

# On the Unity of ‘Number’ in Semantics and Morphology\*

Daniel Harbour

June/July 2004

This paper advances the case that linguistics requires a unified theory of number, serviceable to both semantics and morphology, by proving that the morphological concept of augmentation and the semantic concept of cumulation are near logical equivalents. From this emerge an inventory of number features incorporating the categories ‘paucal’ and ‘unit augmented’, a typology of number systems crosslinguistically, and indication of other areas of likely convergence between semantic and morphological research.

**Keywords:** augmented, cumulative, morphology, number, paucal, semantics

## 1 Introduction

In an extended study of the relationship between morphologically complex agreement and semantically based noun classification, Harbour (2003a) argued that the status quo in number theory—namely, that morphologists and semanticists concentrate on disjoint bodies of fact and develop correspondingly disjoint theories—is untenable: linguistics requires a unified morphosemantic theory of number. The current paper advances this case by observing that morphologists and semanticists have, despite divergent concerns, converged on a single discovery: the morphological notion of augmentation (Noyer 1992, Harbour 2003b) and the semantic notion of strict cumulativity (Krifka 1992) are near logical equivalents.

- (1) a. A predicate,  $P$ , is **augmented**,  $\text{Aug}(P)$ , if and only if  

$$\exists x \exists y [P(x) \wedge P(y) \wedge x \sqsupset y]$$
*(i.e., it is satisfied by two individuals,<sup>1</sup> one containing the other).*
- b. A predicate,  $P$ , is **strictly cumulative**,  $\text{Cum}(P)$ , if and only if  

$$\forall x \forall y [[P(x) \wedge P(y)] \rightarrow P(x \sqcup y)] \wedge \exists x \exists y [P(x) \wedge P(y) \wedge x \neq y]$$
*(i.e., it is satisfied by the join of any individuals, minimally two, that satisfy it).*

After outlining the quite disparate development and use of these concepts in morphology (section 2) and semantics (section 3), I demonstrate their near equivalence (section 4), as formulated in (2) and (3):

- (2) a. A predicate,  $P$ , is **additive**,  $\text{Add}(P)$ , if and only if  

$$\forall x \forall y [[P(x) \wedge P(y)] \rightarrow P(x \sqcup y)]$$
*(i.e., it is satisfied by the join of any individuals that satisfy it).*
- b. A predicate,  $P$ , is **augmented\***,  $\text{Aug}^*(P)$ , if and only if  

$$\text{Add}(P) \wedge \text{Aug}(P)$$
- (3) a. Augmentation entails additivity (for non-cardinality predicates, in morphologically relevant models).
- b. Augmentation\* entails strict cumulativity (in all models).
- c. Strict cumulativity entails augmentation\* (in all models).

Section 5 discusses the theoretical and practical significance of (3) for semantics and morphology, suggesting not only that we need a unified morphosemantic theory of number, but that some topics are immediately relevant to both semantic and morphological research. It also provides a new feature classification of agreement and pronominal categories, incorporating ‘unit augmented’ and ‘paucal’, and a typology of number systems.

## 2 Augmentation

Augmentation originates in descriptions of pronominal and agreement systems found in languages of the Philippines and of Australia’s Arnhem Land (Corbett 2000, whose exposition, pp. 166–169, is followed here). Its motivation lies in the rather odd view of such systems that results from use of the traditional descriptive categories ‘singular’, ‘dual’, ‘plural’, and so on. This was first noted by Thomas (1955) for Ilocano.

- (4) **Table 1**  
Ilocano pronominal forms (traditional categorization)

Person	Singular	Dual	Plural
1 inclusive		-ta	-tayo
1 exclusive	-ko	....	-mi ....
2	-mo	.....	-yo.....
3	-na	.....	-da.....

