

# Number and Markedness: A view from Dagaare

Scott Grimm

Stanford University/ZAS

September 28, 2009

## Background

There is ongoing debate concerning the core semantics of singular and plural forms

Classic Link-style analysis, Exclusive/Strong Plural (Link 1983; Chierchia 1998):

- ▶ Singular denotations are atoms
- ▶ Plural denotations are the closure of atoms under the sum operator ( $\oplus$ ) less the atoms themselves :  
the reference domain of the plural is the set of *sums*

## Background

Problem: inference patterns under negation (“There are no horses at the coral”) and in questions (“Do you have children?”), where an answer about *one or more* is required, the Inclusive/Weak Plural

Krifka (1989) (and later Sauerland (2003) and Sauerland et al. (2005)) take the plural as intrinsically denoting the inclusive reading

- ▶ Singular denotations are atoms
- ▶ Plural denotations are the closure of atoms under the sum operation ( $atoms \cup sums$ ), i.e. the entire semi-lattice structure

# Semantic Markedness

Farkas and de Swart (to appear):

- ▶ following Horn's division of pragmatic labor:
  - ▶ the unmarked singular aligns with an unmarked meaning (*atoms*)
  - ▶ the morphologically marked plural aligns with a marked semantic meaning (*sums of atoms*)

Semantic markedness:

- ▶ atomic reference is the unmarked meaning
- ▶ sum reference (whether inclusive or exclusive) is the marked meaning

# Semantic Markedness

The inclusive plural analysis of English plurals claims that the plural is semantically unmarked, while the singular is more specific

Exclusive plural interpretations, e.g. in affirmative, episodic sentences, arise via pragmatic blocking

# Dagaare and Inverse Number Marking

Dagaare (Gur; Niger-Congo) possesses an inverse number marking system proving problematic for both sides of the debate

- ▶ the cross-linguistic facts, in Dagaare and beyond, are more complicated than if only the singular *or* plural were unmarked, rather markedness is conditioned upon a nominal's level of individuation

# Inverse Number Marking

Marking system for singular and plural distinctions where same marker can mark either singular or plural.

The number marking pattern of Dagaare is demonstrated by the near minimal pair below.

Both nouns share the same stem, yet *-ri* marks the plural interpretation for ‘child’ and the singular interpretation for ‘seed’.

Singular	Plural	Stem	Gloss
bie	biiri	bi-	‘child’
biri	bie	bi-	‘seed’

## Inverse Number Marking: Basic Pattern

The morphophonology does not provide any easy answer:

- ▶ The inverse marking pattern does not appear to be reducible to any coherent noun class pattern
  - ▶ Dagaare's noun class system is in great decay
  - ▶ Many of the vowel endings appear to be epenthetic (see Anttila and Bodomo 2007)
- ▶ This pattern is not triggered by any simple phonological environment, as evidenced by several minimal pairs



# Markedness

- ▶ This pattern is to all appearances a big problem for theories of markedness:
  - ▶ Usually, singular is unmarked and plural is marked (Jakobson, Greenberg's Universal 35)
  - ▶ This is clearly contradicted by the inverse number marking pattern

How can we align the inverse number marking pattern with what is known about markedness and the semantics of number?

# Number Marking in Dagaare

Dagaare, as well as other Gur languages, possesses an interesting number marking system

Three distinct morphological markers:

- ▶ *-ri*, sometimes marking singular, sometime plural
- ▶ *-ree*, a “second plural”, which gives the reading “different kinds of x” or “x in different locations”
- ▶ *-ruu*, a singulative marker  $\approx$  “a piece of”

## The Second (Distributive) Plural

Grammars of Dagaare normally only discuss the second plural *-ree* in relation to mass terms and liquids

The second plural is however very productive, and combines with nouns that are not mass terms:

- ▶ *waa* 'yam' has both a regular plural form *waari* 'yams' and a second plural form *waaree* which designates 'different (kinds of) yams'

# The Singulative

The singulative appears mainly with clearly mass terms as well as aggregates with are particularly close-knit:

muoruu	'blade of grass'	muo	'grass'
kpeeruu	'piece of malt'	kpee	'malt'

Several other languages dispose of a singulative (Breton, Welsh, Shilluk)

## Mass terms

The inverse marking pattern cannot be aligned with a mass/count distinction *tout court*: mass terms fall in a separate paradigm, combining with a distinct distributive plural marker *-nee* and singulative marker *-ruu*, depending on the noun.

