

# Number-based noun classification: The view from Kipsigis

Maria Kouneli  
*Universität Leipzig*

**Abstract:** Nilo-Saharan languages are well-known for their complicated system of nominal number marking, which features a variety of singulative and plural affixes (Dimmendaal 2000). Even though these systems have received some attention in the typological literature, there has been limited theoretical work on their implications for the morphosyntax of number cross-linguistically. The goal of this paper is to fill this gap, by providing an analysis of nominal number morphology in Kipsigis (Nilotic, Kenya), based on data from original fieldwork. First, I show that singulatives in Kipsigis are true allomorphs of singular number, unlike singulatives with a classifier function in languages like Ojibwe (Mathieu 2012). The descriptive term ‘singulative’ is therefore misleading, as it corresponds to two very different types of morphemes. Second, I claim that the tripartite system of number marking of Kipsigis and other Nilo-Saharan languages is due to the classification of nouns into morphosyntactic classes defined by the presence of inherent number features on little *n*; the interaction of these features with interpretable number features on the functional projection Num (Ritter 1991 a.o.) in the post-syntactic component gives rise to the exponence pattern that we observe. Finally, my analysis corroborates the existence of noun classification based on number, which has only been argued for Kiowa-Tanoan before (Harbour 2007). The existence of three number classes in Kipsigis can only be explained by reference to bivalent number features; number-based noun classification systems thus strongly support the view that number features are bivalent and not privative, which is also argued by Harbour (2007, 2011) for Kiowa.

## 1. Introduction

Kalenjin languages<sup>1</sup> have what has been called in the literature a ‘tripartite’ system of number marking (Corbett 2000; Dimmendaal 2000): some nouns are interpreted as singular in their morphologically unmarked form and form their plural by the addition of a plural suffix (1), some nouns are interpreted as plural in their unmarked form and form their singular by the addition of a singulative suffix (2), while a third class of nouns never appear in their unmarked form: they have a singulative suffix in the singular, and a plural suffix in the plural (3). This system is different from that of most Indo-European languages, where nouns usually follow the pattern in (1), that is, they are morphologically unmarked in the singular and marked in the plural.

(1) Plural marking:	kipaw (SG)	kipaw- <b>tiin</b> (PL)	‘rhino’
(2) Singulative marking:	peel- <b>yaan</b> (SG)	peel (PL)	‘elephant’
(3) Singulative/Plural marking:	pata- <b>yaan</b> (SG)	pat- <b>een</b> (PL)	‘duck’

(Endo-Marakwet dialect; Zwarts 2001)<sup>2</sup>

---

<sup>1</sup>Kalenjin is the term used to refer to a dialect cluster of the Southern Nilotic branch of the Nilo-Saharan language family. The Kalenjin languages are spoken in Western Kenya by about 5 million speakers, with Kipsigis being the dialect with the highest number of speakers – approximately 2 million (Eberhard et al. 2020). There is some descriptive work on Kalenjin (e.g., Toweett 1975, 1979 on Kipsigis, Hollis 1909; Creider & Creider 1989; Creider 1982 on Nandi), but limited theoretical work (e.g., Jake & Odden 1979; Bossi & Diercks 2019; Diercks & Rao 2019; Kouneli 2017, 2019a, b on Kipsigis, as well as a series of papers on Nandi: e.g., Creider 1989; Creider & Creider 1983).

<sup>2</sup> Unless otherwise indicated, the data in this paper come from fieldwork on Kipsigis conducted by the author during three field trips to Kenya between 2016 and 2018, with ten male and two female speakers (age range: 18-30). The speakers came from four different Kipsigis-speaking areas in Western Kenya (Bomet, Narok, Kericho, and Nakuru Counties); the speakers had varying degrees of fluency in English and Swahili (the official languages of Kenya), ranging from fluent to poor. All of them had grown up in monolingual Kipsigis regions, with no exposure to English/Swahili before the age of 6. The data were collected in elicitation interviews designed to investigate the

Even though these systems have gained some attention in the typological literature (e.g., Dimmendaal 2000; Di Garbo 2014), there has been limited theoretical work on their implications for the syntax of number cross-linguistically. Grimm (2012, 2018), however, includes data from Turkana (Eastern Nilotic; Kenya) in his theoretical account of tripartite systems of number marking, and he argues that those systems reflect an organization of nouns into semantic classes according to the *Scale of Individuation* in (4). In this type of theory, nouns with singulative marking (as in 2 above) constitute a number category (‘collectives’) that lies between mass and count nouns in the language. The singulative is then analyzed as a classifier-like morpheme that creates units/individuals out of mass or collective nouns, a view that goes back to Greenberg (1978) (recent syntactic accounts of this view of the singulative include Mathieu 2012 and Ouwayda’s 2014 analysis of the singulative in Ojibwe and Lebanese Arabic respectively).

(4) *Scale of individuation* (Grimm 2018: 547)

liquids/substances < granular aggregates < collective aggregates < individuals

The goal of this paper is to provide an analysis of the tripartite system of number marking of the understudied Kipsigis dialect of Kalenjin, based on data from original fieldwork. I show that there is no evidence that nouns that take singulative marking in Kipsigis are collectives, and that the singulative is not an individualizing morpheme. Rather, nouns that take singulative marking behave in the same way as plural count nouns in the language, and the singulative is best seen as an allomorph of singular number. I provide an analysis of the pattern in (1) – (3) according to which nouns are classified into morphosyntactic classes based on uninterpretable number features on little *n*, the nominalizing head in theories like Distributed Morphology (e.g., Marantz 1997). The interaction of these features with interpretable number features on the functional projection Num (Ritter 1991 among many others) gives rise to the pattern seen in (1) – (3).

The analysis of the Kipsigis system developed in this paper highlights that the term ‘singulative’ is misleading: it can refer to either a classifier-like morpheme with individualizing semantics (e.g., in Ojibwe; Mathieu 2012), or to a true allomorph of singular morphology (e.g., in Kipsigis). The latter type of languages illustrate that singular number is not always the morphologically unmarked number form (for at least some nouns). Furthermore, the analysis corroborates the existence of number-based noun classification (i.e., the use of uninterpretable number features to sort nouns into morphosyntactic classes), which has only been argued for Kiowa-Tanoan languages before (Harbour 2007, 2011; Watanabe 2015).<sup>3</sup> As I will discuss in detail in Section 4, the existence of three (as opposed to two) number classes in Kipsigis (seen in 1 – 3) can only be explained by reference to binary, and not privative, number features. This is a point also emphasized by Harbour (2007, 2011) in his analysis of Kiowa, which shows that number-based noun classification systems provide evidence in favor of the existence of binary number features.

The remainder of the paper is structured as follows. In section 2, I describe the morphological expression of nominal number in Kipsigis. In Section 3, I show that nouns that take singulative marking in Kipsigis do not behave like collectives, and that the classification of nouns in number classes is orthogonal to the mass/count distinction in the language. In Section 4, I present the analysis of the tripartite system of number marking in the language. In Section 5, I discuss the implications of my analysis for other languages with number-based noun classification. In Section 6, I conclude.

---

morphology and syntax of noun phrases in the language; a short story was also elicited, and the author was present at numerous spontaneous speech situations due to accommodation with native speakers.

<sup>3</sup> Bantu languages are famous for having a gender system that reflects both noun class and number, but I focus here on languages where number/individuation is the sole property used for classification purposes.

## 2. The Kipsigis number system: morphology

### 2.1. Three number classes

We can think of the Kipsigis system of number marking in terms of three morphological classes of nouns: one class of nouns inflects for number according to the pattern in (1), a second one according to the pattern in (2), and a third one according to the pattern in (3). Once we think of the number system in these terms, we can define the three classes as follows:

(5) *Kipsigis number classes*:

- a. *Inherently singular nouns*: these are nouns that trigger singular agreement in their unmarked form, and form their plural by the addition of a plural suffix. The majority of nouns in the language belong to this category. Example:

peet-u-it → pêetúut	peet-uus-ya-ik → pêetùusyék <sup>4,5</sup>
day-TH-SEC	day-PL-TH-SEC
‘day (SG)’	‘days (PL)’ <sup>6</sup>

- b. *Inherently plural nouns*: these are nouns that trigger plural agreement in their unmarked form, and form their singular by the addition of a singulative suffix. Nouns in this category include many mass nouns, insects and small animals, medium-/large-size animals that usually appear in groups (e.g., elephants), some plants, names for groups of people or names of professions (e.g., ‘the Kipsigis’), agent nominalizations, as well as nouns describing entities that tend to appear in pairs or groups (e.g., the words for ‘teeth’, ‘fruit’, and ‘beads’). Example:

---

<sup>4</sup> The following spelling conventions are used: a) double vowels indicate a long vowel, b) underlined morphemes indicate [–ATR] vowels, while regular font indicates [+ATR] vowels (the language has a system of [ATR] harmony; cf. Lodge 1995; Baković 2000; Nevins 2010), and c) à = Low tone, á = High tone, and â = HL contour tone; these are the three tones in the language, which are used mainly to convey grammatical information, such as Case. Tone transcriptions are included, but there is a limited understanding of how the tone system of Kipsigis works (see Creider 1982 for an analysis of the nominal tonology of the related Nandi dialect).

<sup>5</sup> Most data in this paper are presented in the following format: the underlying form of the morphemes (before phonological processes take place) appears on the left of the arrow, while the surface form of the word appears on the right of the arrow. Tones are presented only in the surface form. The two relevant phonological processes are:

- (i) Vowel coalescence: a + i → ee, e + i → ee, o + i → ee, u + i → uu, i + i → ii  
(ii) Length dissimilation: word-final long vowels with a HL contour tone become short when they follow a long vowel with a HL tone.

Shortening of word-final long vowels sometimes also occurs in polysyllabic nouns, but the conditioning factors are not well-understood. Tonal constraints play an important role, which has also been argued for the same phenomenon in the Nandi dialect of Kalenjin (Dimmendaal 2012). See the Appendix and Kouneli (2019a: Chapter 2) for more data and discussion of vowel shortening in Kipsigis and its relation to Dimmendaal’s (2012) analysis of Nandi.

<sup>6</sup> The following abbreviations are used in glosses: AH Arabic singulative, AG agent nominalization, AN animate, APPL applicative, D dual, DEM demonstrative, FEM feminine, I inverse, IN inanimate, IPFV imperfective, NOM nominative case, PAT patient nominalization, PAST1 current past, PAST2 recent past, PL plural, POSS possessive, Q question particle, REL relativizer, SEC secondary suffix, SG singular, TH thematic suffix, TOP topic marker 1/2/3 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup> person.

ngeend-**yaan**-ta-it → ngéendyaát<sup>7</sup>  
 bean-SG-TH-SEC  
 ‘bean (SG)’

ngeend-a-ik → ngéendéek  
 bean-TH-SEC  
 ‘beans (PL)’

- c. *Numberless nouns*: these are nouns that always appear with a number suffix (singulative in the singular and plural in the plural), and they correspond to what Dimmendaal (2000) calls the replacement pattern. Few nouns belong to this class, such as the words for ‘fish’, ‘socks’, and ‘shoes’. Example:

sigis-**yaan**-ta-it → sigisyáat  
 sock-SG-TH-SEC  
 ‘sock (SG)’

sigis-**iin**-ik → sigisiiník  
 sigis-PL-SEC  
 ‘socks (PL)’

More examples of nouns in each class can be found in the Appendix.

The above brief description of the lexical categories comprising each class leads us to the rough generalization that the inherently plural class includes nouns denoting entities that appear in groups (or pairs) more often than they appear in units. This generalization seems to hold in all languages with a tripartite system of number marking both within the Nilo-Saharan language family (Dimmendaal 2000; Moodie 2016, 2019), and outside of it (Grimm 2012, 2018; Haspelmath & Karjus 2017). The converse though is not true: there are nouns denoting entities that usually appear in groups that do not belong to the inherently plural class, and belong to the inherently singular class instead (for example, both lions and elephants live in herds, but *ngétùndá* ‘lion (SG)’ is inherently singular, while *pèeléek* ‘elephants (PL)’ is inherently plural). As for numberless nouns, there is no clear semantic generalization, though those nouns have similarities to nouns appearing in the inherently plural class. Finally, the inherently singular class comprises the majority of Kipsigis nouns; most of them are count nouns, but a fair amount of mass nouns also belong to this class.

Even though there are semantic generalizations to be made about the classification in (5), there is a high degree of idiosyncrasy in the system. Moreover, there is variation across speakers, and across dialects with respect to class membership. For example, the word for ‘shoe’, *kwèyáat*, belongs to the numberless class in Kipsigis, but to the inherently plural class in Endo-Marakwet (Zwarts 2001). Moodie (2019) reports similar patterns of variation for the Eastern Nilotic language Lopit.

The idiosyncrasy of the system and speaker/dialect variation are typical properties of gender systems in better-studied Indo-European languages. For example, in Greek, like in many other Indo-European languages, nouns are classified into masculine, feminine and neuter gender. There is a semantic core to the classification system with animate nouns being assigned masculine or feminine gender according to their biological sex, but there are a lot of exceptions (e.g., *koritsi* ‘girl’ in Greek is neuter, and not feminine). Furthermore, similarly to the situation in Kipsigis, some words have a different gender in

<sup>7</sup> This extreme reduction cannot be easily explained by phonology only. We know that *-yaan*, followed by *-ta*, are the underlying morphemes, because they surface as such when the demonstrative suffix is attached to the stem instead of the secondary suffix, as in (iii):

(iii) ngeend-yaan-ta-ni → ngéendyaándani  
 bean-SG-TH-DEM  
 ‘this bean (SG)’

Moreover, according to Creider & Creider (1989), the reduction to *-yáat* occurs in the oblique case, but not in the nominative in some dialects. In Kipsigis, however, the reduction always takes place, irrespective of the case of the noun. Daniel Harbour (p.c.) asks whether it is simpler to posit two different allomorphs for the surface form of *-yaan*, instead of positing complex morphophonological rules (such as the ill-understood reduction described here). I do not see why this would not be possible, but I leave the details of exploring the different predictions made by the two approaches as a topic for further research.

different dialects and for different speakers (e.g., *dzaki* ‘fireplace’ is neuter in Standard Modern Greek, but masculine for some speakers).

Kramer (2015: 70) gives the following (relatively uncontroversial) definition of gender (= noun class):

(6) Gender is:

- (i) the sorting of nouns into two or more classes;
- (ii) assigned depending on biological sex, animacy, and/or humanness, for at least some animate nouns;
- (iii) reflected by agreement patterns on other elements (e.g., adjectives, determiners, verbs, auxiliaries)

Kipsigis clearly sorts nouns into two or more classes, thus satisfying (i), which is reflected in agreement patterns (as we will see in more detail in Section 4) – thus satisfying (iii). Whether it satisfies (ii) is more difficult to evaluate because none of the features of (ii) are involved in the Kipsigis classification system. However, if number is added to biological sex, animacy, and humanness as a possible semantic core in noun class systems (which has already been argued for by Harbour 2007, 2011), Kipsigis would satisfy all criteria in (6). For these reasons, I will assume from now on that the Kipsigis classes reflect a noun classification system. This view will be refined in Sections 3 and 4, which discuss the properties of the system in more detail.<sup>8,9</sup>

## 2.2. Morphological expression of number

Having outlined the basic characteristics of the three number classes in Kipsigis, we can now turn to the details of how singular and plural number are morphologically expressed for each class in (5). The discussion in this section is limited to the behavior of morphologically underived nouns, but derived nouns do not show significant differences from the description provided here; derived nouns are discussed in greater detail in Section 4.2.

As can be seen in the examples in (7) – (8), in their unmarked form, nouns of both the inherently singular and inherently plural classes consist of the root, followed by a thematic suffix, followed by what has been called the ‘secondary’ suffix in previous descriptions of the language (Creider & Creider 1989; Towett 1979).

(7) *Inherently singular nouns – unmarked form (singular)*

- |                      |                          |                        |
|----------------------|--------------------------|------------------------|
| a. ser-u-it → sèrúut | b. laak-wa-it → làakwéet | c. k̄ar-i-it → k̄aríit |
| nose-TH-SEC          | child-TH-SEC             | car-TH-SEC             |
| ‘nose’               | ‘child’                  | ‘car’                  |

<sup>8</sup> I do not claim, however, that Kipsigis noun classes are *just* like gender in Indo-European languages; the definition in (6) is a broad definition of noun class systems (covering, for example, both sex-based gender systems of the Indo-European type and complex noun class systems of the Bantu type), and the only claim being made is that the Kipsigis classes reflect such a system. It is clear, however, that different types of noun class systems share certain properties, but differ in other dimensions. Harbour (2007), for example, argues that the number-based noun classification system of Kiowa is more like Bantu and less like sex-based gender; a similar claim could be made for Kipsigis, and we will see in later sections that the Kipsigis number system co-occurs with sex-based gender in some Nilotic languages.

<sup>9</sup> Alexiadou (2019), in an analysis of the cross-linguistic distribution of plural marking in the presence of numerals, claims that nouns should appear in their singular form in those languages that lack both a definite article and a gender system. Nouns in Kipsigis are obligatorily plural in the presence of a numeral, even though the language lacks a definite article. If Alexiadou’s (2019) analysis is correct, this means that Kipsigis has a gender system. This is, thus, an indirect indication that the noun classification system of Kipsigis performs a similar function in the syntax as gender in other languages.

(8) *Inherently plural nouns – unmarked form (plural)*

- a. keel-a-ik → kèelék  
tooth-TH-SEC  
‘teeth’
- b. karat-i-ik → kàràtíik  
blood-TH-SEC  
‘blood’
- c. sugar-u-ik → sùgàrúuk  
sugar-TH-SEC  
‘sugar’

As can be seen in the examples in (7) – (8), the secondary suffix has the form *-it* in the singular and *-ik* in the plural. The meaning of the secondary suffix is complicated. Previous analyses of Kipsigis (Brian & Tucker 1965; Toweett 1975, 1979), as well as analyses of the related dialects of Nandi (Hollis 1909; Creider & Creider 1989) and Endo-Marakwet (Zwarts 2001) report that every noun has a primary form (i.e., a form without the secondary suffix) and a secondary form (i.e., a form with the secondary suffix). Toweett (1975, 1979) and Creider & Creider (1989) point out that the difference in meaning between the two forms is subtle, and suggest that the secondary form presupposes the existence of the noun in question, while the primary form does not. Hollis (1909) claims that the secondary suffix is a definite article in Nandi, while Zwarts (2001) argues that this suffix is a specificity marker in Endo-Marakwet. However, with the exception of the singular demonstrative that will be discussed later, my consultants never use a noun without its secondary suffix (irrespective of specificity or definiteness), and many times they even fail to recognize the ‘primary form’ of the noun as an existing word of the language.<sup>10</sup> I, therefore, hypothesize that the secondary suffix has historically evolved from a specificity marker (and possibly still survives as such in other dialects), but is now simply a nominal marker in Kipsigis (see Greenberg 1978 for possible evolution paths of such markers).

As for thematic suffixes, they usually consist of a vowel or glide + vowel combination, and they cannot be predicted by the phonological shape or semantic content of the nominal root that they attach to.<sup>11</sup> More importantly, they are not correlated with the number class of the noun: for example, the inherently singular noun in (7a), *sèríut* ‘nose’, and the inherently plural noun in (8c), *sùgàrúuk* ‘sugar’, both have the thematic suffix *-u*, despite the fact that they belong to different number classes. The following two tables give a list of the thematic suffixes found with inherently singular and inherently plural nouns in my data, along with an example. The thematic suffixes are presented in order of frequency of occurrence in my fieldnotes.

Table 1: *Thematic suffixes for inherently singular nouns*

<i>Thematic suffix</i>	<i>Example</i>
-a	<u>óorêet</u> (or-a-it) ‘road/clan’
-i	<u>kàríit</u> (kar-i-it) ‘car’
-wa	<u>lâakwéet</u> (laak-wa-it) ‘child/girl’
-u	<u>sèríut</u> (ser-u-it) ‘nose’
-e	<u>chérérêet</u> (cherer-e-it) ‘baby’
-ya	<u>tàríityét</u> (tariit-ya-it) ‘bird’

<sup>10</sup> The primary form of the noun is used for a couple of high frequency nouns, such as the word for ‘child’, in the vocative.

<sup>11</sup> The only loose semantic correlation that could be found was the frequent appearance of the thematic suffix *-wa* in plants’ names.