Observe that there is only one specifically dual form. As this is for the first person inclusive ‘you and I’, the dual is to some extent ‘forced’ on the language—a singular inclusive is impossible. Yet, by adopting [ $\pm$ augmented], one can avoid positing this defective, semantically predictable dual:

- (5) **Table 2**  
Ilocano pronominal forms (revised categorization)

Person	[−augmented]	[+augmented]
1 inclusive	-ta	-tayo
1 exclusive	-ko	-mi
2	-mo	-yo
3	-na	-da

By way of illustration, consider the top two rows. For the first person inclusive, let us take  $P$  to be the predicate ‘includes “I” and includes “you”’. Then, for *-tayo*, [1 inclusive +augmented],  $\exists x \exists y [P(x) \wedge P(y) \wedge x \sqsubset y]$  means that the model has two individuals, both containing ‘I’ and ‘you’, the one individual contained in the other. Minimally, then, the model includes  $\{I, \text{you}, \text{other}\}$  (as this contains an individual,  $\{I, \text{you}\}$ , that contains both ‘I’ and ‘you’), though it may also contain  $\{I, \text{you}, \text{other}_1, \dots, \text{other}_n\}$ , up to arbitrary  $n$ . This is the desired result. By contrast, for *-ta*, [1 inclusive −augmented],  $\forall x \forall y [\neg P(x) \vee \neg P(y) \vee x \not\sqsubset y]$  means that there are no pairs containing both ‘I’ and ‘you’ that are in a containment relation; but, as  $\{I, \text{you}\}$  would be contained by any other individual satisfying  $P$ ,  $\{I, \text{you}\}$  must be the only individual satisfying  $P$ . Again, this is the desired result.

For the first person exclusive, let us take  $P$  to be the predicate ‘includes “I” and excludes “you”’. Then, for *-mi*, [1 exclusive +augmented],  $\exists x \exists y [P(x) \wedge P(y) \wedge x \sqsubset y]$  means that the model has two individuals, both

containing ‘I’ but not ‘you’, with one individual contained in the other. By the reasoning above, the model minimally includes {I, other}, the desired result. And for *-ko*, [1 exclusive –augmented],  $\forall x\forall y[\neg P(x) \vee \neg P(y) \vee x \not\supseteq y]$  means, by the reasoning above, that the model has a unique individual, {I}, satisfying P, again the desired result.

Observe that [ $\pm$ augmented] permits flexibility in cardinality. In Ilocano, [–augmented] sometimes entails cardinality 1, sometimes 2; [+augmented] sometimes entails cardinality 2 or more, sometimes 3 or more. Analogous to Ilocano, but more complex, is Rembarrnga (McKay 1978, 1979), which displays, in traditional terms, singular, dual, trial and plural. However, trial is restricted to first person inclusive (*cf.*, Ilocano’s dual). Here, then [–augmented] can entail cardinality 1, 2, or 3. What is particularly elegant about this trial-free reanalysis is that the forms ending in *-bbarrah* occupy the same part of the ‘paradigm’ (as opposed to *ngakorrbbarrah* being trial and all other *-bbarrah* forms being dual):

(6) **Table 3**  
Rembarrnga dative pronouns

Person	Minimal	Unit Augmented <sup>2</sup>	Augmented
1 inclusive	y $\text{akk}\text{u}$	ngakorr <b>bbarrah</b>	ngakor $\text{u}$
1 exclusive	ng $\text{nn}\text{u}$	yarr <b>bbarrah</b>	yarr $\text{u}$
2	k $\text{u}$	nakor <b>bbarrah</b>	nakor $\text{u}$
3 masculine	naw $\text{u}$	barr <b>bbarrah</b>	barr $\text{u}$
3 feminine	ngad $\text{u}$	barr <b>bbarrah</b>	barr $\text{u}$

It is expressly with this flexibility of cardinality in mind that [ $\pm$ augmented] has been defined. Various versions have been offered: [ $\pm$ restricted] (Conklin 1962), [ $\pm$ others] (Matthews 1972), as well as Noyer’s (*op. cit.*), which is the most robustly typologically tested to date. (1a) is just a notational variant of this, minus the condition that individual must be non-zero. Such a ban is crucial, as, without it, [–augmented],  $\forall x\forall y[\neg P(x) \vee \neg P(y) \vee x \not\supseteq y]$ , is satisfied by the empty set (and one does not want first person inclusive non-augmented, say, to refer to non-entities). However, in another point of contact between the two theories of number, Krifka too must rule out zero elements, witness his  $\neg\exists x\forall y[x \sqsubseteq y]$  postulate. As I am urging a unification of semantic and morphological treatments of number, it will suffice to stipulate the ban once for both.

