PAUCITY, ABUNDANCE, AND THE THEORY OF NUMBER

Daniel Harbour

Queen Mary University of London

Approximative numbers, like paucal and greater plural, can be characterized in terms of a feature, [±additive], concerned with additive closure. The two parameters affecting this feature (whether it is active and whether + and − values may cooccur) also affect the two features that generate nonapproximative numbers. All three features are shown to be derivative of concepts in the literature on aspect and telicity, to have a straightforwardly compositional semantics, and to eschew ad hoc stipulations on cooccurrence (such as geometries and filters). Thus, what is proposed is a general theory of number, free of extrinsic stipulations. Empirically, the theory yields a characterization of all numbers attested crosslinguistically, a combinatorial explanation of Greenberg-style implications affecting their cooccurrence, a natural account of morphological compositionality, and insight into their diachronic sources and trajectories.*

Keywords: features, geometry, (greater) paucal, greater plural, number

1. Introduction. In the domain of count nouns and personal pronouns, the morphological numbers of natural language fall into three descriptive classes: exact, approximative, and catch-all. The exact are singular, dual, and trial, minimal, and unit augmented.[1] The approximative numbers, characterized by their inexact upper and/or lower bounds, are paucal and greater paucal, greater plural, greatest plural, and global plural. Quantities that lie beyond whichever of these a language chooses fall to the catch-all category of plural (or augmented if the only lower numbers are minimal, or minimal and unit augmented).[2] The meanings of the approximative numbers are laid out in more detail below, but, in brief, paucal covers small numbers, greater paucal, slightly larger numbers, greater plural, quantities above those of normal pluralities,

* Any work on the typology of number stands in the shadow of Greville Corbett's seminal (2000) study, and my first acknowledgment must therefore be to him, and to the speakers and linguists on whose tireless work all such research relies. Additionally, I wish to thank David Adger, Matthew Baerman, John Brownie, Colleen Cotter, Stephanie Harves, and Linnaea Stockall for discussion of many different aspects of the current proposals. The article has also benefited greatly from the thoughtful input of two anonymous referees of this journal, as well as editor Greg Carlson and associate editor Heidi Harley. The research reported below was undertaken under the auspices of Project א (Atomic Linguistic Elements of Phi), Subproject צ (Theoretical Structure of Approximatives, their Diversity and Integrity), Arts and Humanities Research Council of the United Kingdom grant AH/G019274/1.

The following abbreviations are used: 1: first person, 2: second person, 3: third person, A: agent, aug: augmented, def: definite, du: dual, excl: exclusive, gr.gr.pl: greatest plural, gr.pl: greater plural, incl: inclusive, min: minimal, mood: mood, nsg: nonsingular, O: object, pc: paucal, pfv: perfective, pl: plural, sg: singular, tma: tense, mood, aspect, tr: transitive, u.aug: unit augmented. As a methodological clarification, I note that I cite Corbett (2000) frequently below in part not to diminish his scholarly contribution and in part because his discussion frequently contains crucial information not present in earlier sources. Nonetheless, in all cases where a language receives more than mere mention, I have returned to primary materials, even if these are not directly reflected in the discussion.

[1] Minimal, first described in Thomas 1955 for Ilocano, is the smallest satisfaction of a given person specification—hence, dual for first-person inclusive, and singular for exclusive, second, and third. Unit augmented, first described in McKay 1978, 1979 for Rembarrnga, is one more than minimal—hence, trial for inclusive, dual otherwise.

[2] All of these distinctions are discussed in Corbett 2000, though 'greatest plural' is not a category he explicitly names and it used tentatively here. Global plural might be taken to be a definite version of some other kind of plural. It is also used tentatively here; see §3.3. Corbett also discusses general number and what we might call 'modes of plurality', such as collectives and distributives, as well as plurals of plurals. See §5.3 on the last and n. 14 on the others. Mass nouns are discussed in §§2.3 and 4.4.

greatest plural, larger numbers still, and global plural exhausts all instances of a given nominal in the domain of discourse.

Besides their semantic variety, approximative numbers raise other fundamental questions that an explanatory theory of linguistic number must answer. On the one hand, they are subject to Greenberg-style implications, such as: a language with greater paucal must have paucal, and: a language with paucal must have singular or minimal. Table 1 presents the full set of such implications, both approximative and nonapproximative, that are, I believe, to be derived.

NOTATION	IMPLICATION
TR → DU	Trial requires dual.
DU → SG	Dual requires singular.
SG → PL	Singular requires plural.
PL → SG/MIN	Plural requires singular or minimal.
U.AUG → AUG	Unit augmented requires augmented.
MIN → AUG/PL	Minimal requires augmented or plural.
AUG → MIN	Augmented requires minimal.
GR.PC → PC	Greater paucal requires (lesser) paucal.
PC → PL	Paucal requires plural.
GR.PL → PL/AUG	Greater (and global) plural requires plural or augmented.

TABLE 1. Implicational universals (following Greenberg 1966, Corbett 2000).

On the other hand, in a good number of languages, approximative numbers are not morphologically primitive. In Motuna, for instance, paucal is frequently composed by crosscutting 'dual' with 'plural', as illustrated in 1–3. (Examples from Onishi 1994:287, my glossing and translation; [o] denotes deleting of the *o*.)

(1) kaa- ru-ti- m[o]-ee
disagree-2- DU-TMA-DU/PC
'you (two) disagree'

(2) kaa- ru-m- m[o]-ee
disagree-2- PC/PL-TMA- DU/PC
'you (few) disagree'

(3) kaa- ru-m- mo- ng
disagree-2- PC/PL-TMA-PL
'you (many) disagree'

In Fula (Anderson 1976), the greater plural is a morphological addition to the normal plural.

(4) pucc- i- (iji)
horse-PL-GR.PL
'(a great many) horses'

Additionally, paucals appear frequently to arise from semantic drift of a numeral reanalyzed as part of a (pro)noun, and greater plurals from concatenation of plural morphemes. Such typological implications, morphological compositions, and diachronic dispositions are all explicanda of a theory of number.

This article attempts to answer these questions using just one feature, [±additive], which divides (pro)nominal lattices into subregions that have an upper bound and those that do not have such a cut-off (terminology that is explained in §3). It is governed by two syntactic parameters, concerning whether Number^0 hosts the feature at all, and, if so, whether it may host [+additive −additive] simultaneously. In addition, the feature is subject to a sociosemantic convention affecting where the cut that it induces is placed (roughly, low for paucals, high for greater plurals).

Faced with the apparent oddity of [+additive −additive], it is natural to ask to what extent the parameters posited for approximative numbers have independent motivation, either within the domain of number or beyond it. The central claim of the current proposal is that exactly the same syntactic parameters as those that govern [±additive] also govern the features that generate the exact numbers and their complements (Noyer 1992, Harbour 2011b). Thus, we have a general theory of number, rather than separate theories of different varieties of number.

In fact, the theory of number proposed here offers quite substantial generality. First, it requires no filters on outputs, no geometric dependencies between features, and no tacit assumptions about order of semantic composition in order to prevent aberrant number values. Such 'constraints' as the analysis needs emerge either from the semantics of the features themselves or from more fundamental constraints on cognitive systems.

Second, the theory exhibits intriguing connections to other areas of linguistic and nonlinguistic cognition. The concepts behind the features used below are not novel, but have been posited in the analysis of masshood, aspect, and telicity. Given that numbers like (greater) paucal, (unit) augmented, and dual have not been research concerns in those domains, the emergence, from such research, of the means to characterize these numbers is quite striking. In addition, the features provide the means of defining a successor-like function and, so, of characterizing the natural numbers. Thus, the theory is general in the sense that it uses the same types of parameters for all numbers, that it is free from extrinsic stipulations, and that it connects with other areas of cognitive science, both within linguistics and beyond it.

The capacity to generate the natural numbers speaks against an alternative approach to grammatical numbers that grounds numbers like dual in quantifiers like *two*. This view harks back to the early days of grammatical theory, where featuralized versions of traditional grammatical labels, like [+dual], basically amounted to featuralized numerals. The approach was abandoned in light of the discovery of the natural classes of numbers (e.g. Hale 1973, Silverstein 1976) that manifest themselves in, among other things, the overlapping exponence of Motuna and Fula above (but see Cowper 2005 for a reconsideration of the approach). More fundamentally, though, viewing grammatical number as the derivative of a given stock of quantifiers shifts the explanatory burden without diminishing it. We still require an explanation of where those quantificational concepts come from. Moreover, quantifiers and numbers are mismatched. There are many more numerals than there are grammatical numbers (there are no quadrals and quinquals, let alone sexals and septals). Conversely, there are grammatical numbers, such as minimal, unit augmented, and, possibly, greater/greatest plural, that do not correspond to any numeral or simplex quantifier-like element. None of these problems arise on the current approach.

Beside such abstract concerns, the theory has considerable empirical coverage. It provides a natural means of deriving Greenberg-style implications between number values (for instance, no greater paucal without paucal, no trial without dual) and predicts a typology of number systems that is thoroughly, though not completely, saturated. Furthermore, the featural identities it affords to approximative and nonapproximative numbers permit a range of patterns of morphological compositionality, whereby one number arises from crosscutting two others or serves as the base from which others are constructed, like the Motuna and Fula examples above. Finally, the account makes interesting predictions about the diachronic rise of higher and lower approximative numbers and when featural (and phrase-structural) reconstrual of numbers can, or must, occur.

These arguments are structured as follows. First, the broad theoretical framework and specific theoretical apparatus assumed below are provided (§2). Then [±additive] is defined, its parametrization is motivated, and all approximative numbers are exemplified (§3). Section 4 embeds the treatment of approximative number in a theory of non-approximative number using the features [±atomic] and [±minimal] (to be defined). Besides showing that the same parameters govern all three features, the section argues for the semantic and cognitive naturalness, or minimality, of apparent constraints on the features and examines connections between analyses of number, mass, aspect, and the integers. We then return to more empirical concerns, exploring the typology of implications and number systems that the theory predicts, the types of morphological relationships between numbers that it makes possible, and the kind of diachronic development to which it implies that languages are predisposed (§5). Finally, §6 compares the account with the feature geometry of Harley & Ritter 2002, arguing that the current proposal offers both theoretical and empirical advantages. Online appendices present a range of formal results germane to the theory.[3]

2. Preliminaries. Fundamentally, I think of grammatical numbers, like dual, augmented, and paucal, as the result of actions by features on lattices. A minimal preliminary to what follows, therefore, is an exposé of the relevant lattices (§2.1) and what it means for features to act on them (§2.2). Beyond that, though, I make use of some specific assumptions about the syntactic structures that the features occupy, on the one hand, and that give rise to the lattices that they act on, on the other; these are coupled with some very general assumptions about how syntax maps to morphemes (§2.3). And to keep representations of semantics, syntax, and parameters distinct, I lay out some notational conventions (§2.4).

I hasten to add, though, that the current treatment of grammatical number is intended to be largely 'platform independent', implementable in any theory committed to the view that some meanings are composed from the meanings of more fundamental elements that correlate with the surface morphology in a systematic, if complex, fashion. The account is phrased in the terms with which I am most comfortable, but all of the results are exportable to frameworks with very different semantic, syntactic, and morphological underpinnings. I present my assumptions and apparatus not to define a credo to which others must subscribe, but to allow readers most easily to translate my account into their preferred vocabulary and, only if necessary, to home in on those areas where such substitution is impossible.

2.1. What number features act on. I take number features to compose with lattices, since abstract features are easier to understand when anchored to the concrete pictures to which lattices lend themselves. The discussion is concerned just with atoms and their combinations and, so, could be rephrased set-theoretically, in terms of power sets and their subsets. The theory is therefore exportable to any framework that has devices equivalent to considering all ways of combining a set of atoms.

The slice of lattice theory used below can be illustrated with pizzas and salads. Pizzas generally consist of a base (with sauce, cheese, herbs) and a variety of toppings. If I have tomatoes, onions, and peppers to choose from, then I have eight possible pizzas: three with one topping (tomato, onion, or peppers), three with two (tomato and onion, tomato and peppers, or onion and peppers), one with everything, and one with none.

[3] The online appendices are available at http://muse.jhu.edu/journals/language/v090/90.1.harbour01.pdf.

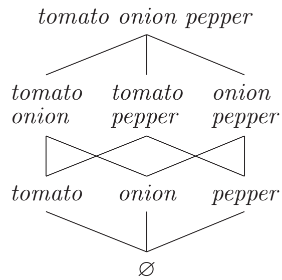


Figure 1. Hasse diagram of the lattice of possible pizza toppings.

These options are laid out in Figure 1 as a Hasse diagram, a common way of visualizing lattices. Here, the rows reflect cardinality: the pizza on the top row has three toppings, all on the next row down have two, all on the next row have one, and the one on the bottom row has none (symbolized by ∅, a toppingless base). Between each row, lines link toppings in subset/superset relation: the element in the top row is linked to every element on the second row, and these two-element sets each link in turn to two one-element sets on the first row, which, again, all link the empty set (a pizza without toppings) on the bottom row. Note that the top element does not link directly to *tomato* on the singleton row, even though *tomato* is one of its ingredients. Instead, the linking proceeds via the elements that contain *tomato* on the intervening row (formally, the diagram depicts the transitive reduction of the set).

'Pizza lattices' are isomorphic to the structures I assume for (numberless) first- and second-person pronouns (Figure 2), supposing a pronominal ontology of a unique speaker, *i*, a unique hearer, *u*, and multiple others, *o*, *o′*, *o″*, … . For instance, Classical Chinese (Norman 1988) *wú* 'I, we' corresponds to the leftmost lattice of Fig. 2: *wú* can refer to the speaker *i* (English *I*) or to any meaning covered by English *we*: the speaker-hearer dyad *iu*, the speaker-hearer dyad plus a third person *iuo*, the speaker and two third persons without the hearer *ioo′*, and so on. Here, *i* is the pizza base, to which other ingredients (*u*, *o*, *o′*) are singly or multiply added. Similarly, *rǔ* 'you', which corresponds to the rightmost lattice, covers the hearer alone *u*, the hearer with a single other *uo*, the hearer with two others *uoo′*, and so on. Again, *u* is the base to which others are added.

