Individuation and Inverse Number Marking in Dagaare

Scott Grimm

Stanford University

sgrimm@stanford.edu

5.1 Introduction

Dagaare (Gur; Niger-Congo) exhibits an initially surprising system for marking number.[1] The basic paradigm is given with respect to the Dagaare words 'child' and 'seed' in (1), showing a near[2] minimal pair where both nouns share the same stem, yet the morpheme *-ri*[3] marks the plural interpretation for 'child' and the singular interpretation for 'seed'.

(1)	Gloss	Singular	Plural	Stem
	'child'	bíé	bíírí	bì-
	'seed'	bírì	bíè	bí-

Upon first view, this pattern would appear arbitrary and unstable in comparison with number marking systems from Indo-European languages—for how would one know if a particular noun is to be marked in the singular or the plural except on a noun-by-noun basis? Inverse number systems are rare but attested at least in North America (Kiowa) and the Pacific (Nehan).[4] To give an example in a mock version of English, the pattern would appear as in (2), where *-s* is an inverse number marker in English:

(2)	Gloss	Singular	Plural
	'child'	child	child-s
	'seed'	seed-s	seed

This paper, based on detailed investigation of the lexicon of Dagaare, argues that the core of Dagaare's number marking system is based on the subtle but pervasive semantic principle of individuation, the propensity for an entity to appear as an individual unit. While individuation has most often been invoked in a binary fashion to account for the different behaviors of count and mass terms, I will show that it is also relevant when restricted to the count domain, where, for the case at hand, Dagaare's morphology appears to be sensitive to different degrees of individuation. These results are relevant for how individuation is viewed in general. Some researchers have made a sharp distinction between individuated and non-individuated (Mufwene 1984; Bloom 1999), although others have viewed individuation as comprised of different sub-properties (Wierzbicka 1988) or as a continuum (Gentner and Borodisky 2001). The view emerging from Dagaare's morphology is compatible with the broader view of individuation where it does not simply distinguish in a binary fashion between count and mass nouns, but operates in a scalar fashion. Under this view, mass terms (*water*) are less individuated than entities typically occurring in collections (*ribs*), which are in turn less individuated than typically singular entities (*dog*). As such, individuation has potential effects throughout a given grammatical number system. This influence extends to the categorization of terms as collectives, common cross-linguistically for e.g. insects and vegetables/grains/fruits, as well as influencing the preference for occurring in the plural or singular, the distinction which will be shown to be crucial for the paradigm in (1). A broader understanding of individuation then provides a common semantic backdrop against which distinct patterns of grammatical number marking can be related, and of which the categories mass and count are simply the most visible endpoints of the scale.

3

A second theme throughout the paper will be examining the Dagaare system from the perspective of markedness, as the observed morphological patterns are unexpected upon the standard view. Usually, the singular is considered unmarked and the plural is considered marked (see Greenberg 1966), an alignment which is clearly contradicted by the inverse number marking pattern. The following sections will also demonstrate that the data from Dagaare is consistent with markedness patterns; however, the universal tendencies are more nuanced than one would expect under the simple alignment between singular/unmarked and plural/marked.

This paper is structured as follows. Section 5.2 discusses some of the basic facts of Dagaare's grammatical number system and then proceeds to establish and test the primary hypothesis of this paper that the distribution of *-ri* correlates with different levels of individuation: nouns unmarked in the singular pattern with highly individuated entities whereas nouns unmarked in the plural pattern with entities which are less individuated and/or tend to appear in groups. This general view is supported both by phenomena within Dagaare, such as the domain of diminutives and dialect variation as discussed in 5.3, and cross-linguistic facts that surface in a wide array of language types, discussed in 5.4. The paper concludes with a formal analysis of the number marker *-ri*, demonstrating that this non-standard method of marking singular and plural distinctions can be aligned with standard semantic theories of number.

4

5.2 The Semantic Basis of Inverse Number Marking in Dagaare

Before presenting evidence for a semantic basis of Dagaare's nominal system, I lay out some of the basic features of Dagaare's nominal system and discuss previous approaches. Additional examples of nouns similar to those in (1) are given in Table 5.1[5], where nouns marked by *-ri* in the plural are shown on the left and those marked by *-ri* in the singular are shown on the right. Table 5.1 also displays instances where *-ri* undergoes assimilation following nasals and liquids, resulting in the allomorphs *-nI* and *-lI*, respectively.

Table 5.1: Number in Dagaare: Marked Plural and Marked Singular Patterns

-V Singular	-rI[6]/-nI Plural	Gloss	rI/-nI Singular	-V Plural	Gloss
tìέ	tììrí	'tree'	lúgrí	lúgó	'prop, pillar'
gbìé	gbèrí	'forehead'	nyágrí	nyágá	'root'
pìέ	pèrí	'basket'	filí	filέ	'sore'
nàŋá	nànní	'scorpion'	íílí	íílὲ	'horn'

The forms in (1) and Table 5.1 can be shown to correspond to singular and plural interpretations by examining their interaction with other elements of the grammar that mark number. Examples (3)-(4) demonstrate distinct number agreement with the singular (*ŋâ*) and plural (*-má*) forms of demonstrative pronouns. In the plural, the agreement prefix further distinguishes between human (*ba-*) and non-human (*a-*) referents.

(3)   a.   bíé   ŋâ
            child.SG DEM.PROX.SG

'this child'

  b. bíí-rí bà-má
    child-PL HUM.PL-DEM.PROX.PL
    'these children'   (HUM = Human Prefix)

(4) a. bí-rì ŋâ
    seed-SG DEM.PROX.SG
    'this seed'

  b. bíê à-má
    seed.PL NHUM.PL-DEM.PROX.PL
    'these seeds'   (NHUM = Non-Human Prefix)

Number words whose value is greater than one also show selection of plural forms and agreement. No agreement is visible, however, in the use of the word designating 'one', *yénì*, which acts as a modifier of the noun, directly attached to the noun stem.[7] Number words designating 'two' or more select for plural nouns and again take plural agreement prefixes. Examples of both nouns from (1) combined with the number words for 'one' and 'two' are given in (5)-(6).

(5) a. bì-yénì (bì- + yénì)
    child-one
    'one child'

  b. bíí-rí bá-yì
    child-PL HUM.PL-two
    'two children'

(6) a. bí-yénì (bí- + yénì)
    seed-one
    'one seed'

b. bíè   à-yí
    seed.PL NHUM.PL-two
    'two seeds'

Much previous work on the nominal system of Dagaare has considered the pattern
in (1) from the perspective of a system of noun classes in Dagaare. Bodomo (1997),
Kropp Dakubu (2005) and Bodomo and Marfo (2006) elaborate systems of noun classes
for Dagaare based upon different singular-plural pairings of nouns.[8] The inverse marking
pattern in these analyses is simply related to a distinction between two different singular/plural
pairings: one is comprised of nouns ending in vowels in the singular and -*ri* in the
plural, while the other is comprised of nouns ending in -*ri* in the singular and vowels
in the plural.

Analyses of the different noun classes in Dagaare are clearly valuable from a diachronic
and comparative perspective. As a Gur language, the nominal system of Dagaare
stems from Proto-Gur, which possessed a highly developed noun class system (Miehe
and Winkelmann 2007) and without a doubt, the pattern observed in (1) historically
derives from a noun class system. At the same time, accounting for the inverse marking
pattern is worthwhile in its own right. If the pattern in (1) were only a minor singular/plural
pairing among many others, it would probably not merit much attention. Yet, in contrast
to related languages, such as Gurenε (Nsoh 2002), which conserve more of the Proto-
Gur system, the noun class system in modern Dagaare is largely decayed. The inverse
pattern of (1) has become the predominant pattern of nominal system, accounting
for over 70% of the nouns in my current database which have singular and plural

7

forms. The two other major singular-plural pairings are for nouns designating humans and for liquids and other typically uncountable nouns which possess a distinct plural, sometimes known as the "second plural" in the Dagaare literature. Thus, accounting for the inverse marking pattern will lead to understanding the principle form of nominal organization in Dagaare.

