

Number and Markedness: A View from Dagaare

Scott Grimm
Department of Linguistics
Stanford University

sgrimm@stanford.edu

Abstract

Semantic theories of number currently debate whether the plural is weak (referring to both collections and individuals) or strong (referring to collections but not individuals). The former view holds that the plural is ‘unmarked’ as it is less specific; for the latter view, the singular is simpler and therefore ‘unmarked’. This paper examines the inverse number marking system of Dagaare (Gur; Niger-Congo), which proves problematic for both sides of the debate. The data demonstrate that the cross-linguistic facts are more complicated than if only the singular *or* plural were unmarked, instead markedness is conditioned upon a nominal’s level of individuation. Applying the logic of both strong and weak plural analyses shows that the strong plural analysis has better empirical traction in such systems.

1 Introduction

Semantic theories of number have long debated the nature of the contrast between singular and plural denotations. Since at least Link (1983), most formal analyses of the semantics of number share a consensus of the basic ingredients of the count domain: (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.

The consensus breaks down, however, on differences concerning (i) the model-theoretic structure of the domain of plurals and (ii) whether the singular or plural is ‘unmarked’. Link (1983) models the plural as denoting the closure of atoms under join 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. Farkas and de Swart (to appear) argue that this proposal has the additional virtue of aligning with morphological markedness patterns, where the singular is usually considered unmarked and the plural is considered marked (Greenberg,

1966; Corbett, 2000). On the other side, Krifka (1989) (and later Sauerland (2003) and Sauerland et al. (2005)), favor a weaker plural motivated by inference patterns under negation and in question-answer pairs such as “Do you have children?-Yes, I have one.”, where an answer about *one or more* is required, thus modeling the plural more weakly as the entire semi-lattice structure. In this approach, the denotation of the plural includes that of the singular and straightforwardly captures the ‘inclusive plural’ reading of *one or more*. Under this analysis, the singular is more specific than the plural, and the plural surfaces as the unmarked number.

This paper contributes to this debate by examining the inverse number marking system of Dagaare (Gur; Niger-Congo), which proves problematic for both sides. Dagaare provides a number system which demonstrates that the cross-linguistic facts are more complicated than if only the singular *or* plural were unmarked, either morphologically or semantically, rather markedness is conditioned upon a nominal’s level of individuation. Section 2 investigates the number system of Dagaare in detail, demonstrating that the distribution of Dagaare’s inverse number marker *-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. Section 3 adduces support for this from cross-linguistic facts that surface in a wide array of language types, including English frequency patterns. The paper concludes by applying the logic of both strong and weak plural analyses which demonstrates that the strong plural analysis has better empirical traction in such systems.

2 The Semantic Basis of Number in Dagaare

Dagaare exhibits an initially surprising system for marking number. The basic paradigm is given for the Dagaare words ‘child’ and ‘seed’ in (1), showing a near minimal pair where both nouns share the same stem, yet the morpheme *-ri* marks the plural interpretation for ‘child’ and the singular interpretation for ‘seed’.

	Gloss	Singular	Plural	Stem
(1)	‘child’	bíé	bíírí	bi-
	‘seed’	bírí	bíê	bi-

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? Such systems are rare but attested at least in North America (Kiowa) and the Pacific (New Ireland) (see Corbett 2000).

Number marking in the nominal system of Dagaare is, from all appearances, not predictable from the phonological form of the stem. This is made clear by sets of minimal pairs, similar to the example in (1), shown in table (1). (Note that *-ri* assimilates before nasals and liquids and capital letters for vowels indicate +/- ATR). This is the

predominant pattern in the Dagaare nominal system, accounting for 60% of the nouns in the current database.

