The prosodic reality of focus projection: evidence from word stress and F0 patterns in Georgian

Lena Borise

Abstract
Based on novel experimental evidence, this paper shows that focus projection/percolation (Contreras 1976; Culicover & Rochemont 1983; Selkirk 1984) – the phenomenon by way of which prosodic prominence on a sub-constituent signals focus on the whole constituent – has a consistent prosodic realization in Georgian. Specifically, some focus projection accounts that are explicitly grounded in syntactic structure (Cinque 1993; Zubizarreta 1998, a. o.) predict that narrow focus on the direct object may be realized in the same way as VP-/broad focus, since in all three scenarios prosodic prominence is realized on the most deeply embedded constituent, the direct object, and projected up the syntactic tree. Accordingly, narrow focus on the subject does not lead to focus projection, given that the subject in such contexts attracts prosodic prominence away from the most deeply embedded constituent. This paper shows that the distribution of focus-induced prominence in Georgian fits with this generalization. Specifically, narrowly focused subjects differ from subjects in broad focus utterances in their F0 patterns and prominence of the stressed syllable, while the prosody of utterances with narrowly focused objects does not differ from VP-/broad focus contexts. The novelty of the current study lies in that narrow foci of both types surface in the immediately preverbal position in Georgian (though it is not in the same structural position; Borise 2019). Nevertheless, the two types of foci have systematically different prosodic realizations. Accordingly, the results presented here show that prosodic structure reflects differences in syntactic structure even when they are not reflected linearly.

Keywords: focus projection/percolation, word stress, syllable duration, F0, prosodic phrasing, Georgian.

***

1 Introduction
This paper reports on an experimental investigation of the prosodic properties of narrow foci in Georgian, a Kartvelian language spoken in the Caucasus. A verb-final language, Georgian is typologically similar to other such languages, in that it has a preverbal focus position, which houses narrow foci (as well as wh-phrases). Syntactically, it has been shown that preverbal foci in Georgian are, in fact, found in situ, and the adjacency with the verb is achieved via ‘altruistic’ movement of any intervening material (Borise 2019). This means that subject and object focus contexts rely on different syntactic structures: while object foci are structural sisters to the verb, subject foci are found higher up in the syntactic tree, and may require ‘altruistic’ movement of any intervening material in order to achieve adjacency with the verb (e.g., movement of the direct object in a transitive clause).

Prosodically, as the novel experimental evidence presented here shows, there are important differences between the realizations of preverbal narrowly focused subjects and direct objects. Specifically, while the prosody of preverbal narrowly focused objects aligns with that of preverbal objects in broad focus conditions, the prosody of narrowly focused subjects is markedly different from that of subjects found in broad-focus utterances. Two distinct prosodic phenomena investigated here – prominence of the stressed syllable and the F0 contour on the constituent in question – allow to reliably track the prosodic difference
between the two argument types. The contrast at hand provides support to focus projection accounts, which predict that the prosodic realization of narrow focus on a direct object (i.e., the most deeply embedded constituent) but not another constituent, such as a subject, may pattern together with the prosodic realization of broad focus (Contreras 1976; Culicover & Rochemont 1983; Selkirk 1984; von Stechow & Uhmann 1986; Reinhart 1995; Cinque 1993; Zubizarreta 1998).

The most important contribution of this paper lies in the fact that all types of foci in Georgian, including subjects and objects, surface in the immediately preverbal position. Nevertheless, narrowly focused preverbal subjects and objects have completely different prosodic realizations. This means that prosodic structure reflects differences in syntactic structure even when they are not reflected in the linear word order.

This paper is structured in the following way. Section 2 provides necessary background for the experimental results discussed in the remainder of the paper. More specifically, Section 2.1 introduces the phenomenon of focus projection and existing theoretical approaches to it, Section 2.2 discusses the syntax of focus in Georgian, and Section 2.3 provides an overview of the existing work on the prosody of information structure in Georgian. Next, Section 3 reports on the results of a production study which was aimed at investigating the prosody of different types of focus in Georgian. First, Section 3.1 introduces the experimental design, and Section 3.2 lays out the generalizations concerning word orders used in different focus contexts. Then, Section 3.3 reports on the stress properties of narrowly focused constituents, which are reflected in syllable duration patterns, and Section 3.4 turns to the F0 contours that accompany focused constituents, as compared to the same constituents found in broad focus conditions. Finally, Section 4 provides a summary of the current results and discusses their theoretical implications.

2 Theoretical background and previous work

2.1 Existing approaches to focus projection

The idea of focus projection (percolation), which also may be thought of as focus ambiguity between a constituent and a sub-constituent, goes back to Chomsky (1972), who observed that a sentence such as (1) is ambiguous in terms of its focus structure: any of the nested bracketed constituents in (1) can be regarded as carrying focus:

(1) He was (warned (to look out for (an ex-convict (with (a red (SHIRT)))))).

The same effect can be illustrated with another example: (2) can be felicitously uttered in response to any of the questions in (3), each of which selects for a focused constituent of different size. One of the main tenets of focus projection accounts is that, in terms of their prosodic properties, the replies to the questions in (3) would be identical or nearly identical. The presence of focus in replies to such questions is manifested by prosodic prominence on the word SOCKS, which serves as the locus of such prominence in the focused constituent, regardless of its size. This means that the prosodic pattern with prominence on the word SOCKS

---

1 Note that this applies to narrow foci found in the preverbal domain – i.e. narrow foci found in the part of the utterance to the left of the finite verb must be left-adjacent to it. Narrow foci can also be found immediately postverbally in Georgian. There are no major semantic or pragmatic differences between preverbal and postverbal foci, but postverbal focus differs from preverbal both in its syntactic structure and prosodic realization. This type of focus is not discussed here. See Skopeteas and Fanselow (2010) for the discourse properties of postverbal focus, Skopeteas and Féry (2010) for its prosodic realization, and Borise (2019) for the syntactic representation of postverbal focus.
in (2) may be interpreted as characterizing broad focus, VP focus, or narrow focus on the direct object (or a subpart of the direct object).

(2) Nini (knit (a pair of (SOCKS))).

(3) a. What did Nini do?
   b. What did Nini knit?
   c. A pair of what did Nini knit?

The notion of focus projection was further developed in much subsequent work (Contreras 1976; Culicover & Rochemont 1983; Selkirk 1984; von Stechow & Uhmann 1986; Reinhart 1995; Cinque 1993; Zubizarreta 1998). Building on the original insight from Chomsky (1972), the main idea that underlies these approaches is that if a sub-constituent of a complex phrase is prosodically marked as focused, the whole phrase may be treated as focus for the purpose of focus interpretation.2

The link between focus projection and the neutral declarative prosodic pattern, typical of broad focus contexts, has been subject to various formalizations. Some of the earliest ones are based on the Nuclear Stress Rule (NSR; Chomsky & Halle 1968), according to which the strongest stress in an utterance (in English or other VO languages) is located on the rightmost constituent. In turn, according to the first focus projection accounts grounded in syntax (Contreras 1976; Culicover & Rochemont 1983; Selkirk 1984, a.o.), the fact that nuclear stress is found on the rightmost constituent is derived not from its utterance-final status but from rules that calculate the relative prominence of adjacent nodes in the syntactic structure. This idea was later brought to fruition in Cinque (1993) and Zubizarreta (1998), who suggested that it is the most deeply embedded constituent (i.e., in transitive clauses, the direct object) that carries sentence-level prosodic prominence; crucially, this formulation of the nuclear stress rule allowed to account for the distribution of nuclear stress in both VO and OV languages.

According to the syntax-based approaches introduced above, focus projection has the following pattern: narrow focus on the direct object may be realized in the same way as focus on the VP and broad focus, given that in all three scenarios prosodic prominence is realized on the most deeply embedded constituent, the direct object, and projected up the syntactic three through a succession of embedding phrases. The size of the constituent carrying narrow focus thus ranges from the direct object alone to the projection that embeds it (VP), to the next embedding projection, such as a vP or TP, and so on, up to broad focus on the whole utterance. Crucially, constituents other than the direct object and the projections on the clausal spine that embed it – for instance, the subject of the clause – do not take part in the phenomenon of focus projection. This means that narrow focus on the subject is incompatible with the prosodic pattern typical of broad focus. This is illustrated in (4) and (5): of the three questions in (5), only (d) can be felicitously answered with the utterance in (4).

(4) NINI (knit (a pair of (socks))).

