

CHAPTER 3: GENDER IN AMHARIC NOMINALS

1 INTRODUCTION

Partially in preparation for the study of gender agreement in Chapter X, in this chapter I examine the gender system of Amharic nominals. I show how natural gender (aka semantic or biological gender, or sex) and grammatical gender (e.g., the arbitrary gender on inanimate objects) both must be part of the analysis of the Amharic gender system, and use Distributed Morphology assumptions about word formation to capture the distinction in a novel way.

In Section 2, the main descriptive facts about gender in Amharic are presented. In Sections 3 and 4, I investigate where gender features are located within DPs in Amharic, arguing that natural gender (aka semantic or biological gender) is part of the feature bundle associated with the nominalizing head *n*, whereas grammatical gender is a diacritic feature on roots. In Section 4, I develop an analysis of gender using licensing conditions that predicts which of the two sources for gender are used for agreement. Previous analyses of gender and the broader implications of the analysis here are discussed in Section 5. Section 6 concludes.

2 GENDER IN NOMINALS

As mentioned in Chapter 1, Amharic has two genders: masculine and feminine. The Amharic system for assigning gender is more reliant on natural gender (also called semantic or biological gender) than many of the more widely-known gender assignment systems (e.g., Spanish, French, Italian, Greek, etc.). For example, there is not a more or less equal division of the set of inanimate nouns into masculine and feminine. The vast majority of inanimate nouns, whether they are concrete or abstract, are masculine.

(1) Masculine Nouns (inanimate)

mot	‘death’	wɪdɪddɪr	‘competition’
kɪbr	‘honor’	bet	‘house’
wänbär	‘chair’	dɪmmɪr	‘total, sum’
dɪŋgay	‘stone’	wäräda	‘district’
kɪbab	‘circle’	gazet’a	‘newspaper’

However, grammatical gender (also called arbitrary or lexical gender) still plays a role in gender assignment. There are a handful of inanimate nominals that are usually treated as feminine.

(2) Feminine Nouns (inanimate)

mäkina	‘car’	s’ähay	‘sun’
azurit	‘whirlpool’	kätäma	‘city’
agär	‘country’	betä krɪstiyän	‘church’

As for animate nouns, their gender is assigned almost exclusively according to natural gender (Leslau 1995: 161ff., Hartmann 1980:278ff., Appleyard 1995:33). Some male and female pairs have different lexical items for each gender. These will be referred to as different root nominals.

(3) Different Root Nominals

abbat	‘father’	ɪnnat	‘mother’
bal	‘husband’	mist	‘wife’
sāw	‘man, person’	set	‘woman’
bäre	‘ox’	lam	‘cow’
wäyftɪn	‘bull calf’	gidär	‘heifer’

Note that the nominal *sāw*, besides being the typical word for ‘man,’ can also refer to a human being whose gender is not known to the speaker (‘person’).

The set of different root nominals is more or less limited to kinship terms and certain domesticated animals. The majority of animate nouns use the same root for either gender, and they will accordingly be referred to as same-root nominals. The gender of these nominals can be determined by gender agreement on associated elements, e.g., the definite marker.

(4)	a. tämari-w student-DEF the (male) student	tämari-wa student-DEF.F the (female) student
	b. muʃarra-w wedding.participant-DEF groom	muʃarra-wa wedding.participant-DEF.F bride
	c. hakim-u doctor-DEF the (male) doctor	hakim-wa doctor-DEF.F the (female) doctor
	d. halafi-w person.in.charge-DEF the (male) person in charge Walta hed12a2	halafi-wa person.in.charge-DEF.F the (female) person in charge Walta hed01a2
	e. wɪʃʃa-w dog-DEF the (male) dog	wɪʃʃa-wa dog-DEF.F the (female) dog

If natural gender is not known, the default gender is masculine, as the following example from Leslau 1995 makes clear.

(5)	hɪs’an-u wänd näw set?
	baby-DEF male is female?
	Is the baby a he or a she?
	(Leslau 1995:164)

The nominal *hɪs’an* ‘baby’ takes the masculine definite article *-u*, despite its natural gender not being known to the speaker. Further evidence that the default gender is masculine comes from data involving indefinite pronouns like ‘nobody.’

- (6) balläfaw sımmtnt betä kristiyan mannımm al-hed-ä-mm
 last week church nobody NEG-go.PF-3MS-NEG
 Last week, nobody went to church.
 (Leslau 1995:122)

The indefinite pronoun *mannımm* ‘nobody’ is animate but has no natural gender. Nevertheless, the agreement on the verb *albedämm* ‘not went’ is masculine (see Roca 1989 for a similar argument for masculine default in Spanish). Since masculine forms are used when gender is unknown (and not just for male natural sex), they can often be translated as gender-neutral, e.g., *säw* as a default masculine is ‘person,’ the nominal *ıdʒ* as a default masculine is ‘child’ (as opposed to ‘boy’).

Exceptionally, certain animals seem to have a feminine default gender in that they are assigned feminine gender when their natural gender is not known (Leslau 1995:166).

- | | | | | |
|-----|--|--|--|---|
| (7) | a. bāk’lo-wa
mule-DEF.F
the mule | b. ayt’-wa
mouse-DEF.F
the mouse | c. k’äbäro-wa
jackal-DEF.F
the jackal ¹ | d. jārärit-wa
spider-DEF.F
the spider |
|-----|--|--|--|---|

If the natural gender of the referent for one of these animals is known, though, it is the natural gender that the definite marker (and other elements) agree with.

- (8) ayt’-u
 mouse-DEF
 the male mouse

Natural gender thus ‘overrides’ grammatical gender. This kind of nominal will be crucial in the analysis of gender to come.

In terms of gender morphology, masculine gender is never morphologically marked (unsurprisingly). Feminine gender is also not universally associated with a particular affix, unlike in other Afroasiatic languages like Egyptian (where feminine gender is marked by a *-t* suffix; see Chapter X) and Hebrew (where feminine gender is marked by one of a set of suffixes; see e.g., Arad 2005). A feminine *-at* suffix has in fact been reconstructed in Proto-Afroasiatic and is one of the hallmarks of Afroasiatic languages in general (Zaborski 1992:37). A descendant of this suffix does remain in Amharic, namely, the suffix *-it*, but its use is not consistent. Some of the inanimate nouns that are feminine end in *-it*, but some do not, and not all words that end in *-it* themselves are feminine (Leslau 1995:163-164, Cohen 1970:74).

- | | | | |
|-----|-----------------------------|-------------------------|------------------------------|
| (9) | <u>Feminine, End in -it</u> | <u>Feminine, No -it</u> | <u>Masculine, End in -it</u> |
| | järärit ‘spider’ | mākina ‘car’ | mädhanit ‘medicine’ |
| | azurit ‘whirlpool’ | agär ‘country’ | särawit ‘army’ |

The suffix *-it*, therefore, is neither a necessary nor a sufficient condition for a noun to be considered feminine.

Moreover, the suffix *-it* does not convert inanimate nominals to feminine grammatical gender. In Hebrew, adding a feminine suffix (*-et, -it*) to an inanimate masculine noun derives a semantically related feminine noun (Ritter 1993).

¹ An informant reported masculine gender as the default for *k’äbäro* ‘jackal,’ but feminine gender as the default for *ayt’* ‘mouse.’ This is perhaps a testament to the exceptional nature of these nominals in that they might not be acquired consistently.

- (10) a. magav magav-et
 wiper towel
- b. maxsan maxsan-it
 warehouse magazine
- (Ritter 1993:796, (2))

For example, *magav* without any suffixes has the meaning ‘wiper,’ but *magav* with a feminine suffix *-et* has a different meaning, i.e., ‘towel.’ In Amharic, adding *-it* to an inanimate masculine noun results in a diminutive interpretation of the nominal (i.e., adding an interpretation that the nominal is small and/or cute, among other readings; see Leslau 1995:167-168), not a new, semantically related nominal.² Moreover, removing the feminine suffix from either *ṣārūrīt* ‘spider’ or *aṣurīt* ‘whirlpool’ does not result in a related masculine noun. There is in fact no such word as *ṣārūr* in Amharic, and *aṣur* is a verbal form (the masculine imperative of the verb *ṣorū* ‘to turn,’ which is morphologically related to *aṣurīt*).

However, *-it* does seem to be capable of converting some **animate** nominals to **natural** female gender. For certain animate nouns, adding *-it* causes the noun to denote a female entity.

- (11) a. ḥdṣ ḥdṣ-it
 boy, child girl
- b. mānāk^{wse} mānāk^{ws}-it³
 monk nun
- c. ṣimagille ṣimagill-it
 old man old woman
- d. muṣṣirra muṣṣirr-it
 groom bride (compare to (4)b)

However, this is not a highly productive process since it is not an option for all animate nouns, e.g., **tāmarit* ‘female student’ and **ḥakīm-it* ‘female doctor.’

Amharic also has a set of gender ‘specifiers’ that indicate natural gender (Leslau 1995:164-166, Cohen 1970:76, Hartmann 1980:279). For humans, the specifiers are *wand* for males and *set* for females.

- (12) a. wānd ayat male grandparent ‘grandfather’
 b. set ayat female grandparent ‘grandmother’

There are a few additional specifiers only for animals which still denote either male or female gender, but the lexical items are different from those used for humans. There seem to be two options for analyzing the gender specifiers. First, they could be adjectives, like ‘male’ and ‘female,’ but more differentiated than in, say, English. Second, they could be nominal classifiers, similar to those found in Mayan languages, Bantu languages and many other language families (see Allan 1977 for an overview of classifier systems). There is

² This statement is true under the assumption that, in a nominal like *bet-it-u* ‘the small house’, *-it* is the feminine/diminutive suffix and not part of the definite article. In Leslau 1995, it is claimed that the feminine definite marker may surface as *-itu* (also *-itwa*), but it is unclear whether these forms are truly feminine definite markers or combinations of *-it* + a definite article. A consultant simply judged all examples where *-it* was added to an inanimate noun ungrammatical, regardless of any subsequent definite marking, which is compatible with the general point here that the feminine suffix cannot ‘convert’ inanimate nominals to feminine gender in Amharic.

³ The final vowels in these nouns are deleted when the *-it* suffix is added in order to avoid hiatus. This is similar to other kinds of nominal suffixes, which also trigger deletion of the final vowel on the stem which they attach to (Leslau 1995:36).

some indication that the adjective analysis is correct. The specifiers exhibit the same morphosyntactic behavior as adjectives (e.g., the definite marker attaches to them; Leslau 1995:65) and they can be predicates of a copular clause (see (5)). Also, most classifier systems operate over several more criteria than gender (animacy, shape, size, etc.) and do not usually co-exist with a masculine/feminine two-gender system. I thus assume the gender specifiers are adjectives, and do not treat them further.

To sum up, it is useful to consider how Amharic fits into Corbett's (1991) classification of the systems of gender assignment in the world's languages. Corbett draws a fundamental distinction between semantic systems of assignment, where most nouns are assigned gender according to semantic principles, and formal systems of assignment, where most nouns are assigned gender according to morphological or phonological principles. Both kinds of systems are found in a variety of languages and language families, and semantic and formal criteria can overlap in a particular language.

Amharic is best described (I believe) as either a "strict semantic" or a "predominantly semantic" system (Corbett 1991:13), where the gender of most nouns is assigned via semantic principles but there are certain sets of exceptions.⁴ It is certainly not the case that phonology or morphology determine the gender of a noun in Amharic -- there are no phonological regularities about which nouns are assigned which gender and the only morphological indication of gender (the *-it* suffix) is neither necessary nor sufficient to deduce gender.

The semantic principles are clear, though. If a nominal refers to a male animate, then it has masculine gender. If a nominal refers to a female animate, then it has feminine gender. These generalizations seem to be virtually exceptionless.⁵ Moreover, for the vast majority of cases, if a noun is inanimate, it is masculine gender (or, better, it lacks gender entirely and masculine is assigned as a morphological default; see analysis below). There are, of course, a small number of exceptions to this principle for inanimates -- there is a 'residue' of inanimate nouns that are assigned feminine gender and presumably simply memorized by the language learner. Finally, if a nominal refers to an animate whose gender is unknown, the nominal is masculine (default) in the vast majority of cases. However, there are a small number of animals for whom feminine is used in this situation and which (again) must be memorized.