-V Singular	-rI/-nI Plural	Gloss	rI/-nI Singular	-V Plural	Gloss
tìé	tìirí	‘tree’	lúgrí	lúgó	‘prop, pillar’
gbìé	gbèrí	‘forehead’	nyágrí	nyágá	‘root’
pìé	pèrí	‘basket’	filí	filé	‘sores’
nàṅá	nànmí	‘scorpion’	ífilí	íflè	‘horn’

Table 1: Number in Dagaare: Marked Plural and Marked Singular Patterns

In previous work on the nominal system of Dagaare, two main approaches have been employed. The first approach establishes noun classes in Dagaare based upon a given noun’s behavior with respect to number marking. In particular, the line of research in Bodomo (1997) and Dakubu (2005), in coarse terms, presents one class as comprised of nouns ending in vowels in the singular and *-ri* in the plural, while another class is comprised of nouns ending in *-ri* in the singular and vowels in the plural. In contrast, Anttila and Bodomo (2009) give a set of phonological generalizations about the morphophonology of Dagaare which are relevant for plural formation. While both approaches have increased the understanding of the organization of Dagaare, neither provides a predictive answer to why certain nouns have *-ri* in the singular while others have *-ri* in the plural. This section explores the hypothesis that general semantic principles play the organizing force in the nominal system of Dagaare, namely those associated with individuation, a principle stating that speakers discriminate between entities which are distinct and countable and those which are non-distinct and uncountable. While this principle has most often been invoked for the syntactic distinctions concerning the differential syntax of count and mass terms, I will show that such a principle also makes predictions in the count domain, and is predictive of the data observed in Dagaare.

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 1994; Wisniewski, Imai, & Casey, 1996).

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 Allan (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 that the noun expresses a singular/plural contrast.

Two other factors relating to individuation were proposed by Wierzbicka (1985), 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 a 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 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 (1985) 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 with one of two phrases 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.¹ 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.

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. 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

¹This distribution is significantly above chance ($p < .05$.)

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, propensity to appear in a group of more than one is also distinct from the factors of distinguishability and interaction: Middleton et al. (2004) showed effects when examining only aggregates all the items were already assumed to be prone to coming in groups, in other words, the inherent plurality of the aggregates was held constant. Thus, distinguishability and interaction are relevant for aggregates in addition to and distinctly from a given entity's propensity to occur in an aggregate or group.

2.2 Individuation and Inverse Marking

The individuation factors discussed immediately above have been argued to independently affect the realization of number marking, primarily in English. It is not unreasonable to suppose that their influence would extend to inverse number marking in Dagaare. This gives rise to a clear prediction: 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 (2):

- (2) [Highly Individuated N] + *-ri* \Rightarrow plural
 [Less Individuated/Inherently Plural N] + *-ri* \Rightarrow 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)
- (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 as singular items 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.

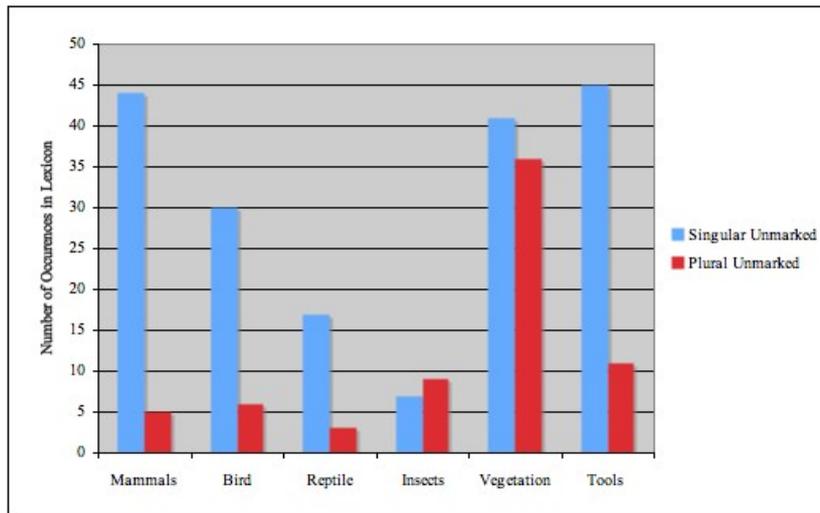


Figure 1: Number Marking Across Semantic Domains

2.3 Results from Fieldwork

To test the predictions elaborated in the preceding section, I conducted fieldwork in Ghana with native speakers to develop 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 research².