(5) a. # What did Nini do?
   b. # What did Nini knit?
   c. # A pair of what did Nini knit?

2 There is no agreement in the literature as to whether the focused constituent has to have the status of a syntactic head in order to be able to project; cf. Selkirk’s (1984) seminal proposal for in favor of this hypothesis, and Büring’s (2006) ‘Unrestricted Vertical Focus Projection’ theory for the opposite view.
d. Who knit a pair of socks?

The focus projection accounts, then, make a clear prediction: phonetically, the prosodic realization of narrow focus on the direct object may match that of the prosodic realization of focus on the higher embedding projections, such as VP, TP, and the full clause. This prediction has been subject to empirical scrutiny, in a number of languages, but the available results vary. In particular, a number of studies on the prosodic realization of focus in English, such as Gussenhoven (1983), Birch and Clifton Jr. (1995), and Welby (2003), found no differences between the prosody of narrow focus on the direct object and broad focus, as focus projection accounts predict. In the perception experiments reported by the authors, listeners did not always reliably distinguish replies to questions eliciting broad focus or VP focus from narrow focus on the direct object. On the other hand, Rump and Collier (1996), and Bishop (2010) found that listeners can reliably tell apart narrow focus on the direct object from broad focus, and Breen et al. (2010) reached the same conclusion based on the phonetic cues used by speakers in a production experiment. One of the main conclusions of the existing studies is that, while broad focus utterances and those with narrow focus on the direct object align with respect to the presence and type of a pitch accent on the direct object, the difference between the two contexts may lie in other categorical measures, such as types of accents found on prenuclear constituents, as well as gradient measures such as duration, intensity and relative F0 values of the constituents in an utterance.

In German, Baumann et al. (2006; 2007) found that speakers produce narrow focus on the direct object differently from broad focus, with duration and patterns of upstep and downstep differentiating the two contexts. In contrast, it was shown for Greek (Gryllia 2009) and Hindi (Patil et al. 2008) that narrow focus on the direct object and broad focus, in terms of their F0 contours, are not reliably different from each other. Note that Greek is a VO language, and Hindi is an OV one, which means that the relative linear positions of the object and the verb indeed do not affect the workings of focus projection, as was proposed by Cinque (1993).

This paper contributes to the ongoing debate about the prosodic reality of focus projection by bringing in evidence form Georgian. In contrast with some of the existing studies, it compares not only the prosodic properties of narrow focus on the direct object and broad focus, but also narrow focus on the subject. The results reported here show that while the prosody of narrow focus on the direct object aligns with that of broad focus, focus on the subject has a markedly different prosodic realization, providing further support for the focus percolation accounts. The contrast between the prosodic realization of the two constituents is especially striking given that Georgian is a language in which the immediately preverbal slot houses all focus constituents (i.e., in the current experiment, both subjects and objects).

### 2.2 The syntax of focus in Georgian

Georgian allows for considerable flexibility of word order, including both OV and VO orders in broad focus declaratives. While most contexts would allow for both OV or VO, there are some register differences: OV is preferred in conservative registers/written styles, while in colloquial registers OV and VO are equally common (Vogt 1971; Skopeteas & Fanselow 2010). This variability has led to some controversy with respect to headedness at the level of VP: most authors take OV to be neutral/underlying (Pochkhua 1962; Aronson 1982; Nash 1995; McGinnis 1997a; 1997b; Harris 2000; Boeder 2005), though some take transitive verbal projections to be head-initial (Tschentkeli 1958; Amiridze 2006) or the order of the verb and object to be unspecified (Harris 1981; Anderson 1984). I am adopting the view that Georgian is underlyingly OV, given the existence of contexts that only allow for OV, such as nominalizations (Borise 2019), contexts that
strongly favor OV, such as native idiomatic expressions (Skopeteas & Fanselow 2010), and the fact that a number of other projections and constructions, such as PPs, possessive constructions and auxiliary constructions, are head-final.

Like many other verb-final languages (cf. Basque, Turkish, Hindi, Persian), Georgian requires for narrow foci (if found in the preverbal domain) and wh-phrases to appear in the immediately preverbal position, as shown in (6). This holds for all types of narrow foci in Georgian, including foci found in replies to wh-questions, contrastive foci found in corrective replies, and constituents modified by focus-inducing particles only and even.

   ‘What did grandma clean yesterday morning?’

   A’: *Gušin dilas ras bebia alagebda?

   B: Gušin dila-s bebia samzareulo-s a-lag-eb-d-a.
   yesterday morning-DAT grandma.NOM kitchen-DAT VER-clean-SF-SM-IPFV.3SG
   ‘Grandma cleaned THE KITCHEN yesterday morning.’

   B’: *Gušin dilas samzareulos bebia alagebda.

Syntactically, it has been shown that preverbal placement of wh-phrases/narrow foci, in languages that have a preverbal focus position, may be derived in more than one way. For instance, it has been suggested that preverbal focus placement in Kashmiri (Bhatt 1999), Persian (Karimi 2008), and Malayalam (Jayaseelan 2001) is derived via a Spec-Head configuration. That is, the focused constituent and the verb are linearly adjacent to each other by virtue of the fact that the focused constituent undergoes A-bar movement to the specifier of a dedicated projection, such as FocP, and the verb is attracted to Foc0, as shown in (7).

(7) \[ \text{FocP} \]
    \[ \text{Focus} \quad \text{FocP} \]
    \[ \text{Verb} \quad \ldots \]

In contrast, it has been proposed that focus-verb adjacency in Basque (Arregi 2002) is achieved with both elements remaining in situ, and ‘altruistic’ movement removing the material that intervenes between them to the periphery of the clause, as in (8). The same mechanism has been proposed for Zulu (Cheng & Downing 2012), in which the position immediately after the verb (IAV) acts as the mirror image of the preverbal focus position in verb-final languages.

---

3 Recall that Georgian also allows for narrow foci to appear in the immediately postverbal position, which is unlike most other verb-final languages. This type of focus is not discussed here.

4 Alternative approaches to the syntax of preverbal focus in Basque have been put forward by Ortiz de Urbina (1989; 2002) and Elordieta (2001).
What is the syntax of the preverbal position in Georgian? Based on a number of syntactic tests, detailed below, preverbal narrow foci in simple clauses in Georgian are interpreted in situ, thus falling into the category of languages that rely on the structure in (8) for the derivation of the preverbal focus position. First, narrow foci are insensitive to island constraints – i.e., they can be embedded in a relative clause or a complex NP – which means that they are not subject to overt or covert movement to the left periphery. This is shown for corrective focus in (B) in (9). In this property, narrow foci contrast with wh-phrases in Georgian (and many other languages), which cannot be embedded in a relative clause – arguably, due to the fact that wh-phrases undergo overt or covert movement to the left periphery, which is disallowed in a relative clause.

(8) XP
    /  
  YP
    /  
Focus ZP      Verb

(9) A: Marik’a-m  i-q’id-a  c’ind-eb-i [RC romeli-c Nino-m
    M.-ERG VER-buy-AOR.3SG socks-PL-NOM which-COMP N.-ERG
    mo-ksov-a].
    PRV-knit-AOR.3SG
‘Marika bought the socks that Nino knitted.’

B: Ara, Marik’a-m  i-q’id-a  c’ind-eb-i [RC romeli-c Nana-m
    no M.-ERG VER-buy-AOR.3SG socks-PL-NOM which-COMP N.-ERG
    mo-ksov-a].
    PRV-knit-AOR.3SG
‘No, Marika bought the socks that NANA knitted.’

Next, narrowly focused constituents align with their broad-focus in-situ counterparts with respect to scope, which suggests that they, too, are found in situ. Specifically, quantified direct objects scope below the position of a low adverb such as seldom. To illustrate, the broad-focus utterance in (10) is more naturally interpreted as describing the situation in which a professor usually calls on more than three students (i.e., rarely calls on less than three; ADV > NUM), as opposed to the situation in which there are less than three students such that the professor rarely calls on them (’NUM > ADV).

(10) Masc’avlebel-i  īšviatad sam-ze  nak’leb  st’udent’-s mo-u-c’od-eb-s.
    teacher-NOM seldom three-on less student-DAT PRV-VER-call-SF-PRS.3SG

---

5 Note that this is not the case for wh-phrases in Georgian, which, while also found in the immediately preverbal position, rely on a Spec-Head configuration instead (Borise 2019).
‘The teacher seldom calls on fewer than three students.’