### 3 Cumulation

The term ‘cumulative reference’ originates with Quine and his treatment of the ontogenesis of reference: ‘*mass* terms like “water” ... have the semantical property of referring cumulatively: any sum of parts which are water is water’ (1960, p. 91). It has come to play a significant role in semantics owing largely to Krifka’s treatment of the interaction of nominal and verbal reference. For instance, Krifka (1992, pp. 33–36) shows that the telic~atelic distinction can be reduced to denial~assertion of strict cumulativeness. Specifically, he defines the notion of having a set terminal point as the formal correlate of telicity and then shows that strict cumulativeness entails atelicity, that is, non-existence of a set terminal point.

This formal result is important. It reduces telicity, a property of events only, to strict cumulativeness, a property that applies equally to events and to objects. This permits a straightforward analysis of the dual fashion in which telicity can arise: either a predicate is inherently telic (e.g., *arrive*), or a predicate that is ordinarily atelic (e.g., *drink*) may become so if its object is non-cumulative (e.g., *a glass of wine*—non-cumulative because a glass of wine plus another glass of wine is no longer just a glass of wine).

(A minor difference should be noted between Krifka’s definition of strict cumulativeness (7) and that assumed here.

$$(7) \quad \forall x \forall y [[P(x) \wedge P(y)] \rightarrow P(x \sqcup y)] \wedge \neg \exists x [P(x) \wedge \forall y [P(y) \rightarrow x = y]]$$

(7) is weaker than (1b), since only (7) is true of predicates that are true of no individuals: (1b)  $\models \exists x \exists y [P(x) \wedge P(y) \wedge x \neq y]$ ; however (7) entails the same proposition only given the auxiliary assumption that the predicate is true of at least one individual:  $\{(7), \exists x P(x)\} \models \exists x \exists y [P(x) \wedge P(y) \wedge x \neq y]$ . This minor difference is irrelevant for current purposes.<sup>3)</sup>

### 4 Equivalence

Clearly, the empirical concerns of the morphologists who devised the notion of augmentation are very different from those of the semanticists who devised the notion of cumulativeness. Nonetheless, I now prove the statements in (3). Formally stated, they are:

- (8) a.  $\text{Aug}(\text{P}) \models_{\mathcal{M}} \text{Add}(\text{P})$  (for any model,  $\mathcal{M}$ , relevant to morphology, and for any non-cardinality predicate,  $\text{P}$ )  
 b.  $\text{Aug}^*(\text{P}) \models \text{Cum}(\text{P})$   
 c.  $\text{Cum}(\text{P}) \models \text{Aug}^*(\text{P})$

To demonstrate (8a), I take non-cardinality predicates to be person and gender predicates. As the discussion of Ilocano illustrates, when morphologists are concerned with semantics (of pronominal or agreement categories), they are generally concerned with groups of people/things and whether they include the speaker, the hearer and/or others. If  $\text{P}$  is a predicate that denotes inclusion or exclusion of ‘I’ or ‘you’ from an individual of arbitrary size, then  $\text{P}$  obeys additivity, as the join of any two individuals of arbitrary size containing ‘I’, say, or excluding ‘you’ and ‘I’, is another individual of arbitrary size with the same property. Similarly, if two individuals consist entirely of feminine individuals or if they contain at least one masculine one, then their joins will also contain only feminine individuals, or at least one masculine one. We saw in section 2, that [+augmented] permits individuals of arbitrary size. So, in morphological models,  $\mathcal{M}$ , augmentation entails additivity, that is,  $\text{Aug}(\text{P}) \models_{\mathcal{M}} \text{Add}(\text{P})$ , for any non-cardinality predicate,  $\text{P}$ .

