no. → σ count ↓	1st σ	2nd σ	3rd σ	4th σ	5th σ
3 σ	*p=0.64 β=0.02 t=0.46*		*p<0.05* β=0.08 t=2.4*		
4 σ	*p<0.001*** β=0.04 t=6.76*	*p<0.001*** β=0.03 t=5.02*		*p<0.001*** β=0.1 t=15.79*	
5 σ	*p<0.001*** β=0.06 t=8.34*	*p<0.001*** β=0.05 t=6.48*	*p=0.052 β=0.02 t=1.94*		*p<0.001*** β=0.1 t=13.12*

Table 18. Results of statistical analysis of F0 for all syllables, with the antepenult as the intercept

σ no. → σ count ↓	1st σ	2nd σ	3rd σ	4th σ	5th σ
4 σ	*p=0.08* *β=0.01* *t=1.76*		***p<0.001**** *β=-0.03* *t=-5.02*	***p<0.001**** *β=0.07* *t=10.94*	
5 σ	***p<0.001**** *β=0.05* *t=6.49*	***p<0.001**** *β=0.04* *t=4.6*		*p=0.052* *β=0.02* *t=1.94*	***p<0.001**** *β=0.09* *t=11.31*

To sum up, the results show that greater duration consistently marks initial syllables, but not penults or antepenults. The intensity values fall gradually throughout the word and are the lowest on the penult and ultima. The F0 contours attest to the presence of low F0 targets on the initial syllable (L*) and the penult (L), and a high one on the ultima (Ha). These F0 results differ somewhat from the typical declarative broad-focus F0 contour, described in Section 2.3 as L* Ha, based on Vicenik & Jun (2014).

5 Supplementary study

5.1 Stimuli and design

The test words in the experiment regularly correspond to APs, as was shown in (8). Accordingly, the isomorphism between words and phrases does not allow for distinguishing between word- and phrase-level prosodic effects. To circumvent this, as well as the potential confound created by the final Ha at the end of a preceding AP, data from a supplementary study was additionally brought in. The study had been conducted to elicit natural word orders and prosodic realizations in different focus contexts, but, importantly for the purposes of the current paper, the data also contains examples of ip-formation and post-focal deaccenting.

The study consisted of 30 scenarios, based on 14 transitive and 16 intransitive verbs, such as *A girl picked flowers last summer* or *Birds returned home last spring,* rendered as picture prompts. Personal names and common nouns (*Mariami*, *children,* etc.) were used as subjects and objects. Lexical items containing no/few voiceless segments were preferred but non-CV syllables and consonant clusters were allowed. For each scenario, five questions were constructed, aimed at eliciting broad, VP-, subject, and object focus, as well as contrastive focus on either subject or object. A sample picture prompt is provided in Figure 12, with the questions listed in (9).

Figure 12. Sample picture prompt

(9)　　a.　　What happened last summer?
　　　　b.　　What did the girl pick last summer?

c.   Who picked flowers last summer?

d.   What did the girl do last summer?

e.   Did the girl pick cherries last summer?'

Participants were presented with picture prompts on a laptop screen (pseudo-randomized), each accompanied by a question, and were tasked with answering the question based on the picture. They were instructed to speak clearly, use natural intonation and avoid single-word replies, but otherwise were free to construct their own responses. Eight native speakers of Georgian participated in the study: two males (M3, M4) and six females (F6-F11), with the age range 20-35 y.o, mean age 26.9 y.o. All speakers were from Tbilisi and had a complete or in-progress university degree. The recordings were performed in Tbilisi, Georgia, using a head-worn microphone. All data were recorded at a sampling rate of 44.100 Hz and 16 bits per sample.

For the purposes of the current paper, four datasets were compiled from the data obtained during the focus study, used here as a supplementary one. First, to ensure full comparability of the results with those from the main experiment, only nominals, two to five syllables long, of CV syllable structure, where onsets consist of a single voiced consonant were considered. Dataset 1 was compiled to investigate the prosody of nouns that form part of an ip together with a modifying adjective. Because of the restrictions on the phonological shape of the noun, only one nominal phrase from the supplementary study was selected – *gogo* 'girl' accompanied by the adjective *mousvenari* 'mischievous, antsy' – which appeared in two scenarios: *A mischievous girl stole cherries last summer* and *A mischievous girl fell from a ladder last week.* In total, the phrase *mousvenari gogo* occurred in the participants' responses 49 times. These occurrences were further restricted to those in which the phrase *mousvenari gogo* occurred (i) not under narrow focus (i.e., in broad, VP-, object, or contrastive object focus), and (ii) not clause-finally (i.e, in SOV or SVO clauses), which yielded 35 occurrences. Of these, in 8 instances the words *mousvenari* and *gogo* were realized as independent APs (cf. Figure 3) and were also excluded. The remaining 27 occurrences of *mousvenari gogo* formed ips (cf. Figure 2), and the realizations of *gogo* were selected for the same analysis as nominals in the main experiment as Dataset 1a.

To supplement Dataset 1a, following a suggestion from a reviewer, nominal phrases that do not adhere to the CV condition on syllable structure were also considered, which allowed for analyzing the nominal phrase *lamaz-ma kalbat'on-ma* 'beautiful-ERG lady-ERG'. Keeping the remaining restrictions in place (the nominal phrase occurring not under narrow focus, not clause-finally, and as a single ip) yielded 33 occurrences of *lamaz-ma kalbat'on-ma* 'beautiful-ERG lady-ERG'. The realizations of the noun *kalbat'on-ma,* extracted from these examples, formed Dataset 1b. Given the variable syllable composition, which can impact duration values, the noun was annotated by segment (in addition to the by-syllable annotation), and only the intensity and F0 values of the vowels were analyzed. Datasets 1a and 1b were analyzed separately.

Dataset 2 was aimed at investigating the prosody of post-focal deaccenting. For this, objects in the context of subject focus or contrastive subject focus, in S_FVOX word orders, were selected ('X' stands for a clause-final adverbial that buffers the object from the right-edge of the clause, to ensure lack of clause-final effects on the object). In Dataset 2a, the same restrictions on the phonological shape of the noun applied as did with Dataset 1a, and only one noun, *banan-ebi* 'banana-PL', was selected for analysis. It occurred as the object in S_FVOX word orders 12 times, of which, in 2 instances, it was not deaccented (cf. Figure 4). These occurrences were excluded, and the remaining 10 instances, in which *banan-ebi* 'banana-PL' was deaccented (cf. Figure 5) were subjected to the same analysis as nominals in the main experiment.

Additionally, following a reviewer's suggestion, nouns that do not have CV syllable structure but otherwise adhere to the same restrictions as those in Dataset 2a were extracted from the results of the supplementary study. Several such nouns were available, all trisyllabic (43 occurrences in total): *alubl-eb-s* 'cherry-PL-DAT', *mankana-s* 'car-DAT', *muraba-s* 'jam-DAT', *murab-eb-s* 'jam-PL-DAT', *q'vavil-i* 'flower-NOM', *q'vavil-eb-s* 'flower-PL-DAT'. They were pooled together, annotated by segment, and the intensity and F0 values of vowels were analyzed (duration was left out, as in Dataset 1a, given the

mismatches in syllable structure). Datasets 2a and 2b were analyzed separately. A summary of the four datasets is provided in Table 19.