In the analysis below, I attempt to capture the Amharic principles of gender assignment. The 'residue' feminine nominals whose gender must be learned will be stipulatively marked feminine. However, for the other nominals, their gender will in fact be derived from their natural gender (in the case of gendered animates) or by having 'masculine' gender be the default morphological form of nominals that lack gender (in the sense of either lacking natural gender for the animates, or not having a gender feature at all in the case of the inanimates). There are many analytical tools that one could use to capture these principles, and the next section is spent narrowing down the options.

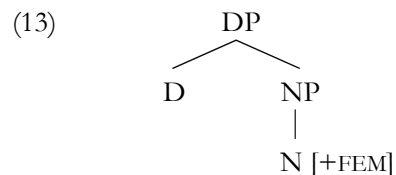
3 THE REPRESENTATION OF GENDER: FIRST STEPS

In most languages with binary masculine/feminine gender systems, the masculine has been argued to be the unmarked or default gender for a variety of empirical reasons (it is not morphophonologically marked in all or most cases, there are more masculine roots in the language, non-nominals that are treated as nouns take masculine gender, etc.; e.g., Roca 1989 and Harris 1991 for Spanish, Levy 1983 for Hebrew, Nelson 1998 for German, etc.; see also discussion in Sauerland 2008). This has led to a general assumption that the gender feature is $[\pm \text{FEM}]$, with feminine gender being the marked option. Since masculine gender is the default in Amharic as well for the vast majority of cases (see above for the empirical details), I will adopt this assumption.

⁴ It is unclear to me whether the number of exceptional nouns in Amharic is enough to nudge it down to only 'predominantly semantic.'

⁵ The major exception is the use of diminutive forms (which are all feminine) to refer to male animates (with some kind of emotional impact: affection, mockery, etc.). The diminutives in Amharic are fascinating, but there is unfortunately not space to discuss them in full. See Leslau 1995:167-169 and some discussion in [Section 6](#).

Intuitively, gender seems to be an inherent property of nominals. The gender of a nominal is generally consistent no matter how it is inflected, e.g. for case, number, definiteness, etc.⁶ This intuition has led to many proposals that the feature [\pm FEM] is in the lexical entry for any given noun, i.e., on the nominal head N in the syntax (see e.g., Harris 1991, 1996, Carstens 2000, Koopman 2006, among many others).

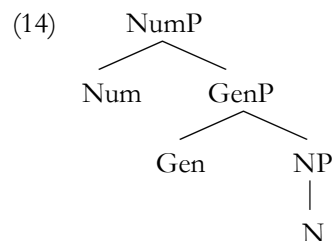


However, it has occasionally been claimed that the gender feature can or must originate elsewhere, and I discuss these proposals in the next section.

3.1 GenP and NumP

In Picallo 1991, it is argued that the gender morphology on a nominal stem heads its own projection, i.e., Gen(der)P. In Ritter 1993, it is proposed that a gender feature can be a part of the Num head within NumP, which typically houses number inflection. I argue that neither of these proposals is supported by sufficient evidence.

In Picallo 1991, it is proposed on the basis of data from Catalan that GenP immediately dominates NP and that its head is the source of gender inflection for all nominals. Unlike Amharic, Catalan assigns gender (at least partially) according to phonological criteria, so that GenP could have (in this theory) some phonological content (e.g., an *-a* suffix for feminine gender, an *-o* suffix for masculine gender). Picallo argues for an articulated DP structure where NumP, which contains number inflection, dominates GenP.



She notes that if N raises through Gen to Num, Gen is successfully predicted to be closer to the stem than Num (Stem-Gen-Num). However, this is not an argument for a gender projection per se. Gender morphology will also be closer to the stem than number morphology if gender is simply a feature on N, and N subsequently moves to Num.

Picallo also notes that when a noun has multiple arguments, the noun precedes all the arguments, e.g., *novelles* ‘novels’ in (15).

- (15) les *novelles* d’en Pere de Nabokov
 the novels of Pere of Nabokov
 (Picallo 1991:283, (7a), NB: Pere is a male name)

She assumes that *en Pere* ‘Pere’ is base-generated in the specifier of GenP and *Nabokov* is base-generated in the specifier of NP. She then argues that the noun raises past these arguments through Gen to Num, and this results in the N-initial order. However, the starting assumptions about argument position here are

⁶ Although there are cases where plural number causes a switch in the gender of a noun, e.g., the so-called ambigenics in Romanian (Farkas 1990, Acquaviva 2008ab) and gender polarity in Somali (Lecarme 2002).

unmotivated. It seems equally likely that *Nabokov* would be the complement of the noun *novelles* and that *Pere* would be in Spec,NP -- in which case, the noun would travel past them both by moving to Num, without a need for GenP at all.

- (16) [NumP Num [NP Pere [N' nouvelles [DP Nabokov]]]]
-

Even if two specifier positions turned out to be required for data like (15), there is no necessary link between the existence of an extra functional projection to house the specifier and that functional projection being the source of gender morphology. Overall, the same predictions could be achieved concerning all the Catalan data without GenP being present at all. Picallo essentially presents arguments for N-to-Num raising in Catalan, but not necessarily for GenP.⁷

Ritter (1993) argues explicitly against GenP⁸, proposing that gender can be fully accounted for as a feature either on N or on Num depending on the language (N for Hebrew, Num for the Romance languages). The linchpin of her argument against GenP for Hebrew is that a structure like (14) indicates that gender is an inflectional affix on the noun, just like number. In Hebrew, though, gender suffixes are clearly derivational, e.g., a new, related nominal can be formed by adding a feminine suffix to a masculine inanimate noun (see Section 1).

However, for Romance languages, Ritter claims that changing the gender of an inanimate noun does not systematically result in a new related noun. She thus concludes that gender is inflectional in Romance, but argues that it is housed in NumP and not GenP for two reasons: (i) plural nouns in Romanian switch genders (the so-called ambigenics), i.e., Num must have gender and number specifications anyway, and (ii) in Walloon, gender and number are spelled out together as the realization of a Num head, separately from a nominal head that is not inflected for either.

I will not take issue here with the claim that gender is in NumP for Romance, although derivational gender morphology seems to be more pervasive in the Romance languages than Ritter reports (see e.g., Lowenstamm 2008 on French) and there may be other ways to account for ambigenics in Romanian (see e.g., Farkas 1990, Acquaviva 2008ab). However, the empirical observations that Ritter uses as the basis of her arguments (besides the fact that inanimates cannot switch genders) do not hold in Amharic. First, there is no gender switching in plural nouns like in Romanian. Adjectival agreement with a plural nominal does not often encode gender distinctions in Amharic, but when it does, it encodes the same gender the nominal has when it is singular.

- (17) a. afrikawi-yan wändɨmm-otʃʃ-atʃʃɨn-ɨmm b. afrikawi-yat ɨhɨt-otʃʃ-atʃʃɨn
 African-M.PL brother-PL-our-TOP African-F.PL sister-PL-our
 our African brothers⁹ our African sisters¹⁰

wändɨmm 'brother' is a masculine noun, and when pluralized and modified by an adjective that ends in the suffix *-awi*, there is plural masculine agreement on the adjective. In the same way, when an adjective that ends in *-awi* modifies the plural feminine noun *ɨhɨt* 'sister,' it takes plural feminine agreement.¹¹

As for the data that Ritter uses from Walloon, a little more detail is needed to show how it is not relevant for Amharic. Ritter follows Bernstein (1991) in assuming that, because nouns are not inflected for

⁷ In more recent work, Picallo (2006, 2007) has developed a different analysis of gender. This analysis is discussed briefly in Section 5 and in Chapter X.

⁸ See also di Domenico 1997 (p. 136; as cited in Picallo 2007:9) and Alexiadou 2004 for more recent approaches to gender that explicitly argue against GenP.

⁹ From <http://ectvamharic.blogspot.com>, accessed 10/1/08.

¹⁰ From <http://etiopiainitalia.blogspot.com>, accessed 10/1/08.

¹¹ It should be noted that plural definite markers and demonstratives do not encode gender, which may misleadingly appear to indicate that the gender of the plural nominal is masculine or not present. The evidence from (21), though, shows that the gender of the nominal is retained, and simply not reflected in the plural forms of these elements.

number in Walloon, they do not move to Num. She also assumes (again following Bernstein) that the feminine plural marker *ès* on adjectives is the realization of a Num head with inherent gender features.

- (18) les belès feyes
the pretty girls
(Ritter 1993:801, (13a))

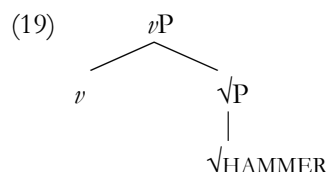
Thus, number and gender features can be realized as one morpheme when the noun is not inflected for number (the *-es* on *fey* ‘girl’ is purely orthographic), and this is a prediction of Ritter’s theory that gender is a feature of Num. However, nouns in Amharic are inflected for number, so the argument from Walloon cannot even get started in Amharic.

As discussed in Section 1, gender assignment in Amharic mostly aligns with natural gender but sometimes aligns with what seems to be inherent, or grammatical, gender (e.g., feminine default animates like *ayt’* ‘mouse’). It is difficult to see how to incorporate any generalizations about natural vs. grammatical gender into a NumP analysis since there is only one gender feature. Moreover, it is proposed in Chapter X that Num is associated only with regular plural morphology, and regular plural morphology never varies according to the gender of the nominal (as it might be expected to if Num also housed gender features). Overall, then, there is no evidence indicating that gender features are located on Num in Amharic.

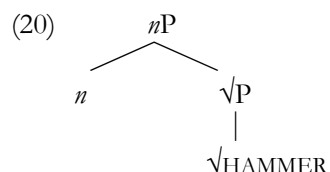
3.2 Gender on the Root or on *n*

In this section, I return to the widely-held intuition that gender features are part of the nominal head N. In Distributed Morphology, the assumption of category-neutral roots has led to a more detailed structure of lexical heads like N. The increase in complexity ultimately allows for a more nuanced analysis of gender in Amharic that makes explicit the relationship between the natural gender (male or female) and the grammatical gender of nominals.

In the Distributed Morphology literature (see e.g., Marantz 1997, 2001, Arad 2003, 2005, Embick and Noyer 2007, Embick and Marantz 2008), the idea has been pursued that all lexical categories are made up of a category-neutral root and a category-determining head.¹² For example, a verb like *hammer* consists of a root $\sqrt{\text{HAMMER}}$ that could theoretically be either a noun or a verb, and a functional head *v* that ‘verbalizes’ it.



This also goes for nouns, which consist of a root and the nominalizing functional head *n*.



(20) results in the nominal *hammer* ‘a tool for pounding nails’ whereas (19) results in the verb *hammer* ‘to pound in (something)’. The upshot of this approach is that there now seem to be two possible heads on

¹² Although it should be noted that the idea was not entirely new (see e.g., van Riemsdijk 1990 on *n*), nor do only Distributed Morphologists subscribe to it (see e.g., Lowenstamm 2008). Cf. also Borer 2005 for a similar approach, although Borer (2005:20-21) argues against the specific Distributed Morphology analysis adopted here.

which the gender feature could be: the root or *n*. In Section 3.2.1, I examine whether the gender feature is located only on the root, and in Section 3.2.2, whether the gender feature is located only on *n*.

3.2.1 Gender on the Root

It is often assumed that roots have no syntactic or semantically active features, i.e., that they do not possess any features that drive syntactic operations or that are interpretable at LF (see e.g., Embick and Noyer 2007:295 for a clear statement of this assumption; see also Borer 2005 for the assumption that roots have no features whatsoever). However, they do have so-called diacritic features, which are non-phonological and encode root-particular quirks like inflection class (see e.g., Embick and Halle 2005:46). Inflection class and gender should be distinguished empirically and theoretically (as influentially argued in Harris 1991; see Alexiadou 2004 for a recent perspective, and Section 5 for discussion). However, it could be that gender would also be a type of diacritic feature that can occur on a root, especially for a language like French where gender is often perceived as an arbitrary property of nominals. However, there are several reasons to think that gender should not only be on the root in Amharic.

Recall that some animate nouns (the same-root nominals) do not change in form depending on gender.