To demonstrate (8b), observe that, if  $\text{Aug}^*(\text{P})$ , then  $\text{Aug}(\text{P})$ , and so there are individuals,  $a$  and  $b$ , such that  $\text{P}(a)$  and  $\text{P}(b)$  and  $a \sqsupset b$ . From  $a \sqsupset b$ , it follows  $a \neq b$ . So, we can write  $[\text{P}(a) \wedge \text{P}(b) \wedge a \neq b]$ . By existential quantification, we have  $\text{Aug}(\text{P}) \models \exists x \exists y [\text{P}(x) \wedge \text{P}(y) \wedge x \neq y]$ . So,  $\text{Aug}^*(\text{P}) \equiv \text{Add}(\text{P}) \wedge \text{Aug}(\text{P}) \models \text{Add}(\text{P}) \wedge \exists x \exists y [\text{P}(x) \wedge \text{P}(y) \wedge x \neq y] \equiv \text{Cum}(\text{P})$ .

To demonstrate (8c), observe that any model for  $\text{Cum}(\text{P})$  will contain two individuals,  $a$  and  $b$ , such that  $\text{P}(a)$ ,  $\text{P}(b)$  and  $\text{P}(a \sqcup b)$ . Since  $a \sqcup b \sqsupset a$ , it follows that  $[\text{P}(a \sqcup b) \wedge \text{P}(a) \wedge a \sqcup b \sqsupset a]$ . So, by existential quantification, we have  $\text{Cum}(\text{P}) \models \exists x \exists y [\text{P}(x) \wedge \text{P}(y) \wedge x \sqsupset y] \equiv \text{Aug}(\text{P})$ . As conjunctions entail conjuncts,  $\text{Cum}(\text{P}) \models \text{Add}(\text{P})$ . So,  $\text{Cum}(\text{P}) \models \text{Add}(\text{P}) \wedge \text{Aug}(\text{P}) \equiv \text{Aug}^*(\text{P})$ .

## 5 Ramifications

It is surely remarkable that morphologists concerned with agreement and pronoun inventories and semanticists concerned with the interaction of telicity with nominal and verbal reference should have converged on two such similar notions as augmentation and strict cumulativity. As augmentation\*

and cumulativity are logically equivalent, one can be dispensed with and the other adopted in morphology and semantics alike. Which?

I suggest that augmented\* be adopted, as it induces a classification of pronominal and agreement categories in terms of its conjuncts, additivity and augmentation, that is superior to that induced by strict cumulativity. In particular, it illuminates the notions of ‘unit augmented’ in (6) and ‘paucal’.

First, observe from the comments following (5) and from the proof of (8a) that the category augmented is [+additive] (in addition, of course, to [+augmented]). Combining [−additive] with [+augmented] yields, I suggest, a paucal. Paucals pick out groups with few members (Foley 1991, p. 111, for instance, says the Yimas paucal generally ranges from 3 to 7); they are a plural-like category with an upper bound. Consequently, a paucal could identify {a, b, c, d} or {a, b, c}. As the former contains the latter, paucals are [+augmented]. However, they are [−additive], by the reasoning of the Sorites paradox: a few plus a few is not necessarily a few.

Now consider unit augmented forms. Second person unit augmented, say, identifies {you, other<sub>1</sub>}, {you, other<sub>2</sub>}, and so on. Clearly, none of these is contained in any other; so, unit augmented is actually [−augmented]. Moreover, unit augmented is non-additive, as {you, other<sub>1</sub>}  $\sqcup$  {you, other<sub>2</sub>} does not contain just ‘you’ and a unique other. Hence, unit augmented is [−additive −augmented]. We can summarize these results as:<sup>4</sup>

(9) **Table 4**  
Typology of agreement/pronoun categories

[±additive]	[±augmented]	Category
+	+	plural, augmented
−	+	paucal
−	−	unit augmented