Table 19. Summary of the four datasets based on the supplementary study

Dataset	Conditions	Test words	n	Segmentation	Measurements
1a	adjective+noun, single ip	*gogo* 'girl'	27	by syllable	duration maximum intensity average F0
1b	adjective+noun, single ip	*kalbat'on-ma* 'lady-ERG'	33	by segment (vowels only)	maximum intensity average F0
2a	post-focal deaccenting	*banan-ebi* 'banana-PL'	10	by syllable	duration maximum intensity average F0
2b	post-focal deaccenting	*alubl-eb-s* 'cherry-PL-DAT' *mankana-s* 'car-DAT' *muraba-s* 'jam-DAT' *murab-eb-s* 'jam-PL-DAT' *q'vavil-i* 'flower-NOM' *q'vavil-eb-s* 'flower-PL-DAT'	43	by segment (vowels only)	maximum intensity average F0

5.2 Dataset 1a (ip-formation)

A typical realization of an ip, formed by the noun *gogo* 'girl', accompanied by the adjective *mousvenari* 'mischievous', is provided in Figure 13, with glosses in (10):

(10)   ((Mousvenari)~AP~ (gogo)~AP~)~ip~   (i-p'ar-av-d-a)~AP~   (alubl-eb-s)~AP~.
        mischievous   girl   VER-find-SF-SM-IPFV.3SG   cherry-PL-DAT
        'A mischievous girl stole (the) cherries.'
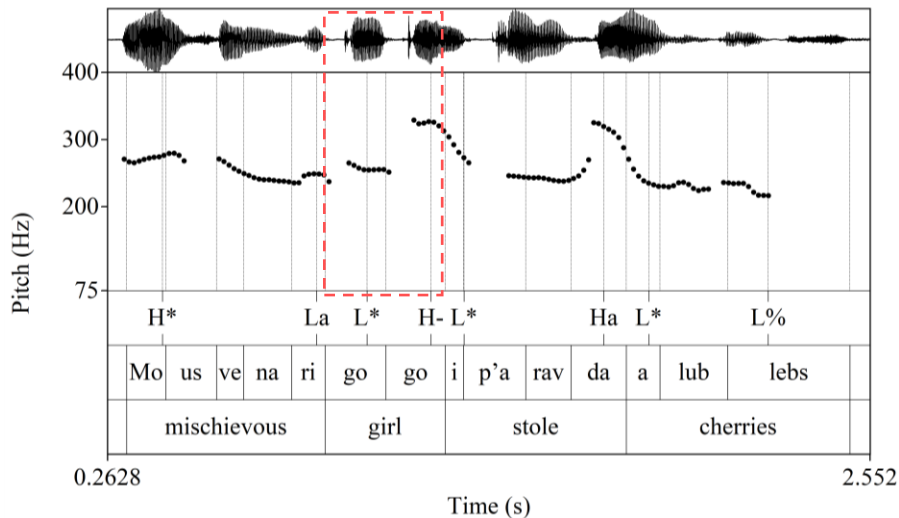

Figure 13. A realization of *gogo* 'girl' as part of an ip

25

Because *gogo* 'girl' is a disyllabic word, and the initial syllable is the only baseline for comparison, the three acoustic parameters (duration, intensity, F0) will be considered side by side. Figure 14 provides the visualization of the acoustic parameters, and the mean values are summarized in Table 20. The durational and F0 values of the two syllables in *gogo* exhibit a pattern comparable to that of disyllabic stimuli in the main experiment (except for the lack of a sharp fall from the preceding Ha on the initial syllable), but not the intensity values: here, the initial syllable is lower in intensity than the second one.

Given the small size of the Datasets 1a-b and 2a-b, no statistical analysis was carried out; while this makes the evidence more tentative, the overall trends are clear from the figures and summary tables.
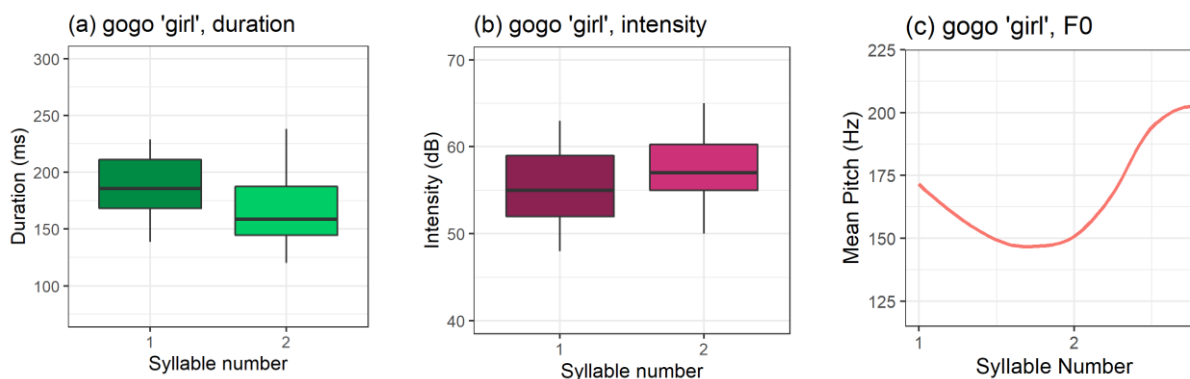

Figure 14. Duration, intensity and an average F0 contour (smoothed at 0.6) of *gogo* 'girl'

Table 20. Duration (ms), maximum intensity (dB), and mean F0 values (Hz), averaged per syllable in *gogo* 'girl'; in brackets, standard deviation is provided.

	1st σ	2nd σ
Duration	186 (27)	166 (32)
Intensity	55.58 (4.6)	57.58 (3.81)
F0	144.79 (57.58)	193.15 (74.03)

5.3   Dataset 1b (ip-formation)

A typical realization of the noun *kalbat'on-ma* 'lady-ERG' and adjective *lamaz-ma* 'beautiful-ERG', when they form a single ip, was provided in (5a) and Figure 2, repeated here as (11) and Figure 15, for convenience.

(11)   ((Lamaz-ma)$_{AP}$ (kalbat'on-ma)$_{AP}$)$_{\textbf{ip}}$ (k'aba)$_{AP}$ (mo-i-zom-a)$_{AP}$.
       beautiful-ERG   lady-ERG   dress.NOM   PV-VER-try-AOR.3SG
       'A beautiful lady tried on a dress.'

Due to variable syllable structure, only vowels (and not full syllables) in *kalbat'on-ma* 'lady-ERG' were considered, and only the (maximum) intensity and (average) F0 per vowel were measured. These parameters are visualized in Figure 16, and the mean values are summarized in Table 21. The pattern of the intensity values is somewhat similar to that obtained for tetrasyllabic words in the main experiment: the initial syllable has the greatest intensity, which then falls throughout the word; in contrast with the main experiment, it rises again on the ultima. The rise on the ultima is also reminiscent of the intensity pattern on *gogo* 'girl' in Dataset 1a. The averaged F0 pattern closely follows the F0 shape found on individual examples, like Figure 15, and that found in Dataset 1a: a flat F0 contour on the first three syllables is