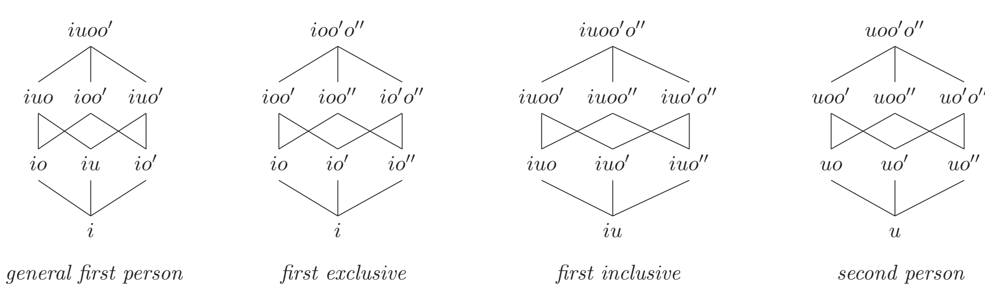


Figure 2. Hasse diagrams of pizza-isomorphic pronominal lattices.

In a language with a clusitivity contrast, such as Aymara (Hardman 2001), the basic pronominal structures are the second, third, and fourth lattices of Fig. 2. For the exclusive (the second lattice), *naya* 'I, we.EXCL', the common base is *i*, to which any number of *o*s can be added—though not *u*, in contrast to Classical Chinese. Instead, for combinations including both the speaker and hearer, there is a separate structure (the third lattice), *jiwasa* 'we.INCL', for which *iu* is the base to which any number of *o*s can be added. (The second person, *juma* 'you', of the rightmost lattice works identically to Chinese *rǔ*.)

More specifically, I take these lattices to represent domains over which variables range. Consider Classical Chinese first person *wú*. Covering English *I* and *we*, it could, as already noted, denote *i*, the speaker, or *iu*, the speaker-hearer dyad, or *io*, the speaker and one other, and so on. However, Chinese speakers obviously could use *wú* to refer to a specific first person, *iu* say; they did not refer to all possible values of *wú* on every occasion of use. The distinction between possible denotations and actual denotation on a particular occasion is captured by treating pronouns as comprising both a domain (the possible) and a variable, which may be bound or closed (the actual).

Thus, the Chinese and Aymara pronouns bind variables that range over the relevant structures of Fig. 2. In both languages, first- and second-person pronominal structures consist of a bottom element, *i*, *iu*, or *u*, that is a subset of everything else in the lattice. As in the pizza lattice, these are like the base, to which 'toppings' are added, giving *io*, *iuo*, uoo', and so on.

Third person differs from first and second in the same way that salads differ from pizzas. We can make salads, like pizzas, from tomatoes, onions, and peppers, but there are only seven possible salads, as opposed to the eight possible pizzas. A pizza without toppings is a margherita, but salad without ingredients is just a plate. Analogously, a first person with no others is just me, but a third person with no others is not a person at all. The lattice of possible salads is shown on the left of Figure 3. Its bottom element is missing. It is isomorphic to the structures for third-person pronouns (Classical Chinese *qí* and Aymara *jupa*) and those for common nouns, like *cat* (Fig. 3). (Technically, these are atomic complete join-semilattices, but, for simplicity, I call them lattices too.)

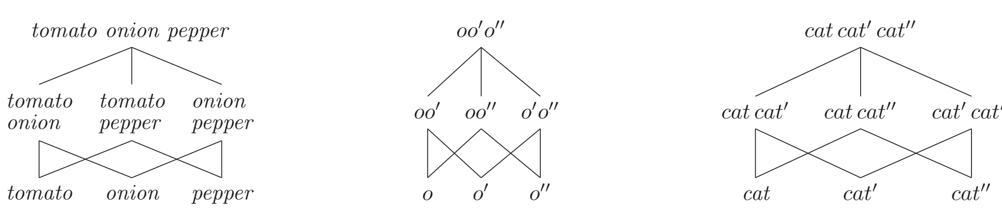


FIGURE 3. Hasse diagrams of the lattice of possible salads, third-person pronouns, and cats.

2.2. FEATURES AS DOMAIN RESTRICTIONS. If lattices are the domains of variables, then numbers are restrictors on variables. For instance, singular restricts variables to range over atoms. Hence, the singular of first person (general or exclusive) is just the atom *i*. A third-person singular is a variable over o, o', o'', and so on (but not over oo', oo'', etc.). Dual, by contrast, restricts variables to elements of the form *xy*—hence, to *iu* for first-person inclusive; to elements of the form io, io', io'' for first-person exclusive; to the union of these for general first person; and so on.

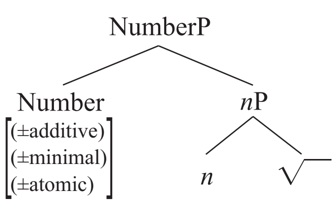
Of course, languages differ as to which numbers they exhibit and how they express them. English and Modern Chinese both distinguish singular from plural. For English, the distinction applies overtly to almost all count nouns (*cat*, *cats*) but is morphologically irregular in the first person (*I*, *we*). In Modern Chinese, the distinction is more restricted, but is morphologically uniform across pronouns (e.g. *wǒ* 'I', *wǒmen* 'we') and those count nouns that express it (*lǎoshǐ* 'teacher', *lǎoshǐmen* 'teachers') (Chappell 1996). Other languages exhibit dual, or paucal, or greater plural, with different ranges and differing degrees of morphological regularity (see Corbett 2000 and Cysouw 2003 for excellent overviews).

To reiterate §1, the challenge for a theory of number is to articulate a set of semantic primitives (features) that apply compositionally to nominal and pronominal lattices alike to yield only attested numbers (e.g. no quadral), and only within attested number systems (e.g. no dual without singular and plural), while at the same time capturing the natural classes attested by morphological compositionality and related diachronic and synchronic phenomena.

2.3. SYNTAX. In proffering my own account, I make some specific assumptions about both the structure that my three bivalent features occupy and the syntax that gives rise to the lattices that they act on.

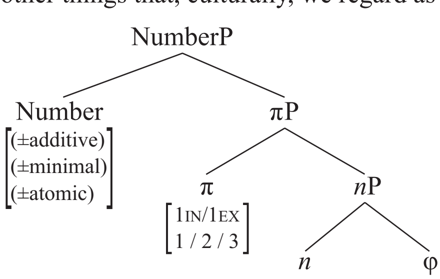
First, I assume that number features are collocated under a single head, which I call Number^0 following Harbour 2007. This head forms part of the extended projection of the nominal, an idea based on Grimshaw 2005 and recently thoroughly explored by Borer (2005a,b, 2013). Following Borer (and Marantz 1997, Adger 2012), I take these extended projections to begin with a category-forming head, in this case, *n*, which dominates an acategorial root. Most simply, the structure for common nouns is as in 5.

(5)



For personal pronouns, I assume an additional projection, π, that specifies person. (I abstract away from the actual feature content of π^0 and use glossing labels here.) I take φ to be a special root denoting the domain of animates (loosely construed to include cats, gods, and other things that, culturally, we regard as our pronominal kin).

(6)



I assume that categories that agree in number with nouns or pronouns generally just replicate the feature content of Number^0 (the mechanism of this relationship is immaterial below).

Following Adger, Borer, and Marantz, I assume that roots name concepts and that n^0 makes concepts 'nouny', structuring them as lattices. However, *n* underdetermines whether that lattice has an atomic stratum or whether its subparts have ever smaller subparts—that is, whether it is count or mass. It is Number that actually introduces the variable[4] and that constrains it, and hence the lattice, to range over atoms and their combinations (a subtlety not reflected in the foregoing figures).[5]

Not all of the features under Number are present for all nouns in all languages. In Classical Chinese and Aymara, Number is empty: all Number does is introduce a variable over individuals and their combinations. In English and Modern Chinese, at most [±atomic] is present. In languages, like Ilocano (Thomas 1955) and Winnebago (Lipkind 1945), with a minimal–augmented number system, only [±minimal] is present. In languages with singular–dual–plural, like Kiowa or Warlpiri, both [±atomic] and [±minimal] are present (Noyer 1992, Harbour 2007, 2011b, Nevins 2011). Languages like Classical Chinese have what is termed 'general number'.[6]

I assume that feature bundles are sets. For instance, [+minimal −atomic] stands for {+minimal, −atomic}. The assumption that number features occupy sets is of importance in §4.2, where it caps the quantity of features that can cooccur (via the axiom of extension). This is a particularly natural assumption in theories where the fundamental structure-building operation is set-theoretic (e.g. Chomsky's 2004 'Merge'). Nonetheless, other assumptions can achieve the same result. Accounts where each feature projects separately as its own head might apply the axiom of extension within the numeration, should they posit one. Alternatively, they might suppose, following Adger 2012, that functional sequences must always go onward and upward, that they cannot reiterate one head or repeat an earlier one.

This last assumption about functional sequences is also adopted below. Specifically, I assume that the only way of having multiple occurrences of Number^0 within the syntax is to begin the whole functional sequence afresh, by reprojecting n^0. This has concrete semantic consequences, however: the first set of number features modifies the first nominal lattice; the second n^0 creates a new nominal meaning from the output of that modification, and that is what the second set of features modifies. Section 5.3 explores these possibilities, and §4.4 examines how a nonlinguistic syntax might circumvent them to generate the integers.

The interpretation of feature bundles is discussed at length in §§3.3 and 4.3. Foreshadowing, a feature bundle, like dual [−atomic +minimal], applies successively to a lattice, L, giving (+minimal(−atomic(L))). The interpretation proceeds by standard semantic means. When Number^0 introduces variables that range over lattices, it produces objects of type $\langle e, t\rangle$. As defined in 20, [±atomic] is of type $\langle e, t\rangle$. So, [±atomic] composes via function modification. As defined in 10 and 21, [±additive] and [±minimal] are of type $\langle\langle e, t\rangle, \langle e, t\rangle\rangle$ and so compose via function application. (Nothing hangs on this type difference; see n. 15.)

This setup may seem to require extrinsic stipulation as to which feature composes with the complement lattice first (given that types permit any feature to compose with a

[4] The role of Number in introducing a variable specifically over individuals may seem redundant in the case of φ, if we conceive of this solely as a pronominal root. In Harbour 2012, however, I allow φ to be surmounted by a spatial head, χ, that generates deictic systems, like *here–there*. Since spatial deictics do not have obvious atomic foundations, the role of Number is nontrivial here too.

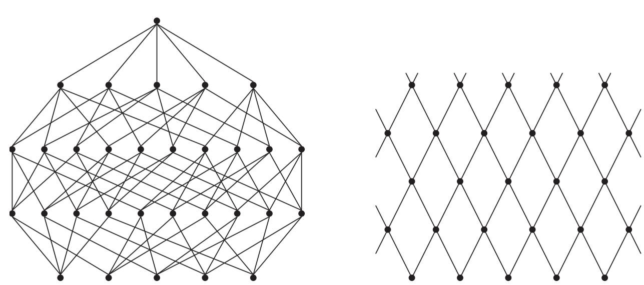
[5] I have little to say about mass nouns in the current article, but clearly Number is not present for mass nouns, which lack a foundational atomic stratum. Nonetheless, I do not rule out the possibility that the features on Number might also play some role in mass nouns. See, for instance, Harbour 2007, Tsoulas 2009, and Ouwayda 2013 for case studies and data.

[6] General number may coexist alongside other number distinctions, like singular–plural (see Corbett 2000 for examples). Facultative use of a number may constitute another variety of parameter, but I do not explore the issue here.

lattice variable or with the $\langle e, t\rangle$ output of such composition). However, §4.3 shows that order of composition can be derived from minimal assumptions about how learners activate features in response to the richness of their input data.

My assumptions about pronunciation are rather permissive. I do not assume (as in much work in DISTRIBUTED MORPHOLOGY, for instance) that morphemes are inserted only at terminal nodes, that terminal nodes may host at most one morpheme, or that morphemes that realize multiple heads are parasitic on head movement. Instead, I assume that any contiguous part of a structure is a legitimate target for pronunciation. At its smallest, exponence may target a single feature, and in numerous places below I assume that distinct exponents target different features within the same Number^0. At its largest, exponence may target a whole subtree. For instance, I assume English *we* realizes a version of the whole tree in 6, or some even larger structure, with Number hosting –atomic and π hosting 1.

2.4. REPRESENTATION AND NOTATION. Having gone through all of these assumptions, it may seem that I demand very specific theoretical commitments from my readers. So, I reiterate that I do not. At its core, the current proposal is about the primitives of number, and most of the discussion, even if couched theory-specifically, is broadly neutral across a wide selection of approaches. This applies to the semantic apparatus, its syntactic embedding, and the correspondence between syntax and sound.



Complete structure for five atoms.

Simplified structure of unspecified size (with five atoms explicit).

FIGURE 4. Full and simplified Hasse diagrams of larger lattices.

As a last preliminary, I lay out some notation. Talk of paucals requires larger lattices than those used so far. However, full Hasse diagrams become very complex when several stories high. For ease of exposition, I use simplified diagrams that show the joins only of some atoms of the lattice. A comparison is given in Figure 4. In addition, I frequently omit diacritics from the *o*s or, even more simply, omit labels all together (as in Fig. 4), using dots as stand-ins for more contentful labels.

Finally, I use brackets as follows. When I talk about features as parametric settings that a language may choose, I use curly brackets: {±atomic, ±additive}. This notation is intended to emphasize that I regard parametric options as freely chosen sets, unrestricted by geometries or filters (§§5.1, 6). Although syntactic feature bundles are also sets, I distinguish them from parametric choices by enclosing them in square brackets and omitting commas between features: [−atomic +additive]. And when presenting semantic formulae, I enclose expressions in parentheses. For instance, (+additive(−atomic(P))) represents (+additive) applied to the result of applying (−atomic) to a person or nominal property P. Sometimes, a syntactic feature bundle may be embedded within a semantic expression, usually to abstract away from order of composition of multiple features. For instance, ([−atomic +additive](P)) represents the application, to P, of the interpretation of [−atomic +additive], whatever its order of composition.

And now we are ready to proceed.