(ADV > NUM; ?NUM > ADV)

Similarly, a narrowly focused constituents in (B) in (11) takes narrow scope as compared to the adverb იშვათ ʻseldom’:

‘How many students does the teacher seldom call on?’

B: Masc’avlebel-i იშვათ  _sam-ze nak’leb st’udent’-s_ mo-u-c’od-eb-s. teacher-NOM  seldom three-on less  student-DAT  PRV-VER-call-SF-PRS.3SG  
‘The teacher seldom calls on fewer than three students.’

(ADV > NUM; *NUM > ADV)

The fact that narrowly focused objects align with their counterparts in broad focus declaratives (which are found in situ) with respect to their scope properties suggests that narrowly focused objects also are found in situ. Similarly, narrow foci take narrow scope with respect to the material in the left periphery, which also signals absence of movement.

Finally, an argument for the in-situ status of Georgian preverbal foci comes from weak cross-over (WCO) effects. WCO effects are independently attested in Georgian between pronominal subjects and quantified objects, as shown in (12). WCO effects are also reported by some speakers in constructions with pronominal subjects and wh-objects, as shown in (13), though see Amiridze (2006: 62) for the view that such constructions do not give rise to WCO effects. Structurally, it is commonly assumed that WCO effects arise from overt or covert movement of a variable or quantified expression over a coindexed lexical item (Chomsky 1976).

(12) *Mis- ma_i  p’ropesor-ma  mo-u-c’-od-a  titeoul  st’udent’-s.  
3.POSS.SG-ERG  professor-ERG  PRV-VER-call-SM-AOR.3SG each  student-DAT  
(ʻHis professor called each student.ʼ)

(13) %Mis- ma_i  kmar-ma _vin_i  agh-u-c’er-a  Giorgi-s?  
3.POSS.SG-ERG  husband-ERG  who.NOM  PRV-VER-write-AOR.3SG  G-DAT  
ʻWhose husband described her to Giorgi?ʼ  
(lit.: Who did her husband describe to Giorgi?)

In contrast with examples in (12) and (13), preverbal narrow foci do not give rise to WCO effects, which suggests that they are not subject to overt or covert movement:

(14) Deida-misi  _mxolod_ Nino-s  xat’-av-s. 
3.POSS.NOM-3.POSS.NOM  only  N.-DAT  paint-SF-PRS.3SG  
ʻHer aunt draws only Nino.ʼ

---

6 When possessive pronouns are used with kinship names in Georgian, the order of the two elements is the opposite from that used in other contexts: noun-pronoun as opposed to pronoun-noun.
To recap, based on the evidence above, narrow foci in Georgian are interpreted in situ. This means that there is no dedicated syntactic projection that narrowly focused preverbal constituents occupy – instead, Georgian relies on a structure such as the one in (8) to derive focus-verb adjacency. In turn, this means that the syntactic structures that underlie the derivation of narrow focus on e.g. a subject and a direct object differ with respect to the amount of structure that lies between the narrowly focused constituent and the verb, as well as, possibly, the need for ‘altruistic’ movement. Specifically, narrow focus on the object does not require any altruistic movement, given that OV is the underlying structure of the VP in Georgian; instead, both broad focus OV contexts and O₃V contexts with narrow focus on the object rely on the same syntactic structure, as illustrated in (15):

\[
(15) \text{Gušin dila-s bebia [VP samzareulo-s a-lag-eb-d-a].}
\]

\[
\text{yesterday morning-DAT grandma.NOM kitchen-DAT VER-clean-SF-SM-IPFV.3SG ‘Grandma cleaned the kitchen yesterday morning.’ / ‘Grandma cleaned THE KITCHEN yesterday morning.’}
\]

In contrast, narrow focus on the subject of a transitive SOV clause would involve ‘altruistic’ movement of the direct object, resulting in an OS₃V structure, as shown in (16).\(^7\) Narrow focus on a subject of an intransitive clause, S₃V would not involve ‘altruistic’ movement, but would still be syntactically different from O₃V contexts given that there would be more empty structure between the narrowly focused constituent and the verb.\(^8\)

\[
(16) \text{[CP Samzareulo-si [VoiceP bebia [VP ti a-lag-eb-d-a].}
\]

\[
\text{kitchen-DAT grandma.NOM VER-clean-SF-SM-IPFV.3SG ‘GRANDMA cleaned the kitchen.’}
\]

That is to say, even though both the O₃V and S₃V sequences are instantiations of the preverbal focus construction, the syntactic derivations that they rely on differ from each other. Specifically, O₃V structures are fully parallel to the broad focus OV counterparts, with the narrowly focused constituent and the verb being structural sisters. In contrast, in S₃V structures, the narrowly focused constituent and the verb are not merged as part of the same phrase, and, in transitive clauses, their adjacency may result from ‘altruistic’ movement. As the remainder of this paper shows, prosodic structure is sensitive to the differences in syntactic structure that underlie the derivation of preverbal narrowly focused objects and subjects. Specifically, while O₃V contexts receive a prosodic realization that is identical to broad-focus OV contexts, S₃V contexts systematically differ from their broad-focus SV counterparts. As discussed in Section 2.1 above, such dichotomy is predicted by focus projection accounts.

### 2.3. Prosodic realization of focus in Georgian

The prosodic realization of narrow focus in Georgian has received some attention in the literature. In particular, the F0 contours that characterize various focus contexts have been the object of several instrumental investigations; other prosodic cues for information structure, such as non-F0 prominence on

---

\(^7\) I am adopting the view that constituents displaced via ‘altruistic’ movement undergo topicalization into the left (or right) periphery. Nothing in the analysis developed here hinges on the type of displacement that non-focal material undergoes, or its landing site.

\(^8\) The position of the subject in Georgian co-varies with case-marking of the subject. Specifically, nominative subjects are generated in Spec, VoiceP, ergative in Spec, VP and dative in Spec, ApplP. They receive case in situ and do not undergo movement to Spec,TP; for details of case assignment in Georgian, see Legate (2008), Nash (2017), Borise (2019) and Thivierge (2019).
the stressed syllable, voice quality, etc. have also been addressed, though less systematically. This section provides an overview of existing work and discusses the motivation for the current study.

One of the first experimental studies aimed at investigating the prosodic realization of different information-structural contexts in Georgian was conducted by Alkhazishvili (1959), who suggested that a Georgian sentence can be divided into “subject” and “predicate” phrases, which can be roughly equated with the notions of topic and focus/comment, or a logical subject and a logical predicate (which do not necessarily correspond to a syntactic constituent). The “predicate”, in Alkhazishvili’s terms, includes the verb and the immediately preverbal focused constituent, while the “subject” includes all other material in a clause, as illustrated in (17)-(18).

(17) \([\text{“Subject”} \; \text{Giorgi-m}] \quad [\text{“Predicate”} \; \text{pex-i} \; \text{ar} \; \text{ga-a-ndzr-i-a}]\).
G-ERG foot-NOM NEG PRV-VER-move-SM-AOR.3SG

‘Giorgi didn’t move’
(Alkhazishvili 1959: 373)

(18) \([\text{“Predicate”} \; \text{Omarašvil-ma} \; \text{da-i-xsn-a}] \quad [\text{“Subject”} \; \text{gač’irvebi-dan} \; \text{samartal-i}]\).
O-ERG PRV-VER-save-AOR.3SG hardship-from court-NOM

‘Omarashvili was the one who led the court out of the difficult situation.’
(Alkhazishvili 1959: 380)

According to Alkhazishvili’s results, the two phrase types also receive different prosodic realizations: within “subject” phrases, each word is typically characterized by rising prosody, while “predicate” phrases carry an overall falling or a rising-falling pitch contour. A falling pitch contour notwithstanding, Georgian speakers often perceive “predicate” phrases as overall more prominent than “subject” ones. Alkhazishvili’s early results were corroborated in later work. Jun, Vicenik and Lofstedt’s (2007) and Vicenik and Jun (2014) in their description of the prosodic properties of Georgian also observe that the intonational pattern characteristic of out-of-the-blue declaratives is a rising F0 contour that applies to each prosodic word/Accentual Phrase (AP), while preverbal focus contexts are characterized by a shallow falling or rising-falling tune that encompasses the focused constituent and the verb, with the fall in F0 becoming sharper on the penultimate syllable of the predicate.