followed by a sharp rise on the ultima (the absence of the F0 fall on the first syllable is, again, due to lack of a preceding Ha).
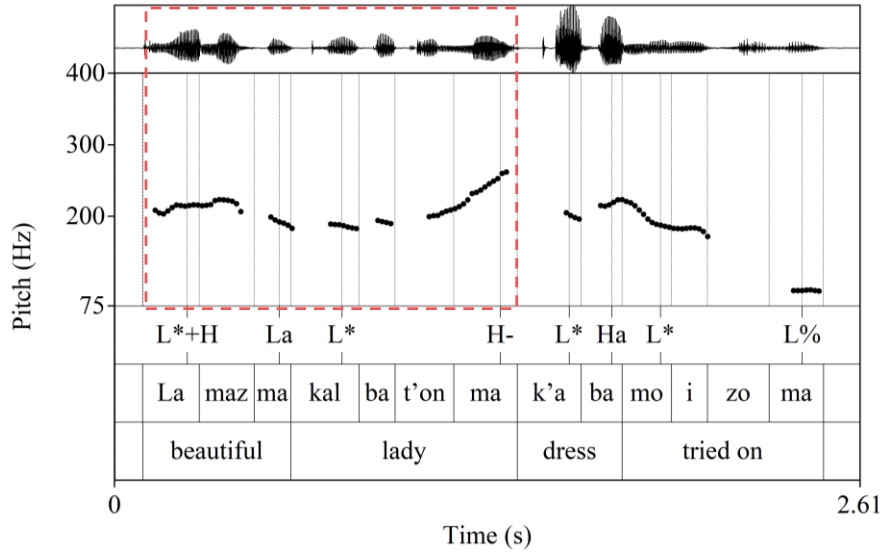

Figure 15. The intonational contour of an utterance containing *lamaz-ma kalbat'on-ma* 'beautiful-ERG lady-ERG' as an ip
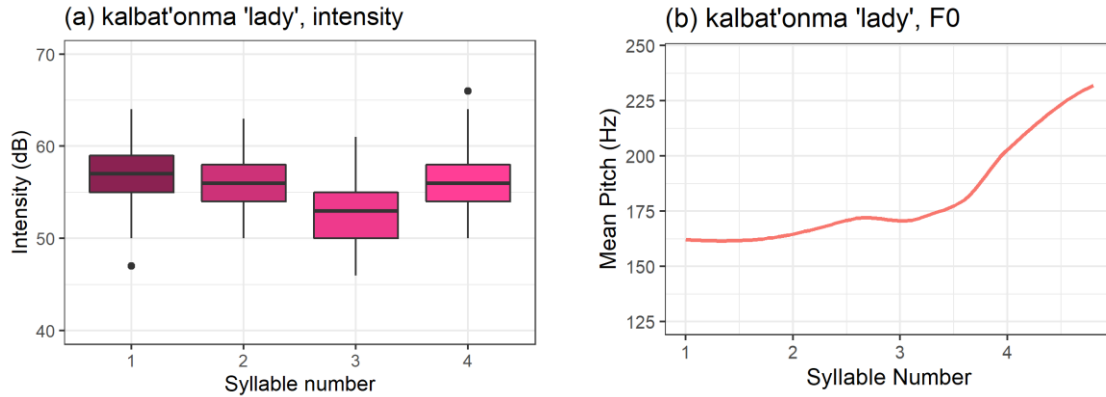

Figure 16. Intensity and an average F0 contour (smoothed at 0.6) of *kalbat'on-ma* 'lady-ERG'

Table 21. Duration (ms), maximum intensity (dB), and mean F0 values (Hz), averaged per syllable in *kalbat'on-ma* 'lady-ERG'; in brackets, standard deviation is provided.

	1st σ	2nd σ	3rd σ	4th σ
Intensity	56.82 (3.7)	56.15 (3.4)	52.91 (3.52)	55.88(3.84)
F0	163.85 (60.56)	161.97 (59.63)	172.56 (62.28)	223.48 (86.18)

5.4 Dataset 2a (post-focal deaccenting)

One of the typical realizations of *banan-ebi* 'banana-PL', when in the domain of post-focal deaccenting, is provided in Figure 17, with glosses and translation in (12). Note that the verb in Figure 17 is deaccented together with the other post-focal material, and does not carry pitch targets, unlike in Figure 5. Examples

27

of both kinds – i.e., those where the verb was part of the focal and post-focal domains – were represented in the sample.

(12) ('Who bought bananas last week?')
      (Manana-m$_F$)$_{AP}$ i-q'id-a   **banan-eb-i**   gasul   k'vira-s.
      Manana-ERG   VER-buy-AOR.3SG   banana-PL-NOM   last   week-DAT
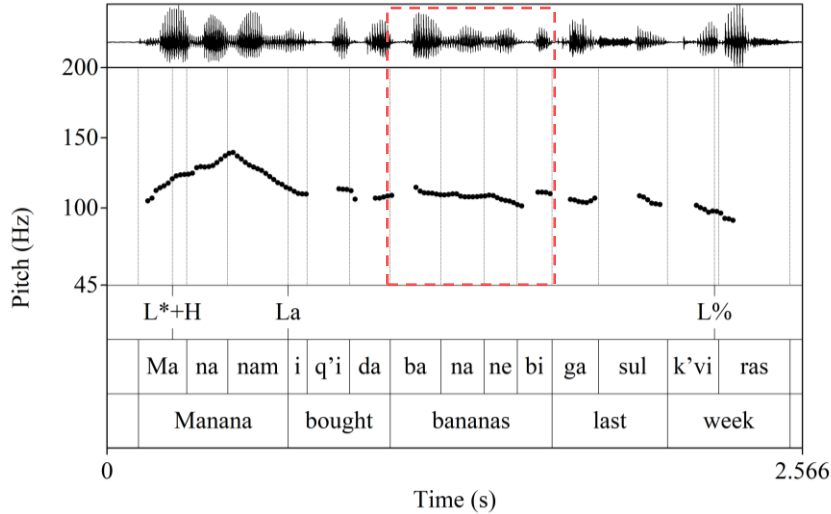   'Manana$_F$ bought bananas last week.'

Figure 17. A realization of *banan-ebi* 'banana-PL' within the domain of post-focal deaccenting

The duration and intensity results are visualized in Figure 18 and an averaged F0 contour is provided in Figure 19.
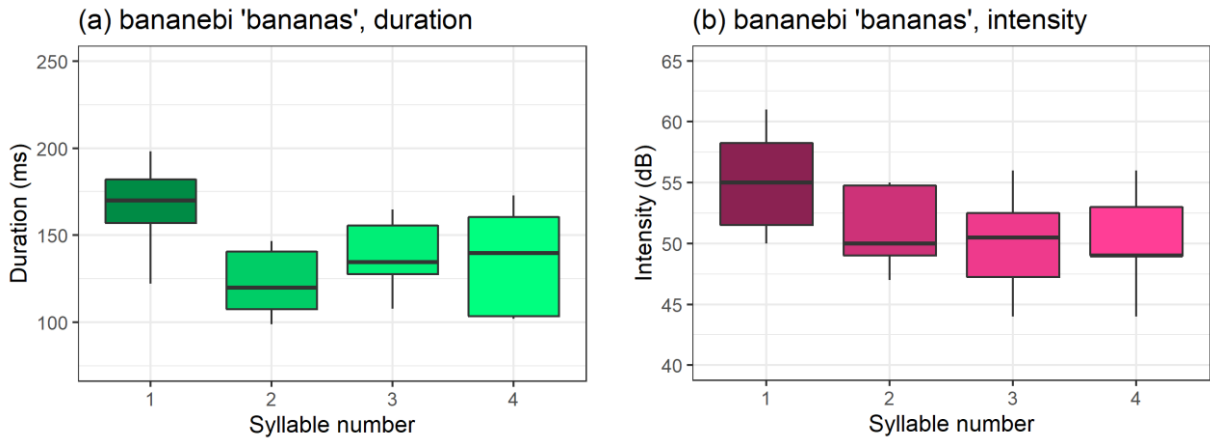