3. A FEATURE FOR APPROXIMATIVE NUMBERS. The starting point for the treatment of approximative number that follows is an observation that goes back to Eubulides of Miletus, a student of Euclid's. In the SORITES PARADOX, from the Greek for 'heap', Eubulides famously observed that a heap minus a smidge is not always a heap. Beginning with the paucal, §3.1 uses the notion of boundedness implicit in the paradox to characterize the difference between paucities and pluralities and to define the feature [±additive]. Section 3.2 then introduces the basic parameter and the convention governing [±additive]: whether languages use the feature (i.e. have approximative number) at all, and, if so, whether the approximative number is paucal, greater plural, or other. Finally, §3.3 introduces the parameter of feature recursion, which permits the specification [+additive −additive] on Number^0. The section explains how composition by function application saves this specification from being contradictory and then examines the types of systems with two approximative numbers that it permits.

3.1. FROM CONCEPT TO FEATURE. In its traditional form, the sorites paradox is phrased in terms of subtraction, focusing on the point when an abundance becomes a paucity. Turning the paradox around, however, we can say that the sum of two pluralities is always a plurality, but the sum of two paucities is not always a paucity. In more technical language, the plural is closed under addition; the paucal is not. This simple truth is the Archimedean point about which one can construct a featural characterization of approximative numbers, beginning with the paucal.

In lattice-theoretic terms, additive closure is expressed as join completeness; that is, if two elements of the lattice are joined together, they yield a third point that is also within the lattice.

(7) P is JOIN-COMPLETE if and only if $\forall x\ \forall y\ ((\mathrm{P}(x) \wedge \mathrm{P}(y)) \rightarrow \mathrm{P}(x \sqcup y))$.
(The sum of any two elements of P is in P.)

By definition, all lattices are join-complete (everything links to the top element in Figs. 1–3). To define a feature that characterizes paucal versus nonpaucal pluralities, we must transform 7 from a statement into a property that can be attributed or denied so as to divide lattices into regions. (Readers wishing to avoid some semantic detail may skip to 10.)

As a first step, we lambda abstract over the predicate, so that we can have a paucal of any given predicate (*cat*, *dog*, *you*, etc.). In addition, we must also lambda abstract over one variable, so as to talk about particular paucities (of cats, dogs, you, etc.).

(8) $\lambda \mathrm{P}\ \lambda x\ \forall y\ ((\mathrm{P}(x) \wedge \mathrm{P}(y)) \rightarrow \mathrm{P}(x \sqcup y))$
(The set of elements that, if in P, stay within P no matter which other element of P you add to them.)

However, this will not split a lattice region into paucal and nonpaucal parts. Rather, it will pick out every *x* in P if P is join-complete, and no *x* otherwise. It is all or nothing, rather than paucal versus plural. Moreover, it is satisfied by elements not in P, since these make the antecedent of the conditional false and so the conditional itself true—which is like saying that a plurality of dogs is an instance of a paucity of cats.

Fixing the first problem, the formula must introduce a subregion of P and then characterize that region as being join-complete or not. In consequence, the formula requires two predicate terms: it takes one predicate P and yields a subregion, Q, of its lattice denotation. Join completeness of the subregion is then what is at issue.

(9) $\lambda P\, \lambda x\, (Q(x) \wedge \forall y\, (Q(y) \rightarrow Q(x \sqcup y)) \wedge Q \sqsubset P)$
(The set of elements of a join-complete proper subregion of P.)

Q is a contextually supplied free variable, and $Q \sqsubset P$ means that everything in Q is in P but not vice versa: $\forall z(Q(z) \rightarrow P(z)) \wedge \neg\forall z(P(z) \rightarrow Q(z))$.

Finally, given that we are defining a property that can be denied (as well as asserted), only the correct parts should fall within the scope of negation. The negation should characterize points that are within the join-incomplete subregion of P. It should not characterize points that, say, are not in the lattice at all (the cat-dog problem). That is, $Q(x)$ and $Q \sqsubset P$ must be beyond the scope of negation, since they force x to be within both Q and P. Hence, we treat the predication of Q and the inclusion of Q in P as presuppositions.

(10) $[\pm\text{additive}] = \lambda P\, \lambda x\, (\neg)\forall y\, (Q(y) \rightarrow Q(x \sqcup y))$
presuppositions: $Q(x)$, $Q \sqsubset P$
(The set of elements of join-(in)complete subregion P.)

The parenthetic negation signifies that ¬ is present for the minus value, absent for plus.

If features could speak, this one would say: ‘Give me a lattice region, and I’ll give you back the set of points that comprise a subregion of that region’. Ask what is so special about that and the feature would respond: ‘Assert me and the subregion will be join-complete: add any point of the subregion to any other and you’ll still be in the subregion. Deny me, and the subregion will be join-incomplete: sometimes adding points will keep you in the subregion, sometimes it won’t’.

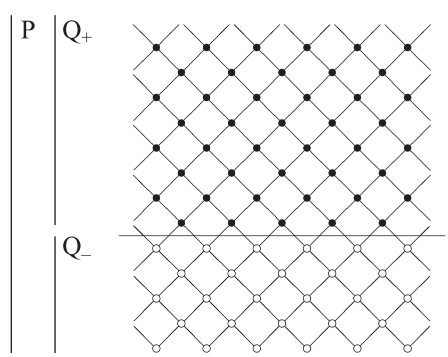


FIGURE 5. A bounded region, Q_-, and an unbounded top, Q_+.

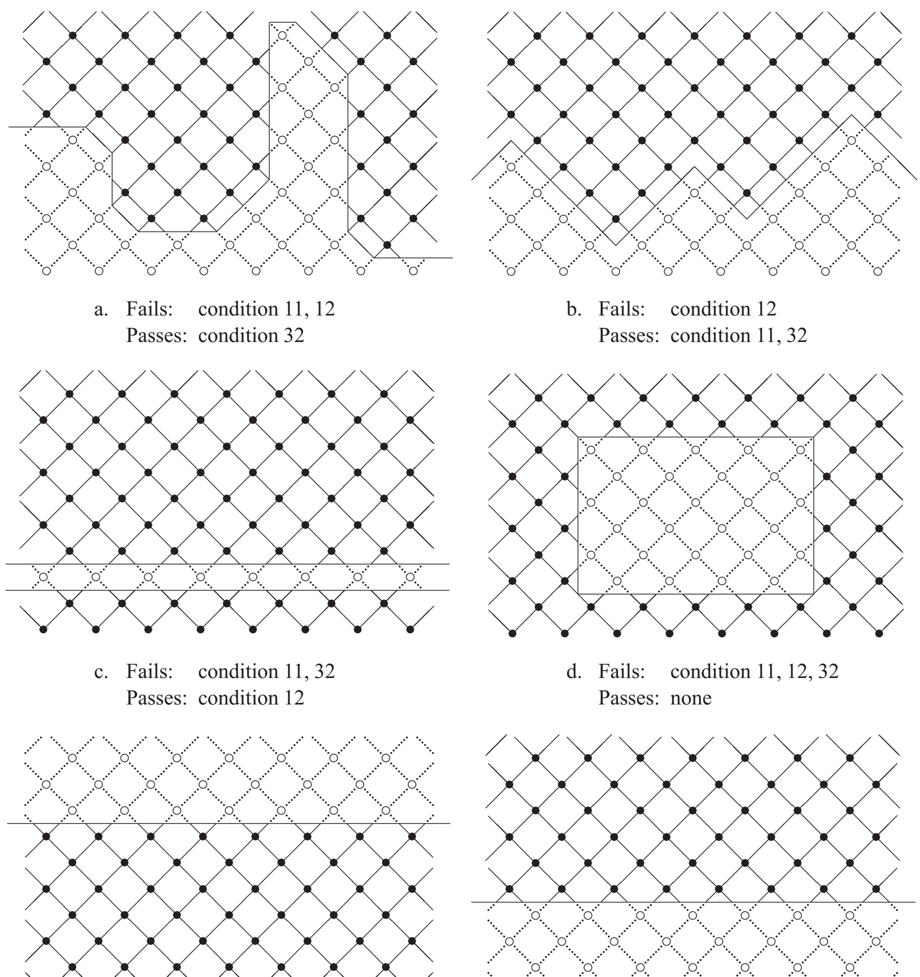
As this paraphrase suggests, (+additive) defines a different region from (−additive). For the latter, the Q that 10 introduces is a bounded region; for the former, Q is an unbounded region. Figure 5 illustrates a plural Q and a paucal Q, for a given P, arising from (+additive(P)) and (−additive(P)), respectively. To emphasize that these are different Qs, they are labeled according to the feature that defines them, as Q_+ and Q_-.[7] It is

[7] This is different from more familiar number features: (+atomic) picks out the atoms of a lattice, and (−atomic) picks out its complement; (+minimal) picks out the lowest stratum of a lattice region, and (−minimal) picks out its complement. Figure 5 represents choices for Q_+ and Q_- for which this complement relation holds, but other values could have been chosen.

My diagrams of paucals also portray a specific upper bound for the paucal, a reflection of my limited drawing skills. In practice, two sources for approximative numbers’ approximativeness are imaginable, vagueness versus variability. They might be inherently vague as to their cut-off point; or, within any model, they might have a specific cut-off point, but with that specific point varying from model to model. See Chierchia 2010 for in-depth discussion of vagueness and nouns.

important, therefore, not to think of Q as the paucal. It can be either a bounded or an unbounded region, depending on the value that introduces it.

Given Fig. 5 and how we arrived at 10, [±additive] appears to be constructed to characterize paucal versus nonpaucal pluralities. However, a moment's reflection shows that it characterizes much else besides. As Figure 6 illustrates, many different Qs are join-incomplete subregions of P but are not paucal. So, if join incompleteness is to yield a characterization of the paucal, rather than of assorted random cuts, we need a means of ruling out, for instance, the arbitrary (Figs. 6a–b), the purely triadic (Fig. 6c), the excessively specific (Fig. 6d), and the not top- but bottom-bounded (Fig. 6e) as legitimate values of Q, leaving only cuts as in Fig. 6f, a paucal where the cut-off point happens to be four. Two conditions achieve this. (Condition 32, discussed in §4.5 below, is included in Fig. 6 for the sake of completeness, but it may be ignored for the moment.)



a. Fails: condition 11, 12
Passes: condition 32

b. Fails: condition 12
Passes: condition 11, 32

c. Fails: condition 11, 32
Passes: condition 12

d. Fails: condition 11, 12, 32
Passes: none

e. Fails: condition 11
Passes: condition 12, 32

f. Fails: none
Passes: condition 11, 12, 32

FIGURE 6. Six bounded (pale) cuts of the lattice. Note: condition 32, discussed in §4.5 below, can be ignored when reading §3.1.

The first is implicit in the definition of [±additive], assuming that the complement of any Q_- is a possible value of Q_+.

(11) COMPLEMENT COMPLETENESS: The complement of Q_- must be join-complete.

Concretely, no point in Q_- can be the join of any point(s) in its complement. In terms of Fig. 6, this means that there cannot be a downward path from a white circle to any black one. Figures 6a,c–e fail on this count. Figure 6e is particularly noteworthy: it illustrates that Q_+ and Q_- of Fig. 5 cannot swap places, even though both are legitimate values for Q. Given that Fig. 6e fails complement completeness, it shows that it is not underdetermined as to which region is which.

The second condition follows from the fungibility implicit in (pronominal) lattices. We can think of the predicate that the atoms of the lattice satisfy as providing an equivalence relation. That is, although we can name different cats, say, or distinguish between *this cat* and *that cat*, none of this information is present in the lattice. All that the lattice does is pick out all cats from the world, as opposed to dogs, books, pianos, and so on. Consequently, to be in the lattice is to be in an equivalence class, and members of an equivalence class can be permuted without affecting the class itself.[8] Such permutation invariance rules out any nonhorizontal cut of the lattice (online Appendix B), such as 6a,b,d.

(12) FUNGIBILITY: Q must be permutation-invariant and, so, must be defined by horizontal cuts of the lattice.

Jointly, conditions 11–12 constrain Q_- to be a region below a single horizontal cut, as in Fig. 6f. This is necessary, but not sufficient, to define the paucal, since one still has to stipulate that the cut defined by Q is low. However, this stipulation is not, I suggest, a flaw, but is intrinsic to what gives approximative numbers their approximateness. The positions permitted for the cut are not inherent to the feature, but must be conventionalized by each speech community that uses [±additive] as a feature on Number^0. If its cuts are conventionally low, [±additive] differentiates between low and middling-to-high plurality; if high, between low-to-middling and high plurality. As we now see, this is a desirable result.

3.2. PARAMETERS 1. We have now moved from join completeness (boundedness, additive closure) as a means of characterizing paucities to a feature that defines subregions of a lattice defined by a single horizontal cut. In so doing, it has led to one parameter and one 'sociosemantic convention'. This section lays these out explicitly and shows that they capture what appears to be the correct range of empirical phenomena.

The parameter is a parameter in the conventional sense (Borer 1983). It concerns whether a particular feature occurs on a particular functional head in a given language.

(13) ACTIVATION PARAMETER (restricted version): [±additive] is (not) a feature of Number^0.

[8] To understand the thinking behind permutation invariance, it is useful to envisage near-indistinguishable objects, like cue balls. Suppose ten cue balls are arranged in a pyramid and we choose (i.e. produce a cut isolating) the one at the top. If someone permutes the balls and produces a new pyramid, then picking one at the top (i.e. repeating the same cut) does not guarantee that we will get the same ball. So, the cut is not permutation-invariant. The only way of getting the same set of balls from a single cut, subject to permutation, is to take every single one. Of course, this example is different from lattices in that, if you take every ball in the pyramid, then you take every pair of balls, every triad, every tetrad, and so on, whereas, in a lattice, one can take the atoms and leave the dyads or vice versa. Nonetheless, the example hopefully helps to clarify what it means to make the same cut under different permutations and for cuts to be permutation-invariant.

In order for a language to have an approximative number, this parameter must be active. (Languages, like English, in which it is inactive, have no approximative numbers.) This underdetermines the position of the cut. So the sociosemantic convention is needed.

(14) Sociosemantic convention: The semantic range of the cut defined by [±additive] is subject to social convention.

That is, whenever the syntactic parameter 13 is active, the language—or, rather, the speech community—must conventionalize its interpretation, since this is constrained, not determined, by the feature itself. This approach permits one feature to characterize three distinct numbers—paucal, greater plural, and global plural (discussed immediately below)—and predicts cooccurrence restrictions between them (§5.1).