Singulative	Mass	2nd Pl.	Gloss
—	kuo	konnee	'water/ (types of) waters'
muoruu	muo	muonee	'blade of grass/grass/grasses'

More typologically common are systems that resemble the inverse marking system, but possess a singular/plural contrast as well as a mass/singulative contrast (Nilo-Saharan)

# Towards a Semantic Approach: Individuation

## Individuation Hypothesis:

- ▶ Cognitive or perceptual qualities influence the grammatical realization of count and mass nouns
  - ▶ count nouns (dog) correlate with individual entities
  - ▶ mass nouns (water) correlate with non-individuated substances
- ▶ Open Question: Do speakers attend to individuation distinctions beyond the well-known count/mass dichotomy?

## Individuation: The Hunch

Nouns possess lexical information, i.e. nouns come with a ‘basic’ number determined by the noun’s semantic properties.

The application of *-ri* gives the inverse value.

[Highly Individuated N] + *-ri*  
⇒ plural

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

# Individuation

Individuation suffers in the same manner as other commonly cited conceptual factors in linguistics, such as animacy and agentivity—far from rigorously defined

- ▶ 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
- ▶ Consider the potential influence of four individuating factors on the realization nominals in Dagaare: animacy (Smith-Stark 1974; Corbett 1996, 2000), ease of distinguishability, manner of interaction (Wierzbicka 1988; Middleton et al. 2004), and “inherently plurality” (Acquaviva 2008)



## Individuation: Animacy

Animacy (relative to some sort of animacy scale ranging from humans to larger than smaller animals), which correlates to a scale of individuation, is known to influence number marking cross-linguistically (Smith-Stark 1974; Corbett 1996, 2000).

- ▶ The higher the entity corresponding to a noun rates on an animacy hierarchy, i.e. the closer to human a noun is, the greater likelihood that the noun is capable of expressing a singular/plural contrast.

The higher the animacy level of the entity, the more likely it will be treated as individuated and unmarked in the singular

## Individuation: Distinguishability

“Distinguishability” as a factor originates in Wierzbicka (1988)

- ▶ entities for which the constituents are more easily distinguishable are more likely to be used as count nouns while those entities for which the constituents are not easily distinguishable will be used as mass nouns.
- ▶ *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 terms of distinguishability.

## Individuation: Distinguishability

Middleton et al. presented subjects with pairs of aggregate displays which varied along two dimensions:

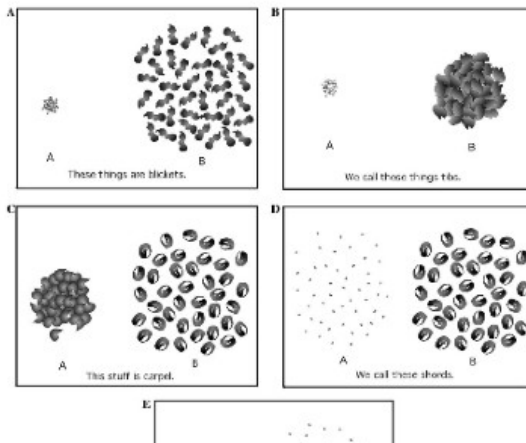
- (i) spatial proximity to other elements (Close versus Apart)
- (ii) size of elements (Large versus Small)

A subject would see a two sets of an element where for one set, for instance, 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.”

## Individuation: Distinguishability

Novel Aggregates Used in Middleton et al. (2004) (reproduced from p. 383)



## Individuation: Distinguishability

Results:

subjects' choice of count or mass terms was very significantly influenced ( $p < .001$ ) by the factor of spatial proximity, but not of the size, of the elements.

## Individuation: Manner of Interaction

Wierzbicka exemplifies “canonical manner of interaction” with examples such as the naming of berries in Polish.

- ▶ Berries are generally count terms because, she claims, people interact with them one by one, viz. picking/eating them
- ▶ Farmers selling berries typically use mass syntax to describe berries since they interact with them in quantities rather than individually

## Individuation: Manner of Interaction

This also receives experimental support from Middleton et al. (2004) with a similar forced choice design (mass vs. count syntax)

When subjects were presented with a novel aggregate—“yellow decorative coarse-grained sugar” in a cardboard box—they majoritarily assigned it a mass phrase (“This is worgle”).

When subjects interacted with the sugar by scooping up individual grains, they majoritarily assigned it a count phrase (“These are worgles”).