Table 2: *Thematic suffixes for inherently plural nouns*

Thematic suffix	Example
-a	sáráméek (sáram-a-ik) 'twins'
-u	sùgàrúuk (sugar-u-ik) 'sugar'
-i	kàràtiik (karat-i-ik) 'blood'

Finally, there are some nouns that do not have a thematic suffix in their unmarked form, in which case the root is followed directly by the secondary suffix. In this case, the secondary suffix has the *-ta* allomorph in the singular and the *-ka* allomorph in the plural:<sup>12</sup>

- (9) a. roop-ta → ròoptá *Inherently singular noun without a thematic suffix*  
rain-SEC  
'rain (SG)'
- b. chee-ka → chèegá *Inherently plural noun without a thematic suffix*  
milk-SEC  
'milk (PL)'

In their marked number form, nouns consist of the root, followed by a singulative/plural suffix, (sometimes) followed by a thematic suffix (which is predictable by the number suffix), followed by the secondary suffix, as can be seen in the examples in (10). Note that the thematic suffix found in the unmarked form of the noun is absent in its marked form; it will be shown, however, that it partially predicts the number suffix selected by the noun.

- (10) a. *Inherently plural noun/singulative marking:*

sig- <b>iin</b> -ta-it → sigiindét	sig-i-ik → sigíik
parent-SG-TH-SEC	parent-TH-SEC
'parent(SG)'	'parents (PL)'

- b. *Inherently singular noun/plural marking:*

oosn-a-it → òosnêet	oosn- <b>oos</b> -ya-ik → òosnòosyék
forest-TH-SEC	forest- <b>PL</b> -TH-SEC
'forest (SG)'	'forests (PL)'

There are a great number of singulative and plural suffixes in the language, while some nouns form their plural by irregular phonological changes or suppletion of the stem. In fact, Corbett (2000), in his typological survey of the morphological expression of number in the world's languages, observes that Nilo-Saharan languages lie at the extreme edge of irregularity in number formation, and considers them a problem for language acquisition due to the apparent lack of any pattern in the formation of singular/plural. A careful examination of the number suffixes in Kipsigis, however, reveals that: a) only a small number of inherently singular nouns have irregular plurals, while no inherently plural noun has an irregular singular,

<sup>12</sup> If the root ends in a [t], then the *-it* form of the suffix is used. Example:

(iv) met-it → métít  
head-SEC  
'head (SG)'

and b) the singulative or plural suffix of the noun can be partially predicted by the thematic suffix of the noun in its unmarked form. Moreover, when the number suffixes are followed by a thematic suffix, this suffix is predictable.

There are two singulative suffixes in the language, *-iin* and *-yaan*, both followed by the thematic suffix *-ta*.<sup>13</sup> The suffix *-yaan* is by far the most productive suffix in the language, while *-iin* is only used with a few human nouns. Examples:

- |  |  |
|--|--|
| <p>(11) chuumb-<b>yaan</b>-ta-it → chûumb-yáat<br/> salt-SG-TH-SEC<br/> ‘one package of salt (SG)’</p> | <p>chuumb-i-ik → chûumbíik<br/> salt-TH-SEC<br/> ‘salt (PL)’</p> |
| <p>(12) sig-<b>iin</b>-ta-it → sigliindét<br/> parent-SG-TH-SEC<br/> ‘parent (SG)’</p>                 | <p>sig-i-ik → sigíik<br/> parent-TH-SEC<br/> ‘parents (PL)’</p>  |

The following table shows the plural suffixes in the language (excluding irregular processes). Most of them are given in the form they have when merged with the (always present) secondary suffix *-ik*, because it is not clear from the phonology where the boundary between the plural suffix and the thematic suffix lies (the last three suffixes in the table are decomposable and, thus, given in their form without *-ik*). Previous descriptions of the language, though, which were composed when primary forms were still in use, state that, with the exception of the plural suffix *-V:s*, which is followed by the thematic suffix *-ya*, all other plural suffixes are followed by the thematic vowel *-i*.

Table 3: Major plural suffixes (excluding irregular/rare processes)

Plural suffix (in the form of ending)	Them. suffix in singular	Example	Gloss
-oosyek	-a	òosnêet – òsnòosyék	‘forest’
-iisyek	-i	kàrîit – kàrîisýék	‘car’
-uusyek	-u	pêetúut – pêetùusýék	‘day’
-oonik	-wa	símàatwét – símàatòoník	‘twin’
-oonok	-wa	àywéet – àònóok	‘axe’
-uunek	-u	éúut – èùnénék	‘hand’
-oy	-a	sâunéet – sâunóok	‘type of tree’
-ay	-a	mógôombéet – mógôombáíik	‘hoe’
-wa	∅	ásîistá – ásîiswék	‘sun’

The thematic suffix of the noun in its unmarked singular form is indicated in the table, and we see that each plural suffix can only appear with nouns that have a specific thematic suffix in the singular (the reverse is not true: nouns with an *-a* thematic suffix, for example, could take three different plural suffixes). Therefore, we see that thematic suffixes are associated with the declension class of the noun. They seem, then, to be very similar to thematic vowels linked to declension class in Indo-European languages. In Spanish, for example, there are three declension classes for nouns, each one of which is associated with a thematic vowel (which could be zero in the third class in the same way that some Kipsigis nouns have a zero thematic suffix). These thematic vowels cannot be predicted by the gender, phonological shape, or semantic content of the nominal root (Roca 1989; Aronoff 1994 among others). These are exactly the

<sup>13</sup> For some nouns, the initial glide of *-yaan* has phonologically predictable allophones (e.g., *-w* after another *w/u* in a diphthong). There is one noun where the phonology is not predictable (*pòobáat* ‘mushroom’ where phonology predicts *pòobcháat*). Finally, there are a couple of exceptional nouns where only the singulative’s thematic suffix *-ta* is present, without *-yaan*. More examples of nouns with different singulative and plural suffixes can be found in the Appendix.



characteristics of the Kipsigis thematic suffixes. I provide a sketch of an analysis of Kipsigis thematic suffixes in Section 4; I assume that they reflect declension class, which is generally agreed not to play an active role in syntactic derivation (Aronoff 1994; Oltra-Massuet 1999; Alexiadou 2004; Oltra-Massuet and Arregi 2005; Embick and Halle 2005; Alexiadou and Müller 2008).<sup>14</sup>

In sum, this is the morphological make-up of underived nouns, depending on their number class:

(13) a. *Inherently singular nouns*

SG: root – thematic suffix – secondary suffix

PL: root – plural suffix – (thematic suffix) – secondary suffix

b. *Inherently plural nouns*

SG: root – singulative suffix – thematic suffix – secondary suffix

PL: root – thematic suffix – secondary suffix

c. *Numberless nouns*

SG: root – singulative suffix – thematic suffix – secondary suffix

PL: root – plural suffix – (thematic suffix) – secondary suffix

### 3. The Kipsigis number system: semantics and the mass/count distinction

Singulative marking exists in a number of languages outside of the Nilo-Saharan family, with Welsh and Breton (Celtic), Arabic and Maltese (Semitic), and Ojibwe (Algonquian), being the best discussed in the theoretical literature (Grimm 2012, 2018; Mathieu 2012; Ouwayda 2014 among others). All previous accounts of the singulative treat it as a sort of classifier, i.e., a morpheme that turns mass or collective nouns into individuals. For example, Mathieu (2012: 653) writes that ‘the singulative is a process by which a collective or a mass noun is turned into a unit’. As for the definition of collective nouns, this term has been used in the literature to refer to a variety of different entities, but in the context of singulatives the term is mostly used for a noun that refers to entities that usually appear in spatiotemporally organized collections (e.g., ants, beans, etc.). In fact, Grimm (2012, 2018) uses data from languages with singulative marking and/or tripartite systems of the Kipsigis type to argue in favor of a non-binary view of the mass/count distinction. He argues that the mass/count distinction is best viewed as a scalar phenomenon, with languages dividing the scale of individuation in (14) in different ways. Some languages have therefore number categories in addition to the well-known count and mass categories.

(14) *The Scale of individuation* (Grimm 2018: 547)

liquids/substances < granular aggregates < collective aggregates < individuals

The picture that emerges from the literature on singulatives is the following: singulatives are classifier-like morphemes, and they attach to a mass or collective nominal base. Crucially, singulatives have not been treated as simple allomorphs of singular number (e.g., see Nurmio 2017 for arguments against treating the Welsh singulative as an allomorph of singular morphology). The question that arises is whether what I call *inherently plural* nouns in Kipsigis are collective nouns, with the singulative being an individuating suffix, which is what Grimm (2012, 2018) has argued for the Nilotic language Turkana.

<sup>14</sup> According to an anonymous reviewer, equating thematic suffixes in Kipsigis to thematic suffixes in Spanish is not methodologically sound, since there is no clear way to show the identity between the two; he/she argues that the non-predictability of the thematic suffix in particular could be due to many reasons, independent of their status as thematic suffixes or not. Even though this is a valid objection, the main point of comparison between the two languages is the fact that these affixes are declension class markers, which, to my knowledge, is an established fact for Spanish and also an empirical fact for Kipsigis: as discussed in this section, the choice of both the correct number allomorph and the correct secondary suffix allomorph is dependent on the thematic suffix of the noun in its unmarked form. To my knowledge, the closest there is to a definition of a thematic suffix is exactly what is being described here for Kipsigis.

In this section, I show that there is no evidence that inherently plural nouns in Kipsigis are collectives, and that the singulative marker is therefore not a classifier. I argue that the Kipsigis singulative is a true allomorph of singular morphology. I show that it has significant differences from previously discussed singulatives, and conclude that the term ‘singulative’ is misleading, as it can refer to two different types of morphemes cross-linguistically: a classifier-like element (of the sort discussed in Grimm 2012, 2018; Mathieu 2012; Ouwayda 2014) or a true allomorph of singular number (as in Kipsigis, and most likely other Nilo-Saharan languages with a tripartite system of number marking).<sup>15</sup> The remainder of the section is structured as follows: in 3.1., I discuss the properties of the singulative in other languages which have motivated the singulative-as-classifier type of analysis of singulatives (which seems to be correct for these languages); in 3.2., I show that the Kipsigis singulative has significant differences from singulatives in the languages discussed in 3.1., and that inherently plural nouns in the language do not have any of the properties of collective nouns discussed in Grimm (2012, 2018); in section 3.3., I discuss the mass/count distinction in the language, which I show to be orthogonal to the classification of nouns into number classes; in section 3.4., I summarize the findings of the previous sections, and conclude that the Kipsigis singulative is a true allomorph of singular number.

### 3.1. Collectives and singulatives

A crucial component of the analysis of singulatives as classifier morphemes is the claim that the nominal base that they attach to is a mass or collective noun. While many diagnostics exist for the identification of mass nouns, the definition of collectives is more elusive. Grimm (2012, 2018) treats them as a number category that lies between mass and count nouns, with different languages dividing the Scale of Individuation in (14) in different ways; he provides a detailed semantic account of collective nouns which combines mereology with topological relations. The details of this analysis are not relevant here, but the properties that collectives have are. More specifically, he discusses the criteria in (15) as possible diagnostics for collective nouns cross-linguistically. Collectives in different languages may have slightly different semantic denotations, which would make them pattern differently for some of the diagnostics below. For example, Welsh collectives satisfy all diagnostics in (15), but Dagaare collectives only satisfy the first two (see Grimm 2012 for details).<sup>16</sup>

(15) A noun is a collective if:

- i. Native speakers have intuitions about the noun denoting a collection (in other words, the collective does not always have the same interpretation as the plural form of a count noun).<sup>17</sup>
- ii. The language has a distinct morphological class for that noun.
- iii. The noun cannot directly combine with a numeral.

A singulative attaches to such a noun yielding a unit interpretation; crucially, in many languages with collectives, the singulative form can then be pluralized. There is, therefore, a distinction in these languages between a morphologically unmarked collective and a morphologically marked plural, which behaves like a plural count noun. This pattern is the strongest piece of evidence in favor of the treatment of collectives as a special class and the analysis of the singulative as a classifier, and is illustrated below with data from Lebanese Arabic and Fox. In Arabic and Fox, gender switch to feminine and animate respectively, turns a mass or collective noun into a count noun, as shown in (16) and (17). Moreover, the count noun that

<sup>15</sup> Acquaviva (2015), in a short overview paper on the typology of singulatives, also notes that the Nilo-Saharan singulative looks like a true allomorph of singular, and treats it differently from the singulative in other languages.

<sup>16</sup> I should clarify that Grimm (2012, 2018) does not call the criteria in (15) ‘diagnostics’. These are, however, the properties that collectives can have in his analysis.

<sup>17</sup> This criterion refers to the frequent claim in Grimm (2012, 2018) that speakers conceptualize ‘collectives’ differently from regular plurals.

results from the addition of the singulative can be subsequently pluralized in both languages, yielding plural count interpretations, as shown in (16c) and (17c).

(16) *Feminine singulative suffix -ah in Lebanese Arabic* (adapted from Ouwayda 2014: 48, 52)

- |    |  |                                    |                               |
|----|--|------------------------------------|-------------------------------|
| a. | ʕaSar-t<br>squeezed-1SG<br>'I squeezed one apple or more/less than an apple.'  | <b>teffeeH</b><br>apple            | Mass interpretation           |
| b. | ʕaSar-t<br>squeezed-1SG<br>'I squeezed one apple or more/#less than an apple.' | <b>teffeeH-ah</b><br>apple-FEM     | Singular count interpretation |
| c. | stre-t      tlat<br>bought-1SG   three<br>'I bought three apples.'             | <b>teffeeH-aat</b><br>apple-FEM-PL | Plural count interpretation   |

(17) *Animate singulative suffix -a in Fox* (adapted from Mathieu 2012: 664)

- |    |  |                               |
|----|--|-------------------------------|
| a. | zhooniyaah-i<br>silver-IN<br>'silver/money'      | Mass interpretation           |
| b. | zhooniyaah-a<br>silver-AN<br>'a coin/a bill'     | Singular count interpretation |
| c. | zhooniyaah-a-ki<br>silver-AN-PL<br>'coins/bills' | Plural count interpretation   |

With this view of collectives and singulatives in mind, we would expect the following to be true in Kipsigis: the language has mass and count nouns, but in addition it also has a class of collective nouns, which are the nouns belonging to the inherently plural class. The singulative morphemes result in a unit interpretation of mass or collective nouns. We would, therefore, expect Kipsigis inherently plural nouns to satisfy at least one criterion in (15). In what follows, I show that Kipsigis inherently plural nouns do not satisfy any of the criteria in (15); rather, they have the same behavior as morphologically marked plurals (of count nouns) in the language.

### 3.2. There is no evidence that Kipsigis has collective nouns

The first criterion for identifying collective nouns in (15) is native speakers' intuitions regarding the interpretation of these nouns: their interpretation is different from that of regular plurals in the language. My consultants have never made any comments indicating that inherently plural nouns are conceptualized differently from plurals of inherently singular nouns, and inherently plural nouns pattern with regular plurals in a number of environments (to be discussed shortly).<sup>18</sup> Importantly, the citation form of all nouns

<sup>18</sup> This does not mean, of course, that such intuitions do not exist, due to well-known difficulties with accessing native speakers' intuitions, as pointed out by an anonymous reviewer. Here, I simply note that no speaker has ever made any comment indicating that these nouns are different. The remainder of the section presents specific contexts of plural use that were tested with native speakers.

in the language (with the exception of inherently plural mass nouns) is singular, morphologically marked or unmarked, as shown in (18).

- (18) a. Citation form of inherently plural *ngéendéek* ‘bean(s)’: morphologically marked singular  
 ngeend-yaan-ta-it → ngéendyáat  
 bean-SG-TH-SEC  
 ‘bean’
- b. Citation form of inherently singular *làakwéet* ‘child’: morphologically unmarked singular  
 laak-wa-it → làakwéet  
 child-TH-SEC  
 ‘child’

We, therefore, see that the citation form of inherently plural nouns is never their morphologically unmarked plural form, which we might expect if this is their base form, where they are interpreted as collectives.<sup>19</sup> This is, in fact, what we see with inherently plural mass nouns, which do have their plural form in citation, as in (19). The citation form of inherently plural nouns in Kipsigis, thus, patterns with that of count nouns (belonging to the inherently singular class) in requiring singular agreement; the morphological marking of singular on the noun (singulative vs. zero) does not matter. Crucially, this is different from the behavior of mass nouns.

- (19) Citation form of inherently plural *péek* ‘water’: morphologically unmarked plural  
 p-a-ik → péek  
 water-TH-SEC  
 ‘water’

Another observation about morphologically unmarked plurals in Kipsigis is that they easily combine with distributive predicates that reference individuals, as shown in (20) for the inherently plural nouns *màuwéek* ‘flowers’ and *pèeléeek* ‘elephants’. These readings are surprising if morphologically unmarked plurals are collectives.<sup>20</sup>

<sup>19</sup> An anonymous reviewer argues that the singular citation form of inherently plural nouns is not a valid argument, since it could be influenced by English or Swahili. First, even if a structure in language X is due to influence from language Y, it does not mean that it is not part of the grammar of language X. Furthermore, in this particular case, influence from English on Swahili on citation forms is unlikely for many reasons. First, early on in elicitation of nouns, I used the plural form of nouns in English when eliciting inherently plural nouns. For example, the first time the word for ‘elephant’ was elicited, I asked my consultant how one would say ‘elephants’ in Kipsigis, in an attempt to minimize the effect of singular from English. The speaker answered with the (marked) singular form of the noun (*pèelyáat*), and even made the comment: “this is actually one elephant, if you want to say many, you’d have to change it to *peeleeek*”. Second, the citation form was singular even for two speakers who had not completed primary school education and had a poor command of both English and Swahili. Third, Dimmendaal (2001), in a study of morphological convergence between Bantu and Southern Nilotic (where Kipsigis belongs) argues that while we see structural convergence between Bantu and Southern Nilotic languages in the verbal domain, no such convergence is observed in the nominal domain.

<sup>20</sup> An anonymous reviewer argues that this is not a valid argument against collective semantics, since even nouns like *team* in English are compatible with distributive predicates. Grimm’s (2012) analysis of collectives treats them as atoms that are organized in a spatiotemporal collection; it seems to me that if atoms are allowed to behave as individuals in the absence of dedicated morphology, as in (20) – (21), it indicates they are not treated as a collection. Furthermore, it is unlikely that nouns like *team* in English constitute the same category as plural collectives in languages with tripartite systems of number marking. For example, Grimm (2012) claims that Welsh has what he calls group nouns (i.e., nouns like *team* in English) that are singular in their unmarked form but trigger plural agreement. These nouns are different from collective/unit nouns in the language, which are plural in their unmarked form and take singulative morphology. Thus, the behavior of nouns like *team* in English is not directly relevant.

(20) Kà-à-gòon-chí-i                      Kibêet    màuwéek àgêengè                      àgêengè.  
 PAST1-1SG-give-APPL-1/2    Kibeet    flowers    one                      one  
 ‘I gave Kibeet the flowers one by one.’

(21) Kò-íit                      pèelèek                      àgéengé                      àgéengé.  
 PAST2-arrive                      elephants.NOM    one.NOM                      one.NOM  
 ‘The elephants arrived one by one.’

In general, what we observe for a number of diagnostics in the language is that what counts for semantic interpretation is the type of agreement (singular or plural) triggered on other elements in the clause; how number is morphologically marked on the noun does not matter. For example, plural agreement is required for reference to kinds (similar to how bare plurals are used for kinds in languages like English). However, how plural is marked on the noun is irrelevant: either morphologically marked plurals of inherently singular nouns or morphologically unmarked plurals of inherently plural nouns are used for kinds, as shown in (22).

(22) a. *Unmarked plural of inherently plural noun*

Pèelèek (peel-a-ik)                      kó                      tyángîik                      ché                      êechèen.  
 elephant-TH-SEC                      TOP                      animals.PL                      REL.PL    big.PL  
 ‘Elephants are big animals.’

b. *Marked plural of inherently singular noun*

Púgùusyék (pug-uus-ya-ik)                      kó                      kárâarán                      éen                      ínyêe  
 book-PL-TH-SEC                      TOP                      good.PL                      for                      you  
 ‘Books are good for you.’

Another common test for the interpretation of plurals cross-linguistically is whether they can have an inclusive interpretation in downward-entailing environments (Krifka 1989; Sauerland 2003; Sauerland et al. 2005 among others). More specifically, plurals in languages like English do not always mean ‘more than one’, but rather they include singularities in their denotation, illustrated with the common example in (23): the question contains the plural noun *children*, but the answer is positive even if John only has *one child*. If the plural were not inclusive (i.e., did not include singularities), the answer to the question in (23a) would only be positive if John had two or more children.

- (23) a. Does John have children?  
 b. Yes. John has one child.