A different perspective on singular and plural formation in Dagaare is provided by Anttila and Bodomo (2009), who provide detailed morphophonological analyses of the Dagaare nominal system. They uncover a range of regularities governing the morphophonology of Dagaare number inflection, which the account presented here is compatible with.

Building upon this previous research, this paper considers the further question of what predictive factors, if any, govern the occurrence of nouns with *-ri* in the singular or *-ri* in the plural. Number marking in the nominal system of Dagaare is, from all appearances, not directly predictable from the phonological form of the stem. Counter-examples are provided by sets of minimal pairs, similar to the examples in (1) and Table 5.1, given in Table 5.2.

Table 5.2: Near Minimal Pairs

Gloss	Stem	Singular	Plural
'wild rat'	kù-	kúó	kúúrí
'hoe'	kù-	kùùrí	kùé
'granary'	bùg-	bùgó	bùgrí
'pillar'	lùg-	lúgrí	lúgó

I now turn to demonstrating that the answer to this question can be found by examining

the meaning of the nouns at issue.

5.2.1 Number Marking and Individuation

In English as well as cross-linguistically, nouns which refer to count entities differ from nouns which refer to mass entities in their morphosyntactic realization. Commonly cited properties include the ability to be modified by certain quantifiers—mass nouns accept quantifiers such as "much" or "little" ("much/little wine") while count nouns accept quantifiers such as "many" and "few" ("many/few books"), yet the converse does not hold for either class ("?much books"/"?few water"). The literature attempting to account for such distinctions is vast and various, yet much of it reacts in one way or another to the principle of individuation. There are of course divergent perspectives on what individuation designates, but generally the thesis relates cognitive or perceptual qualities of objects to the grammatical realization of count and mass nouns. An early view from Quine (1960) held that count syntax provided an apparatus for individuating objects, viz. delimiting the relevant object from others and tracking its spatio-temporal identity, while mass syntax does not. This view leads to positing a sort of correspondence between syntax and entities in the world. On a strong version of this correspondence theory, language users should conceptualize the referents of count nouns as distinct, countable, individuated things and those of mass nouns as non-distinct, uncountable, unindividuated things (Wisniewski et al., 1996, p. 271). Varieties of this distinction have been picked up in the formal semantics literature, e.g. the atomic/non-atomic distinction in Link (1983), as well as in the psycholinguistic literature (e.g. Bloom

9

1994; Wisniewski, Imai, & Casey, 1996; Barner and Snedecker 2005).

While the individuation hypothesis was primarily elaborated in relation to the mass/count divide, it is reasonable to suppose its influence is relevant within the count domain. First, although the divide between count and mass domains is often loosely spoken of as dichotomous, much work following on Allen (1980) has shown that not all countable nouns are created equal. Rather, evidence from interaction between different determiners and quantifiers demonstrates that there are different levels of countability between true count terms and uncountable mass terms. Accordingly, it is plausible that individuation is related to different levels of countability, and in turn, to the nominal morphology of Dagaare.

While individuation is a commonly cited concept, it suffers in the same manner as other commonly cited conceptual factors in linguistics, such as animacy and agentivity, in that individuation is far from rigorously defined. A rigorous definition will not emerge here either, but rather the strategy is to use individuation as a heuristic to gain insight into the nominal structure of Dagaare and consequently into the functioning of inverse number marking. I will consider four factors linked to the individuation hypothesis and their potential influence on the realization of nominals in Dagaare, which I now discuss along with the evidence for considering them relevant.

The first factor, animacy, receives a good deal of independent support. Animacy, relative to some sort of animacy scale ranging from humans to larger then smaller animals which in turn correlates to a scale of individuation, is known to influence number marking cross-linguistically (Smith-Stark 1974; Corbett 1996, 2000). In particular,

the higher the entity designated by a noun rates on an animacy hierarchy, i.e. the closer to human, the greater the likelihood is that the noun expresses a singular/plural contrast.

Two other factors relating to individuation were proposed by Wierzbicka (1988), and subsequently investigated experimentally by Middleton et al. (2004). First, Wierzbicka argues that nouns designating entities for which the constituents are more easily distinguishable are more likely to be used as count nouns, while those nouns designating entities for which the constituents are not easily distinguishable will be used as mass nouns. For example, she argues that *beans* is more likely to be a count term than *rice* since individual beans are in principle easier to distinguish than individual grains of rice. Middleton et al. (2004) examined this hypothesis experimentally, where subjects had to match a nonce count or mass term with one of two graphical displays of novel aggregates which varied in distinguishability. The graphical displays of novel aggregates were sets of 40 elements where "each element was a simple shape with a black-to-white gradient that appeared slightly 3-dimensional and did not obviously resemble the constituents of any familiar aggregate. (p. 382)" They then presented subjects with pairs of aggregate displays which varied along the dimensions or spatial proximity to other elements (*Close* versus *Apart*) and size of elements (*Large* versus *Small*). Thus, a subject would see two sets of an element where for one set, each element was spatially separated from the other and for the other set each element was spatially contiguous with other elements. The subject would then decide which picture aligned with a phrase such as "This is worgel." The general results were that subjects' choices of count or mass terms were significantly influenced (p< .001) by spatial proximity, but

11

not by the size, of the elements. These results are compelling as the design of the experiment using nonce items ensures that such factors are general.

The second factor argued for by Wierzbicka (1988) is the canonical manner of interaction with a given entity. She exemplifies this with examples such as the naming of berries in Polish, generally count terms because, she claims, people interact with them one by one, viz. picking/eating them, while farmers selling berries typically use mass syntax to describe berries since they interact with them in quantities rather than individually. This factor was investigated via novel objects, again by Middleton et al. (2004). They presented subjects with a novel aggregate, "yellow decorative coarse-grained sugar", in a cardboard box, which the subjects then needed to match to one of two phrases presented in count and mass syntax (e.g. "This is worgel/These are worgels"). The experimenters manipulated the mode of interaction with the aggregate. In the baseline condition, the subjects simply observed the material and then were presented with a response sheet to decide which phrase was appropriate. In the interaction condition, the experimenter and the participants used a thin paper-clip implement to scoop up individual grains of the material and insert each grain into a hole of a board distinct from the box containing the material. The participants then were presented with the response sheet to decide which of two phrases was appropriate, one with mass and one with count syntax. The responses for the baseline and interaction conditions were inversely related: a majority of participants in the baseline condition (69%) selected a mass phrase while a majority of participants in the interaction condition (61%) selected a count phrase.[9] While this result is not definitive, it would appear that the mode of interaction with an aggregate can affect the manner by which it is referred to.

12

The final factor I consider is the likelihood of a noun to be "inherently plural", in other words the likelihood that individual referents of a noun canonically appear as a member of a pair or group, as, for example, is the case for paired body parts (e.g. kidneys). Recent work by Acquaviva (2008) has emphasized the distinctive morphosemantic behavior of entities which canonically appear in collectives, duals and other "marked" number categories. While individuation is normally considered only in light of mass/count syntax, it seems probable that entities that canonically appear as a member of a pair or group, as in the case of duals and collectives, are qualitatively different from those which canonically appear as individuals. This distinction is independent from the previous factors. With regard to animacy, Corbett (1996) previously pointed out that dual/collective paradigms are orthogonal to the animacy scale. Very general number marking patterns, such as the occurrence of plural marking, are correlated with the animacy scale: the higher on the animacy scale the referent of a noun falls, the more likely it is to allow plural marking. Yet, nouns which accept or require dual/collective marking do not systematically align with the animacy scale, thus, such a factor is independent of animacy. Further, the factors of distinguishability and interaction were examined by Middleton et al. (2004) by use of stimuli that only presented groups of entities, which in effect held constant whether a particular type of entity was more likely to appear within a group or singly, again making it plausible that the factor of inherent plurality is distinct from distinguishability or interaction.