The remainder of this section illustrates the range of cardinal values that a speech community may settle on in establishing an approximative number. This variation is not limited to the difference between the descriptive labels 'paucal' and 'greater plural', say. It also affects the size of 'paucal' or 'greater plural' in different languages.[9]

Beginning with the paucal, a language in which it has a rather restricted cardinal range is Koasati. Kimball (1991:403, 449) writes that the 'few nouns' that may take a paucal suffix do so for the meaning 'two or three' or 'more than two but less than five or six'.

In Yimas, where the paucal extends to all humans, its numerical range is slightly greater. Foley (1991:216) says it is '[p]rototypically, … three to five individuals', though its meaning is, more generally, 'a few; from three up to about seven, but variable depending on context'.

In Boumaa Fijian, the range is greater still. Dixon (1988:52) says that '[t]here is no fixed number of people below which it is appropriate to use a paucal pronoun and above which a plural should be employed'. However, 'a plural must refer to *more* participants than paucal'. For instance, '[p]aucal is used when addressing one-third of the adult villagers (twenty or so people), but plural when referring the whole village (perhaps sixty adults)'. This almost suggests that the Boumaa paucal has been conventionalized to mean 'proportionally few', as well as 'cardinally few'.

In the context of numerals, languages may conventionalize a very precise cut-off point for the paucal. In Russian, it is limited to complements (in nonoblique cases) of numerals ending in two, three, and four, a cardinal range reminiscent of Koasati. In Arabic (Ojeda 1992) and Biak (Suriel Mofu, p.c.), the cut-off point is ten, after which the plural is used, a fact doubtless related to the languages' decimal counting systems. I assume that [±additive] is, in these languages, a selectional feature on numerals. In Russian, it has only this function. In Biak, it is governed by numerals but may also occur independently (as in n. 9). When Biak paucals occur without numerals, I take it that the strict decimal cut-off no longer applies. If it did apply, then use of the paucal—in situations when nine, ten, or eleven fish, say, suddenly escape from a net containing scores—would rely on an ability to make immediate judgments about exact cardinali-

[9] The reality of the syntactic parameter requires little comment beyond the observation that some languages have approximative numbers and some do not, except that one might wish to see agreement for approximative numbers to demonstrate that [±additive] is really in the syntax, rather than packaged away within the lexical semantics of a quantifier (cf. the distinction between dual and *two*). Some instances are found in Banyun ('Noun phrase modifiers such as adjectives agree, distinguishing the various singular, plural and greater plural classes'; Corbett 2000:31), Yimas (see examples 44–45 below), and Biak, as in (i).

(i) Sinan kovan-sko-ya sko-ra sk- ún wós anine
ancestor our- 3pc-the 3pc-go 3pc-take word this
'Our ancestors went there and took this word' (van den Heuvel 2006:447)

ties that fall well beyond the capacities documented in the literature on subitization that begins with Kaufman et al. 1949.

Corbett (2000:22–25) gives a good number of other values for other languages, including conflicting values for the same language (from different sources). The range of values that the paucal may cover within and across languages strongly suggests that it is correct not to rigidify the cardinal range as part of the inherent meaning of the feature, but to allow different speech communities to subject it to their own conventions.[10]

As the cardinal range of the cut progresses higher, we cease to have a paucal and arrive at what is descriptively termed a greater plural. This is found, for instance, in Fula, for which Anderson (drawing on sources for several dialects, 1976:123) describes '*pucciiji* "a great many horses" ' as a ' "plural of abundance", i.e., an emphatic plural', in contrast to *pucci* 'horses'. Schuh (1998:199) too describes 'plurals of abundance' for Miya. His examples are in 15.

(15) *sə̀bə* ~ *sə̀babáw*	'people' ~	'large number of people'
kùtə ~ *kùtatáw*	'things' ~	'large number of things'
cùw ~ *cùwawáw*	'goats' ~	'large number of goats'

And, describing Mokilese, Harrison (1976:89) writes that 'remote pronouns refer to groups of people, usually large, and most of which are probably not directly present'. For instance, first-person inclusive *kihs* 'is often used to refer to all the people of Mokil, or to the whole human race' (p. 90). In a related vein, remote pronouns are also used in generic habituals (pp. 90, 194).

In characterizing the whole human race, the Mokilese greater plural leads to another value, namely, the global plural. This is the complete join of the lattice and, so, is the highest position for the cut. Global plurality is attested in Kaytetye. Citing correspondence with Harold Koch, Corbett (2000:33) reports that Kaytetye 'nouns split the plural into a normal plural marked with the suffix -*amerne*, and a greater (global) plural ("all the X in the universe of discourse"), marked with the suffix -*eynenge*'. Corbett cites Lydall's (1976) discussion of Hamer as providing another potential example.[11]

Examples of global plurals appear to be fewer and less well described than those of greater plurals, which in turn seem to be fewer and less well described than paucals. Corbett concludes: 'The evidence [for numbers above plural] is limited, but it comes from a variety of languages and sources, sufficient to indicate that there is an interesting phenomenon that deserves study. More examples with careful descriptions of their meanings would be welcome' (2000:30). I certainly agree that it would be helpful to have more descriptions of circumstances of actual usage, rather than listings of a few isolated forms with approximate meanings, since such examples frequently do little

[10] For instance, Ojeda (1992:318), whose concern is solely with Arabic, builds the cardinal bound of ten into the denotation of the paucal itself, a move that prevents a crosslinguistically general account of the paucal, let alone other approximative numbers, and is subject to the above-noted oddity of subitizability.

Besides their range in number, paucals, as well as other approximative numbers, vary in which nouns they may or must qualify: the first person in Chocktaw (Broadwell 2006) and Tukang Besi (Donohue 1999), humans in Yimas (Foley 1991), humans and optionally nonhumans in Boumaa (Dixon 1988), and mostly inanimates in Avar (Corbett 2000). I do not address this variation here. It may be too great to deserve theoretical explanation.

[11] Ouwayda (2013) presents comparable readings in the domain of mass nouns. Bittner and Hale (1995) describe plural and global plural as arising from indefinite versus definite construals of a particular nominal modifier in Warlpiri. It is, therefore, possible that global plural is not a different number, but the definite counterpart of (greater) plural. My claim here is merely that, if an approximative number, then global plural is explicable in terms of [±additive]. Nothing commits one to this analysis, however, and the account does not suffer if the global plural turns out to be illusory.

more than tantalize. Nonetheless, the range of examples and systems available supports, I believe, the claim made here, that having approximative number is a parameter and its range is a convention.

3.3. Parameters 2. In addition to the basic syntactic parameter (whether [±additive] occurs on Number^0 at all), I propose that a full theory of approximative number requires a second syntactic parameter, which I term 'feature recursion'. Simply put, this permits opposite values of a feature to cooccur on a head.

(16) Feature recursion parameter (and notation): Both values of [±F] may cooccur on head X^0. (Features so parametrized are starred, [±F]*.)

That is, X^0 may be specified as [+F], as [−F], or as [+F −F]. (See §§4.1 and 5.1, as well as immediately below, for further discussion.) I use the term 'recursion' because, in a sense, (±F) acts on itself, or more precisely, on the output of its own action. Of course, each action corresponds to a different value of ±F, but the sense of recursiveness should still be apparent.

At first glance, [+F −F] seems contradictory, given that (−F) = ¬(+F). However, [±additive] composes with its argument via function application, not function modification. So, ([−additive +additive](P)) is not interpreted as (−additive(P) ∧ +additive(P)), which is indeed contradictory, but as one of either (−additive(+additive(P))) or (+additive(−additive(P))).

Thinking informally, one quickly sees that (+additive(−additive(P))) is unsatisfiable. It takes a join-incomplete part of P and 'looks for' the join-complete proper subregion within it (see online Appendix B).

The alternative formula, (−additive(+additive(P))), by contrast, is satisfiable. After P is divided into join-incomplete and join-complete subregions, (−additive) divides the join-complete subregion, (+additive(P)), into join-incomplete and join-complete sub-subregions. There are consequently two bounded regions, one stacked on top of the other. Figure 7 represents these as partitions. Q_- is the nonadditive subregion arising from (−additive(P)) and Q_+ the unbounded region from (+additive(P)). Q_+ is split into a second join-incomplete subregion, Q′, by (−additive(+additive(P))), and this induces a new, smaller, join-complete complement region, $Q_+\backslash Q'$ (that is, Q_+ minus the subregion Q′).[12]

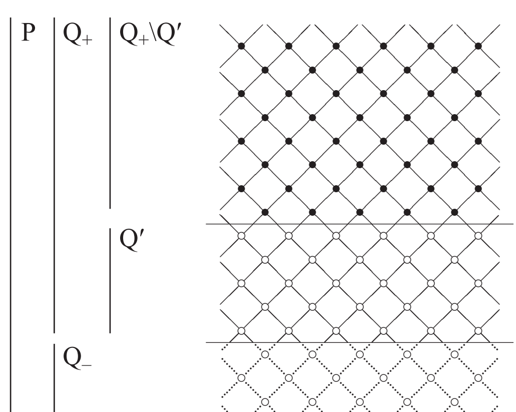


Figure 7. A lesser bounded region, Q_-, greater bounded region, Q′, and an unbounded top, $Q_+\backslash Q'$. Note: the portion of a number system illustrated is paucal (Q_-), greater paucal (Q′), and plural ($Q_+\backslash Q'$). Higher placement of one or both cuts is possible, as discussed below.

[12] The smaller join-complete region cannot be defined by (−additive(−additive(P))), as [−additive −additive …] falls foul of the axiom of extension (§5.1).

We can verify this interpretation in slightly more formal terms by expanding the outer term in (−additive(+additive(P))). Writing the presuppositions associated with the formulae underneath them (as shown on the right), we have 17, where $Q' \sqsubset$ +additive(P) arises from setting P' = +additive(P) in $Q' \sqsubset P'$.

(17) (−additive(+additive(P)))

$$= [\, \underbrace{\lambda P'\, \lambda x\, \neg\forall y\, (Q'(y) \rightarrow Q'(x \sqcup y))}_{Q'(x),\ Q' \sqsubset P'}\,]\, \underbrace{(+\text{additive}(P))}_{Q(x),\ Q \sqsubset P} \qquad \underbrace{\text{formula}}_{\textit{presuppositions}}$$

$$= \underbrace{\lambda x\, \neg\forall y\, (Q'(y) \rightarrow Q'(x \sqcup y))}_{Q(x),\ Q'(x),\ Q' \sqsubset +\text{add}(P),\ Q \sqsubset P}$$

The formula characterizes the set of elements, x, in a proper subregion, Q', of the join-complete subregion, (+additive(P)), of P, and asserts that Q' is join-incomplete. (For further detail, fully expanding $Q' \sqsubset$ +additive(P), and thus explicitly introducing the Q_+ of two paragraphs above and Fig. 7, see online Appendix C, which also addresses the accumulation of presuppositions.)

What feature recursion permits, therefore, is a system with two approximative numbers. Languages that exploit this parameter must again conventionalize the cuts that characterize the two approximative numbers. The following examples give an idea of the range of variation.

First, if both cuts are low, then we have two paucals, one lesser, the other greater. Such a system arises in Sursurunga. Previously described as having trial and quadral (in addition to singular, dual, and plural; Hutchisson 1986), the language has since been shown by Corbett, in correspondence with Hutchisson, to have two paucals. Corbett writes (2000:26–27):

> Besides being used for four, the quadral has two other uses. First, plural pronouns are never used with terms for … kinship pairs like uncle–nephew/niece … and the quadral is then used instead for a minimum of four, not exactly four (Hutchisson 1986:10). The second additional use is in hortatory discourse; the speaker may use the first person inclusive quadral, suggesting joint action including the speaker, even though more than four persons are involved. These two special uses account for most instances of the quadral. … The forms might be better described as paucal.
>
> … The trial will be used for three. But, it is also used for small groups, typically around three or four, and for nuclear families of any size. It is therefore not strictly a trial, rather it could be labelled a paucal (an appropriate gloss would be ‘a few’). … we therefore have two paucals, a (normal/lesser) paucal … and a greater paucal.

This is the type of system represented in Fig. 7 (minus singular and dual, that is, taking the lowest stratum of the lattice to represent triads, not atoms or dyads).

Second, if one of the cuts is low and the other high, then we have a paucal and a greater plural. Corbett exemplifies this with Mele-Fila (via personal communication with Ross Clark), which has singular–dual–paucal–plural–greater-plural. It should be noted that Clark (1975:7) describes the plural pronouns as terminating at about fifteen or twenty, with the greater plural used for larger numbers. Clearly, then, greater plural in Mele-Fila is not as great as in other languages (§3.2): it is greater than the ranges given for Yimas paucal and the Sursurunga greater paucal, but is within the range of the Boumaa Fijian paucal. Such interlinguistic overlap between ‘paucal’, ‘greater paucal’, ‘greater plural’, and the like reminds us that these are descriptive labels, not the objects of explanation. They are simply different conventionalizations of the cuts that [±additive] induces.[13]

[13] Additionally, it is important to understand numerically specified bounds in their ethnographic context (cf. the ‘decimalized’ paucal of Arabic and Biak, §3.2). In our highly numeric culture, which produces *Bil-*

Lastly, both cuts might be high. In honesty, I would have thought this an unusable parameter setting. If plurals range from paucity to normal levels of abundance and if the greater plural measures superabundance, then the greatest plural ought to measure 'super-duper-abundance', numbers beyond usual expectations of what is beyond usual expectations. The notion sounds almost paradoxical. However, some languages seem to evidence the contrast. First, for Banyun, Sauvageot (1967:227–28) gives the following forms, two of which are 'unlimited'.

(18)	ha-sɔl	ba-sɔl	gu-sɔl	ti-sɔl
	SG-bubu	PL-bubu	GR.PL-bubu	GR.GR.PL-bubu
	'bubu'	'bubus'	' "unlimited" bubus'	' "unlimited" bubus'

Despite the identical glosses, the meanings are distinct. Though *gu-sɔl* itself means 'unenumerable' (*pas chiffrable*), *ti-sɔl* 'expresses a higher degree of unlimitedness'.

Similarly, in Warekena, an Arawakan language of Brazil and Venezuela, the following plurals are available (Aikhenvald 1998:302; see also n. 13).