Figure 18. Syllable duration and intensity in *banan-ebi* 'banana-PL'
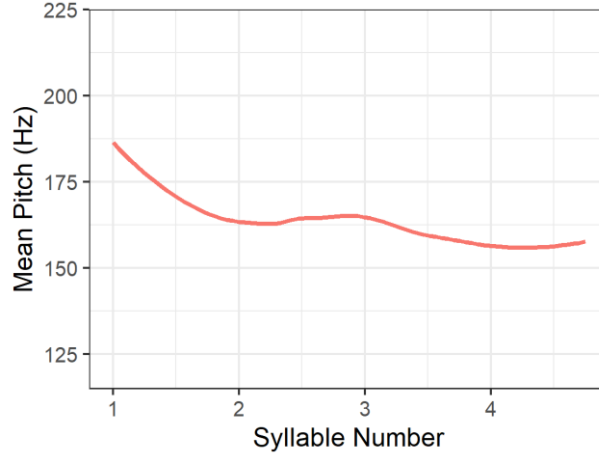
28

Figure 19. Average F0 contour on *banan-ebi* 'banana-PL'; smoothed at 0.6; each tick mark corresponds to the left edge of a corresponding syllable (e.g., '1' is the left edge of the first syllable).

Duration, maximum intensity, and mean F0 values, averaged per syllable, are provided in Table 22. The duration results are comparable to those obtained for tetrasyllables in the main experiment, in that the initial syllable has greater duration than all other syllables, though there is more variability in the durational values of the subsequent syllables. The intensity values show a similar trend to those of tetrasyllables in the main experiment, in that the initial syllable has the greatest intensity, and the final two syllables have similar intensity values. With respect to F0, the fall in F0 within the initial syllable is due to the presence of a high final boundary Ha on the preceding verb in some of the stimuli (cf. Figure 5). There is no evidence for there being F0 targets.

Table 22. Duration (ms), maximum intensity (dB), and mean F0 values (Hz), averaged per syllable in *banan-ebi* 'banana-PL'; in brackets, standard deviation is provided.

	1st σ	2nd σ	3rd σ	4th σ
Duration	165 (25)	122 (18)	139 (19)	135 (29)
Intensity	55.1 (4.2)	51.1 (3.35)	49.8 (3.8)	50.3 (3.88)
F0	162.85 (52.84)	162.12 (50.79)	166.74 (50.35)	149.6 (48.35)

5.5 Dataset 2b (post-focal deaccenting)

The words that comprised Dataset 2b (*alubl-eb-s* 'cherry-PL-DAT', *mankana-s* 'car-DAT', *muraba-s* 'jam-DAT', *murab-eb-s* 'jam-PL-DAT', *qvavil-i* 'flower-NOM', *qvavil-eb-s* 'flower-PL-DAT'), being deaccented postfocally, received realizations similar to that of *banan-ebi* 'banana-PL' in Dataset 2a. This was illustrated for *muraba-s* 'jam-DAT' in (6b) and Figure 5, repeated as (13) and Figure 20, for convenience.

(13) ('Who made jam last week?')
   ((Bebia**F**)AP   (a-ket-eb-d-a)AP)ip   **muraba-s**   ts'ina   k'vira-s.
   grandma.NOM   PV-make-SF-SM-IPFV.3SG   jam-DAT   last   week-DAT
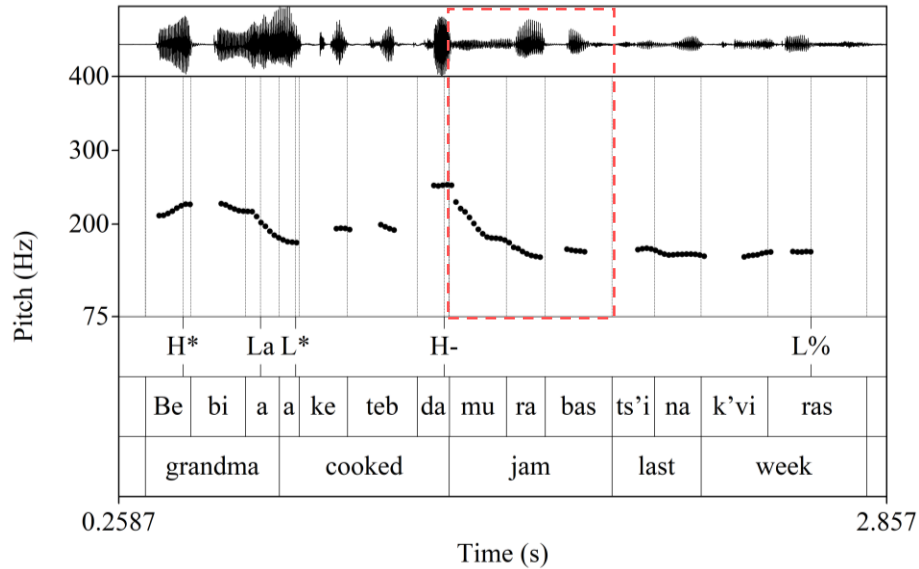   'Grandma**F** made jam last week.'

Figure 20. The intonational contour of an utterance with narrow focus on *bebia* 'grandma', with post-focal deaccenting on *muraba-s* 'jam-DAT'

As in Dataset 1b, due to mismatches in syllable structure, the acoustic parameters of individual vowels, not syllables, were considered, and included only intensity and F0 measurements. They are visualized in Figure 21, and the summary of the mean values is provided in Table 23. As in the previous datasets, intensity values are greatest on the initial syllable and drop on subsequent syllables. The F0 contour is low and flat, presenting no evidence for F0 targets.
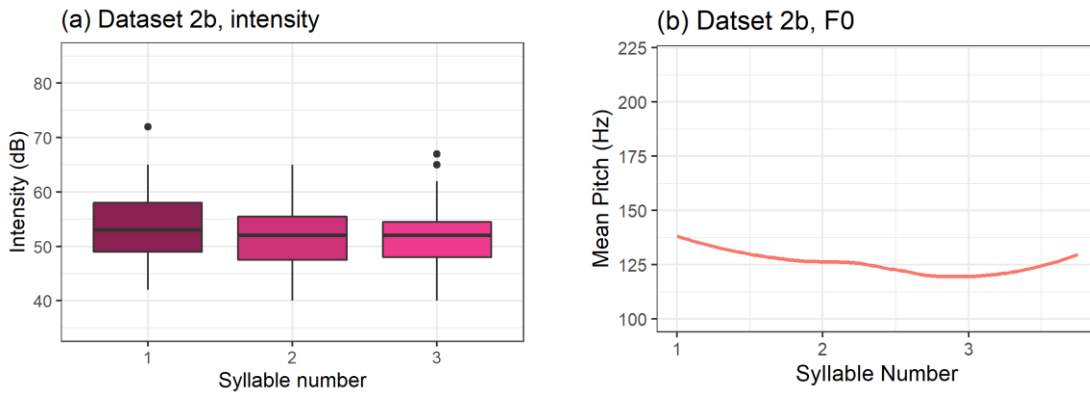

Figure 21. Intensity and an average F0 contour (smoothed at 0.6) of words in Dataset 2b.

Table 23. Duration (ms), maximum intensity (dB), and mean F0 values (Hz), averaged per syllable in the words in Dataset 2b; in brackets, standard deviation is provided.