Many important generalizations about the prosodic realization of information structure in Georgian were established in the work of Skopeteas, Asatiani, Féry, and Fangselow. Skopeteas, Féry, and Asatiani (2009) show that even though Georgian allows for considerable freedom of word order, it is restricted by information-structural factors, such as the requirement for narrowly focused expressions to appear in the immediately preverbal position; they also note the existence of postverbal focus in Georgian. Based on experimental results, Skopeteas and Fangselow (2010) show that contrastive foci can be realized both preverbally and postverbally. In turn, Skopeteas and Féry (2010), based on another instrumental investigation, show that preverbal and postverbal foci have systematically different prosodic (F0) realizations. With respect to the prosodic realization of exhaustivity, Skopeteas and Féry (2011) conclude that it is signaled by pitch expansion combined with increased duration and breathy voice on the first (stressed) syllable of the exhaustively focused constituent. Finally, Skopeteas and Féry (2016) report on several instrumental studies and address a number of topics. They conclude that utterance-initial constituents in broad-focus contexts may carry variable F0 contours, and narrowly focused constituents are characterized by greater duration of stressed syllables as compared to broad-focus contexts. They also provide a detailed
investigation of F0 contours characteristic of preverbal foci (utterance-initial and utterance-medial ones), as well as postverbal foci.

As the above shows, many of the important properties of the prosodic realization of different types of focus in Georgian have been addressed in the literature. Nevertheless, it is worth pointing out that the existing generalizations are based on experimental techniques that did not involve spontaneous production by native speakers. In particular, Skopeteas, Féry, and Asatiani (2009) used read speech in their investigation of the interaction between word order and prosodic contours, and instructed the participants “to put emphasis on the information under question” (Skopeteas, Féry & Asatiani 2009: 110). Skopeteas and Féry (2010; 2011; 2016) did not use read speech, but they asked the participants to memorize the sentences to be used as responses for the experimental questions, which also does not constitute spontaneous production. At the same time, it has been well-documented that the prosodic characteristics of read speech are considerably different from those of spontaneously produced speech, including spectral characteristics of sounds (Nakamura, Iwano & Furui 2008), the amount of variation in the observed F0 contours (Lieberman et al. 1985), placement of accents, phrase boundaries and pauses (Howell & Kadi-Hanifi 1991; Ayers 1994), and type of accents used (Hedberg & Sosa 2008). This means that the conclusions that are drawn from non-spontaneously produced speech may not be applicable to spontaneous speech, the mode of language production/comprehension that is, arguably, most commonly encountered and used by speakers.

The existing work on the prosody of information structure in Georgian also contains conflicting analyses of certain phenomena. For instance, Skopeteas & Féry (2010) and Asatiani & Skopeteas (2012) conclude that preverbal focused constituents are prosodically grouped together with the verb and separated by a prosodic boundary from the material preceding the focus+verb prosodic unit: X(YFV). In contrast, Skopeteas & Féry (2016) propose an alternative analysis of the focus facts, according to which preverbal foci are phrased separately from the verb and grouped together with the preceding material: (XYF)V. The former analysis aligns with the account of prosodic phrasing described as part of the general intonational profile of Georgian by Vicenik & Jun (2014). According to their analysis, prosodic grouping in Georgian – in a range of contexts, including preverbal focus constructions – is signaled by the lack of a high final boundary tone on the first of the two constituents that are prosodically phrased together. In contrast, Skopeteas & Féry (2016) take prosodic phrasing in narrow focus contexts in Georgian to rely on prosodic cues different from those described by Vicenik & Jun (2014), or those used in non-focal contexts more generally.

These two issues – the fact that the existing conclusions are based on non-spontaneous speech and the availability of incompatible analyses – raise the following questions. Are the patterns observed in the available data consistent and representative of all speech modes? Are the analytical approaches to prosodic phenomena based on specific criteria and used consistently for phenomena of similar profiles? These two questions are what guided the current investigation. Compared to the previous studies, it is based on the largest dataset which includes all major types of focus: subject, object, VP, whole sentence. Most importantly, in contrast with most previous studies, the participants in the current experiment were asked to provide their own responses to questions based on picture prompts, which led to greater naturalness of the data.

On the theoretical front, the tonal inventory proposed for Georgian by Vicenik & Jun (2014) is adopted and built upon in the current analysis. As described above, lack of a high final boundary tone between two constituents is taken to be the cue that signals prosodic grouping. Accordingly, prosodic phrasing in the
current analysis is determined based on this highly specific phonetic notion. The results reported here provide support for the approach according to which preverbal subject foci undergo prosodic grouping with the following verb: (S₅V). In contrast, preverbal object foci behave differently: they are separated from the following verb with a high final boundary tone: (O₅)(V). Overall, the current results highlight the fact that experimental techniques that rely of different speech modes may lead to different results. They also demonstrate the need for more consistent use of phonetic cues for establishing the phonology of intonation of a given language.

3 Methodology and results
3.1 Stimuli and design
Eight native speakers of Georgian participated in the study: two males (M3, M4) and six females (F6-F11). All speakers were natives of Tbilisi, with a complete or in-progress university degree, with the age range 20-35 y.o, mean age 26.8 y.o. The recordings were performed in Tbilisi, Georgia, using a Shure SM10A (head-worn, close-range) microphone and a Zoom H4n recorder. All data was recorded at a sampling rate of 44.100 Hz and 16 bits per sample.

During the experiment, the participants were presented with a series of picture prompts that appeared on a laptop screen. Each prompt was accompanied by a statement introducing the main participant of the event in the picture, and a question about it. Both the statement and the question were presented in Georgian orthography, and also provided as an embedded pre-recorded soundtrack, with a native speaker of Georgian reading out the statement and the question. The audio recording was provided in order to make answering the question shown on the screen more natural. A sample picture prompt is provided in Figure 1, with the accompanying statement and question (transliterated) given in (19). The participants were asked to listen to the pre-recorded statement and question, and answer the question based on what they see in the picture. They were instructed to speak clearly and use natural intonation but avoid single-word replies.

![Sample picture prompt](image.png)

Figure 1. Sample picture prompt used in Experiment 1.

(19) Es mebaduri-a. Ra da-i-č’ir-a mebadur-ma
DEM fisherman.NOM-be.3SG what PRV-VER-catch-AOR.3SG fisherman-ERG
šaršan zapxul-ši?
The semi-controlled design of the study allowed speakers to have a certain degree of freedom in their responses while maintaining some control over the lexical and phonological variables. Consequently, the experimental design allowed for making generalizations not only about the prosodic realization of different types of focus, but also about the speakers’ choices of syntactic constructions (word orders) employed in different focus contexts. The latter is addressed in Section 3.2.

The stimuli were designed in such a way as to capture possible syntactic and/or prosodic variability between different constituents carrying focus. Thirty verbs (14 transitive, 9 unergative, and 7 unaccusative) were used in the study, based on the verbs, thirty situational statements were built. Personal names (Mariami, Giorgi, etc.) and common nouns (a fisherman, a ghost, children, etc.) were used as subjects and, for the transitive verbs, also as direct objects. A temporal adjunct was added to each situational statement, in order to provide additional segmental material that may buffer the focal material from phrase-initial or -final prosodic processes, such as initial glottalization or final lengthening. In order to maximize the likelihood of collecting responses with easily analyzable F0 contours, lexical items containing no or few voiceless segments were used. However, naturalness of the stimuli was taken to be no less important than the phonetic make-up, and some better fitting lexical items containing voiceless segments were chosen over fully voiced counterparts that were a poorer contextual fit. A sample situational statement is given in (20):

(20) *Mabadur-ma da-i-č’ir-a zvigen-i šaršan zapxul-ši.*
    fisherman-ERG PRV-VER-catch-AOR.3SG shark-NOM last_year summer-LOC
    ‘The fisherman caught a shark last summer.’

Each of the situational statements was then turned into five questions, aimed at eliciting broad focus over the whole reply, as shown in (21), narrow focus on the direct object (22), subject (23), and the VP (24), and contrastive focus on one of the constituents (subject, object, or the verb) (25):

(21) *Ra mo-xd-a šaršan zapxul-ši?*
    what PRV-happen-AOR.3SG last_year summer-LOC
    ‘What happened last summer?’

(22) *Ra da-i-č’ir-a mebadur-ma šaršan zapxul-ši?*
    what PRV-VER-catch-AOR.3SG fisherman-ERG last_year summer-LOC
    ‘What did the fisherman catch last summer?’