- (21) a. hakim-u doctor-DEF ‘the male doctor’
 b. hakim-wa doctor-DEF.F ‘the female doctor’

If gender were encoded on the root, there would have to be two roots for all the same-root nominals, one [-FEM] and the other [+FEM].

- (22) √HAKIM [-FEM]
 √HAKIM [+FEM]

This might be acceptable if only a small number of nominals were ambiguous, but ambiguity is in fact quite common. As noted above, different-root nominals tend to be limited to either kinship terms or select domestic animals, whereas the ambiguous nominals can refer to almost every other kind of animate human and animal (e.g. *profesor* ‘professor,’ *täkäsaf* ‘defendant,’ *ämmät* ‘cat’). An even more troubling consequence of having two separate roots for ambiguous nominals would be that the two roots would not be morphologically related. There would be a large amount of roots in Amharic that would be coincidentally identical in terms of morphophonology and meaning, except for their gender. This kind of massive repetition of information is clearly undesirable.

Another problem with a root approach is that root diacritics are assumed to be non-phonological. There is no process of vocabulary insertion whereby the diacritic feature is given morphophonological content at Vocabulary Insertion, i.e., root diacritics are not directly spelled out. For the most part, gender is not marked morphologically on the noun in Amharic. However, there is the feminine *-it* suffix, which optionally is added to certain animate, female nominals (see (11)). If this suffix is to be associated with [+FEM], the assumption that root diacritics are non-phonological will have to be abandoned, considerably complicating the status of features on roots.

Finally, recall that certain animates have feminine grammatical gender if their natural gender is not known.

- (23) ayt’-wa
 mouse-DEF.F
 the mouse whose gender is unspecified (or the female mouse)

However, they can have masculine gender if the referent for the animal is known to have natural masculine gender.

- (24) ayt'-u
 mouse-DEF
 the male mouse

Thus, feminine grammatical gender is 'overridden' by masculine natural gender, and this overriding is very difficult to treat in a root-based approach to gender. Even if each of these animals were associated with two roots, one masculine and the other feminine (as discussed w.r.t. (22)), there is no straightforward way to ensure that the feminine root is inserted when the natural gender is not known (or that the masculine root is **not** inserted when the natural gender is not known). Essentially, a root approach cannot capture the interplay between natural and grammatical gender that the facts demonstrate. I conclude for all these reasons that a root-based approach to gender is not viable.

3.2.2 Gender on *n*

Similar problems are faced by an analysis of gender where the gender feature is only on *n*. To the best of my knowledge, it has not been specifically proposed in previous literature that gender is a diacritic feature on the root (at least insofar as gender has been treated as distinct from inflectional class). However, the idea that gender is a feature on *n*, the 'nominalizing' head, has received some support. Working with data from French and Yiddish, Lowenstamm (2008) assumes that *n* has gender features, with there being as many versions of *n* in a particular language as the language has genders. Kihm (2005) proposes that *n* is where Class is located, Class being a kind of supercategory that includes gender, noun class as found in Niger-Congo languages, and numeral classifiers like those found in Chinese (see also Ferrari 2005 for an approach very similar to Kihm's).

Acquaviva (2008b) argues against diacritic features on roots in general, and lays out explicitly how *n* having a gender feature relates to roots. Roots must have licensing conditions, such that they can only be inserted in the context of an *n*[-FEM] or *n*[+FEM].¹³ For example, in Amharic, the roots associated with *mākina* 'car' or *set* 'woman' could only be inserted in the context of *n*[+FEM]. As Acquaviva (2008:9) points out, this allows for a simple treatment of same-root nominals, e.g., *tāmari* 'student' in Amharic. They simply have no licensing conditions and are thus compatible with either gender.

However, the Amharic animals that take feminine gender when the natural gender is not known (e.g., *ayt'* 'mouse,' see (23) and (24)) are still difficult to deal with, mostly because the *n* account only allows for one dimension of gender: natural or grammatical. To see why this is so, assume that the gender feature on *n* corresponds to natural gender for animates, [+FEM] being female and [-FEM] being male. When the natural gender is unknown, assume that *n* lacks a gender feature altogether.

- | | | |
|------|-----------------|-----------------------|
| (25) | <i>n</i> [+FEM] | Female Natural Gender |
| | <i>n</i> [-FEM] | Male Natural Gender |
| | <i>n</i> | No Natural Gender |

When there is no gender feature on *n*, there can be no gender agreement between the nominal and the categories it usually agrees with, e.g., the definite marker. These categories then must be spelled out in their least marked form, i.e., masculine in Amharic (see (5)).

To give a specific example, the vocabulary items which can spell out the Amharic definite marker in (26).

- | | |
|------|--------------------------------------|
| (26) | D, [DEF], [+FEM], [-PL] ↔ <i>-na</i> |
| | D, [DEF] ↔ <i>-n</i> |

¹³ Acquaviva (2008) assumes that roots are inserted post-syntactically, in accordance with the Late Insertion hypothesis of Distributed Morphology (and earlier theories) but against more recent work in Distributed Morphology (see e.g., Embick 2000, Embick and Noyer 2007:296). See Section 4.2 for further discussion.

In Distributed Morphology, the Subset Principle (Halle 1997) governs the insertion of vocabulary items.

(27) **Subset Principle**

- i) The phonological exponent of a vocabulary item is inserted into a position if the item matches all or a subset of the features specified in that position.
- ii) Insertion does not take place if the vocabulary item contains features not present in the morpheme.
- iii) Where several vocabulary items meet the condition for insertion, the item matching the greatest number of features specified in the terminal morpheme must be chosen. (Halle 1997:428)

Looking again at (26), if the definite marker has feminine and singular features (presumably obtained via agreement with the nominal), then it must be spelled out as *-wa*, since *-wa* matches a greater number of its features than *-u*. The vocabulary item *-u* is underspecified and hence the elsewhere case -- every non-feminine and/or non-singular definite marker will be spelled out as *-u*. So, when there is no gender feature on *n* (and thus no gender feature transferred to D), the ‘masculine’ form *-u* must be inserted since *-wa* contains a feature that would not be present on D: [+FEM]. Under this account, then, ‘default’ gender (when there is no gender feature on D) is never predicted to be feminine. However, this is incorrect for animals like *ayt* ‘mouse’ that are feminine when their natural gender is not known. Since there is no other source of gender in this analysis, there is simply nowhere for the feminine gender assigned to the relevant animals to have come from.¹⁴

Essentially, the feminine default animals present a case where both grammatical gender and natural gender will need to be referred to in the analysis. However, this is very difficult to accomplish if there is only one element that the gender feature attaches to, i.e., either the root or *n*. I conclude that both the root and the *n* analyses are unsatisfactory with respect to the Amharic gender system.

4 THE REPRESENTATION OF GENDER : GRAMMATICAL GENDER AND NATURAL GENDER

There remains (at least) one more option for analyzing the gender system of Amharic: a gender feature on **both** *n* and the Root, encoding the difference between natural gender and grammatical gender respectively. Having two sources for gender may seem to complicate the gender system. However, in Amharic, both natural gender and grammatical gender are clearly manifested in the data, especially with respect to the ‘feminine default’ animals. Natural gender even has its own morphophonological realization for certain roots (*-it*). In the remainder of this section, I lay out the fundamentals of an analysis that treats both grammatical and natural gender, and develop the details of the analysis using licensing conditions on roots (similar to the *n* analysis above).

4.1 Fundamentals of the Analysis

As stated above, the heart of the present proposal for the Amharic gender system is that both the root and *n* have a gender feature [+FEM]. However, on the root, the gender feature corresponds to grammatical gender, indicating the arbitrary sorting of all the roots in a language into two types. I assume that feminine roots are marked [+FEM] and this feature is uninterpretable on roots. However, masculine roots do not need to have a gender feature. As shown in (26), the vocabulary items inserted for agreeing elements (e.g., definite markers) can be structured such that a feminine marker is inserted in a context marked specifically as feminine (and singular), and the ‘masculine’ marker is inserted everywhere else. There is thus no need for a [-FEM] feature on roots. Note also that the sole morphophonological realization of gender in Amharic is associated with natural (not grammatical) gender, i.e., the gender feature on the root does not have

¹⁴ Note that the vocabulary items in (26) successfully predict the form of the definite marker in a wide range of contexts (to be discussed in more detail in Chapter 4), including the masculine default gender that was shown above in (5). Similar rules can also be constructed for the other categories that participate in gender agreement, namely, demonstratives, certain kinds of adjectives, etc. (see Chapter 4).

any morphophonology in line with previous assumptions about features on roots (see discussion in Section 3.2.1).

On *n*, the gender feature is interpretable and corresponds to natural gender, that is, biological sex. I assume *n* comes in three varieties with respect to this feature.

- (28) *n* [+FEM] Female natural gender
n [-FEM] Male natural gender
n No natural gender

Note that there is a three-way contrast between having a positive value for the feature (female), having a negative value for the feature (male) and not having the feature at all (i.e., inanimate or natural gender unknown). Despite this, the same underspecified vocabulary items for the definite marker from above may be used, and they are repeated in (29).

- (29) D, [DEF], [+FEM], [-PL] ↔ *-na*
D, [DEF] ↔ *-n*

Suppose that *n*[-FEM] enters into an agreement relation with D, resulting in the feature bundle: (D, [-FEM]). The vocabulary item *-na* cannot be inserted in this case because it contains a feature not present in the morpheme, i.e., the feature [+FEM]. However, since the vocabulary item *-n* is underspecified (i.e., does not have very many features), it may be inserted when D has a [-FEM] feature. Hence, all agreeing elements that contain [-FEM] will be spelled out as the underspecified masculine forms. Overall then, nominals which lack a gender feature and those which have a [-FEM] feature will be spelled out with ‘masculine’ Vocabulary Items.

Since there may be two gender features per *n*P (one on *n* and one on the root), it is necessary to determine which feature is used for the purposes of agreement, i.e., which feature the definite marker, demonstratives, etc. agree with, and how this can be captured. Empirically, if a nominal has natural gender, its natural gender is the agreeing gender.

- (30) a. *abbat-u* father-DEF ‘the father’
b. *tāmari-wa* student-DEF.F ‘the (female) student’

However, if a nominal lacks natural gender (whether natural gender is unknown or whether the nominal is inanimate), its **grammatical** gender is the agreeing gender.

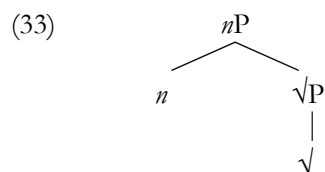
- (31) a. *mākina-wa* car-DEF.F ‘the car’
b. *agār-itu* country-DEF.F ‘the country’
c. *ayt’-wa* mouse-DEF.F ‘the mouse whose gender is unspecified’ (or the female mouse)

Essentially, the grammatical gender ‘emerges’ when there is no natural gender.

In terms of the *n* and root analysis developed here, these generalizations can be stated as in (32).

- (32) **Gender Principle**
a. If *n* has a gender feature, the agreeing gender is the gender of *n*.
b. If *n* has no gender feature, the agreeing gender is the gender of the root.

However, simply stating them and appending them to the analysis seems stipulative. It would be better if these generalizations were derivable from some independently established principle(s), and I argue that they are. Consider the syntactic structure that has been assumed for a nominal.



For many nominals, the agreeing gender is natural gender and the feature that corresponds to natural gender is on n . So, the initial generalization is perhaps better stated as the agreeing gender is the gender of the topmost terminal node in the structure corresponding to the noun ($n + \sqrt{\quad}$), i.e., n .

As shown in (31), sometimes the grammatical gender of a nominal is the agreeing gender, but crucially, this only occurs if the nominal has no natural gender. Recall that I have proposed that grammatical gender is found on the root. With this in mind, the generalizations about agreeing gender can be stated in terms of a search process looking down the tree from above nP : the agreeing gender is the gender that is present on the highest terminal node in nP . When present, natural gender must be the agreeing gender since it is on n and higher than the root. However, if n lacks gender, the search process can continue downward and find the grammatical gender on the root. The approach here is very reminiscent of an unvalued probe searching downward in a tree for a valued goal to enter into the Agree relationship with, i.e., it is very reminiscent of agreement from a Minimalist perspective. The formal account of gender agreement will be dealt with in Chapter 4, but for now, it is sufficient to note that the generalizations concerning the agreeing gender can most likely be derived from the well-known principles of Agree, which is a very welcome result. For the sake of brevity, I will refer to the generalizations about agreeing gender as the Gender Principle in the following discussion.