5.2.2   Individuation and Inverse Marking

The individuation factors discussed immediately above have been argued, primarily in the case of English, to independently affect the realization of number marking. If individuation is grammatically relevant, a clear hypothesis emerges in considering the inverse number system of Dagaare: the more likely the entity is to be viewed as individuated, the more likely the singular noun will be unmarked and *-ri* will mark the plural; and conversely, the more likely the entity is to be viewed as coming in groups or non-individuated, the more likely the plural noun will be unmarked and *-ri* will mark the singular. The information about a noun's individuation level therefore would be lexical information. Nouns would come with a basic number, determined by its semantic properties, while application of *-ri* gives the inverse value. This can be schematically pictured as in (7):

(7)   [Highly Individuated N] + *-ri* ⇒ plural

           [Less Individuated/Inherently Plural N] + *-ri* ⇒ singular

The considerations of individuation lead straightforwardly to testable predictions. If individuation has an effect on the distribution of *-ri*, one should observe distributional asymmetries in the appropriate semantic domains. In part guided by how transparent verification in a lexicon would be, I explored four relevant predictions:

  (i) Nouns for higher-level (more salient) animals are more likely to be unmarked in the singular than nouns for insects (animacy)

14

(ii) Nouns for trees should be unmarked in the singular in comparison to nouns for vegetation (distinguishability)

(iii) Nouns for tools should be more likely to be unmarked in the singular than the converse (one canonically interacts with them individually)

(iv) Nouns for body parts which inherently come in pairs or groups should be more likely to be unmarked in the plural than not; while nouns for body parts which inherently come in single units should be more likely to be unmarked in the singular than not (inherently plural)

I now turn to the results of fieldwork which bear on these hypotheses.

5.2.3 Results from Fieldwork

To test the predictions elaborated in the preceding section, I conducted fieldwork in Ghana with native speakers to develope a wordlist to determine the behavior of inverse number marking. The findings below are based on a wordlist of nearly 1500 words which I compiled during my field research.[10]

As the hypotheses involved distributions over semantic domains, I coded each word for (relatively transparent) semantic domains, where possible. The chart in Fig. 5.1 displays the results with respect to hypotheses (i)-(iii). The x-axis displays various semantic domains while the y-axis displays the number of lexicon entries. The dark-shaded regions show the number of lexicon entries in a given semantic domain with the singular unmarked, while the light-shaded regions show the number which are

unmarked in the plural and marked by -*ri* in the singular. For instance, the category of *mammal* shows 43 entries in the lexicon that are unmarked in the singular and 5 entries which are unmarked in the plural and marked in the singular by -*ri*. In these counts, I excluded derived forms, since they follow their own patterns which tends to obscure any generalization. Nominal derivation will be discussed in section 5.2.4.

Figure 5.1 demonstrates reliable asymmetries visible across the semantic domains. Nouns for higher-level animates, namely mammals, birds and reptiles are typically unmarked in the singular; however, the majority of nouns for insects have a plural that is unmarked. Similarly, nouns for trees are typically unmarked in the singular, while most nouns for vegetation are unmarked in the plural. Nouns for tools, which were hypothesized to be individuated as a result of the typical manner with which one interacts with them, also showed strong tendency towards being unmarked in the singular.

Figure 5.2 shows similar results for the fourth hypothesis, viz. nouns for body parts which inherently come in pairs or groups should be more likely to be unmarked in the plural while nouns for body parts that inherently come as singular items should be more likely to be unmarked in the singular. The x-axis displays whether the noun is inherently singular, e.g. the term for *head* where canonically humans only have one, or inherently dual/plural, e.g. *eye* or *rib* where canonically humans have two and multiple of each, respectively. Again the y-axis displays the number of items in the lexicon for each category.

16

Figure 5.1: Number Marking Across Semantic Domains
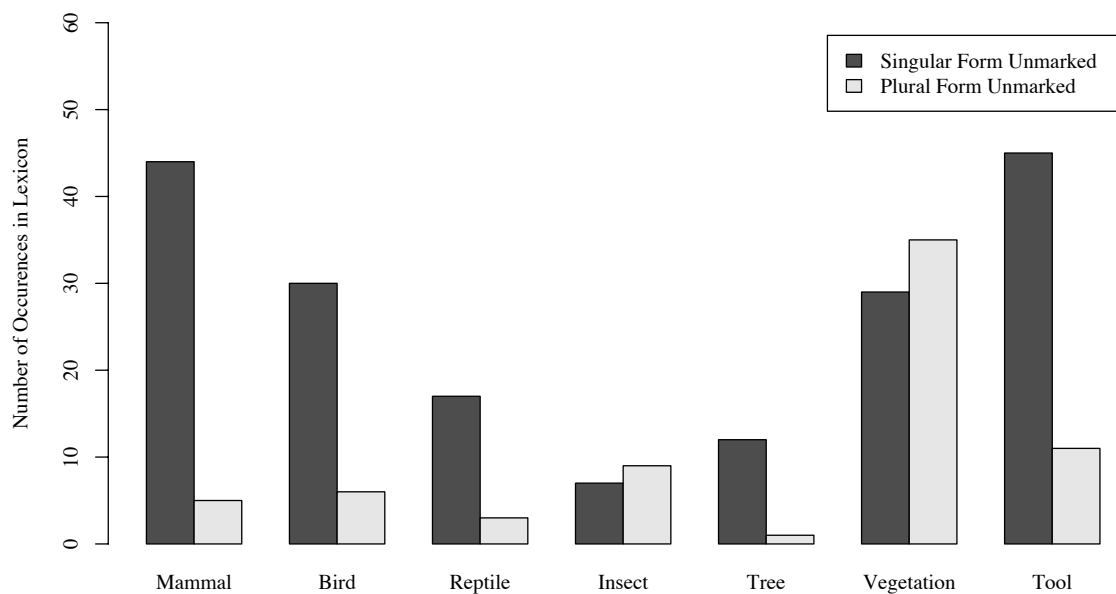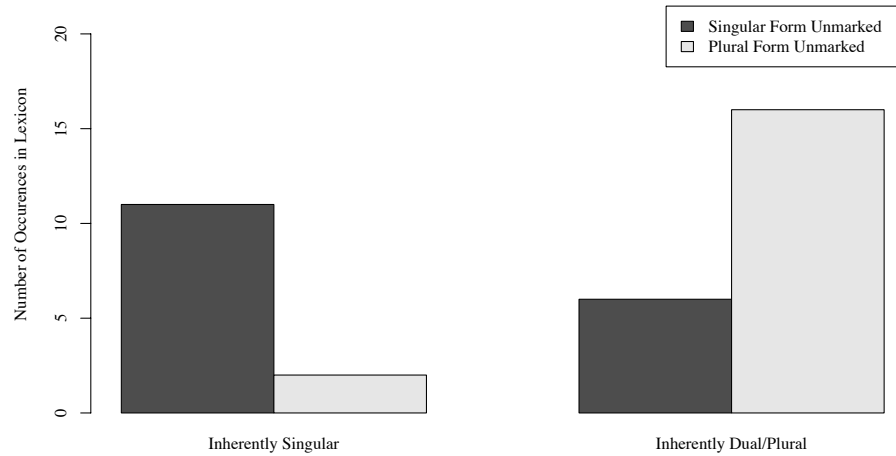
Figure 5.2: Number Marking and Inherent Plurality in the Domain of Body Parts

5.2.4 Discussion

The above results indicate that Dagaare morphology is sensitive to the degree of individuation for the referent of a noun, i.e. *-ri* marks the singular when a noun is considered to be low in individuation/inherently plural, otherwise it marks the plural. Section 5.2.1 laid out a number of hypotheses which made specific predictions about particular semantic domains. When the individuation hypothesis is applied systematically to the lexicon of Dagaare, it uncovers many exceptions; however, most of these demonstrate semantic sub-regularities, or result from practices elsewhere in the grammar, viz. derivational morphology, or historical considerations. I next discuss each of these in turn.