(23) *Vin da-i-č’ir-a zvigen-i šaršan zapxul-ši?*
    what PRV-VER-catch-AOR.3SG shark-NOM last_year summer-LOC
    ‘Who caught a shark last summer?’

(24) *Ra ga-a-k’et-a mebadur-ma šaršan zapxul-ši?*
    what PRV-VER-do-AOR.3SG fisherman-ERG last_year summer-LOC
    ‘What did the fisherman do last summer?’

(25) *R vapexa da-i-č’ir-a mebadur-ma šaršan zapxul-ši?*

---

9 No significant interaction between verb type and prosodic realization of focus or word order employed was subsequently detected; therefore, different verb types are discussed together in the remainder of the paper.
octopus.NOM  PRV-VER-catch-AOR.3SG  fisherman-ERG  last_year  summer-LOC
‘Did the fisherman catch an octopus last summer?’

The resulting full set of questions equaled 150 (30 situational statements x 5 focus types). However, the speaker who was pre-recorded pronouncing the questions failed to pronounce some of them naturally, and the final set that was used in the experiment consisted of 110 questions. Therefore, each speaker provided replies to 110 questions, each of them uttered once.

After eliminating disfluent replies from the participants (due to pauses, errors, etc.), the final dataset consisted of 817 replies. A breakdown of the complete dataset by speaker and focus type is provided in Table 1 (subscript \( F \) stands for focus, \( CF \) for contrastive focus).

<table>
<thead>
<tr>
<th>Table 1. Replies by speaker and focus type.</th>
</tr>
</thead>
<tbody>
<tr>
<td>speaker</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>( F6 )</td>
</tr>
<tr>
<td>( F7 )</td>
</tr>
<tr>
<td>( F8 )</td>
</tr>
<tr>
<td>( F9 )</td>
</tr>
<tr>
<td>( F10 )</td>
</tr>
<tr>
<td>( F11 )</td>
</tr>
<tr>
<td>( M3 )</td>
</tr>
<tr>
<td>( M4 )</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

3.2 Word orders used in focus constructions
Since the participants provided spontaneous replies to the experimental questions, there was considerable variability in sentence structures employed in the replies. In the complete dataset, all replies were coded for their constituent structure in the following way: \( S = \) subject, \( O = \) object, \( V = \) verb, \( X = \) temporal adjunct, \( No = no (ara) \). The resulting list of structures that occurred 10 or more times in the final dataset (with all focus types), broken down by focus type and verb type, is provided in Table 2.

Table 2 allows for some informative generalizations. First, narrowly (non-contrastively) focused objects \( (O_F) \) appear preverbally and postverbally with almost equal frequency: 16 narrowly focused objects are found in \( SVO_F \) and \( SO_F V \) clause types, respectively; similarly, \( XSVO_F \) clauses were used in constructions with narrowly focused objects \( (O_F) \) 6 times, and \( XS_O_F V \) clauses 5 times. The only clause structure that favors preverbal placement of narrowly focused objects and does not have a \( VO_F \) counterpart is \( SXO_F V \), used 6 times. The picture is markedly different for contrastively focused objects \( (O_{CF}) \). Here, there is a strong preference for postverbal placement of contrastively focused constituents has been reported for Georgian (Skopeteas & Fanselow 2010). Narrowly focused subjects behave differently from narrowly focused objects: both contrastively and non-contrastively focused subjects \( (S_{CF}, S_F) \) are found exclusively in the immediately preverbal position. It is unclear why postverbal narrowly focused subjects did not occur in the current dataset, despite being attested in the elicitation setting, as well as in the literature.
Broad focus contexts and VP focus contexts ([VP]F) typically align in their syntactic realization. One noticeable difference between the two is found with respect to clause types that allow for adjunct material to appear between the subject and the rest of the clause, such as SXOV, SXV, and XSXV contexts: these are more commonly found in [VP]F contexts, as opposed to broad focus contexts: SX[OV]F with transitive verbs, and SX[V]F, and XSX[V]F with intransitive verbs.

**Table 2. Clause structures per verb type and focus type; the highlighted values are addressed in the discussion.**

<table>
<thead>
<tr>
<th>Verb type</th>
<th>Clause type</th>
<th>Focus type (redoing for the paper, Sept 2019)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>[VP]F</td>
<td>broad</td>
</tr>
<tr>
<td>transitive</td>
<td>SV</td>
<td></td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>XSV</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>NoSV</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SVO</td>
<td></td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>SOV</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>XSVO</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>SXOV</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>XOSV</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>NoSVO</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>SVOX</td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>intransitive</td>
<td>SV</td>
<td></td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>XSV</td>
<td></td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>SVX</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>SXV</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>XSXV</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>NoSV</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NoSVX</td>
<td></td>
<td>20</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>160</td>
</tr>
</tbody>
</table>
be directly related to the speaker’s proficiency in Russian, though it likely has some sociolinguistic significance (recall that there is a preference for OV in more formal contexts, and VO in less formal ones).

### Table 3. Individual speaker preferences for SOV or SVO.

<table>
<thead>
<tr>
<th>Clause type</th>
<th>Focus type</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>F6</td>
</tr>
<tr>
<td>VO type</td>
<td>broad</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>[VP]F</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>O_F</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>O_CF</td>
<td>5</td>
</tr>
<tr>
<td>OV type</td>
<td>broad</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>[VP]F</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>O_F</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>O_CF</td>
<td></td>
</tr>
</tbody>
</table>

### 3.3 Duration of the stressed syllable

Other than allowing for making generalizations about speakers’ preferences for word orders in different information-structural contexts, the dataset introduced in Sections 3.1 and 3.2 was used for the acoustic analysis of the prosodic properties that characterize different types of focus – notably, the duration of the stressed syllable and the F0 contour spanning the narrowly focused constituent. The prosodic data obtained was annotated in Praat (Boersma & Weenink 2018); average duration and F0 of each syllable, as well as F0 at four fixed points in a syllable (left edge, 1/4, 2/4, 3/4) were measured using a modified Praat script by Elvira-García (Elvira-García 2014); before that, the correct syllabification was established with three native speakers of Georgian. The syllable duration results are presented and discussed in Section 3.3, and F0 results are addressed in Section 3.4.

While the existence and nature of word stress in Georgian have been a matter of debate for a considerable amount of time (Selmer 1935; Zhghenti 1953; 1959; 1963; 1964; Alkhazishvili 1959; Tschenkeli 1958; Vogt 1971; Tevdoradze 1978, a.o.), recent experimental evidence shows that Georgian has fixed initial stress, cued by syllable duration (Vicenik & Jun 2014; Borise & Zientarski 2018; Borise 2019). Specifically, it has been shown that the initial syllable in Georgian is always characterized by greater duration than all subsequent syllables, and this durational effect cannot be explained by initial strengthening, since it does not stem from the greater prominence of the initial consonant alone (Borise, submitted); cf. Barnes (2008), who argues for two main sources of greater duration of the initial syllable, stress and initial strengthening.

With respect to narrow focus, it has been observed for a variety of languages that the acoustic effect of stress (which may rely on various acoustic means) becomes more prominent if a word carries narrow focus. For instance, if word stress is cued by duration or F0 movement, this effect may become more pronounced when the word is narrowly focused. In particular, the increase in duration of the stressed syllable under

---

10 In Praat, the following settings were used for the F0 analysis: pitch range 75-500 Hz for the female speakers and 50-450 Hz for the male speakers, voicing threshold = 0.6, octave jump cost = 0.6.
narrow focus has been documented for English (Xu & Xu 2005) and German (Braun & Ladd 2003; Baumann et al. 2007; Kügler & Genzel 2009), and the increase in the duration of the onset and coda of the stressed syllable under narrow focus has been observed in Dutch (Hanssen, Peters & Gussenhoven 2008). More generally, an increase in the duration of the word carrying narrow focus is reported for Swedish (Allwood 1974), English (Cooper, Eady & Mueller 1985; Eady & Cooper 1986; Eady et al. 1986; Sityaev & House 2003), German (Baumann, Grice & Steindamm 2006; Féry & Kügler 2008; Kügler 2008) and Korean (Jun & Kim 2007).11

In contrast with these languages, according to the current results, the picture is more complex in Georgian: here, the stressed syllable receives greater duration, as compared to a broad focus context, only in a subset of narrow focus conditions. In particular, narrowly focused subjects in the preverbal position are marked by greater duration of the initial syllable, but the same does not hold for narrowly focused preverbal direct objects. Consequently, duration of the stressed syllable cannot be treated as a consistent acoustic marker of narrow focus. From a theoretical point of view, the observed pattern of prosodic focus marking – special prosody being used for focused subjects but not focused direct objects – is consistent with focus projection accounts, according to which the prosodic realizations of broad focus and narrow focus on the direct object are predicted to be identical or nearly identical, while the prosody of narrowly focused subjects is expected to differ from them.

