Somali as a Tone Language

David Le Gac

1 Normandie Université, UR, DySoLa, France
david.legac@univ-rouen.fr

Abstract

Since [1]’s seminal paper, Somali has usually been considered to be a tonal/pitch-accent language. Recently, phonologists have cast doubt upon the pertinence of such a prosodic class, arguing that pitch-accent languages do not form a coherent category with distinctive criteria and can be reduced to either stress/accent or tonal languages. This paper outlines a tonal analysis of Somali. It aims at showing that the multiple pitch patterns observed in different classes of words and syntactic contexts find a more accurate account within an approach using tonal features only.

Index Terms: pitch-accent, tone, Somali, focus, morphology

1. Introduction

The question of whether the pitch patterns observed in Somali, a Cushitic language spoken in Somalia and parts of Djibouti, Kenya and Ethiopia, are of a tonal and/or accentual nature is still a matter of controversy. [2] and [3] considered Somali as primarily a tone language; [4] took an opposite view and emphasized the accentual nature of the pitch contours. In his seminal paper, [1] put forward the idea that Somali is actually a tonal accent language. In this intermediate approach, the underlying forms have neither accent nor tone. Accents are assigned to the penultimate or final vowel by morphological rules. After the application of accent shift and deleting rules, the surviving accents receive invariant high tones. Thereafter, the tonal (or pitch) accent analysis was adopted in its broad outlines by [5]–[13]. However, in recent papers, Hyman himself ([14]) and [15] have called into question the tonal accent analysis of Somali, along with that of other languages, such as Japanese or Basque, for instance. More generally, they discuss the suitability of a category of ‘tonal or pitch accent languages’, though these authors defend opposite views.

[15] argues that Somali and probably many of the so-called ‘pitch accent’ languages are fundamentally accentual. He bases his claim on a remark by [1, p. 194], who wrote: “Rather than having a process accent → tone → pitch, it is possible to go directly from the accent specifications to the pitch integers [heights] themselves.” [15] thus asserts that Somali has no tones at the phonological level, but accents, whose acoustic correlate is pitch. This accounts for the distributional restrictions of the high pitches within the Somali words – at most, one accent per word, on the final or penultimate mora –, which are typical of accent/stress. Nonetheless, some Somali word categories may have no prominence at all. Yet it is widely accepted that, in accentual systems, every lexical word has at least one accented syllable. [15] then claims that Somali testifies to the fact that one of the four parameters he proposes to account for the various accentual systems found in languages (the ‘Default’ parameter) does need not to be active; this is a property of the so-called ‘pitch-accent languages’.

If one wishes to maintain the tonal accent hypothesis, then it has to be shown that the high pitches are not the mere acoustical ‘reflexes’ of underlying accents and that tones and accents are both independent and necessary in Somali. This is what [10]–[12] suggest in a series of works dealing with the intricate tonal alternations appearing in nouns (N). To explain why the Ns of [8]’s declensions (D) 6 and 7 undergo tonal shifts, while those of the other Ds do not, he hypothesizes that the D6/7 Ns have two accents (*) for morphological reasons. To account for the fact that there is only one high tone (H) per word, he adopts the framework of [16] and proposes that tonal-accent H tones are boundary tones associated with each prosodic word (ω) – one H per ω – and secondarily link to an available *. Furthermore, Hωs may interact with tones associated with higher prosodic constituents (‘a’, ‘i’), which causes the tonal shifts in D6/7 nouns. Conversely, the Ns of other Ds (D1–4) have no underlying * in plural; Hω is then not realized except under specific conditions (focalization, prosodic heading within DetPs). As seen, Hωs exists independently from * and does not always require an * to surface. Likewise, in a paper addressing Somali wordwood and its relationship to prosodic structure, [13] crucially distinguish the tonal aspects from the accentual behaviors of words.