(19)	abida-pe	abida-nawi	abida-penawi
	pig-PL	pig-GR.PL	pig-GR.GR.PL
	'pigs'	'very many pigs'	'very many pigs indeed, so many one cannot count them'

These forms raise obvious questions about morphological compositionality (Aikhenvald segments the rightmost form into *pe-nawi*, the concatenation of other two), addressed in §5.2. For the moment, it suffices to observe that Banyun and Warekena, with Mokilese and Sursurunga, fill out the range of empirical options that a language with feature recursion on [±additive], and hence two approximative numbers, may exhibit.

4. The nature of the theory of number. The foregoing analysis of approximative number has focused only on featural identity. It has said nothing about the kind of theory that best embeds a feature like [±additive] and the parameters governing it, nor has it addressed the cooccurrence restrictions or requirements on approximative numbers. This section presents such a theory and argues that it is extremely minimal.

Harbour 2011b presents an account of the numbers singular, dual, trial, plural, minimal, unit augmented, and augmented in terms of just two features, [±atomic] and [±minimal].

(20) [±atomic] = $\lambda x\ (\neg)\mathrm{atom}(x)$

(21) [±minimal] = $\lambda \mathrm{P}\ \lambda x\ (\neg)\neg\exists y(\mathrm{P}(y) \wedge y \sqsubset x)$
presupposition: $\mathrm{P}(x)$

Informally, (+atomic) characterizes the atomic stratum of a lattice region, if it exists, and (+minimal) collects the bottom layer of any given lattice region, the set of elements that have no subelements within the given region; (−atomic) and (−minimal) are their respec-

lionaire magazine, what counts as numerically few or many is probably quite out of keeping with the quantities and counting practices common in most nonindustrial cultures. A good example of this is Warekena (§5.3), for which Aikhenvald (1998:302) cites one speaker's characterization of the difference between normal and greater plurals as PL 'many … , about 2–6', GR.PL 'very many … , about 40–50', and GR.GR.PL 'very many … indeed, so many one cannot count them'. (Presumably, 'many … , about 2–6' means as 'many but possibly as few as just 2–6'. Aikhenvald confirms that it is not a paucal, but ranges over 'quite large groups' (Corbett's phrase, 2000:37–38). It is used for bones of a tapir and of a dead person (Aikhenvald 1998:301, 310), where 'maximally six' is implausible.) Given that Warekena numerals are borrowings (*kwatru* 'four', *syete* 'seven'), counting probably has cultural connotations that differ from those of potential *Billionaire* readers, and this should be kept in mind when interpreting what counts as, say, 'greater plural' in Mele-Fila.

tive complements. (Despite their seeming similarity, both features are necessary in order to characterize, among other things, the singular–plural versus minimal–augmented number systems of Svan and Winnebago. See Noyer 1992, Harbour 2011b for extensive discussion. On the type disparity, see n. 15.)

The parameters affecting [±atomic] and [±minimal] are simply those affecting [±additive].

(22) ACTIVATION PARAMETER (general version): [±additive]/[±minimal]/[±atomic] is (not) a feature of Number^0.

(23) FEATURE RECURSION PARAMETER: Both values of [±F] may (not) cooccur on Number^0.
(Features so parametrized are starred, [±F]*.)

That is to say, all there is to the theory of number is a feature inventory and free combinatorics. There are no implicational relations between features, no geometries, no filters on outputs. As this section shows, such 'constraints' as the theory needs arise naturally from its mathematical underpinnings or from broader principles of cognitive organization.

A theory with the features [±additive], [±atomic], and [±minimal], plus feature recursion, generates all and only the number values attested crosslinguistically.

(24) NUMBER VALUES ATTESTED AND GENERATED: general, singular, dual, trial, plural, minimal, unit augmented, augmented, paucal, greater paucal, greater plural, global plural

Of course, this does not exhaust what must be said about number.[14] But current coverage is not insignificant, and so the theory that yields it warrants closer inspection.

This section examines five aspects of the theory of number and the place of [±additive] in it. The first three are more narrowly concerned with number alone. The last two consider how the proposed theory of number relates to broader issues of cognition.

Section 4.1 shows that feature recursion is not unique to [±additive], but is shared by [±minimal]. (The section also shows that all parameters and conventions apply to all features unless independently ruled out.) Section 4.2 addresses the issue of potential overgeneration of numbers and shows that the correct constraint is a basic axiom of set theory, the axiom of extension, to which the theory is already subject. Lastly, §4.3 considers the order of semantic composition for systems that activate multiple features. Rather than additional parameters being required to govern semantic composition, the order is shown to follow straightforwardly from the feature semantics: whenever the input data are sufficient to prompt the learner to activate an extra feature, the data are also sufficient to rule out one order of composition, or else the order is immaterial. Consequently, order of composition can be inferred from the feature definitions and the richness of the primary linguistic data.

Moving to more general issues, §4.4 observes that the features deployed for nominal number have all been posited elsewhere, in the literatures on mass nominals and on aspect and telicity. Not only does this demonstrate a fundamental unity within the family of features {±additive, ±atomic, ±minimal}, but it also shows that these semantic primitives constitute a deep nexus between what are, superficially, quite distinct linguistic domains: (pro)nominal number, masshood, aspect, telicity, and the integers.

[14] On kinds of plurals, like collective and distributive, see Lasersohn 1995, Ojeda 1998, Harbour 2011e. On iterated number, such as plurals of plurals, see Ojeda 1992 and §5.3 below. On an approach to mass nouns using the current features, see Harbour 2007, which also discusses pluralia tantum, and §4.4. On the semantics of associatives, see Nakanishi & Tomioka 2004. Further high-level pretheoretic discussion and other references are found in Acquaviva 2008 and, of course, Corbett 2000.

Section 4.5 continues this theme by looking at an apparent disparity between [±additive] and [±atomic], [±minimal]. Only the last two can function as the sole number feature of a language. This makes activation of [±additive] look like a dependent parameter, requiring prior activation of other features, and so this subsection is thematically related to §§4.1–4.3. The restriction on [±additive] follows from general principles of cognition, concerning convexity of basic meanings (defined in §4.5). So, the theory remains free of stipulation and, as in §4.4, reveals the connection between this domain of linguistic theory and broader issues of cognitive structure.

4.1. Feature recursion. To show that we have a unified theory of approximative and nonapproximative number, the two parameters governing [±additive] must apply to [±atomic] and [±minimal] too, or, if they fail to apply, then this must follow naturally from the underlying theory itself and should not amount to stipulative sidestepping of untidy consequences. Here, we confine attention to the feature recursion parameter given in 16. Sections 4.5 and 5.1 address the activation parameter.

Noyer (1992) showed that {±minimal} generates the number system minimal–augmented, and {±minimal, ±atomic} generates singular–dual–plural. Building on this, Harbour 2011b shows that feature recursion yields {±minimal*} (minimal–unit-augmented–augmented) and {±minimal*, ±atomic} (singular–dual–trial–plural). Consider trial, (+minimal(−minimal(−atomic(P)))), as a concrete example. First, (±atomic) partitions P into atomic and nonatomic strata. Next, (+minimal) collects the lowest stratum of the region it applies to, here, the dyads. Complementarily, (−minimal (−atomic(P))) contains everything triadic or larger. When, by feature recursion, (+minimal) applies to (−minimal(−atomic(P))), the lowest layer that it collects contains just triads. That is, we have derived the trial.

Although feature recursion applies to both [±additive] and [±minimal], it does not apply to [±atomic]. This is for a principled reason, however. Unlike the other features, (±atomic) does not 'operate' on lattice regions, but simply characterizes the atomic stratum of a region, should it exist. In consequence, $(\alpha\text{atomic}(\bar{\alpha}\text{atomic}(\text{P}))) = \lambda x\ (\text{P}(x) \wedge \text{atom}(x) \wedge \neg\text{atom}(x))$ is empty: no x is both atomic and nonatomic (with respect to a single predicate P). Nothing prevents language learners from positing {±atomic*}, but they gain no number distinctions by doing so. This, alone, is the reason that this option is, in practice, ruled out.[15]

The significance of this type of result for the treatment of approximative numbers is, most basically, that it shows feature recursion to be not a quirk of [±additive], but a parameter that applies to other feature sets to yield numbers that have fallen beyond the semantic means of various previous proposals. Feature recursion derives the unit augmented, trial, and greater paucal, as well as number systems with two approximative numbers, via a single parameter, without the addition of new features.

4.2. Limits of feature recursion. A natural question, given the existence of feature recursion, is whether any other parameters affect [±additive] and [±minimal] and, in particular, what rules out other potential numbers, such as the 'dyad augmented' and 'quadral', characterized by the feature bundles below.

[15] Inapplicability of feature recursion to [±atomic] is independent of the fact that it is of a different type from [±additive] and [±minimal]. For type parity, we can take [±atomic] = $\lambda\text{P}\ \lambda x\ (\neg)\ \text{atom}(x)$, presupposition P($x$). Hence, we have (i), where $*A \in \{\neg A, A\}$.

(i) $$[\alpha\text{atomic}]([\bar{\alpha}\text{atomic}](\text{P})) = [\lambda\text{P}'\ \underbrace{\lambda x\ *\text{atom}(x)}_{\text{P}'(x)}]\ (\underbrace{\lambda x\ \neg *\text{atom}(x)}_{\text{P}(x)}) = \underbrace{\lambda x\ *\text{atom}(x)}_{\neg *\text{atom}(x),\ \text{P}(x)}$$

This picks no x in P, because it asserts ¬atom(x) while presupposing atom(x), or vice versa.

(25) *[+minimal −minimal −minimal] (dyad augmented)
(26) *[+minimal −minimal −minimal −atomic] (quadral)

The issue is not that these feature specifications fail to yield the appropriate meanings. On the contrary, if the specifications are licit, then the semantics yields these values as a matter of rote. In fact, absence of a dyad augmented is particularly striking: if singular plus singleton and singular plus dyad both exist (dual and trial), and if minimal plus singleton exists (unit augmented), then minimal plus dyad is also expected. The absence is not clearly geographic: languages within northern and northwestern Australia exhibit both unit augmentation and trial. So, 25 and 26 must be ruled out by other means.

Given that subitization (§3.2) peters out around three, where cardinally exact numbers also end, one might seek the source of the limits there. However, subitization makes dyad augmented seem quite reasonable. One sometimes has to think explicitly about counting oneself when counting a group of which one is part. This suggests that, to use a first inclusive dyad augmented, one would only need to count the hearer and two others, which is well within the subitizable range. Indeed, if one does not need to count the hearer, then we would even expect there to be a triad augmented.

Instead of treating these numbers as unthinkable, I suggest that they are unrepresentable. The device that ensures this is not an external constraint on the current theory but part of its mathematical underpinnings. As mentioned in §2.3, I assume that feature bundles are the most primitive of mathematic collections, namely, sets. As such, they are subject to the Zermelo Fraenkel axioms and, in particular, to the axiom of extension, which entails 27 (see online Appendix D).

(27) $\{a, a\} = \{a\}$

So, [+F −F] is the maximum complexity that specification of a single feature can attain. Adding in extra copies of the same feature in an attempt to enrich the semantics is redundant, like pouring water into a full glass.[16]

This makes several predictions about the maximal quantity of approximative and nonapproximative values that a number system may have. Detailed discussion is deferred until the more general issue of typology is addressed (§5.1). For the moment, let us observe, simply, that 27 predicts trial and unit augmented to be the highest exact numbers that can be attained without recourse to numerals, and that a language may have at most two approximative numbers. According to the typological literature, these predictions are correct (Corbett 2000, Cysouw 2003).

The pertinent point is, again, that such constraints as the system needs are not arbitrary or extrinsic, but emerge internally, from the formal underpinnings of the theory, and apply equally to all (number) features, deriving one set of predictions about approximative numbers from [±additive], and another about nonapproximative numbers from [±minimal].

4.3. Order of composition. When feature recursion was introduced (§3.3), it was pointed out that there is no need to stipulate an order of composition between (−additive) and (+additive). Rather, semantics effectively rules one order of composition out. This fact is not a peculiarity of [±additive], but holds of the system more generally. That

[16] Recall that §2.3 laid out some related alternatives that still cap number systems at the same level of complexity. One can deny that features occupy sets and have them project individually, as part of the functional sequence. If this approach is coupled with a minimalist numeration, and if the numeration is also regarded as a set, then the axiom of extension simply applies there. Alternatively, instead of appealing to a numeration, one can follow Adger's arguments (2012) that functional sequences cannot repeat a single head or return to an earlier one. (Given feature recursion, these alternatives would require +F and −F to be distinct functional projections.)

is to say, whenever the data are rich enough to prompt the learner to activate multiple features, they are also rich enough to determine the order of composition (because one order is redundant), or else the order of composition is immaterial and learners might adopt either.

The significance of this result bears emphasis. A potential peril of multiple features is that unattested interpretations might arise if the features are composed in the 'wrong' order. The current theory is free of this problem, however, in that, no matter the order of interpretation, no unattested systems arise. The only 'error' a learner might make is to posit an order of composition that causes a richer feature system to produce an inventory of numbers that could be achieved with a poorer system. Such aberrant grammars do not arise given two reasonable assumptions about how learning proceeds: that learners only ever activate an *n*th feature when they observe that their previous $n-1$ features do not account for all their input data, and that they try out different orders of composition and maintain a feature set cum order of interpretation only if it improves their match with the input data.[17] A learning procedure that conforms to these constraints rules out semantically redundant orders of composition since these fail to improve the match with observed data.

Consider first (+minimal(−minimal(P))), the case most similar to the one just mentioned. Whereas the plus-minus order yields the unit augmented or trial (depending on the value of P), in the reverse order, it fails to pick out any elements at all. In (−minimal(+minimal(P))), (−minimal) takes the region of (+minimal) elements, which, by definition, lack subelements in the region, and 'looks for' those with such subelements. So, as with [±additive]*, the choice between orders of composition can be left to the learning procedure.

This principle extends to orders of composition between distinct features. Consider (±atomic(±minimal(P))). As per §4.1, (+minimal) gathers the bottom layer of P. So, (−minimal(P)) cannot contain any atoms of P. As a result, (+atomic(−minimal(P))) is empty, and (−atomic(−minimal(P))) is simply the whole of (−minimal(P)). Similarly, (±atomic(+minimal(P))) yields the same distinctions as (+minimal(P)) alone.[18] Hence, if the data are rich enough to cause the learner to activate [±atomic] in addition to [±minimal], then they are also rich enough to cause the learner not to posit an order of composition, (±atomic(±minimal(P))), that makes [±atomic] redundant and vitiates the motivation for its activation.

