

Inverse Number in Dagaare*

Scott Grimm
University of Rochester
scott.grimm@rochester.edu

Abstract

This chapter examines the inverse number system in Dagaare (Gur; Niger-Congo). Inverse number systems possess a number morpheme which for some nouns encodes the plural interpretation while for others it encodes the singular interpretation. This chapter argues that a principled lexical semantic classification underlies the inverse number strategy in Dagaare, guiding whether for a particular noun the inverse morpheme codes the singular or the plural interpretation. The chapter further explores the functional grounding of inverse number, in terms of frequency and individuation, and presents a formal semantic account of the inverse number system.

*I would like to thank the editors and an anonymous reviewer for helpful comments on a draft of this chapter. I would also like to thank Adams Bodomo for his early and continued encouragement of my work on Dagaare, as well as Arto Anttila for his early, generous help. A special thanks goes to Mark Ali at the University of Education in Winneba, Ghana who devoted an enormous amount of time to working with me on Dagaare and without whom this research would have never come to fruition. I would also like to thank the many others who provided helpful input on this work, particularly Jeff Good, Tom Güldemann, Paul Kiparksy, Beth Levin, Frank Seifart, and Tom Wasow; as well as the various audiences at venues where this work has been presented. This material is in part based upon work supported by the National Science Foundation under Grant No. 1023289 ‘Number and Individuation: Nominal Semantics in Dagaare’. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

1 Introduction

Dagaare, a Gur language spoken in Northern Ghana, manifests an “inverse” or “polarity” system for marking number.¹ The example in (1) shows the inverse pattern with the Dagaare words ‘child’ and ‘seed’. While the word forms constitute a near minimal pair where both nouns share similar stems, the morpheme *-ri* marks the plural interpretation for ‘child’ but the singular interpretation for ‘seed’.

	Gloss	Singular	Plural	Stem
(1)	‘child’	bíé	bíí-rí	bì-
	‘seed’	bí-rì	bíè	bí-

While *-ri* is a suffix and can be attached to both plural and singular stems, the final vowel in the contrasting singular or plural forms, in this example *-e*, is epenthetic, due to a requirement that nouns be bimoraic, and is determined by the quality of the stem vowel (Anttila and Bodomo, 2009, 59). Dagaare possesses two levels of tone, high and low, although the morpheme *-ri* itself is, however, unmarked for tone. A further phonological process of high vowel lengthening occurs in the plural form of ‘child’ *bíírí*, which appears to be lexically conditioned. See Anttila and Bodomo (2009) for further discussion of the phonology of nominals in Dagaare.

To better illustrate the pattern and abstract away from the phonological processes in Dagaare, a parallel example is given for a mock version of English in (2), for which *-s* is an inverse number marker in English:

	Gloss	Singular	Plural
(2)	‘child’	child	child-s
	‘seed’	seed-s	seed

This pattern is strikingly different from grammatical number systems as found in English, Spanish and Italian. In nominal systems of languages such as English, countable nouns in their base form² encode a singular value while the plural value is explicitly encoded with a plural morpheme, as in *dog/dogs*. I will refer to a noun’s BASIC NUMBER VALUE, or BASIC NUMBER, as that which is present in its base form. For instance, English *dog* has a basic number value of singular, or is ‘basic singular’.

¹The Dagaare data comes from the Central Dialect in and around the town of Jirapa, Ghana. The bulk of the data in this section were gathered over the course of 7 months of fieldwork conducted by the author in 2008 and 2011. For more general information on Dagaare grammar, see Bodomo (1997) and Bodomo (2000).

²I follow Corbett (2000, 139) in using the term ‘base form’ to designate that which can no longer be reduced of inflectional material.

Inverse systems are remarkable in that there is no single basic number value affiliated with nominal base forms: Some nominal base forms are basic singular, as in English, while others are basic plural. This trait of inverse number systems is shared with languages that feature singulative morphology — languages which possess a singulative morpheme encoding a singular value in addition to a plural morpheme encoding a plural value, such as occurs in languages of the Celtic or Nilo-Saharan language groups. (See chapter XX for discussion of singulatives.) In parallel to the Dagaare example (1), for systems with both plural and singulative morphemes, the basic number value varies according to the noun, with the non-basic value encoded by the plural or singulative morpheme.

The distinguishing characteristic of inverse systems is that the *same* morpheme is responsible for encoding both singular and plural values. In Dagaare, this morpheme is *-ri*. Thus, pairs of words which are “inverses” are mirror images of each other in terms of their basic number.

It is important to separate out true inverse number systems from those systems with spuriously homophonous endings in plural and singular — for the mere homophony of word endings involved in encoding singular and plural values is not sufficient to establish an inverse system.³ In the German examples in (3), the singular form of the noun *Affe* ‘monkey’ ends with *e*, which is identical with the plural ending *-e* of the noun *Hund* ‘dog’.

	Gloss	Singular	Plural
(3)	‘monkey’	Affe	Affe-n
	‘dog’	Hund	Hund-e

While an ending in *-e* is shared by the singular *Affe* and the plural *Hunde*, this occurs with the base form for *Affe* and the non-base form for *Hund*, and only for *Hund* does *-e* constitute an inflection. The inflectional number morphemes across these nouns, *-n* and *-e*, both appear on the plural form and, furthermore, are not the same, and thus this does not constitute an inverse number system.

Other inverse number systems are found in Kiowa (Tanoan family, spoken in Oklahoma) and Nehan (Western Oceanic, spoken in Papua New Guinea). For discussion of Kiowa, see Watkins and Harbour (2008, 2011). Nehan is discussed both in Corbett (2000) and Baerman (2007), which also discuss number marking reversals generally.