Before proceeding to discuss the licensing conditions on roots in this analysis, a brief digression is necessary on the formal aspects of the gender feature system just presented. Following convention, I have been using the term ‘feature’ to refer to what is more technically a feature specification, i.e., a pairing of a feature (FEM, PL, etc.) with a value [+/-, in these cases]. For the root and n , I have claimed that most roots and one type of n lack a gender feature altogether. To be more specific, the claim is that in the bundle of features that comprise this particular type of n , and in the relevant roots, there is no such feature [FEM] present (similar to how n lacks a [DEF] feature, or a tense feature, etc.). However, another way this claim can be interpreted is that these roots and this type of n have a 0 value for the gender feature [0FEM], i.e., they are unspecified for gender. Redundancy rules at PF could then be posited to assign the unmarked feature value to these elements, i.e., [-FEM], masculine gender.

However, there are two reasons I will not pursue an unspecified approach to gender here. First, the redundancy rule will assign masculine gender to each nominal that either (a) inherently has no natural gender (i.e., all the inanimates) or (b) whose natural gender is unknown. As will be discussed in more detail below, inanimate nominals are only licensed under the n that that is [0FEM] (the n that has no gender feature in terms used below) since they do not have natural gender. The redundancy rule will then assign this n masculine gender. According to the Gender Principle, then, all inanimate nominals will trigger (only) masculine agreement. This is clearly incorrect.

- (34) *mākina-wa*
 car-DEF.F
 the car

The feminine inanimate nominal *mākina* ‘car’ appears with a feminine definite marker. However, the unspecified analysis incorrectly predicts that, since the redundancy rule assigns masculine gender to n , these animals should be realized as masculine.

Moreover, under certain syntactic theories (e.g., Legate 2002, Pesetsky and Torrego 2007), unvalued features like [0FEM] cause the derivation to crash unless they are valued (undergo Agree) before the derivation

is sent to the interfaces.¹⁵ Unfortunately, the unvalued gender feature on certain inanimate roots will not always be capable of being valued. Since inanimates do not have natural gender, their unvalued gender feature will always appear in the context of a *n* that has a similarly unvalued gender feature. The unvalued gender feature on *n* cannot value the unvalued gender feature on the root, and no other category immediately presents itself as a possible source for valuing the root's gender feature. The syntactic theories in question also cannot just be set aside -- not only is Pesetsky and Torrego 2007 rapidly becoming the standard way of thinking about features in Minimalism, but the systems established in both Legate 2002 and Pesetsky and Torrego 2007 prove to be extremely useful in accounting for both gender and number in Amharic (see later in this chapter, and Chapter X).

An unspecified analysis could possibly overcome the obstacles laid out briefly here. However, there is another way to configure the features that has none of these obstacles, i.e., if the gender feature is simply not a part of the feature structure of the *n* in question, and similarly not found on the relevant set of inanimate roots. Since the two options seem otherwise identical, I will not delve further into an unspecified analysis of gender.

4.2 Licensing Conditions

So far, the analysis consists of (a) assumptions about the distribution of gender features and (b) the (independently derivable) Gender Principle. The final piece is the licensing conditions on roots, like the licensing conditions from the *n* analysis of gender discussed above. It is easiest to see how they come into play by going through each type of nominal in turn, and I begin with inanimate nominals.

Inanimate roots are specified as either [+FEM] or having no gender feature, with the majority having no gender feature. This is an intuitively satisfactory result -- the vast majority of inanimates are masculine and can thus be unmarked, and the few feminine nouns are clearly marked as special. Inanimates are licensed only in the context of the *n* that has no gender features, since they do not have natural gender. The Gender Principle then correctly predicts that the agreeing gender will be that of the root, i.e., a [+FEM] root will result in feminine agreement. To be clear, though, for 'masculine' inanimates (which have no gender feature) it is more accurate to say that there is simply no gender agreement and the agreeing elements are in the default form.

Different-root nominals have strict licensing conditions, e.g., *abbat* can only be licensed under *n*[-FEM], resulting in masculine agreeing gender, and *ɨmat* can only be licensed under *n*[+FEM], resulting in feminine agreeing gender. In contrast, the same-root nominals (e.g., *tāmari* 'student') have no licensing conditions: they may be inserted under *n*[+FEM] (resulting in female natural gender and feminine agreeing gender), under *n*[-FEM] (resulting in male natural gender and masculine agreeing gender) or *n* that has no gender features (resulting in no natural gender and default, i.e., 'masculine,' agreeing gender).

Certain different-root nominals that have masculine agreeing gender can refer either to male entities or to entities whose natural gender is unknown. For example, as mentioned above, the nominal *säw* can mean either 'man' or 'person' (human of either gender) as opposed to its feminine counterpart *set* which can only mean 'woman.' This phenomenon is present in other languages, cf. German *mensch* and Spanish data in Roca 1989. In terms of the licensing conditions analysis, it means that nominals like *säw* are licensed under both *n*[-FEM] and *n* such that they always have masculine agreeing gender but can refer to either male entities or entities whose natural gender is unknown.

The feminine default animals, e.g. *ayt* 'mouse,' are easily accounted for under this analysis. The roots for these animals are all [+FEM], but they crucially have no licensing conditions. When such a root combines with *n*[+FEM], female natural gender and feminine agreeing gender is the result. When it combines with *n*[-FEM], male natural gender and masculine agreeing gender is the result. The key case is when these roots combine with *n* that has no natural gender -- in this case, the animal's natural gender will be interpreted as unknown, but since the root is specified for feminine gender, the Gender Principle dictates that the agreeing gender will be feminine.

¹⁵ This seems to conflict with the (much older) idea that morphophonological redundancy rules fill in the values of unspecified features. To the best of my knowledge, though, I do not know of any research that addresses this issue.

Finally, to account for the *-it* suffix that indicates female natural gender (see (11)), one can simply assume that $n[+FEM]$ (female natural gender) may be spelled out as *-it* in the context of certain roots. A summary of the assumptions and predictions of the licensing conditions analysis is below.

- (35) **Inanimate Nominals**
 Root: [+FEM] (e.g., *mākina* ‘car’) or no gender feature (e.g., *bet* ‘house’)
 Licensed under: n
 Agreeing gender = grammatical (root) gender
- (36) **Different-Root Nominals**
 Root: no gender features (e.g., *ħħ* ‘sister,’ *wändħmm* ‘brother’)
 Licensed under: $n[+FEM]$ or $n[-FEM]$ depending on the root
 Agreeing gender = natural gender
- (37) **Same-Root Nominals**
 Root: no gender feature (e.g., *tāmari* ‘student’)
 No licensing conditions
 Agreeing gender = natural gender or masculine (default)
- (38) **Feminine Default Nominals**
 Root = [+FEM] (e.g. *ayt* ‘mouse’)
 No licensing conditions
 Agreeing gender = natural gender, or feminine gender if natural gender not known (default)
- (39) $n[+FEM] \leftrightarrow -it / _ [\sqrt{LID3}, \sqrt{MÄNÄK^wSE...}]$

The licensing conditions analysis is built on, naturally, licensing conditions, and these should be discussed in more detail before moving on. In Acquaviva 2008b, the licensing conditions on roots are assumed to be conditions on the insertion of roots post-syntactically. However, this raises a problem. The gender feature on the root for a feminine word like *mākina* ‘car’ should be present in the syntax since it can participate in an agreement relation with a definite marker, demonstrative, etc. It may be that such agreement is completely post-syntactic, but that cannot be assumed a priori.

Interestingly, it has been concluded in a prominent strand of research on roots that they are not inserted post-syntactically, but present in the syntactic derivation (Embick 2000 and subsequent work by Embick and colleagues, e.g., Embick and Noyer 2007). Embick (2000) raises the possibility that roots in the syntax are only represented as labels with whatever syntactic features they happen to have, e.g. $\sqrt{365}[+FEM]$ or $\sqrt{294}$ (cf. recent work by Borer). The label would then be phonologically realized post-syntactically. The point of this kind of representation is that it keeps the phonological representation of the root (which is never needed by the syntax) out of the syntactic component.

This idea is workable for the licensing conditions analysis developed here. The gender feature would be visible in the syntax for agreement, and roots would be inserted post-syntactically per contextually-specified Vocabulary Insertion rules like (40)a-d.

- (40) Licensing Conditions on Vocabulary Insertion Rules
- | | |
|--|---------------------------------|
| a. $\sqrt{292} \leftrightarrow abbat / n[-FEM] \text{ ____}$ | Different-Root Nominal |
| b. $\sqrt{546} \leftrightarrow t\ddot{a}mari$ | Same-Root Nominal |
| c. $\sqrt{365}[+FEM] \leftrightarrow m\ddot{a}kina / n \text{ ____}$ | Inanimate Nominal |
| d. $\sqrt{140}[+FEM] \leftrightarrow ayt'$ | Feminine Default Nominal |

In the Vocabulary Insertion Rule in (40)a for a different-root nominal, Root 292 is spelled out as *abbat* in the context of an *n*[-FEM]. Since I assume this is the only Vocabulary Insertion Rule for Root 292, the effect is that Root 292 can only be spelled out in the context of a *n*[-FEM]. (40)b is the Vocabulary Insertion rule for a same-root nominal which has no licensing conditions; it can be realized as *tämari* regardless of what kind of context it is in. (40)c is the insertion rule for an inanimate feminine nominal. The root has a [+FEM] feature, and it is realized only in the context of a *n* that lacks a natural gender feature. Finally, (40)d is the insertion rule for a feminine default nominal which has a feminine feature on the root, but no licensing conditions. The licensing conditions, therefore, which I have been discussing in abstract terms, are actually contextual restrictions on Vocabulary Insertion rules for roots.

Alternatively, it is possible that the licensing conditions are active in the semantics. To take a specific example, it may be that some restriction in the semantic component rules out combinations of inanimate roots and nominalizing heads that have natural gender. In Arad 2005, roots are interpreted in the Encyclopedia in the context of whatever category-defining head they have been merged with in the syntax. So, it may be that there is no Encyclopedia listing for, say, a given inanimate root when it is combined with a *n* that has natural gender, rather than there being no vocabulary item for such an element as in the analysis above.

Whichever interface the licensing conditions are captured at, the licensing conditions analysis has several advantages. It generates all the data without any repetition of roots (as in the root analysis of gender) and successfully treats the feminine default nominals (unlike the *n* analysis¹⁶). It captures intuitions about the grammatical gender of same-root nominals (they do not really have any) and the grammatical gender of feminine default animals (definitely present). However, there is one remaining issue to be worked out.

4.3 Gender and Interpretability

The licensing conditions analysis raises questions about the interpretability of gender features, and the following section is an extended digression on how these questions can be addressed. The gender feature on *n* corresponds to natural gender, and is thus part of the meaning of the nominal and interpretable. However, the gender feature on the root has no semantic impact, and thus should be uninterpretable.¹⁷ Under standard Minimalist assumptions (see e.g., Chomsky 2000, 2001, 2004), uninterpretable features must be checked/valued and then deleted before the derivation is sent to LF or else the derivation will crash. It is

¹⁶ It is possible to construct a successful analysis of gender on *n* if the gender feature comes in two different types on *n*: interpretable (natural gender) and uninterpretable (grammatical gender). However, this analysis requires very complicated and stipulative licensing conditions. For example, a nominal like *ayt'* mouse will be licensed under interpretable [+FEM] (female natural gender), interpretable [-FEM] (male natural gender) but also uninterpretable [+FEM] (default feminine gender). In the analysis developed above, all that needs to be said is that the root associated with *ayt'* has a [+FEM] feature - this root then combines freely with different types of *n* to generate the observed gender facts. However, for those who are firmly against any kind of features being on roots, an analysis where gender is only on *n* represents a viable alternative to the analysis above that maintains the fundamental distinction between interpretable and uninterpretable gender features and has natural gender features in the syntax.