	1st σ	2nd σ	3rd σ
Intensity	53.88 (6.23)	51.3 (5.91)	51.81 (6.37)
F0	142.77 (52.26)	142.34 (52.36)	152.13 (64.98)

Let us summarize the results of the supplementary study (Datasets 1a-b and 2a-b). When embedded into an ip, an AP/lexical word is still characterized by greater duration of the initial syllable and a falling-rising or rising F0 pattern. The intensity marking of the initial syllable was no longer there in the disyllabic

30

gogo 'girl' but is present in the longer *kalbat'onma* 'lady'. In turn, after undergoing post-focal deaccenting, lexical words carry a flat F0 contour, but initial syllables are still marked by greater duration and intensity.

6 Discussion

Let us bring together the results of the main experiment and the supplementary study. The most consistent finding is that the initial syllable has greater duration than all subsequent syllables, in words of all syllable counts and in all phrasal contexts. The fact that this is the case in words that form individual APs (the main experiment), APs that are part of an ip (Dataset 1a-b), and words that do not form an AP, as in post-focal deaccenting (Dataset 2a-b), demonstrates that the durational effect is independent of phrasal prosodic phenomena. By itself, though, this does not yet mean that the durational effect is necessarily linked to word-level stress. Alternatively, greater duration of the initial syllable can also result from initial strengthening – a phonetic process that applies to left edges of prosodic domains and is independent of stress. However, the two phenomena have different phonetic signatures: initial strengthening affects the duration of the absolute initial segment but does not extend to the vowel in the initial syllable, while word stress can contribute to greater duration of the initial syllable on the whole (Fougeron & Keating 1997; Byrd, Krivokapić & Lee 2006; Barnes 2008). Therefore, the relative contribution of the consonant and vowel to the duration of the initial syllable was also investigated in Section 4.1. The results in Table 9 show that the greater duration of the initial syllable results from greater duration of both the consonant and the vowel, as compared to their counterparts in the second syllable. This would be unexpected on account of initial strengthening alone, which only targets the absolutely domain-initial (here, word-initial) segment. With initial strengthening as an alternative explanation ruled out, the most fitting interpretation for the durational effect on the initial syllable is that it marks word-level stress. Note that the current findings contrast with Selmer's (1935) results, who found that, in disyllables and trisyllables, the initial and final syllables were of equal duration. This discrepancy is likely due to the fact that Selmer's stimuli were not embedded in carrier phrases and were subject to phrase-final lengthening.

The initial syllable is also often (though not always) marked by greatest intensity, as compared to all other syllables, but there is also a more pronounced difference in intensity values between each syllable and the subsequent one. At the same time, intensity rises on the ultima in Datasets 1a and 1b – i.e., in those cases where the ultima carries an ip-final high H- tone. This suggests that intensity may also mark the location of word-level stress, but, unlike duration, intensity values are affected by and 'parasitic on' the phrasal prosodic context.

In terms of F0 properties, the results from the main experiment and Datasets 1a and 1b align: in both, individual words carry an F0 contour that steeply rises on the final syllable (and may include a fall on the initial syllable, if preceded by a Ha). This aligns with the existing analysis of Georgian prosody (Vicenik & Jun 2014), according to which individual APs in Georgian carry an L* Ha tonal contour. Likewise, ip-final APs in Datasets 1a and 1b carry an L* H- contour (similar in shape, but with the final boundary tone marking the right edge of the ip). Additionally, the data discussed here presents evidence for another low tonal target on the penultimate syllable, the turning point between the falling and rising parts of the contour. Note that the F0 results are comparable to Zhghenti's (1953; 1958) results, who found that F0 values drop on the penult of his stimuli. The fact that the ultima, in Zhghenti's results, did not have high F0 values, unlike in the current study, is likely due to the fact that his stimuli were produced in isolation and carried a low phrase-final tone. Finally, the words in Datasets 2a and 2b presented no evidence for tonal targets. This is expected in the context of post-focal deaccenting and, most importantly, shows that F0 targets in Georgian are phrasal/post-lexical in nature.

The results from the main experiment provide evidence that the penultimate syllable in Georgian is reserved for F0 targets that are part of the right-edge intonational make-up of a phrase, which is independent from stress (cf. Gordon 2014). The next question is what type of an F0 target it is. The inventory of F0 targets available in Autosegmental-Metrical theory includes pitch accents, phrase accents, and boundary tones (Liberman 1975; Bruce 1977; Pierrehumbert 1980; Ladd 1983; Gussenhoven 1984; Pierrehumbert &

Beckman 1988, a.o.). Pitch accents are anchored to syllables that carry a degree of stress, while boundary tones are aligned with edges of prosodic domains. It is unlikely that the F0 target on the penults is a pitch accent: pitch accents align with stressed syllables, but there is no evidence from the distribution of duration and intensity that penults in Georgian carry stress. It is equally unlikely that the F0 target on the penults is part of a complex final boundary tone, 'crowded out' to the penultimate syllable. This is because final syllables in Georgian can accommodate bitonal boundary tones (e.g., HL%) without placing one of the tones on the penult (Vicenik & Jun 2014; Borise 2017).

The remaining prosodic category is a phrase accent. As discussed in Section 2.3, penultimate syllables in Georgian, in certain informational-structural conditions (questions and narrow focus contexts), carry a low phrase accent L. I suggest that the resemblance between phrase accents, described in the literature, and the low F0 target found on penultimate syllables in the current study is not accidental. Given their identical distribution, the most parsimonious approach to the two F0 targets is to treat them as two subtypes of the same phenomenon. This, in turn, means that the distribution of the phrase accent in Georgian is broader than has been described before and is not limited to predicates in questions and narrow focus contexts. The exact information-structural conditions that may influence its distribution merit further investigation.

If viewed against the backdrop of the existing literature, which allowed us to hypothesize that Georgian stress may be initial, penultimate, or antepenultimate, the current results demonstrate that Georgian has initial word-level stress, while the penult houses a phrasal tonal target. Notably, the results provide no evidence for antepenultimate stress. This is consistent with Alkahzishvili (1959: 402), who, similarly, found no acoustic evidence for antepenultimate stress.

These conclusions have implications for the theory of word stress. Recall that stress in Georgian is morphophonologically 'inert': speakers do not have consistent intuitions about stress placement, and stress placement is not subject to regular variation in declensional/conjugational paradigms. The fixed nature of Georgian stress, as advocated here, may be a contributing factor to this 'inertia'. Languages with fixed stress are known to have a weaker acoustic expression of stress than languages with variable/contrastive stress placement (Cutler 2005; Dogil 1999; Fónagy 1966; Janota 1967; Jassem 1962; Rigault 1970). Speakers of languages with fixed stress have weaker intuitions about stress placement (including in other languages) – the so called 'stress-deafness' (Dupoux, Peperkamp & Sebastián-Gallés 2001; Peperkamp & Dupoux 2002; Peperkamp, Vendelin & Dupoux 2010, a.o.). The fact that Georgian falls into the category of languages with fixed stress, therefore, may explain the reported 'weakness' of Georgian stress, and lack of consistent intuitions on behalf of the speakers. More broadly, these results suggest that other languages described as having 'weak' stress may, in fact have fixed, morphophonologically 'inert' but still identifiable stress (instead of not having the category of stress at all).