Semantic Sub-regularities

Nouns that do not conform to the general trend of the domain often display semantic sub-regularities. One instance from the animate domain is that most of the nouns for insects unmarked in the singular are for insects capable of causing harm (e.g. scorpion, wasp, spider). In the domain of tools, while the vast majority of nouns were marked by *-ri*, or one of its allomorphs, in the plural, exceptions included nouns such as *fúmìní* (sg) / *fúmìné* (pl) 'needle' or *mírì* (sg) / *míê* (pl) 'rope/string', both of which are consistent with the individuation hypothesis. While 'needle' would appear a strong candidate for canonically appearing in a collection, the canonical rope for Dagaare speakers consists of two strands twisted together, which again would be consonant with inherent plurality. (Cf. English 'twine' which contains the root for 'two'.)

Derived Forms

Several types of derivational processes are in effect in Dagaare, which cloud the homogeneity of any semantic generalizations. In particular, the general number marking trend of a semantic domain may not be realized by a particular member of the domain if its singular-plural pattern is simply due to its derivational history.

Nouns formed by compounding take the number pattern associated with the final element of the compound. While in the domain of nouns for insects, singulars are typically marked by *-ri*, three instances are given in Table 5.3 where the noun is formed by compounding, and the noun's number marking follows that of the final member of

the compound.

Table 5.3: Compound Forms

Singular	Plural	Gloss	Derivational Source
bʊ́ndʊ̀náá	bʊ̀ndʊ̀nnì	'insect' (general)	bʊ̂n 'thing' (n.) + dòn 'bite' (v.) + -aa (agentive suffix)
kyànkyàmpʊ́nàà	kyànkyàmpʊ́nnì	'firefly'	kyankyan 'regularly' (adv.) + pʊ̀n 'flash' (v.) + -aa (agentive suffix)
sɛ́ŋkpɔ̀glí	sɛ́ŋkpɔ̀glɔ̀	'bedbug'	sɛ̌ŋ 'bed' (n.) + kpɔ̀glí 'lump' (n.)

Similarly, exceptions arise in the domain of humans. Nouns for humans normally
fall into a distinct singular-plural pairing, taking the plural marker *-ba*; however, some
do not, especially those that designate human roles ('doctor', 'lord', 'strong person',
etc.). Most of these items exhibit marking by *-ri* for the plural, as one would expect
in the human domain under the individuation hypothesis. Yet, there is again a set
of examples for which the singular is marked by *-ri*, or an assimilated form thereof, a
sampling of which is provided in Table 5.4. These examples all are related to verbs,
from which they are ostensibly derived.[11] Thus, derivational history can again account
for the semantic misalignment of these lexical items under the individuation hypothesis.

Table 5.4: Humans Marked in the Singular by *-ri*

Singular	Plural	Gloss	Derivational Source
gbɛ́rí	gbɛ́ɛ́	'cripple'	gbɛ́rí(v.) 'to cripple'
sèní	sɛ́nímɛ́	'girlfriend'	sὲŋ (v.) 'to befriend'
zɔ̀lí	zɔ́lɔ̀	'fool'	zɔ́lí(v.) 'to be foolish'

Etymological Source

A final source of misalignment between the present semantic hypothesis and the lexical items of Dagaare is provided by their etymology.

Borrowing Number marking for words borrowed into Dagaare appear to be influenced by phonetic similarity, as can be seen from the example *lɔɔrì* (sg) / *lɔ̀ɛ̀* (pl) 'truck; lorry', clearly a borrowing from the English word *lorry*. The Dagaare singular form *lɔɔrì* corresponds to the English singular form *lorry*, and the Dagaare plural is formed on analogy with words for which the stem is vowel final and the singular is marked by -*ri*. A similar explanation can be found for the term *tìrí* (sg) / *tìɛ́* (pl), borrowed from Akan (Mark Ali, p.c.). Table 5.5 shows three Dagaare words which display the pattern which is extended to *lɔɔrì* and *tìrí* by analogy.

Table 5.5: Borrowing and Paradigmatic Analogy

Stem	Singular	Plural	Gloss	
gbɛ-	gbérì	gbɛ́ɛ̀	'leg'	
lɛ-	lɛ̀rí	lɛ́ɛ̀	'bead'	
lo-	lórí	lóé	'meteorite'	
loo-	lɔɔrì	lɔ̀ɛ̀	'truck ; lorry'	> English
tɪr-	tìrí	tìɛ́	'spoon'	> Akan

Semantic Shift Some of the lexical entries which mark their singular with -*ri* have primary meanings that appear to be at odds with the individuation hypothesis, yet they retain traces of an antecedent meaning which accords with the hypothesis. An

instance of this phenomenon is *yírì* (sg) / *yíè* (pl) 'house'. The physical entity 'house' would appear to be a canonical example of an individuated entity. But, as often is the case, the canonical gloss obscures the full range of meanings associated with the word. While synchronically the most common use of *yírì* is to designate 'house', it has antecedent meanings as 'compound' (Durand 1953) as well as 'family' or 'family members' (Mark Ali, p.c.). In its uses for 'compound' and 'family', *yírì* designates something closer to a collection of entities, i.e. units of the compound/family members, and thus is more aligned with the notion of inherently plural/collective entities.

5.3   Language Internal Correlates

The preceding section offered a principled explanation for the organization of Dagaare's nominal systems. In doing so, appeal was made to an assumption that Dagaare is sensitive to degrees of individuation and to collections as opposed to individuals. This section adduces further data from Dagaare that validates the implications of this assumption, and as such increases the plausibility that Dagaare is organized in such a fashion.

5.3.1   Dialect Variation

I have argued that individuation plays an organizing role in the choice of nominal inflection in Dagaare. Given that degrees of individuation are akin to a scale-structure, one would expect to see dialect variation in the mid-region of the scale, i.e. entities which are not clearly individuated or group-like would be predicted to vary. Adams

Bodomo (p.c.) has noted that there are instances where the direction of number marking differs among dialects, as shown for the noun stem *pi-* 'rock' in Table 5.6.

Table 5.6: Dialect Variation: Variation in Directionality of Marking

Singular	Plural	Gloss
pììrí	pìé	'rock' (Central Dialect)
pìé	pììrí	'rock' (Southern Dialect)

While nouns such as 'human' and 'rib' are naturally associated with individual and collective interpretations, respectively, items such as 'rock' could in principle be associated with either individual rocks or collections of rocks. Such claims must be assessed through further research, yet even so, aligning number formation with the propensity towards individuation provides an explanation for variation where purely morphophonological considerations would be hard-pressed to do so.

A second source of variation is found in the choice between *-ri* and a singulative marker, *-ruu*, which Dagaare employs to designate "a piece of" for a limited set of nouns. The singulative appears mainly with clear mass terms as well as aggregates which are particularly close-knit, as shown in Table 5.7, where due to the nasal environment the singulative marker is realized as *-nuu*.

Table 5.7: Singulative Paradigm

Singular/Base	Gloss	Singulative	Gloss
mʊ́ɔ́	'grass'	mʊ́ɔ́nʊ́ʊ́	'blade of grass'
súnnì	'gum'	súnnúú	'piece of gum'

Several words in Dagaare, as shown in Table 5.8, vary across dialects between whether

a noun marks its singular form with *-ri* or *-ruu*. This dialect variation in turn supports
the main hypothesis that *-ri* marks the singular for objects which are inherently plural.
The use of the two different markers implies that there is overlap between inherently
plural and mass/aggregate terms.

