On the abstractness of the phonological component in a piece-based morphological theory

Nicola Lampitelli

Université Paris 7

Draft, December 2009

1. Introduction

A central issue in piece-based morphological theories (Halle & Marantz 1993, Embick & Halle 2005, Embick & Noyer 2007, Marantz 1997, 2001, 2007) is the realizational process which associates a form to morphemes. In such theories, the morphemes are features bundles which receive phonological content in the post-syntactic component of the Grammar, i.e. PF. This process is called Vocabulary Insertion and the phonological exponents which are inserted at each terminal are called Vocabulary Items (VI).

Crucially, for a given feature more than one VI may in principle match one of its values with the ones in the bundle. In such a case, competition occurs between these VI’s and the most specified one wins, following the Subset Principle (Halle 1997). Note that the only possible competition in these terms is at node level, and only at this point VI’s are listed and checked at PF (cf. Embick & Marantz 2008). A classic example is English nominal plural morpheme, which has at least three allomorphs:

(1) English plural allomorphs

a. [+pl] ↔ Ø / √MOUSE; √FOOT, etc.. (closed list)

b. [+pl] ↔ /ən/ / √OX; √CHILD, etc.. (closed list)

c. [+pl] ↔ /s/ / elsewhere (open list, plus all loans)

The allomorph [s] being the less specified, it is inserted only and only if a given root does not appear in either list (1.a) or (1.b). These are cases of contextual allomorphy.

Another crucial aspect of Vocabulary Insertion device is that this occurs post-syntactically, as by standard Late Insertion hypothesis. Consider that the allomorph (1.c) wins the competition because, say, the root √DOG is inserted. The sequence /dɔɡ + s/ must surface as [dɔgz]. Now, these theories claim that the phonological component has rules which can account for such phonological conditioned allomorphy.

The fundamental architecture of this realizational device is that it splits standard morphophonology in two components of the Grammar: one syntactic and the other post-syntactic. A question arises once one considers such an architecture: what is the exact border between phonological insertion syntactically (and locally) conditioned and phonological insertion phonologically conditioned? In other words, why do we postulate that the realization [z] of the VI [+pl] ↔ /s/ is a matter of phonological rules applying at PF whereas the insertion of either /s/ or /ən/ allomorphs is a matter of competition at syntactic node level?

In principle, one could argue that the border is unclear and thus movable, either up or down, i.e. the competition is only syntactic (no ‘standard’ phonology at all) or there is no competition anymore as the syntax sends its features bundle at PF without ‘exceptions’ (the phonology computes the representations it receives from syntax and creates actual surface forms). This paper explores the latter hypothesis.1

---

1 As for the first hypothesis, it turns out to be similar to Ferrari’s (2005): a model where the syntax does the entire job in word formation.
The hypothesis this paper concentrates on concerns the idea that the whole phonological computation is made at PF where each terminal is spelled-out without undergoing the kind of competition shown in (1). If so, a one-to-one correspondence between a given feature and a VI must be postulated, i.e. in the case of English plural morpheme, the unique VI would be [+pl] ↔ /s/ which is the less specified entry.

Consequently, phonological rules apply in order to give the actual surface form corresponding to a given bunch of structure. Consider again the English plural morpheme: once the VI /s/ is inserted, phonological rules will compute transformations into surface forms, i.e. [z] in the context of √DOG. Nevertheless, how can we account for cases such VI’s (1.a) and (1.b)? These are clearly not phonologically related to underlying /s/. Should the Grammar incorporate a phonological rule applying to /s/ in the context of √OX and transforming this segment into /sn/?

Two answers can be provided. The former, which is theory-external, calls for a different phonological approach, a theory whose scope may be more global and hence the generative power superior. Such a framework could arguably be Optimality Theory (Prince & Smolensky 1993), where the mapping of input /s/ into output [sn] is certainly possible, given some particular markedness constraints (ranked above faithfulness constraints).

However, I do not believe this is the correct path. First of all, recent OT implementations have restricted the global power of standard theory and thus the mapping of /s/ into [sn] is not possible anymore. Neither OT-Candidate Chains (OT-CC) (McCarthy 2006a, 2006b) nor Harmonic Serialism (HS) (McCarthy 2008) can generate such an output, because of some local and serial restrictions on both markedness and faithfulness constraints. Secondly, this solution is theory-external and I do not argue for a new theoretical framework nor for a considerably new theory of grammar architecture. This approach would correspond to a constraint-based model where lexical insertion occurs in the phonological component as in Wolf’s theory (2008, To appear). Wolf assumes that word formation is computed by ranked constraints, as OT does, but he argues that allomorphy selection is made in the phonology.

Back to our /əs/ → /n/ derivation. The second possible answer, which is theory-internal, must include reference to allomorphy. One could argue that the form of the morpheme [+pl] is [əs] in the context of √OX. Such an approach is discussed by Embick’s (2010) localist theory of allomorphy. In these terms, allomorphy is triggered only locally, i.e. when two morphemes are syntactically adjacent. In the case in handle, the unique spell-out for [+pl] feature is /s/. Once phonology tries to compute √OX and /s/ into the surface form *[oksəs], the allomorphy rule [+pl] → [ən]/√OX intervene. Crucially, my hypothesis is that allomorphy rules apply before phonological rules but both are in the PF component of the grammar.

I illustrate this model below:

(2) Abstract Phonology after Syntax

Syntactic derivations

<table>
<thead>
<tr>
<th>Spell-out of terminals (one-to-one correspondence)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF</td>
</tr>
<tr>
<td>a. allomorphy rules (AR)</td>
</tr>
<tr>
<td>b. phonological rules (PR)</td>
</tr>
</tbody>
</table>

Note that [ə] can be claimed to be epenthetic; but still, it /s/ → /n/ /√OX conceivable on strict phonological terms?
Before continuing the discussion, let me clarify that this hypothesis makes a prediction on the complexity of the phonological component. The fact of having a one-to-one correspondence between a given terminal node and a VI entails that the form of VI’s must be more “abstract”, i.e. phonologically more distant from surface forms. At the meantime, the phonological component has more job to do, as far as this model is compared to standard Distributed Morphology (henceforth DM) assumptions.

In this sense, roots and morphemes become pure abstract entities which do not have irregularities (or which are not sensitive to irregularities). The spell-out procedure provides phonological content to each of them and later it computes all sort of rules to get the actual surface form.3

Standard work in DM such as Halle & Vaux (1998) has already looked at morphemes on a more abstract phonological ground, for example decomposing pl. Dat and Abl ending -bus into two morphemes (-bu- and -s) to get more generalizations in describing Latin declension. This paper explores the same path arguing for a more radical distinction of the labor between the syntax and phonology.

Going back to English nominal plural morphemes, I propose the following architecture to account for its allomorphs:

(3) **English noun structure at PF**

```
   #     #     #
  n   #   n   s  
  √  CAT n  s  
  #   #
  /kæt+s/

 AR ---
 PR1 --- /dɔɡz/
 PR2 ---
    [kæts]  [dɔɡz]  [ɔksən]
```

The details of the structures and of how heads are created at PF are discussed throughout three case-studies. The crucial idea is that in all cases, a single root and a single plural morpheme are inserted. Surface divergences are a matter of either AR’s or PR’s which apply post-syntactically. These rules have the same status in the grammar and they are ordered, as in standard rule-based systems.

This hypothesis clearly does not have the pretension to be universal, in the sense of current generative theory. Instead, this paper crucially proposes an interpretation of a domain where the background theory, namely DM, is less clear. As noticed in the introduction, this site is at the border between the syntactic domain and the phonological component. The revised model shown in (2) and (3) is defended throughout the following sections. I first discuss what kind of phonological approach I assume, then three case-studies are presented to support the general idea of the paper. These involve Italian, Bosnian and Somali, three unrelated languages.

---

3 Although I do not want to claim that this hypothesis has a psychological reality, it is interesting to note what Stockall & Marantz (2006) have discovered on a neurological ground about root abstractness: even in the case of irregular forms such as the pair teach - taught speakers access to a √TEACH entity. Cf. Boudelaa & Marslen-Wilson (2004) for evidence of such a claim based on Arabic data.
3. What phonology?

Assuming the model depicted in (2), rises the question whether any phonological approach can be in fact used within this theory. So far I have been using a general rule-based approach, in claiming that PR take the form of “rules” in the sense of Chomsky & Halle (1968) (SPE). For instance, PR1 and PR2 in (3) are rules of the form \( /x/ \rightarrow [y] /___\).

As far as I understand it, DM assumes that the kind of phonology applying after the syntactic derivation is exactly rule-based. This is quite clear in works such as Halle & Vaux (1998) where a set of readjustment rules applying late in the derivation account for surface form of both Latin and Armenian declension-number-case endings. Calabrese (1998) follows the same analytical path to account for Italian nominal endings. Or, Halle & Nevins (2008) assume a syntactic approach to word formation while discussing rule-based phonological approach to Slavic declensions.

Two major alternatives come to my mind: a constraint-based approach (OT), which has been briefly mentioned earlier and the Government Phonology (Kaye et al. 1990, Kaye 1995) (GP).

As for the former, classic OT could clearly solve the problem of the mapping \( /n/ \rightarrow [s] \), as already mentioned. But again, this theory has a too much power in mapping input onto possible outputs, as McCarthy (2006a) acknowledges. In these terms, one could argue that [kæt] is the output for the input /dɔg/. Clearly, this seems too powerful and OT-internal debate has noticed that, too. New trends in OT, such as OT-CC and HS cannot be helpful, neither. The former crucially proposes that in the Candidate-Chains, a PRIORITY-constraint is active, and it can see the whole derivation, obliging two (or more) constraints to be ordered, in a fashion that reminds ordered rules. The latter instead restricts even more the global power of classic OT and proposes that the derivation must be serial (like in rule-based models) and harmonic (unlike rule-based phonology): the second property of derivations is the central idea of HS.