## Inherent Plurality

Acquaviva (2008) has emphasized the distinctive morphosemantic behavior of entities which canonically appear in collectives, duals and other “marked” number categories.

Individuation is normally considered only in light of mass/count syntax, but 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

Appears independent of the previous three factors:

- ▶ dual/collective paradigms is orthogonal to the animacy scale (Corbett 1996)
- ▶ distinguishability and interaction are relevant for aggregates when all else is held constant

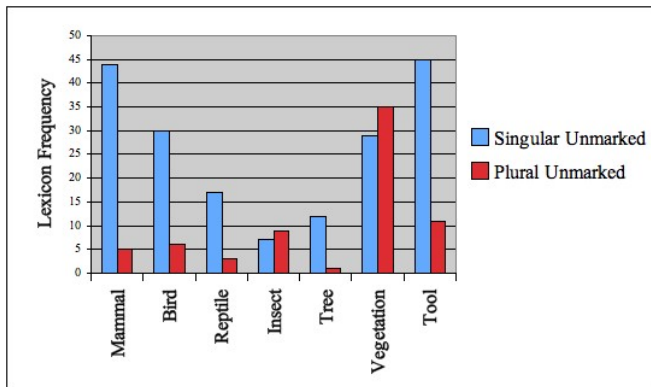


## Validation Across Semantic Domains

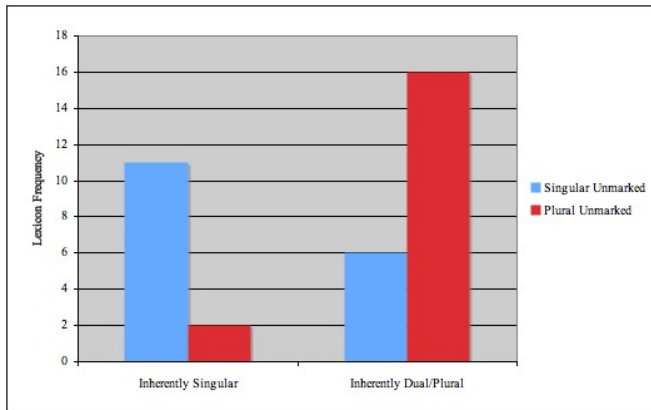
If individuation has an effect on the distribution of *-ri*, one should observe distributional asymmetries in the appropriate semantic domains. Four relevant predictions would be the following:

- (i) Nouns referring to larger (more salient) animals should be more likely to be unmarked in the singular than insects
- (ii) Nouns referring to trees should be unmarked in the singular in comparison to vegetation
- (iii) Nouns referring to tools more likely to be unmarked in singular (canonically interact with them individually)
- (iv) Nouns referring to body parts in pairs/groups are more likely to be unmarked in the plural while non-paired/grouped body parts are more likely to be unmarked in the singular

## Results from Fieldwork



# Validation for Body Parts



# Validation Across Semantic Domains

Reliable asymmetries are visible across the semantic domains:

- ▶ higher level animates, trees, tools and canonically individual body parts are typically unmarked in the singular
- ▶ insects, vegetation and inherently plural body parts have a majority of nouns for which the plural is unmarked

Bottom line:

- ▶ Dagaare morphology is sensitive to noun's degree of individuation/inherent plurality
- ▶ *-ri* marks singular when a noun is considered to be less individuated/inherently plural, otherwise marks the plural

## Language Internal Correlates: Domains of the Diminutive

Dagaare shows it is sensitive to collections/inherently plural distinction in the domains of its two diminutives: *-lee* ('small') and *-biri* ('seed')

Stem	Gloss	Diminutive	Resultant Noun (Sg)	Gloss
ba	'dog'	-lee	balee	'puppy'
naa- <i>ɔ</i>	'cow'	-lee	naalee	'calf'
gangaa	'drum'	-lee	gangalee	'alto drum'
gbee-	'leg/foot'	-biri	gbebiri	'toe'
nu-	'hand'	-biri	nubiri	'finger'
ko-	'funeral'	-biri	kobiri	'funeral song'
yel-	'say, tell'	-biri	yelbiri	'word'
baal-	'sick'	-biri	ballumbiri	'germ'

## Language Internal Correlates: Domains of the Diminutive

-lee and -biri are not equivalent:

- ▶ -lee is a standard diminutive
- ▶ -biri has an additional component yielding a collective sense—only results in small collections (toes, fingers, germs (“sick seeds”))