As the hypotheses involved distributions over semantic domains, I coded each word for (relatively transparent) semantic domains, where possible. The chart in figure 2.3 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 blue-shaded regions show the number of lexicon entries in a given semantic domain with the singular unmarked, while the red-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.

Figure 2.3 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, nouns for insects generally 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.

²I would like to acknowledge Arto Antilla and Adams Bodomo for generously permitting me to incorporate elements from their wordlist, Anttila and Bodomo (2006)

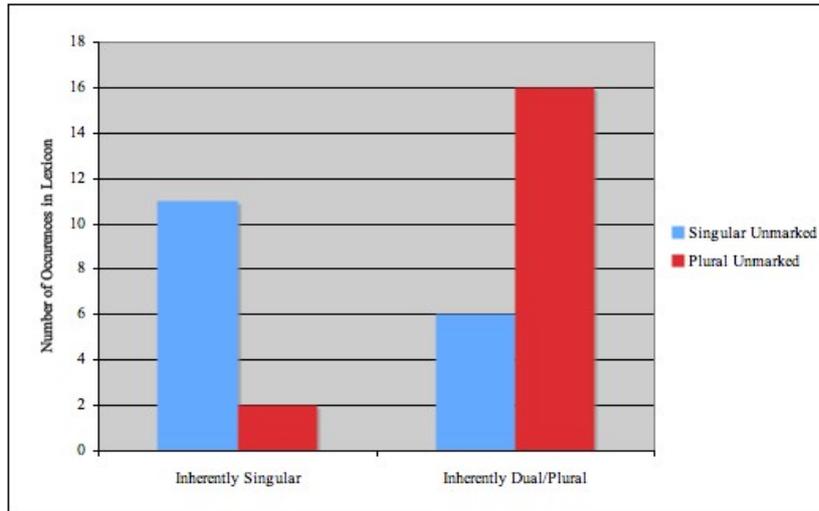


Figure 2: Number Marking and Inherent Plurality in the Domain of Body Parts

Figure 2.3 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.

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 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, often the marking results from practices elsewhere in the grammar, viz. derivational morphology. Frequently, 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* in the plural, exceptions included nouns such as *fúmìní* (sg) / *fúmìné* (pl) ‘needle’, which are clear candidates for canonically appearing in collections, or not participating in the initial assumption that one interacts with them individually. Additional apparent counter-examples result from semantic shift in the history of the lexical item. One instance

is *yírì* (sg.) / *yíè* (pl.), which synchronically designates ‘house’. While this would be an apparent example of an individuated entity, the word has antecedent collective uses meaning ‘compound’ (Durand 1953) as well as ‘family’ or ‘family members’ (Mark Ali, p.c.) aligning more closely with the notion of inherently plural/collective entities.

Viewing individuation as an organizing force in the choice of nominal inflection in Dagaare makes further predictions for dialect variation. 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. Bodomo (2004) notes that there are instances where the direction of number marking differs among dialects, as shown for the noun stem *pi-* ‘rock’:

Singular	Plural	Gloss
pìrì	pìé	‘rock’ (Central Dialect)
pìé	pìrì	‘rock’ (Southern Dialect)

Table 2: Dialect Variation: Variation in Directionality of Marking

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 clearly mass terms as well as aggregates with are particularly close-knit, as shown in table 2.4.

Singular/Base	Gloss	Singulative	Gloss
múó	‘grass’	móónúó	‘blade of grass’
súnnì	‘gum’	súnnúú	‘piece of gum’

Table 3: Singulative Paradigm

Several words in Dagaare, as shown in table in 2.4, vary 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.

Variant	Singular	Plural	Gloss
I:	váálí	váálá	‘rubbish’
II:	váálúú	váálá	‘rubbish’
I:	kómmírí	kómmié	‘tomato’
II:	kómmírúú	kómmié	‘tomato’

Table 4: Dialect Variation: Singular vs. Singulative

3 Cross-Linguistic Correlates

Additional support for the assumptions underlying the hypothesis that individuation underlies the organization of Dagaare’s nominal number marking system can be adduced from cross-linguistic data. Section 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. Although necessarily cursory due to space, this section will point to three cross-linguistic correlates of the unmarked plural in Dagaare in three domains: structuring nominal paradigms, morphological processes and text frequency. Despite different encodings languages may make, all these different systems seem to make similar divisions on a scale of individuation.