Table 5.8: Dialect Variation: Singular vs. Singulative

Variant	Singular	Plural	Gloss
I:	váálí	váálá	'rubbish'
II:	váálúú	váálá	'rubbish'
I:	kómmírí	kómmíé	'tomato'
II:	kómmírúú	kómmíé	'tomato'

5.3.2   Domains of the diminutive

Another area of Dagaare morphology where sensitivity to the collections/inherently
plural distinction occurs is in the domains of its diminutive morphemes. Dagaare
actually makes use of two separate diminutives, *-lee* ('small') and *-biri* ('seed') illustrated
in Table 5.9, each of which highlights different associated meanings. In the instance of
-lee, the application of the diminutive marker results in the meaning 'small/young',
while the application of *-biri* derives lexical items that are associated with the base
noun, often through a part/whole relation. Of particular interest is that *-biri* is used
predominately for describing entities which come in collections, viz. toes, fingers, germs/bacteria
("sick seeds"), beads, bullets, ribs, words ("speech seeds"), stars, bees ("honey seeds").
Uses of *-biri* occasionally describe items which do not come in collections, but which
are associated with the base noun through relations such as 'small' or 'part of', as well

24

as through the literal meaning 'seed of', but these are comparatively rare.[12]

Table 5.9: Diminutives in Dagaare

Diminutive	Stem	Gloss	Derived Noun (Sg)	Gloss
	bà	'dog'	bàléé	'puppy'
-lee	náábʊ̀	'cow'	nààléé	'calf'
	gàŋg áá	'drum'	gàŋgàléé	'alto drum'
	gbɛ́-	'leg/foot'	gbɛ́bìrí	'toe'
	nú-	'hand'	núbìrí	'finger'
	sí-	'honey'	síbírí	'bee'
-biri	málfà-	'gun'	málfàbírì	'bullet'
	yèl-	'say, tell (v.)'	yèlbírì	'word'
	bààl-	'sick'	bààlʊ́mbírì	'germ'

Languages more often employ just one diminutive from which all the diminutive senses are derived (see the survey in Jurafsky 1996). An example is provided by Ewe, which, like Dagaare, has a diminutive which is taken from the word for 'child', -ví; however, -ví applies across both senses seen in Dagaare, resulting in derivations for 'young/small X' as well as items that come in collections ('toes'), as seen in Table 5.10 (data from Heine and Kuteva 2009).

Table 5.10: The Diminutive in Ewe (data from Heine and Kuteva 2009)

Stem	Gloss	Derived Noun (Sg)	Gloss
ŋútsu	'man'	ŋútsu-ví	'boy'
yevú	'European'	yevú-ví	'young European'
nyi	'cow'	nyi-ví	'calf'
kpé	'stone'	kpé-ví	'small stone'
du	'small village'	du-ví	'small village'
afɔ	'foot, leg'	afɔ-ví	'toe'

In contrast to the more common pattern, where one morpheme is generalized to

many related senses, Dagaare marks the collective sense separately from other diminutive functions. In sum, *-lee* and *-biri* are not equivalent : *-lee* is closer to a standard diminutive generally meaning 'young' or 'small' while the employment of *-biri* most often yields a collective sense, highlighting another area where Dagaare is sensitive to a distinction between individuated entities and collections.

5.4   Cross-Linguistic Correlates

Another source of support for the assumptions underlying the primary hypothesis that individuation underlies the organization of Dagaare's nominal number marking system can be adduced from cross-linguistic data. Section 5.2 identified a set of semantic domains that are typically unmarked in the plural. In the same manner in which one expects certain features of the mass terms to be consistent across languages, viz. not accepting cardinal terms without a measure term, one would expect the behavior of the nouns associated with semantic domains unmarked in the plural to have parallel behavior across languages. In this section, I will demonstrate that cross-linguistic correlates to the unmarked plural in Dagaare surface in an array of language types, demonstrating that, despite different encodings, these systems seem to make similar divisions on a scale of individuation.

5.4.1 Nominal Paradigms

Semantic domains similar to those discussed in section 5.2 are cross-linguistically relevant for collectives and duals (see discussion in Acquaviva 2008). One example that accords quite well with the findings in Dagaare is the collective/singulative class in Welsh, discussed in Stolz (2001). Welsh disposes of a singular/plural distinction for count nouns just as in, say, English, where the plural is morphologically marked; however, for select semantic domains, a collective interpretation is morphologically unmarked, while a singular interpretation is signalled by a singulative marker *-yn* or *-en*. For instance, the Welsh term for 'flea' is *chwain*, a collective term whose singulative form is *chwann-en*. Of interest are the semantic domains where this holds: small animals and insects, vegetables/cereals/fruits, body parts ('ribs', etc.), and what Stolz terms "uncountable substance", essentially granular mass terms ('turf', 'embers', 'sand'), all of which accord with the semantic domains seen as unmarked plurals in Dagaare.[13] A similar division is in effect for languages with nominal class systems, e.g. Swahili (Contini-Morava 2000) and Lingala (Mufwene 1980), where some noun classes appear to be unmarked in the plural. Once again, the relevant semantic domains are strikingly similar to those in play for Dagaare, e.g. vegetation, pairs and collectives.

5.4.2 Morphological Behavior

The semantic domains discussed in section 5.2 also manifest unexpected behavior with respect to morphological processes. Tiersma (1982) noted that classes of nouns for

entities that "naturally occur in pairs or groups" tend to show surprising behavior with respect to morphological leveling, borrowing and double plural formation. Tiersma claims that "when a referent of a noun naturally occurs in pairs or groups, and/or when it is generally referred to collectively, such a noun is locally unmarked in the plural." Local markedness then is a semantic notion of markedness, which may effect how nouns are morphologically marked. This characterization aligns well with the distributional patterns established for Dagaare and converges with the individuation hypothesis. This section reviews the evidence from morphological processes which support the notion of "unmarked plurals".

Morphological Leveling   Tiersma (1982) provides a detailed discussion of instances of paradigm leveling, a form of analogical change whereby a paradigm regularizes, as in the change from Early Modern English *reach/raught* to *reach/reached*. Tiersma notes that while morphological paradigms typically level towards the unmarked members of the paradigm, in certain cases nominal paradigms level in favor of the plural stem. This is exemplified in Modern Frisian which exhibits "breaking" in conservative forms of singular/plural pairs, i.e. an alternation between falling and rising diphthongs, shown in Table 5.11 for the pairs /ie/ and /ji/ and /oa/ and /wa/. This alternation is undergoing leveling, where typically the falling diphthong, which is associated with a singular stem, is generalized; however, for a small number of items, the leveling generalized instead towards the rising diphthong, associated with the plural stem. As the selections of nouns in Table 5.11 indicate, this occurs in instances where the referent of the noun "naturally occurs in pairs or groups".

Table 5.11: Morphological Leveling in Frisian (from Tiersma 1982)

Conservative	Innovative	Gloss
kies (sg)/ kjizzen (pl)	kjizze (sg) / kjizzen (pl)	'tooth'
toarn (sg)/ twarnen (pl)	twarne (sg) / twarnen (pl)	'thorn'

Borrowing   As with leveling, borrowing typically proceeds by taking the unmarked singular stem; however, there are cases, and not surprisingly in the same semantic domains, in which the plural form is borrowed in preference to the singular. A clear example is provided by Welsh borrowings from English shown in Table 5.12, discussed in Stolz (2001), where the borrowed plural form from English serves as the basic term which can then be inflected for the singulative.