7 Conclusion

Based on instrumental evidence, this paper identified and provided interpretation for some of the acoustic cues that are regularly utilized in Georgian prosody: duration, intensity, and alignment of F0 targets. First, it established that initial syllables in Georgian words are marked by greater duration than subsequent syllables. This is the case in words two to five syllables long, when they form independent APs, APs that are embedded in an ip, or are subject to deaccenting/dephrasing in the post-focal portion of the utterance. After excluding initial strengthening as an alternative explanation, this effect is best phonologically interpreted as cuing stress, fixed on the initial syllable. This result aligns well with the existing literature on Georgian prosody, which consistently lists the initial syllable as the locus of prosodic prominence and a possible stress locus. The intensity results support the evidence form duration, in that initial syllables both in independent APs and in deaccented words often have greater intensity than all subsequent syllables – though the intensity values also drop throughout the word, which questions the link between intensity and stress. Also, in contrast with duration, the intensity values are sensitive to phrasal prosodic context: e.g., ip-final syllables have consistently higher intensity. This shows that, while also acting as a cue for word stress, intensity-marking interacts with prosodic phrasing. With respect to F0 effects, the results confirmed the

presence of a low pitch accent on initial syllables and a high final boundary tone at the right edges of APs, as described in the literature. They also showed that penultimate syllables carry a low intonational F0 target. Based on what is known about the phrasal prosody of Georgian, this F0 target is best analyzed as a phrase accent. Finally, the results showed that post-focal deaccenting eliminates F0 targets (but not durational or intensity effects). This demonstrates that F0-marking in Georgian is reserved for phrasal prosody and is not intrinsic to stress-marking. Overall, the findings contribute to distinguishing word- and phrase-level prosodic phenomena in Georgian and resolving the conflicting views that surround word stress placement in the language.

Abbreviations

In addition to abbreviations from the Leipzig Glossing Rules, the following are used in this article for language examples: AOR aorist, AUG augment, PR pre-root vowel, PV preverb, SF stem formant, SM stem marker, TS thematic suffix, VER version vowel. Abbreviations in the body text include AP Accentual Phrase, ip Intermediate Phrase, IP Intonational Phrase, AM Autosegmental-Metrical, DOI Digital Oject Identifier.

References

Akhvlediani, Georgij S. 1949. *Zogadi ponet'ik'is sapudzvlebi*. Tbilisi: Tbilisi State University Publishing.

Alkhazishvili, Archil. 1959. Porjadok slov i intonacija v prostom povestvovateljnom predloženii gruzinskogo jazyka [Word order and intonation in simple declarative clauses in Georgian]. *Fonetika* 1. 367–414.

Aronson, Howard. 1990. *Georgian: a reading grammar*. Bloomington, IN: Slavica Publishers Inc.

Asatiani, Rusudan & Stavros Skopeteas. 2012. Information structure in Georgian. In Manfred Krifka & Renate Musan (eds.), *The Expression of Information structure*, vol. 5, 127–158. Berlin ; Boston: De Gruyter. http://dx.doi.org/10.1515/9783110261608.127.

Barnes, Jonathan. 2008. *Strength and weakness at the interface: Positional neutralization in phonetics and phonology*. Berlin ; New York: Mouton de Gruyter. http://dx.doi.org/10.1515/9783110197617.

Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1). 1–48. https://doi.org/doi:10.18637/jss.v067.i01.

Beckman, Mary E. 1986. *Stress and non-stress accent* (Netherlands Phonetic Archives 7). Dordrecht: Foris Publ.

Berg, Helma van den. 1995. *A grammar of Hunzib*. München-Newcastle: Lincom Europa.

Boersma, Paul & David Weenink. 2021. *Praat: doing phonetics by computer [Computer program]*. http://www.praat.org/. (25 March, 2021).

Bokarev, Evgenij A. 1967. Gunzibskij jazyk [The Hunzib language]. In Evgenij A. Bokarev & Ketevan V. Lomtatidze (eds.), *Jazyki narodov SSSR. Iberijsko-kavkazskie jazyki*, vol. 4, 472–487. Moscow: USSR Academy of Sciences Publishing.

Borise, Lena. 2017. Prosody of Focus in a Language with a Fixed Focus Position: Evidence from Georgian. *Proceedings of WCCFL* 34. 89–96.

Borise, Lena. 2019. *Phrasing is Key: The Syntax and Prosody of Focus in Georgian*. Harvard University PhD Dissertation.

Borise, Lena. 2020. Word Stress in the Languages of the Caucasus. In Maria Polinsky (ed.), *Handbook of the Languages of the Caucasus*, 757–782. Oxford; New York: Oxford University Press.

Borise, Lena & Xavier Zientarski. 2018. Word stress and phrase accent in Georgian. *Proceedings of TAL 6* 207–211. http://dx.doi.org/10.21437/TAL.2018-42.

Bruce, Gösta. 1977. *Swedish word accents in sentence perspective*. Lund University PhD Dissertation.

Bush, Ryan. 1999. Georgian Yes-No Question Intonation. *Phonology at Santa Cruz* 6. 1–11.

Byrd, Dani, Jelena Krivokapić & Sungbok Lee. 2006. How far, how long: On the temporal scope of prosodic boundary effects. *The Journal of the Acoustical Society of America* 120(3). 1589–1599.

Chikobava, Arnold. 1942. Maxvilis sak'itxisatvis dzvel kartulši (ts'inasts'ari moxseneba). II. [On the question of stress in Old Georgian (a preliminary report). II.]. *Sakartvelos SSR mecnierebata ak'ademiis moambe* 3(3). 296–303.

Cutler, Anne. 2005. Lexical stress. In David B. Pisoni & Robert E. Remez (eds.), *The Handbook of Speech Perception*, 264–289. Blackwell.

De Jong, Kenneth & Bushra Adnan Zawaydeh. 1999. Stress, duration, and intonation in Arabic word-level prosody. *Journal of Phonetics*. Elsevier 27(1). 3–22.

Del Giudice, Alex, Ryan Shosted, Kathryn Davidson, Mohammad Salihie & Amalia Arvaniti. 2007. Comparing methods for locating pitch "elbows." In *Proceedings of the XVI th International Congress of Phonetic Sciences*, 1117–1120.

Dilley, L., S. Shattuck-Hufnagel & M. Ostendorf. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24(4). 423–444. https://doi.org/10.1006/jpho.1996.0023.

Dirr, Adolf. 1904. *Grammatik der modernen georgischen (grusinischen) Sprache.* Vienna: Hartleben.

Dogil, Grzegorz. 1999. Baltic languages. In, 385–392. Berlin: Mouton de Gruyter. http://journals.cambridge.org/abstract_S1470542702000223. (3 April, 2016).

Dupoux, Emmanuel, Sharon Peperkamp & Núria Sebastián-Gallés. 2001. A robust method to study stress "deafness." *The Journal of the Acoustical Society of America* 110(3). 1606–1618.

Edwards, Jan, Mary E. Beckman & Janet Fletcher. 1991. The articulatory kinematics of final lengthening. *the Journal of the Acoustical Society of America*. Acoustical Society of America 89(1). 369–382.

Elvira-García, Wendy. 2014. *Prosodic data extraction*. http://stel.ub.edu/labfon/sites/default/files/prosodic_data-extraction-2.0.praat.

Eriksson, Anders & Mattias Heldner. 2015. The acoustics of word stress in English as a function of stress level and speaking style. In *16th Annual Conference of the International Speech Communication Association (INTERSPEECH 2015), Dresden, Germany, September 6-10, 2015*, 41–45.

Everett, Keren. 1998. The acoustic correlates of stress in Pirahã. *Journal of Amazonian Languages* 1(2). 104–162.