Inverse number systems provide a rather unique vantage point from which to examine some of the central questions about how grammatical number systems are structured. Corbett (2000, 161) succinctly summarizes a frequent intuition in the literature that an inverse number morpheme encodes the “less expected” form.

³Thanks to Patricia Cabredo Hofherr for suggesting discussion of this point.

Understanding in more precise terms what exactly is meant by “more expected” or “less expected” for a base noun form and how that relates to number marking is however a significant challenge, as it is in other domains.⁴

The question of what underlies inverse number systems can in turn be related to other, more well-studied questions in the grammatical number literature. A core question in the count/mass literature is to determine what conditions cause a base form to be countable or non-countable in a given language. A variant of this question arises for inverse number systems, as well as singulative systems: What conditions cause a base form to be basic singular or basic plural? Grimm (2012b, To Appear) argues that both questions find a common, general explanation through understanding how individuation plays out in nominal semantics — although it must be emphasized that this is a very open and active area of research.

A different set of questions arise from considering the formal analysis of inverse number morphology, or inverse morphology more generally. To what extent is the semantics of inverse number morphemes consistent with what is generally assumed for number morphemes, e.g. the plural? To what extent can possible analyses of *-ri* or other inverse number morphemes bear on controversies in the understanding of plurality?

This chapter first presents Dagaare’s nominal system and inverse number paradigm in section 2. Section 3 demonstrates that whether a base form is basic singular or basic plural can be related to lexical semantic classification. Section 4 investigates the functional grounding of the lexical semantic classification, considering two hypotheses, frequency and individuation, along with the relation between the two. Finally, section 5 provides a formal semantic account of *-ri*. Section 6 concludes and discusses directions for further research.

2 The Nominal System of Dagaare

In this section, I lay out the basics of Dagaare’s nominal system. I first present a more substantial set of nouns exemplifying the inverse number paradigm. I then show that the singular and plural forms correspond to singular and plural interpretations through examining behavior with demonstratives and cardinal numerals. Finally, I discuss two other number morphemes, a distributive plural and a singulative, present in Dagaare.

Table 1 and Table 2 present other nouns belonging to the inverse paradigm already given in (1). Nouns with *-ri* in the plural are given in Table 1 and

⁴See Haspelmath (2006) for an insightful discussion of similar issues related to the understanding of “markedness”, noting that the term “unmarked” has had a range of interpretations, one of which is “more frequent”.

nouns with *-ri* in the singular are given in Table 2. Some further remarks on the phonology of Dagaare are in order.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 /ɔ/ as well as for /e/ and /ɛ/, but I have followed Bodomo’s (1997) 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. I use capital letters for vowels to indicate +/- ATR value. Furthermore, *-ri* assimilates following nasals and liquids, resulting in the allomorphs *-nɪ* and *-(l)ɪ*, respectively, in the following examples.

-V Singular	-rɪ Plural	Gloss
tɪ́é	tìrɪ́	‘tree’
gbìé	gbèrɪ́	‘forehead’
pìé	pèrɪ́	‘basket’
nàŋá	nànní	‘scorpion’

Table 1: Number in Dagaare: Marked Plural Pattern

rɪ Singular	-V Plural	Gloss
lúgrɪ́	lúgó	‘prop, pillar’
nyágrɪ́	nyágá	‘root’
fɪ́lɪ́	fɪ́lé	‘sore’
íflɪ́	íflè	‘horn’

Table 2: Number in Dagaare: Marked Singular Patterns

The forms in (1) and Tables 1 and 2 can be shown to correspond to singular and plural interpretations by examining their interaction with other elements of Dagaare that express number. First, demonstrative pronouns differentiate singular and plural referents in Dagaare. Accordingly, the base noun forms that are basic singular and those that are basic plural values manifest distinct number agreement with the singular (*ŋá*) and plural (*-má*) forms of demonstrative pronouns. This is

⁵Again, see Anttila and Bodomo (2009) for detailed morphophonological analyses of the Dagaare nominal system.

demonstrated in examples (4) and (5) where the different noun forms from (1) are shown in combination with the different forms of the demonstrative pronoun. (In the plural, the agreement prefix further distinguishes between human (*ba-*) and non-human (*a-*) referents.)

- (4) a. bíé ηâ
 child.SG DEM.PROX.SG
 ‘this child’
 b. bí-rí bà-má
 child-PL HUM.PL-DEM.PROX.PL
 ‘these children’ (HUM = Human Prefix)
- (5) a. bí-rì ηâ
 seed-SG DEM.PROX.SG
 ‘this seed’
 b. bíè à-má
 seed.PL NHUM.PL-DEM.PROX.PL
 ‘these seeds’ (NHUM = Non-Human Prefix)

Similarly, 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, in particular an adjective, and directly attaches to the noun stem.⁶ 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 (6)-(7).

- (6) a. bì-yénì (bì- + yénì)
 child-one
 ‘one child’
 b. bí-rí bá-yì
 child-PL HUM.PL-two
 ‘two children’
- (7) a. bí-yénì (bí- + yénì)
 seed-one
 ‘one seed’
 b. bíè à-yí
 seed.PL NHUM.PL-two

⁶In Dagaare, adjectives compound with the noun stem and then the adjective supplies its own singular/plural suffix patterns.