[14], on the other hand, emphasizes that ‘obligatoriness’ is a core property of accent. Accordingly, the fact that words can be unaccented demonstrates that Somali is not an accentual system. In addition, the mora is the tone bearing unit in Somali, while another widely acknowledged property of stress systems is that the stress-bearing unit is the syllable. Regretting his own 1981 proposal, [14] claims that Somali is, in fact, a restricted tone system in which the high pitches are the phonetic realizations of H tones. He even suggests that the cumulative and non-obligatory tones of Somali could be analyzed as grammatically conditioned boundary H% tones. More generally, [14] claims that pitch-accent is not a coherent notion, but rather, a “pick and choose” among the properties that characterize prototypical tone vs. prototypical accent. He therefore recommends that “no language should be analyzed with pitch-accent. A tonal analysis is always possible”. ([14, p. 661]).

The aim of this paper is to test [14]’s recommendation and propose a tonal analysis of Somali. As mentioned above, one of the intriguing problems of Somali is the complex tonal alternations that affect most word categories and especially Ns. [1] was a first attempt to give a principled account of these alternations in a ‘classical’ generative framework using ordered rules. [5], [6], [8], [10]–[13] considered accents as fundamental units as well, even though tones could endorse a more important role. Now, how can these alternations be accounted for without any accent? [14] actually left this issue open. In this paper, I will show that, indeed, tone features are needed to explain the varying pitch patterns that characterize Ns and other word categories in Somali. On the other hand, accent will be proven to be unable to capture these pitch variations and should be removed from the phonological representation of words.
This paper is based on the data provided by [3], [8], [9], [17] and in particular by [10], who recorded two speakers from Somalia (Mogadishu, Hargeisa) and studied the Somali prosody on an instrumental and experimental basis. The theoretical background used to analyze the data is that of Government Phonology (GP, [18], [19]), Distributed Morphology (DM, [20], [21]) and Phase Theory (PT, [22], [23]). GP and DM have recently achieved promising results in studying Somali morphology ([24]–[29]). Likewise, PT provides a new way of understanding the manner in which a sentence is assigned prosodic elements ([30]–[32]).

2. H tone assignment rules

As mentioned above, Ns are traditionally classified into declensions (D), whose number may be different according to the criteria taken into account by the authors. This paper follows the D system of [8], [9], which relies on gender, plural formation and case marking. Somali has four cases that may influence the pitch patterns of the Ns, Adjs and Dets: (i) the vocative (VOC), (ii) the nominative (NOM), (iii) the genitive (GEN), and (iv) the absolutive (ABS), which is the ‘default’ case and corresponds to words that are not in the other three cases. One of the particular features of Somali is that NOM and GEN case markers, both tonal and segmental, always appear on the last element of a DetP, whatever its category. For instance, in the NOM DetP, wasir-a-arrimá-h-a dibád-d-u (‘the foreign minister’ lit. ‘minister-the-ABS matters-the-ABS abroad-the-NOM’), the NOM marker -a is suffixed to the last Det -d- of the DetP, while the first two DETs receive the ABS -a suffix (here, the GEN case is indicated by position only). In addition, for some Ds, the ABS case has to be distinguished from the ‘premodificer’ form (PREM) as far as the pitch patterns are concerned (see section 3); this form is used when a N is not the last element of a DetP.

Let us begin with the singular Ns of the first five Ds. In the ABS case, there is a tonal contrast between D1 feminine and D2-5 masculine Ns. When the Ns have at least two moras, fem. Ns have a final high pitch, while masc. Ns exhibit a penultimate one. This is illustrated by the following minimal pairs: inan ‘girl.D1.fsg’ vs. inan ‘boy.D2.msg’, beér ‘garden.D1.fsg’ vs. beér ‘river.D2.msg’. In [1] and [6] analyses, a final or penultimate accent is assigned by rule according to gender and D; in a tonal approach, the accent is to be simply replaced by a high (H) tone.