Kipsigis plurals can also have an inclusive interpretation, as shown in (24) – (25). Importantly for our purposes, the morphological marking of plural on the noun does not matter: either morphologically marked or morphologically unmarked plurals are compatible with an inclusive interpretation. This is, therefore, one more area where unmarked plurals behave like marked plurals in the language.

(24) *Unmarked plural of inherently plural noun*

Context: We are looking for cockroaches in the house; you are looking behind the couch. I ask you:

Í-géer-é                      sólòbêek (solop-a-ik)-í?  
 2SG-see-IPFV                      cockroach-TH-SEC-Q  
 ‘Do you see cockroaches?’

Context (continued): ... you see a cockroach behind the couch and you respond:

Êe, á-géer-é      àgêengè.  
 yes 1SG-see-IPFV    one  
 ‘Yes, I see one.’

(25) *Marked plural of inherently singular noun*

Context: We are talking about Kibeet and I want to know if he has children. We have the following conversation in Kipsigis:

Tíny-è              Kíbêet              làagóok (laak-oy-ik)-í?  
 have.3-IPFV      Kibeet.NOM      child-PL-SEC-Q  
 ‘Does Kibeet have children?’

Êe, tíny-è              Kíbêet              làakwéet (laak-wa-it)      àgêengè.  
 yes have3-IPFV      Kibeet.NOM      child-TH-SEC              one  
 ‘Yes, Kibeet has one child.’

Finally, numerals freely combine with morphologically unmarked plurals of inherently plural nouns (but not with mass nouns, a point that will be discussed in detail in the next section), just like morphologically marked plurals, illustrated in (26). In addition to showing that morphologically unmarked plurals pattern with morphologically marked plurals for yet one more diagnostic, (26) shows that Kipsigis inherently plural nouns do not satisfy the criterion in (15iii), discussed in Grimm (2012, 2018). Furthermore, singulative forms cannot be pluralized in Kipsigis, making it different from Arabic and Algonquian languages (cf. examples 16 –17). As (26a) shows, there is no three-way distinction between ‘collective’ - singular count – plural count interpretation, and the morphologically unmarked form already qualifies as a plural count noun. To the best of my knowledge, this is true of all Nilo-Saharan languages with a tripartite system of number marking: singulative forms cannot be pluralized.

(26) a. *Unmarked plural of inherently plural noun*  
 peel-a-ik              somok → pèeléek sómòk  
 elephant-TH-SEC three  
 ‘three elephants’

b. *Marked plural of inherently singular noun*  
 laak-oy-ik              somok → làagóok sómòk  
 child-PL-SEC      three  
 ‘three children’

To sum up, this section has shown that inherently plural nouns in Kipsigis do not satisfy the diagnostics in (15i) and (15iii) of the previous section: there is no evidence for collective interpretations, and they can freely combine with numerals. Regarding (15i), I have argued that inherently plural nouns in their morphologically unmarked form pattern with morphologically marked plurals for a number of tests. Morphologically marked plurals of inherently singular nouns qualify as plural count nouns for all previous accounts of singulatives and collectives, including Grimm (2012, 2018). Therefore, we can conclude that inherently plural nouns in their unmarked form behave in the same way as plurals of count nouns in the language.<sup>21</sup> The next section strengthens this claim: I present a number of diagnostics for the mass/count

<sup>21</sup> According to an anonymous reviewer, while I have shown that inherently plural nouns in Kipsigis do not behave like collectives in Welsh, I have not shown that they cannot have Grimm’s (2012) denotation for Dagaare collectives, shown in (v). This denotation is a disjunctive one (i.e., nouns in Dagaare can have one of two interpretations). Since there is no evidence for collective semantics in Kipsigis, it seems redundant and entirely stipulative to posit a

distinction in the language, and I show that both the inherently singular and the inherently plural class include both mass and count nouns. In other words, there is a binary mass/count distinction that cross-cuts the morphological types of number marking. I argue that this in turn shows that Kipsigis does not satisfy criterion (15ii) for collectives, i.e., the claim that collectives constitute a distinct morphological class in the language.

Before discussing the mass/count distinction in the language, it is worth pointing out that the numberless class (i.e., those nouns that are morphologically marked in both the singular and the plural) is mysterious in a theory where singulatives are morphemes that attach to a collective noun, since there is no morphologically unmarked ‘collective’ to begin with. Even though this class is relatively small in Kipsigis, it is robust in other Nilo-Saharan languages with a tripartite system of number marking (e.g., Moodie 2019 reports that in Lopit this class contains the same number of nouns as the inherently plural class). I come back to this class in the next section.

### 3.3. The mass/count distinction in Kipsigis

I show in this section that the mass/count distinction in Kipsigis is orthogonal to the division of nouns into inherently singular and inherently plural classes: a number of tests that distinguish between count and mass nouns in the language reveal that both the inherently singular and the inherently plural class contain a mix of count and mass nouns. I do not give examples from the numberless class, which only includes count nouns.

First, count nouns, but not mass nouns, freely combine with numerals. Crucially, the nouns in (27) and (29) belong to the inherently plural class, to the exclusion of (28), which belongs to the inherently singular class. However, the nouns in (27) and (28) are compatible with numerals, unlike the noun in (29). This shows that the mass/count distinction is independent of the morphological number class of the noun. We find a similar pattern with the numeral ‘one’, which is compatible with either marked or unmarked singular count nouns, but not with unmarked singular mass nouns.

(27)	Á-mách-è 1SG-want-IPFV ‘I want three elephants.’	pèelék (peel-a-ik) elephant-TH-PL	sómòk. three	Inherently plural – count
(28)	Á-mách-é 1SG-want-IPFV ‘I want three books.’	púgùusyék (pug-uus-ya-ik) book-PL-TH-SEC	sómòk. three	Inherently singular – count
(29)	*Á-mách-è 1SG-want-IPFV ‘I want three waters.’ <sup>22</sup>	peek (p-a-ik) water-TH-SEC	sómòk. three	Inherently plural – mass

Second, the adjective *oo* (in its singular form) is interpreted as ‘big’ when it modifies a singular count noun, but as ‘a lot’ when it modifies a singular mass noun (remember that mass nouns in Kipsigis are either inherently singular or inherently plural, but always morphologically unmarked with respect to

---

disjunctive denotation. Furthermore, Grimm (2012) argues that there is morphological evidence for (v) in Dagaare, which is however absent in Kipsigis, as explained in the next section. Finally, it should be noted that Grimm (2012) is not the only analysis of the Dagaare system, and there are alternatives that do not posit collective semantics (Anttila & Bodomo 2009). It is beyond the scope of this paper, though, to evaluate the different analyses for Dagaare.

(v)  $xO \in [\text{CLUSTER}_{\text{PC}} \vee \text{SUM}]$  (adapted from Grimm 2012: 154)

<sup>22</sup> This is ungrammatical even when a coerced portion reading is intended. However, as we will see later, some mass nouns are compatible with numerals in their unmarked plural form with a portion/measure reading.

number). Again, this is independent of the morphological class of the noun as shown by the pattern in (30) – (32).

- |   |                                      |                             |
|---|--------------------------------------|-----------------------------|
| (30) <i>méesêet</i> ( <i>mees-a-it</i> )<br>table-TH-SEC<br>'a big table'                 | <i>né</i> <i>òo</i><br>REL.SG big.SG | Inherently singular – count |
| (31) <i>pèelyáat</i> ( <i>peel-yaan-ta-it</i> )<br>elephant-SG-TH-SEC<br>'a big elephant' | <i>né</i> <i>òo</i><br>REL.SG big.SG | Inherently plural – count   |
| (32) <i>púywêet</i> ( <i>puy-wa-it</i> )<br>dust-TH-SEC<br>'a lot of dust'                | <i>né</i> <i>òo</i><br>REL.SG big.SG | Inherently singular – mass  |

Third, the quantifiers *tyan/tyaana* 'how much/how many' are compatible with singular mass nouns, plural mass nouns, and plural count nouns, but the quantifier *ata* 'how many' is only compatible with plural count nouns. The independence of the mass/count distinction and the inherently singular/inherently plural distinction is nicely illustrated by the pair in (33) – (34): the nouns *chèegá* 'milk' and *nèegá* 'goats' are both inherently plural and have exactly the same morphological make-up. However, only the latter is compatible with the count quantifier *ata*.

- |  |  |   |                           |
|--|--|---|---------------------------|
| (33) <i>kò-i-lú</i><br>PAST2-2SG-drink   | <i>chèegá</i> ( <i>chee-ka</i> )<br>milk-SEC | <i>ché</i> <i>tyàn/ *átà</i><br>REL.PL how-many | <i>ámùt?</i><br>yesterday |
| 'How much milk did you drink yesterday?' |  |   |                           |
| (34) <i>í-géer-é</i><br>2SG-see-IPFV     | <i>nèegá</i> ( <i>nee-ka</i> )<br>goat-SEC   | <i>ché</i> <i>tyàn/átà?</i><br>REL.PL how-many  |                           |
| 'How many goats do you see?'             |  |   |                           |

Fourth, mass nouns in Kipsigis are incompatible with shape adjectives (Quine 1960), unlike count nouns, irrespective of the noun's number class:

- |  |                             |
|--|-----------------------------|
| (35) <i>Múgûl</i> <i>kíptúulíit</i> ( <i>kíptuul-i-it</i> ).<br>round.SG ball-TH-SEC.NOM<br>'The ball is round.' | Inherently singular – count |
| (36) <i>*Múgûl</i> <i>rôoptà</i> ( <i>roop-ta</i> ).<br>round.SG rain-SEC.NOM<br>'The rain is round.'            | Inherently singular – mass  |
| (37) <i>*Múgûl-èen</i> <i>pèek</i> ( <i>p-a-ik</i> )<br>round-PL water-TH-SEC.NOM<br>'The water is round.'       | Inherently plural – mass    |

Finally, inherently singular mass nouns are incompatible with plural suffixes, and inherently plural mass nouns are incompatible with singulative suffixes, unless the noun is coerced into a portion reading. However, not all mass nouns are compatible with number morphology, and there is some speaker variation. For example, the singular mass noun *rôoptá* (*roop-ta*) 'rain' has no plural form for two speakers, but another three speakers are perfectly comfortable with the plural *rôobwék* (*roop-wa-ik*), with the meaning 'long rainy seasons'. The nouns *chèegá* (*chee-ka*) 'milk' and *pèek* (*p-a-ik*) 'water' are examples of mass nouns without



a singular form. Furthermore, the marked (singular or plural) form of mass nouns has a standard portion reading – not a flexible one. For example, in (38) the singular form of the inherently plural *sùgàríúuk* ‘sugar’ can only mean ‘a pack of sugar’ and not ‘a grain of sugar’ or ‘a spoon of sugar’. The behavior of Kipsigis mass nouns with number morphology is, therefore, very similar to that of English mass nouns: some mass nouns tolerate coercion better than others, and the coerced meaning is usually that of a standard portion.<sup>23</sup>

- (38) a. *sugar-u-ik* → *sùgàríúuk*                      b. *sugar-yaan-ta-it* → *sùgàryáat*  
           sugar-TH-SEC                                      sugar-SG-TH-SEC  
           ‘sugar’    ‘a pack of sugar/\*a grain of sugar/\*a spoon of sugar’

It is worth pointing out that in the case of inherently plural mass nouns, the unmarked form is ambiguous between a mass interpretation and a plural portion reading. For example, (38a) can also mean ‘packs of sugar’ as illustrated in (39), where it is compatible with the count adjective ‘round’. These data are hard to capture in a theory that treats the singulative as a classifier in Kipsigis. If the singulative were a classifier, we would expect it to be present in (39) where the mass noun *sùgàríúuk* ‘sugar’ has count syntax. At the very least, we would expect to find a plural suffix. However, we do not find either, which is in fact predicted by the analysis argued for in this paper; we will come back to these data in Section 4.

- (39) *Múgùl-èen sùgàríúuk*.  
       round-PL    sugar.NOM  
       ‘The packs of sugar are round (in shape).’

In brief, a series of robust diagnostics show that the mass/count distinction in Kipsigis is similar to the mass/count distinction in more familiar languages like English, and there is no evidence that the language has a number category in addition to the mass and count categories. I, therefore, conclude that the language has three morphological classes of number marking (inherently singular, inherently plural, and numberless nouns), and a binary mass/count distinction that is orthogonal to the noun classification system.

Such a conclusion might still look suspicious due to the semantic coherence of the inherently plural class (i.e., nouns that appear more often in pairs/groups), pointed out in Section 2.3. For example, referring to the last diagnostic (incompatibility with number morphology for some mass nouns), an anonymous reviewer writes: “If inherently plural nouns which do not take a singulative form a different class (namely, a class of mass nouns), then this would seem to align well with the prediction of Grimm (2012). [...] If there are very few inherently singular mass nouns and “numberless” nouns, then the bulk of the nouns could be argued to match up to a scalar pattern, namely, inherently plural mass < inherently plural/singulative < inherently singular/plural.”

Thus, according to the reviewer, there are three number classes in Kipsigis: inherently plural mass nouns (incompatible with singulative morphology), inherently plural nouns (compatible with singulative morphology), and inherently singular nouns (compatible with plural morphology). These three ‘morphological’ classes align with ‘semantics’: mass nouns < collective nouns < count nouns. Such a view, however, faces serious empirical and conceptual challenges with respect to the Kipsigis data.

<sup>23</sup> According to an anonymous reviewer, my claim that some mass nouns cannot take number morphology contradicts my definition of the three classes in (5), where I write that inherently plural nouns form their singular with a singulative suffix, and inherently singular nouns form their plural with a plural suffix. I disagree: a) many mass nouns *do* take number morphology when coerced; calculating the type of morphology is impossible without reference to the inherently singular or plural classes, and b) for those nouns that are incompatible with number morphology (e.g., *péek* ‘water’), their morphological class is still visible: they trigger singular/plural agreement in their unmarked form (depending on their class), and (more importantly for the reviewer’s objection) speakers are able to form the singular or plural form of a given noun with the right morphology (singulative or plural suffix), but they simply judge it as ungrammatical for semantic reasons. This means that speakers still group these nouns with other inherently singular/plural nouns morphologically, but they treat them differently with respect to semantics.

First, this view crucially relies on treating singular mass nouns and the numberless class as exceptional (which is what Grimm's 2012, 2018 analysis also predicts for Turkana). These nouns are, however, by no means exceptional. Inherently singular mass nouns in Kipsigis are as frequent as plural mass nouns in my sample (the exhaustive list of mass nouns can be found in the Appendix). Furthermore, the existence of both singular and plural mass nouns in the related language Turkana is acknowledged by Grimm (2018) himself, and is a well-established fact in the typology of number in Nilo-Saharan languages. For example, it is discussed in the detailed descriptions of the related languages Turkana and Lopit (Dimmendaal 1983 and Moodie 2016, 2019 respectively), while Dimmendaal (2000: 230), in his seminal work on number in Nilo-Saharan, writes:

Whereas certain mass nouns referring to concrete entities, as well as abstract nouns, are inherently plural in Bari and Turkana, others are inherently singular. When forms are cognate between the two languages--that is, when they go back to a common ancestral form--their inherent singular or plural nature tends to correspond, as with 'milk' [...] Also, the cognate forms for 'thirst' are inherently singular (without any corresponding plural) in both Turkana and Bari [...]. However, where roots referring to identical concepts are not cognate, they may be inherently plural or singular depending on the language [...]. Such differences in conceptualization are due, it would seem, to the etymological origin of the term in question.

A similar conclusion can be reached for the numberless class. This class is a stable class across the Nilo-Saharan family (cf. Dimmendaal 2000), with the exact number of nouns in the class varying by language. For example, in a sample of 446 nouns in the Eastern Nilotic language Lopit, Moodie (2019: 60) reports that 19% of the nouns belong to the inherently plural class and 18% to the numberless class. Even though the class is smaller in Kipsigis (see Appendix for details), it is a productive noun class in the language. For example, loanwords can enter the class: the words *sig̣isyáat* 'sock' (a loanword from English), and *inḍiisyáat* 'banana' (a loanword from Swahili) are numberless nouns.

One could claim that facts from other Nilotic languages are irrelevant for the debate of whether numberless nouns are exceptional in Kipsigis. However, my analysis (to be further developed in the next section) can explain the existence of such nouns, despite their small number in the language, and can be easily extended to those languages where these classes are larger. The alternative view, however, is forced to treat these classes as exceptions. An analysis that can account for all the data is, I think, preferable to one that has to treat a large number of nouns as exceptions.

The classification advocated for by Grimm (2012, 2018) and the reviewer faces additional issues. First, as discussed in detail in sections 3.2. and 3.3, most inherently plural nouns (i.e., all non-mass inherently plural nouns) pattern with marked plurals (i.e., count nouns in Grimm and the reviewer's classification) for a number of tests. There is no evidence for their treatment as collectives semantically. Second, the definition of the first two classes on the scale ('inherently plural, compatible with singulative' vs. 'inherently plural, incompatible with singulative') cannot explain the first four mass/count diagnostics, without postulating further classes. Let me explain: the first four diagnostics (i.e., all diagnostics apart from compatibility with singulative morphology) clearly show two types of inherently plural nouns, which I call 'mass nouns' and 'count nouns'. My 'count nouns' are all compatible with singulative morphology, and thus belong to the reviewer's 'inherently plural, singulative morphology class'. There are two 'classes' of (my) mass nouns however: those that are compatible with singulative morphology (in their coerced form), and those that aren't. In the reviewer's view of morphology, this would lead us to three classes of what I call inherently plural nouns (= one class): inherently plural nouns that cannot take a singulative (= mass nouns), inherently plural nouns that can take a singulative and pattern with 'mass nouns' for the four diagnostics (=mass nouns that can take a singulative), and inherently plural nouns that can take a singulative and pattern with count nouns (=collectives).

This seems to be missing a generalization, namely that these nouns are the same in terms of morphology. The nouns in (40) – (41) and (42) – (44) respectively belong to the same class morphologically. *Pèléléek* and *mùchêléléek* in (40) – (41) both take a low-toned *à* thematic suffix, which in turn, takes the *-it* allomorph of the secondary suffix; they form their singular with the same singulative and thematic suffix (*-yaan* and *-ta*). Both their plural and their singular form are existing words in the language.

For the grammar to generate these forms, reference to morphological classes is crucial: these two nouns form a class morphologically. We see a similar pattern in (42) – (43). These nouns (which also happen to be similar phonologically) take a zero thematic suffix, which in turn, takes the *-ta* allomorph of the secondary suffix. As discussed in Section 2, nouns with a zero suffix take the *-wa* plural suffix, which is what we see in (42b) – (43b). Again, for the singular and plural forms to be generated, reference to the notion of morphological class is crucial: these nouns form a class morphologically. However, semantically it is the nouns in (40) – (42) and (41) – (43) that form a class: the former behave as count nouns for the mass/count diagnostics in the language, while the latter behave as mass nouns for the same diagnostics.

(40) a. pèel-à-ík → pèeléek elephant-TH-SEC ‘elephants’	b. pèel-yaan-ta-it → pèelyáat elephant-SG-TH-SEC ‘elephant’	‘collective’
(41) a. mùchêel-à-ík → mùchêeléek rice-TH-SEC ‘rice’	b. mùchêel-yaan-ta-it → mùchêelyáat rice-SG-TH-SEC ‘a bag of rice’	mass noun
(42) a. môok-tá → môoktá throat-SEC ‘throat’	b. môok-wá-ík → môokwék throat-PL-SEC ‘throats’	count noun
(43) a. ròop-tá → ròoptá rain-SEC ‘rains’	b. ròop-wá-ík → ròobwék rain-PL-SEC ‘long rainy seasons’	mass noun

I, therefore, conclude that the mass/count distinction in Kipsigis is orthogonal to the tripartite system of number marking in the language, a view that is probably correct for Turkana as well, since those features of the Kipsigis number system that are problematic for Grimm (2012, 2018) are also present in Turkana (e.g., inherently singular mass nouns, the numberless class, compatibility of inherently plural mass nouns with numerals; Dimmendaal 1983). However, it should be noted that this does not mean that Grimm (2012, 2018) is wrong for the other languages he discusses (e.g., Maltese), where there is indeed evidence for the presence of collectives and for the analysis of the singulative as a classifier (cf. Section 3.1.). Furthermore, parts of Grimm’s (2012, 2018) general theory are, in fact, compatible with the analysis presented in this paper, since it could explain lexicalization patterns (e.g., the lower the noun is on the Individuation Scale, the more likely it is to combine with the nominalizing head of the inherently plural class), and it could offer an insight into the historical development of the system. That analysis, however, cannot account for the synchronic facts: the status of the singulative as an allomorph of singular number, and not a classifier, and the lack of collective semantics in the synchronic grammar of Kipsigis.