However, OT-type approaches treat the abstractness of underlying form as pure accidents, as everything depends on how many violation marks each candidate falls to show. We do need to constrain an abstract representation and we also need to constrain the kind of prediction our theory does. None of the most recent approaches within the OT paradigm is able to do so.4

Ségéral & Scheer (2001) address exactly the question as to how abstract is phonology and its representation. Their answer lays on the assumption that autosegmental representations allow for an easy explanation to otherwise opaque phenomena. They argue for the existence of virtual geminates in German and Somali. This leads us on the latter alternative on the nature of phonological approach. The general program is referred to as Autosegmental Phonology (Goldsmith 1978) but it has been elaborated and refined into the more recent GP approach.

Work within GP has shown that independent principles on the architecture of template make strong predictions on surface forms, the consequence being a more abstract level of representation. For instance, Kaye (1990) demonstrates that central high vowel [i] in Moroccan Arabic shows a totally predictable behavior if GP is assumed; Lowenstamm (1991) argues for phonologically long vowels in Semitic even if these are short at surface. The most constrained hypothesis within GP paradigm has been proposed by Lowenstamm (1996) who argues a strict CVCV.. templatic tier in the representation of each word.

Along these lines, Lowenstamm (2008) designed a model which is in the tradition of GP but which assumes a syntactic approach to word formation. The crucial idea is that the CVCV.. tier may (possibly) be associated to each syntactic terminal. Only items which have either a C or a V to anchor on can surface. The prediction which is made deals with the nature of spell-out. In standard DM, VI’s are phonological exponent of abstract morphemes (feature matrices). These VI’s are segments in SPE sense. On other hand, in Lowenstamm’s (2008) terms VI’s are autosegments.

4 The following question seems to me perfectly legitimate: do we really need a completely different model (global and parallel with some local and serial restrictions) applying only to the phonology?
which can be associated to a CVCV.. template. Parametric variation lays on the nature of spell-out, which has different natures, as the case-studies show.\(^5\)

Spell-out will be dealt with in deeper detail in section 6 after discussing the first two case-studies, which are the focus of the following section.

4. Case studies: Italian and Bosnian

In this section I look at Italian and Bosnian nouns.\(^6\) Both languages show interesting phenomena occurring at the last vocalic position on nouns. In particular, I deal with the concept of “final vowel” and I argue for decomposition of such a vowel in primitive Elements in the sense of Kaye et al. (1985). The phonological representation that follows, allows for more generalizations and thus a more abstract architecture of the post-syntactic component. As far as syntactic structure is involved, this approach points to an economy of VI’s.

4.1 Italian nouns

Italian nouns are an interesting challenge for morphological theory because of two particular aspects: (i) the vocalic alternation between singular and plural and (ii) the presence of clearly different vocalic patterns relying singular to plural (o-i, a-e, e-i, etc.). Both phenomena contrast with the general behavior of Romance Languages where plural marker is generally consonantal (/s/ in Spanish, Portuguese, Catalan, etc.) and no prediction can be made on the form of final syllable on nouns.

Observations (i) and (ii) deal with a well-known feature of Italian nouns: they must all end in a vowel, loanwords being generally the only exceptions to this pattern.\(^7\) It is important to note that this final vowel on nouns is generally unstressed. Whenever the final vowel is stressed then (i) all the five vowels in the phonological system of the language are allowed to be in that position, and (ii) the noun is invariable in number marking (the same vowel in sg. and pl.). Put in a different way: oxyton nouns are invariable in Italian:

\[(4) \text{Oxyton nouns (invariable)}\]

\[
\begin{align*}
a. [\text{čit.ta}] & \quad \text{‘town’} & b. [\text{kaf.fe}] & \quad \text{‘coffee’} \\
c. [\text{sup.pli}] & \quad \text{‘fried rice’} & d. [\text{o.blo}] & \quad \text{‘porothle’} \\
e. [\text{vir.tu}] & \quad \text{‘virtue’} \\
\end{align*}
\]

(The stressed vowel is underlined)

On the other hand, if the final vowel isn’t stressed, then (i) the set of possible vowels appearing in this position is restricted, and (ii) the noun is never invariable as far as number is concerned.\(^8\)

---

\(^5\) Faust & Lampitelli (2009) is an attempt in deriving the difference between Hebrew non-concatenative morphology and Italian concatenative one from a difference on root and functional morpheme spell-outs.

\(^6\) As Browne 1993 and Corbett & Browne 2008 point out, Serbo-Croatian is one single literary language which has three major dialect groups: Čakavian, Kajkavian and Štokavian. The disaggregation of Yugoslavia and the political issues related to this fact, made each independent Republic make distinctions within the standard language. As my informant comes from Bosnia-Herzegovina, I will refer to Serbo-Croatian as simply Bosnian and my data will be mostly based on Stokavian dialect.

\(^7\) Some examples of loanwords are: film ‘movie’, sponsor ‘sponsor’, pullman ‘bus’, computer ‘computer’, etc..

\(^8\) There are few exceptions to this general rule, but the majority of this kind of words are Greek loans, such as analisi (sg. and pl.) ‘analysis’, crisi (sg. and pl.) ‘crisis’ and so on, cf. Ferrari (2005) for a complete survey.
(5) Italian variable nouns

<table>
<thead>
<tr>
<th>(A)</th>
<th>sg. gloss</th>
<th>gender</th>
<th>(B)</th>
<th>pl. gloss</th>
<th>gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>lupo</td>
<td>‘wolf’</td>
<td>M</td>
<td>lupi</td>
<td>‘wolves’</td>
</tr>
<tr>
<td>b.</td>
<td>rosa</td>
<td>‘rose’</td>
<td>F</td>
<td>rosé</td>
<td>‘roses’</td>
</tr>
<tr>
<td>c.</td>
<td>poeta</td>
<td>‘poet’</td>
<td>M</td>
<td>poeti</td>
<td>‘poets’</td>
</tr>
<tr>
<td>d.</td>
<td>ala</td>
<td>‘wing’</td>
<td>F</td>
<td>ali</td>
<td>‘wings’</td>
</tr>
<tr>
<td>e.</td>
<td>cane</td>
<td>‘dog’</td>
<td>M</td>
<td>cani</td>
<td>‘dogs’</td>
</tr>
<tr>
<td>f.</td>
<td>nave</td>
<td>‘ship’</td>
<td>F</td>
<td>navé</td>
<td>‘ships’</td>
</tr>
</tbody>
</table>

These data illustrate that only four alternating vowels appear in final position (Vfin): [a], [e], [i] and [o]. We could easily argue that Vfin’s enforce some phonological requirements on final syllables, as notice above: normally nouns do not end in a consonant. However, this seems to be false, as three different vowels appear in (5.A) and if it was only epenthesis to enforce its presence, we should logically expect only one and the same vowel throughout the paradigm.9

In DM terms, a possible and logical solution to this puzzle is a list of VI’s, which compete for insertion: for each morpheme, the more specified item wins. The following entries can be proposed:

(6) VI’s for Italian nouns

(A) singular markings
   a. [+sg] ↔ a / √class 2 F and √POET, etc.. (list)
   b. [+sg] ↔ o / √class 1 M
   c. [+sg] ↔ e / elsewhere

(B) plural markings
   a. [+pl] ↔ e / √class 2 F
   b. [+pl] ↔ i / elsewhere

These VI’s are certainly possible spell-outs for Vfin’s in Italian. However, there must be a reason as to why three vowels appear at sg. whereas only two at pl. (not the same, by the way), or why [u] is banned from appearing in that position in the language. What do these vowels exactly reveal us? Just the fact that they can vary depending on gender and number?

It turns out that the Theory of Elements (Kaye et al. 1985) (henceforth TE) makes interesting predictions on the internal composition of vowels: each surface vocalic segment is in fact the result of a complex fusion between basic Elements, namely /A/, /I/ and /U/. In TE terms, the five vowels appearing in unstressed position in Italian are defined as follows:

(7) Decomposed vowels
   a. [a] = /A/
   b. [i] = /I/
   c. [u] = /U/
   d. [e] = /I.A/
   e. [o] = /U.A/

Now, if we replace each Vfin in (5) by the corresponding algorithm in (7), we get the following results:

9 Consider also that some final-hiatus words exist: [ma.re.a] ‘tide’; [nu.cle.o] ‘nucleus’; [bo.a] ‘float’; etc.. In other words, the Italian case is not as Spanish one, as described in Harris (1995).
Two facts are strikingly evident: (i) the element /A/ appears in all occurrences at sg. whereas (ii) the element /I/ appears in both occurrences of Vfin at pl. Clearly, TE gives us the advantage of a deeper understanding precisely because further generalizations can be made on the distribution of Vfin’s themselves. In the case at hand, it is tempting to associate the element /A/ with the sg. marking and the element /I/ with the pl. one. If so, VI’s in (6) do not have any reasons to hold anymore, because as far as number is involved, the degree of predictability is now established.

What about the second element appearing in other occurrences of Vfin in (8)? In Italian, we can predict the gender of a noun having the pattern a-e (cf. 5.b): it is always F, whereas a noun having the pattern o-i (cf. 5.a) is almost always M (I only know an exception: mano - mani ‘hand(s)’, F).