Flemming, Edward Stanton. 2018. Modeling 'elbows' in F0 contours: Phrase accents in English. In.

Fletcher, Janet. 2010. The prosody of speech: Timing and rhythm. In William J. Hardcastle, John Laver & Fiona E. Gibbon (eds.), *The Handbook of Phonetic Sciences*, 521–602. 2nd edition. Wiley-Blackwell. https://doi.org/10.1002/9781444317251.ch15.

Fónagy, Ivan. 1966. Electrophysiological and acoustic correlates of stress and stress perception. *Journal of speech and hearing research* 9(2). 231–244.

Fougeron, Cécile & Patricia A. Keating. 1997. Articulatory strengthening at edges of prosodic domains. *The Journal of the Acoustical Society of America* 101(6). 3728–3740. http://dx.doi.org/10.1121/1.418332.

Gachechiladze, Givi R. 1968. K voprosu ob udarenii v gruzinskom stixe v sopostavlenii s anglijskim [On the question of stress in Georgian verse as compared to English]. In Jiří Levý & Karel Palas (eds.), *Teorie verše. II, Sborník druhé brněnské versologické konference, 18-20 October 1966*, 89–112. Brno: Universita J.E. Purkyně.

Gamzatov, Magomed-Rasul E. 1975. *Fonetika gunzibskogo jazyka [Phonetics of the Hunzib language].* Moscow: Georgian SSR Academy of Sciences, Institute of Linguistics.

Garellek, Marc & James White. 2015. Phonetics of Tongan stress. *Journal of the International Phonetic Association*. Cambridge University Press 45(1). 13–34.

Gordon, Matthew. 1995. Acoustic properties of primary and secondary word-level stress in Estonian. In *Poster presented at the 130th meeting of the Acoustical Society of America, St. Louis.*

Gordon, Matthew. 2003. The phonology of pitch accent placement in Chickasaw. *Phonology* 20. 173–218.

Gordon, Matthew. 2004. A phonological and phonetic study of word-level stress in Chickasaw. *International Journal of American Linguistics* 70. 1–32.

Gordon, Matthew. 2005. An autosegmental/metrical model of Chickasaw intonation. In Sun-Ah Jun (ed.), *Prosodic typology I: the phonology of intonation and phrasing*, 301–330. Oxford: Oxford University Press.

Gordon, Matthew. 2014. Disentangling stress and pitch-accent: a typology of prominence at different prosodic levels. In Harry G. Van der Hulst (ed.), *Word stress: theoretical and typological issues*, 83–118. Cambridge: Cambridge University Press.

Gordon, Matthew & Timo Roettger. 2017. Acoustic correlates of word stress: A cross-linguistic survey. *Linguistics Vanguard* 3(1). https://doi.org/10.1515/lingvan-2017-0007.

Gorgadze, Sergi. 1912. *Kartuli ts'q'obilsit'q'vaoba [Georgian poetic writing]*. Tbilisi: n.p.

Grice, Martine, D. Robert Ladd & Amalia Arvaniti. 2000. On the place of phrase accents in intonational phonology. *Phonology* 17(02). 143–185. http://dx.doi.org/10.1017/S0952675700003924.

Gudava, Togo. 1969. Maxvilis adgilisatvis megrulši [Stress placement in Megrelian]. In *Giorgi Akhvlediani Festschrift*, 106–111. Tbilisi: Tbilisi University Press.

Gussenhoven, Carlos. 1984. *On the grammar and semantics of sentence accents*. Dordrecht: Foris. http://dx.doi.org/10.1515/9783110859263.

Harnsberger, James D. & Jasmeet Judge. 1996. Pitch range and focus in Hindi. *The Journal of the Acoustical Society of America* 99(4). 2493–2500. https://doi.org/10.1121/1.415628.

Hewitt, Brian George. 1995. *Georgian: A structural reference grammar*. Vol. 2. Amsterdam; Philadelphia: John Benjamins Publishing. http://dx.doi.org/10.1075/loall.2.

Hock, Hans Henrich. 1988. Initial strengthening. *Phonologica* 101–110.

Huh, Woong. 1985. *Kwuke Umunhak [Korean Phonology]*. Second edition. Seoul: Saym Munhwasa.

Ioseliani, Platon. 1840. *P'iruel-dats'q'ebitni k'anonni kartulis ghramat'ik'isa*. Tbilisi: I. da D. Arzanovta St'amba.

Isakov, Isak A. & Madzhid Sh. Khalilov. 2012. *Gunzibskij jazyk (Fonetika. Morfologija. Slovoobrazovanie. Leksika. Teksty) [The Hunzib language (Phonetics. Morphology. Word building. Vocabulary. Texts)]*. Makhachkala: Gamzat Tsadasa Institute for Language, Literature and Arts of the Dagestan Scientific Centre of the Russian Academy of Sciences.

Janashvili, Mose. 1906. *Kartuli gramat'ik'a*. Tbilisi: Elekt'rombečdavi amx. "Šroma."

Janota, Přemysl. 1967. An experiment concerning the perception of stress by Czech listeners. *Acta Universitatis Carolinae-Philologica, Phonetica Pragensia I* 45–68.

Jassem, Wiktor. 1962. *Akcent języka polskiego* (Prace Jezykoznawcze). Vol. 31. Wroclaw: Zaklad Narodowy im. Ossolińskich.

Jun, Sun-Ah. 1993. *The Phonetics & Phonology of Korean Prosody*. OSU PhD Dissertation.

Jun, Sun-Ah. 1995. A phonetic study of stress in Korean. *The Journal of the Acoustical Society of America* 98. 2893–2893. http://dx.doi.org/10.1121/1.414317.

Jun, Sun-Ah. 1998. The accentual phrase in the Korean prosodic hierarchy. *Phonology* 15. 89–226.

Jun, Sun-Ah, Chad Vicenik & Ingvar Lofstedt. 2007. Intonational Phonology of Georgian. *UCLA Working Papers in Linguistics* 106. 41–57.

Kibrik, Alexander E. & Sandro V. Kodzasov. 1990. *Sopostavitel'noje izučenije dagestanskix jazykov: Imia. Fonetika [A comparative study of Dagestanian languages: Nominal categories. Phonetics]*. Moscow: Moscow University Publishing.

Ladd, D. Robert. 1980. *The structure of intonational meaning*. Bloomington, IN: Indiana University Press.

Ladd, D. Robert. 1983. Phonological features of intonational peaks. *Language* 59(4). 721–759. http://dx.doi.org/10.2307/413371.

Ladd, D. Robert. 1996. *Intonational phonology*. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511808814.

Lee, Ho-Young. 1990. *The structure of Korean prosody*. University College London PhD Dissertation.

Lee, Hyun-Bok. 1973. A Phonetic Study of the Accent in Korean. *Mullidaehakpo (Seoul National University)* 19. 1–16.

Lehiste, Ilse. 1966. *Consonant quantity and phonological units in Estonian*. Bloomington, IN: Indiana University Press.

Lehiste, Ilse. 1972. The timing of utterances and linguistic boundaries. *The Journal of the Acoustical Society of America* 51(6B). 2018–2024. http://dx.doi.org/10.1121/1.1913062.

Liberman, Mark Y. 1975. *The intonational system of English*. MIT PhD Dissertation.

Machač, Pavel & Radek Skarnitzl. 2009. *Principles of phonetic segmentation*. Praha: Epocha.

Marr, Nikolaj J. 1925. *Grammatika drevneliteraturnogo gruzinskogo jazyka [A grammar of the Old Georgian literary language]*. Leningrad: Russian Academy of Sciences Publishing.