¹⁷ It has occasionally been claimed that grammatical gender is interpretable (most notably in Dowty and Jacobson 1988, but see also Pesetsky and Torrego 2007: fn. 31 and references therein). See Legate 2002:2-3 for some criticism of this idea.

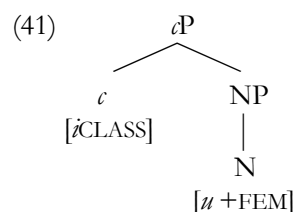
also assumed that checking/valuation happen via agreement (more specifically, the Minimalist relation Agree). The main problem for present purposes is that the uninterpretable gender feature on roots (grammatical gender) is not always found with its interpretable counterpart (natural gender) available to check it before the syntactic derivation is sent to LF. For example, inanimate roots are only licensed under the ν that lacks a gender feature entirely.

In the remainder of the section, I go through several different solutions to the problem of an uninterpretable gender feature, ultimately following Legate 2002 (and, to a lesser degree, Pesetsky and Torrego 2007) in claiming that **unvalued** features (and not uninterpretable features) cause a derivation to crash.

4.3.1 Post-Syntactic Agreement and Class Features

A potential solution to the problem of an uninterpretable gender feature in the syntax is to say that the problematic feature is not present in the syntax at all, i.e., that grammatical gender is a purely post-syntactic phenomenon. Under this approach, roots would have no gender features in the syntax, but post-syntactically certain roots (feminine inanimates) would trigger the insertion of a [+FEM] feature (presumably adjoined to the root itself).¹⁸ An immediately worrying aspect of this proposal, though, is that the gender feature on a root has no morphophonological exponence (ever), so it is suspicious for it to be specifically inserted in the morphophonological component. The only noticeable effect of the gender feature on a root is in fact the agreement that it controls, and this predicts that, if the gender feature is inserted post-syntactically, the gender agreement that is controlled by grammatical gender (i.e., inanimates, feminine default) must be post-syntactic as well. This is an interesting prediction to test, and it requires more attention than there is room for here. It will be explored in-depth in Chapter X (see also Picallo 2007:3-4). For now, I proceed as if the prediction were false and consider other possibilities.

Another kind of solution to the gender/interpretability problem is that the uninterpretable gender on roots is in fact checked/valued in every case, i.e., that there is some head distinct from ν (natural gender) that can check/value the gender feature on roots. In Picallo 2006, 2007, just this kind of approach is developed.¹⁹ Picallo proposes a unified analysis of gender systems (e.g., Romance languages) and noun classifier systems (e.g., some Bantu languages) whereby both gender and noun classifiers are associated with a functional category ϵ which contains an interpretable feature [CLASS] (cf. Kihm 2005, Ferrari 2005). In gender systems, the class category ϵ takes an NP complement whose head N has an uninterpretable gender feature.



Framed in terms of present assumptions, ϵ P will be below ν P and NP is equivalent to the root. The class category ϵ enters into an agree relationship with N/the root and checks/values its uninterpretable gender feature. The relationship between ϵ and N is similar to the relationship between T and V in English in certain frameworks (e.g., Pesetsky and Torrego 2007) -- T is inserted with interpretable tense features, and V is

¹⁸ Gender from this perspective would be a dissociated feature, a feature only inserted at PF (similar to agreement nodes; see Embick 1997, Embick and Noyer 2007:309).

¹⁹ This approach has also been adopted in previous research concerning uninterpretable features on roots. In Embick 2000, it is proposed that certain roots in Latin (the deponent verbs) have an uninterpretable syntactic feature [PASS] that triggers certain morphosyntactic effects associated with the passive. An interpretable version of [PASS] is also found on ν , and presumably uninterpretable [PASS] on the root and interpretable [PASS] on ν can enter an agreement or checking relation so that the uninterpretable [PASS] does not persist to the LF interface.

inserted with uninterpretable tense features that are checked/valued by T and ultimately spelled out on V. Put another way, [+/- FEM] on N is the formal exponent of the interpretable class feature.

However, the *DP* analysis has a major shortcoming. The [CLASS] feature on *ε* is not obviously interpretable, i.e., it is unclear what (if any) contribution it makes to the semantics of the nominal. Picallo (2007) suggests that *ε* is related to non-linguistic processes of entity categorization. It is arguably the case in noun classifier systems that the sorting of nouns into classes conveys something about their meaning of the nominal in terms of size, shape, etc. Recall, though, that the only type of gender relevant to the original problem here is **grammatical** gender -- natural gender is of course interpretable (and Picallo deliberately does not treat natural gender (2007:5-6)). The sorting of nouns into kinds of grammatical gender does not generally convey anything systematic about their semantics, which Picallo recognizes (2007:5-6 and fn. 7), but interprets as only meaning that gender is uninterpretable on the noun. However, it is very unclear how the class feature is interpretable if it “translates to the grammatical system processes of entity categorization” (Picallo 2007:6) and that process of entity categorization has no semantic content. I conclude that the *DP* analysis is not viable.

4.3.2 Questioning Assumptions: Legate 2002

The other kind of solution I consider is of a different sort -- it questions the assumptions that lead to the problem. What if it is not the case that all uninterpretable features must be deleted? Or what if they at least do not have to be deleted by Agree? Two recent proposals for revamping the feature system in Minimalism independently take this approach, and I discuss the first (Legate 2002) in this section.

To start with some background, though, Chomsky (2000, 2001, 2004) suggests that the deletion of uninterpretable features is what motivates the existence of the Agree relation. He takes it as a given that uninterpretable features must be deleted before LF, and then proposes that the only way to do so is via Agree (see e.g., Chomsky 2004:113). He also proposes that uninterpretable features always enter the derivation unvalued (see e.g., Chomsky 2004:116), i.e., without a +/- value (in most cases) for the feature in question. It is unvalued features, then, that are probes and the Agree relation serves the purpose of valuing them w.r.t. a valued version of the same feature.

However, there are several less than ideal aspects about this set of proposals. As Pesetsky and Torrego (2007) note, the obligatory pairing of uninterpretability and unvaluation seems arbitrary: “Why should the *lexicon* couple such distinct properties of lexical items as interpretability (‘Does the item have a message to send to the semantics?’) and valuation (‘Are any syntactically relevant properties of the lexical item left unspecified?’)” (Pesetsky and Torrego 2007: 267). Moreover, uninterpretable features must be deleted before LF, but somehow reinstated (or preserved) in PF because they often surface morphophonologically (e.g., subject agreement is phi features on T). To account for this, Chomsky proposes an operation which deletes features that have been valued during the phase from the part of the derivation that continues to LF, but not from the part of the derivation continuing to PF. As Legate (2002) points out, this operation must look back to a previous stage of the derivation, and this kind of looking-back is usually taken to be a serious flaw in the analysis.

In Legate 2002, a solution to both of these problems is developed. Legate first argues that almost all the syntactic features that have been called interpretable thus far in the literature are actually uninterpretable. Her examples of this includes grammatical gender (understood as part of the interpretable phi-features on a DP in Minimalism, but uninterpretable as discussed above) and tense features on T. With respect to the latter case, she draws on work by semanticists on tense and aspect (e.g., Heim 1994, von Stechow 2002) that documents how the morphological expression of tense/aspect can differ from the semantics of tense/aspect.

Legate then proposes a reconfiguration of the feature system, as laid out in (42).

(42) Types of Features: Legate 2002

Morphosyntactic Features:

uninterpretable, unvalued
uninterpretable, valued

Semantic Features:

interpretable

Uninterpretable features can be either valued or unvalued. Unvalued features are probes and must Agree with a valued version of the same feature or they will cause a crash at PF (presumably because PF cannot spell out a feature that is unvalued). Legate labels this set of features the “morphosyntactic” features since they drive the syntactic computation and are morphologically realized.

In contrast, valuation is not relevant for interpretable features. They do not participate in syntactic operations/relations and are not morphologically expressed, hence they are labeled purely as “semantic.” Uninterpretability is therefore no longer needed to motivate Agree (Agree exists simply to value features) and there is no longer an arbitrary requirement that uninterpretable features be unvalued. As for the relationship between uninterpretable features and the interfaces, Legate proposes that all the morphosyntactic (uninterpretable) features are deleted in the branch of the derivation that continues to LF. There is no need to look back and see which were valued.

Legate’s proposals not only address some problems inherent in Chomsky 2000, 2001 and 2004, but also provide a basis for why grammatical gender in Amharic (+FEM on certain roots) does not trigger a derivation crash. The crucial fact is that the grammatical gender feature is valued (as is grammatical gender in most languages). So, under Legate’s system, it does not cause a crash at PF and is simply deleted before interpretation by the semantics interface.

However, Legate’s system raises some serious questions about both the nature of semantic features and the relationship between morphosyntactic and semantic features that make it less appealing. Consider natural gender features, which are the interpretable counterpart of grammatical gender features. Since they are interpretable, they seem to be part of the semantic set of features for Legate, i.e., syntactically inactive. However, if gender agreement is syntactic, natural gender features must participate in an Agree relation with demonstratives, determiners, etc., which seems to indicate that they are morphosyntactic features, i.e., uninterpretable.

The distinction between morphosyntactic and semantic features seems almost too sharp --- it predicts that, for any feature that participates in agreement (or any feature that is even morphologically expressed!), it either (a) will not have a semantic impact or (b) there will be a second instance of that feature that does have a semantic impact, which seems to be an unnecessary repetition of information. Legate notes that extragrammatical pressures (acquisition, diachrony) may motivate “compatibility” between semantic and morphosyntactic features, but it is unclear how this compatibility should be captured. Nevertheless, the broader characteristics of Legate’s system are insightful: unvalued features are what must be dealt with via Agree before the derivation is sent to the interfaces, and uninterpretable features can simply be deleted in a global fashion on the way to the semantic component. These characteristics are what enable grammatical gender to be part of the derivation without causing a crash.

4.3.3 Questioning Assumptions: Pesetsky and Torrego 2007

These insights can be preserved, but with fewer semantic issues, in a version of the feature system proposed in Pesetsky and Torrego 2007 (henceforth P&T). P&T modify the feature system, but in a slightly different way than Legate 2002. To eliminate the arbitrary relationship between uninterpretability and unvaluation, they fully separate valuation and interpretability and claim that all four resulting combinations are available as different types of features.

(43) Types of Features: Pesetsky and Torrego 2007

- a. uninterpretable, unvalued
- b. uninterpretable, valued
- c. interpretable, unvalued
- d. interpretable, valued

Along with Chomsky (and Legate), they propose that unvalued features are probes and trigger Agree, the novelty being that interpretable unvalued features in P&T can also be probes.

The key part of P&T for present purposes concerns, once more, the motivation for Agree. P&T go along with Chomsky in claiming that Agree is motivated by uninterpretability, i.e., uninterpretable features must enter into an Agree relation with an interpretable version of the same feature as precondition for being deleted. This is not helpful with respect to grammatical gender in Amharic. Since grammatical gender is uninterpretable, P&T predict that it must enter into an Agree relation with an interpretable version of the gender feature, i.e., natural gender. However, as discussed above, this is not always possible.

However, recall Legate's (2002) perspective. In her system, it is **unvalued** features that must enter into an Agree relation as a precondition for morphological realization at PF. Otherwise, they cause a crash. This seems to be the mirror image of a system that relies on **uninterpretable** features to motivate Agree, and Pesetsky and Torrego even briefly note (fn. 17) that "if PF interpretation cannot apply to an element that bears an unvalued feature, consequences similar to those discussed in the text would follow" (P&T 07:274). In fact, with respect to the data discussed later in P&T (raising, *that*-omission), the unvalued approach makes the exact same predictions as the uninterpretable approach to Agree. This is primarily because P&T still maintain that only unvalued features are probes -- hence, every instance of Agree involves an unvalued feature and thus Agree can be straightforwardly motivated by the needs of that feature.

Note also that, in P&T, it must be stipulated that Agree is a precondition for deletion. However, this is not necessary. Since uninterpretability and valuation are separate properties of features, the semantics can look directly at features and see whether they are interpretable or not without any need for a look-back mechanism to see which were valued. P&T even propose that the semantics essentially attempts to interpret all features and then deletes those which are uninterpretable (P&T 2007:290).