The typology can be expanded to include the traditional singular and dual using the feature [±singular], which asserts that cardinality equals 1. Harbour (2003a, pp. 87–89) shows that singular is [+singular −augmented] and dual, [−singular −augmented]. Singular and dual cannot be inherently specified for [±additive]: singular is [−additive] for third person, [+additive] otherwise; dual is [+additive] for first person inclusive, [−additive] otherwise. Failure fully to crossclassify may make the feature inventory [±singular], [±additive], [±augmented] seem somewhat redundant—especially, also, as

there are several ways to specify, for instance, plural: [−singular], [−singular +augmented], [−singular +additive +augmented], ... I suggest that this accurately reflects the partial overlap between number systems crosslinguistically. By activating different subsets of number features, we can characterize such overlapping systems as singular-dual-plural and singular-paucal-plural, say. I propose the following typology (examples from Corbett 2000).

(10) **Table 5**  
Typology of number systems

Language	Categories	Features
Pirahã	no number distinctions	[ ]
Russian	singular, plural	[±sg]
Ilocano	minimal, augmented	[±aug]
Unacquirable <sup>5</sup>	non-additive, additive	[±add]
Upper Sorbian	singular, dual, plural	[±sg ±aug]
Bayso	singular, paucal, plural	[±sg ±add]
Rembarrnga	minimal, (unit) augmented	[±add ±aug]
Yimas	singular, dual, paucal, plural	[±sg ±add ±aug]

It follows from paucal’s being a composite of features that no language can have paucal~non-paucal as its only number distinction: singular~non-singular must also be distinguished. Similarly, unit-augmented~non-unit-augmented cannot be a language’s only number distinction. Corbett’s study number systems suggests that these predictions are correct.

To my knowledge, no feature classification has been previously given for paucal or unit augmented (beyond the unenlightening [±paucal]).<sup>6</sup> However, we lose this new insight, and the typologies it implies, if we adopt cumulative in place of augmented\*. For, then  $\exists x \exists y [P(x) \wedge P(y) \wedge x \neq y]$  replaces  $\exists x \exists y [P(x) \wedge P(y) \wedge x \sqsupset y]$ . Since both paucal and unit augmented satisfy the former condition, we lose any ready means of distinguishing them.

What follows from semanticists’ adopting morphologists’ augmented\*? It would be a mistake, after the initial bout of search-and-replace subsides, simply to return to business as normal. The convergence of morphological and semantic research indicates that the fields are closer than their disparate subject matter suggests, close enough, in fact, to demand a unified theory of number, as argued more generally by Harbour (2003a). If this conclusion is correct, then it is likely that further points of contact or convergence



will emerge and this will hold practical implications for morphologists and semanticists alike.

For morphologists, a practical implication is that, before one posits a new morphological feature, one should first search the semantic literature for kin concepts; many morphological concepts (paradigmatic dimensions) are likely not *sui generis*. This has just been illustrated for unit augmented and paucal, which are easily characterized once one has additivity and augmentation. (See Ojeda (1998) for a semantically rooted, morphologically insightful study of distributives and collectives.)

For semanticists, a practical implication is that paradigmatic distinctions are testing grounds for semantic concepts. This too was illustrated above. The paucal is especially interesting as its use is determined both by ‘absolute size of the group being referred to’ and by ‘relative size, i.e. whether the group being referred to is contrasted with some larger group within which it is subsumed’ (Crowley 1982, p. 81, as cited by Corbett 2000, p. 24). One and the same form is, then, ambiguous between *few* and *few of* readings.

A similar point of contact is provided by associative plurals. The Japanese morpheme *-tati*, for instance, creates ‘a non-uniform plural whose extension can include entities that are not in the extension of the common noun to which *-tati* is attached’ (Nakanishi and Tomioka 2004). Non-uniformity is illustrated by *tati*-modified proper names (ibid.):

- (11) Taro-tati-wa moo kaetta  
Taro-TATI-TOP already went home  
‘The group of people represented by Taro went home already’

The role of *Taro* in *Taro-tati* is strongly reminiscent of the role of first person in the first person plural: not multiple first persons, but multiple persons including the first. Similarly, *Taro-tati* does not mean multiple Taros, but multiple persons including Taro. Masculine gender is the same in some languages (Philippe Schlenker, p.c.) (e.g., Romance, Semitic): masculine plural agreement need not indicate multiple masculine things, but multiple things including a masculine one. This means that the interaction of personhood with plurality is not a quirk of person *per se* but is an instance of a more widespread semantic phenomenon, of interest to morphologists and semanticists alike, one that cannot be satisfactorily treated by theory of number that is solely morphological or solely semantic.