### 3.4. Interim Summary

In this section, I have argued that there is no evidence that the syntax and semantics of number in Kipsigis are special. More specifically, there is no evidence for the existence of ‘collective’ nouns, and all diagnostics point towards a binary distinction between mass and count nouns. The behavior of count and mass nouns in the language can be captured by any theory of the mass/count distinction that has been proposed for English, since the mass/count diagnostics return exactly the same results for both English and Kipsigis.<sup>24</sup>

<sup>24</sup> It is, of course, possible that a more detailed investigation of the distinction in Kipsigis will reveal differences that show that the behavior of mass nouns is not identical in the two languages, but there is nothing in the current paper that would suggest the need for a new theory of mass nouns in Kipsigis.

It is also clear that the mass/count distinction is orthogonal to the morphological expression of number, which is dictated by the existence of three morphological number classes: inherently singular, inherently plural, and numberless nouns. These follow rough semantic generalizations, but this does not undermine their morphological nature: biological sex-based gender systems also follow semantic generalizations (gender of animate nouns reflects biological gender) but, to my knowledge, no theoretical account of gender claims that the classification of inanimate nouns into genders is semantic in nature. In the domain of number-based noun classification, Harbour (2007, 2011) makes a conceptually similar claim for Kiowa-Tanoan languages: number features on Num are responsible for semantic interpretation, while uninterpretable features on Class (the head that classifies nouns into distinct classes) are relevant for morphosyntax (the agreement pattern that a given noun triggers when interpreted as singular, plural, or dual), and only reflect lexical semantic tendencies.

In conclusion, there is no reason to believe that the difference between the morphological number system of English and Kipsigis lies in the semantics of number and/or the mass/count distinction in the two languages. In the remainder of the paper, I present an analysis that accounts for the difference between the two languages by postulating (uninterpretable) number features on *n* in Kipsigis, which are absent in English.

#### 4. The analysis

In this section, I first lay out my theoretical assumptions in 4.1., and I discuss why the standard theory of nominal number, with number features on NumP alone, cannot account for the pattern of number marking in Kipsigis. Then, I present my analysis in three steps: in 4.2, I argue in favor of number features on *n* in Kipsigis that divide nouns into three number classes; in 4.3, I show how the interaction of the number features on *n* with those on Num can account for the tripartite pattern of number marking; in 4.4, I briefly discuss how number agreement works in the language.

##### 3.1. Theoretical assumptions

In my analysis of the Kipsigis number system, I will be adopting the assumptions of Distributed Morphology (DM) (Halle & Marantz 1993). More specifically, I assume that word formation takes place in the syntax, in the same way that phrases are built. The input to syntax is not words or morphemes, but rather bundles of features and category-less roots that do not have any phonological material at the moment they enter the derivation. Once the syntactic operations are complete, morphological operations can manipulate feature bundles at the terminal nodes. Vocabulary items are then inserted in each terminal node according to the Subset Principle (Halle 1997), which in brief specifies that the phonological exponent of a vocabulary item can be inserted if the item contains all or a subset of the features present at the terminal node, while at the same time the item must have no feature that is absent from the node.

A DM assumption of particular importance to the topic of this paper is that lexical categories are composed of a categorizing head, and a category-neutral root (Marantz 1997, 2001; Arad 2003, 2005; Embick and Noyer 2007; Harley 2014 among others). For example, nouns are built by merging a nominalizing head (little *n*) with a category-neutral root.

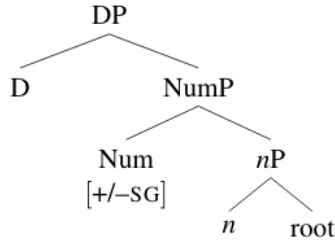
Following Ritter (1991), Carstens (1991), Bernstein (1993), among many others, I assume that number features are hosted in the functional projection NumP.<sup>25</sup> These are the number features that are

---

<sup>25</sup> An anonymous reviewer asks whether an analysis with number features on Div (cf. Borer 2005) would work. For my purposes, it is not crucial whether the relevant number features are on Div or Num: what matters is that interpretable number features are binary and that they are hosted on a functional projection above little *n* and below D. There are, however, analyses of the singulative in Ojibwe (Mathieu 2012) and Lebanese Arabic (Ouwayda 2014) that argue that Div can host singulative morphology. Even though the singulative in Kipsigis will be argued to spell out a [+SG] Num node (in certain contexts), it differs from the singulative in Ojibwe and Lebanese Arabic in being a true singular allomorph, and not a classifier (cf. Section 3). Thus, even though Borer's (2005) Div can be used instead

interpreted semantically at LF. I follow Harbour (2007) in assuming that number features are bivalent and have [+/-SG] values (later implementations use the features [+/-atomic]; Harbour 2011, 2014). The former characteristic is crucial for the analysis, but the latter is not; [+/-PL] features would have the same predictions, and the choice simply depends on which number value we want to treat as semantically marked. Since the language only displays a singular-plural distinction, however, most semantic theories of number are compatible with the Kipsigis data, as long as they accept that the features are bivalent. The structure of a simple DP in Kipsigis is, then, the following:

(44) *DP structure*



For the purposes of this paper, I assume that D is occupied by the secondary suffix in Kipsigis. This suffix has historically evolved from a specificity marker, and it is in complementary distribution with the demonstrative suffix in the singular. However, there is no strong evidence showing that this suffix is in D, and more research is needed to understand its behavior. The position of the secondary suffix and/or content of D in the language, though, is not crucial for the analysis to be presented in this paper, which works as long as the secondary suffix is in a position higher than NumP.

The exact role and syntax of thematic suffixes is left as a topic for further research, but for the purposes of this paper, I will treat them on a par with theme vowels in Romance languages (their similarities were briefly discussed in section 2.2). The DM consensus for theme vowels in Romance is that they are inserted post-syntactically as adjoined nodes to *nP* and/or NumP (see Kramer 2015: 235-243 for an overview of previous analyses and for a detailed analysis of Spanish theme vowels). Thematic suffixes in Kipsigis appear after the root (or after the nominalizing suffix in the case of nominalizations), similarly to theme vowels in Spanish. However, unlike Spanish where theme vowels appear before the plural suffix, thematic suffixes in Kipsigis are placed after the singulative or plural suffix if they are present.<sup>26</sup> In this case, the thematic suffix of the noun in its unmarked form is absent; only one thematic suffix per noun is overt at any time. Moreover, the thematic suffix of a noun in its unmarked form is dependent on the root (or nominalizing suffix in the case of nominalizations), but the form of the thematic suffix present after number suffixes is predictable by the number suffix, not the root. Therefore, in (45), the inherently plural noun *pèelèek* ‘elephants’ has an *-a* thematic suffix in its unmarked plural form, which is absent in the marked singular form. In this case, the singulative suffix *-yaan* is followed by the thematic suffix *-ta*.

(45) a. peel-a-ik → pèelèek  
elephant-TH-SEC  
‘elephants’

b. peel-yaan-ta-ni → pèelyáandáni  
elephant-SG-TH-DEM  
‘this elephant (SG)’

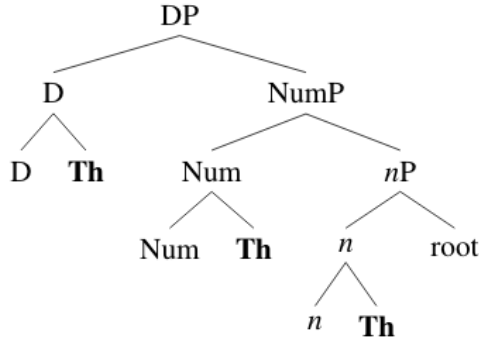
---

of Num in my analysis, the particular implementations of Borer’s system in Mathieu (2012) and Ouwayda (2014) cannot be adopted for the Kipsigis singulative.

<sup>26</sup> As was discussed in section 2, it is not entirely clear whether a thematic suffix is present between the plural suffix and the secondary suffix due to their phonological coalescence (the presence of the thematic suffix after singulatives is straightforward). However, previous analyses of the language, when the morphophonology was more transparent, postulate the presence of a thematic suffix following all plural suffixes in the language.

A theory of theme vowels that can (with a small modification) account for the basic facts in Kipsigis is Oltra-Massuet (1999) and Oltra-Massuet & Arregi's (2005) analysis of Catalan and Spanish theme vowels. According to this theory, a theme node is inserted post-syntactically adjacent to every functional node (including categorizing heads) because of a well-formedness condition on words. Extending this theory to Kipsigis, thematic nodes are inserted post-syntactically to the functional heads in (44) above, as shown in (46).

(46) *Post-syntactic thematic node insertion*



In Oltra-Massuet & Arregi's (2005) theory of Spanish, thematic nodes are present for all functional heads, as shown in (46), and the exponence of each node is determined by the closest head (usually the head it adjoins to) due to locality conditions on contextual allomorphy. This can explain why in Kipsigis the thematic suffix is dependent on the root in the unmarked form of the noun, but on the particular number suffix in the marked form. Moreover, for Oltra-Massuet & Arregi (2005), only one node in (46) is pronounced. For Spanish, this node is the one below Num (this ensures that the highest little *n* thematic node is pronounced, but the Num thematic node is not). However, as was already discussed, in Kipsigis the situation is a bit different, with the thematic suffix appearing after the number suffix, and with the noun-specific thematic suffix being absent in this case. Therefore, in both languages only one of the nodes is pronounced.<sup>27</sup> However, the choice of the node to be pronounced is different in the two languages. This implies that the choice of which thematic node to pronounce is subject to parametric variation. In Kipsigis, the thematic node to be pronounced is the node adjacent to Num. In the absence of an overt number suffix though, we saw that the form of the thematic suffix is dependent on the root. We will see in the next section that given the analysis of number morphology outlined in this paper, there are two possible analyses for this observation.<sup>28</sup>

Moving on to why the standard approach to number cannot account for the Kipsigis pattern, we return to the structure in (44). This structure is relatively uncontroversial, and number suffixes cross-linguistically are generally seen as the spellout of number features on Num. For example, in English it is accepted by most that the plural suffix *-s* is the exponent of a [–SG] Num head, while zero is the elsewhere case (spelling out a [+SG] Num head or nouns not specified for number, such as mass nouns). If we do not make any modifications to the ‘standard’ approaches to the morphosyntax of number, we would have to postulate the following spellout rules for number suffixes in Kipsigis:

<sup>27</sup> Both Spanish and Kipsigis also have athematic nouns, i.e., nouns without a thematic vowel. It is an open question whether this is a zero allomorph of the thematic suffix or whether there is something fundamentally different about the structure of these nouns.

<sup>28</sup> Kramer (2015) shows that the post-syntactic insertion of theme nodes is not enough to account for the distribution of theme vowels in Spanish; declension class features/diacritics are also needed. It is almost certain that these features are also necessary for a complete analysis of Kipsigis thematic suffixes, since the presence of different plural suffixes, which are always correlated with the thematic vowel in the unmarked form of the noun, can be seen as a declension class system. However, the morphosyntax of thematic suffixes, as well as the details of the particular plural suffix to be chosen by a noun, are not the focus of this paper.

(47) *Spellout rules for number morphology in Kipsigis (preliminary version)*

- a. Num[–SG] → plural suffix / {inherently singular nouns, numberless nouns}
- b. Num[–SG] → ∅ / {inherently plural nouns}
- c. Num[+SG] → ∅ / {inherently singular nouns}
- d. Num[+SG] → singulative suffix/ {inherently plural nouns, numberless nouns}

However, this approach, which employs the standard DP structure in (44) and contextual allomorphy (in the form of the spellout rules in 47), fails to capture a number of important characteristics of the Kipsigis number system. First, in the rules in (47), the number class of a noun is crucial in predicting the number allomorph that the noun will take, but nothing in the system accounts for why specific nouns belong to a specific number class. In other words, the system does not explain the semantic generalizations that characterize the number classes (i.e., nouns denoting entities that usually appear in groups tend to belong to the inherently plural class). A related problem is the fact that the zero exponence of Num[+SG] or Num[–SG] is accidental and arbitrary in (47), but ideally our theory should explain why nouns of each number class appear unmarked in a number value that is related to their semantics. Second, in Kipsigis, nouns of the numberless class have the same singulative suffix as the inherently plural nouns in the singular, and the same plural suffix as the inherently singular nouns in the plural. Similarly to the problem of zero exponence, the rules in (47) do not explain why this should be the case for numberless nouns – the identity of the singulative and plural suffixes with those of the inherently plural and inherently singular nouns respectively is simply an accident. Finally, the rules in (47), without further modifications, cannot account for the number morphology of mass nouns in the language: mass nouns in Kipsigis are either singular (48) or plural (49), depending on the number class they belong to. They never appear with a number suffix (singulative for the inherently plural ones and plural for the inherently singular ones) when they have their typical mass interpretation.

- (48) roop-ta → ròoptá  
rain-SEC  
'rain'

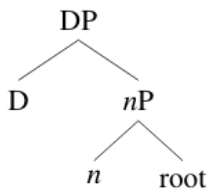
*Inherently singular mass noun*

- (49) karat-i-ik → kàràtiik  
blood-TH-SEC  
'blood'

*Inherently plural mass noun*

It is generally accepted that mass nouns are number-neutral, and in most languages they have unmarked number morphology (e.g., mass nouns in English appear in the singular, which is the unmarked form of the noun). In many syntactic approaches to the mass/count distinction, it is assumed that one correlate of the number neutrality of mass nouns and their appearance in the unmarked number form is the lack of a NumP projection in their extended projection (Borer 2005; Harbour 2007; Kucerova & Moro 2012 among others). If we assume that mass nouns have the structure in (50), there is no straightforward way to explain why some mass nouns trigger singular agreement, while others trigger plural agreement in Kipsigis.

(50) *Structure of the DP – mass nouns*



In brief, if we assume that number features are always placed on NumP as in (44), we have to assume the spellout rules in (47), which fail to capture important generalizations of the Kipsigis number system. A better analysis should be able to formalize the semantic generalizations of each number class, and derive the morphological exponence of number from those semantic generalizations. Such an analysis is outlined in the remainder of this section.

#### 4.2. Noun classes defined by number features on *n*

As was discussed in section 2.1, the Kipsigis noun classification system has similarities to gender in other languages, and I will argue that an analysis of the number classes along the lines of previous analyses of gender can explain the tripartite system of number marking in the language. Gender usually plays an active role in syntactic derivation (e.g., it determines agreement), which is why most previous analyses of gender postulate the existence of gender features within the DP. Any analysis of gender, then, should elaborate on the nature of these features and on their exact position in the syntactic structure. Regarding the first question, there is disagreement in the literature on whether these features should be privative or bivalent, interpretable or uninterpretable, but it is generally accepted that the features responsible for gender in a given language reflect the semantic notion based on which classification takes place. For example, it is generally accepted that in a sex-based gender system of the Indo-European type, some sort of [FEM] or [MASC] features should be at play. As for the position of these gender features, there have been a number of different proposals. Among the proposals that see gender as a property of nouns, there are those that postulate gender features on the nominalizing head *n* (e.g., Ferrari 2005; Lowenstamm 2008; Acquaviva 2009; Kramer 2015), and those that see gender as an inherent property of the nominal root (e.g., Alexiadou 2004; Carstens 2010); the latter usually include a mechanism that translates this property of nouns into syntactic features that can participate in agreement. There are also proposals that place gender features on various projections within the DP: Ritter (1993) suggests that Num, in addition to number features, can also host gender features, Piccolo (1991) proposes the existence of a functional projection GenP responsible for gender features, while Steriopo & Wiltschko (2010) suggest that gender features can be distributed on a variety of positions within the DP (the root, the nominalizing head, and D).

I suggest that in the case of Kipsigis the noun's inherent features are number features, which are hosted on the nominalizing head, little *n*. Number features on *n* have been suggested in a number of previous studies, but not in the form of features used robustly for noun classification (e.g., Alexiadou 2011 for Greek; Kramer 2009, 2016 for Amharic; Acquaviva 2008 for lexical plurals in various languages; Lecarme 2002 for Somali). There are three possible types of *n* in the language, which divide nouns into the three number classes:

##### (51) *Kinds of n in Kipsigis*

- a. *n*[+SG]: inherently singular nouns
- b. *n*[-SG]: inherently plural nouns
- c. plain *n* (no number features): numberless nouns

Each noun in Kipsigis belongs to one number class only, which means that a nominal root can appear in the context of only one of the nominalizing heads in (51). This means that there must be a mechanism in the grammar responsible for matching a root with the right nominalizing head. Within the DM literature, this mechanism usually has the form of various types of licensing conditions for roots. The data presented in this paper do not support or contradict any particular theory of root licensing, and the reader is referred to Acquaviva (2009) and Kramer (2015) for different implementations of licensing conditions of roots under nominalizing heads with particular features.

Turning to the choice of little *n* as the locus of inherent number features in Kipsigis, there are a number of arguments in favor of this position. First, even though most syntactic categories in the language are inflected for number, only nouns show the tripartite system of number marking (in other words, only



nouns are divided into number classes). For example, adjectives, which show a morphological distinction between singular and plural, are unmarked in the singular, and mark the plural with a plural suffix, as seen in (52). Crucially, there is no adjective that is unmarked in the plural, but marked in the singular, which is the case for a whole class of items in the nominal domain.

- (52) a. *làakwéet nè tóròor*  
 girl.SG REL.SG tall.SG  
 ‘a tall girl’
- b. *làagóok sómòk chè tóròor-èen*  
 girl.PL three REL.PL tall-PL  
 ‘three tall girls’

Since these number classes are limited to the syntactic category of nouns, it is expected for class features to be associated with the categorizing head that turns roots into nouns, i.e., little *n*.

Second, as Kramer (2015) points out in her argument in favor of gender features (such as [+/-FEM] for sex-based gender systems) on *n*, the Kipsigis number class system is root-specific (i.e., the number class of a certain noun is idiosyncratic and depends on the root), and exhibits paradigmatic gaps in the allowable combinations of *n* and roots (i.e., not all roots are possible with all types of *n* and vice versa), which are often cited as characteristics of the relationship between the root and a categorizing head (Marantz 2001).

The third and strongest argument in favor of class features on *n* comes from the behavior of derived nominals with respect to number morphology. With the exception of a couple of nominalizing prefixes, derived nominals in the language are formed by the addition of a nominalizing suffix to a verbal or adjectival stem. This nominalizing suffix is followed by one of the thematic suffixes also encountered with common nouns (the thematic suffix in this case is dependent on the nominalizing suffix), followed by a secondary suffix, as shown in (53).

- (53) *cham-an-a-it* → *chámànéet* (patient nominalization, verb *chám* ‘to love’)  
 love-PAT-TH-SEC  
 ‘loved one (SG)’

Interestingly, derived nominals follow the tripartite system of number marking: some nominalizing suffixes turn verbs/adjectives into inherently singular nouns, while others turn them into inherently plural nouns. For example, patient nominalizations, like the one in (53) above, are derived by the addition of the suffix *-an* to a verbal stem. This suffix forms a singular noun, which forms its plural by the addition of the *-oosyek* plural suffix (also used with common nouns; cf. table 3), as seen in (54). It, therefore, has the same behavior as inherently singular nouns.

- (54) *cham-an-oos-ya-ik* → *chámànòosyék*  
 love-PAT-PL-TH-SEC  
 ‘loved ones (PL)’

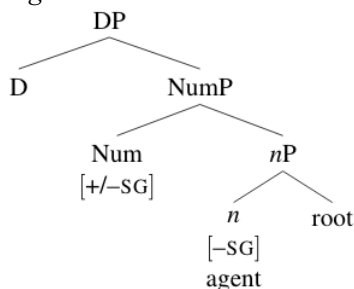
Agent nominalizations, on the other hand, are derived by the addition of a segmentally null [+ATR] suffix to the verbal stem, which forms a plural noun, as seen in (55). This noun forms its singular by the addition of the singulative suffix *-iin* (followed by the thematic suffix *-ta*), which is the same singulative suffix used with human common nouns (cf. example 12). We see, thus, that agent nominalizations have the same behavior as inherently plural nouns.

- (55) *choor-∅<sub>[+ATR]</sub>-iin-ta-it* → *chòoríndét* *choor-∅<sub>[+ATR]</sub>-i-ik* → *chòoríik*  
 steal-AG-SG-TH-SEC steal-AG-TH-SEC  
 ‘thief (SG)’ ‘thieves (PL)’

Verbs and adjectives are not divided into number classes, and therefore the inherent number value of derived nominals must come from the number value of the nominalizing suffix. The following quote from Toweett (1975: 59) illustrates the fact that these suffixes are specified for number: “the functional and number suffixes coincide; to separate them is not realistic” (where ‘functional’ refers to nominalizing suffixes).