On the other hand, if a noun has the well attested pattern e-i it can be either M or F, without possible and logical prediction. Note also that the adjectives follow the same two patterns: one group distinguishes the gender by the opposition -o.M vs. -a.F, while the second one does not have such a distinction, as all its items end in -e. Hence -e is unspecified for gender. This parallelism allows us to consider that nouns too are divided into two major groups (or class):

(9) Groups of nouns

a. group 1  
lupo ‘wolf’ M  rosa ‘rose’ F
b. group 2  
cane ‘dog’ M  nave ‘boat’ F

The division in two groups as in (9) constitutes a way of tracking what ingredients exactly a speaker needs to build a well-formed noun in Italian. Recall the decomposed Vfin’s in (8) and the two strong generalizations on number marking. Now, knowing that a noun belongs to a given group is the only piece of information we need in order to correctly predict the shape of the final vowel. Let us take √LUP, we know it belongs to group 1 (or class 1). This means that once this roots merges with a n-head bearing a [-F] gender feature, the predicted spell-out is the element U. This combines with sg. A and the result is the correct surface [o]. Crucially, the presence of a lexical element (U, A or I) determines the surface form in combination of number marking elements. I call such a lexical vocalic element Root Element (RE).

Two cases are still unaccounted for. These are the examples in (5.c) and (5.d) all displaying the pattern a-i. Only two F nouns exist in this group, arma ‘weapon’ and ala ‘wing’. I claim these nouns do not display any RE, they’re bare roots which merge to n-head: thus their Vfin take the shape of number marking vowels, [a] and [i].

For sake of expositing clarity, I illustrate the whole underlying structure of the nouns in (6) in what follows:

(10) Underlying structures of Italian nouns

<table>
<thead>
<tr>
<th>noun</th>
<th>root</th>
<th>RE</th>
<th>sg</th>
<th>root</th>
<th>RE</th>
<th>pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. lupo ‘wolf’</td>
<td>√LUP + U + A</td>
<td></td>
<td>√LUP + U + I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. rosa ‘rose’</td>
<td>√ROS + A + A</td>
<td></td>
<td>√ROS + A + I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. cane ‘dog’</td>
<td>√KAN + I + A</td>
<td></td>
<td>√KAN + I + I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. nave ‘boat’</td>
<td>√NAV + I + A</td>
<td></td>
<td>√NAV + I + I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. poeta ‘poet’</td>
<td>√POET + zero + A</td>
<td></td>
<td>√POET + zero + I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. ala ‘wing’</td>
<td>√AL + zero + A</td>
<td></td>
<td>√AL + zero + I</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The decomposition of Vfin’s has brought us along the lines of the evidence that the formation of a noun in Italian follows a clear predictable path. First, the root bearing its RE information merges with a n°, then with #°. This structure is sent to PF, where a complex head is created. At this stage of the derivation, a Th node is inserted to both n° and #° as a general requirement of the language. Spell-out occurs now and provides phonological content to each terminal of the complex head.

Recall the hypothesis the current paper investigates: the syntax produces “regular” structures with feature matrices and send all this material to PF, where the correct order of morphemes is established and where spell-out occurs. For each matrix of features at the terminal nodes of the complex head, one single VI corresponds, all other surface differences being computed later on. The general case is illustrated below:

(11) Syntactic structures and complex heads in Italian

a. numP
   num
   n
   CVCV..
   l u p

b. num
   n
   Th
   [-pl]
   CV
   U
   CVCV..
   l u p

In addition, observe the list of VI’s for the nominal paradigm:

(12) VI’s for Italian nouns

a. [-pl] ↔ A
b. [+pl] ↔ I
c. [-F] ↔ U
d. [-M] ↔ A

In the case of group 2 nouns, these neutralize the gender element into /I/. As by hypothesis, surface differences are accounted for by later processes, either phonological or contextually morphological. In this case, we can assume that an allomorphy rule applies to gender element in the context of a given list of roots (group 2), as illustrated below:

(13) Allomorphy rule

a. U, A → I /√[group 2]_

The rule can apply because the root is “seen” by the adjunct-to-n° Th (√/U/A) and then it can trigger the allomorphy. This situation mirrors what indeed happens in the case of agreement within the DP. The definite article, has only two genders and thus two morphological distinct forms (il and la ‘the.M-F’), not three or more as nouns seem to have.

What has just been said enhances a further prediction. If something holds between the root and the elements U or A, then the rule (13) should fail to apply. This seems actually the case when a
diminutive morpheme attached to the root. This kind of morpheme is quite productive in Italian and it takes several forms, -ino, -etto and -ello being the most used. Whatever form and whatever group of the base, the diminutivized nouns always displays either [o] for M or [a] for F. Consider the noun nave ‘boat’. A small boat, navina, shows regular F morphology, the form *navine being ungrammatical. The structure of a diminutivized noun is shown above:

(14) Diminutive structure

```

```

The context of application of rule (13) is the adjacency of a root of the type of √NAV and the RE. In the situation shown in (14) this clearly does not happen and thus the allomorphy rule fails to apply. The result is the spell-out of VI (12.d) without further computation.

Before turning to Bosnian, I want to discuss the status of CVCV.. tier within the theory presented in this paper. So far, it seems that the presence of such a phonological abstract level is justified only as a theoretical assumption, but no independent evidence as been proved. Consider loan consonant-final nouns in Italian, such as film ‘movie’, bar ‘café’, sponsor ‘sponsor’ etc.. Their final syllables are odd in Italian and crucially never found in native lexicon. Having a CVCV.. tier in the representation of each word means that nouns like film have a free V-slot at the end of their underlying template:

(15) Underlying CVCV.. template

```

Why none of these nouns accept any overt inflectional morphology marking? The fact of being loan maybe corresponds to a lack of RE, which is logically possible, but this does not entail the lack of overt number marking. In a closely related language, Spanish, plural /s/ always attaches to nouns, either origin they have (cf. bar vs. bares ‘café(s)’). But in Italian this is never the case. The puzzle becomes clearer if we consider that root template is unavailable to inflectional morphology (cyclicity of derivation? cf. Kaye 1995) and that a syllable CV is underlyingly attached to n°. The two elements fusion and are then associated to this syllable. This makes the difference between film-type nouns and poeta-type nouns: the former do not have any RE or CV either, whereas the latter crucially have a CV syllable in their n° and thus display overt number morphology. Hence, the status of CVCV.. tier is independently demonstrated. As a consequence, the choice of GP over strict rule-based approach takes its advantage.
As a partial conclusion, I claim that Italian nouns show that a model where a more abstract phonology is responsible for surface forms instead of arbitrary lists of VI’s is logically conceivable. In addition, the model of GP adopted to account for the phonological processes strongly suggests that whatever the output of syntax is, only material associated to either a C or a V is effectively pronounced at surface. The next section investigates the behaviour of Bosnian, which provides more evidence on this ground.

4.2 Bosnian nouns

This section provides an analysis of Bosnian nouns following the same path as for Italian ones. Crucially, Bosnian has three genders, neuter (NEU) being marked on determiners, too. In addition, Bosnian systems has an overtly marked case system, formed by five distinct forms (plus vocative, which will not be considered). Dative and locative are syncretic throughout the paradigm.

There are at least two different approaches in classifying the nouns: I use a gender-based distinction, as exemplified below (cf. Browne 1993, Corbett & Browne 2008):

(16) Bosnian nouns

<table>
<thead>
<tr>
<th></th>
<th>group 1</th>
<th>group 2</th>
<th>group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Masculine (M)</td>
<td>Feminine (F)</td>
<td>Neuter (NEU)</td>
</tr>
<tr>
<td></td>
<td>sg.</td>
<td>pl.</td>
<td>sg.</td>
</tr>
<tr>
<td>a. NOM</td>
<td>okvir</td>
<td>okvir-i</td>
<td>kuć-a</td>
</tr>
<tr>
<td>b. GEN</td>
<td>okvir-a</td>
<td>okvir-ā</td>
<td>kuć-ē</td>
</tr>
<tr>
<td>c. DAT/LOC</td>
<td>okvir-u</td>
<td>okvir-ima</td>
<td>kuć-i</td>
</tr>
<tr>
<td>d. ACC</td>
<td>okvir-(a)\textsuperscript{10}</td>
<td>okvir-e</td>
<td>kuć-u</td>
</tr>
<tr>
<td>e. VOC</td>
<td>okvir-e</td>
<td>okvir-i</td>
<td>kuć-o</td>
</tr>
<tr>
<td>f. INSTR</td>
<td>okvir-om</td>
<td>okvir-ima</td>
<td>kuć-ōm</td>
</tr>
</tbody>
</table>

However, an other possible way to distinguish declensions is by looking at the genitive singular form of a given noun. This results in three groups as in Hammond (2005:105-177) where masculine and neuter nouns having [a] as genitive singular fall in group 1; feminine (and some masculine) nouns having [e] instead fall in group 2 and finally feminine nouns having [i] fall in group 3. This classification, though, does not allow for a one-to-one correspondence between the gender and declensional paradigm and therefore it is not interesting for my proposal.\textsuperscript{11}

The two most striking facts are that (i) the only phonologically zero case is the M sg. NOM, and (ii) case markers are almost exclusively formed by vowels (hence Vfin, too). The only consonant allowed in this context is /m/ which appears in INSTR (sg. and pl.) and pl. DAT/LOC.\textsuperscript{12} It seems that Bosnian uses the vocalic alternation as the unique morphological device to express gender, number and syntactic case, exactly as Italian does (without showing case marking, though).

This interesting similarity is even more reinforced by a further observation on Vfin’s in both languages: an alternating inflectional vowel must be unstressed, otherwise no alternation is possible. As noticed above, Italian oxyton nouns are invariable in number, whereas Bosnian oxyton nouns (which are all loanwords and M) do show inflectional morphology but the stem is unaffected:

\textsuperscript{10} M nouns are marked by -a in sg. ACC only when the referent is animate.

\textsuperscript{11} Standard Serbo-Croatian has distinctive long-short vocal oppositions as well as a pitch accentuation as illustrated by Matešić (1970): there is a falling and a rising accent which can both be short or long. Then vowels vary according to length and pitch. In particular, falling short accent occurs only on the first syllable of a word, hence monosyllabic ones can have only a falling accent. But no accent can occur in the final vowel. In addition, Magner & Matejka (1971) demonstrated that this system is not so well preserved for many speakers which tend not to distinguish length on unstressed vowels (among other facts which are not relevant for the present purpose). For these reasons, I will not be concerned with such a phonological issue anymore and the length given in (1) has only illustrative purposes.

\textsuperscript{12} The case ending -ju exists, too.
(17) **Oxyton nouns in Bosnian**

a. [biro] ‘office (M NOM sg)’  
b. [biro:i] ‘offices (M NOM pl)’  
c. [tabure] ‘stool (M NOM sg)’  
d. [tabure]a ‘of the stool (M GEN sg)’

(The stressed vowel is underlined)

This situation illustrates that in Bosnian too, a CV syllable is active. It must be associated to K° instead of n° or #°: this explains why all nouns display case marking (this is not generally the case in Slavic, e.g. Russian has some indeclinable nouns).

Now let me go back two the fact that Bosnian tolerates a surface zero Vfin, namely the M sg NOM ending. Why is this possible? If there is a one-to-one correspondence between gender and a given group as in (16), this entails that this zero is underlingly the addition of three null morphemes:  

(18) **Null Morphemes**

a. Gender: M gender is marked by zero.  
b. Number: sg. is marked by zero.  
c. Case: NOM is marked by zero.

Standard assumption is that M is marked by a yer-vowel which does not surface as it is in a weak position (cf. Jakobson 1948, Baily & Nevin 2008, Halle & Marantz 2008 and Halle & Nevin 2008). The point is that in the case of Bosnian, the only zero at the surface is the M sg NOM, and it may not be an accident, as Despîć (2009) notices too.

Like for Italian nouns, we could simply list the VI’s and the rules accounting for the cases of syncretism between the forms, as group 1 and group 2 sg GEN. Instead, consider again TE and its prediction on the underlying structure of vocalic segments. The decomposed paradigm of Vfin’s is shown below:

(19) **Decomposed vocalic case endings**

<table>
<thead>
<tr>
<th></th>
<th>group 1</th>
<th>group 2</th>
<th>group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>F</td>
<td>NEU</td>
</tr>
<tr>
<td>sg.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pl.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sg.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pl.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. NOM</td>
<td>zero</td>
<td>I</td>
<td>A</td>
</tr>
<tr>
<td>b. GEN</td>
<td>A</td>
<td>A</td>
<td>A.I</td>
</tr>
<tr>
<td>c. DAT/LOC</td>
<td>U</td>
<td>I(ma)</td>
<td>I</td>
</tr>
<tr>
<td>d. ACC</td>
<td>(A)</td>
<td>A.I</td>
<td>U</td>
</tr>
<tr>
<td>e. INSTR</td>
<td>A.U(m)</td>
<td>I(ma)</td>
<td>A.U</td>
</tr>
</tbody>
</table>

By hypothesis in (18), the column corresponding to sg M is transparently formed by case markings, whereas the line of NOM shows either number or gender or the combination of both markings.

As for gender, I follow Lowenstammm’s (2008) decomposition where NEU is marked by the feature [-gen(der)] whereas M and F are marked [+gen, -F] and [+gen, +F], respectively.

Finally, I reinterpret Halle’s (1997) decomposition to manage case morphemes, as the original system does not provide a matrix for either LOC nor INSTR whereas ABL does not appear in Bosnian:13

---

(20) Case features

<table>
<thead>
<tr>
<th>Case</th>
<th>Oblique</th>
<th>Structural</th>
<th>Superior</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOM</td>
<td>- oblique</td>
<td>+ structural</td>
<td>+ superior</td>
</tr>
<tr>
<td>ACC</td>
<td>- oblique</td>
<td>+ structural</td>
<td>- superior</td>
</tr>
<tr>
<td>GEN</td>
<td>+ oblique</td>
<td>+ structural</td>
<td>- superior</td>
</tr>
<tr>
<td>DAT/LOC</td>
<td>+ oblique</td>
<td>α structural</td>
<td>+ superior</td>
</tr>
<tr>
<td>INSTR</td>
<td>+ oblique</td>
<td>- structural</td>
<td>- superior</td>
</tr>
</tbody>
</table>

The consequence of these successive hypotheses is that we can now list the VI’s for each gender, number and case, stemming from the crucial observations in (18). This is schematized below:

(21) VI’s for Bosnian nouns

a. [-pl] ↔ zero
d. [+gen, -F] ↔ zero
e. [+gen, +F] ↔ A
f. [-oblique, +structural, +superior] ↔ zero
g. [-oblique, +structural, -superior] ↔ A
h. [+oblique, +structural, -superior] ↔ A
i. [+oblique, α structural, +superior] ↔ U
j. [+oblique, -structural, -superior] ↔ A.U

The crucial idea of this paper is in fact that a correct analysis of linguistic facts must tend towards a reduction of the ingredients a speaker must handle to produce grammatical forms. (21) shows exactly this situation: instead of positing lists of VI’s in competition among them, this approach proposes the unique and fixed list shown below. Of course, surface reality displays “exceptions”, but these must and can be accounted for by either TE-internal mechanism or DM-internal allomorphy rules, as shown for the Italian system.

In addition, such an approach entails that each morpheme projects its own category, disallowing for such a processes as fusion of nodes, as proposed for Latin by Halle & Vaux (1998) or for Slavic in general by Halle & Nevins (2008). This does not mean that surface forms display three distinct morphemes. Instead, this reinforces the abstractness of the phonological component in the Grammar, whereas the syntax works in a more transparent way. In the rest of the section, I show how this approach accounts for Bosnian nouns.

I propose to concentrate first on NOM. Note that NEU nouns can end in [e], too i.e. srce ‘heart’. This ending affects all direct case but only at sg. I recast the data including this fourth group:

(22) Decomposed V.fin for NOM

<table>