Taking stock, the discussion began with the question of how to account for grammatical gender, an uninterpretable feature that does not crash the derivation. A solution can be found by drawing on two related proposals: Legate 2002 and P&T. Legate 2002 contains the crucial insight that unvalued features are what drives Agree, but the feature system proposed therein is not sketched out fully. P&T is compatible with this approach to Agree (although the approach is not overtly adopted), and proposes a better-motivated feature system. Both Legate 2002 and P&T reap the benefits of separating valuation and interpretability by having the semantics (or an operation that transfers the syntactic derivation to the semantic component) identify uninterpretable features directly and delete them.

Under these combined assumptions, an uninterpretable but valued feature like grammatical gender poses no threat. It is not a probe (i.e., it is not unvalued), so it need not participate in Agree. It is valued, so it will not crash the derivation when it is sent to PF. Although it is uninterpretable, it will be removed by whatever mechanism removes such features at or slightly before the semantic component.

4.4 The Briefest of Summaries

A successful analysis of gender in Amharic nominals has been developed in this section that relies on licensing conditions and different locations for natural gender and grammatical gender features within a nominal. However, what are its wider implications and how does it compare to other analyses of gender? These questions are explored in Section 5.

5 PREVIOUS ANALYSES AND BROADER ISSUES

In this section, I begin by discussing some previous analyses of gender that rely on both natural gender and grammatical gender. This leads to a comparison between the analysis here and the previous analyses in terms of the role of the lexicon in any analysis of gender and the presence of natural gender features during the syntactic derivation. It closes with some discussion of several issues related to having a gender feature be on *n*, the so-called nominalizing head that can be associated with nominal inflectional class.

5.1 Literature Review

In the morphosyntactic literature, there is a tendency to either conflate natural gender and grammatical gender or to simply deal with one of them. From a Minimalist perspective, gender is presumably one of the bundle of phi-features that are interpretable on a nominal but uninterpretable on, say, T. However, gender is not semantically interpretable on an inanimate nominal (although see fn. 17). On the other hand, natural gender on an animate nominal is as much a part of the semantic content as number and person (see e.g., Cooper 1983, Heim and Kratzer 1998 for an explicit account of gender for English pronouns). Sometimes, though, natural gender is deliberately set aside from morphosyntactic investigations (Bernstein 1993:117, Picallo 2007:5-6).

The licensing condition analysis formally encodes both natural gender and grammatical gender on *n* and the root, respectively, as the Amharic data requires. Although it is not a common approach, there is some precedent for treating both natural and grammatical gender together, and I review several previous analyses in this vein: Roca 1989, Harris 1991, Riente 2003, Ralli 2002, and Alexiadou 2004.²⁰

Harris 1991 and Roca 1989 are morphological accounts of gender in Spanish nominals. The primary thesis of Harris 1991 is that natural gender, grammatical gender and inflectional class are related but independent properties of nominals, and Roca 1989 comes to roughly the same conclusion. The analyses both rely on the properties of lexical entries and on lexical rules to generate the facts. All inanimate feminine nouns, for example, are marked with an *f* in their lexical entries. The *f* leads to feminine agreement morphology on associated elements like determiners, verbs, etc. In Harris 1991, nouns that have natural female gender are mostly generated in the lexicon from nouns with male natural gender (a central claim Harris makes is that almost all human-referring nouns in Spanish are ‘mated,’ i.e. come in male-female pairs). A separate rule then marks female natural gender nouns with the feminine *f* so that the grammar ultimately treats female natural gender and feminine grammatical gender nouns alike. A similar rule is proposed in Roca 1989 that directly links female natural gender with feminine morphological marking.

In Riente 2003, the focus is on how nouns in Italian are assigned grammatical gender. Riente proposes that gender assignment happens in one of two ways. First, a noun can be inherently specified for gender in the lexicon (and this gender specification is not affected by natural gender, e.g., *guardia* ‘guard’ is always feminine no matter who it refers to). Second, nouns that are not inherently specified for gender are assigned feminine grammatical gender by a redundancy rule if they have a female referent (and masculine default otherwise).

(44) Gender Assignment in Italian

[]_{GEN} → f / _____ female referent
(Riente 2003:7, (9))

Female natural gender thus leads to feminine morphology, similar to Roca 1989 and Harris 1991, although it is unclear at what point in the derivation Riente’s gender assignment rule applies.

Alexiadou (2004) comes to similar conclusion as Riente, although with a broader empirical scope. After examining data from Romance languages, Greek and Hebrew, she concludes that gender is lexically specified in some nouns and not specified in others. She proposes that the nouns where it is not specified are all human, and that they are assigned grammatical gender via an agreement process with the natural gender of their referents. This is a different technical device than Riente uses (agreement rather than a redundancy rule), but it is the same idea. Alexiadou does not go into detail on the mechanics of the agreement, noting

²⁰ See also research on gender and agreement in HPSG for another approach that treats both natural gender and grammatical gender (see e.g., Pollard and Sag 1994: Ch. 2, Wechsler and Zlatić 2003). There have also been several recent conference papers (given as this dissertation was being written!) that have claimed that different types of gender correspond to different projections within the DP, similar to the approach taken here (Steriopo and Wiltschko 2008, Matushansky 2009). I hope to address the similarity of these proposals to the ones made here when public versions of these papers become available.

only that it may be licensed by a specific syntactic configuration or in the Numeration (although it is unclear where or how the natural gender of the referent would be represented in the structure or the Numeration). Ralli (2002) takes the same kind of approach again, although she focuses only on Greek. In her analysis, all nominals have gender features in the lexicon, but some are fully specified and some are underspecified. Nominals with underspecified gender features have the features specified via feature co-occurrence rules.

These analyses all share certain key characteristics. First, they all make a distinction between nominals that are lexically (fully) specified for gender, and nominals that have their agreeing gender assigned to them via some kind of mechanism. Nominals that are lexically specified include inanimates and epicene nominals (epicene nominals refer to animates but do not vary in gender according to the natural gender of their referents, e.g. *guardia* ‘guard’ in Italian as mentioned above). Nominals whose gender is assigned are usually those whose gender corresponds to the natural gender (e.g., *ragazzo/ragazza* ‘boy/girl’ in Italian), and the assignment mechanism usually assigns them their agreeing gender according to the natural gender (e.g., Alexiadou’s agreement process between a nominal and the natural gender of its referent, Harris and Roca’s assignment of *f* to nouns that are [+female] in the lexicon, etc.). There are two major ways, then, in which these analyses differ from the analysis here: (i) they rely on gender specification in the lexicon for certain nouns (ii) they do not represent natural gender in the syntax, and instead assign (or derive) the gender feature that appears in the syntax from wherever else natural gender may be in the grammar. I discuss each of these differences in turn.

5.2 A Non-Lexical Analysis of Gender

The analyses above all assume that, for some nominals, gender is specified in the lexicon. More crucially, they assume there is a contrast between these nominals and other nominals which are not specified for gender and have gender assigned to them. This contrast is also found in Amharic between, say, a feminine inanimate noun (fixed, idiosyncratic gender) and a same-root nominal like *tāmari* ‘student’ whose agreeing gender is variable and always equivalent to its natural gender.

It might seem difficult to capture this contrast without a lexicon, but all that is really needed is a way to mark gender on certain roots, and a way to assign gender to other (gender-free) roots. In the analysis here, gender is stipulated to be a feature on certain nominal roots, and this corresponds to ‘lexically specified’ gender. Nominals whose gender varies depending on natural gender do not have any gender features on their roots, and gender is ‘assigned’ to them insofar as they can combine with a *n* that has [+FEM], [-FEM] or no gender feature whatsoever. The details of how the treatment of natural gender here differs from the other analyses is dealt with immediately below. The main point is just that it is not necessary to have a centralized lexicon in order to capture the gender system of a language. The theory of Distributed Morphology can be viewed as an extended experiment into what consequences result from ‘distributing’ the lexicon throughout the grammar, and gender can now be added to the list of phenomena that can be successfully accounted for in this way.

5.3 Natural Gender in the Syntax

None of the analyses above have natural gender features represented in the syntax. Instead, when natural gender is the agreeing gender for a particular nominal, some kind of mechanism (lexical rule, redundancy rule, agreement) converts natural gender features into the kind of features that are present in the syntax. In the present analysis, though, natural gender features are present in the syntax on *n*. From the perspective of minimizing the number of operations the grammar has to go through, the extra step ‘converting’ natural gender seems unwarranted. If natural gender is the agreeing gender, why not have it be present in the syntax so it can trigger agreement directly? In this section, I consider several reasons why the previous analyses have taken this approach and argue against them, ultimately concluding that natural gender should be present in the syntax. The section ends with some discussion of the semantics of natural gender.

Many of the analyses above divide natural gender and grammatical gender into two different types of features: [+/- FEMALE] for natural gender and [+/-FEM] for grammatical gender. This may lead to the idea that a natural gender feature and a grammatical gender feature would require separate sets of agreeing

elements that would be coincidentally homophonous. In other words, [+FEMALE] must be ‘converted’ to [+FEM] to trigger insertion of [+FEM] agreeing forms -- or else there will be a lot of redundancy among the vocabulary items (or spell-out rules).²¹

However, the separation of natural gender features and grammatical gender features in this way is a false dichotomy. In Minimalism, there are many paired interpretable and uninterpretable versions of the same feature on different morphemes. For example, there may be a [+PL] feature on T and on a DP -- interpreted as plural in the DP, and not interpreted on T. In the licensing conditions analysis, it is proposed that gender -- a single [+/-FEM] feature -- works in the same way. It comes in uninterpretable (root) and interpretable (*n*) versions. On the root, it has no semantic impact, but on *n*, it indicates female or male natural gender. The Vocabulary Items associated with agreeing elements can all contain this single feature (regardless of interpretability), and thus the same morphological realization is predicted for natural and grammatical gender. To give a concrete example, consider again the definite marker.

- (45) D, [DEF], [+FEM], [-PL] ↔ *-va*
 D, [DEF] ↔ *-u*

Both an uninterpretable [+FEM] feature on a root and an interpretable [+FEM] feature on *n* will lead to a [+FEM] feature on D. This in turn will be spelled out as *-va*. Thus, there is no need for an analysis of gender to keep natural and grammatical gender strictly separate *qua* features.

However, there may be an empirical reason why previous analyses have kept natural gender away from the syntax. In all of the languages of inquiry for these analyses, grammatical gender plays a much larger role than natural gender in agreement (and agreement is usually assumed to be syntactic). For example, there are epicene nouns in Italian and Spanish where grammatical gender seems to ‘override’ natural gender (see e.g., *guardia* ‘guard’ (always feminine) or *pilota* ‘pilot’ (always masculine) in Italian, and the list of animals with fixed grammatical gender in Harris 1991:41; see also Alexiadou’s conclusion that in Romance, Greek and Hebrew, only human-referring nouns vary in gender). In Amharic, though, natural gender is typically the agreeing gender, with grammatical gender ‘emerging’ only when natural gender is demonstrably absent, as with the feminine default animals like *ayt* ‘mouse.’

It is possible to account for languages with a greater reliance on grammatical gender in the licensing conditions analysis by saying, at the broadest level, that the role of *n*[+FEM] and *n*[-FEM] is significantly reduced in Italian, Spanish, etc. After all, epicene nouns like *guardia* in Italian and animals whose gender is fixed do not convey anything about the natural gender of their referents.²² The vast majority of nouns in Italian, Spanish, etc. would then only be licensed under *n*. The remaining nouns (human-referring nominals with variable gender) could combine with either *n*[+FEM] or *n*[-FEM], resulting in female or male natural

²¹ It should be noted this is my (mis?)interpretation of these analyses. I am merely speculating about why they may have separated the two sets of features.

²² This raises the question of how to deal with nominals like *mädchen* ‘young girl’ in German which is epicene in that its gender is fixed (neuter) but certainly conveys information about the natural gender of the referent. I assume that this information is contained in the Encyclopedia entry for *mädchen*. The Encyclopedia entry for every noun contains non-linguistic knowledge about the noun, e.g., the Encyclopedia entry for *dog* contains information like ‘has four legs, canine, barks,’ etc. I assume the Encyclopedia entry for *mädchen* accordingly contains something like ‘is a young female.’ Under Arad’s (2005) point of view, Encyclopedia entries are assigned to combinations of root and categorizing head. Licensing conditions on gender (like those discussed above) can be conceived of as falling out from these combinations, e.g., there would be no Encyclopedia entry for a root like *mmat* ‘mother’ combined with a *n*[-FEM]. From this perspective, a nominal like *mädchen* simply has an Encyclopedia entry for (i.e., it is licensed under) a combination of root and plain *n*.