The same freedom of order of composition extends, by identical reasoning, both to {±additive, ±minimal} and to {±additive, ±minimal, ±atomic}. Consider, first, the order in which (±additive) is the outer feature. Both for (±additive(−minimal(P))) and for (±additive(−minimal(−atomic(P)))), (±additive) partitions the (−minimal(P)), or (−minimal(−atomic(P))), region into an approximative number and plural. If [±atomic] and [±minimal] take any other values, then they pick out regions (singular, dual, minimal)

[17] I assume that learners test uniform orders of composition, that is, (αF(βG(…))) for all values α, β, rather than, say, (+F(βG(…))) but (βG(−F(…))) for +F versus −F. Correspondingly, I assume that order of interpretation is invariant in the mature grammar (unless the input data are peculiar in a way that I cannot imagine).

[18] Specifically, by 12, (+minimal(P)) contains all atoms of the lattice or none. Moreover, it follows from the definition of [±minimal] that, if (+minimal(P)) contains the atomic stratum (*a*, *b*, *c*, …), then it cannot contain anything else (*ab*, say), as that something else will contain an atom ($a \sqsubset ab$) and so will be (−minimal). So, if (+minimal(P)) is the atomic stratum, then (+atomic(+minimal(P))) = (+minimal(P)) and (−atomic(+minimal(P))) is empty; and if (+minimal(P)) is not the atomic stratum, then (−atomic(+minimal(P))) = (+minimal(P)) and (+atomic(+minimal(P))) is empty. Either way, (±atomic) is redundant if composed in this order.

that are too small for [±additive] to divide further. So, this order of composition results in familiar, and desirable, systems.

For orders where (±additive) is not the outer feature, I assume, anticipating §4.5 on convexity, that (±additive) does not compose directly with the predicate P. This assumption does not hide any monsters (see Table 2); it merely shortens the exposition. With it, we confine attention to (±minimal(±additive(±atomic(P)))) and omit (±minimal(±additive(P))), (±minimal(±atomic(±additive(P)))), and so on.

Within this order, the simplest cases to consider arise when (±additive) takes the value minus. These yield only numbers already derived above. For instance, (−additive(−atomic(P))) is a paucal that begins after the singular. So, (+minimal(−additive(−atomic(P)))) will, again, gather the lowest stratum of this paucal, the dyads, yielding the dual.

However, when (±additive) takes the value plus, it creates the greater plural region above a plural or the plural region above a paucal. If (±minimal) attempts to compose after this region has been formed, (+minimal) will gather the region's bottom stratum, and (−minimal) will characterize its complement. These numbers, together with feature-recursed counterparts, are laid out in Table 2. Whether one takes them to be usable depends on whether one views approximative as vague or as variable. Either way, no unattested systems result.

If one regards approximative numbers as vague, then these numbers are incoherent. A vague quantity plus an exact quantity is still a vague quantity. So, one cannot specify what counts as exactly one more than a paucity, or exactly one more than abundance. If one could, Eubulides of Miletus would never have posed his paradox about heaps.

Alternatively, one can suppose that, on any particular occasion of use, that is, within any semantic model, each approximative number corresponds to a specific cut of the lattice. (This is what Fig. 7 and its like depict, but chiefly because I cannot think how to draw vague cuts.) This does not help to create a coherent system if the approximative number is greater plural. Greater plural is used precisely when counting is impracticable, so the bottom, (+minimal), stratum of a greater plural is 'precisely the quantity at which things become massively imprecise'. This brings us back to the paradox of Eubulides. Even if usable, I contend that its occurrence would be so severely limited as to make its acquisition impossible (when, in the normal course of events, would we need to refer to such a quantity?).

If, however, (±additive) induces a low cut by sociosemantic convention, then, under some quite particular assumptions, the system that results is, arguably, coherent. Specifically, suppose that the range of paucal–plural overlap is five to seven (in this scenario, three and four are obligatorily paucal; recall, from four paragraphs above, that this system has a featurally unambiguous dual). So, when plural, (+additive(−atomic(P))), covers five and higher, then (+minimal) picks out pentads; when plural covers six and higher, it picks out hexads; and for a plural covering seven and higher, heptads. The resulting system has a paucal from three to seven, and another approximative number overlapping with the upper range of the paucal, from five to seven. Descriptively, this is a (rather restricted) greater paucal.[19] Thus, this order of composition derives a version of singular–dual–paucal–greater-paucal–plural, without recourse to feature recursion.

[19] The greater paucal is, furthermore, facultative, as it is wholly parasitic on five to seven also being in the range of the paucal. A shift away from facultativity would require featural reanalysis (in terms of ±additive*). I leave aside how learnable and hence stable—and hence how findable in searches through grammars—this system would be. Facultativity and greater paucals are both rather rare.

For completeness, we should consider the further effect of adding feature recursion to this last system. The result would be coherent on paper, but probably unusable. Having (+minimal(−minimal(+additive(−atomic(P))))) yields a number that ranges from six to eight. I suspect that this overlap, which lies well beyond the subitizable zone, together with the likely infrequency of signal, would make it unlikely for a learner to acquire two separate (sets of) number morphemes, α for five to seven and β for six to eight (in addition to γ for the paucal three to seven and δ for five and above), and more likely that α and β would be taken as free variant allomorphs over the range five to eight. (Again, see Table 2.) The reverse order of composition between (−minimal) and (+minimal) is incoherent, as discussed earlier in this section: minimal regions do not contain nonminimal subregions. Nonetheless, in the unlikely event that a learner did develop this feature system, with feature recursion on (±minimal) as the outer feature, the learner would still know to compose (−minimal) before (+minimal).

FEATURE SPECIFICATION	NUMBER
(+min(−add(−atomic(P))))	dual
(+min(−min(−add(−atomic(P)))))	trial
(−min(−add(−atomic(P))))	paucal/plural above dual/trial
(+min(+add(−atomic(P))))	‘few/many + 1’
(+min(−min(+add(−atomic(P)))))	‘few/many + 2’
(+min(−add(+add(−atomic(P)))))	‘few/many + 1’
(+min(−min(−add(+add(−atomic(P))))))	‘few/many + 2’

TABLE 2. Order of composition (±minimal(±additive(P))). Note: the left-hand column assumes that (±additive) always composes with (±atomic) (§4.5). Omission of (−atomic) would yield minimal and unit augmented for dual and trial in the first four rows, numbers that are attested in any event.

The preceding discussion requires some clarification. Aberrance of ‘few/many + 1/2’ does not completely rule out (±minimal(±additive(…))) as an order of composition. By way of contrast, consider the system {±minimal, ±atomic}. This generates [+minimal +atomic], which is unusable: it characterizes the set of atoms with atomic subelements. However, this rules out use only of the plus-plus specification itself. The whole feature set is licit and generates the robustly attested singular–dual–plural system. Similarly, the problem of interpreting the bottom rows of Table 2 rules out use only of those feature specifications. Ignoring aberrant values leaves us with two means of composing {±additive, ±minimal$^{(*)}$, ±atomic}, depending on whether ±additive or ±minimal composes last, but both yield the same number systems (e.g. singular–dual–(trial–)paucal–plural). Presumably, learners either prefer grammars that do not produce redundant numbers, or are free to choose either order of composition.

The foregoing covers all major cases. So, it shows that the system does not require extrinsic constraints on how a learner composes features. If learners only ever activate additional features in order to create additional semantic distinctions, and therefore eschew orders that make newly activated features semantically redundant, then correct orders of composition emerge naturally.

Concretely, (±atomic) always composes first, if active; the order of composition for (±additive) and (±minimal) is, potentially, immaterial, unless the cut induced by (±additive) is high and unless learners prefer systems with fewer redundant feature combinations, in which case (±additive) should be outermost; and, under feature recursion, (−additive) and (+minimal) compose after their opposite value. Thus the maximally complex orders are as shown in 28.

(28) ORDER OF COMPOSITION
a. (±additive(±minimal(±atomic(P))))
or possibly (±minimal(±additive(±atomic(P))))
b. (+minimal(−minimal(…)))
c. (−additive(+additive(…)))

Simpler orders arise by omitting features from 28a. For instance, languages without [±additive] have the order (±minimal(±atomic(P))).

4.4. ASPECT AND INTEGERS. The preceding parallelisms have been narrowly focused on the features that constitute number systems and on the naturalness of the constraints that govern them. These principles are not the only factors that underline the unity of [±additive] with [±atomic] and [±minimal]. One such factor is the connection of number features to concepts in the literatures on masshood (Quine 1960, Cheng 1973, Link 1983) and aspect and telicity (Krifka 1992, Ramchand 1997). Another is the emergence of the means for defining a successor-like function and hence for generating the integers, a foundational issue in mathematical psychology. Both of these point to a much more fundamental nexus between the theory of number and the underlying structure of grammatical and cognitive systems.

Researchers in the domain of mass nouns and aspect and telicity have had little concern—nor obvious reason to have concern—with the featural identity of such numbers as dual or trial, unit augmented or augmented, or paucal or greater plural. It is therefore quite noteworthy that the three features that generate all and only the typologically attested number values have been independently posited by researchers whose concerns lay elsewhere.

Naturally, the notion of atomicity, which characterizes singular–plural number systems, was bound to be posited by researchers, like Quine, interested in the 'referential apparatus' of English, or like Link, interested in the similarity of mass nouns to plurals. However, it is genuinely surprising to find a telic notion such as strict cumulativity containing both [±additive] and [±minimal].

To see, informally, how this arises, consider 29 (see online Appendix E for formal details).

(29) a. Dierk ate sauerkraut for an hour.
b. ??Dierk ate sixty portions of sauerkraut for an hour.

The infelicity of durative adverbials with telic predicates as in 29b is well known. According to Krifka (1992:42), '[t]he underlying reason is that durative adverbials presuppose that the verbal predicate they are applied to is strictly cumulative'. That is, in current terminology, 29a is felicitous because it satisfies [+additive] (sauerkraut plus sauerkraut is still sauerkraut, and so an eating of sauerkraut abutting an eating of sauerkraut is just one (long) eating of sauerkraut), and it satisfies [+minimal] (every portion of sauerkraut contains a subelement that is sauerkraut, and so eating that subserving of sauerkraut is a subevent of eating sauerkraut). By contrast, 29b is infelicitous because it is [−additive] (an eating of sixty portions of sauerkraut followed by another such eating is not simply an eating of sixty portions of sauerkraut) and [−minimal] (a proper subevent of eating sixty portions of sauerkraut involves eating strictly fewer than sixty portions, and so is not an event of eating sixty portions).

Striking as the correspondence is between aspect and the like and number features, further investigation is required that is beyond the scope of this article. Two obvious questions are whether [±additive] and [±minimal] can do all the work of their conceptual kin in Krifka's and related theories (see e.g. recently Borer 2013), and whether [−F

+F] specifications enjoy any use in domains like aspect. But these outstanding questions should not distract from the main point: that the formal unity of [±additive] with previously proposed number features extends beyond the domain of number.

A second extension of the use of number features is as a means of generating the integers. Though this application excludes inherently approximative [±additive], it does make further use of feature recursion. We have already discussed the feature composition of singular, dual, and trial. Expressing their denotation slightly differently, we can write them as in 30.

(30) $(+\text{minimal}(+\text{atomic}(\text{P}))) = \{x : \text{P}(x), |x| = 1\}$
$(+\text{minimal}(-\text{atomic}(\text{P}))) = \{x : \text{P}(x), |x| = 2\}$
$(+\text{minimal}(-\text{minimal}(-\text{atomic}(\text{P})))) = \{x : \text{P}(x), |x| = 3\}$

Russell (1919) calls each such equivalence class of pluralities 'the number of a class' and defines natural numbers, like one, two, three, in terms of them.

The axiom of extension prevents any further additions of [+minimal] to Number^0, and, within the syntax of natural language, once Number^0 has projected, the next head in the functional sequence projects and the semantics moves on (§§2.3, 4.2). However, we can imagine a different syntax-like system, in which the only head is Number^0 and the only root is semantically bleached, providing, simply, nondescript countable entities. Then phase after phase can add extra values of [−minimal] nonredundantly, and, semantically, this system expresses nothing other than exact quantities and their complements. The result is 31, where $(-\text{min}^n(x)) = (-\text{min}(-\text{min}^{n-1}(x)))$ and $(-\text{min}^0(x)) = x$.

(31) A SUCCESSOR-LIKE FUNCTION
$(+\text{min}(-\text{min}^n(-\text{atomic}(\text{P})))) = \{x : \text{P}(x), |x| = n + 2\}$ for $n \geq 0$

This defines the numbers two, three, four, … . Taking the definition of one in 30 delivers the whole of the natural numbers.

As with the connection to aspect, this brief foray into mathematical psychology raises questions of its own. The interested reader is referred to Harbour 2011c for discussion of the nonuniqueness of functions like 31, for a derivation of Wynn's (1990) stages of 'one-knowing', 'two-knowing', and so on, and for other issues in the psychological and epistemological literature, such as LeCorre and Carey's (2007) overview.[20] Though these lie beyond the scope of an investigation of the theory of linguistic number, especially one where the primary focus is on approximative numbers, they nonetheless demonstrate the depth and breadth of the resources that the theory offers.

4.5. CONVEXITY AND AN APPARENT DISPARITY. Thus far, the emphasis has been on the similarities between [±additive] and the other number features, [±atomic] and [±minimal]. There is, however, one difference between them: {±additive} is not a legitimate number system, but {±atomic} and {±minimal} are. This, I suggest, has its root in what others have argued to be a fundamental fact about the organization of meaning and cognition, namely that the meanings of atomic terms must be convex (to be defined below).