<thead>
<tr>
<th>Masculine</th>
<th>Feminine</th>
<th>Neuter</th>
</tr>
</thead>
<tbody>
<tr>
<td>sg.</td>
<td>pl.</td>
<td>sg.</td>
</tr>
<tr>
<td>I</td>
<td>A.I</td>
<td>A.U - A.I</td>
</tr>
<tr>
<td>zero</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

M and F V.fin’s are totally transparent: they combine the given VI’s into actual surface forms. As for NEU, more explication is needed. First, notice that an element A is strangely present at sg. If U or I mark the gender (lexical marking), why should these nouns be marked by a further morpheme?

---

14 Browne (1993) claims that this is due to a palatalizing effect of the last radical consonant, namely [ts], [s], [dž], [š] and [ž]. Hammond (2005) does not acknowledges such a situation but the example she provides for group 10.d is polje ‘field’ which has a palatal in the last radical position ([j]). However, I am aware of at least a counterexample: more ‘sea’. Either way, the system I provide accounts for NEU nouns.
This becomes clear if we consider the whole paradigm. NEU display NOM-ACC syncretism (a normal fact in Indo-European), and ACC is crucially marked by the element A (cf. 19.d and 21.g). This means that in the context of roots triggering NEU, the spell-out of NOM and ACC is the same, i.e. the feature [superior] is neutralized. A similar situation actually holds for M nouns, as they display syncretism between NOM and ACC, too, but only when the referent is inanimate: in this case, the zero morpheme marks both cases. Still, we need an explanation for NEU pl, as I₁[+pl] must appear as well as the REU or I. This becomes clear when the whole case system is discussed, later on.

Building on the structure provided for Italian in (11), I propose the following for Bosnian nouns:

(23) Syntactic structures and complex heads in Bosnian

\[ \text{(23)} \]

\text{Crucially, the adjunction of Th nodes to each functional head at PF accounts for the presence of distinct morphemes in the structure. Again, it is the later application of phonology and eventually of morphological allomorphy rules which map these ingredients into actual surface forms. In the case of kuće ‘houses’ F pl NOM, the situation is quite simple, as represented in (23).}

The table (24) shows the whole paradigm for nouns in Bosnian: each case contains three morphemes, as by hypothesis. Double-framed cases need further clarifications, as the expected surface form is not compositionally the fusion of the underlying elements.

(24) Underlying morpheme-sequences

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOM</td>
<td>GEN</td>
<td>NEU</td>
</tr>
<tr>
<td>sg.</td>
<td>pl.</td>
<td>sg.</td>
</tr>
</tbody>
</table>

It is interesting to note that almost all plurals present a problematic case. Nevertheless, a crucial distinction must be made between forms which can be accounted for by TE itself (white framed).
and those which require a deeper effort by the theory, namely the application of specific allomorphy rules.

I focus on the former, first. In the four cases at study, the element I surfaces alone, despite the fact that other elements are at play in the underlying structure. TE explains that in a five-vowel language as Bosnian, element I and element U are represented on the same abstract level, thus allowing for the five combinations shown above in (7). This prohibits front rounded vowels such as [y] or [o] from existing in such a system. For whatever phonology external reason, if I and U are forced to be close at the same level of representation (the templatic tier CV), the phonology formally disallows fusion between these two elements (or more of the same kind): the result is that either one or the other surfaces alone. In the case at hand [i] is the output vowel, as expected by a general trend: Italian pl. above are exactly the same.\(^{15}\)

The latter group of problematic forms, instead, clearly requires specific rules which apply in the phonological component. The following allomorphy rules can be postulated:

\[(25) \text{Allomorphy rules (not exhaustive)}\]

a. AR1: Element \(\rightarrow\) zero /\_A\[\text{+oblique, +structural, -superior}\]
b. AR2: Element U \(\rightarrow\) zero /\_I\[pl\]
c. AR3: A\[-oblique, +structural, +superior\] \(\rightarrow\) zero /\_√\[inanimate referents\]

AR1 applies in the context of GEN, which has been recognized to have particular status in Slavic (cf. Bailyn & Nevins 2008 for a recent analysis). In their terms, however, GEN inherits its particular status from the form of the theme. Such a solution could be easily inserted within the theory presented here, as:

“distribution of genitive plural endings will depend on the form of the stem (that is, on the phonological nature of the them) and the genitive plural ending emerges with no reference necessary to class or gender information per se [...]” (Bailyn & Nevins 2009:256)

I will return on this point in section 5.

As for allomorphy rules, AR2 erases the RE U in NEU pl. Consequently, a rule of the type of AR1 applies and erases I\[pl\].

AR3 is interesting as it applies in sg. whenever its context is satisfied (adjacency of the root), but crucially it cannot ever apply at pl., as the context is blocked by the intervening I\[pl\]. As predicted, M pl ACC is [e] in all cases, regardless the form of sg.

Other and more specific rules must be postulated to account for light-grey framed F cases, namely GEN, DAT-LOC and ACC singular. Further analyses must be dedicated in order to determine the exact nature of such rules.

5. Paradigm(s)

The theory presented here brings us to the discussion about the linguistic status of paradigms: are these an invention made by linguists or are they a reality speakers use to access the correct output form in producing an utterance?

Throughout the paper, I used the term “paradigm” without any theoretical relevance, but it seems clear from the presentation of the main ideas that I do not consider paradigms as a relevant reality for the production of an actual surface form. The approach presented here is incompatible with such a view, as processes applying in the phonological component do not have any formal skill to know as to why things are as they are at that particular stage of the derivation. Put in another way, the

\(^{15}\) Cf. Bosnian loan *biro* ‘office’ from French *b[y]reau*. 

constitution of the paradigm shown in (16) is an accident as far as its shape is considered. This is the result of the application of syntactic rules which build the complex head and then the result of some phono-morphological rules which apply locally and contextually. For instance, AR1 in (25) applies to whatever element is adjacent to an element /A/ marking GEN. It does not do any reference to the fact that a given stem belongs to the paradigm of M nouns or NEU or to the paradigm of o-stem nouns, etc.

In this sense, I agree with Bobaljik (2008) when he asserts that:

“To predict the surface form of a word, it is sufficient to know the constituent pieces of that word, their hierarchical arrangement, and the general phonology of the language. Reference to other members of that word’s paradigm is neither needed nor possible.” (Bobaljik, 2008:X)

Consider again the table in (24). If paradigms exist as an independent morphological reality, we’d expect that a given form could be built in doing reference to another specific form. For instance, we could assert that knowing the NOM of a given form in combination with its corresponding GEN would be the way of generating the whole paradigm. But still, we should postulate a system as the following:

(26) Paradigmatic influence in Bosnian nouns?

In (26), the form of case endings is predictable whenever one knows a given NOM-GEN pair but this does not mean that the surface forms are derived from NOM-GEN ones. In this approach, we know that for a given NOM-GEN pair, another given list of forms corresponds, but this does not say anything about the fact that form can share some morphemes or some phonological properties or have the same underlying representations or properties. It is even more an accident if forms are syncretic or easily recognizable as to belonging to a particular group. Such an approach is lexicalist in the sense that it postulates the existence of lexical knowledge in determining the exact path of morphological processes such as inflection. Burzio’s (1998) analysis of Italian verb stem allomorphy is an interesting paradigmatic approach to word formation. His analysis accounts for stem allomorphy by means of a one-to-one correspondence between output forms, which become thus a prominent factor of lexical knowledge. Speakers do need to know the behavior of a given paradigm in order to compare all other formations to the former. This approach is computed within OT where, again, not only the system can overgenerate forms but also it can see the whole derivation process while creating a new form. Burzio’s schematized system (1998: 97) recalls the one I depicted in (26), where the idea of paradigm must be pertinent along the derivation. It seems to me that the more constraints in analyzing hold, the more accurate predictions are made.

Let me go back to Bosnian nouns. There is a fourth small group of nouns in the language that I have been skipping since the beginning of section 4.2. These nouns are all F and their NOM is zero, as M one: that is they do not have any RE associated to. However, the other Vfin’s are different from M ones, but crucially only at sg. In the terms of the current approach, this means that AR’s are active, and they have the following forms:

(27) More on allomorphy: stvar-type nouns

a. AR4: Element[+oblique, α structural, α superior] → zero /√[list(x)]
b. AR5: Element[-oblique, α structural, α superior] → 1 /√[list(x)]
If this approach is correct, whenever an intervening morpheme stands between the roots in the list(x) and the case morpheme, no allomorphy should be observed. This is actually the case, as the pl. of this group of nouns, follow the regular pattern:16

(27bis) Underlying forms for stvar-type nouns:

<table>
<thead>
<tr>
<th></th>
<th>sg.</th>
<th>pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>NOM</td>
<td>ø.ø.ø</td>
</tr>
<tr>
<td>b.</td>
<td>GEN</td>
<td>ø.ø.I</td>
</tr>
<tr>
<td>c.</td>
<td>DAT/LOC</td>
<td>ø.ø.I</td>
</tr>
<tr>
<td>d.</td>
<td>ACC</td>
<td>ø.ø.ø</td>
</tr>
<tr>
<td>f.</td>
<td>INSTR</td>
<td>ø.ø.I/U</td>
</tr>
</tbody>
</table>

6. Spell-out