Some authors cast doubt on the fact that gender governs the location of the H tone. D1 fem. Ns receive a NOM -i suffix and a GEN -V1Vd morpheme, whereas masc. Ns do not. This led [25], [29], [33] to suggest that fem. Ns are suffixed by an empty [CV] morpheme to which a H tone is assigned. In the NOM case, all Ns may undergo gender polarity in plural, whereby a fig. N becomes masculine in plural and vice versa). [25], [29] provided strong evidence that the underlying form of the plural suffix is the same for all four of these Ds. This assumption is supported by the fact that D1-4 plural Ns exhibit similar pitch patterns, which contrast with those of the singular Ns in precise contexts ([3], [10], [12]). As shown in Figure 1, when a N is a contrastive focus (FOC) ([10], [12]) or is in the PREM context, a high pitch shows up; in D1-4 pl. Ns, this high pitch always takes place on the pl. suffix. In the NOM case, D1-4 Ns, singular and plural, are realized with a declining low pitch up to the end of the word ([10], [34]). On the other hand, ABS D1-4 plural Ns are characterized by a mid pitch up to end of the word ([3], [10]), while the D1-4 singular Ns present a H tone as seen in section 2.

Figure 1: time-normalized and averaged F0 curves of singular and plural D1-4 Ns (data from [10])

To capture these contrasts, I propose that the pl. suffixes are marked by a latent H tone, which remains unassociated in the ABS case. Following the general rules set down in section 2,
this latent H tone deletes the preceding H(s) in the word. As pl. Ns have no associated H, this gives an unspecified/mid H0 at the acoustic level (see [10] for details). In the FOC/PREM contexts, following [10], I propose that the latent H tone associates with the plural morpheme, which triggers the observed pitch rises. The H tone associates in this context in order to instantiate prosodic dominance relationships, expressed by pitch prominence and downstep, a phenomenon described by [1], [8], [10], [11], [34], whereby a H tone is lowered after another H tone within a given domain. [10]–[12] have shown that FOC in Somali, as in other languages, is characterized by higher pitch at the sentence level and followed by downstepped H tones. In the PREM context, prosodic dominance relationships reflect syntactic ones that take place between a head N and subsequent Det, Adj(s) or N(s). Accordingly, when there is no subsequent element to dominate and hence to downstep, i.e. when D1–4 pl. Ns are alone in a DetP and not focalized, no H is to be associated and realized.

The alternating analysis, namely, that the pl. morphemes would be toneless and a H tone inserted under FOC/PREM conditions, would not explain why no tone emerges on the root in the ABS case, while toneless suffixes like the INF -i maintain H tone on the root. Furthermore, some words never display any H tone, even in the FOC and PREM contexts, such as the ‘clitic’ pronouns (e.g. aan ‘I’) and definite Det +/i/-a/ ‘the’. It can be therefore argued that this kind of word has no underlying H tone, whereas a H tone must be present in D1–4 pl. Ns in a latent form. A latent H tone is likely to characterize other words like Adj(s) ended by –oon; further investigations should be carried out to determine to what extent latent H tones are present in Somali morphology. In any case, latent tones interacting with other tone features constitute the first evidence of a tonal approach for Somali; the NOM case provides further support for such an approach.

In light of the acoustic data, I propose that the NOM case is marked by a morphemic low tone, LNOM, that appears at the end of the last element of a DetP. LNOM then de-associates the left adjacent H tone in the DetP, generating the low and declining F0 curves displayed in Figure 1. The appearance of the suffix -i in D1 Ns, Adj(s) and Dets supports the morphemic status of LNOM, namely, LNOM and -i are likely to be both the phonological exponents of a NOM morphological feature anchoring to the end of DetP. Again, this proposal puts forward the idea that a NOM pitch pattern results from an interaction between tone features only. An accentual approach is unable to capture the pitch contrasts observed in D1–4 Ns. Further support for a tonal analysis is provided by the tonal alternations of D6/7 Ns.