Compare with languages such as Ewe which has a diminutive from the word for ‘child’, *-vi* applies across both senses seen in Dagaare

Stem	Gloss	Resultant Noun (Sg)	Gloss
yevu	‘European’	yevu-vi	‘young European’
nyi	‘cow’	nyi-vi	‘calf’
kpe	‘stone’	kpe-vi	‘small stone’
afɔ	‘foot, leg’	afɔ-vi	‘toe’

## Language Internal Correlates: Dialect Variation

The picture argued for above would also predict variation for cases where an entity type could in principle be seen as unmarked in the singular or plural

Dagaare provides such cases:

<b>Singular</b>	<b>Plural</b>	<b>Gloss</b>	<b>Dialect</b>
piiri	pie	'rock/stone'	Central Dialect
pie	piiri	'rock/stone'	Southern Dialect

## Dialect Variation

Some nouns show dialectal variation between the singular being marked by *-ri* or the singulative *-ruu*

*kommiruu* and *kommiri* are attested for singular form of *kommie* 'tomato'

This dialectal 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 terms



# Cross-Linguistic Correlates

Cross-linguistic correlates to the unmarked plural in Dagaare surface in an array of language types

Similar semantic domains are relevant for collectives and duals in a number of languages, e.g. Breton (see Acquaviva 2008), Welsh (Stolz 2001)

Also for languages with nominal class systems, e.g. Swahili (Contini-Morava 2000) or Lingala (Mufwene 1980), some classes have been argued to be unmarked in the plural

- ▶ Relevant semantic domains are strikingly similar to Dagaare: vegetation, pairs and collectives, etc.

Despite different encodings, these systems all seem to make similar divisions along a scale of individuation

## Cross-Linguistic Correlates: Welsh

The same lexical semantic classes appear to be at issue in Welsh's number marking system, which possesses both a singular/plural distinction and a singulative/collective distinction ('leaf': *deil-en* singulative/*deil* pl.)

Semantic domains for singulative/collective (Stolz 2001):

- ▶ Small animals/insects
- ▶ Mid-sized animals coming in herds, swarms, etc.
- ▶ Vegetation/cereals/fruit

# Local Markedness

Tiersma (1982) noted that classes of nouns which “naturally occur in pairs or groups” tend to show surprising behavior wrt:

- ▶ leveling: a morphological paradigm levels in favor of the plural stem, e.g. in Frisian:

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

- ▶ double plurals: older plurals have been reanalyzed as singular units which leads to the addition of another plural marker

# Local Markedness

borrowing: where the plural stem is borrowed in preference to the singular

- ▶ Welsh borrowings from English:

*ffigys-en* (singulative) / *ffigys* (collective) < engl. *figs*

*gwsbery(s)-en* (singulative) / *gwsberys* (collective) < engl.  
*gooseberries*

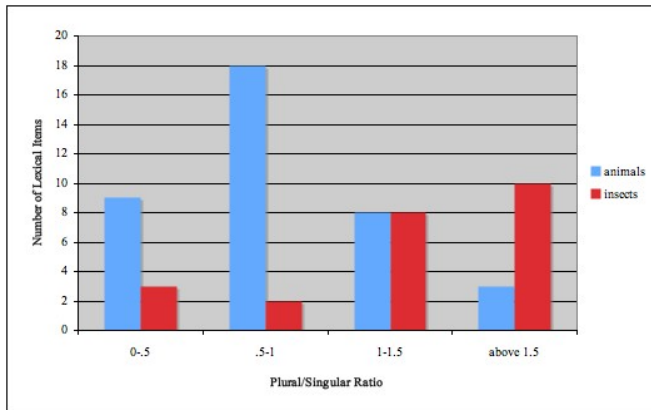
# Evidence from English Frequency Patterns

One would expect to see “unmarked plurals” reflected even in languages which do not display any morphological evidence of such a pattern

One prediction is that such classes of nouns should be unmarked in the plural in terms of frequency

- ▶ Examined frequencies from the COBUILD corpus (18 million words) provided by CELEX
- ▶ Calculated plural-to-singular ratio for animals and insects (basic terms and those consistent with the vocabulary of Dagaare)

# Evidence from English Frequency Patterns



## A Formal Account of *-ri*

Can align the data from Dagaare with two formal analyses, based on the exclusive and inclusive plural approaches

Classic Link-style analysis (Exclusive Plural):