Nominal Paradigms Semantic domains similar to those discussed in section 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 marked; however, for select semantic domains, a collective interpretation is unmarked, while a singular interpretation is marked by a singulative marker *-yn* or *-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. 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.

Morphological Processes The semantic domains discussed in section 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 dou-

ble plural formation. For instance, morphological paradigms typically level towards the unmarked members of the paradigm, i.e. normally the singular stem, yet in certain cases nominal paradigms level in favor of the plural stem, which Tiersma (1982) exemplifies with Modern Frisian.

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 3, discussed in Stolz (2001), where the borrowed plural form from English serves as the basic term which can then be inflected for the singulative.

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

Table 5: Borrowing in Welsh

Tiersma infers from such examples 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.” This characterization aligns well with the distributional patterns established for Dagaare and converges with the individuation hypothesis.

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 inasmuch 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 figure 3, where the x-axis represents the number of lexical items and the y-axis represents the ratio of the token frequency of plurals to the token frequency of singulars. 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 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”.

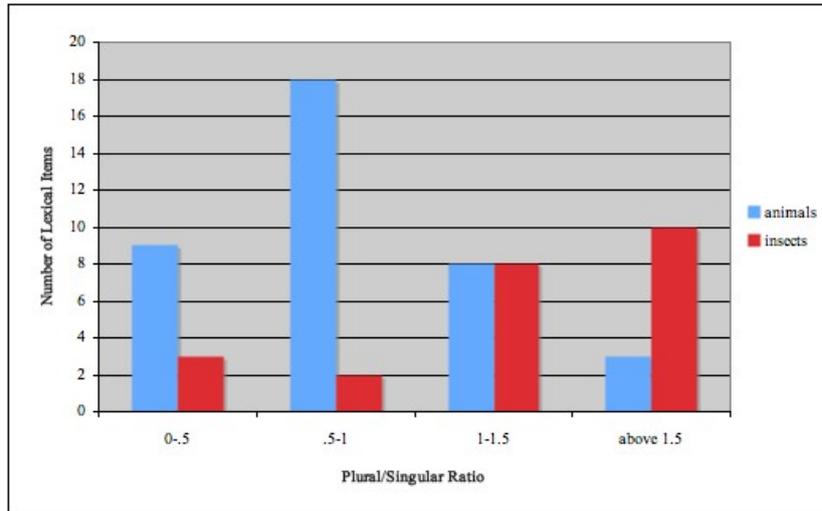


Figure 3: Number Marking Frequency Patterns in English for *Animals* and *Insects*

4 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 mass terms. Having secured an understanding of the Dagaare's nominal system, the number marking system can be related to formal models of number and be brought to bear on the controversy among the different analyses of the plural discussed in section 1.

Recall that the exclusive plural analysis, as in Link (1983), models the plural as denoting the closure of atoms under join (\oplus) less the atoms themselves (*sums* – *atoms*), thus the denotation of the plural excludes that of the singular, while the inclusive plural analysis models the plural as denoting both atoms and their closure under join (*sums* \cup *atoms*). In order to evaluate these proposals 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 analyses to data from Dagaare shows the exclusive plural analysis makes the better predictions for 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, the semantics of inverse number marking is relatively straightforward: *-ri* is simply treated as a form of negation of the lexical denotation of the base. 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)³. 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) with respect 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 4, 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.