4. D6/7 tonal shifts

D6/7 Ns are characterized by particular morphological features. While Somali sg. Ns usually end with a consonant (or an epenthetic -i, CC# being illegal, e.g. ábit ‘uncle:D2.msg’), sg. D6/7 Ns are suffixed by -o (hooyo ‘mother:D6.fsg’) or -e (bara ‘teacher:D7.msg’). In plural, specific suffixes are added: -aýin in D6 (hooyo ãýin ‘mothers.mpl’) and -aýall in D7 (barayaal ‘teachers.fpl’); N.B. any -e/o suffix changes into -a/o before C). More importantly for our concern, as displayed in Figure 2, sg. D6/7 Ns present specific pitch patterns, and all authors agree with their description, with the exception of some details ([1], [3], [6], [8]–[10], [34], [35]): (i) in the NOM, they exhibit a penultimate H tone and not a low pitch as in other Ds, and (ii) they present a penultimate or final H in the ABS, but an invariant final H in the PREM form, whereas the location of H in the other Ds is invariant in both ABS and PREM. How can we account for these tonal patterns? [1], [3], [5]–[9] resolve this issue by introducing specific accentual patterns for D6/7 Ns and/or accent shifting rules (e.g. *°–°–°*). Although this achieves descriptive adequacy, it does not help us to determine the causes of D6/7 tonal alternations.

![Figure 2: time-normalized and averaged F0 curves of singular D7 Ns (data from [10]).](image)

I will assume that D6/7 Ns actually have two underlying H tones: one associated with the stem, the other with the suffix -e/o, e.g. [hooyo-ød], [baře-ød]. But, for reasons I will discuss later, the H of -e/o does not delete the preceding H(s), as does the H of the plural suffixes, for instance. In the NOM case, LNOM thus delinks the first H leftwards, i.e. the H of -e/o, as it does in the other declensions, but the H of the stem remains, deriving [hooyo-ød] and [baře-ød]. In this way, the NOM phonetic HL pattern is not only generated, but the NOM case is also reduced to a unique and hence more understandable process – ‘LNOM delinks the first H leftwards’ – applying to all NOM Ns.

Furthermore, this analysis implies that if the D6/7 NOM penultimate H tone results from the association of a L tone, then the penultimate H appearing in other contexts might also be due to an observable L. This is just what happens in certain ABS contexts. As shown by [10]–[12], contrastive focus in Somali is usually followed – as in many other languages – by a L tone that reaches the bottom of speaker pitch range; after [10]–[12], I will name it LFOC and consider it as an intonation morpheme. LFOC may anchor to the -e/o suffix and then delinks the left-adjacent H tone, as does LNOM. In isolation, D6/7 Ns are reported to exhibit either a final or a penultimate H. I presume that, in the latter case, LFOC is chosen by the speaker to express some emphasis and is associated with the -e/o suffix.

In sum, the ‘two-H hypothesis’, instead of, say, an accent shift, explains the D6/7 idiosyncratic pitch patterns. It has revealed that the NOM case is always marked by LNOM, which triggers the deletion of preceding Hs and has helped us to identify LFOC; the penultimate H thus results from a unique factor, a morphemic L tone, LNOM or LFOC. Three questions nonetheless need to be answered: (i) when they are contrastive focuses, D1 sg. Ns and pl. Ns have a final H ([10], [12]); why is this final H not delinked by LFOC? (ii) As stated above, the H of -e/o underlyingly does not dissociate the preceding H(s), while other suffixed Hs do; why? And (iii) how is the final phonetic H tone derived in the PREM or ABS forms? The answer to the first question is straightforward. We have seen in section 3 that the latent H tone in plural Ns shows up to express prosodic prominence and to trigger downstep. If LFOC deleted the final Hs in D1, dominance relationships could not be established. In other words, a prosodic head H tone cannot be delinked by any other tone. To answer to questions (ii) and (iii), we need to address the internal morphological structure of words in Somali and to determine how the tones surface.
5. Structure of words and phase theory