- ▶ Singular denotations are atoms
- ▶ Plural denotations are the closure of atoms under the sum operator ( $\oplus$ ) less the atoms themselves

## An Exclusive Plural Account of *-ri*

Given that Dagaare disposes of lexical singulars and lexical plurals, the semantics of *-ri* is straightforwardly treated as a form of negation of the lexical denotation of the base

- ▶ A formal semantic update of Wonderly 1954's analysis of Kiowa, see also Bach (2008)



## An Exclusive Plural Account of *-ri*

Given that (i) the base of the noun has a denotation of the entire space generated by the atoms and their sums ( $atoms \cup sums$ )

and that (ii) the degree of individuation determines whether a noun is considered lexically plural or singular

- ▶ *-ri* can be modeled as the operation of complementation ( $C$ ) with respect to the domain of the base.
  - ▶ *-ri* applied to a lexically singular noun will yield a plural denotation
  - ▶ *-ri* applied to a lexically plural noun will yield a singular denotation

# An Exclusive Plural Account of *-ri*

Lexically Singular

$\llbracket \text{bi-} \rrbracket := \lambda x(\text{CHILD}(x))$

$\llbracket \text{bi-} \rrbracket + \text{ri}$

$(\llbracket \text{bi-} \rrbracket)^C$

$[\lambda x(\text{CHILD}(x))]^C$

$\lambda x[(\text{CHILD}(x))^{\oplus} - \text{CHILD}(x)]$

$= \text{PL}(\text{bi-})$

Lexically Plural

$\llbracket \text{bi-} \rrbracket := \lambda x((\text{SEED}(x))^{\oplus} - \text{SEED}(x))$

$\llbracket \text{bi-} \rrbracket + \text{ri}$

$(\llbracket \text{bi-} \rrbracket)^C$

$[\lambda x((\text{SEED}(x))^{\oplus} - \text{SEED}(x))]^C$

$\lambda x[\text{SEED}(x)]$

$= \text{SG}(\text{bi-})$

# Inclusive Plural

- ▶ Singular denotations are atoms
- ▶ Plural denotations are ( $atoms \cup sums$ ), i.e. the entire semi-lattice structure.

## An Inclusive Plural Account of *-ri*

The same inferences motivating the inclusive plural analysis 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.

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 (a suggestion by Uli Sauerland).

*-ri* can be uniformly analyzed as the closure of the space under join and meet.

# An Inclusive Plural Account of *-ri*

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

## Behavior Under Negation

These two hypotheses can be distinguished by behavior under negation, which was one of the primary motivations for the inclusive plural analysis.

In Dagaare, as in English, the negation of the plural interpretation excludes the truth of the singular.

- (1) John doesn't have children.  
→ False when John has one child.

Inclusive plural analysis predicts that only for nouns marked by *-ri*, as it results in the completion of the space, will negation correctly exclude both the singular and plural.

## A Formal Account of *-ri*: Behavior Under Negation

Acceptability patterns from Dagaare martial against the inclusive plural analysis: only the unmarked form is acceptable for inherently plural terms:

- (2) N        da    ba    da    bie/\*biri        (zaa)  
      1st.pro NEG buy Past seed.PL/seed.SG (any)  
      I didn't buy (any) seeds.

## A Formal Account of *-ri*: Behavior Under Negation

While the number markedness patterns of Dagaare differ from those of English and other IE languages, once one fixes a singular or plural interpretation, the behavior is parallel.



## Conclusion

The cross-linguistic facts, in Dagaare and beyond, are more complicated than if only the singular *or* plural were unmarked, rather markedness is conditioned upon a nominal's level of individuation

While inverse number marking seems surprising at first sight, under closer inspection it is a clever exploitation of relatively universal markedness patterns

- ▶ Less individuated/inherently plural entities are unmarked in the plural
- ▶ applying the logic of both strong and weak plural analyses shows the strong plural analysis has better empirical traction in such systems

# Thank you

Thanks to:

Mark Ali, Adams Bodomo, J.A. Saanchi for discussion of Dagaare data

Arto Anttila, Eve Clark, Paul Kiparsky, Beth Levin, Uli Sauerland and Tom Wasow for discussion

Stanford Center of African Studies for their generous support

## Plurals in Questions

Licensing of inclusive reading depends on specific expectations (examples from Farkas and de Swart):

- (3) Does Sam have a Roman nose/#noses?
- (4) Does a worm have eyes/#an eye?