At first glance, one might expect {±additive} to yield a number system contrasting just paucal and plural. This is correct for third person, but, for first and second, {±addi-

[20] A general issue that arises when numbers are defined in relation to countable objects is what happens when one runs out of things to count. If there are only N things in the universe (or atoms in the lattice), then numbers above N are nondistinct. A traditional move in foundational mathematics is to start counting sets, an infinity of which can be constructed. A compelling alternative, due to Kitcher (1984), is that arithmetic, and other mathematical theories, describe the capacities of an idealized ('competent') agent, one not subject to our contingent physical limitations, such as finite memory or access to a finite number of countable objects.

tive} would generate minimal–paucal–plural, or, in the absence of a clusivity contrast, singular–paucal–plural. To see why, consider the first person (exclusive). This lattice consists of a bottom element *i*, the first-person atom; a dyadic stratum, containing elements of the form *io*, where *o* is a third person; a triadic stratum, containing elements of the form *ioo′*, where *o* and *o′* are distinct third persons; and so on. This lattice in fact has two [+additive] regions. One is, as before, the join-complete upper portion, and the other is the first-person atom itself (Figure 8). This follows because permutation invariance applies nontrivially only to atoms of which there are multiple tokens, like *o* or *cat*. Given the uniqueness of *i*, it is the only member of its equivalence class. And, of course, any unique element is closed under join, as $x \sqcup x = x$, for all x. It is, therefore, its own join-complete region, defined by a single horizontal cut.

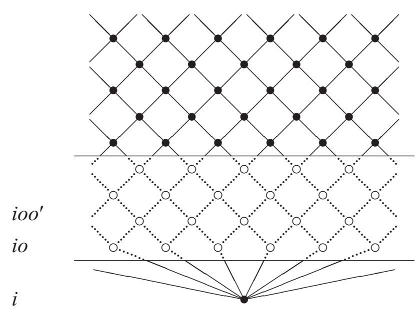


FIGURE 8. The first person and [±additive].

Corbett observes that languages often show more number distinctions for first and second persons than for third. Even so, attestation of the {±additive} system is questionable: if it occurs, it has nothing like the frequency of {±atomic} (singular–plural) and {±minimal} (minimal–augmented), to judge by absence of comment in typological literature.

The source of this disparity is, I suggest, a condition on convexity. That is, the requirement that an atomic term denote a convex cut of the semantic field to which it belongs, where a cut is convex if the path (set of intermediary values) between any two points in the region (or complement) it defines is wholly contained within the region (or complement). If we understand 'basic meaning' to mean 'the meaning of an atomic term', then we can state this more straightforwardly.

(32) CONVEXITY CONDITION: Basic meanings must be convex.

Gärdenfors (2004) has argued for this on extensive logical, theoretical, and empirical grounds, incorporating, for instance, terms for spatial reference and relations (following Zwarts 1995). More generally, it can be shown, using replicator dynamics, that the only evolutionarily stable partition of a semantic space is as a collection of convex cuts (a so-called Voronoi tessellation; see Jäger 2008 for discussion within a linguistic context), and, so, this result extends also to replicator-dynamics-based studies of other semantic fields, such as color (Regier et al. 2007), another topic that Gärdenfors addresses.

Convexity has been well studied for lattices (e.g. Grätzer 1978).

(33) DEFINITION: CONVEXITY FOR LATTICES: A lattice region L is convex if and only if $c \in L$ whenever $a, b \in L$ and $a \sqsubseteq c \sqsubseteq b$.

Various cuts of the lattice to be excluded as values for Q of [±additive] are, in fact, nonconvex (e.g. Figs. 6a,c,d). Indeed, in lattices without a bottom element (Fig. 3, or upper portions of structures in Fig. 1), any nonconvex region either fails permutation invariance, contra 12, or has a join-incomplete complement, contra 11, and thus is inadmissible as Q (online Appendix F). So, in addition to Gärdenfors's, Jäger's, and others' arguments for the relevance of convexity, there are factors internal to the current approach 'conspiring' to rule out nonconvex cuts of the lattice.

Applying this to the case we were just considering, the first person, we can see from Fig. 8 that [±additive] produces a nonconvex cut of the lattice: between the [+additive] first-person atom and any [+additive] first-person plural, there must lie a [−additive] first-person paucal.

If convexity is a condition governing the way in which semantic spaces may be partitioned, then it is straightforward to assume that learners are sensitive to this and, in consequence, do not activate [±additive] as a sole number feature. Moreover, they may use this knowledge to constrain the order of semantic composition of feature bundles containing [±additive], as assumed in §4.3 and Table 2.[21]

5. Empirical analysis. Having considered the theoretical foundations of the proposed account of (approximative) number, we now turn to its empirical support, which, I suggest, is substantial. Section 5.1 examines the typology of number systems that the account predicts and shows how it derives universals in the style of Greenberg 1966. Section 5.2 shows how the feature composition of the approximative numbers captures facts akin to the well-known composed dual of Hopi (Hale 1997). And §5.3 considers two patterns of diachronic development involving reconfigurings of the featural identity of number morphemes, including the well-known phenomenon of trial-to-paucal shift in Austronesian.

5.1. Typological analysis. Harbour 2011b provides the following rubric of explanation for Greenbergian implications.

(34) Typological implication schema: If category A must cooccur with category B in a (domain of a) language, then the parameter setting for A generates B.

That is, if a feature system is rich enough to yield A = trial, then it is rich enough to yield B = dual, say. This section shows that the implicational schema extends to Greenberg-style implications involving approximative numbers (though these were not analyzed in Harbour 2011b). Table 1 above (§1) summarizes the implicational universals that the proposed feature inventory and parameter system derive.[22]

To begin with, I illustrate du → sg,pl (Harbour 2011b). Featurally, dual is [+minimal −atomic]: (+minimal) gathers the lowest layer of the nonatomic region comprising

[21] The aim here is not to rule out the singular–paucal–plural system of Bayso and the minimal–paucal–plural system of Mebengokre. Rather, it is to constrain how they arise: from {±additive, ±atomic} and {±additive, ±minimal}, not from {±additive} alone.

[22] On the current theory, Greenberg's 'universals' arise from two factors: the set of underlying distinctions that a language makes through its choice of features, and the reflection of those underlying distinctions in its surface forms. Surface reflection of underlying distinctions depends on the lexical resources that a language possesses. It is reasonable to entertain the possibility that, as the morphological system of a language atrophies, it might, for a period, attest an inventory of surface forms that violate one of Greenberg's 'universals'. This is unproblematic for the theory offered here, which is concerned with the underlying inventory. That said, however, I am unaware of any exceptions. Evans and Levinson (2009) claim a counterexample to du → sg,pl, discussed immediately below, but data cited in the *New Scientist* (Harbour 2010:30) refute this; see Harbour 2011d for discussion.

dyads, triads, tetrads, and so on, and so captures only dyads. The smallest system where this specification occurs is {±atomic, ±minimal}. This also permits three other feature-value combinations. In [−minimal −atomic], (−minimal) captures the part of the nonatomic region not taken by the dual and, so, is plural (triadic and larger). In [+minimal +atomic], (+atomic) captures just the atomic stratum, and (+minimal) redundantly specifies that they should not have subelements (the definitional criterion of atoms) and, so, is singular. And [−minimal +atomic] corresponds to no number, because there are no atomic elements with atomic subparts. We have therefore derived that a system with the featural resources for dual has singular and plural as well.

Implications PC → PL and GR.PL → PL/AUG follow similarly, though not identically. A language with approximative number must activate [±additive]. Section 4.5 argued that [±additive], if the sole number feature, would induce either singular/minimal–paucal–plural, singular–plural–greater-plural, or minimal–augmented–greater-plural, for first and second person. This delivers the implications. The same section argued, however, that {±additive} is not a legitimate number system because it induces nonconvex cuts. Consequently, the smallest feature systems with an approximative number are {±additive, ±atomic} and {±additive, ±minimal}, which also have singular and plural, and minimal and augmented/plural, respectively.[23]

Feature recursion is a second source of typological implications. Any system with feature recursion has all the numbers of the same system without feature recursion. This derives, for instance, TR → DU. Trial has the specification (+minimal(−minimal (−atomic(P)))). So, it arises only in systems at least as large as {±minimal*, ±atomic}. Languages with feature recursion do not have to use it in every single feature specification, and, as a result, a language with {±minimal*, ±atomic} also has all number values generated by {±minimal, ±atomic}, that is, singular, dual, and plural. Thus, a language with the featural resources for the trial has dual (and singular and plural) too.

Implication GR.PC → PC follows identically. Featurally, greater paucal is (−additive (+additive(P))). So, it arises only in systems of the form {±additive*, … }. These also have the semantic yield of {±additive, … }, one value of which is (−additive(P)), the paucal. Analogous implications affect any language with two approximative numbers.

An alternative way of laying out the typological predictions of the account is by expanding the full range of parametric combinations. This is shown in Table 3.[24] Though well populated, the typology suffers some lacunae. Curiously, these arise more in systems of middle than maximal complexity, that is, in {±F*, ±G$^{(*)}$} more than {±F*, ±G$^{(*)}$, ±H}. The remainder of this section shows that the lacunae have (at least) two plausible sources and, so, do not threaten the account.

The two simplest unattested systems result from adding feature recursion to one of the features in {±additive, ±minimal}. My impression from typological surveys and my

[23] The specification for [±additive], if present in the singular/minimal, depends on person. Third person is [−additive −atomic] or [−additive −minimal], because there are multiple distinct third-person atoms and their join is never atomic. First- and second-person singular/minimal are, by contrast, join-complete, and so correspond to [+additive −atomic] or [+additive −minimal]. Given that the nonthird-person specifications for singular/minimal and plural both involve [+additive], it is important to clarify that this does not constitute a nonconvex cut of the lattice by [+additive]. Rather, [±atomic] or [±minimal] make the first cut of the lattice, convexly, and then [±additive] convexly cuts each of these convex subregions.

[24] Language sources for Table 3: Pirahã: Everett 1986; Svan: Tuite 1995; Winnebago: Lipkind 1945; Rembarrnga: McKay 1978; Kiowa: Harbour 2007; Bayso: Corbett & Hayward 1987; Fula: Anderson 1976; Mebengokre: Silva 2003; Mokilese: Harrison 1976; Yimas: Foley 1991. For the other languages, see Corbett 2000 (which is more up to date than the primary sources; similarly, Anderson 1976, though not a primary source, discusses data across several dialects). Where possible, example languages have a clusivity contrast.

PARAMETER SETTING[a]	NUMBER SYSTEM	EXAMPLE
{ }	no number	Pirahã
{±atomic}	singular, plural	Svan
{±minimal}	minimal, augmented	Winnebago
{±minimal*}	minimal, unit augmented, augmented	Rembarrnga
{±minimal, ±atomic}	singular, dual, plural	Kiowa
$\{\pm\text{additive}, \pm\text{atomic}\}_1$	singular, paucal, plural	Bayso
$\{\pm\text{additive}, \pm\text{atomic}\}_2$	singular, plural, greater plural	Fula
{±additive, ±minimal}	minimal, paucal, plural	Mebengokre
{±minimal*, ±atomic}	singular, dual, trial, plural	Larike
{±additive, ±minimal*}	minimal, unit augmented, paucal, plural[b]	—?
{±additive*, ±atomic}	singular, plural, greater and greatest plurals	Banyun
{±additive*, ±minimal}	minimal, paucal, greater paucal, plural[b]	—?
{±additive*, ±minimal*}	minimal, unit aug., paucal, gr. paucal, plural[b]	—?
$\{\pm\text{additive}, \pm\text{minimal}, \pm\text{atomic}\}_1$	singular, dual, paucal, plural	Yimas
$\{\pm\text{additive}, \pm\text{minimal}, \pm\text{atomic}\}_2$	singular, dual, plural, greater plural	Mokilese
{±additive, ±minimal*, ±atomic}	singular, dual, trial, paucal, plural	Marshallese
$\{\pm\text{additive*}, \pm\text{minimal}, \pm\text{atomic}\}_1$	singular, dual, paucal, greater paucal, plural	Sursurunga
$\{\pm\text{additive*}, \pm\text{minimal}, \pm\text{atomic}\}_2$	singular, dual, paucal, plural, greater plural	Mele-Fila
{±additive*, ±minimal*, ±atomic}	singular, dual, trial, paucal, gr. paucal, plural[b]	—?

TABLE 3. A typology of number systems.

[a] $\{F\}_1$, $\{F\}_2$ denote instantiations of {F} with different approximative numbers, that is, with different sociosemantic conventions as to cut-off points. Systems with two approximative numbers, {±additive*, … }, might have paucal and greater paucal, or paucal and greater plural, or greater plural and greatest plural. Examples of multiple instantiations are given wherever possible. (Global plural is excluded from the typology because it is so scantly attested.)

[b] Or 'plural, greater plural' for 'paucal, plural', and either 'paucal, plural, greater plural' or 'plural/augmented, greater plural, greatest plural' for 'paucal, greater paucal, plural'.

own grammar trawling is that, among two-parameter systems, dual is substantially more common a third number than approximative numbers and unit augmented (i.e. systems like Bayso, Mebengokre, and Rembarrnga are less frequent than systems like that of Kiowa). This might suggest a cognitive bias toward perception of dyads over paucities (and, possibly, a functional pressure resulting from what humans more regularly need to talk about). If so, then we plausibly expect an even stronger bias among three-parameter systems toward those with dual. This makes the two lacunae not wholly surprising.

This argument might weigh more against {±additive*, ±minimal} than {±additive, ±minimal*}, as the latter (minimal–unit-augmented–paucal–plural), although it has no dual, does permit reference to all dyads (via minimal for first inclusive, via unit augmented for other persons). Unit augmentation is very restricted, however, both quantitatively and geographically. In Cysouw's typological sample (2003:234–36), it is found in the Gunwinyguan family and two Burarran languages of northern Australia, and in two languages of Papua New Guinea and one of Eastern Papua; Bowern (2004) adds some Nyulnyulan languages of northwestern Australia.[25] Given such narrow attestation of [±minimal]* without [±atomic], the chances of attesting {±additive, ±minimal*} are, plausibly, low for independent reasons.