‘two seeds’

Two other morphemes bear on nominal number in Dagaare. First, both countable and non-countable nouns in Dagaare may optionally combine with a distributive plural suffix *-ree*, which designates “different types of X” or “X in different locations”. Examples of nouns with the distributive plural are given in Table 3.⁷

Noun	Dist. Pl.	Gloss
kùó	kònnéé	‘water; types of waters; water in different locations’
múó	múónéé	‘grass; types of grasses; grass in different locations’

Table 3: Distributive Plural in Dagaare

Second, Dagaare possesses a distinct singulative morpheme *-ruu* designating ‘a piece of’. This morpheme also has an ATR allomorph, *-rvv*, and assimilates before nasals (*-nuu/-nvv*) and liquids (*-luu/lvv*). This morpheme applies to fewer than 100 nouns, primarily granular aggregate nouns or foodstuff nouns (*sérímáání* ‘pepper’, *kpéé* ‘malt’, *múó* ‘grass’, *kámáání* ‘corn’). There also appears to be a use of *-ruu* among some speakers to indicate a packaging reading, i.e. where *kpéérúú* designates one sack of malt rather than a piece.

Altogether, Dagaare’s number morphology isolates four classes of nouns, laid out in Table 4. I now turn to examining the differing lexical semantic classes corresponding to nouns with basic singular or basic plural values.

3 The Lexical Semantic Basis for the Inverse Pattern in Dagaare

I now turn to considering what grammatical organizing principle accounts for the inverse pattern in Dagaare, developing an argument that the inverse pattern

⁷The distributive plural, however, does not “coerce” non-countable nouns to countable nouns, as nouns with the distributive plural are unable to combine with cardinal numbers, as shown in (i).

- (i) *kònnéé à-yí
water.DISTPL NHUM.PL-two
‘two waters’

Grammatical Category	Gloss	Singular	Plural	Singulative	Dist. Pl.
Basic Singular	‘child’	bíé	bíírí	—	bíírée
Basic Plural	‘seed’	bírì	bíè	—	bíréé
Non-Countable/Singulative	‘grass’		mòó	móóróó	móónéé
Non-Countable	‘water’		kòó	—	kònméé

Table 4: Grammatical Number Categories in Dagaare

reflects lexical semantic contrasts among nominals. I first set aside the idea that the pattern could be derived solely from the morphophonological properties of Dagaare’s nominal system. Table 5 provides two sets of nouns which, in terms of their morphophonology, are minimal pairs and yet nonetheless form inverse pairs. As such, these sets of minimal pairs provide counter-examples to any purely morphophonological explanation.

Table 5: Near Minimal Pairs

Gloss	Stem	Singular	Plural
‘wild rat’	kù-	kúó	kúúrí
‘hoe’	kù-	kùùrí	kùé
‘granary’	bùg-	bùgó	bùgrí
‘pillar’	lùg-	lùgrí	lùgó

Within the Gur linguistic tradition, Dagaare’s inverse pattern has been understood within the context of noun classes, by analogy to a large number of other nominal systems in Niger-Congo languages with a rich inventory of noun classes. Bodomo (1997), Dakubu (2005) and Bodomo and Marfo (2006) use the different singular-plural pairings of nouns to construct noun classes for Dagaare.⁸ The basis for the inverse pattern in these analyses mostly reduces 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.

⁸This 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 (4)-(5): singular, human plural and non-human plural.

It is very likely that Dagaare’s inverse system historically derives from a noun class system. Dagaare stems from Proto-Gur, which possessed a highly developed noun class system (Miehe and Winkelmann, 2007). Yet, in contrast to many other Gur languages — such as Gurene (Nsoh, 2002) — that have preserved more of the Proto-Gur system, Dagaare no longer has a robust noun class system. Instead, the inverse pattern of (1) has become the dominant pattern in the nominal system, accounting for over 70% of countable nouns, i.e. nouns with singular and plural forms (as estimated in Grimm 2012a from a lexicon (\sim 1500 words) compiled in the field).⁹ The only other remaining major singular-plural pairings are (i) for nouns designating humans with a \emptyset /*-ba* contrast, as in *dɔ́ɔ́/dɔ́bá* ‘man’/‘men’ and (ii) for liquids and other typically uncountable nouns, which lack a (standard) plural inflection, as was seen for *kùɔ́*.

Acknowledging that the Dagaare inverse pattern relates to a noun class system still leaves open the question of what organizing principle, if any, underlies the pattern. Grimm (2012a) hypothesized that the coding asymmetry of the inverse pattern reflects a semantic asymmetries among different nouns. For nouns designating entities more likely to appear as a single unit, it is more likely that the base form has a singular value and *-ri* codes the plural. Conversely, the more likely the entity is to be viewed as coming in groups or “non-individuated”, the more likely that the base form has a plural value and *-ri* codes the singular.

Grimm (2012a) supported this hypothesis showing an effect of different semantic domains on the asymmetrical coding of the inverse number pattern, shown through counts in a field lexicon. In particular, different semantic domains were hypothesized to contrast in their likelihood of canonically referring to single referents or multiple referents, e.g. large mammals (single) vs. insects (multiple) or trees (single) vs. vegetation (multiple). Such asymmetries were found as shown in Figure 1, which gives counts

⁹To be sure, there remain prominent morphophonological asymmetries which would appear to follow from the form of the stem. Yet, these asymmetries can primarily be linked to either patterns of derivational morphology or to the same semantic contrast that is being tracked in this paper, but just manifested in a, presumably historically prior, noun class contrast. For an instance of the first sort of example, CVC stems ending in η nearly always have *-ri*, in its assimilated form, in the plural, as in *bij/binni* ‘pen’/‘pens’ (for animals). Yet, a stem-final *-η* most often signals a nominalizing locative suffix, and thus the asymmetry for these stems is most likely due to regularities of derivational processes. A different sort of example comes from *m*-final CVC stems which nearly always have *-ri* in the singular, as in *pání/pámá* ‘doorway cover made of straw’. These nouns do appear to be related to a noun class pairing, namely the *dí/-a* (5/6) pairing present in Gur languages (Manessy, 1975), and the 5/6 class more broadly in Niger-Congo. Yet this particular noun class pairing is often noted, as in Mufwene (1980), to be interesting in that the plural form appears to be more basic than the singular form, and accordingly dovetails with the same semantic contrast investigated in this section. I exclude from discussion other minor classes, such as *támmú/támá* ‘bow’, since they constitute only a very small number of nouns, on the same level as *-en* pluralization in English (*oxen*).