This section is an attempt in generalizing furthermore the theory proposed insofar. A radical consequence of such view is that the border between the application of syntax and phonology is the spell-out. Before that, no information on phonological material is associated to the terminals whereas right after occurrence of spell-out, the root and all the morphemes are filled by phonological content. Bendjaballah & Haiden (2008) propose a typology of the possible shapes a spell-out can have (six in their terms). Throughout this paper, I have assumed that such an operation can associate four different types of phonological content to terminals, as (28) illustrates:

(28) Types of spell-outs

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>Segmental</td>
<td>/XY/</td>
<td>CV</td>
</tr>
<tr>
<td>b.</td>
<td>Templatic</td>
<td>/XY/</td>
<td>CV</td>
</tr>
<tr>
<td>c.</td>
<td>Both</td>
<td>/XY/</td>
<td>CV</td>
</tr>
<tr>
<td>d.</td>
<td>Silence</td>
<td>zero</td>
<td></td>
</tr>
</tbody>
</table>

Italian number markings are examples of the kind (28.a), whereas Bosnian number sg. is of kind (28.d). The type (28.b) has a good sample in NOM case in Bosnian, as only the syllable CV is spelled-out (a property of Th adjoined to K° more than of NOM case, but the question overruns the scope of this article). Finally, roots are good examples of (28.c) kind of spell-out. Crucially, languages displaying non-concatenative morphology must allow for a templatic and segmental spell-out for roots, hence disallowing interdigitation of morphemes and thus preserving the unity of the roots.

Another issue raised by my hypothesis deals with cases where it is hard to trace a unique underlying form for the whole inflectional paradigm. Both Italian and Bosnian seem relatively simple cases to analyze within this approach, basically because every inflectional ending has vocalic shape throughout all genders, numbers and syntactic cases. What does the theory predict about languages having more complex systems, such as plurisyllabic inflectional morphemes or suprasegmental phenomena, such as tones?

This is the focus of following section.

---

16 Again, the only problem is created by GEN pl. which shows unexpected [i] instead of predicted [e] or, by application of AR1, [a] (we should discuss the correct rule order). But GEN pl. are impostors.
7. A third case study: the Somali nominal system

7.1 Departing analyses

I propose to briefly concentrate on Somali nominal system which represents an old challenge for grammatical theory because of two interesting phenomena: (i) the presence of a tonal accent (Hyman 1981), and (ii) the presence of gender polarity within the nominal system.

This system is complex and it has not been completely understood. As for the polarity, for instance, two major proposals have been made. Lecarme (2002) established a syntactic account which consists basically in the idea that plural suffixes come with their inherent gender and with categorial information (they never attach to categories other than nouns), just as derivational suffixes do. The second proposal deals with a CVCV.. phonology approach and makes the strong claim that gender polarity does not exist: it is only apparent. Godon (1998), Bendjaballah (2001) and Barillot (2002) propose that the polarity is the result of adding a CV-syllable at the right edge of a noun, which entails the shift of the accent and thus the [+F] interpretation (M nouns are paroxyton whereas F are oxyton at surface). Both analytical paths seem to recognize that the tonal accent and the gender polarity are closely related, but none clearly states why the tone seems to disappear at pl. and in cases other than Abs.17

In fact, Somali does not only show these two phenomena in its nominal system, it also has other two interesting features. First, the existence of declensional patterns have traditionally been recognized (Andrzejewski 1964, Hyman 1981, Puglielli 1981 and Saeed 1993). These are normally based on tonal patterns and other morpho-phonological properties. Secondly, Somali has three morpho-syntactic cases: Abs(olutive), Nom and Gen (plus Voc). As Saeed (1993:138) notes, Abs is the most basic form morphologically and is the only form which cannot be predicted by another case form, say Nom. The opposite is true: knowing an Abs form for a given noun allows the prediction of the form of all other cases. What is the syntactic role of Abs, though? Puglielli & Frascarelli (2005, 2007) propose to consider that Abs is assigned to those nouns being syntactically in a Small Clause constituent, a traditionally non-argument position. This explains why subjects in Somali take Nom case only when they are not focused, in all other instances they are at the Abs. I will return on this point in the following sub-section 7.2.

Again, Saeed (1993:138) claims that Abs case is an oblique one, “which nouns show when they are direct or indirect objects, or adverbials, or as we shall see later, when they are focused.” Moreover, consider that Nom is characterized by the absence of tonal accent, and as far as F nouns ending in a consonant are involved, by a [i] marking, which attaches to the right. But whenever a nouns phrase is marked by such a case, the [i] marking occurs only once, at the rightmost edge of the phrase. Finally, Gen implies tonal shift: high tone is moved one syllable rightwards (hence it is almost always on final vowel). F nouns add the morpheme -eed (or -yeed) (some exceptions are excluded, though).

For sake of clarity, I show a concise view of the cases just presented in (29):

(29) Somali case markings

a. Abs: unpredictable form (see declensions tables below).
b. Nom: absence of tone. F nouns allow the insertion of [i] to the rightmost element of the NP.
c. Gen: the tone moves one mora rightwards. F nouns add [-eed] at the right edge.

17 My examples are taken from Godon (1998) as she worked with a native speaker informant and thus she could verify her data.
d. examples (from Godon 1998:27)

<table>
<thead>
<tr>
<th></th>
<th>Absolutive</th>
<th>Nominative</th>
<th>Genitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>i.</td>
<td>xòog</td>
<td>xoog</td>
<td>xoóg</td>
</tr>
<tr>
<td></td>
<td>force (M)</td>
<td>force (M)</td>
<td>force (M)</td>
</tr>
<tr>
<td>ii.</td>
<td>géri</td>
<td>gerí</td>
<td>mindyééd</td>
</tr>
<tr>
<td></td>
<td>giraffe (M)</td>
<td>giraffe (M)</td>
<td>kinfe (F)</td>
</tr>
<tr>
<td>iii.</td>
<td>mindí</td>
<td>mindi</td>
<td>naageéd</td>
</tr>
<tr>
<td></td>
<td>kinfe (F)</td>
<td>kinfe (F)</td>
<td>woman (F)</td>
</tr>
<tr>
<td>iv.</td>
<td>naág</td>
<td>naagi</td>
<td>naág</td>
</tr>
<tr>
<td></td>
<td>woman (F)</td>
<td>woman (F)</td>
<td>woman (F)</td>
</tr>
<tr>
<td>v.</td>
<td>aábbe</td>
<td>aabbe</td>
<td>aabbé</td>
</tr>
<tr>
<td></td>
<td>father (M)</td>
<td>father (M)</td>
<td>father (M)</td>
</tr>
</tbody>
</table>

Note that the unpredictable forms are the Abs ones: these must be known and there is no other way to produce a well-formed noun in this case. On the other hand, Nom and Gen forms can be derived by Abs in a totally predictable way, as explained above.

For this reason, I suggest to concentrate first on Abs nouns. A lexical tonal accent is traditionally recognized to be associated to this case. In addition, seven major and distinct groups are distinguished, depending on the position of the tone, the gender, specific suffixes and finally gender polarity. Basic Somali nouns are shown below, following Saeed (1993:129-137):

(30) Somali nouns

### a. Singular

<table>
<thead>
<tr>
<th></th>
<th>Form</th>
<th>Gen</th>
<th>Tone</th>
<th>Ex.</th>
<th>Premodifier</th>
<th>Ex.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 a</td>
<td>*-o</td>
<td>F+</td>
<td>v v #</td>
<td>naág</td>
<td>v v - ta</td>
<td>naágta</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td></td>
<td></td>
<td>galáb</td>
<td></td>
<td>galábta</td>
</tr>
<tr>
<td>2 a</td>
<td>*-e</td>
<td>M+</td>
<td>v v #</td>
<td>albááb</td>
<td>v v - ka</td>
<td>albáába</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td></td>
<td></td>
<td>daríiq</td>
<td></td>
<td>daríiq</td>
</tr>
<tr>
<td>3</td>
<td>C(V)V,CV,C</td>
<td>M</td>
<td>v v #</td>
<td>ilig</td>
<td>v v - ka</td>
<td>iligga</td>
</tr>
<tr>
<td>4</td>
<td>C(V)VC</td>
<td>M</td>
<td>v (v) #</td>
<td>míís</td>
<td>v (v) -ka</td>
<td>míiska</td>
</tr>
<tr>
<td>5</td>
<td>*CVC</td>
<td>M</td>
<td>v v #</td>
<td>mádax</td>
<td>v v - ka</td>
<td>mádaxa</td>
</tr>
<tr>
<td>6</td>
<td>-o</td>
<td>F</td>
<td>v v #</td>
<td>hóóyo</td>
<td>v v - ta</td>
<td>hóoyáda</td>
</tr>
<tr>
<td>7</td>
<td>-e</td>
<td>M</td>
<td>v v #</td>
<td>báre</td>
<td>v v - ta</td>
<td>báráha</td>
</tr>
</tbody>
</table>

### b. Plural

<table>
<thead>
<tr>
<th></th>
<th>Form</th>
<th>Gen</th>
<th>Tone</th>
<th>Ex.</th>
<th>Premodifier</th>
<th>Ex.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 a</td>
<td>-o</td>
<td>M</td>
<td>--</td>
<td>naago</td>
<td>v v - ka</td>
<td>naagáha</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td></td>
<td></td>
<td>galbo</td>
<td></td>
<td>galbáha</td>
</tr>
<tr>
<td>2 a</td>
<td>-Co</td>
<td>F</td>
<td>--</td>
<td>albaabbo</td>
<td>v v - ta</td>
<td>albaabáda</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td></td>
<td></td>
<td>daríiqyo</td>
<td></td>
<td>daríiqyáda</td>
</tr>
<tr>
<td>3</td>
<td>CV(V)Co</td>
<td>M</td>
<td>--</td>
<td>ilko</td>
<td>v v - ka</td>
<td>ilkáha</td>
</tr>
<tr>
<td>4</td>
<td>CV(V)CaC</td>
<td>M</td>
<td>--</td>
<td>míisas</td>
<td>v v - ka</td>
<td>míisáska</td>
</tr>
<tr>
<td>5</td>
<td>o-óyin</td>
<td>F</td>
<td>v v #</td>
<td>mádáx</td>
<td>v v - ta</td>
<td>mádáxda</td>
</tr>
<tr>
<td>6</td>
<td>a-yaal</td>
<td>M</td>
<td>v v #</td>
<td>hóoyoóyin</td>
<td>v v - ka</td>
<td>hóoyoóyinka</td>
</tr>
<tr>
<td>7</td>
<td>a-yaal</td>
<td>F</td>
<td>--</td>
<td>barayaal</td>
<td>v v - ta</td>
<td>barayaásha</td>
</tr>
</tbody>
</table>

Tonal accent is marked only when high, as in Hyman (1981). Somali nouns are written in the proper national orthography which is basically phonological. In addition, double-framed cases point to gender polarity.

Note that not all authors agree on the number of declensions. For instance, Andrzejewski (1964) claims that these are eight instead of seven and he defines them on strict basis of tonal patterns, hence permitting a sg. noun in a declension and its corresponding pl. in another one given their different tonal behavior. As Godon (1998) follows Saeed (1993) and as I use her central hypothesis.
on feminine formation, I adopt Saeed’s classification, too. Moreover, recall that in my approach the paradigmatic behavior is only an accident: the fact that a root is associated to group 2 or group 5 is a mere consequence of some formal phonological facts but not a consequence of lexical paradigmatic properties.\textsuperscript{18}