Table 5.12: Borrowing in Welsh

Singulative	Collective	Source
ffigys-en	ffigys	figs
gwsbery(s)-en	gwsberys	gooseberries
pys-en	pys	peas

Double Plurals   The formation of double plurals occurs when older plurals have been reanalyzed as singular units, leading to the addition of another plural marker. Tiersma (1982) provides examples from West Frisian which disposes of two productive plural markers, *-en* and *-s*. The examples in Table 5.13 show both markers have been applied. The referents here again involve an element of inherent plurality, which ostensibly provides a motivation for the reanalysis, as the plural form is the more basic, or "locally unmarked" form.

29

Table 5.13: Double Plurals in West Frisian (from Tiersma 1982)

Singular	Plural	Gloss
boei	boeijens	'handcuff'
lears	learzens	'boot'
reed	redens	'skate'
trep	treppens	'stair'

5.4.3 Evidence from English Frequency Patterns

Throughout my discussion of Dagaare, and the above morphological patterns from other languages, I have referred to "unmarked plurals". This term has been appropriate insomuch as these plurals have had less morphological material as opposed to a clearly suffixed singular, and thereby qualifying as basic. Yet, for an element to be "unmarked" has another implication in terms of text frequency, as in Greenberg (1966) where unmarked forms were shown to have greater text frequency than marked ones. If the semantic domains I have discussed are truly unmarked, one would expect to see "unmarked plurals" in languages which do not display any morphological evidence of such a pattern reflected in terms of text frequency. In order to evaluate this prediction, I examined frequencies for the semantic domains of *animal* and *insect* from the COBUILD corpus (18 million words) provided by CELEX. Using basic terms and terms consistent with the vocabulary of Dagaare, I calculated the plural-to-singular ratio for these two domains, shown in Fig. 5.3, where the x-axis represents the ratio of the token frequency of plurals to the token frequency of singulars and the y-axis represents the number of lexical items. The graph indicates that there is a clear trend for insect terms to have a plural/singular ratio greater than 1, i.e. insect terms occur more frequently in the plural, while animal

Figure 5.3: Number Marking Frequency Patterns in English for *Animals* and *Insects*
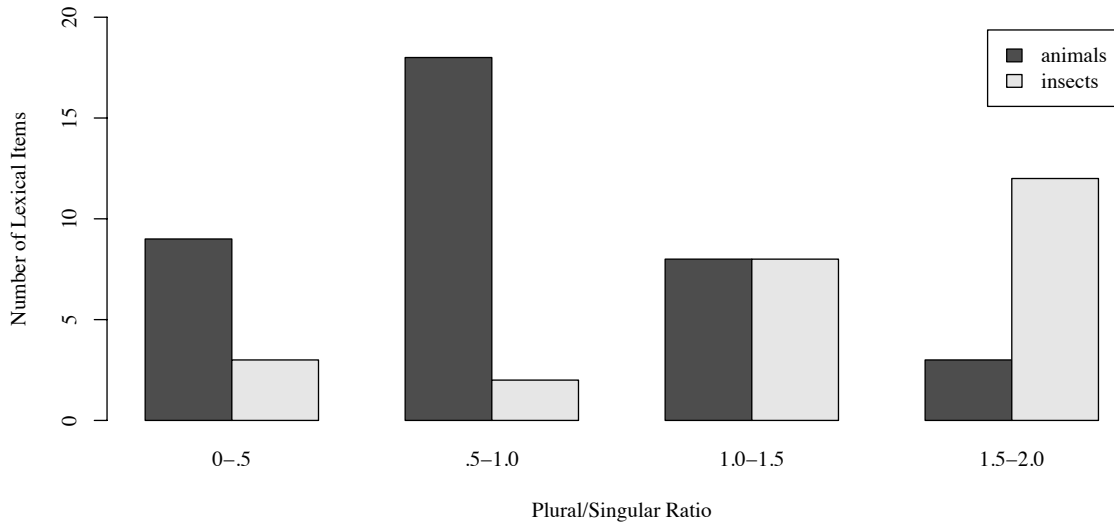
terms tend to have a plural/singular ratio less than 1, i.e. animal terms occur more frequently in the singular. This finding lends additional support to the arguments about morphological patterns in this section, as well as to the assumption that the plural of certain domains as "unmarked".

5.5 A Formal Account of *-ri*

I have established that Dagaare is sensitive to the degree of individuation and inherent plurality associated with the referents of nominal elements. Indeed, this should not come as a surprise since such facts have been consistently assumed with respect to the count/mass divide. Here, I have made the case that individuation is matter of degree,

sensitivity to which pervades the nominal system and is not limited to partitioning nouns into count and mass terms. This section will demonstrate that once inherent plurality, and thereby singularity, become available as lexical information, the semantics of inverse number marking can be aligned with formal models of the semantics of number. Further, the empirical data from Dagaare bears upon a controversy among different popular analyses of the plural.

Most analyses of the semantics of number since at least Link (1983) base their model of the count domain on two ingredients: (i) a set of atomic objects in some domain A, which correspond to individual entities such as a dog or a chair and (ii) a domain E, where $E \subseteq A$, containing sets generated from the atomic objects in A, which is structured by a part-whole or subset relation relating the sets of atoms. The singular of a noun denotes the atomic entities for which the noun is a true description, while the plural of a noun denotes, at least, sets of atomic entities for which the noun is a true description. The whole domain, $E \subseteq A$, possesses the structure of a join semi-lattice. Beyond the basic agreement on these two ingredients, however, there has been substantial debate as to what is contained in the domain of the plural.[14]

One line of proposals, originating with Link (1983), models the plural as denoting the closure of atoms under join ($\oplus$) less the atoms themselves, thus the denotation of the plural excludes that of the singular. In this treatment, the singular (atomic) denotation is simpler, and thus the unmarked form.

Many researchers, including Krifka (1989) and Sauerland et al. (2005), have noted that this account is problematic for environments such as negation or in questions.

For instance, if the plural designates *two or more*, then the statement *Ed didn't see horses* would be true if Ed had seen only one horse, but that is clearly not the desired result. Such facts have been taken to motivate a weaker plural designating *one or more*, whereby the denotation of the plural includes that of the singular, in formal terms corresponding to the entire semi-lattice structure of the denotation of a noun. Under this analysis, the singular is more specific than the plural, and the plural surfaces as the unmarked number.

The two sides of this debate differ on whether (i) the singular or plural is unmarked and (ii) the model-theoretic structure of the domain of plurals—whether the plural is better modeled as intrinsically exclusive ($sums - atoms$) or inclusive ($sums \cup atoms$). As to the first point, the data from Dagaare reviewed above reveal a more intricate picture. Dagaare demonstrates that the cross-linguistic facts are more complicated than if only the singular *or* plural were unmarked, rather markedness is conditioned upon a nominal's level of individuation. In order to evaluate the second point in light of the data from Dagaare, I consider two analyses, one consistent with the exclusive and the other consistent with the inclusive plural. Applying the logic of both exclusive and inclusive plural analyses to data from Dagaare shows the exclusive plural analysis has better empirical traction in inverse number marking systems.

The Exclusive Plural   Assuming inherent plurality, and thereby singularity, becomes available as lexical information, as argued in the above sections, and assuming the exclusive plural analysis, the semantics of inverse number marking is relatively straightforward: *-ri* is simply treated as a form of negation of the unmarked number value for the noun

33

in question. This is an intuitive version of the function of inverse number marking, and is in essence a formal semantic update of the analysis of Kiowa in Wonderly (1954).[15] Further assuming along with Ojeda (1998) that the base or root of the noun has a denotation of the entire space generated by the atoms and their sums ($atoms \cup sums$), i.e. the base is compatible with singular and plural individuals, then *-ri* can be modeled as the operation of complementation (C), relativized to the domain of the base. The degree of individuation determines whether a noun is considered lexically plural or singular, whereupon *-ri* applied to a lexically singular noun will yield a plural denotation, while if *-ri* is applied to a lexically plural noun, it will yield a singular denotation.