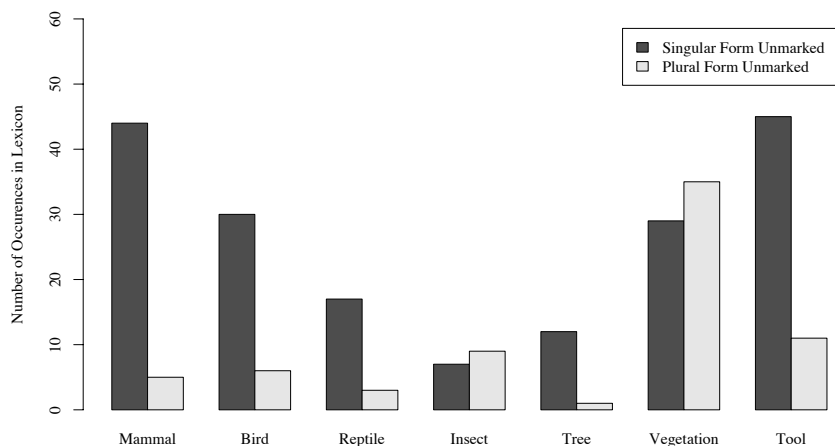


Figure 1: Number Marking Across Semantic Domains

in the lexicon for the domains of animals, insects, trees, vegetation, and tools. The dark-shaded bars indicate basic singular nouns, while the light-shaded bars indicate basic plural nouns, e.g. for *mammal*, 43 entries are basic singular and 5 are basic plural.¹⁰

Broad trends can be observed in Figure 1. Nouns which designate higher-level animates (mammals, birds and reptiles) are basic singular; however, the majority of nouns for insects are basic plural. Similarly, nouns for trees are typically basic singular, while most nouns for vegetation are basic plural. By way of comparison, Figure 1 includes data for nouns designating tools, which also could be reasonably hypothesized as entities which typically occur singly. These nouns too showed a strong tendency towards being basic singular.

An asymmetry in the domain of body parts is displayed in Figure 2. Nouns designating body parts which inherently come in pairs or groups, e.g. *eye* or *rib*, are more likely to be basic plural; yet nouns designating body parts which are inherently singular, e.g. *head*, are more likely to be basic singular. In sum, the distribution of lexical items in these semantic domains supports the hypothesis that when a noun designates an entity which is likely to appear singly, *-ri* encodes the plural, while when a noun designates an entity which is likely to appear in multiples, *-ri* encodes the singular.

¹⁰These counts exclude derived forms since they follow their own patterns, as will be discussed below.

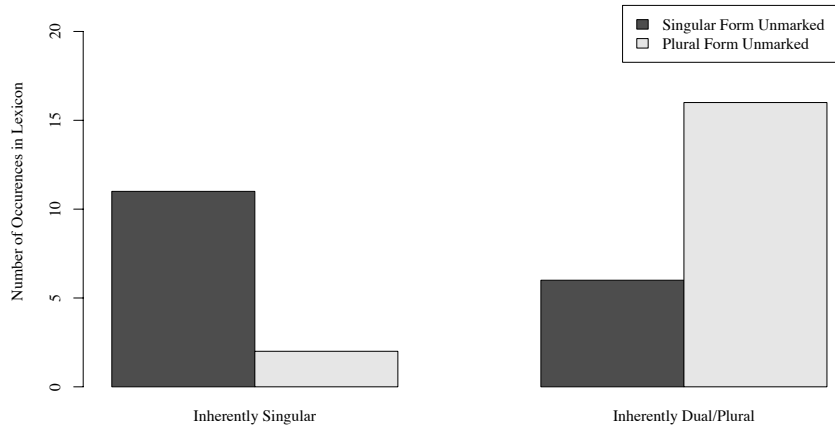


Figure 2: Number Marking and Inherent Plurality in the Domain of Body Parts

The nature of these asymmetries implicate that there is a lexical semantic categorization at work in the nominal system of Dagaare: Nouns describing entities typically occurring singly are basic singular while those describing entities typically occurring in groups (or more generally nonsingly) are basic plurals. As with most forms of lexical semantic categorization, there is a certain amount of conventionalization, historical residue and fuzzy boundaries in the Dagaare system, yet the broad tendency is clear.

Grimm (2012a) discusses further correlates of the hypothesis as well as various exceptional cases. One strain of additional evidence comes from observing dialectal variation in the distribution of *-ri*. Several words in Dagaare, as shown in Table 6, vary across dialects as to whether a noun encodes its singular form with *-ri* or the singulative morpheme *-ruu*. Since the singulative appears mainly with granular aggregate terms, this dialectal variation supports the hypothesis that *-ri* marks the singular for entities which are inherently plural.

Systematic exceptions to the pattern come from derived nominals. For instance, agentive nominalization is derived via a suffix *-aa*, which in turn comes specified for its own basic number value, namely singular. Thus, the number categorization for the word *bámbááláá* ‘small fish, fingerling’, which solely from its meaning might be expected to be a likely candidate for basic plural, is determined by the suffix, whereby the plural is formed with *-ri*, *bámbáálí*.