Observe that sg. nouns have a tone, either on the last mora (F) or on the penultimate (M). Only group 6 nouns (F) do not follow this pattern. Crucially, groups 6 and 7 display nouns with a clear distinct suffix, -e and -o. Lecarme (2002) shows that these suffixes are instances of n° (for instance, deverbal nouns always take such a suffixes) and she also shows that these suffixes have inherent gender. I will treat group 6 and 7 later and I concentrate first on groups 1 to 5.

As for plural, gender polarity affects all groups except 3 and 4. In addition, nouns in groups 1 to 4 and in group 7 do not have any tonal accent. Another important factor is that pl. are generally marked by specific suffixes, as -o (cf. groups 1, 2 and 3) or reduplication applies (group 4).

Godon (1998:45-76) deeply analyzes how plural works in Somali. Her most important results crucially deal with an abstract phonology which can account for a large amount of surface-unrelated phenomena.

In particular, she shows that neither tonal accent shift nor reduplication are the proper way to pluralize a noun in Somali. Note that only M nouns can be pluralized by moving the tone and conversely, moving a tone can also trigger the change of gender:

\textbf{(31) Tonal accent shift}

\begin{tabular}{lcc}

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. plural</td>
<td>éy</td>
<td>a dog</td>
</tr>
<tr>
<td></td>
<td>ey</td>
<td>some/many dogs (male, female or both)</td>
</tr>
<tr>
<td>b. feminine</td>
<td>éy</td>
<td>a male-dog</td>
</tr>
<tr>
<td></td>
<td>ey</td>
<td>a female-dog</td>
</tr>
</tbody>
</table>

As for reduplication, Godon demonstrates that it applies only in the case of templatic roots having the form /C1C2C2/ where the second C2 cannot surface at sg, but only at pl (cf. group 4). This situation mirrors the case of Semitic deaf verbs. Hence, reduplication is just a consequence of a templatic constraint and not a pluralizing phenomenon.

What is the pl morpheme then? Godon proposes that this is a morpheme /A/, an element in the sense of Kaye et al.’s (1985) theory. This accounts for the plural of groups 1, 2, 3 and 4, and it has an allomorph [o] when it surfaces at the word boundary. I show this situation in (32):

\textbf{(32) Plural morpheme in Somali (cf. Godon 1998:64)}

\begin{tabular}{llllll}
|          |          |          |          |          |          |
| a. CVCVCV | b. CVCVCVCV | c. CVCVCV | d. CVCVCVCV |          |
| a        | a        | a        | i        |          |
| A        | A        | A        | A        |          |
| (U)      | (U)      | (U)      | (U)      |

naago (1)  inammo (2)  galbo (3)  miisas (4)

This is an important result, as goes towards the hypothesis I am defending throughout this paper. A unique spell-out for pl. accounts for at least four surface-unrelated groups of Somali nouns, as shown below:

\textsuperscript{18} I refer the reader to all cited sources for more details and a complete justification as to whether there are seven or eight classes (or three as in Hyman’s (1981) account).
(33) VI for plural

Plural is marked by morpheme /A/ in the case of groups 1, 2 and 3/4.

But there is an other analysis which is fundamental for my approach. Godon also shows that nouns can either have a template or not in the representation of their root. If they have one, this holds the group 3 and 4, which in GP framework have the same underlying representations, as shown below:

(34) Identity of groups 3 and 4

<table>
<thead>
<tr>
<th>UR</th>
<th>sg.</th>
<th>pl.</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. group 3</td>
<td>/C_1(v)vC_2øC_3/</td>
<td>C_1(v)vC_2vC_3</td>
<td>gärab</td>
</tr>
<tr>
<td></td>
<td>gärøb/</td>
<td>gärøb</td>
<td>gärøb</td>
</tr>
<tr>
<td></td>
<td>/garøb/</td>
<td>gärøb</td>
<td>gärøb</td>
</tr>
<tr>
<td>b. group 4</td>
<td>/C_1(v)vC_1øC_i/</td>
<td>C_1(v)vC_i</td>
<td>míis</td>
</tr>
<tr>
<td></td>
<td>/miisøs/</td>
<td>míisøs</td>
<td>míisøs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>míis</td>
<td>míisas</td>
</tr>
</tbody>
</table>

On the other hand, nouns in groups 1, 2 and 5 do not have any particular constraint as far as their template is concerned, hence they form an open class (which is confirmed by the fact that group 2 nouns are the most numerous and almost all M nouns can be pluralized as in group 5, even if they lexically display other strategies): loans fall into it.

To sum up the generalization made so far, we have a unique plural morpheme /A/ (group 5 still needs an explanation) and we have a parameter on root template, all other things being equal. In the representation (32), of course, groups 1 and 2 receive a CVCCV.. later in the derivation, as otherwise the segments couldn’t be pronounced.

Let us focus on the way language creates a feminine noun. Godon (1998) makes a strong claim on how F is created in the language: she argues for an underlying CV syllable attached at the right edge of a M noun which is responsible for the tonal accent shift and thus for the F interpretation. In this view, the tonal difference between M and F nouns is only surface apparent: it becomes a consequence of the presence of a phonological abstract syllable.

The analysis proceeds as follows. Observe the nouns without a template, i.e. groups 1, 2 and 5. F gender seems to be the only constraint for a noun in group 1, whereas the opposite situation occurs for group 2, where nouns are M. Crucially, nouns having either the suffix -o or -e cannot be in these groups. Finally, group 5 displays nouns having a penultimate tone and being all M.

Now observe the distribution of gender throughout singular and plural:

(35) Gender polarity distribution

<table>
<thead>
<tr>
<th></th>
<th>a. singular</th>
<th>b. plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>group 1</td>
<td>feminine</td>
<td>masculine</td>
</tr>
<tr>
<td>group 2</td>
<td>masculine</td>
<td>feminine</td>
</tr>
<tr>
<td>group 3/4</td>
<td><strong>masculine</strong></td>
<td><strong>masculine</strong></td>
</tr>
<tr>
<td>group 5</td>
<td>masculine</td>
<td>feminine</td>
</tr>
</tbody>
</table>

If a noun have a template at sg. then no gender polarity is observed (cf. group 3/4). In addition, groups 2, 3/4 and 5 point to a correlation between pl. gender and templatic constraints at sg, as whenever a template is specified, then pl. gender is M. Conversely, if no template is specified, the pl. will be F.

However group 1 points to the contrary. It behaves as it had a template as its pl. is M, despite the fact the sg. is F and without a template. Nevertheless, these are the only nouns to be F at sg.

Moreover, consider that pl. M nouns are formed by simply adding the morpheme /A/ (cf. group 1 and 3/4) whereas pl. F nouns either display geminated consonants at the boundary between the root
and the suffix in addition to the morpheme /A/ (group 2) or they do not show the regular pl. suffix (group 5):

(36) *M* and *F* plural

<table>
<thead>
<tr>
<th>sg.</th>
<th>pl.</th>
<th>pl. morpheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>naág</td>
<td>F</td>
<td>naago</td>
</tr>
<tr>
<td>M/F</td>
<td></td>
<td>/A/</td>
</tr>
<tr>
<td>gára</td>
<td>M</td>
<td>garbo</td>
</tr>
<tr>
<td>M/F</td>
<td></td>
<td>/A/</td>
</tr>
<tr>
<td>albaa</td>
<td>M</td>
<td>albabbo</td>
</tr>
<tr>
<td>F</td>
<td></td>
<td>/A/ + germination of final /b/</td>
</tr>
<tr>
<td>màdax</td>
<td>M</td>
<td>madàx</td>
</tr>
<tr>
<td>F</td>
<td></td>
<td>tonal accent moves one mora</td>
</tr>
</tbody>
</table>

The correlation between the form of plural and its gender is strikingly evident: plural is *M* when the morpheme /A/ alone suffices. On the other hand, plural is *F* when either /A/ goes with a gemination or tonal accent shift occurs without the insertion of /A/. It seems that *F* nouns hide something which does not let us have a whole vision on the phenomenon.

Hence, Godon concludes that:

“Il apparaît donc inutile de concevoir le genre des noms en somali en terme de polarité de genre. Si le genre du pluriel n’est pas attribué par polarité, c’est qu’il dépend du type de formation du pluriel utilisé”. (Godon *ibidem* : 73)

The gender polarity is only apparent, as it is a secondary effect of a deeper phenomenon. The question is now to discover the exact nature of such a phenomenon.

The hypothesis Godon makes is that *F* gender is marked by a CV syllable merged at the right edge of the root. A crucial argument to support such a claim is the gemination observed at the right edge of the root in plural forms of the group 2 (cf. 36.c):

(37) *Gemination at plural entails feminization (group 2)*

<table>
<thead>
<tr>
<th>a. singular</th>
<th>b. plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>b b r</td>
<td>b b r</td>
</tr>
<tr>
<td>CVC CVC</td>
<td>CVC CVC + CV</td>
</tr>
<tr>
<td>a u</td>
<td>a u A</td>
</tr>
<tr>
<td>M</td>
<td>F</td>
</tr>
</tbody>
</table>

*baabiuur*                      *baabuurro*                     “truck(s)"

Crucially, plurals of groups 1 and 3/4 do not show this gemination, but only the morpheme /A/, which has been stated to be the proper plural marking. Nouns in group 2 are *F* at pl. because of the presence of the final underlying CV.

Godon (*ibidem*: 74-75) also observes the behavior of the tonal accent when the noun is in premodifier form, i.e. determined by the article /ka/ (/ta/ for *F*). Table (30) shows both sg and pl of determined nouns. Note that in pl premodifier forms, all nouns have the tonal accent in their last mora, regardless the gender. On the other hand, only sg *F* nouns show the tonal accent in their last mora, regardless the gender.

---

19 Both /t/ and /k/ undergo phonological processes which transform their surface forms depending on adjacent consonants. I will not analyze such a phenomenon here.
mora when in premodifier form. Indeed, nouns in groups 6 and 7 also show a final mora tonal accent. These nouns behave differently, as noticed, but crucially are suffixed both at sg and pl.

It has been shown that pl are either suffixed or F nouns (cf. 36). In addition, (37) shows that F nouns are also suffixed by adding a CV. Thus, the generalization below follows:

(38) *Tonal accent in premodifier forms*

a. Both suffixed and F nouns have a final mora tonal accent;
b. M sg. nouns have a penultimate mora tonal accent.

This generalization points to the fact that a F noun behaves like a suffixed noun, hence allowing for the hypothesis that feminization actually means “add a CV syllable at the right edge of the root”, as claimed by Godon. Thus, in terms of the present account, the F morpheme is represented by the following VI:

(39) *F morpheme*

F gender is marked by a suffixed CV.

A consequence of the hypothesis in (39) is that, as Godon (*ibidem*: 82-85) briefly sketches, the tonal accent can be seen as always occurring on the penultimate mora, regardless the gender of the noun. Consider group 5 pl (36.d): it seems that the fact of moving rightwards the tonal accent implies the changing of number (and gender). If we posit the presence of a simple CV syllable in the representation of pl and a rule assigning the tonal accent to the second mora from the right, it follows that that form must be F:

(40) *Group 5 plural*

A. singular  
2 1 -  
TA m d x  
| | |  
CVC

| a M  
| m d x  
| | |  
CVC

B. plural  
2 1 -  
TA m d x  
| | |  
CVC

| a F  
| | |  
CVC

|mádax  

Tonal accent shift is marked by an underlined V. Final V’s do not count as morae. These are the general lines of Godon’s analysis, which is, as far as I know, the most complete one on the morpho-phonology of Somali nouns. The following sub-section illustrates my proposal, which takes the same analytical path as for Italian and Bosnian nouns, cf. *supra* 4.

7.2 Syntactic structures and VI’s for Somali nouns

Godon (*ibidem*: 84) clearly claims that the tonal accent is not a gender marking and that it is always on the second mora from the right. This result follows from the analysis she provides in her dissertation.
I propose that TA is in fact the case marking. At Abs, it is assigned to the second mora on the right; whereas at Nom (non-focused subject), it is phonetically zero and, finally, it is assigned to the first mora on the right when the noun is at Gen. This hypothesis is in contrast with Saeed (1999) when he declares that:

“If the absolutive case of a noun is basic and inherent, the other cases are best seen as processes which operate on this basic form.” (Saeed 1999: 141)

In the present approach, it makes no difference to know where AT falls at Abs, as the three cases are all derived by a unique underlying form: specific VI’s assign to each case its particular AT.

Now, recall that the gender is expressed by the absence or presence of a CV syllable, thus allowing the VI in (41). Recall also that the number is expressed by either zero (sg) or /A/ (pl), except than in group 5. As, by hypothesis, I claim that the morpheme /A/ is the unique pl. spell-out, in the case of group 5 too. This is shown below, too:

(41) VI’s for gender and number in Somali

a. gender
   [+F] ↔ suffixed CV
   [-F] ↔ zero

b. number
   [+pl] ↔ A
   [-pl] ↔ zero

Crucially, VI [+pl] ↔ A predicts that group 5 pl should have the form *madaxyo (like group 2 pl: dariiq vs. dariiqyo) instead of expected madax ‘heads’. Clearly, pl. /A/ shows an allomorph in the context of group 5 roots. This situation can be formalized by the following allomorphy rule, allowing for the merge between group 2 and 5:

(42) Pl allomorphy

AR1: /A/ [+pl] → zero / √ (list of roots)

If pl morpheme becomes phonetically zero, the gemination of the last radical consonant (or yod insertion when the consonant cannot geminate) cannot anymore apply as final CC# clusters are not tolerated in Somali. The question is why TA surfaces in group 5 pl whereas it does not in all other pl: a possible answer deals with the actual nature of case assignment in Somali.

We have seen gender and number VI’s. But we need a third distinction in order to define the groups 1 to 5. Interestingly, the root shows a lexical difference, as it has already been shown while discussing Godon’s analysis. A Somali nominal root can either be templatic or template-less. The former case corresponds to group 3/4, as shown in (34). This means that, contrarily to Italian and Bosnian where roots are always spelled-out with a template, Somali displays two different kinds of roots:

(43) VI’s for roots

a. √ (list A) ↔ /XYZ/  groups 1, 2, 5
b. √ (list B) ↔ /WQP/  group 3/4

| | CVCV

Hence, two parameters (one on templaticity of roots, the other on gender) trigger the existence, at surface level, of four different noun groups. The table (44) illustrates this situation:
(44) Parameters on root and gender

<table>
<thead>
<tr>
<th>root</th>
<th>gender</th>
<th>example</th>
<th>gloss</th>
<th>group</th>
<th>pl gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. templatic</td>
<td>M</td>
<td>illig</td>
<td>3/4</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>b. no template</td>
<td>F</td>
<td>naåg 'woman'</td>
<td>1</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>c. no template</td>
<td>M</td>
<td>màdax 'head'</td>
<td>5, 2</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>d. templatic</td>
<td>F</td>
<td>does not exist</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Crucially, group (44.d) does not exist, as its nouns would have two templates: one given by the root and another one given by F. Godon shows, as we have seen, that this is not possible: only one lexical template is allowed in a noun.

As we have all the ingredients for the construction of a noun, it is possible to show the syntactic structure and its corresponding complex head that underlie Somali nouns. This is shown below, for *albáab* ‘door’:

(45) Syntactic structures and complex heads in Somali nouns

\[\begin{align*}
\text{a.} & \quad \text{KP} \\
& \quad \text{K} \\
& \quad \text{[ABS]} \\
& \quad \#P \\
& \quad \# \\
& \quad \mathbf{n} \quad \sqrt{} \\
& \quad [-\mathbf{pl}] \\
& \quad [-\mathbf{F}] \\
& \quad \mathbf{albaab} \\
& \quad \mathbf{n} \quad \sqrt{} \\
& \quad \mathbf{Th} \quad \mathbf{TA} \\
& \quad (2^{\text{nd}}) \\
\text{b.} & \quad \text{K} \\
& \quad \text{K} \\
& \quad \text{Th} \\
& \quad \# \\
& \quad \mathbf{n} \quad \sqrt{} \\
& \quad \mathbf{zero} \\
& \quad \mathbf{Th} \\
& \quad \mathbf{zero} \\
& \quad \mathbf{CVCVCVCVCVC} \\
& \quad \mathbf{a l b a a b} \\
\end{align*}\]

c. [albáab] ‘door’ M, sg. ABS.

The template is inserted at PF, as a general requirement of the root to be audible. Two crucial questions still need an answer. The first one deals with the representation of ABS case (what is its feature-matrix?) and its phonetic representation at pl (it seems to disappear at pl ABS, cf. (30.b)). The second question involves the appearance of a CV at pl, thus allowing a F interpretation (i.e. gender polarity): this phenomenon occurs in groups 2 and 5 (cf. 44.d).

Let me answer to the latter question, first. Note that a single CV is inserted in the case of pl only in groups 2 and 5 which correspond to the set in (44.d). This set is exactly the one where no template appears at sg, hence Godon claims that there is a general constraint on template in nouns. This solution is in the good path but it is incomplete: why there should be a constraint on the obligatory presence of a template, as Somali does not seem to be as templatic as Arabic or Hebrew are, for instance. Instead, I propose that this is the result of the application of an allomorphy rule, which triggers the spell-out of a CV syllable, as it is shown below:

\[\begin{align*}
\text{a.} & \quad \text{KP} \\
& \quad \text{K} \\
& \quad \text{[ABS]} \\
& \quad \#P \\
& \quad \# \\
& \quad \mathbf{n} \quad \sqrt{} \\
& \quad [-\mathbf{pl}] \\
& \quad [-\mathbf{F}] \\
& \quad \mathbf{albaab} \\
& \quad \mathbf{n} \quad \sqrt{} \\
& \quad \mathbf{Th} \quad \mathbf{TA} \\
& \quad (2^{\text{nd}}) \\
\end{align*}\]
(46) Groups 2 and 5 allomorphy

AR1: zero $\rightarrow$ CV / $\sqrt{\_}/A/^{+[pl]}$ (list which corresponds to groups 2 and 5)

As the presence of a CV entails a [+F] interpretation, those roots triggering AR1 allow for F interpretation at plural.

As for the former question, the situation is more complicated. My hypothesis is that ABS is marked by a TA on the second mora from the right. This is regular at sg, but in pl. only group 5 seems to follow such a pattern, all other forms displaying only low tones (cf. 30.b). In a more recent monography, Saeed (1999) slightly modifies his representations for pl. displaying a final -o (cf. 30.b): here (ibidem: 59-66) pl. morpheme always bears high tone, i.e. TA:

(47) Plurals from Saeed (1999)

\begin{itemize}
  \item a. gacán ‘hand’ F
  \item b. gacmó ‘hands’ M group 1
  \item c. ólól ‘flame’ M
  \item d. ololló ‘flames’ F group 2
\end{itemize}

This shows how complicated is to state how the tone behaves in detail. However, note that whenever the noun is followed by the determiner (cf. premodifier forms in 30), TA appears exactly on pl morpheme (which transforms again into /a/, as it is not anymore final):

(48) Isolated forms vs. premodifier forms

\begin{itemize}
  \item a. gacmó ‘hands’ M
  \item b. gacmáha ‘the hands’ M group 1
  \item c. ólolló ‘flames’ M
  \item d. ololláda ‘flames’ F group 2
\end{itemize}

Forms (48.b) and (48.d) are regularly stressed by TA on second mora from the right, whereas the other two forms are stressed on the last mora. If we consider the data from Saeed (1993) and Godon (1998), we can postulate an allomorphy rule which erases TA when A[pl] is adjacent. Conversely, if we accept the data from Saeed (1999), we have to postulate a rule which shifts TA on pl. morpheme.

To recap, consider the whole set of VI’s for Somali nouns which accounts for the whole “paradigm” in (30), although integrated by a set of allomorphy rules:

(49) VI’s for Somali nouns

\begin{itemize}
  \item a. [-F] $\leftrightarrow$ zero
  \item b. [+F] $\leftrightarrow$ CV (cf. 39)
  \item c. [-pl] $\leftrightarrow$ zero
  \item d. [+pl] $\rightarrow$ /A/ (subject to PR1: /A/ $\rightarrow$ [o] /__#), (cf. 33)
  \item e. [ABS] $\rightarrow$ H on 2\textsuperscript{nd} mora from the right
  \item f. [NOM] $\rightarrow$ zero (maybe allomorph /i/ for F nouns)
  \item g. [GEN] $\rightarrow$ H on the last mora (maybe /eed/ too)
\end{itemize}

To conclude, let me just mention how Godon’s analysis clearly points to the type of approach I am presenting in this paper. The decomposition of Somali nouns in basic pieces is an invitation to consider each noun as built in the syntax by cyclic merge of functional heads. Phonology does the rest of the job.

8. Conclusion
The theory presented here shows that an abstract phonological component in a piece-based approach to word formation is possible. The most important consequence is the economy on VI’s and hence a minor weight of competition driven by the Subset Principle. The character of such a hypothesis is based on evidence from three unrelated languages, Italian, Bosnian and Somali.