Representative derivations are given in Table 5.14, demonstrating that this analysis clearly secures the desired interpretations. In prose, for lexically singular nouns, the application of *-ri* gives the complement of the denotation of a singular noun, viz. the complement of the relevant set of atoms. The value returned is the sums formed from the atoms, less the atoms themselves, which is in turn exactly the value of the noun's plural denotation. For lexically plural nouns, the application of *-ri* gives the complement of the denotation of a plural noun, viz. the complement of the relevant set of sums. The value returned is the atoms which form the sums, which is in turn exactly the value of the noun's singular denotation.

The Inclusive Plural An alternate analysis, which is consistent with weak plural analyses, models *-ri* as designating the completion of the space.[16] The weak plural analysis of English plurals claims that the plural is unmarked, denoting closure under join, while the singular, designating atoms, is more specific. When the plural form is

Table 5.14: Derivations of Lexically Singular ('child') and Plural ('seed') Nouns with the Exclusive Plural

Lexically Singular	Lexically Plural
$[\![$bi-$]\!] := \lambda x(CHILD(x))$	$[\![$bi-$]\!] := \lambda x((SEED(x))^{\oplus} - SEED(x))$
$[\![$bi-$]\!]$ + ri	$[\![$bi-$]\!]$ + ri
$([\![$ bi- $]\!])^{C}$	$([\![$ bi- $]\!])^{C}$
$[\lambda x(CHILD(x))]^{C}$	$[\lambda x((SEED(x))^{\oplus} - SEED(x))]^{C}$
$\lambda x[(CHILD(x))^{\oplus} - CHILD(x)]$	$\lambda x[SEED(x)]$
= PL(bi-)	= SG(bi-)

used, the singular interpretation is excluded by pragmatic blocking. The same inferences found in negative and interrogative contexts in English which motivate the weak plural analysis were also elicited in Dagaare, thus one could think to analyze *-ri* when marking the plural just as the English plural is analyzed, designating closure under join, with the singular interpretation disallowed by blocking.

The inclusive plural analysis can be extended to *-ri* by positing a parallel analysis for nouns which are lexically plural. In this case, the application of *-ri* to lexically plural nouns must also yield the entire semi-lattice, whereupon the plural interpretation is disallowed by blocking. This can be achieved through modeling *-ri* as forming the ideal of the denotation of the plural, the set of sums. A subset of a lattice is an ideal if it is a non-empty downward closed set[17] which is also closed under join. When applied to the set of sums, closure under join is already satisfied while downward closure results in including all the atoms.

Ideal formation will also secure the desired result for lexically singular nouns, whereby *-ri* can be uniformly analyzed as the closure of the space under ideal formation. Downwards

closure is already satisfied since by definition atoms have no proper parts, and closure under join results in the inclusion of all the sums. The singular interpretation then is disallowed by blocking. Representative derivations are given in Table 5.15, where ^{Cl} represents the ideal formation closure operator. In prose, for lexically singular nouns, the application of -*ri* gives the closure of the denotation of a singular noun, which is the entire semi-lattice. For lexically plural nouns, the application of -*ri* gives the closure of the denotation of a plural noun, which is again the entire semi-lattice.

Table 5.15: Derivations of Lexically Singular ('child') and Plural ('seed') Nouns with the Inclusive Plural

Lexically Singular	Lexically Plural
$[\![\text{bi-}]\!] := \lambda x(\text{CHILD}(x))$	$[\![\text{bi-}]\!] := \lambda x((\text{SEED}(x))^{\oplus} - \text{SEED}(x))$
$[\![\text{bi-}]\!] + \text{ri}$	$[\![\text{bi-}]\!] + \text{ri}$
$([\![\text{bi-}]\!])^{Cl}$	$([\![\text{bi-}]\!])^{Cl}$
$[\lambda x(\text{CHILD}(x))]^{Cl}$	$[\lambda x((\text{SEED}(x))^{\oplus} - \text{SEED}(x))]^{Cl}$
$\lambda x((\text{CHILD}(x))^{\oplus})$	$\lambda x((\text{SEED}(x))^{\oplus})$
= inclusive(bi-)	= inclusive(bi-)

Behavior under negation would demonstrate whether such a suggestion was feasible, for in many languages, such as English and as was elicited in Dagaare, negation of the plural always also excludes the truth of the singular. Extending the inclusive plural analysis to -*ri* in this way predicts that the negation of the form marked by -*ri* should exclude the truth of both singular and plural. The example in (8) shows that this turns out not to be the case (Adams Bodomo p.c.):

(8)　　n　　dà　bá　dà　bíe　　(zàà)
　　　1st.pro NEG buy Past seed.PL (any)
　　　I didn't buy (any) seeds.

The same entailment patterns hold in Dagaare as in English and in (8) the negated plural also indicates that the speaker did not buy a single seed; however, the form negated is not marked by *-ri*, but rather is the unmarked form. As the data does not align with the logic of the inclusive plural analysis, this analysis must be rejected in favor of the exclusive plural analysis.

In concluding this section, it is noteworthy that the ability to align the number system of Dagaare with standard semantic accounts indicates that the semantics of number in Dagaare are similar to better described languages, in other words Dagaare speakers may communicate similar meanings, it is just the means of expression, here the morphology, that differs.

5.6   Conclusion

This paper has demonstrated that number marking in Dagaare, and more generally inverse number marking, which while at first sight surprising, under closer inspection is a clever exploitation of relatively universal markedness patterns, namely less individuated/inherently plural entities which are semantically unmarked in the plural are also morphologically unmarked in the plural. Once the lexical generalizations are clear, the formal implementation is straightforward. Such systems reinforce the point that number marking comprises more than simply marking a dichotomy between reference to atoms or reference to sums, rather the canonical properties of the referents themselves influence how the marking is achieved. Articulating the precise connections between the lexical properties

of nouns and number realization remains a fertile area for further exploration.

Bibliography

Paolo Acquaviva. *Lexical Plurals: A Morphosemantic Approach.* Oxford, 2008.

Keith Allen. Nouns and countability. *Language*, 56(3):41–67, 1980.

Arto Anttila and Adams Bodomo. A Dagaare-English Word List. Unpublished Manuscript (2006 Revision), 2006.

Arto Anttila and Adams Bodomo. Prosodic morphology in Dagaare. In Fiona McLaughlin, Eric Potsdam, and Masangu Matondo, editors, *Selected Proceedings of the 38th Annual Conference on African Linguistics (ACAL 38)*, pages 56–68. Cascadilla Proceedings Project, Somerville, Massaschusetts, 2009.

Emmon Bach. Parochial and universal semantics: Semantic typology and little studied languages. Handout for pre-ALT Workshop on Language Documentation and Linguistic Typology, 2007.

Emmon Bach. On morphosemantics: the internal meanings of words. Paper presented at the Copenhagen Linguistic Circle, 2008.

Matthew Baerman. Morphological reversals. *Journal of Linguistics*, 43:33–51, 2007.

D. Barner and J. Snedeker. Quantity judgments and individuation: Evidence that mass nouns count. *Cognition*, 97:41–66, 1996.

Paul Bloom. Possible names: The role of syntax-semantics mappings in the acquisition of nominals. *Lingua*, 92:297–329, 1994.

Paul Bloom. The role of semantics in solving the bootstrapping problem. In R. Jackendoff, P. Bloom, and K. Wynn, editors, *Language, logic, and concepts: Essays in memory of John Macnamara*. MIT Press, Cambridge, MA, 1999.

Adams Bodomo. *The Structure of Dagaare*. CSLI, Stanford, 1997.