Table 6: Dialect Variation: Singular vs. Singulative

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

While the nominal system in Dagaare appears to manifest lexical semantic categorization amongst different entity types, the semantic distinction between basic singular and basic plural nouns does not appear to be productive any longer, as can be seen in borrowings into Dagaare. Number marking for words borrowed into Dagaare may be driven by phonetic similarity, as can be seen from the example *lǒrí* (sg) / *lǒ* (pl) ‘truck; lorry’, clearly a borrowing from the English word *lorry*. The Dagaare singular form (as determined by its agreement behavior) *lǒrí* corresponds to the (British) 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*, as in *lǒrí* (sg) / *lǒ* (pl) ‘bead’. A similar explanation can be found for the term *tírí* (sg) / *tí* (pl) ‘spoon’, borrowed from Akan (Mark Ali, p.c.).

I take it as established the inverse number system in Dagaare reflects principled lexical semantic categorization, although this is no longer active in the formation of new nouns. I now turn to considering what might underpin this categorization.

4 The Basis for “Expected” Number: Individuation and Frequency

The last section argued that the marking reversal in the number system of Dagaare correlates with a contrast among semantic classes of nouns, those which typically designate single referents and those which typically designate multiple referents. One can go a step further and question why such a contrast should be found among the different semantic domains. The broader literature on grammatical number offers two main responses: The contrast reduces to a contrast in usage frequency or, alternately, to a conceptual contrast among nouns. Much of the discussion of the two explanations are based on data from languages with singulatives, yet the predictions made for inverse number systems would be analogous.

The explanation from usage frequency is straightforward. Nouns differ as to

the frequency with which an entity is spoken of as referring to multiple referents as compared to the frequency with which it is spoken of as referring to a singular referent. On this line of explanation, the singular form of a noun will be basic due to its higher rate of occurrence referring to singular entities in comparison with the plural, and vice versa for nouns where the plural is basic, since it is more economical to encode the more frequent value as the basic value (see Haspelmath and Karjus 2017 and references therein). For instance, since, by hypothesis, one more often speaks of multiple ants rather than a single ant, it is more economical to use a word-form for which the plural interpretation is basic. Early work by Greenberg (1966) supported this view showing that zero-coded forms have greater textual frequency than overtly coded ones.

Grimm (2012a) shows that there is a plausible correlation between the coding asymmetries in Dagaare and frequency asymmetries. Since measuring frequency of occurrence in Dagaare is not currently feasible, Grimm (2012a) approaches this indirectly, reasoning that if semantic domains relevant to “basic plurals” in Dagaare are such that speakers more frequently refer to multiple referents than singular referents, then even in languages which do not possess morphologically unmarked plurals should still display asymmetries in the same semantic domains in terms of textual frequency. Grimm (2012a) examined frequencies for English nouns in the semantic domains of animal and insect from the COBUILD corpus (18 million words) provided by CELEX (Baayen et al., 1996). Figure 3 shows the plural-to-singular ratio for these two domains, where the vocabulary items selected were basic nouns, i.e. non-scientific terms like *bee*, and nouns which had correspondents in the vocabulary of Dagaare. The x-axis in Figure 3 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 terms tend to have a plural/singular ratio less than 1, i.e. animal terms occur more frequently in the singular. Similar evidence is provided by Baayen et al. (1997) who investigate grammatical number morphology and the phenomenon of “local markedness” (Tiersma, 1982) in Italian from a psycholinguistic perspective. More recent corpus evidence is provided by Haspelmath and Karjus (2017) who show that nouns which typically designate multiples, as for paired body parts, fruits, small animals, etc., have a higher rate of frequency in the plural compared with randomly selected nouns.

The second line of explanation for coding asymmetries invokes the conceptual distinction of INDIVIDUATION — a term used in cognitive science and particularly psycholinguistics to describe the propensity to designate an individual as opposed to, e.g., a collection of entities or an amorphous mass. Individuation has been long

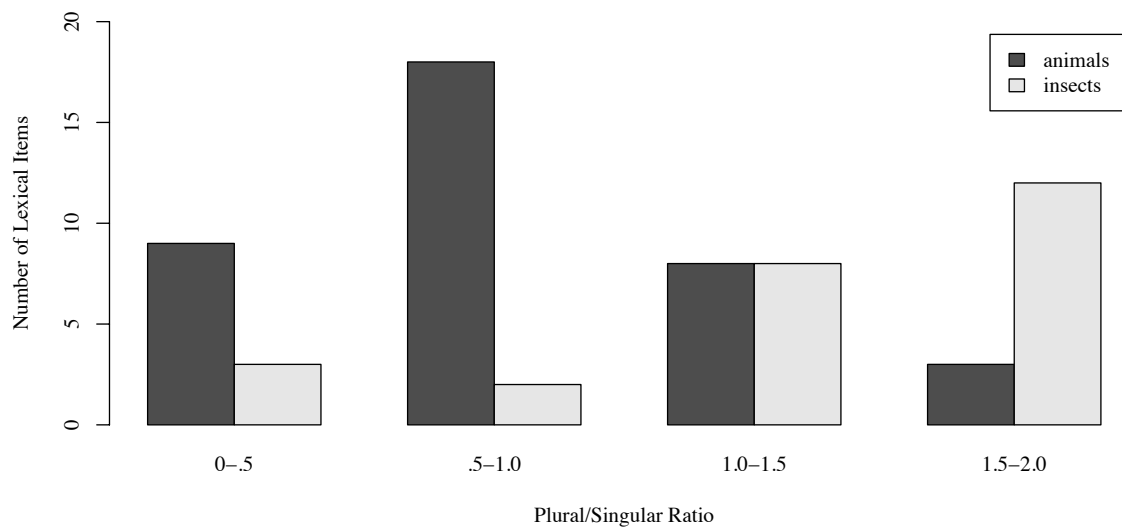


Figure 3: Number Marking Frequency Patterns in English for nouns in the semantic domains *animals* and *insects*

Noun Type	liquids/ substances	granular aggregates	collective aggregates	individual entities
Coding	0	0/Singulative (<i>-ruu</i>)	0/Singular (<i>-ri</i>)	0/Plural (<i>-ri</i>)
Example	kùó	mùó/mùóróó	bíri/bíè	bíé/bíírí
Gloss	‘water’	‘grass’	‘seed’	‘child’

Table 7: Noun Types in Dagaare Ordered by Degree of Individuation

invoked by a range of researchers in the literature on the count/mass distinction, who argue that the grammatical realization of count and mass nouns reflect the cognitive and/or perceptual properties, i.e. properties related to individuation, of the entities which the nouns describe. On one formulation, 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).