## Notes

\* Thanks are due to Christian List, and to David Adger, Susana Béjar and Philippe Schlenker.

<sup>1</sup>An individual is an atom or set of atoms, or, equivalently, a lattice point.

<sup>2</sup>On the feature composition of ‘unit augmented’ (‘minimal’ plus one other), see section 5.

<sup>3</sup>Krifka abbreviates the property in (7) as SCUM(P) and uses CUM(P) for (2a), which I have termed ‘additivity’ in order to avoid using ‘strictly cumulative’ and ‘cumulative’ in rapid alternation.

<sup>4</sup>I suggest that the fourth possibility, [+additive –augmented], is unacquirable. It would pick out, under varying person features, {I, you}, {I}, {you} and would be undefined otherwise. This is equivalent to restriction of [ $\pm$ augmented] to first and second person, an instantiation of a general crosslinguistic pattern: languages frequently make a number distinction only for upper parts of the animacy hierarchy (Corbett 2000). If the generalization is part of UG, then I suggest that [+additive –augmented] would be ‘misacquired’ as an instance of it. So, its exclusion from (9) is justified.

<sup>5</sup>By (8a), all person/gender categories would be [+additive]; so, there would be no evidence that [ $\pm$ additive] alone is active.

<sup>6</sup>Noyer (1992, pp. 198–199) attributes the difference between unit augmented and (normal) augmented to ‘functional inference’. Though the current system more tightly constrains the meaning of unit augmentation, it does not derive that it is augmented by one: the ‘dual augmented’ satisfies the same feature specifications. I leave this issue open.

## References

- Conklin, Harold. 1962. Lexicographic treatment of folk taxonomies. In F. W. Householder and S. Saporta, eds., *Problems in Lexicography*, Publication of the Indiana University Research Center in Anthropology, Folklore, and Linguistics 21, 119–141, Bloomington Indiana.
- Corbett, Greville. 2000. *Number*. Cambridge: Cambridge University Press.
- Crowley, Terry. 1982. *The Paamese Language of Vanuatu*, volume B 87 of *Pacific Linguistics*. Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University.
- Foley, William. 1991. *The Yimas Language of New Guinea*. Stanford, CA: CSLI Publications.
- Harbour, Daniel. 2003a. *Elements of Number Theory*. Cambridge, MA: MITWPL.
- Harbour, Daniel. 2003b. The Kiowa case for feature insertion. *Natural Language and Linguistic Theory* 21:543–578.
- Krifka, M. 1992. Thematic relations as links between nominal reference and temporal constitution. In Ivan Sag and Anna Szabolcsi, eds., *Lexical matters*, 29–53, Stanford, CA: CSLI.
- Matthews, Peter. 1972. Huave verb morphology: Some comments from a non-tagmemic viewpoint. *International Journal of American Linguistics* 38:96–118.
- McKay, Graham. 1978. Pronominal person and number categories in Rembarrnga and Djebbbana. *Oceanic Linguistics* 17:27–37.
- McKay, Graham. 1979. Gender and the category *unit augmented*. *Oceanic Linguistics* 18:203–210.
- Nakanishi, Kimiko and Tomioka, Satoshi. 2004. Japanese plurals are exceptional. *Journal of East Asian Linguistics* 13:141–179.
- Noyer, Rolf. 1992. Features, positions and affixes in autonomous morphological structure. Ph.D. thesis, MIT.
- Ojeda, Almerindo E. 1998. The semantics of collectives and distributives in Papago. *Natural Language Semantics* 6:245–270.

Quine, Willard Van Orman. 1960. *Word and Object*. Cambridge, MA: MIT Press.

Thomas, David. 1955. Three analyses of the Ilocano pronoun system. *Word* 11:204–208.