Adopting the relatively uncontroversial assumption that nominalizing suffixes are the spell-out of a nominalizing head *n* that merges with an *xP* to form a noun (Arad 2003; 2005 among others), the inherent number value of suffixes implies that *n* carries number features in Kipsigis. For example, agent nominalizations in Kipsigis have the structure in (56), where a nominalizing suffix with agent semantics and a [–SG] feature merges with a *vP*.

(56) *Agent nominalizations*



We do not find any derived nominals in the numberless class, i.e., the class where nouns are marked in both the singular and the plural. Moreover, the numberless class is small in Kipsigis: even though my data do not represent an exhaustive list of nouns in the language, only ten nouns in my field notes belong to this class. These observations suggest that nominalizing heads without any number features (plain *n* in 23) are not common in Kipsigis. However, there is nothing in the theory that would explain why the numberless class should be less productive than the inherently singular and inherently plural classes. This is a welcome result, because in other Nilo-Saharan languages with the same three number classes discussed in Dimmendaal (2000), the numberless class seems to be more productive than in Kipsigis, as was already discussed in Section 3.3. For example, Dimmendaal (1983), in his detailed study of Turkana (Eastern Nilotic; Kenya), shows that the numberless class contains many nouns, and also gives various examples of derived nominals that belong to this class. Therefore, the small size of the numberless class in Kipsigis is probably an accident.

Moving on to the nature of the features involved in classification in Kipsigis, I follow Harbour (2007) in assuming that number features on little *n* are uninterpretable, in the sense that they are not assigned a semantic interpretation at LF. In fact, if they were interpretable, there would be a semantic clash at LF with the interpretable number features on Num. Such a view forces me to adopt the assumption that uninterpretable features that are not ‘checked’ by an interpretable counterpart on another head do not lead the derivation to crash, as is assumed in standard Minimalism (Chomsky 2001 among others); rather, it is unvalued features that lead the derivation to crash (Harbour 2007; Carstens 2010; Kramer 2015).

The choice of [+/-SG] features as the inherent features at play is motivated by the semantics of the Kipsigis noun classes. As was discussed in Section 2.1., the Kipsigis noun classes follow certain semantic generalizations: nouns that usually appear in units tend to belong to the inherently singular class, while nouns that usually appear in groups tend to belong to the inherently plural class. This automatically follows from the [+SG] and [–SG] features, because irrespective of the details of the semantics of these features, the former are associated with unit semantics, while the latter are associated with sum semantics in all theories (see Harbour 2007, 2011, 2014 for specific semantic implementations). As for the numberless class, there is no clear generalization, which is consistent with the absence of number features on little *n* for these nouns.

A clarification is, however, needed for the notion of *semantic coherence*. As has been pointed out at various points throughout the paper, my claim about the Kipsigis classes is that they follow mere semantic

*tendencies*; that is, there is no clear one-to-one correspondence between the lexical semantics of a noun and the number class that it belongs to. Most notably, mass nouns could belong to either the inherently singular or the inherently plural class. However, it is still true that the majority of nouns in a given class will follow the semantics of their little *n* in (51). The fact that there are exceptions is not a problem for this theory. First, the fact that number features are uninterpretable means that there is no semantic clash at LF if, say, a mass noun combines with a [+SG] little *n*, as is the case with inherently singular mass nouns. Second, gender features cross-linguistically often come in semantic and formal flavors: e.g., feminine in German is usually used to refer to female entities, but is also used for inanimate nouns which cannot be female. The exact relationship between the semantic and formal use of these features is a matter of debate, but it shows that gender features of the type proposed here for Kipsigis need not have a strict semantic interpretation (see Hammerly 2018 for an interesting proposal that derives gender (un)interpretability from a single type of features). Therefore, the number features proposed here for Kipsigis noun classification capture the rough semantic tendencies, without imposing a strict correspondence between lexical semantics and morphological noun class, which is exactly what we observe in the data.<sup>29</sup>

Finally, the existence of three number classes (as opposed to two) can be explained by binary, but not by privative features. Thus, the Kipsigis pattern strongly supports theories of number that posit binary features (e.g., Harbour 2007, 2011, 2014), and goes against theories that argue for privative number (e.g., Nevins 2011).

#### 4.3. The tripartite system of number marking

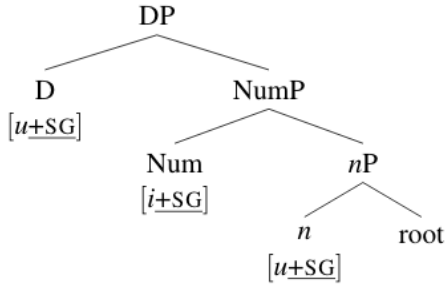
We have seen so far that the nominalizing head has uninterpretable number features ([+SG] for inherently singular nouns, [−SG] for inherently plural nouns, and none for numberless nouns), and that the functional projection NumP has interpretable number features [+/-SG], which determine whether a noun is to be interpreted as singular or plural at LF.<sup>30</sup> Moreover, D is occupied by the secondary suffix, which agrees with Num in number (more details on this in the following section). This means that we have the following possible structures for the singular and plural of nouns from the three classes.

<sup>29</sup> It should be noted that even though Harbour's (2007) analysis of Kiowa (to be further discussed in section 5) is similar to my analysis in having uninterpretable binary number features on little *n*, the semantic coherence of the Kiowa classes seems to be stronger than that of the Kipsigis classes. This difference between Kiowa and Kipsigis could be analogous to the difference between strict semantic gender languages like Tamil, and languages like Spanish, where gender is assigned formally as well as semantically. Another possible explanation for the divergence of Kipsigis and Kiowa in this domain is the difference in feature inventories between the two: Kiowa (and Jemez) make more number distinctions than Kipsigis, which might play a role in the feature combinations that are available for noun classification.

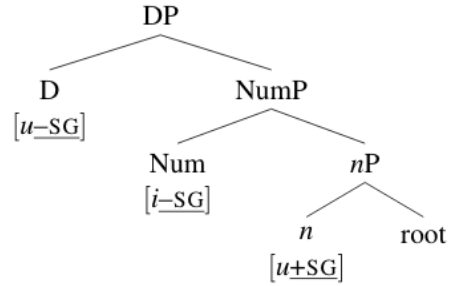
<sup>30</sup> The possibility of two different types of number features (on little *n* and Num) has been suggested before (e.g., Heycock & Zamparelli 2005). However, the Kipsigis data are different insofar as every noun in the lexicon must appear with a numbered *n* (or plain *n* for numberless nouns), with the choice being determined by the particular root in question. This, I believe, is also at the heart of the difference between languages like English or Greek and Kipsigis: while number features on little *n* have been proposed for these languages as well (e.g., Acquaviva 2008; Alexiadou 2011; Kouneli 2018), these features are either idiosyncratic (as in the case of lexical plurals) or perform a specialized semantic function (as in Heycock & Zamparelli 2005 or Kouneli's 2018 account of pluralized mass nouns in Greek). In Kipsigis, on the other hand, each nominal root must appear in the context of a little *n* with number features; in other words, numbered nominal heads provide an exhaustive classification of the lexicon, unlike numbered *n*'s in English and Greek. 'Exhaustive' classification is, of course, reminiscent of the fact that each noun must have a gender (masculine, feminine, or neuter), and the question of why Kipsigis is different from English is similar to the question of why some languages have a gender system but others don't. While this is a very important question, providing a satisfactory answer is beyond the scope of this paper. It seems relevant that even languages like English show signs of what Tiersma (1982) calls 'local markedness' with nouns that correspond to the Kipsigis inherently plural nouns; thus, Kipsigis has grammaticalized a distinction that has more subtle effects in languages like English.

(57) *Inherently singular nouns*

a. *Singular*

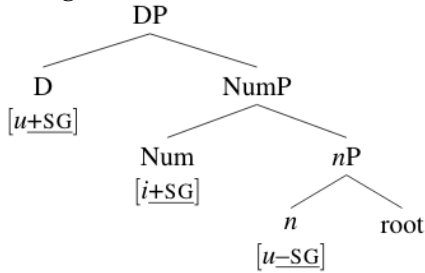


b. *Plural*

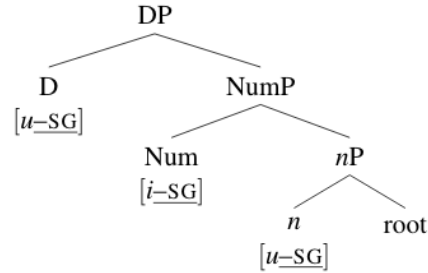


*Inherently plural nouns*

c. *Singular*

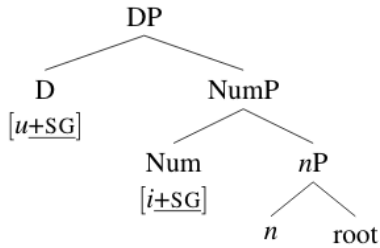


d. *Plural*

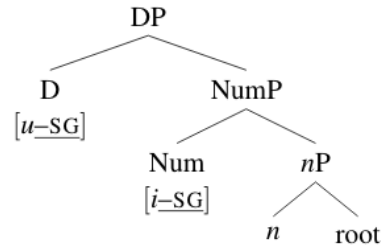


*Numberless nouns*

e. *Singular*



f. *Plural*



With the syntactic structures in place, we need to define the Vocabulary Insertion (VI) rules to account for the exponence of number morphology for each class of nouns. First, the nominalizing head has no overt exponent for underived nouns. Second, number suffixes in the language spell out the Num head: a [+SG] head is spelled out as a singulative suffix, and a [-SG] head as a plural suffix. Third, a [+SG] D head is spelled out as the singular form of the secondary suffix, while a [-SG] D head is spelled out as the plural form of the secondary suffix:

(58) *VI rules for the Kipsigis DP:*

- a.  $n \leftrightarrow \emptyset$
- b.  $\text{Num}[+SG] \leftrightarrow \text{singulative suffix}$
- c.  $\text{Num}[-SG] \leftrightarrow \text{plural suffix}$ <sup>31</sup>

<sup>31</sup> There are a variety of singulative and plural suffixes in the language. I assume that the choice of a particular suffix is dependent on the *nP*, with  $\text{Num}[+/-SG]$  having different allomorphs depending on the *nP* involved.

- d. D[+SG]  $\leftrightarrow$  *it*
- e. D[-SG]  $\leftrightarrow$  *ik*

The rules in (58) can easily account for all the cases where number is marked (structures b, c, e, and f in 57 above). In all these cases, there is a singulative or plural suffix present, which spells out the respective Num head. What about (57a) and (57d) though, i.e., the unmarked forms of inherently singular nouns and inherently plural nouns respectively? We observe that (57a) and (57d) are the only two structures where the number features on *n* are exactly the same as those on Num: there are two adjacent [+SG] features in the singular of inherently singular nouns and two adjacent [-SG] features in the plural of inherently plural nouns. I suggest that for reasons of morphological dissimilation (which is common cross-linguistically; see Nevins 2012 for an overview of such morphological dissimilation phenomena), when two identical [ $\alpha$ SG] features are structurally adjacent, the highest terminal node hosting this feature is deleted via the post-syntactic operation of *obliteration*, which has been proposed independently to account for allomorphy in the *g-/z-* constraint in Basque (Arregi & Nevins 2006, 2012), and for contextual allomorphy in Romance clitics (Calabrese 2011; Pescarini 2010).

#### (59) *Obliteration*

Delete an [ $\alpha$ SG] Num node when it dominates an [ $\alpha$ SG] node.

As a result of (59), which applies post-syntactically, the Num node in (57a) and (57d) is deleted before VI, and no number suffix is inserted, hence the unmarked form of the noun.<sup>32</sup> It might seem counterintuitive that it is the node with interpretable number features that is deleted, but it captures the intuition that a noun does not inflect for number in the number value that is already included in its lexical meaning. A complete derivation for the singular and plural form of an inherently plural noun is shown in (60) – (61). In (60), the marked form of the noun, all terminal nodes in the structure are assigned phonological material at VI. In (61), on the other hand, the Num node has the same number features as *n* and is, thus, deleted post-syntactically via the operation of obliteration. As a result, no vocabulary item is inserted in this node.<sup>33</sup>

As for the exponence of thematic suffixes, they are the spell-out of theme nodes inserted post-syntactically; the only node to be pronounced though is the one adjacent to Num (the details were discussed in 4.1.). In the unmarked form of the noun, the only overt thematic suffix is the one predicted by the root.

<sup>32</sup> Obliteration is different from *Impoverishment*, which is the operation commonly used in DM for morphological dissimilation phenomena. Impoverishment deletes a feature from a terminal node, but a vocabulary item matching the resulting features has to be inserted in that terminal node at VI, while obliteration deletes the terminal node altogether and no vocabulary item is inserted. In this particular case, I prefer obliteration because there is no overt material on Num in the unmarked forms of inherently singular and inherently plural nouns. However, impoverishment could still be used: one could say that the [+SG] or [-SG] feature on Num is deleted, and the vocabulary item ‘inserted’ for a feature-less Number node is zero. A short note on terminology is also in order. I have adopted here Arregi & Nevins’ (2012) terminology, who distinguish between *Impoverishment* and *Obliteration*. Harbour (2003) uses the terms *Impoverishment at the node* vs. *Impoverishment of the node* for the same distinction, while Daniel Harbour (p.c.) informs me that Bonet (1991) had already discussed the distinction: *Impoverishment* is deletion of a feature that leaves a node intact, versus deletion of the node itself.

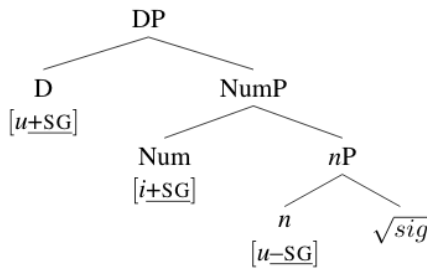
<sup>33</sup> An anonymous reviewer points out that obliteration of the Num node is not the only way to explain why only one VI is used in those cases where the number features on *n* and Num match. He/she mentions Embick’s (2010) *Pruning* and Radkevich’s (2010) *Rebracketing* operations as possible alternatives. In principle, both of these operations can get the facts right. However, what is crucial in those cases where number features on Num and little *n* match is that the only VI inserted is, in fact, zero (i.e., the Num node is usually spelled out as a singulative/plural suffix, but in that case there is no overt number affix). I believe that obliteration of the node captures this generalization in a more straightforward way. Furthermore, Pruning and Rebracketing are general morphological rules, while obliteration – as a form of impoverishment – is more specific to prohibitions of particular feature combinations (in this case, it is motivated by a ban against adjacent identical number feature specifications).

Given the analysis outlined here, there are two possible explanations for this fact. The first possibility is that the insertion of theme nodes is ordered after the operation of obliteration. As a result of the deletion of the Num node, there is no theme node inserted in this position, and the node adjacent to little *n* is the one that is spelled out overtly instead. The second possibility is that the theme node is inserted adjacent to Num before obliteration takes place, and it is the node that is spelled out at VI. However, because the Num node is deleted before VI, the allomorph chosen at VI for the Num theme node is determined by the next closest element, which in this case is the *n* – root. More data are needed to determine which analysis is the right one, and I will prefer the first one in this paper. The ordering of operations and specific VI rules are illustrated in the derivations in (60) – (61).

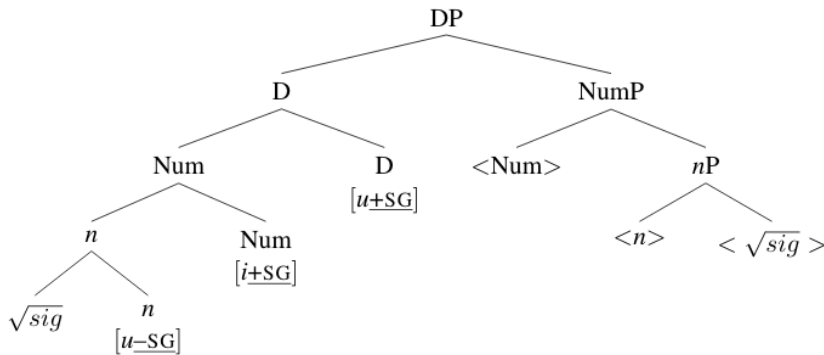
(60) *Singular form of inherently plural noun sigiik ‘parents’*

a. Surface form:  
sig-iin-ta-it → sigiindét  
parent-SG-TH-SEC  
‘parent’

b. Syntactic structure<sup>34</sup>



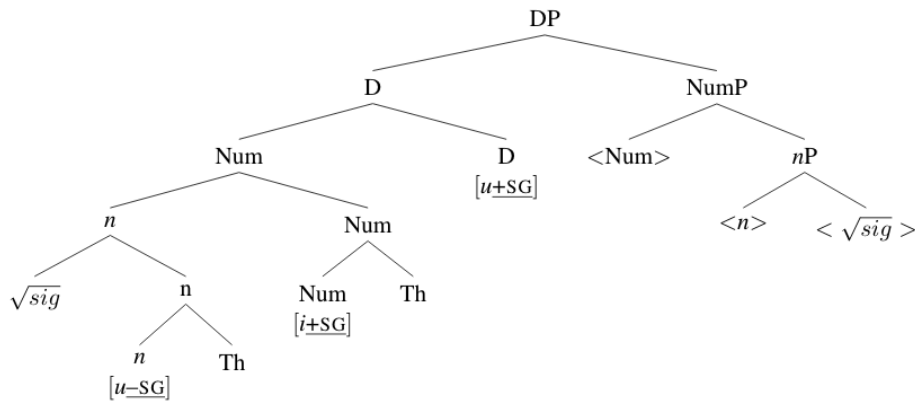
c. Word creation (=complex head) via Head Movement<sup>35</sup>



<sup>34</sup> D enters the derivation with an unvalued [SG] feature, which is valued through Agree via a mechanism that is outlined in section 4.4. An observant reader might wonder why [+SG] on D is not deleted via obliteration because it has an identical number value as Num, which is adjacent. I assume that there is a fundamental difference between D's number features and those on Num and *n*, which is due to the fact that the latter come out valued from the lexicon, while the former get their value via Agree.

<sup>35</sup> Daniel Harbour (p.c.) asks whether Head Movement is important for my analysis, since there are alternative ways to get morphemes to form a word (e.g., spanning). What is crucial for my analysis is the existence of number features on *n* for noun classification, and the deletion of the Num node when its features match those on *n*. As far as I am concerned, any theory of word formation that is compatible with these two assumptions is a possible alternative.

d. Post-syntactic operations: Insertion of theme nodes<sup>36, 37</sup>



e. Vocabulary Insertion:

$\sqrt{sig} \leftrightarrow sig$

$n[-SG] \leftrightarrow \emptyset$

$Num[+SG] \leftrightarrow iin$

$Th(Num) \leftrightarrow ta$

$D[+SG] \leftrightarrow it$

(no other theme node in the structure is pronounced)

f. Phonological operations: sigiindét

(61) *Plural form of inherently plural noun sigiik ‘parents’*

a. Surface form:

sig-i-ik  $\rightarrow$  sigiik

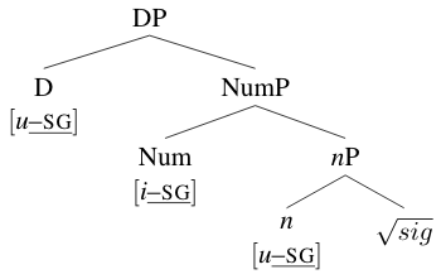
parent-TH-SEC

‘parent’

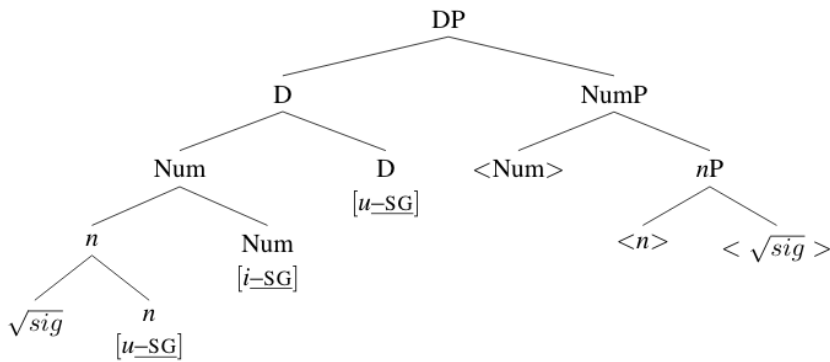
<sup>36</sup> I do not represent here the insertion of a theme node on D, but according to Oltra-Massuet’s theory (1999), there is a theme node inserted for each functional head.