In order to determine whether the duration of the stressed syllable is affected by the focused vs. non-focused status of the word, the following analysis was carried out. The object of comparison was the duration of initial syllables in nouns that occupy the same position in clauses of the same syntactic type, but with different focus properties. Accordingly, the clause types considered included OV and SV sequences; other segmental material that preceded or followed these sequences in the participants’ replies to the experimental questions was trimmed off. The total counts of items in the obtained dataset are provided in Table 4.

Note that the discrepancy in total item counts between the two conditions, OV and SV, is due to the semi-controlled nature of the experiment and the fact that the participants were free to choose a preverbal or postverbal focus construction in their response. While they placed narrowly focused objects preverbally and postverbally with almost equal frequency, narrowly focused subjects in the current dataset were found exclusively preverbally, as shown in Table 2. The reason for such skewing is unclear, given that postverbal narrowly focused subjects are attested in the literature (e.g. Skopeteas & Fanselow 2010) and in an elicitation setting. Similarly, different total counts of n-syllabic items are due to the fact that the participants were, to a degree, free to choose what lexical items to use. Since the statistical analysis was based on a mixed-effects model, with SYLLABLE COUNT as a random factor, these discrepancies do not affect the results of the analysis.

Table 4. The dataset used for the investigation of the duration of the stressed syllable under different focus conditions.

<table>
<thead>
<tr>
<th>Clause type</th>
<th>Arguments</th>
<th>Syllable count</th>
<th>n</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>SV</td>
<td>subjects</td>
<td>3 σ</td>
<td>211</td>
<td>306</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 σ</td>
<td>95</td>
<td></td>
</tr>
</tbody>
</table>

11 On the other hand, there is also work suggesting that the acoustic cues that are used for the realization of word stress and those that are used for the prosodic realization of narrow focus are orthogonal (due to their functional load). In particular, Vogel, Athanasopoulou & Pincus (2016) show, for Spanish and Greek, that word stress is cued primarily (though not exclusively) by F0, while contrastive focus is cued primarily (though not exclusively) by duration.
Statistical analysis of the syllable duration data was performed using the `glmer` function in the `lme4` package for R (R Core Team 2017). For each of the clause types (SV and OV), a mixed-effects model with DURATION (of the initial syllable) as the dependent variable, FOCUS TYPE as the fixed factor and random factors SPEAKER, ITEM, CLAUSE SUBTYPE (such as, e.g. SVO, SV or XSV for subjects) and SYLLABLE COUNT (with a random intercept for each of those predictors, but no random slopes) was run. In the model, the broad focus condition acted as the intercept and was taken to be the baseline that other types of focus are compared with. As the results below show, greater duration of the initial syllable is correlated with narrow focus only for preverbal subjects, both contrastive and non-contrastive, but not narrowly focused direct objects in the preverbal position.

Specifically, with respect to subjects, the duration of the stressed (initial) syllable was significantly greater in narrowly focused subjects (non-contrastive and contrastive; p=0.02 and p=0.01, respectively) than in subjects found in broad focus utterances, but not in any of the other conditions tested, as shown in Figure 2 and Table 5.

![Figure 2](image.png)

Figure 2. Average syllable durations in subjects in SV-type clauses with different focus types.

<table>
<thead>
<tr>
<th>OV</th>
<th>objects</th>
<th>1σ</th>
<th>2σ</th>
<th>3σ</th>
<th>4σ</th>
<th>5σ</th>
<th>92</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>17</td>
<td>31</td>
<td>6</td>
<td>34</td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Average durations of the initial (stressed) syllable in subjects in SV sequences with different information-structural properties; asterisks mark values significantly different from the intercept (broad focus condition).
As compared to preverbal subjects, the results for preverbal objects were markedly different. In contrast with the results for subjects in SV sequences, there were no significant differences between the durations of stressed syllables in direct objects in OV sequences with various information-structural properties (broad focus, narrow focus on the object or the VP). The results for objects are provided in Figure 3 and Table 6. Note that the OV sequences did not include subject focus conditions, since all focused subjects in the current dataset were found in the preverbal position – that is, in SV sequences – as discussed in Section 3.2. Similarly, contrastively focused objects were found exclusively postverbally in the current dataset, which means that OV sequences included only non-contrastively focused objects. These two factors lead to there being fewer information-structural contexts tested in OV contexts than in SV contexts.

![Figure 3](image-url)

**Figure 3.** Average syllable durations in objects in OV-type clauses with different focus types.

<table>
<thead>
<tr>
<th>Focus type</th>
<th>n</th>
<th>1st σ, ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>broad</td>
<td>83</td>
<td>181</td>
</tr>
<tr>
<td>O</td>
<td>14</td>
<td>161</td>
</tr>
<tr>
<td>OCF</td>
<td>7</td>
<td>134</td>
</tr>
<tr>
<td>S</td>
<td>95</td>
<td>192*</td>
</tr>
<tr>
<td>SCF</td>
<td>34</td>
<td>216*</td>
</tr>
<tr>
<td>[VP]F</td>
<td>73</td>
<td>160</td>
</tr>
</tbody>
</table>

**Table 6.** Average durations of the initial (stressed) syllable in objects in OV sequences with different information-structural properties.
To sum up, the results above show that the durational properties of the stressed (initial) syllable in Georgian are correlated with the expression of focus only in a subset of preverbal focus conditions. Recall that, based on the existing studies, our starting assumption was that narrow focus on a constituent affects the realization of the stressed syllable in that constituent, making it more prominent, with the phonetic realization of such focus-induced prominence (F0 value, duration, etc.) being language-specific. For Georgian, it has been independently established that syllable duration marks word stress, fixed on the initial syllable. Contrary to the starting assumption, the data in the current section showed that duration of the stressed syllable is not a reliable cue for narrow focus in Georgian. Specifically, greater duration of the stressed syllable was found only in a subset of the data inspected—in narrowly focused preverbal subjects. In contrast, preverbal narrowly focused direct objects were not marked by greater duration of the initial syllable, as compared to broad focus conditions.

In theoretical terms, this is exactly the distribution of focus prominence that is predicted by focus projection accounts: while the prosodic marking of narrowly focused subjects differs from the prosody of broad focus, the prosody of narrow focus on the direct object may be identical or nearly identical with the latter. In the next section, another acoustic cue that is used in the prosodic marking of narrow focus in Georgian is discussed: F0 contours that signal patterns of prosodic phrasing.

3.4 F0 values
Before turning to the F0 properties of narrow focus contexts, let us first briefly consider the F0 contour typical of broad focus utterances, since in the remainder of this section it acts as the benchmark that narrow focus contexts are compared to. It has been established in the literature that the neutral intonational contour found in broad focus declaratives in Georgian consists of a series of Accentual Phrases (APs), each carrying a rising F0 contour (Jun, Vicenik & Lofstedt 2007; Vicenik & Jun 2014), as shown in Figure 4, with the gloss and translation of the example provided in (26). The sharp final rise in F0 at the end of one prosodic word (high final boundary tone Ha) and a steep fall on the following one (low pitch accent L*) are indicative of a prosodic boundary between adjacent APs. This neutral intonational contour is found in out-of-the-blue declaratives and broad-focus utterances, such as replies to a question What happened?.

(26) Giorgi-s mosc’on-s dzalian lamaz-i gogo Tbilisi-dan.
    Giorgi-DAT like-PRS.3SG very beautiful-NOM girl-NOM Tbilisi-from
    Giorgi likes a very beautiful girl from Tbilisi.
Figure 4. Typical broad-focus intonation in Georgian; each word is an AP with overall rising intonation; downstep applies to each successive high final boundary tone Ha (Borise 2017: 92).