Adams Bodomo and Charles Marfo. The morphophonology of noun classes in Dagaare and Akan. *Studi Linguistici e Fonologici Online*, 2006. `http://www.humnet.unipi.it/slifo/`.

Ellen Contini-Morava. Noun class as number in Swahili. In E. Contini-Morava and Y. Tobin, editors, *Between Lexicon and Grammar*. John Benjamins, Amsterdam, 2000.

Greville G. Corbett. *Gender*. Cambridge University Press, Cambridge, 1991.

Greville G. Corbett. Minor number and the plurality split. *Rivista di Linguistica*, 8 (1):101–122, 1996.

Greville G. Corbett. *Number*. Cambridge University Press, Cambridge, 2000.

B. A. Davey and H. A. Priesltey. *Introduction to Lattices and Order, Second Edition*. Cambridge University Press, Cambridge, 2002.

Donka F. Farkas and Henriëtte de Swart. The semantics and pragmatics of plurals. *Semantics and Pragmantics*, 3(6):1–54, 2010.

D. Gentner and Lera Boroditsky. Individuation, relational relativity and early word learning. In M. Bowerman and S. Levinson, editors, *Language acquisition and conceptual development*, pages 215–256. Cambridge University Press, Cambridge, 2001.

Joseph Greenberg. *Language Universals.* Mouton, The Hague, 1966.

Tom Güldemann. Noun categorization systems in Non-Khoe lineages of Khosian. *Afrikanistische Arbeitspapiere*, 63:5–33, 2000.

Daniel Harbour. *Morphosemantic Number: From Kiowa Noun Classes to UG Number Features.* Springer, 2008.

Bernd Heine and Tania Kuteva. The genesis of grammar: on combining nouns. In Rudolf Botha and Henriëtte de Swart, editors, *Language evolution: the view from restricted linguistic systems.* LOT, 2009.

Daniel Jurafsky. Universal tendencies in the semantics of the diminutive. *Language*, 72:533–578, 1996.

Manfred Krifka. Nominal reference, temporal constitution and quantification in event semantics. In J. van Benthem R. Bartsch and P. von Emde Boas, editors, *Semantics and Contextual Expression*, pages 75–115. Foris Publication, Dordrecht, 1989.

M. E. Kropp Dakubu. *Collected Language Notes on Dagaare Grammar*. Collected Language Notes v.26. Institute of Language Studies, Legon, Ghana, 2005.

Godehard Link. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In R. Bauerle, C. Schwarze, and A. Stechow, editors, *Meaning, use, and interpretation of language*, pages 302–323. de Gruyter, Berlin, 1983.

Erica L. Middleton, Edward J. Wisniewski, Kelly A. Trindel, and Mutsumi Imai. Separating the chaff from the oats: Evidence for a conceptual distinction between count noun and mass noun aggregates. *Journal of Memory and Language*, 50:371–394, 2004.

G. Miehe and K. Winkelmann, editors. *Noun Class Systems in Gur Languages, Vol. I Southwestern Gur Languages (without Gurunsi)*. Rüdiger Köppe Verlag, Cologne, 2007.

Salikoko S. Mufwene. Number, countability and markedness in Lingala *li- ma-* noun class. *Linguistics*, 18:1019–1052, 1980.

Salikoko S. Mufwene. The count/mass distinction and the english lexicon. In *Parasession on lexical semantics*, pages 200–221, Chicago, 1984. Chicago Linguistics Society.

Avea Nsoh. Classifying the nominal in the Gurenɛ dialect of Farefare of Northern Ghana. *Journal of Dagaare Studies*, 2:83–95, 2002.

Almerindo Ojeda. The semantics of collectives and distributives in Papago. *Natural Language Semantics*, 6:245–270, 1998.

W. V. O. Quine. *Word and Object.* MIT Press, Cambridge, MA, USA, 1960.

Uli Sauerland, Jan Anderssen, and Kazuko Yatsushiro. The plural is semantically unmarked. In S. Kepser and M. Reis, editors, *Linguistic Evidence: Empirical, Theoretical, and Computational Perspectives*, pages 413–434, Berlin, 2005. Mouton de Gruyter.

T. Cedric Smith-Stark. The plurality split. In M.W. La Galy, R.A. Fox, and A. Bruck, editors, *Papers from the 10th regional meeting*, pages 657–671. Chicago Linguistics Society, 1974.

Thomas Stolz. Singulative-collective: Natural morphology and stable cases in Welsh number inflexion of nouns. *Sprachtypologie und Universalienforschung*, 54(1):52–76, 2001.

Peter Tiersma. Local and general markedness. *Language*, 58:832–849, 1982.

Laurel J. Watkins. *A Grammar of Kiowa.* PhD thesis, University of Kansas, 1980.

A. Wierzbicka. *The Semantics of Grammar.* John Benjamins, Amsterdam, 1988.

E. J. Wisniewski, M. Imai, and L. Casey. On the equivalence of superordinate concepts. *Cognition*, 60:269–298, 1996.

William J. Wonderly. Number in Kiowa: Nouns, demonstratives and adjectives. *International Journal of American Linguistics*, 20:1–7, 1954.

Notes

[2]The pairs differ in tone and in that the plural of 'child' has a vowel which appears to have undergone lengthening. Anttila and Bodomo (2009) propose that Dagaare phonology

makes use of lexically conditioned high vowel lengthening in such cases.

[3]As marked in the examples, Dagaare possesses two levels of tone, high and low. The morpheme *-ri* itself is, however, unmarked for tone.

[4]For discussion of Kiowa, see Watkins (1980) and Harbour (2008). Nehan is discussed both in Corbett (2000) and Baerman (2007), both of whom provide general discussions of number marking reversals.

[5]Vowels in Dagaare distinguish whether advance tongue root (ATR) is present or not. The standard orthography for Dagaare collapses the representations for the vowels /o/ and /ʊ/ as well as for /e/ and /ɪ/, but I have followed Bodomo (1997)'s orthographic system in which these distinctions are made. For the case at hand, this permits clearly seeing that *-ri* and *-rɪ* are allomorphs which differ only due to ATR-harmony.

[6]Capital letters for vowels indicate +/- ATR.

[7]In Dagaare, modifiers such as adjectives are compounded with the noun stem and the adjective then supplies its own singular/plural suffix patterns.

[8]This classification, although formed from singular-plural pairings, is not a gender system in the sense of Corbett (1991) or Güldemann (2000), where genders are established based on agreement classes. Dagaare has little concord phenomena, and the only three agreement classes that can be established are those seen in examples (3)-(4): singular, human plural and non-human plural.

[9]This distribution is significantly above chance (p<.05).

[10]I would like to acknowledge Arto Anttila and Adams Bodomo for generously permitting me to incorporate elements from their wordlist, Anttila and Bodomo (2006).

[11]Dagaare seems to possess a semi-regular process whereby nouns can be zero-derived

from verbs and adjectives, although this subject needs further research.

[12]According to my current database, twenty nouns derived with *-biri* result in some sort of collective sense similar to those nouns shown in Table 5.9, six derived nouns indicate the literal 'seed of', which again result typically in a collection of entities, while the associated senses of 'small' and 'part of' occur only with two nouns each.

[13]Granular mass terms in Dagaare have several realizations, but most often only have a form marked by *-ri*, as in *bigiri* 'hemp' or *zɔŋɪ* 'flour'.

[14]See Farkas and de Swart (2010) for extensive discussion of the debate.

[15]This line has also been developed independently in Bach (2007) and Bach (2008) for Kiowa.

[16]I am indebted to Uli Sauerland for suggesting this line of analysis.

[17]A formal definition of downward closure in a lattice is if L is a lattice and J is a non-empty subset, then $a \in L$, $b \in J$ and $a \leq b$ imply $a \in J$ (Davey and Priestley 2002, p. 44).