This may point up a weakness of the licensing conditions analysis: does it miss relationships between encyclopedia entries and natural gender features on *n*? More specifically, it seems that if the encyclopedia entry for a root + *n* like ‘mother’ contains the information ‘is female,’ it should somehow be related to the fact that the *n* in this case can only be [+FEM]. I hope to address this in future work, but in the meantime note that this is a problem only for a small subset of the nominals in Amharic, namely, some of the different-root nominals (different-root nominals that can have default gender like *säw* as well as same-root nominals and all the inanimates do not have similar correlations between encyclopedia entry and natural gender) and that such correlations clearly do not hold in all cases, to wit, *mädchen*.

In general, pronominal agreement seems to have a more direct connection with natural gender (aka the gender of the discourse referent). The gender of a pronoun very often corresponds to the natural gender of its referent even if the antecedent of the pronoun is an epicene nominal with ‘fixed’ gender (see e.g., the Romanian data in Farkas and Zec 1995). This can be accounted for by having pronouns agree directly with discourse referents (or more specifically, by having their features correlated with those of the discourse referent; as in Farkas and Zec 1995) or by having gender features combine with pronouns freely but trigger a presupposition failure if the context maps the index for that pronoun to a referent with the wrong gender (Heim and Kratzer 1998). Regardless, it would be interesting to investigate how to connect the licensing conditions analysis with an explicit account of the syntax and semantics of pronominal gender agreement, but that is for a different dissertation.

5.4 Predictions and Consequences of Gender on *n*

In this section, I discuss some of the predictions and consequences of having a gender feature on *n* w.r.t. other phenomena that have been associated with *n*. Specifically, I look at how gender affects the use of *n* in Amharic to capture (a) inflectional class and (b) nominalization (derivational processes that convert verbs, adjectives or nouns to (other) nouns).

5.4.1 Inflection Class

Amharic nominals do not have inflection class in the Indo-European sense, in that they cannot be sorted into two to five types depending on how they are inflected for number, case, gender, etc. However, they may have inflection class in the sense that Amharic, as a root and pattern language, is conventionally analyzed as forming nouns from a consonantal root plus one of a large number of nominal patterns (see e.g., the descriptions in Hartmann 1980, Leslau 1995, etc.). For example, there is a consonantal root /l b s/ which means ‘to wear’ when combined with a verbal pattern. This same root can be combined with one or more nominal patterns to form semantically related nouns, e.g. ‘clothing’ and ‘cover (of a book).’

- (47) Root: / l b s/
 Verb: läbbäsä ‘wear’
 Noun₁: ḥbs ‘clothing’
 Noun₂: ḥbas ‘cover (of book)’

However, across all the Semitic languages, noun formation via root and pattern morphology is far less systematic than verb formation. For example, Arad (2005) observes that in Hebrew there are approximately seven verbal patterns, but over fifty nominal patterns (including patterns that have suffixes and prefixes). Arad also observes that several hundred nouns fall outside of those fifty nominal patterns and are not built from a consonantal root for a variety of reasons (mostly because they are old and new loan words). Moreover, unlike their verbal counterparts, most nominal patterns do not convey anything about the semantics of the noun (which leads Arad (2005: Ch.2) to propose that they only exist to turn consonantal roots into pronounceable strings).

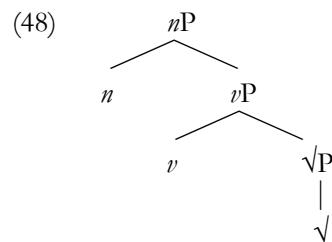
In Amharic, nominal patterns are similarly difficult to pin down. Undoubtedly, there are some productive deverbal noun formations that are associated with particular patterns, e.g., the infinitive, the instrumental, etc. (see Fulass 1966 and below). However, the rest of the nominals display a bewildering variety of patterns which do not convey anything systematic about the semantics of the nominal (as detailed partially in Leslau 1995:261ff. and more fully in Hartmann 1980:223-238). Among the triradicals alone, there are approximately fifty different patterns, not counting prefixal and suffixal patterns (although it is worth noting that most of the patterns vary in terms of vowels and not in terms of syllable structure). This raises the question of whether the nouns are genuinely formed by root and pattern morphology or whether the nouns are listed with (at least) their vowels and then their morphophonology is constrained by general phonotactic and phonological restrictions.

This question is broader than present concerns, but inflection class is relevant to the preceding discussion of gender in so far as, if nouns in Amharic do have patterns, the pattern might be in *n*. Semitic languages have been taken as paragons of the root-and-categorizing-head approach: the consonantal root is the root and the pattern is the category defining head (for at least some cases; see e.g., Marantz 2001, Arad 2003, 2005 for detailed explication of this approach). Under the analysis above, natural gender is in *n*, which might predict that certain patterns are associated with certain natural genders. However, nominal patterns in Amharic are completely insensitive to gender, in that there is no correlation between the gender of a nominal and its pattern. That would mean there would have to be two versions of every pattern used to derive nouns that can have natural gender: one with a [-FEM] feature and one with a [+FEM] feature. This is obviously undesirable. Moreover, a significant strand of work (see e.g., Roca 1989, Harris 1991, Aronoff 1994, Alexiadou 2004) argues that inflection class and gender are related but not the same, i.e., they should not be part of the same morpheme.²⁴

There seem to be two obvious solutions to this problem. First, the structure of the nominal could be more complex than originally thought, i.e., the nominal pattern could be in *n* and natural gender could be in its own separate projection. This option is not aesthetically pleasing, but perhaps workable. However, the other option would be that Amharic nominals are not inflected via root and pattern morphology, and this option seems the most viable. Intuitively, the variation among non-deverbal nominals seems too great to result from any productive root plus pattern combinations, even more than in the other Semitic languages. This intuition remains to be substantiated with statistical evidence, but in Chapter Number, the issue resurfaces and again, the facts seem to indicate that nominals are not inflected via root and pattern morphology (i.e., nominal roots contain vowels in Amharic).

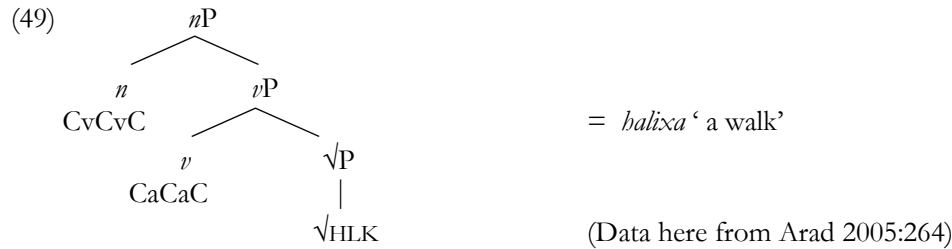
5.4.2 Nominalizations

Besides inflection class, *n* is often assumed to nominalize other categories, e.g., to form nouns from verbs, adjectives, and other nouns. To form a deverbal noun, for example, it is claimed within the framework that I have been assuming here that a *n* is merged with a *v*P (see e.g., Marantz 2001, Arad 2003, 2005).



The root plus *v* constitute the verb, and this verb is nominalized by the addition of a *n*P on top. A specific example from Hebrew is in (49).

²⁴ More specifically, Roca (1989:23-25) claims that while there are correlations between the gender of a nominal and its theme vowel in Spanish, the correlations are unidirectional. Theme vowel can be predicted from gender, but not vice versa. This is highly compatible with an analysis of gender on *n* given the analysis of theme vowels proposed in some recent work within Distributed Morphology (e.g., Embick and Halle 2005, Embick and Noyer 2007). In the DM approach, theme vowels are inserted post-syntactically and adjoined to the categorizing head, i.e., *n* in this case. The insertion of theme vowel would therefore be conditioned by its local environment -- what gender feature is on *n* (or on the root, the complement to *n*).



The root \sqrt{HLK} plus a v result in the verb (vP) *halax* ‘walk.’ Merging a n as sister to the vP nominalizes the verb, giving it a new pattern (whose vowel quality is different from but apparently dependent on the vowel quality of the verb -- hence the vowels are unspecified in the pattern in the tree). The resulting deverbial noun is *halixa* ‘a walk.’

In the analysis of gender developed here, n comes in different flavors: with a [+FEM] gender feature, with a [-FEM] gender feature, or plain (without a gender feature). These gender features correspond to the natural gender of the nominal that results when a n is combined with a root. Do any of the n 's that form nouns from verbs, adjectives and other nouns also have natural gender features?

In Amharic, there are many different nominalization strategies: Fulass (1966) lists twelve different patterns of deverbial nouns and ten suffixes that convert adjectives or nouns to other nouns (although it is quite possible that some of these nouns are derived directly from a root and not from a previously-formed verb, adjective, noun, etc.). Many of these nominalizations result in nouns that do not have natural gender. For example, infinitives, ‘instrumental’ deverbial nouns and nouns formed via the suffix *-inna* all result in abstract concepts or inanimate objects.

- (50) Infinitives
- a. säbbärä ‘break’ → mäsbär ‘breaking, to break’
 - b. sämma ‘listen’ → mäsmat ‘listening, to listen’
 - c. ayyä ‘see’ → mayät ‘seeing, to see’
 - d. bälla ‘eat’ → mäblat ‘eating, to eat’

- (51) Instrumentals
- a. hedä ‘go’ → mähedʒa ‘destination’
 - b. täk’ämmät’ä ‘sit’ → mäk’k’ämmätʃ’a ‘seat’
 - c. k’äzzäfä ‘row’ → mäk’zafya ‘paddle’
 - d. t’ärrägä ‘sweep’ → mäträgya ‘broom’

- (52) -inna Nominals
- a. sänäf ‘lazy’ → sɪnf-ɪnna ‘laziness’
 - b. nɪs’uh ‘clean’ → nɪs’h-ɪnna ‘cleanliness’
 - c. k’omat’a ‘leper’ → k’umt’-ɪnna ‘leprosy’
 - d. gäbäre ‘farmer’ → gɪbr-ɪnna ‘agriculture’

I assume that the n converting from verb, adjective or other nominal to noun in these cases is the plain n that lacks gender features altogether.²⁵

²⁵ It is possible that one or more of these particular nominals are derived directly from a root, as opposed to a previously-formed verb, adjective or noun. Determining which kind of derivation is the case for a particular nominalization requires semantic and morphophonological evidence that goes beyond the scope of this section. If some of them are root-derived, it is not a problem at all - the root would combine with a plain n .

However, some nominalizations result in animate nominals that can have natural gender. For example, the suffix *-ännä* can derive professions.

(53) *-ännä* Nominals

- a. *gṯmb* ‘stone’ → *gṯmb-ännä* ‘mason’
 b. *gāzet* ‘a ‘newspaper’ → *gāzet-ännä* ‘journalist’
 c. *fārās* ‘horse’ → *fārās-ännä* ‘horseman’

Also, the so-called participles (deverbal nouns) can also refer to animates.²⁶

(54) Participles

- a. *s’afä* ‘write’ → *s’afi* ‘writer’
 b. *dʒämmärä* ‘begin’ → *dʒämmari* ‘beginner’
 c. *gaggärä* ‘bake’ → *gagari* ‘baker’
 d. *nädda* ‘drive’ → *nädʒi* ‘driver’

All these derived nominals are same-root nominals in terms of gender, i.e., nominals that can refer to either male or female animates without a change in form (this is not unexpected since both the male and female forms are derived from the same root). For example, *särrat-ännä* ‘worker’ can refer to male or female workers, as shown by the agreement on the definite markers in the examples in (55).

- (55) a. *särrat-ännä-w* b. *särrat-ännä-wa*
 work-N-DEF work-N-DEF.F
 the male worker the female worker
 Fulass 1966:88 Ethiopian Reporter, 10/15/2008²⁷

The same goes for participles, e.g. *märi* ‘leader’ derived from *märä* ‘lead’ can be either masculine or feminine depending on the natural gender of the referent.