Although such a dichotomous view of individuation is discussed widely in the psycholinguistic literature, Grimm (2012b, To Appear) argues, based on the data from Dagaare and from singulative systems, that treating individuation as dichotomous is too coarse, instead proposing a scalar approach to individuation. Thus, nouns may possess different degrees of individuation, which may be aligned on a ‘scale of individuation’. Dagaare, for instance, morphologically recognizes four types of nouns, previously shown in section 2 in Table 4, which can be aligned with different degrees of individuation: in particular, and in increasing degrees of individuation, liquids and substances (no standard plural), granular aggregates (which combine with *-ruu*), collective aggregates (basic plural, *-ri* in singular), and individuals (basic singular, *-ri* in plural). This is shown in Table 7 where the morphological contrasts are given for each noun type along with an example.

Returning to the coding asymmetry of interest, an explanation based on individuation would claim that for nouns high in individuation — those designating individuals — the plural value is less expected and accordingly coded by *-ri*, while for nouns which are lower in individuation while still countable — those designating collective aggregates — the singular value is less expected and accordingly coded by *-ri*.

The explanations of frequency and individuation are not mutually exclusive, rather to all appearances they are most likely interrelated. There is limited evidence that the conceptual distinction may be the primary cause and frequency a secondary effect, although given the current state of knowledge this is an open question.

First, in discussions of languages with singulatives, researchers have noted

that frequency alone is not sufficient to account for the implicit classification of nouns by number morphology. Stolz (2001), in a in-depth analysis of the class of singulative-collective nouns in Welsh argues explicitly that whether a noun belongs to the singular-plural class or the singulative-collective class is not merely “a matter of quantity” (p. 65), i.e. frequency of occurrence. Rather, the behavior of the entities plays a role, and for animate entities a distinguishing characteristic of nouns falling into the collective/singulative class is their “characteristic way of living together in swarms, flocks, herds and shoals” (p. 65). Similar remarks are ventured in Dimmendaal’s (2000) analysis of the collective/singulative class in Turkana (Nilo-Saharan, spoken in Kenya).

A different form of evidence favoring individuation as more primary than frequency comes from artificial language learning experiments conducted in Kurumada and Grimm (2017), studying optional plural marking. In these experiments, participants learn to produce a miniature artificial language comprising of 12 novel nouns and one novel verb. During a word-learning phase, the novel nouns were designed to correspond to two classes: Those which learners usually (75% of the time) see coming as individuals or usually see coming as a collection. During a sentence learning phase, an optional plural marker followed the nouns which had plural reference 2/3 of the time, which corresponds to three times more frequently for collective nouns as for individual nouns. Despite the frequency bias of plural marking in comparison with collective nouns built into the experiment design, the production results show a significant interaction between the use of the plural marker and the noun type: Participants produced the optional plural marker *less* for the collective nouns than for the individual nouns, presumably since plurality would be more inferrable from the learned meaning of the collective nouns. This effect held regardless of the type of novel entities that participants were exposed to. In the first experiment, the ‘individual’ class of novel nouns consisted of novel mammal-like animals and the ‘collective’ class of novel insect-like entities, while in the second experiment, both classes consisted of geometrical shapes. Thus, from these experiments, it does not appear that speakers simply generalize from the more frequent form of use, but that how speakers think of the meaning of a noun may influence number coding patterns as well. It should be borne in mind that this experiment lends at best indirect evidence supporting individuation: Not only does it address optional plural marking, rather than an inverse or singulative system, but artificial language learning results require accepting a set of assumptions before inferring that they bear directly on a given natural language. Despite these limitations, such experimental work points to a fruitful direction to probe the causal factors underlying the shape that number systems may take.

5 Formal Analysis of the Inverse Number Morpheme

I now turn to the formal analysis of inverse number morphology in Dagaare. Two concerns guide the discussion of the formalization. The initial concern of course is securing the correct truth conditions. A further concern, however, is to provide an account which connects the analysis of *-ri* to what is known about the analysis of number markers in general, i.e. how does *-ri* relate to other accounts of the semantics of the plural or the singulative? This question also assumes importance from the vantage point of the diachronic development of an inverse number system. If the inverse number pattern can be traced to diachronic change by which a singular-plural pattern and a collective-singulative pattern are reanalyzed jointly as an inverse pattern, then it should be reflected in the analysis of *-ri*.