Following [25], I propose that Ns have the following syntactic structure: [[[Root] n]ₙ-Th]ₙ Num]ₙw. ‘Little n’ is a nominalizing functional head and Num the functional head for number. As for Th, it is a thematic node adjoined to the functional head n; it is realized as the theme vowel in Latin, for instance ([21]). As seen in section 2, tonal information is part of vocabulary entries; roots have a H tone, while functional nodes may be assigned a (latent) H tone or not. For instance, the D₂ pl. N qoraall ‘writings’ (qor ‘write’ + nominalizing -aad ‘-ing’ + -C₂ ‘-’ indicates the latent H) is structured as follows: [[[qo]Root -arah]ₙ-Th]ₙ Num]ₙw. The plural form is obtained by stipulating that the H tone of a functional head deletes the preceding Hs; in singular, this gives the expected form qorráal ‘writings, texts’.

As for D₆/₇ Ns, I propose that -e/o suffixes are neither derivational nor inflectional suffixes, but the expression of the Th node. -e is often described as an agentic/instrumental suffix (e.g. bár ‘to teach’ +é → bare ‘teacher’). But, when suffixed to an N-/ stem, it creates Ns meaning, ‘owner of N’ (e.g. dukanl ‘shop owner’); moreover, it is also suffixed to many Ns without verbal and independent roots (e.g. aabbe ‘father’, tse ‘table’, txc ‘gossip’). Likewise, the -o suffix is abundantly found with any kind of stem. It appears at the end of Ns without verbal and independent roots (byé ‘water’, hooy ‘mother’); it is suffixed to nominalizing morphemes such as -nim- ‘the state of N’sqon ‘food’ D₂/msg. → doqonnimó ‘foolishliness.D₆.fsg.’, -toóy- ‘N-do m/hood’ boqor ‘king’ D₂/msg. → boqortooyoóyin ‘kingdom.D₆.fsg.’); finally, borrowed words ending with -o/a in the original language, usually Italian, regularly belong to D₆ (basoó ‘pasta’, Maayó ‘may’). In other words, it can be argued that -e/o suffixes have no meaning per se and seem to convey gender for Ns only; agentic meaning is to be interpreted as a mere byproduct induced by the concatenation of a verbal root with -e. I therefore propose the following representations for D₆/₇ Ns: for instance, D₇ msg. N bare = [[[bår]Root -arah]ₙ-Th]ₙ Num]ₙw and D₆ fsg. doqonnimó = [[[doqon]Root -nim-]ₙ-Th]ₙ Num]ₙw. Not being associated with a functional head, the H of -e/o cannot delete the preceding tones; the H of the root or the nominalizing suffix thus remains at this stage of the derivation. In plural, however, -oï/d/ïn are inserted in the Num head position and delete preceding tones as expected, deriving harvayád ‘teachers’, boqortooyoóyin ‘kingdoms’.

In order to derive the D₆/₇ phonetic final H tone, I will adopt recent proposals by [22], [30], [31], [36], [37], according to which the syntactic derivation is ‘chunked’ into phases. The phrase vP, headed by ‘little v’, and the phrase CP, headed by the C, correspond to phases. To this, I will add the phrase DetP after [31]. Within a phase, lexical material is inserted and constituents may move up to higher phase-internal syntactic positions. At the end of a phase, the complement of the phrasal head is spelled out, so it is this category that is given a phonological form.