- (56) a. *budṯn märi-w* b. *budṯn märi-wa*
 team leader-DEF team leader-DEF.F
 the male team leader the female team leader
 Walta yeka23a03 Walta meg18a07

This is straightforward to account for in the licensing conditions analysis. The *n*’s that corresponds to *-ännä* and participializing morphology come in different flavors in that they have different gender features on them, just like the *n* that combines with roots. There are no licensing conditions between the previously formed verb or noun and the nominalizing *n*. At first glance, then, having gender features on *n* causes no difficulties for treating nominalized verbs, adjectives and nouns in Amharic. Even better, the fact that these nominalizations are all same-root nominals is in fact predicted if licensing conditions are assumed only to hold between roots and categorizing heads (i.e., they are generalizations about the distribution of roots) and not between categorizing heads and non-root projections they may attach to.²⁸

²⁶ To again address the root-derived vs. word-derived distinction, I follow the descriptive literature in assuming that these nominalizations are derived from other nouns and verbs, respectively.

²⁷ <http://www.ethiopianreporter.com/content/view/3026/54/>. Also, note that the *-t-* between *särra* ‘work’ and *-ännä* is epenthetic.

²⁸ If licensing conditions are in fact the epiphenomenal result of the presence/absence of Encyclopedia entries (e.g., there is no Encyclopedia entry for the combination of a root like ‘mother’ and a *n*[FEM]), this result truly falls out. Encyclopedia entries are pairings of roots and categorizing heads, not categorizing heads and other types of projections.

However, there is one snag. Recall that $n[+FEM]$ is spelled out as the suffix *-it* for certain same-root nominals.

- (57) a. $l\acute{d}\mathfrak{z}$ $l\acute{d}\mathfrak{z}$ -it b. $m\acute{a}n\acute{a}k^{w}se$ $m\acute{a}n\acute{a}k^{w}s$ -it
 boy, child girl monk nun

However, for most same-root nominals referring to females, $n[+FEM]$ has no morphophonological realization. The same holds true for nominalizations. Most derived nominals do not take the *-it* suffix (see examples above), but there is a crucial exception: denominal nouns formed via the suffix *-awi* do take an *-it* suffix.²⁹

- (58) a. $ityop'p'iy$ -awi $ityop'p'iy$ -aw-it
 an Ethiopian man an Ethiopian woman
 (Fulass 1966:114)³⁰

- (59) a. ... $ityop'p'iy$ -awi-w Abäbä Ymer ye-Las Vegas maratón-in a $\mathfrak{f}änaäfa$
 ...Ethiopia-N-DEF Abebe Ymer of-Las Vegas marathon-ACC won
 The Ethiopian Abebe Ymer won the Las Vegas marathon.
 Walta tir27a01

- b. $b\acute{a}$ -set-ot $\mathfrak{f}t\mathfrak{f}$ maratón w $\acute{d}d\acute{d}d\acute{t}$ $ityop'p'iy$ -aw-it-wa Werk'ne \mathfrak{f} Tola...
 in-woman-PL marathon competition Ethiopia-N-F-DEF.F Werknesh Tola...
 In the women's marathon, the Ethiopian Werknesh Tola...
 Walta hed24a1

The problem that these nominalizations pose is that the nominalizing morphology (e.g., the suffix *-awi*) and the gender morphology (*-it*) seem to be two different Vocabulary Items. All things being equal, they should correspond to two separate morphemes (bundles of features) in the morphosyntax. However, the analysis here predicts that the features they correspond to will be on a single morpheme: n .

There are many potential ways in which this problem could be addressed. First of all, at least some languages are not like Amharic in that gender morphology and nominalization morphology are not realized as separate Vocabulary Items. French, for instance, has a rich set of nominalization suffixes that also express gender (*-eur/-euse*, *-teur/-trice*, *-ant/-ante*, *-oir/-oire*, *ier-ière*, etc.). Also, note the phrase “all things being equal” in the paragraph above. The relationship between the bundle of morphemes manipulated by the morphosyntax and the phonological realization of that bundle (the Vocabulary Item) can be complex and indirect. This is what motivated the Separationist Hypothesis in the first place, and it is also why Distributed Morphology posits certain operations that can manipulate morphemes before Vocabulary Items are inserted, e.g., Fission, Fusion, Impoverishment, etc. It may be that, in Amharic, a particular operation (probably Fission) applies to split off the gender feature from n post-syntactically. Finally, it should be noted that this has not been so far a

²⁹ Leslau (1995:162) also reports that two participles can take a feminine *-t* suffix when they have female referents (*aq'q'abi* ‘male custodian, one who grinds grain for the Host,’ *aq'q'abit* ‘female custodian, one who grinds grain for the Host,’ *mäggabi* ‘monk responsible for the administration of food, pastor’ *mäggabit* ‘nun responsible for the administration of food’). However, I believe the vast majority of participles do not take any feminine suffixes (no participles with feminine suffixes were found in the Walta Corpus), rendering *aq'q'abit* and *mäggabit* exceptional. It is possible that these two participles have been reanalyzed as derived from a root instead of from a previously-formed verb, and thus would be like typical same-root nominals. Leslau claims that the feminine suffix in these cases is retained from Ge'ez, and these forms in particular might have retained it since they are both (at least etymologically) religious terms and Ge'ez is the language of the church.

³⁰ It is unclear whether the *-i* preceding *-t* in the examples belongs to the suffix *-awi* or the feminine suffix *-it*. Leslau (1995:162) assumes that it belongs to *-awi* but it seems more likely that the feminine suffix here is *-it* as per the other same-root nominals.

thorough investigation of nominalization in Amharic. Many observations remain to be explained: does *-it* only attach to nominalizations that end in *-i* (see fn. 29)? If the *-ami* nominals are the only set of nominals relevant for this problem, are they best treated as exceptional in some way (e.g., by triggering Fission)? The answers to these questions will surely affect how to capture the complex relationship between nominalization and gender in Amharic.

To wrap up this section, I have taken a quick look at the implications of having a gender feature be on *n* for inflectional class and for nominalizations in Amharic. It was suggested that nominals in Amharic do not have inflectional class, an idea that remains to be confirmed but seems intuitively correct and in line with other root and pattern languages. Nominalizations were fairly straightforward to account for assuming *n* can have natural gender, with the exception of the *-ami* nouns where gender seems to surface as a separate vocabulary item from the nominalizing morphology.

6 CONCLUSION

In this chapter, an analysis was developed of the Amharic gender system where natural gender and grammatical gender are explicitly formally separated: natural gender is reflected in an interpretable gender feature on *n* and grammatical gender in an uninterpretable gender feature on roots. The analysis is successful in accounting for how nominals are assigned gender as well as what gender is used for gender agreement in Amharic. From a broader perspective, the analysis is a viable account of gender that does not appeal to the lexicon, and it holds promise in accounting for different types of gender systems that may rely on natural gender to a greater or lesser extent (as opposed to some current analyses of gender which have difficulties accounting for a natural gender-based system as in Amharic).

There are a few empirical issues that have been left open in the above discussion. First, Amharic has a productive diminutive-forming operation, alluded to briefly in Section 2.1, which results in any number of interpretations including cuteness, smallness and affection. In terms of its morphosyntax, the diminutive is closely related to feminine gender in that it can be signaled by an *-it* suffix (on nominals or adjectives) or simply by feminine agreement on verbs, definite markers, etc -- even if the referent has male natural gender. The diminutive is in fact the only phenomenon that can ‘override’ natural gender in terms of the agreeing gender, which suggests it may be present syntactically in a functional head that dominates *n* (see Lowenstamm 2008 for exactly this kind of proposal for French). Future work will hopefully make the connection between the proposals here and the diminutives explicit.

It should also be noted that the inanimate nominals that are classified as feminine here are, in fact, **usually** treated as feminine. The distribution patterns are subtle and complex and will require (I think) detailed statistical study to sort out -- the frequency with which these nominals are treated as feminine appears to vary per nominal (some are always treated as feminine, some less consistently) and perhaps according to other factors (countries are consistently treated as feminine in the Walta corpus, but a consultant judged grammatical treating a country name as masculine). Could some of the feminine nominals have two roots, one [+FEM] and one not? This idea seems unpleasant but there are a couple of important mitigating factors, namely, how small the number of feminine inanimates really is (fifteen to twenty is my best guess, not including country names, and bear in mind that some of them **are** treated consistently feminine) and how Amharic (in my impression) appears to be losing grammatical gender as a language. An intermediate period of flux where some roots are listed as both with grammatical gender ([+FEM]) and without it seems reasonable.³¹

³¹ There is a crucial distinction to be made here. It was argued earlier in this chapter that having two roots (one masculine and one feminine) for certain nominals was a severe enough offense to knock an analysis out of the running. However, the circumstances were different both times. First, w.r.t. same-root nominals (see Section 3.2.1), there are many, many more same-root nominals in the language (and more are being added even now as loan words, e.g., *profesor* ‘professor’) than there are feminine inanimate nouns, resulting in much greater inefficiency and redundancy. Second, w.r.t. the feminine default animals (see Section 3.2.1), it was argued that is difficult to ensure that the feminine root is always used in the default gender case and either root when natural gender is known.

Accordingly, I would like to finish this chapter with some discussion from a diachronic perspective. What does it mean to say that Amharic is losing grammatical gender, and what was gender like at an earlier stage of the language? To answer the first question, it is illustrative to look at the history of gender in English. It is universally acknowledged that in the past millennium, English changed from a language that relied on grammatical gender and natural gender, to a language that relies almost exclusively on natural gender (i.e., in Modern English, inanimate nominals are not sorted into two or more arbitrary genders, pronouns and nominals referring to animate nominals have the natural gender of their referents; for recent perspectives on the shift, see Curzan 2003 and Platzer 2005). It is commonly believed that the loss of gender morphology on both nouns and modifiers caused or at least greatly abetted the loss of grammatical gender. The basic idea is that without morphological cues about nominal gender it is difficult to determine (i.e., acquire) what the gender of a nominal is (or whether it has gender at all).

The Ethio-Semitic language Ge'ez (spoken during the Axumite empire, first written down around the 4th century BCE, now the liturgical language of the Ethiopian Christian church) is not a direct ancestor of Amharic; it belongs to the North branch of Ethio-Semitic (with Tigre and Tigrinya) whereas Amharic is part of the South branch (with Harari and the Gurage languages). However, it offers some tantalizing clues about what an earlier stage of Amharic might have been like, and suggests that changes have happened that facilitated the loss of grammatical gender in Amharic.

Ge'ez had both masculine and feminine gender, but it is difficult to ascertain whether there were more feminine inanimate nouns than there are in Amharic.³² It is clear, though, is that the gender agreement system was much richer in Ge'ez. Adjectives generally agreed in gender (Lambdin 1978:68), unlike in Amharic where only a handful of adjectives do so. Also, verbal paradigms were richer w.r.t. gender in Ge'ez. In Amharic, across all the verb forms, verbs with plural number do not have separate gender forms, i.e., there is no third person masculine plural or third person feminine plural form. However, this was not true in Ge'ez -- many (if not all) of the verbs show distinct feminine and masculine forms in the plural (see e.g., the perfect verbal paradigm in Lambdin 1978:50, the imperfect verbal paradigm in Lambdin 1978:144). The list continues: cardinal numbers, relative pronouns, and participles all agreed in gender in Ge'ez but do not in Amharic.

The loss of gender agreement from Ge'ez to Amharic could perhaps have triggered the beginning of a shift away from grammatical gender, similar to the loss of agreement in the history of English. Granted, there are many caveats here: again, Ge'ez is not the direct ancestor of Amharic and it remains to be established whether the gender system for nominals in Ge'ez relied more heavily on grammatical gender (if it had, for example, epicene nominals). Nevertheless, it is highly suggestive that Ge'ez featured significant gender agreement that seems to have all but disappeared from Amharic, and that the gender system in Amharic is currently very much based on natural gender.

³² Lambdin (1978:26-27) notes that the gender of certain inanimate nouns seems to vary in Ge'ez, but this statement should be taken with a grain of salt. Lambdin himself observes that gender usage varies across texts (meaning that gender may be consistent for a particular author) and that some feminine nominals appear to be expressive in use (i.e., he does not separate diminutive from feminine nominals) -- both these factors could seriously inflate the number of words whose gender seems to vary.