<sup>37</sup> Daniel Harbour (p.c.) asks whether Theme nodes are necessary. The short answer is no, since there are alternative ways to get the morphological facts right. For example, he suggests the following: there might exist morphemes that realize the category head in addition to some that realize features on that head. Thematic suffixes would express just Num, while singulative/plural suffixes would express [+/-SG] features. I use Th nodes here for the simple reason that they have been used in DM to account for thematic suffixes in Romance, which have some similarities to thematic suffixes in Kipsigis. Thematic suffixes in Kipsigis and Romance are not, however, identical, and it is possible that an analysis along the lines of the one suggested by Daniel Harbour (p.c.) is better for the Kipsigis facts. Further research into the behavior of thematic suffixes in Kipsigis (and possibly other Kalenjin dialects, where their form also depends on Case; Creider & Creider 1989) can help distinguish between the two analyses for thematic suffixes.

b. Syntactic structure:

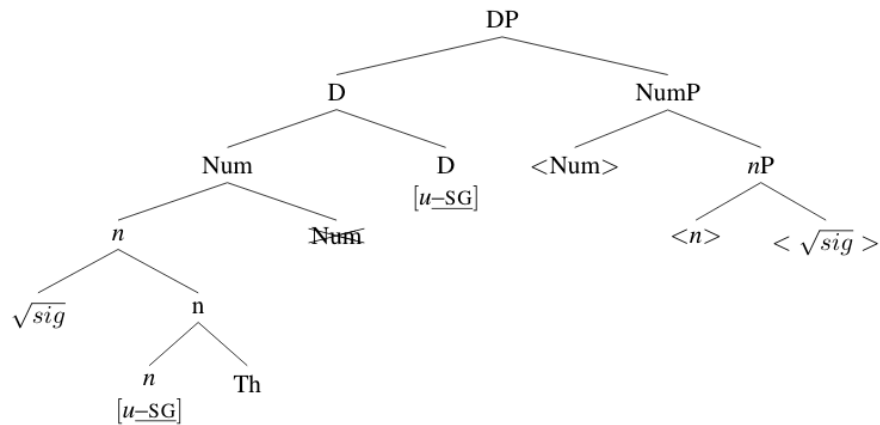


c. Word creation (=complex head) via Head Movement



d. Post-syntactic operations:

- i) Obliteration of the Num node
- ii) Insertion of theme nodes



d. Vocabulary Insertion:

$\sqrt{sig} \leftrightarrow sig$

$n[-SG] \leftrightarrow \emptyset$

$Th(n) \leftrightarrow i$

$D[+SG] \leftrightarrow ik$

(no other theme node in the structure is pronounced)

e. Phonological operations: sigiik



Let's now turn to the number morphology of mass nouns. As was discussed in 4.1, it is generally accepted that mass nouns lack a NumP projection. Since number suffixes are always the spellout of Num in Kipsigis, it is predicted that mass nouns will appear without a number suffix, that is mass nouns in Kipsigis are predicted to be morphologically unmarked with respect to number (in their mass interpretation). However, since *n* can host number features in the language, it is predicted that mass nouns can be either plural or singular in their unmarked form, depending on whether the features on *n* are [-SG] or [+SG]. This prediction is borne out: mass nouns are unmarked when they have their typical mass interpretation (they could be marked only when their meaning is coerced into that of a portion/unit, in which case we can assume the presence of Num), and belong to either the inherently plural or the inherently singular class, as shown in (62) – (63).

- (62) puy-wa-it → púywêet *Inherently singular mass noun*  
dust-TH-SEC  
'dust'
- (63) karat-i-ik → kàràtiik *Inherently plural mass noun*  
blood-TH-SEC  
'blood'

If we assume that Num is syntactically present when mass nouns are coerced into a count interpretation (e.g., Borer 2005), the analysis can also explain why mass nouns are morphologically unmarked even when coerced, a point briefly discussed in section 3.3. For example, the unmarked plural of the inherently plural noun *sùgàrúuk* 'sugar' is ambiguous between a mass and a count interpretation. In my analysis, Num is absent when the noun has a mass interpretation, while it is syntactically present when it has a count interpretation. For an inherently plural mass noun like *sùgàrúuk* 'sugar', this means that there will be no plural morphology in the mass interpretation (Num is absent, and hence no VI is inserted), and in the plural count interpretation Num[-SG] is syntactically present, but it is deleted post-syntactically because it bears the same features as the [-SG] *n*. A singulative suffix will be present in the singular count interpretation, because in this case the Num[+SG] node is not deleted (no matching features).

One might wonder whether we find mass nouns that belong to the numberless class, which is a possibility given the analysis outlined here. However, our theory predicts that a numberless mass noun would be unmarked for number, which means it would look identical to inherently singular mass nouns on the surface. The only way to differentiate between a numberless mass noun and an inherently singular mass noun is through their morphological behavior when coerced into a count reading: a numberless mass noun would have a singulative suffix in the singular count interpretation and a plural suffix in the plural count interpretation, while an inherently singular mass noun would be unmarked in the former case, and would have a plural suffix in the latter. I have not found any examples of numberless mass nouns in my data. However, according to Dimmendaal (2000), some Nilo-Saharan languages (which have the same three number classes as Kipsigis) have a class of nouns that behave exactly like numberless mass nouns are expected to behave in our theory, but further research on these languages is needed to determine the exact interpretation and countability status of nouns like (64a). As you can see in the Shatt (Daju; Sudan) example below, the unmarked form of the noun for 'teeth' has what looks like a collective interpretation (64a), the form with a singulative suffix has a singular count interpretation (64b), and the form with a plural suffix has a plural count interpretation (64c). Therefore, the analysis presented in this paper correctly predicts the existence of this class of nouns in languages with a Kipsigis-type number system.

- (64) a. nyix b. nyix-te c. nyix-ke  
'set of teeth' 'tooth(SG)' 'teeth(PL)'  
[example from Tucker & Bryan 1966: 235 via Dimmendaal 2000: 242]

To sum up, nouns in Kipsigis belong to different classes depending on the number features present on *n*, and appear unmarked when they merge with a Num head that has the same number features as *n* because in this case Num is deleted post-syntactically for economy reasons. The operation of obliteration formalizes the intuition that nouns in Kipsigis appear unmarked in the number value that is already included in the nominal meaning. Therefore, the analysis presented here not only captures the semantic generalizations characterizing the Kipsigis noun classes, but also links these semantic generalizations to the markedness pattern that we observe in number marking. Furthermore, this analysis can account for the number morphology of mass nouns and derived nominals in the language. In the following section, I briefly discuss how number agreement works in the language.

#### 4.4. Number agreement

In the theory presented in this paper, number features are present on both the nominalizing head *n* and the functional head Num, and these features can sometimes be in conflict with each other (in the plural of inherently singular nouns and in the singular of inherently plural nouns). This raises the question of how the features in each position affect agreement (within and outside the DP), which is especially interesting given the increasing interest in mixed agreement patterns with hybrid nouns in various languages (Landau 2016; Despić 2017; Kucerova 2017 among many others). We will see, however, that the Kipsigis agreement system can be easily explained by standard Minimalism assumptions about Agree. Agreement with mass nouns, however, shows that both the features on little *n* and those on Num can take part in agreement in the language, which indicates that we can find different agreement patterns in other languages with number-based noun classification if there are differences in the structure of the DP of those languages and/or in the way Agree works cross-linguistically. I show in section 5 how Kiowa and Jemez are examples of languages with a number-based classification system, but an agreement pattern different from Kipsigis.

In Kipsigis, a number of elements agree with the head noun in number. We have already seen, for example, that the secondary suffix has a singular and a plural form. Number agreement reflects the semantic number of a noun (whether a noun is semantically interpreted as singular or plural), and not its inherent number value. In other words, the number class that a noun belongs to and the morphological expression of number for that noun (singulative vs. plural marking) do not have any effect on agreement. For example, the inherently plural noun *pèelék* ‘elephants’ and the inherently singular noun *làakwét* ‘child/girl’ belong to different number classes, but take the same singular demonstrative suffix *-ni* when interpreted as singular despite the fact singular is marked with a singulative in the former case (65a), but is unmarked in the latter (65b). Similarly, they both take the plural demonstrative suffix *-chu* when interpreted as plural, despite the different kind of plural marking (66).

- (65) a. peel-yaan-ta-**ni**                      (ageenge) → pèelyáandá**ni** (àgêenge)  
          elephant-SG-TH-DEM              one  
          ‘this (one) elephant’

- b. laak-wa-**ni**                      (ageenge) → làakwà**ni** (àgêenge)  
          child-TH-DEM              one  
          ‘this (one) child’

- (66) a. peel-a-ik-**chu**                      (somok) → pèelée**chù** (sómòk)  
          elephant-TH-SEC-DEM              three  
          ‘these (three) elephants’

- b. làak-oy-ik-**chu**                      (somok) → làagóo**chù** (sómòk)  
          child-PL-SEC-DEM              three  
          ‘these (three) children’

The secondary and demonstrative suffixes are not the only elements that agree with nouns in number in Kipsigis. Within the DP, relativizers (67), adjectives (68), and possessive suffixes (69) also agree with the head noun in number (possessive suffixes agree in person and number with the possessor, and in number with the possessee).

(67) *Number agreement with relativizers*

- a. làakwéet      **nè**      á-chám-é  
 girl.SG      REL.SG      1SG-like-IPFV  
 ‘the girl that I like’
- b. làagóok **chè**      á-chám-é  
 girl.PL REL.PL 1SG-like-IPFV  
 ‘the girls that I like’

(68) *Number agreement with adjectives*<sup>38</sup>

- a. làakwéet      nè      tóròor  
 girl.SG      REL.SG tall.SG  
 ‘a tall girl’
- b. làagóok sòmòk chè      tórôor-èen  
 girl.PL three REL.PL tall-PL  
 ‘three tall girls’

(69) *Number agreement with possessive suffixes*

- a. laak-wa-it-nyuun → làakwéenyuun  
 girl-TH-SEC-POSS1SG.SG  
 ‘my girl’
- b. laak-oy-ik-chuuk → làagóochhùuk  
 girl-PL-SEC-POSS1SG.PL  
 ‘my girls’

As for elements outside the DP, predicative adjectives and nominals agree in number with their DP subject (70 – 71). Verbs agree with their subject in person and number, but there is no morphological distinction for number in the 3<sup>rd</sup> person. There are three verbs, however, that are suppletive in the singular and plural, such as the verb ‘to run’, shown in (72).

(70) *Number agreement with predicative adjectives*

- a. Tórôor làakwèet.  
 tall.SG girl.SG.NOM  
 ‘The girl is tall.’

<sup>38</sup> Adjectives in the language are introduced by the same element that introduces relative clauses. Adjectives, however, have different properties from (stative) verbs, and constitute a distinct morphosyntactic category (Kouneli 2019a: Chapter 4).

- b. Tórôor-**èen**      làagôok.  
           tall-**PL**            girl.**PL.NOM**  
           ‘The girls are tall.’

(71) *Number agreement with predicative nominals*

- a. Kâanéetîndét      Kíbêet.  
           teacher.**SG**      Kibeet.**NOM**  
           ‘Kibeet is a teacher.’
- b. Kâanéetîik      Kíbêet      ák      Chèebêet.  
           teacher.**PL**      Kibeet.**NOM**      and      Cheebeet  
           ‘Kibeet and Cheebeet are teachers.’

(72) *Subject-Verb agreement*

- a. Lábât-í      Kíbêet.  
           run.3SG -IPFV      Kibeet.**NOM**  
           ‘Kibeet is running.’
- b. Rúày      Kíbêet      ák      Chèebêet.  
           run.3PL.IPFV      Kibeet.**NOM**      and      Cheebeet  
           ‘Kibeet and Cheebeet are running.’

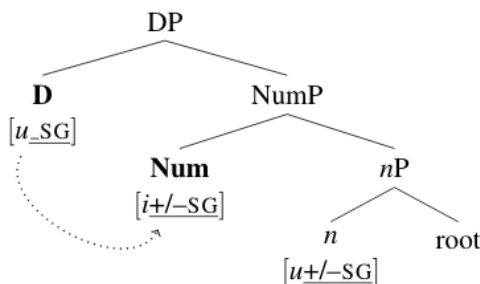
In standard Minimalism (e.g., Chomsky 2000, 2001, 2004), Agree is a syntactic operation where a head with uninterpretable features (the probe) scans its c-commanding domain for a head with interpretable features (the goal); the goal then values the probe’s uninterpretable features. In the analysis presented in this paper, where the difference that matters is that between valued and unvalued features, we can define Agree as an operation where a probe is a head with unvalued features, which scans its c-commanding domain for a head (the goal) with a valued instance of these features. The goal must value the probe’s features for the derivation not to crash, but once these features get a value, the search for a goal ends. This means that in case there are two heads in the probe’s c-commanding domain that could act as a goal, it is only the closest head to the probe that will value the probe’s features.

In the case of number agreement in Kipsigis, all agreeing elements are merged above NumP.<sup>39</sup> For example, we have already seen that the secondary suffix is in D. In any Kipsigis DP, Num and *n* enter the derivation with valued instances of [SG] (interpretable for Num, uninterpretable for *n*, but crucially both are valued). D, however, enters the derivation with an unvalued number feature, and, therefore, acts as a probe. It searches downwards for a goal, and the closest head in its c-commanding domain that has a valued instance of number features is Num. Num then values the number features of D via Agree. Since the features

<sup>39</sup> As will become clear in the remainder of this section, my theory predicts that anything merged below Num in Kipsigis should agree with the number features on *n* (i.e., with the morphological number class of the noun), and not with ‘semantic’ number on Num. Cross-linguistically, some adjectives are merged below Num. For some (e.g., Cinque 2005), all attributive adjectives whose source is not a reduced relative clause are merged below Num. However, there is independent evidence in Kipsigis that all adjectives in the language are reduced relative clauses that are merged above Num (Kouneli 2019a). As there are no modifiers merged below Num in Kipsigis, we cannot test the prediction that those modifiers would agree with the number features on little *n*. It is interesting to test this prediction in languages with a Kipsigis-type system and low adjectives, but further research is needed to identify such languages.

of D are now valued, the search for a goal is over, and it does not matter that *n* also has valued number features.<sup>40</sup>

(73) *Agreement with D*



Therefore, the standard approach to Agree can easily account for the agreement patterns that we observe: we see that agreeing elements in Kipsigis always agree with ‘semantic’ number, i.e., the features on Num, which follows from the fact that Num is always higher than *n*, and, thus, closer to any potential probe. However, the analysis outlined here, with valued number features on both Num and *n*, and standard Agree, makes the prediction that the number features of *n* can value a probe’s unvalued features in the absence of Num. This prediction is borne out: mass nouns, which lack a NumP projection as has already been discussed, trigger singular or plural agreement on other elements in the DP and the clause, depending on which number class they belong to. For example, even though ‘rain’ and ‘blood’ have the same interpretation with respect to number in (74) (they are number-neutral and have a mass interpretation), the former triggers singular agreement on the relativizer and the adjective because it belongs to the inherently singular class of nouns, with a *n*[+SG], while the latter triggers plural agreement on the relativizer and the adjective because it belongs to the inherently plural class, with a *n*[−SG].

(74) *Number agreement – mass nouns*

- a. púywêet nè      píriir  
dust.SG REL.SG red.SG  
‘red dust’
- b. kàràtiik chè      píriir-èn  
blood.PL REL.PL red-PL  
‘red blood’

The fact that the number features on little *n*, which are the features used for noun classification in the language, can participate in agreement under the right circumstances (i.e., the absence of Num, an intervening node with the same kind of features) brings us back to Kramer’s (2015: 70) definition of gender, repeated in (75) below. More specifically, number features in Kipsigis sort nouns into two or more classes (75i), and are sometimes reflected in agreement (75iii). However, because agreement is not always visible (due to the interaction with interpretable number features on Num), and because it is not clear how to modify (47ii), I opt for the more generic term ‘noun classification’ for the Kipsigis system.

<sup>40</sup> There are a variety of alternative analyses of nominal concord (see Norris 2017a, b for a comprehensive overview). However, all analyses use some sort of locality condition, which means that as long as all agreeing elements are merged above NumP, all theories will predict that they will agree with the features on Num, and not with those on *n*. As we will see in section 5, however, there are languages (Kiowa-Tanoan) in which D agrees with both the features on Num and those on *n*. Even though the exploration of this topic is beyond the scope of this paper, a successful theory of concord should be able to explain why the Kipsigis and Kiowa-Tanoan agreement patterns are both possible systems.

(75) Gender is:

- (i) the sorting of nouns into two or more classes;
- (ii) assigned depending on biological sex, animacy, and/or humanness, for at least some animate nouns;
- (iii) reflected by agreement patterns on other elements (e.g., adjectives, determiners, verbs, auxiliaries)

In the next section, I discuss the implications of the analysis presented here for number-based noun classification more generally.

## 5. Towards a typology of number-based noun classification

One of the main claims of this paper is that the tripartite system of number marking found in Kipsigis reflects a noun classification system based on inherent number features. My analysis, thus, corroborates the existence of number-based noun classification, which had only been reported for Kiowa-Tanoan languages before (Harbour 2007, 2011). The goal of this section is to briefly discuss other systems that might be analyzed as number-based noun classification, and to give some thoughts on how they can be integrated in discussions of gender and DP syntax more generally. The discussion is short, and no concrete solution is provided; its goal is, rather, to suggest avenues for further research.

The tripartite system of number marking described here for Kipsigis is very common in Nilo-Saharan languages (cf. Dimmendaal 2000, 2014), and it is also attested in some languages from other families spoken in the same geographical area as Nilo-Saharan: Afroasiatic languages of East Africa, such as Sidaama (Kramer & Anbessa 2019), the Kordofanian language Katcha (whose genetic affiliation is a matter of debate; Turner 2016, 2018) spoken in Sudan, as well as the language isolate Laal spoken in Chad (Lionnet 2016, 2017). Most of these languages are severely understudied (and virtually unstudied in the theoretical literature), which is why further research is needed before understanding the intricacies of their nominal system. It is clear, however, that the tripartite nature of the system is robust in all of these languages. The existence of three number classes (with what seem to be similar characteristics) means that the basic part of my analysis (i.e., nouns sorted into classes based on binary number features on *n*) can be extended to these languages. However, we also find differences from Kipsigis in some of these languages (especially the ones outside of the Nilo-Saharan family), which indicates that the details of the analysis will necessarily differ from language to language.

Some of the known differences are minor and are, in fact, predicted by the analysis presented here for Kipsigis. For example, as mentioned in various points of the paper, membership to the numberless class varies by size (e.g., it is small in Kipsigis but large in Lopit), and some languages (e.g., Shatt) might have mass nouns that belong to this class, unlike Kipsigis. Moreover, different languages express plural and singular differently (e.g., prefix vs. suffix) or determine plural affix allomorphy differently (e.g., thematic affixes are absent from most of these languages, and other factors – such as mora counting – play a role in determining the correct plural allomorph). Such differences are morphological in nature and independent of the presence of number features on little *n*. Other differences are more significant: the singulative can be used as a diminutive in some Nilo-Saharan languages, while some languages (e.g., Maba) exhibit general number forms in addition to the tripartite system of number marking (cf. Dimmendaal 2014). Furthermore, the tripartite system of number marking is also attested with adjectives in some languages (e.g., Turkana; Dimmendaal 1983), while Nuer (Western Nilotic; South Sudan) is, to my knowledge, the only language that exhibits the same pattern with verbs (Reid 2019). It is an open question how these differences are to be accounted for.<sup>41</sup>

---

<sup>41</sup> Thanks to an anonymous reviewer for bringing these facts about Nilo-Saharan languages to my attention. The same reviewer points out that the singulative can have an abundance reading in Lopit, reported in Moodie (2016). Moodie (2019), however, concludes that the abundance reading is not restricted to marked singulars, but also emerges with unmarked singulars as long as the noun in question triggers feminine gender agreement.

Another significant difference from Kipsigis, which is important for the implications of these systems for noun classification/gender more generally, is the fact that the tripartite system of number marking co-occurs with sex-based gender in some languages. Examples of such languages are languages of the Eastern Nilotic branch of Nilo-Saharan (e.g., Turkana; Dimmendaal 1983), Afroasiatic languages (e.g., Sidaama; Kramer & Anbessa 2019), and Katcha (Turner 2016, 2018). This clearly indicates that sex-based and number-based noun classification can co-occur in the same language. Under a little *n* analysis of gender (Kramer 2015), this means that sex-based gender features (e.g., [+/-FEM]) and number features can both be present on little *n* in a given language. Going back to the discussion in Sections 2 and 4, the picture that emerges is one where different types of noun classification exist. Certain aspects of noun classification are independent of the semantic criterion used for classification: both sex-based and number-based noun classification play an important role in agreement, and in both cases nominal lexical items are exhaustively sorted into classes. In theoretical terms, both sex-based and number-based features are present on the same syntactic head (little *n* in my analysis). The semantic feature involved though can have an effect on other aspects of the grammar. For example, number-based classification interacts with interpretable features on the functional head Num in a way that obscures agreement with *n* in Kipsigis (but not in Kiowa, as we will see shortly), while sex-based gender gives rise to similar effects when grammatical gender interacts with semantic (also called ‘natural’) gender.