Additionally, Georgian allows for optional grouping of two or more APs, typically lexical items that form a syntactic constituent, such as a noun and a modifying adjective, into an intermediate phrase (ip) (Jun, Vicenik & Lofstedt 2007; Vicenik & Jun 2014). APs phrased together do not have a high boundary tone between them and instead carry a single rising F0 contour over both prosodic words, which ends in a high final boundary tone of the ip (H-). Compare the APs dzalian ‘very’, lamazi ‘beautiful’, and gogo ‘girl’ in Figure 4, which are not phrased together, to the single ip lamazma kalbat’onma ‘beautiful lady’ in Figure 5.

(27) Lamaz-ma kalbat’on-ma k’aba mo-i-zom-a.
beautiful-ERG lady-ERG dress.NOM PRV-VER-try-AOR.3SG
‘A beautiful lady tried on a dress.’
Figure 5. The fact that the noun phrase lamazma kalbatonma ‘beautiful lady’ forms a single prosodic constituent (ip) is manifested by the lack of a final boundary tone on lamazma ‘beautiful’ (Borise 2019).

The data discussed in the remainder of this section shows that the function of prosodic grouping in Georgian goes well beyond such optional cases. Specifically, prosodic grouping, akin to that shown in Figure 5, plays a crucial role in the prosodic realization of preverbal narrow focus in Georgian. As the data below shows, preverbal subject foci exhibit no evidence for a prosodic boundary between themselves and the following verb; instead, a focused subject and the verb are realized with a single falling F0 contour, which signals that the two elements have undergone prosodic grouping. This is illustrated in Figure 6, with the gloss and translation of the example provided in (28):
(28) *Gasul ghame-s Mariam-i sadil-ob-d-a.*
    last night-DAT M.-NOM dine-SF-SM-IPFV.3SG
    ‘Last night MARIAMI had dinner’

The picture is different with preverbal object foci, which are separated from the following verb with a high final boundary tone Ha, in a parallel fashion to preverbal objects found in broad focus conditions. This is illustrated in Figure 7, with the glosses and translation of the example provided in (29):

(29) *Šaršan gazapxul-ze Ilona q’vavil-eb-s a-grov-eb-d-a.*
    last_year spring-in I.NOM flower-PL-DAT PRV-pick-SF-SM-IPFV.3SG
    ‘Last spring Ilona was picking FLOWERS’
In order to analyze the F0 dimension of narrow focus marking in Georgian, F0 was measured in each word at four points per syllable (left edge, 1/4, 2/4, 3/4), in order to allow for reliable tracking of the alignment and shape of F0 movements within a word. A problem for a direct comparison between the individual items, however, lay in the fact that they are not of equal syllable count, which makes aligning the items with each other, in order to compare their F0 properties, problematic. When investigating syllable duration data in Section 3.3, the issue of unequal syllable count was obviated by only targeting initial syllables, regardless of syllable count in a word, and including SYLLABLE COUNT as a random factor into the mixed-effects model. The same cannot be done with respect to F0, however, since F0 targets can be both left- and right-aligned in a given prosodic domain, so choosing one or the other word/phrase edge in order to align the stimuli of unequal syllable count with each other can obscure generalizations about the location of intonational F0 targets. To illustrate, in a language that carries F0 targets both at the left and right edges of a word/phrase, aligning stimuli of varying syllable counts with each other, by either left or right edge, would not allow to capture the distribution of F0 targets (syllables of interest shown by bolding in (30)):

(30) a.  .sorted  b.  sorted
    sorted  sorted
    sorted sorted

Georgian is exactly the kind of a language that is shown in (30): it carries intonational F0 targets on the penult and ultima, and may also intonationally mark the left edge of a word/phrase (Jun, Vicenik & Lofstedt 2007; Vicenik & Jun 2014; Borise 2017; 2019). Therefore, in order to obviate the alignment problem for the purposes of the statistical analysis of the F0 data, each word/AP in the dataset was reduced to the three syllables that are known to carry F0 targets: the initial syllable, penult and ultima. Word-medial syllables, the loci of tonal interpolation between the edge-aligned F0 targets, were discounted; cf. also Skopeteas &
Féry (2016), who used the same technique for a similar comparison of prosodic words of unequal syllable length.\footnote{Note that Skopeteas & Féry (2016) in their study on Georgian prosody used three final syllables (antepenult, penult, ultima), as opposed to two final syllables used here. The decision to discount the antepenult as a potential F0 locus was made based on the fact that no durational or F0 effects were found on the antepenult in the previous studies (Alkhazishvili 1959; Borise & Zientarski 2018; Borise 2019).}

Consequently, each resulting item in the final dataset was trisyllabic and contained only the initial syllable (henceforth coded ‘1’), penult (coded ‘-2’) and ultima (coded ‘-1’) of the original word. Therefore, e.g. all SV sequences for the purposes of statistical analysis are reduced to six syllables: S1, S-2, S-1, V1, V-2, V-1. For instance, an utterance Mariami sadilobda ‘Mariami had dinner’ was reduced to (Ma)S1 (a)S-2 (mi)S-1 (sa)V1 (lob)V-2 (da)V-1. As was the case in the analysis of syllable duration data in Section 3.3, SV and OV sequences, with various information-structural properties, were extracted from the final dataset; other segmental material from each response was discounted. Because the stimuli had to be at least three syllables long, in order to adhere to the syllable coding technique used here, SV and OV containing mono- and disyllabic words were discounted as well. The total counts of the stimuli are provided in Table 7; note that, as was the case for the syllable duration data discussed in Section 3.3, the discrepancy between the total counts of SV and OV constructions is due to the fact that VO constructions were used by the participants with almost equal frequency as OV ones, while VS constructions were not.

Table 7. The dataset used for the investigation of the F0 properties of different focus conditions.

<table>
<thead>
<tr>
<th>Clause type</th>
<th>Arguments</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>SV</td>
<td>subjects</td>
<td>306</td>
</tr>
<tr>
<td>OV</td>
<td>objects</td>
<td>71</td>
</tr>
</tbody>
</table>

Based on the available literature (Skopeteas & Féry 2010; 2016; Vicenik & Jun 2014) and individual examples, such as the ones in Figure 6 and Figure 7, the F0 property that may signal narrow focus in Georgian is the absence of a high boundary tone between the preverbal focused constituent and the verb. Therefore, with reference to the syllable-coding technique used in the current dataset, F0 properties of syllables S-1 and O-1 in SV- and OV-sequences, respectively, are the most informative ones for the two types of preverbal foci, since they are the ones that may or may not carry the final high boundary tone Ha. Therefore, the mid-syllable F0 value (1/2 into the syllable) for each S-1 and O-1 syllable was used in the statistical analysis, given that F0 peaks are most frequently aligned with the middle point of a syllable. The analysis of the F0 values was performed using a mixed effects model, with the help of the lmer function in the lme4 package for R (R Core Team 2017). For the SV and OV clause types, a model with the dependent variable F0_1/2 (of the syllable of interest), fixed factor FOCUS TYPE and random factors SPEAKER, ITEM, and CLAUSE SUBTYPE was run; the broad focus condition acted as the intercept and was taken to be the baseline that other types of focus were compared with.

Based on the existing literature on the prosodic realization of narrow focus, the null hypothesis is that utterances containing narrow foci systematically differ in their prosodic realization from broad-focus utterances, given that prosody is commonly used to mark narrow focus. However, as the results below show, and in line with the syllable duration data presented in Section 3.3 above, the null hypothesis was found to hold only for a subset of narrow foci in Georgian. Specifically, it was found that narrow focus on preverbal
subjects, both contrastive and non-contrastive, is manifested by a lack of a high final boundary tone on the subject. The same is not true of preverbal narrowly focused objects, which are separated from the following verb by a high boundary tone, in the same way direct objects are in broad focus contexts, and contexts with narrow focus on the VP.

First, consider the prosodic realization of subjects in SV sequences with various information-structural properties, as summarized in Figure 8 and Table 8. According to the statistical results, the F0 realization of the final syllable in subject focus contexts, both contrastive and non-contrastive, is significantly different from the broad focus conditions ($p<0.01$ for both). None of the other contexts significantly differ from the intercept (broad focus).

![Figure 8](image)

**Figure 8.** Averaged F0 contours in SV sequences with various types of focus, with each word reduced to three syllables ($1=S_1$, $2=S_2$, $3=S_3$, $4=V_1$, $5=V_2$, $6=V_3$), smoothed at 0.2. On the x-axis, each tick mark corresponds to the onset of the respective syllable.