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

Table 6: Derivations of Lexically Singular (‘child’) and Plural (‘seed’) Nouns with the Exclusive Plural

The Inclusive Plural An alternate analysis⁴, which is consistent with weak plural analyses, models $-ri$ as designating the “completion of the space”, viz. $-ri$ is the operation of closure under join and meet. The weak plural analysis of English plurals claims that the plural is unmarked, denoting closure under sum, while the singular, designating atoms is more specific. When the singular form is used, the plural interpretation is excluded by pragmatic blocking. The same inferences motivating the weak plural analysis in English were elicited in Dagaare, thus one could analyze $-ri$ when marking the plural similarly to the English plural, designating closure under join, with the singular interpretation disallowed by blocking. By parity, and to give $-ri$ a uniform interpretation, for lexically plural nouns where $-ri$ marks the singular, it must also yield the entire semi-lattice, viz. closure under meet, with the plural interpretation disallowed by blocking. Thus, $-ri$ is uniformly analyzed as the closure of the space under join and

³This line has also been developed independently in Bach (2007) and Bach (2008) for Kiowa.

⁴I am indebted to Uli Sauerland for suggesting this line of analysis.

meet. Representative derivations are given in table 4, where Cl represents a 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.

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

Table 7: Derivations of Lexically Singular (‘child’) and Plural (‘seed’) Nouns with the Inclusive Plural

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. This proposal predicts that the form marked by $-ri$ should always be the one excluding the truth of both singular and plural. Example (3) (Adams Bodomo p.c.) shows that this turns out not to be the case:

- (3) N bá dà dà biè/*birí (zàà)
 1st.pro NEG buy Past seed.PL/seed.SG (any)
 I didn’t buy (any) seeds.

The same entailment patterns hold in Dagaare as in English and in (3) 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.

5 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 widespread markedness patterns, namely less individuated/inherently plural entities are unmarked in the plural. Once the lexical generalizations are clearly established, the formal implementation is straightforward.

Number marking systems such as that of Dagaare make it evident that the marking of singular or plural comprises more than simply marking a dichotomy between ref-

erence to atoms or reference to sums, rather the canonical properties of the referents themselves influence how the marking is achieved. In parallel, the data examined here make it clear that the notion of markedness must be relativized with respect nominal semantics—considering either the singular or the plural as ‘unmarked’ across-the-board leaves much unexplained, both in Dagaare and cross-linguistically as discussed in section 3. Articulating the precise connections between the lexical properties of nouns and number realization remains a fertile area for further exploration.

Acknowledgements

This paper would not have been possible without generous efforts of the Dagaare scholars Mark Ali, Adams Bodomo, and J.A. Saanchi. In particular, Dr. Ali devoted much time to discussing Dagaare vocabulary with me and Dr. Bodomo has provided encouragement and aid at every step of the project. I would also like to thank the following for discussion and comments on various drafts: Matt Adams, Arto Anttila, Eve Clark, Jeff Good, Paul Kiparsky, Beth Levin, Uli Sauerland and Tom Wasow. The field research on Dagaare was made possible by the generous support of the Stanford Center of African Studies. The usual disclaimers apply.

References

- Acquaviva, P. (2008) *Lexical Plurals: A Morphosemantic Approach*. Oxford University Press, Oxford.
- Allan, K. “Nouns and Countability”. *Language*, 56(3):41–67, 1980.
- Anttila, A. and A. Bodomo (2009) “Prosodic Morphology in Dagaare”, in F. McLaughlin, E. Potsdam and M. Matondo, eds., *Selected Proceedings of the 38th Annual Conference on African Linguistics*, Cascadilla Proceedings Project, Somerville, MA, 56–68.
- Anttila, A. and A. Bodomo (2006) “A Dagaare-English Word List”, Unpublished Ms., Stanford University and University of Hong Kong.
- Bach, E. (2007) “Parochial and Universal Semantics: Semantic Typology and Little Studied Languages”, Handout for pre-ALT Workshop on Language Documentation and Linguistic Typology, Paris, France, September 25, 2007.
- Bach, E. (2008) “On Morphosemantics: The Internal Meanings of Words”, Paper presented at the Copenhagen Linguistic Circle, October 21, 2008.
- Bloom, P. (1994) “Possible Names: The Role of Syntax-Semantics Mappings in the Acquisition of Nominals”, *Lingua*, 92, 297–329.
- Bodomo, A. (1997) *The Structure of Dagaare*, Stanford Monographs in African Languages, CSLI, Stanford, CA.