The possibility of different types of classification features (e.g., sex and number features) co-occurring in the same language shows that multiple class systems of the Bantu type might be amenable to a feature decomposition analysis (as opposed to the commonly used arbitrary Class 1/2/3 etc. features), where at least number and humanness/animacy features are relevant. It is well-known that individuation plays a role in the Bantu noun class system (Mufwene 1980; Maho 1999; Contini-Morava 2000 among others), and Harbour (2007) shows that a number-based noun classification system is more similar to the Bantu system than it is to the sex-based gender system (his conclusions being confirmed here by those languages where sex-based and number-based classification coexist). Accounting for the Bantu system is beyond the scope of this paper, but it is worth noting that Katcha, whose genetic affiliation (Nilo-Saharan or Niger-Congo) is debated, exhibits mixed properties between a Kipsigis-type and a Bantu-type system, and might, therefore, provide clues as to what the connection is. More specifically, the language has a tripartite system of number-marking of the Kipsigis type, as well as masculine and feminine gender. The language also has a third gender, which has been called a ‘third’, ‘neuter’, and even a ‘plural’ gender in the literature (cf. Turner 2016). Katcha is interesting in displaying gender agreement, but not number agreement (Turner 2016, 2018). However, the generalization (with minor exceptions) is that nouns in their morphologically unmarked form (singular or plural) trigger masculine, feminine, or neuter agreement, while they trigger neuter agreement in their morphologically marked (singular or plural) form. Thus, what we see is that the sex-based gender features interact with the number features in a way that a uniform gender-number agreement system (reminiscent of Bantu) emerges. Further research into systems with complex (sex-based) gender and number interactions (see Di Garbo 2014 for a typological study of such interactions, as well as Mous 2008 for complex interactions in Afroasiatic languages of East Africa) is needed to determine the implications of such systems not only for the analysis of Bantu, but also for the syntax of noun classification more generally.

The behavior of Katcha briefly described here is different from the behavior of most Nilo-Saharan languages where sex-based gender co-occurs with number-based classification. For example, in Turkana, number and gender agreement are independent of each other, with nouns following the tripartite system of number marking and elements in the clause agreeing with both number (singular vs. plural) and gender (masculine vs. feminine vs. neuter). It is, therefore, possible that these languages are the same at the little *n* level (i.e., the same type of features are present on little *n*), but differ in the way agreement works. This, I think, is also the main difference between the Kipsigis system and that of Kiowa (the only other language with number-based classification for which a complete theoretical analysis is available), to which I turn next.

Kiowa has a three-way number distinction: singular, dual, and plural. There is no number morphology on the noun (with the exception of the inverse marker, to be discussed shortly), and the number

(76) X!óú            ø/e/gya-dǝǝ                                  Harbour (2011: 563)  
stone         3SG/3D/3PL-be  
‘It’s a stone/two stones/some stones.’

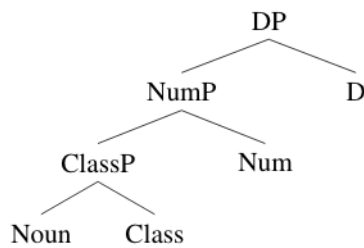
Furthermore, nouns of some number classes bear a suffix called the ‘inverse’ in one or more of the three number values, and trigger inverse number agreement on the verb when they have this suffix. For example, in (77), the noun *śópíit* ‘fish’, which belongs to the *SDI* class (where *I* stands for ‘inverse’), triggers the same kind of agreement seen in (76) above when in the singular and the dual, but when in the plural, it has the inverse suffix *-dś* and the verb has an *e-* agreement prefix, which is different from the *-gya* plural agreement prefix seen in (76).

Apart from this complicated system of number agreement, Harbour (2007, 2011) makes the observation that each one of the nine Kiowa classes includes nouns that share certain semantic characteristics related to number and individuation. For example, the *SSS* class contains non-granular mass nouns. As a result, he argues that nouns in the language are divided into classes based on number features hosted on the projection Class (which corresponds to the nominalizing head in my analysis). Since Kiowa has a more complicated number system than Kipsigis (including the distinction for dual number), Harbour uses the features [+/-SG], [+/-augmented], and [+/-group] to define the different noun classes, as well as the number distinctions on Num.<sup>42</sup> In this paper, I put those features aside, and give examples that involve only [+/-SG], which are the number features available in Kipsigis, where we only see a singular vs. plural distinction and three number classes.

<sup>42</sup> Harbour (2011, 2014) uses the feature [+/-atomic] (roughly) instead of [+/-SG]. The details are not relevant for the discussion here.



(78) *Kiowa DP*



Assuming that both ClassP and Num have number features in this structure, the complicated pattern of number agreement in Kiowa can be accounted for. More specifically, D has unvalued number features, and both ClassP and Num take part in agreement. Putting aside the technical details of Harbour's (2007, 2011) agreement system for now, when the features on ClassP and Num are the opposite of each other, D has a conflicting number specification, which is spelled out as the inverse suffix, and triggers inverse agreement.

Let's look at how the system works for the singular and plural of the noun *áá* 'stick', which is an *IDP* noun. The number specification for the *IDP* class is [–SG] on Class. In the singular, Num will be specified as [+SG], as shown in (79a), while in the plural, Num will be [–SG], as shown in (79b).

(79) *Agreement in the Kiowa DP*



Harbour (2007, 2011) assumes that no feature specification for a head with uninterpretable features that need to be checked means overspecification for both values of a feature, with the value that is not checked in the syntax being deleted at LF. Therefore, D is overspecified for both [+SG] and [–SG]. In (79a), D enters an agreement relation with both ClassP and NumP because they are equidistant. Then, D checks both [+SG] and [–SG]; this specification is spelled out as the inverse suffix, and triggers inverse agreement on the verb. In (79b), on the other hand, both ClassP and NumP have [–SG] features. As a result, D only checks its [–SG] features, and [+SG] is deleted. The [–SG] specification on D triggers plural agreement on the verb.

Harbour's (2007, 2011) analysis of number in Kiowa-Tanoan, in conjunction with the analysis of Kipsigis provided in this paper, shows that there is a variety of possible systems of number agreement and morphological exponence for languages with number-based noun classification. More specifically, even though Kiowa and Kipsigis are similar in using number features to sort nouns into classes, they also have important differences. In Kiowa, agreeing elements agree with number features on two heads (both Class and Num), while in Kipsigis they can only agree with number features on one head (the closest one in the structure, which is usually Num). The possibility for conflicting number specifications of a head in Kiowa is confirmed by the existence of the inverse marker, which is absent in Kipsigis.<sup>43</sup> Furthermore, Kiowa

<sup>43</sup> Even though the discussion has focused on Kiowa, number-based noun classification is a feature of other Kiowa-Tanoan languages as well, and there is significant variation within the family (see Harbour 2017 for an overview). For example, Southern Tiwa has a simpler system than Kiowa, without the dual and with fewer classes (at least for

lacks number morphology on the noun itself, while Kipsigis has rich number morphology on the noun. Therefore, noun classes in Kiowa are mainly reflected on agreement with other elements, while noun classes in Kipsigis are mainly reflected on number morphology. These two types of systems show that there is a range of possible systems of syntactic and morphological expression of number-based classification, and Kiowa and Kipsigis can serve as a basis for the analysis of other languages that seem to use number in the classification of their nouns.

## 6. Conclusion

In this paper, I have provided an analysis of the morphological expression of nominal number in Kipsigis. I have argued that the best way to account for the tripartite system of number marking in the language is by postulating number features on the nominalizing head *n* that sort nouns into three classes. The number features on *n* can capture the semantic coherence of the Kipsigis number classes, and can be used to derive the pattern of morphological marking of number. Furthermore, I have shown that the noun classes of Kipsigis have properties of noun classification systems cross-linguistically, which supports the idea that number features can be used in noun classification in some languages. Kiowa is one of the languages that has been claimed to use number features in noun classification, and I have provided a brief comparison of its system to Kipsigis, which can serve as a basis for the analysis of other languages with similar classification systems. Moreover, I have suggested that the use of number features on *n*, together with other gender features such as animacy or humanness features, might be able to account for gender systems with multiple noun classes, such as the ones found in Bantu. Finally, the proposal outlined in this paper adds to a body of research that shows that number features can be present on different parts of the nominal extended projection (Lecarme 2002; Acquaviva 2008; Wiltchko 2008; Alexiadou 2011; Kramer 2016).

## References

- Acquaviva, Paolo. 2008. *Lexical Plurals*. Oxford: Oxford University Press.
- Acquaviva, Paolo. 2009. Roots and lexicality in Distributed Morphology. *York Essex Morphology Meeting* 2. 1–21.
- Acquaviva, Paolo. 2015. Singulatives. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen, Franz Rainer (eds.), *Word-formation: an international handbook of the languages of Europe*. Volume 2, 1171–1182. Berlin: de Gruyter.
- Alexiadou, Artemis. 2004. Inflection class, gender and DP-internal structure. In Gereon Müller, Lutz Gunkel, and Gisela Zifonun, eds, *Explorations in Nominal Inflection*. Berlin: Mouton. 21–50.
- Alexiadou, Artemis. 2011. Plural mass nouns and the morphosyntax of number. In Mary Byram Washburn et al., (eds), 33–41. *Proceedings of the West Coast Conference on Formal Linguistics* 28. Somerville, MA: Cascadia Press.
- Alexiadou, Artemis. 2019. Morphological and semantic markedness revisited: The realization of plurality across languages. *Zeitschrift für Sprachwissenschaft* 2019 38.1. 123–154.
- Alexiadou, Artemis, and Gereon Müller. 2008. Class features as probes. In Bachrach, Asaf and Andrew Nevins, (eds.), *Inflectional Identity*, 101–155. Oxford: Oxford University Press.

---

inanimate nouns), bringing it closer to Kipsigis. Heck & Richards (2010), in the appendix of their paper, provide an analysis for that pattern (see also Noyer 1992), but it is not clear how to compare it to Kipsigis due to the complex agreement system in the clausal domain (which is absent in Kipsigis). According to Heck & Richards' (2010) analysis, however, Southern Tiwa lacks an inverse marker of the Kiowa type in the nominal domain. Since both Kipsigis and Southern Tiwa lack a dual (at least for inanimate nouns in the Tiwa case), this raises the question of whether the availability of an inverse in the DP is in any way related to the dual. I leave this as a topic for further research. Thanks to Daniel Harbour for bringing the Southern Tiwa data to my attention.

- Anttila, Arto, and Adams Bodomo. 2009. Prosodic morphology in Dagaare. In McLaughlin, Fiona, Eric Potsdam, and Masangu Matondo (eds.), *Selected Proceedings of the 38th Annual Conference on African Linguistics (ACAL 38)*. Cascadia Proceedings Project, Somerville, Massachusetts.
- Arad, Maya. 2003. Locality constraints on the interpretations of roots. *Natural Language and Linguistic Theory* 21. 737–778.
- Arad, Maya. 2005. *Roots and Patterns: Hebrew Morphosyntax*. Dordrecht: Springer.
- Aronoff, Mark. 1994. *Morphology by Itself: Stems and Inflectional Classes*. Cambridge, MA: MIT Press.
- Arregi, Karlos, and Andrew Nevins. 2007. Obliteration vs. impoverishment in the Basque g-/z constraint. *University of Pennsylvania Working Papers in Linguistics* 13.1. 2.
- Arregi, Karlos, and Andrew Nevins. 2012. *Morphotactics: Basque Auxiliaries and the structure of spellout*. Dordrecht: Springer.
- Baković, Eric. 2000. *Harmony, Dominance, and Control*. Ph.D. Dissertation, Rutgers University.
- Bonet, Eulalia. 1991. *Morphology after syntax: Pronominal clitics in Romance*. Ph.D. Dissertation MIT.
- Bossi, Madeline, and Michael Diercks. 2019. V1 in Kipsigis: Head movement and discourse-based scrambling. *Glossa: A Journal of General Linguistics*, 4(1), 65.  
DOI: <http://doi.org/10.5334/gigl.246>
- Calabrese, Andrea. 2011. Investigations on markedness, syncretism and zero exponence in morphology. *Morphology* 21.2. 283–325.
- Carstens, Vicki. 1991. *The morphology and syntax of determiner phrases in Kiswahili*. Ph.D Dissertation, UCLA.
- Carstens, Vicki. 2010. Implications of grammatical gender for the theory of uninterpretable features. In Michael Putnam, (ed.), *Exploring Crash-Proof Grammars*, 31–57. Amsterdam: Benjamins.
- Chierchia, Gennaro. 1998. Reference to kinds across language. *Natural Language Semantics* 6.4. 339–405.
- Chomsky, Noam. 2000. Minimalist inquiries, the framework. In Roger Martin, David Michaels and Juan Uriagereka, (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz, (ed.), *Ken Hale: A Life in Language*, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2004. Beyond explanatory adequacy. In Adriana Belletti, (ed.), 104–131. *Structures and Beyond*, 104–131. New York: Oxford University Press.
- Cinque, Guglielmo. 2005. Deriving Greenberg's Universal 20 and its exceptions. *Linguistic inquiry* 36.3. 315–332.
- Cinque, Guglielmo. 2010. *The syntax of adjectives: A comparative study*. Vol. 57. Cambridge, MA: MIT Press.
- Contini-Morava, Ellen. 2000. Noun class as number in Swahili. In Contini-Morava, Ellen, and Y. Tobi, (eds), *Between Grammar and Lexicon*, 3–30. Amsterdam: John Benjamins.
- Corbett, Greville. 1991. *Gender*. Cambridge: Cambridge University Press.
- Corbett, Greville. 2000. *Number*. Cambridge: Cambridge University Press.
- Creider, Chet A. 1982. *Studies in Kalenjin nominal tonology*. Vol. 3. Reimer.
- Creider, Chet A. 1989. *The syntax of the Nilotic languages: themes and variations*. Vol. 9. D. Reimer.
- Creider, Chet A., and Jane Tapsuei Creider. 1983. Topic: comment relations in a verb-initial language. *Journal of African Languages and Linguistics* 5.1.1–15.
- Creider, Chet A., and Jane Tapsuei Creider. 1989. *A grammar of Nandi*. H. Buske.
- Despić, Miloje. 2017. Investigations on mixed agreement: polite plurals, hybrid nouns and coordinate structures. *Morphology* 27.3: 253–310.
- Diercks, Michael, and Meghana Rao. 2019. Upward-oriented complementizer agreement with subjects and objects in Kipsigis. In Clem, Emily, Peter Jenks, and Hannah Sande, (eds.), *Theory and description in African Linguistics: Selected papers from the 47th Annual Conference on African Linguistics*, 369–393. Berlin: Language Science Press.

- Di Garbo, Francesca. 2014. *Gender and its interaction with number and evaluative morphology*. Ph.D. Dissertation, Stockholm University.
- Dimmendaal, Gerrit J. 1983. *The Turkana language*. Dordrecht: Foris.
- Dimmendaal, Gerrit J. 2000. Number marking and noun categorization in Nilo-Saharan languages. *Anthropological Linguistics* 42.2. 214-261.
- Dimmendaal, Gerrit J. 2001. Language shift and morphological convergence in the Nilotic area. *Sprache und Geschichte in Afrika* 16/17: 83-124.
- Dimmendaal, Gerrit J. 2012. Metrical structures: A neglected property of Nilotic (and other African language families). *Studies in Nilotic Linguistics* 5: 1-26.
- Dimmendaal, Gerrit J. 2014. Pluractionality and the distribution of number marking across categories. In Storch, Anne and Gerrit J. Dimmendaal (eds.), *Number: Constructions and Semantics. Case studies from Africa, Amazonia, India and Oceania*, 57-75. Amsterdam and Philadelphia: John Benjamins.
- Eberhard, David. M., Gary F. Simons, and Charles D. Fennig, (eds.) . 2020. *Ethnologue: Languages of the World. Twenty-third edition*. Dallas, Texas: SIL International.
- Embick, David. 2010. *Localism versus globalism in morphology and phonology*. Vol. 60. Cambridge, MA: MIT Press.
- Embick, David, and Morris Halle. 2005. On the status of stems in morphological theory. In Twan Geerts, Ivo von Ginneken, and Haike Jacobs (eds.), *Romance Languages and Linguistic Theory 2003*, 37-62. Amsterdam: Benjamins.
- Embick, David, and Rolf Noyer. 2007. Distributed morphology and the syntax/morphology interface. In Gillian Ramchand and Charles Reiss, (eds.), *The Oxford Handbook of Linguistic Interfaces*. 289-324. Oxford: Oxford University Press.
- Ferrari, Franca. 2005. A syntactic analysis of the nominal systems of Italian and Luganda: how nouns can be formed in the syntax. Ph.D. Dissertation, New York University.
- Greenberg, Joseph., H. 1978. How does a language acquire gender markers. *Universals of human language*, 3. 47-82.
- Grimm, Scott. 2012. *Number and individuation*. Ph.D. Dissertation. Stanford University.
- Grimm, Scott. 2018. Grammatical number and the scale of individuation. *Language* 94.3. 527- 574.
- Halle, Morris.1997. Distributed morphology: impoverishment and fission. In Benjamin Bruening et al. eds. *MIT Working Papers in Linguistics 30: Papers at the Interface*. Cambridge, Mass.: MITWPL. 425–449.
- Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Ken Hale and Samuel Jay Keyser (eds), *The View from Building 20*, 111-176. Cambridge, Mass.: MIT Press.
- Hammerly, Christopher. 2018. Limiting gender. In Mathieu, Éric, Myriam Dali, and Gita Zareikar (eds.), *Gender and Noun Classification*. Oxford: Oxford University Press.
- Harbour, Daniel. 2003. The Kiowa case for feature insertion. *Natural Language & Linguistic Theory* 21.3. 543-578.
- Harbour, Daniel. 2007. *Morphosemantic Number: From Kiowa Noun Classes to UG Number Features*. Vol. 69. Springer Science & Business Media.
- Harbour, Daniel. 2011. Valence and atomic number. *Linguistic Inquiry* 42. 561-594.
- Harbour, Daniel. 2014. Paucity, abundance, and the theory of number. *Language* 90.1. 185-229.
- Harbour, Daniel. 2017. Kiowa-Tanoan languages. *Oxford Research Encyclopedia of Linguistics*. DOI: 10.1093/acrefore/9780199384655.013.27
- Haspelmath, Martin, and Andres Karjus. 2017. Explaining asymmetries in number marking: Singulatives, pluratives, and usage frequency. *Linguistics* 55.6. 1213-1235.
- Harley, Heidi. 2014. On the identity of roots. *Theoretical Linguistics* 40. 225–276.
- Heck, Fabian, and Marc Richards. 2010. A probe-goal approach to agreement and non-incorporation restrictions in Southern Tiwa. *Natural Language & Linguistic Theory* 28.3. 681-721.