**Table 8.** Averaged F0 values for the syllables of interest ($S_1$) in SV sequences with various types of focus; asterisks mark values significantly different from the intercept (broad focus condition).

<table>
<thead>
<tr>
<th>Focus type</th>
<th>n</th>
<th>$S_1$, Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>broad</td>
<td>83</td>
<td>248</td>
</tr>
<tr>
<td>O</td>
<td>14</td>
<td>272</td>
</tr>
<tr>
<td>O_{CF}</td>
<td>7</td>
<td>312</td>
</tr>
<tr>
<td>S</td>
<td>95</td>
<td>221*</td>
</tr>
<tr>
<td>S_{CF}</td>
<td>34</td>
<td>210*</td>
</tr>
<tr>
<td>[VP]_{F}</td>
<td>73</td>
<td>267</td>
</tr>
</tbody>
</table>
In turn, Figure 9 and Table 9 provide a summary of the F0 properties of OV sequences found in different information-structural contexts. As already noted in Section 3.3, OV sequences are considered in the context of fewer focus conditions given that narrowly focused subjects in the current dataset were exclusively preverbal, and, as such, not found in OV sequences, thereby excluding S_F and S_CF from consideration in the set of OV sequences. Similarly, contrastively focused objects, O_CF, were found only postverbally, and, as such, do not appear in OV contexts either. As is clear from Figure 9, all three remaining focus conditions considered – broad focus, [VP]_F and O_F – closely align in their prosodic realization: in particular, in all of them, there is a high boundary tone between the preverbal object and the verb, with the heights of the F0 peaks being nearly identical across the three conditions. None of the contexts turned out to be significantly different from each other.

![Figure 9](image_url)  
*Figure 9.* Averaged F0 contours in OV sequences with various types of focus, with each word reduced to three syllables (1=O_1, 2=O_2, 3=O_3, 4=V_1, 5=V_2, 6=V_1), smoothed at 0.2. On the x-axis, each tick mark corresponds to the onset of the respective syllable.

<table>
<thead>
<tr>
<th>Focus type</th>
<th>n</th>
<th>O-1, Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>broad</td>
<td>21</td>
<td>187</td>
</tr>
<tr>
<td>O_F</td>
<td>20</td>
<td>190</td>
</tr>
<tr>
<td>[VP]_F</td>
<td>30</td>
<td>188</td>
</tr>
</tbody>
</table>

*Table 9.* Averaged F0 values for the syllables of interest (O-1) in OV sequences with various types of focus.

To sum up, the F0 results discussed in the current section show that while preverbal narrowly focused subjects undergo prosodic grouping with the following verb, which is manifest in the lack of a prosodic
boundary between them, the same is not true of narrowly focused preverbal objects. The latter are not prosodically grouped with the following verb and instead are marked by a high final boundary tone, as preverbal objects in broad focus and VP-focus contexts also are. Overall, the F0 results align with the syllable duration results discussed in the preceding section: namely, both phenomena, duration of the stressed syllable and the F0 contour, identify contexts with narrowly focused preverbal subjects as distinct from subjects found in broad-focus utterances, and identify narrowly focused preverbal objects as identical or nearly identical with their counterparts in broad focus conditions.

4 Summary and theoretical implications

The current results allow for making a number of generalizations that have implications for the nature of the syntax-phonology interface. Before delving into them, let us briefly recap the results. As was shown in Section 3.3, the duration of the initial (stressed) syllable is correlated with narrow focus on subjects, in that the initial syllable of a narrowly focused subject is longer than in broad focus contexts. Notably, the same generalization does not apply to narrowly focused preverbal objects: their prosodic realization is similar to that of broad focus utterances, with the focused constituent and the verb separated by a high boundary tone.

In turn, Section 3.4 showed that the F0 contour that spans narrowly focused subjects is significantly different from that found on subjects in broad focus contexts: specifically, in the former the subject is prosodically grouped together with the verb, while in the latter it is separated from it by a high final boundary tone. With narrowly focused objects, the picture is different: in their F0 properties, they align with both [VP]F contexts and broad-focus contexts. Therefore, both prosodic phenomena, duration of the stressed syllable and the F0 contour, are used to mark subject focus contexts, but neither differentiates object focus contexts from broad focus ones. This is not unexpected, however: the fact that the realization of narrow focus on the object, but not the subject, patterns together with that of broad focus is expected from the point of view of focus projection accounts.

The current results should not be interpreted as showing that the prosodic realization of broad focus and narrow focus on the object are completely identical though. Instead, they only show that there is no systematic difference between the two focus contexts with respect to the two prosodic phenomena considered here, the duration of the stressed syllable and the F0 contour. Though these acoustic cues are among the most important ones, and those likely to be correlated with the expression of focus, they are not the only ones. Specifically, it has been shown that other acoustic cues, such as intensity, as well as duration of words and silences between them, may also take part in marking narrow focus (Breen et al. 2010). Similarly, the prosodic realization of the parts of the utterance other than the focused constituent, such as post-focal reduction in F0 movements, has been shown to be the main correlate of narrow focus in some languages (cf. Patil et al. 2008 for Hindi). Therefore, it is not implausible that a difference between the prosodic realizations of broad foci and object foci in Georgian may lie in these other factors, which remain to be investigated.

The syllable duration and F0 results discussed here contrast with those reported in Skopeteas and Féry (2016), who found no notable differences between the behavior of preverbal subjects and objects. Specifically, they report greater duration of the stressed syllable to be a consistent correlate of narrow focus for both preverbal subjects and objects. With respect to F0, they found both subject and object narrow foci to be phrased separately from the following verb. It is not immediately clear what the discrepancy between the two studies may be attributable to, though experimental set up is a likely contributing factor. Recall that
the participants in Skopeteas and Féry’s study memorized and pronounced replies to experimental questions that were prepared in advance, while the participants in the current study produced spontaneous replies when prompted by a question and a picture. It is likely that the contrast between narrowly focused subjects and objects obtained in the current study is characteristic of an experimental set-up that is more consistent with spontaneous language production.

Analytically, the contrast in the prosodic realization of narrow focus on subjects and objects reported here is exactly the one expected from the point of view of focus projection accounts. Recall that, according to these accounts, broad focus, [VP]- and [O]-contexts may have the same prosodic realization, which in English is manifested as prosodic prominence/nuclear stress on the direct object. Certain syntax-based focus projection accounts further specify that this prosodic prominence is associated with the direct object as the most deeply embedded constituent, and projects from that position up the syntactic tree via each embedding phrase, such as VP, vP, TP, etc. In contrast, subjects do not act as the most embedded constituents; therefore, the prosodic prominence found on subjects in subject focus contexts does not take part in focus projection. Consequently, the prosody of subject focus contexts is incompatible with focus on a larger embedding phrase. Therefore, the results obtained here are fully expected from an analytical point of view, and provide empirical support for the notion of focus percolation.

What is novel about the current study is that it attests to the existence of a contrast between subjects and objects with respect to their focus projection properties in a language with a dedicated preverbal focus position. Specifically, it shows that the contrast in question – the one predicted by focus projection accounts – is present even in a language in which subject and object foci are linearly found in the same position. At the same time, recall from Section 2.2 that, syntactically, narrow foci in Georgian are found in their in-situ positions. Accordingly, even though linearly found in the same position, object foci are sisters to verbs, while subject foci are merged higher in the syntactic tree and, in a transitive clause, require ‘altruistic’ movement of the direct object in order to achieve adjacency with the verb. When these syntactic facts are considered together with the current prosodic results, they show that prosodic structure is sensitive to differences is syntactic structure even when these structural syntactic differences are not reflected in the linear word order.

**Abbreviations/glosses**

**Acknowledgments**
I thank Maria Polinsky, Jonathan Bobaljik, Lauren Clemens, Kevin Ryan, Marcel den Dikken, Balázs Surányi, Stavros Skopeteas, Rusudan Asatiani, Léa Nash, Kriszta Szendrői, Byron Ahn, and Karlos Arregi for their helpful comments and suggestions on the earlier versions of this paper. Many thanks are due to the Georgian consultants who provided the prosodic data and acceptability judgements used here, for their patience and generosity with their time. I am also grateful to Nastassia Kotava, who prepared the experimental prompts, and the research assistants who annotated the prosodic data. All remaining errors are my responsibility.
Funding information
This project was supported by a Graduate Research Travel Grant from Davis Center for Russian and Eurasian Studies at Harvard University.

Competing interests
The author has no competing interests to declare.

References