In the spirit of [24], I will assume that NumP merges with Det to create DetP. As NumP is the complement of Det, it is spelled out in DetP phase. It is during this phase that the H of the functional heads n and Num deletes the preceding Hs; the H of the adjoined Th node, however, does not and is retained together with the penultimate H tone in subsequent phases (vP and CP). In the latter, LNUM and LLOC are inserted at the end of a DetP and may delink the DetP final H tone (i.e. Hₙ₂ or any other H tone). This means that the D₆/7 ABS or PREM final H tone is to be generated at the end of the syntactic derivation. Accordingly, I propose that (i) the final D₆/7 H shows up within the very last spellout domain, (ii) it derives from an OCP-like constraint prohibiting two Hs in the same spellout domain in a DetP phase, and (iii) the final H, which marks the right edge of a DetP spellout domain, is maintained. In addition, within this last spellout domain, (i) ‘phonetic’ L tones are inserted; in particular, L tones are inserted just before the ‘tonal-accent’ H tones to produce pitch rises (see [10]); (ii) some H tones move rightwards to conform to the well-established constraint that prohibits a H tone from appearing before the penultimate mora in a word (cf. keéni → keéni ‘to bring’), and (iii) H tones may spread rightwards to generate boundary ‘H%as’ at the end of the pre-focus phrases as described by [10].

In addition to this, phase theory might explain the tone patterns of DetPs and IP. The fact that NumP is spelled out before Det explains why the latter is tonally independent from the preceding N (Det is suffixed to N in Somalı). For instance, the demonstrative -dán in the plural DetP qoraalládán ‘these texts’ has its own H tone, which does not delete the preceding tones. According to [38], [39], CIP is syntactically headed by one of the focus particles (FP) wáa, aváa or báa; the FP is followed by subject and object pronouns and verbal ‘adpositions’ (ADP) – i.e. ‘pre-verbal morphemes which semantically govern oblique elements of the predication’ cf. [8, p. 94]; DetPs are excluded from IP and are realized in adjoined positions to IP or CP, where they constitute independent prosodic and casual domains; this is illustrated in (1). v-gúrígurí-g[á one of the vP you us-from bring-2.Pra.
‘You bring it for us from the house.’

SUBJ (aad 2sg.) and OBJ (na 1pl.) pronouns are toneless (cf. section 3), while FP and ADPs (here ‘for’ and ‘at’ from) are lexically assigned a H tone. Remember that vP is a phase; within this phase, the V and SUBJ clitics move up to higher nodes in IP ([38], [39]), while the OBJ pronoun and the ADPs, as complements of V, are spelled out together. That is why the SUBJ pronoun, the sequence noogá (< naa-ï-ká) and the V constitute autonomous prosodic domains. The underlying na-ï-ká then undergoes the same late H tone deletion rule as D₆/₇ Ns – and other segmental rules –, giving noogá; the SUBJ pronoun surfaces without tone and merges with the intoned FP (wáa+aad → wáaid).

To conclude this paper, it is worth considering the tone patterns of the Vs. In section 2, I have assumed that V roots and tense suffixes are assigned a H tone (cf. keéni! ‘bring (it)’!; keen-daan SUBJ.3pl.). Then why is the finite V keentaa toneless in (1)? I will argue that Vs are actually submitted to the same tone rules and derivation as Ns. That is, keentaa derives from keéni-t-aa; as -aa is a head in IP (TP), it deletes the H of the root keéni. Tense suffixes usually bear a H tone except when the subject DetP is marked by LNUM in the NOM case. In Somali, a subject focused by an FP is not in the NOM case, but in the ABS one, and hence, exhibits a H tone; so does the V when it agrees with a focused subject ([1], [3], [7]–[10], [38], [39]). When the V finishes a relative clause at the end of a DetP, it exhibits a H tone if the DetP is in the ABS case, but it does not if the DetP is in the NOM one, and hence, it is suffixed by LNUM. In view of these facts, I therefore propose that toneless Vs are due to the suffixation of the LNUM, which deletes the last H tone of the V – i.e. the tense suffix H – as it does for Ns. Acoustic data from [10] confirm the occurrence of a L tone at the end of V. This proposal admittedly implies that LNUM is a special ‘case marker’; it confirms, however, that the complex pitch alternations of Somali are, in fact, governed by a small number of general principles that are only based on tonal features.

Acknowledgements

This work was supported by the “Région Haute-Normandie” (GRR CSN Projet “PPM” – convention n°15B07705GR).
6. References

[27] X. Barillot, “Morphophonologie gbaritique et information consonantique latente en somali et dans les langues est-