The findings in section 3 support analyzing countable nouns in Dagaare as having base forms with basic singular or basic plural denotations. I treat nouns as lexically specified for this information. The nouns in Dagaare which are basic singular can be analyzed in the same manner as countable nouns in English such as *dog*, denoting atomic entities. The nouns which are basic plural in contrast denote sum entities. I give representative lexical entries for basic singular and basic plural nouns in (8) and (9), respectively.¹¹

$$(8) \quad \llbracket bie \rrbracket := \lambda x[\text{Child}(x) \wedge x \in \text{ATOM}]$$

$$(9) \quad \llbracket biri \rrbracket := \lambda x[\text{Seed}(x) \wedge x \in \text{SUM}]$$

Grimm (2012a,b) analyzes *-ri* as a form of negation of the basic number value for the noun in question. This is an intuitive version of the function of inverse number, and is in essence a formal semantic update of the analysis of Kiowa in Wonderly (1954).¹² More technically, *-ri* is modeled as the operation of relative complementation (C), as given in (10), where the complementation operation is relativized to the domain. Thus, *-ri* applied to a basic singular noun will yield a plural denotation, while *-ri* applied to a basic plural noun will yield a singular denotation.

$$(10) \quad \llbracket -ri \rrbracket := \lambda P.P^C$$

Representative derivations are given in Table 8, demonstrating that this analysis secures the right truth conditions. In prose, for lexically singular nouns, the

¹¹These representations are simplified in that they do not speak to the collective nature of basic plural nouns. See Grimm (2012b) for fuller representations where the denotational space of basic plural nouns in Dagaare includes mereotopologically “clustered individuals” in addition to sum individuals, but not atomic individuals.

¹²This line has also been developed independently in Bach (2007, 2008) for Kiowa.

Basic Singular	Basic Plural
$\llbracket \text{bi-} \rrbracket := \lambda x[\text{Child}(x) \wedge x \in \text{ATOM}]$	$\llbracket \text{bi-} \rrbracket := \lambda x[\text{Seed}(x) \wedge x \in \text{SUM}]$
$\llbracket \text{bi-} \rrbracket + \text{ri}$	$\llbracket \text{bi-} \rrbracket + \text{ri}$
$(\llbracket \text{bi-} \rrbracket)^C$	$(\llbracket \text{bi-} \rrbracket)^C$
$(\lambda x[\text{Child}(x) \wedge x \in \text{ATOM}])^C$	$(\lambda x[\text{Seed}(x) \wedge x \in \text{SUM}])^C$
$\lambda x_O[\text{Child}(x) \wedge x \in \text{SUM}]$	$\lambda x_O[\text{Seed}(x) \wedge x \in \text{ATOM}]$
$= \text{PL}(\text{bi-})$	$= \text{SG}(\text{bi-})$

Table 8: Derivations of Individual and Collective Aggregate Nouns

application of *-ri* gives the complement of the denotation of a singular noun, viz. the complement of the relevant set of atomic individuals. 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.¹³ The value returned is the atoms which form the sums, which is in turn exactly the value of the noun’s singular denotation.

It is informative to contrast the derivations in Table 8 with the semantic representation of plural and singulative morphemes. Assuming the analysis of Link (1983), the plural is represented as denoting the closure of atoms under join (\oplus) less the atoms themselves. Grimm (2012b) provides an analysis of the singulative in Welsh, given in (12): Singulative morphology applies to a collective base (here simplified to just requiring sum reference) and returns the atomic parts of the sum entities in the extension of the base form.¹⁴

$$(11) \quad \llbracket \text{PL} \rrbracket := \lambda P \lambda x. P_{\text{ATOM}}[P^*(x) - \text{ATOM}(x)] = \lambda P \lambda x. P_{\text{ATOM}}[P^*(x) \wedge x \in \text{SUM}]$$

$$(12) \quad \llbracket \text{SG} \rrbracket := \lambda P \lambda x. P_{\text{SUM}}[x \leq P \wedge x \in \text{ATOM}]$$

It is straightforward to see that the analysis of *-ri* as taking the complement relative to the base is connected to the meanings of plural and singulative morphology,

¹³Here I am treating the model-theoretic structure of the domain of plurals as “exclusive” (*sums – atoms*) rather than “inclusive” (*sums \cup atoms*). (See ch XX for a discussion of the contrast between inclusive and exclusive plurals.) Grimm (2012a) shows that an inclusive analysis is not viable for Dagaare (see Grimm 2012a for details).

¹⁴See Grimm (2012b) for an full analysis of the semantics of collective nouns and singulative morphology.

since as defined PL takes atoms to sums and SG takes sums to atoms. That is, the relative complement operation can be rewritten as the disjunction of plural and singulative operators. As PL and SG apply to disjoint domains, *-ri* plausibly arose through reanalysis of two separate markings in complementary domains as a single pattern.

There is some limited evidence that the analysis of *-ri* as a relative complement operator could extend to inverse morphology more generally. Recent work on Korean by Jeong (2017) shows that a similar treatment can be extended to a case of morphological reversal in the verbal domain. The Korean suffix *-i* applies to verbs, resulting for some verbs in causativization, but for other verbs inchoativization. Jeong (2017) argues that the effect of *-i* depends on whether the verb stem canonically designates a spontaneously occurring event or an externally caused event. Similar to the analysis just given, *-i* is analyzed as a complementation operator.

A different analysis of inverse number is provided for Kiowa in Harbour (2008, 2011). Under that treatment, the inverse number morpheme does not have, in itself, semantic value or meaning, but rather spells out feature incompatibility. In this account, nouns come with a basic number categorization, represented as a feature bundle under a Class projection, responsible for converting a root into noun. At the same time, a Number head also contains features designating a number specification which are valued in the syntax. If the feature set in Class (“the basic number”) conflicts with the feature set in Number, this conflict is spelled out as the inverse number morpheme. To exemplify, the noun for ‘fish’ in Kiowa has a basic number which consists of singular and dual reference (notated [*-augmented*] in Harbour’s system). Should Number possess a feature set relating a plural value (notated [*-singular, +augmented*]), a clash arises, and the form is spelled out with the inverse marker.

It is worth noting in passing that this account would have to be extended in some fashion to treat Dagaare. If *-ri* manifested a feature clash related to number, this would also be predicted to arise for the granular aggregate nouns which combine with the singulative *-ruu*: Their basic number consists of reference to aggregates, which would clash with the required singular values.