- Bodomo, A. (2000). *Dagaare*. Languages of the World Materials No. 165. Lincom Europa, Munchen.
- Bodomo, A. (2004) *A Dagaare-Cantonese-English Lexicon for Lexicographical Field Research Training*, Rüdiger Köppe Verlag, Cologne.
- Chierchia, G. (1998) "Plurality of Mass Nouns and the Notion of 'Semantic Parameter'" in S. Rothstein, ed., *Events and Grammar*, Kluwer, Dordrecht, 53–103.
- Contini-Morava, E. (2000) "Noun Class as Number in Swahili" in Ellen Contini-Morava and Yishai Tobin (eds.) *Between Grammar and Lexicon*, John Benjamins, Amsterdam/Philadelphia.
- Corbett, G. (1996) "Minor Number and the Plurality Split". *Rivista di Linguistica* 8, 101–122.
- Corbett, G. (2000) *Number*, Cambridge University Press, Cambridge.
- Durand, J. (1953) *Dagaare-English Dictionary*, St. John Bosco's Press, Navrongo, Ghana.
- Farkas, D. and H. de Swart. (to appear) "The Semantics and Pragmatics of Plurals", *Semantics and Pragmatics*.
- Greenberg, J. (1966) *Language Universals*. Mouton, The Hague.
- Krifka, M. (1989) "Nominal Reference, Temporal Constitution and Quantification in Event Semantics", in R. Bartsch, J. van Benthem, P. von Emde Boas, eds., *Semantics and Contextual Expression*, Foris Publication, Dordrecht, 75–115.
- Link, G. (1983) "The Logical Analysis of Plurals and Mass Terms: A Lattice-Theoretical Approach", in R. Bauerle, C. Schwarze, and A. Stechow, eds., *Meaning, Use, and Interpretation of Language*, de Gruyter, Berlin, 302–323.
- Middleton, E., E. Wisniewski, K. Trindel, and M. Imai (2004) "Separating the Chafe from the Oats: Evidence for a Conceptual Distinction Between Count Noun and Mass Noun Aggregates", *Journal of Memory and Language*, 50, 371–394.
- Mufwene, S. (1980) "Number, Countability and Markedness in Lingala *li-/ma-* Noun Class", *Linguistics*, 18, 1019–1052.
- Ojeda, A. (1998) "The Semantics of Collectives and Distributives in Papago", *Natural Language Semantics*, 6, 245–270.
- Quine, W. (1960) *Word and Object*, MIT Press, Cambridge, MA.
- Sauerland, U. "A New Semantics for Number", in R. Young and Y. Zhou (eds.), *Proceedings of SALT 13*, CLC Publications, Cornell University, Ithaca, NY.
- Sauerland, U., J. Anderssen, and K. Yatsushiro (2005) "The Plural Is Semantically Unmarked", in S. Kepser and M. Reis, eds., *Linguistic Evidence: Empirical, Theoretical, and Computational Perspectives*, Mouton de Gruyter, Berlin, 413–434.

- Smith-Stark, T. (1974) "The Plurality Split", in M. La Galy, R. Fox, and A. Bruck, eds., *Papers from the 10th Regional Meeting*, Chicago Linguistics Society, Chicago, 657–671.
- Stolz, T. (2001) "Singulative-Collective: Natural Morphology and Stable Cases in Welsh Number Inflexion of Nouns", *Sprachtypologie und Universalienforschung*, 54, 52–76.
- Tiersma, P. (1982) "Local and General Markedness". *Language*, 58, 832–849.
- Wierzbicka, A. (1985) "Oats and Wheat: The Fallacy of Arbitrariness", in J. Haiman, ed., *Iconicity in Syntax*, John Benjamins, Amsterdam, 311–342.
- Wisniewski, E., M. Imai, and L. Casey (1996) "On the Equivalence of Superordinate Concepts", *Cognition*, 60, 269–298.
- Wonderly, W. (1954) "Number in Kiowa: Nouns, Demonstratives and Adjectives", *International Journal of American Linguistics*, 20, 1–7.