Michael, Lev. 2010. The Interaction of Tone and Stress in the Prosodic System of Iquito (Zaparoan). *UC Berkeley PhonLab Annual Report, 6(6)* 57–79.

Müller, Gabriele. 2005. *Frageintonation im Georgischen*. University of Cologne MA thesis.

Peperkamp, Sharon & Emmanuel Dupoux. 2002. A typological study of stress "deafness." *Laboratory phonology* 7. 203–40.

Peperkamp, Sharon, Inga Vendelin & Emmanuel Dupoux. 2010. Perception of predictable stress: A cross-linguistic investigation. *Journal of Phonetics* 38(3). 422–430.

Pierrehumbert, Janet. 1980. *The phonetics and phonology of English intonation*. MIT PhD Dissertation.

Pierrehumbert, Janet B. & Mary Beckman. 1988. *Japanese tone structure*. Cambridge, MA: MIT Press.

R Core Team. 2020. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/.

Rahmani, Hamed, Toni Rietveld & Carlos Gussenhoven. 2018. Post-focal and factive deaccentuation in Persian. *Glossa: a journal of general linguistics* 3(1). 13. https://doi.org/10.5334/gjgl.328.

Reichel, Uwe D. & Nele Salveste. 2015. Pitch elbow detection. In Günther Wirsching (ed.), *ESSV, Katholische Universität Eichstätt-Ingolstadt, Germany, 25.–27. March 2015.*, 143−149. Dresden: TUD Press.

Remijsen, Bert & Vincent J. Van Heuven. 2005. Stress, tone and discourse prominence in the Curaçao dialect of Papiamentu. *Phonology*. Cambridge University Press 22(2). 205–235.

Rigault, André. 1970. L'accent dans deux langues à accent fixee: le français et le tchèque. *Studia Phonetica* 3. 1–12.

Robins, Robert H. & Natalie Waterson. 1952. Notes on the phonetics of the Georgian word. *Bulletin of the School of Oriental and African studies* 14(01). 55–72. http://dx.doi.org/10.1017/S0041977X00084196.

Rudenko, Boris T. 1940. *Grammatika gruzinskogo jazyka [A grammar of the Georgian language]*. Moscow, Leningrad: USSR Academy of Sciences Publishing.

Selmer, Ernst W. 1935. Georgische Experimentalstudien. In *Avhandlinger utgitt av det Norske Videnskaps-Akademi i Oslo. II. Historisk-filosofisk klasse*, 1–55. Oslo: I Kommisjon Hos Jacob Dybwad.

Shosted, Ryan K. & Vakhtang Chikovani. 2006. Standard Georgian. *Journal of the International Phonetic Association* 36(2). 255–264. http://dx.doi.org/10.1017/S0025100306002659.

Skopeteas, Stavros & Gisbert Fanselow. 2010. Focus in Georgian and the expression of contrast. *Lingua* 120(6). 1370–1391. https://doi.org/10.1016/j.lingua.2008.10.012.

Skopeteas, Stavros & Caroline Féry. 2010. Effect of narrow focus on tonal realization in Georgian. *Proceedings of Speech Prosody 5*.

Skopeteas, Stavros & Caroline Féry. 2014. Prosodic cues for exhaustive interpretations: a production study of Georgian intonation. In Nino Amiridze, Tamar Reseck & Manana Topadze Gäumann (eds.), *Advances in Kartvelian Morphology and Syntax*, 79–100. Bochum: Universitätsverlag Brockmeyer.

Skopeteas, Stavros & Caroline Féry. 2016. Focus and intonation in Georgian: Constituent structure and prosodic realization. Potsdam University, ms.

Skopeteas, Stavros, Caroline Féry & Rusudan Asatiani. 2009. Word order and intonation in Georgian. *Lingua* 119(1). 102–127. http://dx.doi.org/10.1016/j.lingua.2008.09.001.

Skopeteas, Stavros, Caroline Féry & Rusudan Asatiani. 2018. Prosodic separation of postverbal material in Georgian. In Evangelia Adamou, Katharina Haude & Martine Vanhove (eds.), *Information structure in lesser-described languages: Studies in prosody and syntax*, vol. 199, 17–50. Amsterdam ; Philadelphia: John Benjamins.

Sluijter, Agaath M. C. & Vincent J. van Heuven. 1996. Spectral balance as an acoustic correlate of linguistic stress. *The Journal of the Acoustical Society of America*. Acoustical Society of America 100(4). 2471–2485. https://doi.org/10.1121/1.417955.

Tevdoradze, Izabela. 1978. *Kartuli enis p'rosodiis sak'itxebi [Prosodic issues in the Georgian language]*. Tbilisi: Tbilisi State University Publishing.

Tevdoradze, Izabela. 2005. K'itxviti c'inadadebebis melodik'uri surati kartul salit'erat'uro enaši [The melodic makeup of interrogative sentences in literary Georgian language]. *T'ip'ologiuri dziebani* 5. 85–89.

Tschenkeli, K. 1958. *Einführung in die georgische Sprache*. Zürich: Amirani Verlag.

Vicenik, Chad & Sun-Ah Jun. 2014. An Autosegmental-Metrical analysis of Georgian intonation. In Sun-Ah Jun (ed.), *Prosodic Typology II*, 154–186. Oxford: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199567300.003.0006.

Vogel, Irene, Angeliki Athanasopoulou & Nadya Pincus. 2016. Prominence, contrast, and the functional load hypothesis: An acoustic investigation. In Jeffrey Heinz, Rob Goedemans & Harry G. van der Hulst (eds.), *Dimensions of phonological stress*, 123–167. Cambridge: Cambridge University Press.

Vogt, Hans. 1936. *Esquisse d'une grammaire du géorgien moderne* (Norsk Tidsskrift for Sprogvidenskap 9-10). Kristiania: A.W. Brøggers Boktrykkeri.

Vogt, Hans. 1971. *Grammaire de la langue géorgienne*. Oslo: Universitetsvorlaget.

Williams, Briony. 1999. The phonetic manifestation of stress in Welsh. In Harry G. van der Hulst (ed.), *Word Prosodic Systems in the Languages of Europe*, 311–334. Berlin: Mouton de Gruyter.

Zhghenti, Sergi. 1953. Kartuli salit'erat'uro enis akcent'uaciis dziritadi sak'itxebi [Main questions of accentuation in Georgian literary language]. *Iberiul-k'avk'asiuri enatmecniereba* 5. 125–163.

Zhghenti, Sergi. 1958. Kartuli dialekt'ebis c'armotkmis taviseburebata šesc'avlis zogierti sak'itxi [Some questions about the study of pronunciational properties of Georgian dialects]. *St'alinis saxelobis Tbilisis Saxelmts'ipo Universit'et'is šromebi* 69. 253–264.

Zhghenti, Sergi. 1959. Eksperimental'no-fonetičeskoje izučenije akcentuacii gruzinskogo jazyka [An experimental phonetic study of accentuation in the Georgian language]. *Fonetika [Fonetičeskij sbornik]* I. 69–108.

Zhghenti, Sergi. 1963. *Kartuli enis rit'mik'ul-melodik'uri st'rukt'ura [The rhythmic-melodic structure of Georgian language]*. Tbilisi: Tsodna.

Lena Borise
Hungarian Research Center for Linguistics
Benczúr utca 33
Budapest 1068 Hungary
lena.borise@nytud.hu