6 Outlook

This chapter has argued that inverse number in Dagaare, and coding asymmetries more generally, fit into larger trends within grammatical number systems. The coding asymmetry present in inverse systems is conditioned on lexical semantic patterns, which in turn are grounded in contrasting levels of individuation among

the nouns at issue. Thus, inverse number systems should not occur in languages haphazardly, but only as an asymmetry between noun types contrasting in different degrees of individuation. Much remains to be done to put this suggestion on firmer footing. As indicated in section 4, there is wide range of empirical work to be carried out to uncover the functional basis of coding asymmetries. If, as put forth here, effects of individuation play a role, direct, experimental evidence of such influence would bolster the case greatly.

A different set of studies are needed to examine, to the extent possible, the historical development of inverse number systems. For the case of Dagaare, it is plausible that the inverse number system developed from an earlier noun class system as a result of phonological change. Such a scenario would conform to Baerman's (2007) suggestion that reversals in general arise through diachronic change. Baerman (2007, 58) suggests that inverse systems develop from a "change that brings about a distribution of forms within a paradigm which superficially looks like a reversal" which then is reanalyzed as a reversal, whereupon it is extended by analogy. A large scale study of nominal classification in Gur languages could show if this is a plausible route of explanation for Dagaare. If the inverse system developed from separate singular-plural and collective-singular patterns, recognizing the equivalence of *-ri* as a relative complementation operator and the disjunction of a plural and singulative operator establishes a plausible route for semantic reanalysis.

References

- Arto Anttila and Adams Bodo. 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)*. Cascadia Proceedings Project, Somerville, Massachusetts, 2009.
- Harald Baayen, Cristina Burani, and Robert Schreuder. Effects of semantic markedness in the processing of regular nominal singulars and plurals in Italian. In Geert Booij and Jaap van Marle, editors, *Yearbook of Morphology 1996*, pages 13–33. Kluwer Academic Publishers, 1997.
- R.H. Baayen, R. Piepenbrock, and L. Gulikers. *CELEX2*. Linguistic Data Consortium, Philadelphia, PA, 1996.
- Emmon Bach. Parochial and universal semantics: Semantic typology and little studied languages. Handout for pre-ALT Workshop on Language Documentation and Linguistic Typology, Paris, France, 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.
- Adams Bodomo. *The Structure of Dagaare*. CSLI, Stanford, 1997.
- Adams Bodomo. *Dagaare: Languages of the World Materials No. 165*. Lincom Europa, München, Germany, 2000.
- Adams Bodomo and Charles Marfo. The morphophonology of noun classes in Dagaare and Akan. *Studi Linguistici e Filologici Online*, 4:205–244, 2006.
- Greville G. Corbett. *Gender*. Cambridge University Press, Cambridge, 1991.
- Greville G. Corbett. *Number*. Cambridge University Press, Cambridge, 2000.
- M.E. Kropp Dakubu. *Collected Language Notes on Dagaare Grammar*, volume 26 of *Collected Language Notes*. Institute of Language Studies, Legon, Ghana, 2005.
- Gerrit J. Dimmendaal. Number marking and noun categorization in Nilo-Saharan languages. *Anthropological Linguistics*, 42(2):214–261, 2000.
- Joseph Greenberg. *Language Universals*. Mouton, The Hague, 1966.
- Scott Grimm. Individuation and inverse number marking in Dagaare. In Diane Massam, editor, *Count and Mass Across Languages*. Oxford University Press, Oxford, 2012a.
- Scott Grimm. *Number and Individuation*. PhD thesis, Stanford University, 2012b.
- Scott Grimm. Grammatical number and the scale of individuation. *Language*, To Appear.
- 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, Dordrecht, 2008.
- Daniel Harbour. Valence and atomic number. *Linguistic Inquiry*, 42(4):561–594, 2011.

- Martin Haspelmath. Against markedness (and what to replace it with). *Journal of Linguistics*, 42(1):25–70, 2006.
- Martin Haspelmath and Andres Karjus. Explaining asymmetries in number marking: Singulatives, pluratives and usage frequency. *Linguistics*, 2017.
- Sunwoo Jeong. Markedness-sensitive semantics for causative alternation affixes: The case of Korean. Ms. Stanford University, 2017.
- Chigusa Kurumada and Scott Grimm. Communicative efficiency in language production and learning: Optional plural marking. In G. Gunzelmann, A. Howes, T. Tenbrink, and E. J. Davelaar, editors, *Proceedings of the 39th Annual Conference of the Cognitive Science Society*, pages 2500–2505, Austin, TX, 2017. Cognitive Science Society.
- Godehard Link. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow, editors, *Meaning, use, and interpretation of language*. De Gruyter, Berlin, 1983.
- Gabriel Manessy. *Les langues oti-volta. Classification généalogique d'un groupe de langues voltaïques*. SELAF, Paris, 1975.
- Gudrun Mieke and Kerstin Winkelmann, editors. *Noun Class Systems in Gur Languages, Vol. I Southwestern Gur Languages (without Guronsi)*. Rüdiger Köppe Verlag, Cologne, 2007.
- Salikoko S. Mufwene. Number, countability and markedness in Lingala LI- MA-noun class. *Linguistics*, 18:1019–1052, 1980.
- Avea E. Nsoh. Classifying the nominal in the Gurenɛ dialect of Farefare of Northern Ghana. *Journal of Dagaare Studies*, 2:83–95, 2002.
- 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 Watkins. *A Grammar of Kiowa*. University of Nebraska Press, Lincoln.
- Edward J. Wisniewski, Mutsumi Imai, and Lyman 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.