- Heycock, Caroline, and Roberto Zamparelli. 2005. Friends and Colleagues: Plurality, Coordination, and the Structure of DP. *Natural Language Semantics* 13. 201-270.
- Hollis, Alfred Claud. 1909. *Nandi: Their Language and Folklore*. Negro Universities Press.
- Kihm, Alain. 2005. Noun class, gender and the lexicon–syntax–morphology interfaces: a comparative study of Niger-Congo and Romance languages. In Guglielmo Cinque and Richard S. Kayne, eds, *The Oxford Handbook of Comparative Syntax*. Oxford: Oxford University Press. 459–512.
- Kouneli, Maria. 2017. The morphological expression of number in Kalenjin. In Andrew Lamont & Katerina Tetzloff (eds.), *Proceedings of the 47th Annual Meeting of the North East Linguistic Society (NELS)*, Vol.2: 167-176. Amherst: GLSA.
- Kouneli, Maria. 2018. Plural marking on mass nouns: Evidence from Greek. In Myriam Dali, Eric Mathieu & Gita Zareikar (eds.), *Gender and noun classification*. Oxford: Oxford University Press.
- Kouneli, Maria. 2019a. *The syntax of number and modification: An investigation of the Kipsigis DP*. Ph.D. Dissertation, New York University.
- Kouneli, Maria. 2019b. Determiner spreading and modification in Kipsigis. In *Proceedings of Chicago Linguistic Society (CLS)* 54.
- Kramer, Ruth. 2009. *Definite markers, phi-features, and agreement: a morphosyntactic investigation of the Amharic DP*. Ph.D. Dissertation. University of California, Santa Cruz.
- Kramer, Ruth. 2015. *The morphosyntax of gender*. Oxford: Oxford University Press.
- Kramer, Ruth. 2016. A split analysis of plurality: number in Amharic. *Linguistic Inquiry* 47.3. 527-559.
- Kramer, Ruth, and Anbessa Teferra. 2019. Gender switch in Sidaama. *Brill's Journal of Afroasiatic Languages and Linguistics*. DOI: <https://doi.org/10.1163/18776930-01102006>
- Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. In Bartsch, R. (ed.), *Semantics and contextual expressions*, 75-116. Dordrecht: Foris.
- Kučerová, Ivona. 2017.  $\phi$ -features at the syntax-semantics interface: Evidence from nominal inflection. *Linguistic Inquiry*. 1-77.
- Kučerová, Ivona, and Anna Moro. 2012. On the Structural and Typal Differences between Mass and Count Nouns: Evidence from Romance. In *Proceedings of Sinn und Bedeutung*. Vol. 16.
- Landau, Idan. 2016. DP-internal semantic agreement: A configurational analysis. *Natural Language & Linguistic Theory* 34.3: 975-1020.
- Lecarme, Jacqueline. 2002. Gender “polarity”: theoretical aspects of Somali nominal morphology. In Paul Boucher and Marc Plénat, eds. *Many Morphologies*. Somerville, Massachusetts: Cascadia Press. 109-141.
- Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. *Formal semantics: The essential readings*. 127-146.
- Lionnet, Florian. 2016. *Subphonemic teamwork: A typology and theory of cumulative coarticulatory effects in phonology*. Ph.D. Dissertation, UC Berkeley.
- Lionnet, Florian. 2017. A theory of subfeatural representations: the case of rounding harmony in Laal. *Phonology* 34.3. 523-564.
- Lodge, Ken. 1995. Kalenjin phonology and morphology: A further exemplification of underspecification and non-destructive phonology. *Lingua* 96.1. 29-43.
- Lohitare, L.D., D.L.B. Lohammarimoi, D. T. Peter, and P.L. Joseph. 2012. *Didinga Grammar Book*. Juba, South Sudan: SIL publications.
- Lowenstamm, Jean. 2008. On little n,  $\sqrt{\text{ }}$ , and types of nouns. In Jutta Hartmann, Veronika Hegedüs, and Henk van Riemsdijk, eds, *Sounds of Silence: Empty Elements in Syntax and Phonology*. Amsterdam: Elsevier. 105–144.
- Maho, Jouni. 1999. *A Comparative Study of Bantu Noun Classes*. Gothenburg: Acta Universitatis Gothoburgensis.
- Marantz, Alec. 1997. No escape from syntax. *University of Pennsylvania Working Papers in Linguistics* 4. 201–225.
- Marantz, Alec. 2001. Words. Ms, MIT.
- Mathieu, Éric. 2012. Flavors of division. *Linguistic Inquiry* 43.4. 650-679.

- Moodie, Jonathan. 2016. Number marking in Lopit, an Eastern Nilotic language. In Payne, Doris L., Sara Pacchiarotti and Mokaya Bosire (eds.), *Diversity in African languages: selected papers from the 46<sup>th</sup> Annual Conference on African Linguistics*, 397-416. Berlin: Language Science Press.
- Moodie, Jonathan. 2019. *A grammar of the Lopit language*. Ph.D. Dissertation, University of Melbourne.
- Mous, Maarten. 2008. Number as an exponent of gender in Cushitic. In Erin Shay and Zygmunt Frajzyngier, (eds.), *Interaction of Morphology and Syntax: Case Studies in Afroasiatic*. Amsterdam: Benjamins. 137-160.
- Mufwene, Salikoko S. 1980. Number, countability and markedness in Lingala LI/-MA-noun class. *Linguistics* 18.11-12, 1019-1052.
- Nevins, Andrew. 2010. *Locality in Vowel Harmony*. Cambridge, MA: MIT Press.
- Nevins, Andrew. 2011. Multiple agree with clitics: Person complementarity vs. omnivorous number. *Natural Language & Linguistic Theory* 29.4: 939-971.
- Nevins, Andrew. 2012. Haplological dissimilation at distinct stages of exponence. In Trommer, Jochen, ed. *The morphology and phonology of exponence*. Oxford: Oxford University Press. 84-116.
- Norris, Mark. 2017a. Description and analyses of nominal concord (Pt I). *Language and Linguistics Compass* 11.11: e12266.
- Norris, Mark. 2017b. Description and analyses of nominal concord (Pt II). *Language and Linguistics Compass* 11.11: e12267.
- Noyer, Robert Rolf. 1992. *Features, positions and affixes in autonomous morphological structure*. Ph.D. Dissertation, Massachusetts Institute of Technology.
- Nurmio, Silva. 2017. Collective Nouns in Welsh: A Noun Category or a Plural Allomorph? *Transactions of the Philological Society* 115.1: 58-78.
- Oltra-Massuet, Maria Isabel. 1999. *On the notion of theme vowel: a new approach to Catalan verbal morphology*. MA thesis, MIT.
- Oltra-Massuet, Isabel, and Karlos Arregi. 2005. Stress-by-structure in Spanish. *Linguistic Inquiry* 36. 43-84.
- Ouwayda, Sarah. 2014. *Where Number Lies: Plural marking, numerals, and the collective-distributive distinction*. Ph.D. Dissertation. University of Southern California.
- Pescarini, Diego. 2010. Elsewhere in Romance: Evidence from clitic clusters. *Linguistic Inquiry* 41.3. 427-444.
- Picallo, M. Carme. 1991. Nominals and nominalization in Catalan. *Probus* 3. 279-316.
- Quine, Willard Van Orman. 1960. *Word and object*. Cambridge, MA.
- Radkevich, Nina V. 2010. *On location: The structure of case and adpositions*. Ph.D. Dissertation, University of Connecticut.
- Reid, Tatiana. 2019. *The phonology and morphology of the Nuer verb*. Ph.D. Dissertation, University of Surrey.
- Ritter, Elizabeth. 1991. Two functional categories in noun phrases: evidence from Modern Hebrew. In Susan D. Rothstein, ed. *Perspectives on Phrase Structure: Heads and Licensing*. Syntax and Semantics 25. San Diego: Academic Press. 37-62.
- Ritter, Elizabeth. 1993. Where's gender? *Linguistic Inquiry* 24. 795-803.
- Roca, Iggy M. 1989. The organisation of grammatical gender. *Transactions of the Philological Society* 87. 1-32.
- Sauerland, Uli. 2003. A new semantics for number. In Young, R. and Zhou, Y. (eds.), *Proceedings of SALT 13*, 258-275. Ithaca: CLC Publications.
- Sauerland, Uli, J. Andersen, and K. Yatsushiro. 2005. The plural is semantically unmarked. In Kesper, S., and M. Reis (eds.), *Linguistic evidence: Empirical, Theoretical, and computational perspectives*, 413-434. Mouton de Gruyter.
- Steriopolo, Olga, and Martina Wiltschko. 2010. Distributed GENDER hypothesis. In G. Zybatow et al. (eds.), *Formal Studies in Slavic Linguistics: Proceedings of the Formal Description of Slavic Languages 7.5*. New York: Lang. 155-72.

- Stolz, Thomas. 2001. Singulative-collective: Natural morphology and stable classes in Welsh number inflexion on nouns. *STUF-Language Typology and Universals* 54.1. 52-76.
- Tiersma, Peter Meijes. 1982. Local and general markedness. *Language*: 832-849.
- Toweett, Taaita. 1975. *Kalenjin nouns and their classification*. MA thesis, University of Nairobi.
- Toweett, Taaita. 1979. *A study of Kalenjin linguistics*. Nairobi: Kenya Literature Bureau.
- Tucker, Archibald Norman, and Margaret A. Bryan. 1965. *Noun classification in Kalenjin: Nandi-Kipsigis*. School of Oriental and African Studies, University of London.
- Turner, Darryl. 2016. *Morphosyntax of Katcha nominals: a Dynamic Syntax account*. Ph.D. Dissertation, University of Edinburgh.
- Turner, Darryl. 2018. The interaction of number and gender in Katcha. *Studies in African Linguistics* 47.1&2. 129-153.
- Watanabe, Akira. 2015. Valuation as deletion: inverse in Jemez and Kiowa. *Natural Language & Linguistic Theory* 33.4: 1387-1420.
- Zwarts, Joost. 2001. Number in Endo-Marakwet. *Advances in Nilo-Saharan Linguistics: Proceedings of the 8th Nilo-Saharan Linguistics Colloquium, University of Hamburg*. 281-294.

**Acknowledgments:** I'm grateful to Sharon Chelangat, Nehemiah Cheruiyot, Bonface Kemboi, Donald Kibet, Emmanuel Kiprono, Wesley Kirui, Kipyegon Kitur, Amos Ngeno, Robert Ngeno, Mercy Chep, Victor Mutai, and Nathan Rotich for their work as language consultants. I'd also like to thank Stephanie Harves, Richard Kayne, Ruth Kramer, Alec Marantz, and Gary Thoms for guidance and feedback on the project, as well as the audiences at NYU, University of Leipzig, NELS 2016, and BLS 2019, where parts of this paper were presented. All errors are my own.

#### **Author contact information:**

Maria Kouneli  
 Institut für Linguistik, Philologische Fakultät  
 Beethovenstraße 15  
 Leipzig 04107  
 Germany  
 maria.kouneli@uni-leipzig.de

#### **Appendix**

In this appendix, I provide additional data on nominal classes in Kipsigis based on my fieldnotes. There are five tables, organized along two dimensions: morphological number class and count vs. mass. Thus, there are five tables: Table 1 includes inherently singular count nouns, Table 2 inherently plural count nouns, Table 3 inherently singular mass nouns, Table 4 inherently plural mass nouns, and Table 5 numberless (count) nouns. The inherently singular count class contains more than half of the nouns in the language, and Table 1 contains just few examples (I capped it at 30 words). Table 2 contains about half of the inherently plural nouns in my notes. Tables 3 – 5 are exhaustive lists of those nouns in my notes. This does not mean that they represent an exhaustive list of nouns in the language, but I hope that my sample is representative. In Tables 1 and 2 (which are not exhaustive lists), I tried to include examples of nouns from various lexical semantic categories and nouns that take different thematic and number suffixes. Tones are transcribed for the majority of nouns, with the exception of a few nouns that were obtained by native speakers through e-mail (tone is not part of the spelling system).

The tables give the gloss and singular and plural forms of a given noun; additionally, the thematic suffix of the noun in its unmarked form is given, along with the singulative/plural suffix that it has in its marked form. Here you will find examples of more singulative/plural suffixes than in the main text, as well as more allomorphs of the singulative *-yaan*. The tables also include information on whether length dissimilation of the final long vowel takes place. As mentioned in footnote 5 of the main text, final long

vowels are shortened when they follow another long vowel in certain environments. In disyllabic nouns, the rule is simple, but it is not clear what the conditioning factor is in polysyllabic nouns, with tonal constraints playing an important, but ill-understood, role (see Kouneli 2019a for Kipsigis and Dimmendaal 2012 for the related dialect Nandi). ‘Not Applicable’ (NA) is used when the final long vowel is preceded by a short vowel, ‘Yes’ when the final vowel is shortened when following a long vowel, and ‘No’ when the final vowel is not shortened when following a long vowel.

Nouns were included in the mass vs. count noun depending on their behavior with diagnostics for the mass/count distinction outlined in Section 3 (the diagnostics are not given here, but were conducted for all of the nouns in the tables below). Their marked (singular or plural) form is given for those nouns that can be coerced, and a column is added with the interpretation of the coerced form. It should be noted that abstract nouns (e.g., *kárâarindá* ‘beauty’) pattern with mass nouns, and the vast majority belong to the inherently singular class. These are not included in the tables below, with the exception of a couple of words for illustration purposes.

Finally, a column indicates whether a given noun is a loanword, and if so, what the word is in the language of origin.

**Table 1: Inherently singular nouns (count)**

Gloss	SG	PL	Th	PL suffix	LD	Loan
machete	pâangéet	pâangóok	-à	-oy	No/No	Swahili <i>panga</i>
crocodile	tíngóongéet	tíngóongóok	-à	-oy	No/No	
finger	síiyéet	síiyóok	-à	-oy	No/No	
pumpkin	móongéet	móongóosyék/moongóok	-à	-oosyek/-oy	No/Yes/No	
spider	kíprórógéet	kípróróróosyék	-á	-oosyek	NA/Yes	
buffalo	sáaéet	sáaéenfk	-á	-eenik	No/Yes	
leg	kèrèengét	kèrèengáaik	-a	-ay	Yes/No	
knife	róotwéet	róotóok	-wà	-oy	No/No	
twins	símaatwéet	símaatóonfk	-wà	-oonik	No/Yes	
friend	chóorwéet	chóoróonóok	-wà	-oonook	No/No	
leather rope	àanwéet	áanóok	-wà	-oy	No/No	
axe	áywéet	áonóok	-wà	-oonook	NA/No	
mountain	tùlwéet	tùlonóok	-wà	-oonook	NA/No	
village	kóokwéet	kóokwàatinwéek	-wà	-atinweek	No/NA	
knee	kùutùndá	kùutùngáaik	zero	-ay	NA/NA	
rib	kàràastá	kàràswéek/kàràasáaik	zero	-wà/-ay	NA/NA/No	
father	kwāandá	kwāanlisyék	zero	-iisiek	NA/Yes	
throat	móoktá	móokwék	zero	-wà	NA/Yes	
corpse	múustá	múuswék	zero	-wà	NA/Yes	
muscle	wàléeldá	wàléelwéek	zero	wà	NA/No	
monkey	chèeréréet	chèeréréenfk	-è	-eenik	No/Yes	English <i>car</i>
car	kàrít	kàrísyék	-ì	-iisiek	NA/Yes	
scale	ràatlíit	ràatlíisiek	-ì	-iisiek	NA/Yes	
brain	kùundít	kùundjisiek	-í	-iisiek	Yes/Yes	
torch	tóochít	tóochisiek	-í	-iisiek	Yes/Yes	
arm	èúut	éúunék	-ù	-uunek	NA/Yes	
day	pèetúut	pèetúusyék	-ù	-uusiek	No/Yes	
nose	sèrúut	sérúunék	-ù	-uunek	NA/Yes	
donkey	sígiryéet	sígiróok	-yà	-oy	NA/NA	
monkey	tísyéet	tísóok	-yá	-oy	NA/NA	



**Table 2: Inherently plural nouns (count)**

Gloss	SG	PL	Th	SG suffix	LD	Loan
enemy	pũnyáat	pũnfik	-i	-yaan	No/No	
lung	pwáanyáat	pwáanfik	-i	-yaan	No/No	
star	kěechěgyáat	kěechěgik	-f	-yaan	No/Yes	
fly	káalyàangyáat	káalyáangik	-i	-yaan	No/Yes	
pepper (veg.)	pilibilyaat	pilibiliik	-i	-yaan	NA/NA	Sw. <i>pilipili</i>
bean	ngéendyáat	ngéendéek	-à	-yaan	No/No	
elephant	pèelyáat	pèeléek	-à	-yaan	No/No	
hair	sũmyáat	sũmék	-à	-yaan	No/No	
twins (animals)	sárámyáat	sáráméek	-à	-yaan	NA/NA	
fruit	lógóyáat	lógóék	-à	-yaan	NA/NA	
type of fruit	núgyáat	núgúk	-ù	-yaan	NA/NA	
flea	kĩmityáat	kĩmíték	-à	-yaan	NA/NA	
parent	siglindét	sigfik	-i	-iin	Yes/NA	
cockroach	sólóbcháat	sólóbéek	-á	-yaan	NA/NA	
flower	màuwáat	màuwék	-à	-yaan	NA/NA	Sw. <i>ua</i> (pl: <i>maua</i> )
wild olive tree	èemilityáat	èemilitik	-i	-yaan	No/Yes	
guest/visitor	tãandét	tãéek	-à	-ta	Yes/NA	
mushroom	pòobáat	pòobéek	-à	-(?y)aan	No/No	
tooth	kéeldéet	kéelék	-à	-ta	No/No	
cloud	pòoldét	pòolík	-f	-ta	Yes/Yes	
bead	sonoyaat	sonoyéek	-a	-yaan	NA/NA	
white person	chũmiindét	chũmbéek	-à	-iin	Yes/No	
type of worm	sàbityáat	sàbitik	-i	-yaan	No/Yes	
red ant	piirechyaat	piirechiik	-i	-yaan	NA/NA	
louse	siryáat	sirék	-à	-yaan	NA/NA	

**Table 3: Inherently singular nouns (mass)**

Gloss	SG	PL	Th	PL suffix	LD	PL meaning	Loan
dust	púywéet	-	-wá	-	NA	NA	
dirt	sĩmdá	sĩmwéek	zero	-wá	NA/NA	lots of dirt	
rain	ròoptá	ròobwék	zero	-wá	NA/Yes	long rainy seasons	
type of salt	ngèendá	-	zero	-	NA	NA	
oil	mwàaytá	mwàanfik	zero	irreg.	NA/No	packs/types	
mist	kibungweereet	-	-a	-	No	NA	
ugali (food)	kĩmnyéet	kĩmáfik	-yà	-ay	NA/NA	servings	
(animal) fat	sũundéet	sũundéetfik	-à	-eetik	No/Yes	lots of fat	
meat	pèendá	pànyéek	zero	irreg.	NA/NA	no difference	
fire	máat	mástinwéek	?	irreg.	NA/NA	sources of fire	
wind	kóorĩistá	-	zero	-	NA	NA	
type of poison	ngwáanéet	-	-á	-	No	NA	
sand	ngàaynéet	-	-à	-	No	NA	
anthrax	puraastá	-	zero	-	NA	NA	
time	kasaartá	-	zero	-	NA	NA	

**Table 4: Inherently plural nouns (mass)**

Gloss	SG	PL	Th	SG suffix	LD	SG meaning	Loan
honey	kũmyáat	kũmfik	-i	-yaan	No/No	package	
blood	kàràtyáat	kàràtfik	-i	-yaan	NA/NA	drop of blood	
soda ash	-	mũnyéek	-á	-	NA	NA	
tea	chãaiyáat	chãaifk	-i	-yaan	No/No	tea serving/plant	Sw. <i>chai</i>
soup	-	sũutéek	-à	-	No	NA	
water	-	péek	-à	-	NA	NA	
millet	-	péek	-á	-	NA	NA	
milk	-	chèegá	zero	-	NA	NA	
sugar	sũgàryáat	sũgàrúk	-ù	-yaan	NA	package/dose	Eng. <i>sugar</i>
rice	mũchèelyáat	mũchèelék	-à	-yaan	No/No	grain/package	Sw. <i>mchele</i>
salt	chũumbyáat	chũumbfik	-i	-yaan	No/No	package	
ash	òryáat	òréek	-à	-yaan	NA/NA	no difference	
sand	-	ngũgũnyéek	-á	-	NA	NA	
alcohol/beer	-	máywéek	-(w)á	-	NA	NA	
flour	-	pũsyéek	-á	-	NA	NA	
maize	pàandyáat	pàandék	-à	-yaan	No/No	cob/package	
coal	-	nèséek	-à	-	NA	NA	
mucus	-	suberiik	-i	-	NA	NA	

**Table 5: Numberless nouns (count)**

Gloss	SG	PL	SG suffix	PL suffix	LD	Loan
fish	ìrchìrlyáat	ìrchìrènfk	-yaan	-eenik	NA/Yes	
shoulder	tìgfkcháat	tìgfgáfiik	-yaan	-ay	NA/NA	
hunter	kiplágáatyáat	kiplágáatiinfk	-yaan	-iinik	No/Yes	
piece of skin	írfryáat	írfèenf	-yaan	-eenik	NA/Yes	
guinea fowl	térkékyáat	térkéèenf	-yaan	-eenik	NA/Yes	
sock	sìgìsyáat	sìgìsiinfk	-yaan	-iinik	NA/Yes	Eng. <i>sock</i>
shoe	kwèeyáat	kwèeyòosyék	-yaan	-oosya	No/Yes	
ankle	kwàaryáat	kwàariisyék	-yaan	-iisya	No/Yes	
banana	índfìsyáat	índfìsiinfk	-yaan	-iinik	No/Yes	Sw. <i>ndizi</i>
type of tree	kùryáat	kùrèenf	-yaan	-eenik	NA/Yes	