The last observation carries over directly to {±additive*, ±minimal*}. If the chances of finding unit augmented with one approximative number are low, and if multiple ap-

[25] Cysouw (2003:234) suggests that Kayapó, a Jê language of Brazil, may exhibit unit augmentation, but the description in Silva 2003 portrays this language, which she calls Mebengokre, as exhibiting minimal–paucal–plural.

proximative numbers occur less frequently than single approximative numbers do, then, naively, the chances of finding unit augmented with multiple approximative numbers are even lower. So, again, this has the appearance of a contingent lacuna, rather than a flaw in the account. Similarly, if multiple approximative numbers are rare and if trials are too, then, naively, trials with multiple approximative numbers are expected to be even rarer. So, absence of the maximally complex system {±additive*, ±minimal*, (±atomic)} also appears to be contingent.[26]

In the last few paragraphs, the discussion has begun to touch on markedness, both in parameters per se and in relation to possible biases in perception. However, these issues lie beyond current confines. I conclude simply that the lacunae in the typology do not seem overly worrisome and observe that the typological implication schema developed in Harbour 2011b for nonapproximative numbers extends straightforwardly to approximative numbers.

5.2. MORPHOLOGICAL ANALYSIS. One of the major morphological motivations for semantically compositional number features, as opposed to the likes of [±paucal] or [±dual], which just 'featuralize' traditional descriptive labels, is the phenomenon of morphological compositionality, that is, cases where one number is built out of one or more others. The Hopi dual (see 35–37 below) is probably the most publicized case of this, but the phenomenon is more general and affects several approximative numbers. Two different versions of the phenomenon are presented and analyzed below. (Recall, from §2.3, that I assume a permissive pronunciation procedure, in which multiple exponents may realize individual features under Number^{0}, or in which a single exponent may realize the whole of Number^{0} together with person, and other, features.)

Morphological compositionality supports the current analysis in that the feature inventory captures the correct classes. The account does not require every possible compositional relationship to be attested, since factors beyond the current study may constrain the structure of the lexicon. Nor does it rule out that some languages might attest some quite ad hoc morphological relationships. The support therefore lies in the fact that much of what is possible is attested and that most of what is attested is easily captured. (Section 5.3, by contrast, examines the limits of the account.)

COMPOSITIONAL NUMBER 1. Hale et al. 1990:255 gives the following triplet of examples, illustrating the morphological composition of the Hopi dual.

(35) Nu' wari. +atomic +minimal
1SG ran.SG/DU
'I ran.'

(36) 'Itam wari. −atomic +minimal
1DU/PL ran.SG/DU
'We (two) ran.'

(37) 'Itam yu'tu. −atomic −minimal
1DU/PL ran.PL
'We (more than two) ran.'

[26] Feature recursion itself is rather restricted as a parameter. Systems with trial or multiple approximative numbers appear to be confined to Austronesian languages and to the non-Austronesian languages of New Guinea and the surrounding islands (Cysouw 2003:197) and to some languages of the Daly and Nyulnyulan families in Australia (Dixon 1980, Corbett 2000, Bowern 2004). Use of feature recursion on multiple features does arise in Kiowa-Tanoan (Harbour 2007, 2011e), but not within Number^{0}. If number systems with feature recursion on multiple features remained unattested in a sample of 50,000 languages, we might wonder whether feature recursion cannot be so parametrized. However, there is little point speculating on inadequate evidence.

As the underlining shows, the dual shares its verb with the singular and its pronoun with the plural. The feature inventory easily captures this. The form of the pronoun, *nu' ~ 'itam*, is sensitive to the value of [±atomic], the form of the verb, *wari ~ yu'tu*, to that of [±minimal].

A striking case of such composition arises for approximative numbers in Mele-Fila (Clark 1975, Corbett 2000:35). The systems that crosscut are the articular (singular–paucal–plural) and pronominal (singular–dual–plural–greater-plural). Thus, when a speaker uses only an article, or only a pronoun, the precise number intended is underdetermined. A purely paucal reading arises when the dual–paucal article and the paucal–plural pronoun cooccur. And a purely plural reading arises when the paucal–plural pronoun cooccurs with the plural–greater-plural article.[27]

<table>
<tr><th></th><th>SINGULAR</th><th>DUAL</th><th>PAUCAL</th><th>PLURAL</th><th>GR.PLURAL</th></tr>
<tr><td>article (DEF)</td><td>te</td><td colspan="2">ruu</td><td colspan="2">a</td></tr>
<tr><td>pronoun</td><td>aia</td><td>raaua</td><td colspan="2">raateu</td><td>reafa</td></tr>
<tr><td>[±atomic]</td><td>+</td><td>−</td><td>−</td><td>−</td><td>−</td></tr>
<tr><td>[±minimal]</td><td>+</td><td>+</td><td>−</td><td>−</td><td>−</td></tr>
<tr><td>[±additive]</td><td>−</td><td>−</td><td>−</td><td>− +</td><td>+</td></tr>
</table>

TABLE 4. Composed number in Mele-Fila.

Table 4 shows the system and the feature values that make this possible. Because plural is both [+additive] and [−additive], it belongs to two natural classes defined by this feature. In the articles, it shares *a* with the greater plural. Both are [+additive]. In the pronouns, it shares *raateu* with the paucal. Both are [−additive].[28]

Motuna, a non-Austronesian language of Bougainville, has the number system singular–dual–paucal–plural, but only (some) nominals express paucal via dedicated morphology. Everywhere else, the paucal is syncretic either with dual or with the plural, according to Onishi's (1994) analysis. Given how extensive the verb system is, this means that, in most cases, paucal meaning arises from crosscutting of two verbal morphemes, one dual–paucal, the other paucal–plural. This is possible because, in the four-value Motuna system, paucal shares [−additive] with dual and [−minimal] with plural.[29] Examples 38–40 are repeated from 1–3 above, with the feature specifications added.

[27] A similar pattern might arise in Koasati, where nouns are maximally singular–paucal–plural but verb suppletion is maximally singular–dual–plural (Kimball 1991:322, 403, 449). I have not spotted any crosscutting examples, however.

[28] If *a* is the [+additive] form of the article, then the obvious claim to make is that dual–paucal *ruu* is the [−additive] form. This raises the question of why *a* rather than *ruu* realizes the plural (which contains both + and − values; presumably *ruu* is blocked from the singular because there are singular-specific forms). If + is the more marked value, then the blocking relation is immediate (see Nevins 2011 for recent discussion of markedness within a similar set of assumptions).

Otherwise, *ruu* must be taken to be an elsewhere form. However, ad hoc stipulation of elsewhere forms is unappealing, as it must be repeated for three distinct sets of exponents: besides the definite, shown in Table 4, the indefinite and the definite diminutive articles distinguish, respectively: singular *se*, *ti*; dual–paucal *na*, *rii*; and plural–greater-plural *ne*, *mii*. (The indefinite diminutive distinguishes just singular *sina* and nonsingular *riina*.) A more economical response, typical within distributed morphology, is to invoke impoverishment of [−additive] in the context of [+additive] on articles. This prevents insertion of [−additive] forms (*ruu*, *na*, *rii*) and forces insertion of the [+additive] ones (*a*, *ne*, *mii*). Though still stipulative, this at least unites the three cases, if alternative approaches are unavailable.

[29] An interesting feature of this language is that a dual–paucal suffix can trigger surface changes in a preceding paucal–plural morpheme (just as it triggers vowel deletion in the TMA marker *mo*). This sometimes increases the distinctness of the paucal from both dual and plural. See Onishi 1994:Ch. 13 for extensive exemplification and discussion of the processes involved.

(38) kaa- ru-ti- m[o]-ee +minimal −additive
disagree-2- DU-TMA- DU/PC
'you (two) disagree'

(39) kaa- ru-m- m[o]-ee −minimal −additive
disagree-2- PC/PL-TMA- DU/PC
'you (few) disagree'

(40) kaa- ru-m- mo- ng −minimal +additive
disagree-2- PC/PL-TMA-PL
'you (many) disagree'

Additionally, in some parts of the kinship system, dual, paucal, and plural all collapse (Onishi 1994:98). This behavior can be captured by virtue of their all being [−atomic]. (Bittner and Hale (1995), citing personal communication from Mary Laughren, note a similar neutralization of the [−minimal] numbers paucal and (greater) plural in Warlpiri in certain contexts.)

Thus, approximative-number analogues of the Hopi dual composition are easily captured.

COMPOSITIONAL NUMBER 2. A different type of morphological compositionality, possibly more common than the Hopi type, involves one number serving as the base form for another. For instance, in Gahuku (Deibler 1976, especially pp. 23–34), the (syncretic) nonfirst-person dual comprises the nonfirst plural plus a dual-specific morpheme. The compositionality holds across a variety of tenses and moods, with the morpheme order PERSON.NUMBER-(DU)-MOOD. Examples include the indicative (PL *a-ve*, DU *a-si-ve*), the future interrogative (PL *a-he*, DU *a-si-he*), the future with switch reference (PL *i-ko*, DU *i-si-ko*), the topic mood (PL *e-moq*, DU *e-si-moq*), and contrafactual apodoses (PL *a-line*, DU *a-si-line*). (The same exponent of dual extends into the first person, but there is an additional allomorphic overlay in mood, which I leave aside here.) The features that underlie this relationship are shown for the indicative in 41.

(41) a- (si)-ve −atomic (+minimal)
2/3.NSG-DU- MOOD

This is, again, something we see for approximative numbers of various sizes.

Bonan and Fula provide straightforward examples. Bonan exhibits singular–paucal–plural. Wu (2003:333) suggests the decomposition in 42.[30]

(42) more-(ghu)-la (−additive) −atomic
horse- PC- PL
'(a few) horses'

The configuration is possible because paucal and plural are both [−atomic] and the paucal morpheme is the pronunciation of [−additive].

Fula has singular–plural–greater-plural. The greater plural is obtained by suffixing *Vji* to a plural noun, as seen in 43, repeated from 4 above (Anderson 1976).

(43) pucc- i- (iji) −atomic (+additive)
horse-PL- GR.PL
'(a great many) horses'

[30] Fried (2010) (brought to my attention by Matthew Baerman and Greville Corbett, p.c.) describes the language as having dual as well (see his n. 25, p. 44, for a brief comment on the discrepancy). This does not affect the issue of compositionality. Indeed, Fried (2010:68) decomposes the dual suffix *ghala* into *ghar* 'two' and the plural *la*, something with which the account has no difficulty (see, for instance, Sursurunga and Mokilese below).

Analogously to Bonan, the configuration is possible because plural and greater plural are both [−atomic], and the greater plural morpheme is the pronunciation of [+additive].

Similar arrangements of morphemes arise also in richer number systems. In Yimas, which has singular–dual–paucal–plural, paucal marking on the verb involves, for some arguments, adding a suffix to a form already marked for plurality. This is shown in 44 and 45 for a first-person agent and a second-person object, respectively (Foley 1991: 220; if the paucal suffix is removed and the perfective allomorphy is appropriately adjusted, the verbs are simply plural (cf. Foley 1991:205, 217)).

(44) pu- kay- cay-c- (ŋkt) −minimal (−additive)
3PL.O-1PL.A-see- PFV-PC
'we (few) saw them'

(45) pu- kra- tay- c- (ŋkt) −minimal (−additive)
3SG.A-2PL.O-see-PFV-PC
'he saw you (few)'

The configuration of morphemes in 44–45 is possible because paucal and plural are both [−minimal] and the paucal is the pronunciation of [−additive] in the context of [−minimal].

	SINGULAR	DUAL	PAUCAL	GR.PAUCAL	PLURAL
1INCL	—	git-ar	git-tul	git-hat	git
1EXCL	iau	gi-ur	gim-tul	gim-hat	gim
2	iáu	ga-ur	gam-tul	gam-hat	gam
3	i/on/ái	di-ar	di-tul	di-hat	di

TABLE 5. Sursurunga emphatic pronouns.

Two complementary variations on this pattern are seen in Sursurunga and Mokilese. The nonsingular numbers of these languages are dual–paucal–greater-paucal–plural, and dual–plural–greater-plural, respectively. However, in Sursurunga, the largest, plural, serves as the base to which other number morphemes attach (Table 5), whereas in Mokilese, it is the smallest nonsingular number, dual, that has this property (predicted as impossible by Krifka 2007; Table 6). The analysis is the same in both cases: the form that serves as the base for all others realizes simply [−atomic] and the relevant person features, and the suffixes realize the additional features of the other numbers.

In Sursurunga, for instance, the basic form of the first-person inclusive [−atomic] is *git*. Additionally, [+minimal] is pronounced as *ar* (hence dual *gitar*), [−additive] as *tul* (hence paucal *gittul*), and [−additive +additive] as *hat* (hence greater paucal *githat*);[31] plain *git* is left just for the plural. In Mokilese, the basic form of the first inclusive [−atomic] is *kisa*. Additionally, [−minimal −additive] is pronounced as *i* (hence plural *kisai*), and [−minimal +additive] as ^{i} (which raises the preceding vowel and is linearized before a nonnasal; hence greater plural *kihs*, where *ih* represents a long *i*); plain *kisa* is left just for the dual.

	SINGULAR	DUAL	PLURAL	GR. PLURAL
1INCL	—	kisa	kisa-i	kihs (kisa-^{i}) kimi
1EXCL	ngoah	kama	kama-i	(kama-^{i}) kimwi
2	koah	kamwa	kamwa-i	(kamwa-^{i})
3	ih	ara/ira	ara-i/ira-i	ihr (ara/ira-^{i})

TABLE 6. Mokilese emphatic pronouns.

[31] If the [−atomic] form ends in *m*, then the dual suffix *ar* becomes the locus of the labial, surfacing as *ur*, with *m* itself vanishing (cf. Hutchisson 1985).

The examples of Hopi- and Gahuku-style compositionality for approximative numbers do not exhaust all theoretical possibilities nor the data available. They are, however, indicative of the range, and they suggest that [±additive] places approximative and other numbers into the correct classes, which exponence can, and does, exploit.

5.3. DIACHRONIC ANALYSIS. We now turn to two cases of diachronic development. Both address how approximative numbers can arise from reconstruing the featural identity of other morphemes. The first concerns how the greater plural can arise from the plural of a plural and involves reconstrual not just of features but of phrase structure. The second concerns how the paucal can arise from nonapproximative numbers (like the trial) and involves featural reconstrual without a change in phrase structure.

GREATER PLURAL: FEATURAL AND PHRASE-STRUCTURAL RECONSTRUAL. In terms of the patterns of morphological composition examined in §5.2, the two greater plurals of Warekena present a potential problem. The suffixes in 19 have the feature specifications in 46.

(46)	GR.GR.PL	penawi	[+additive −atomic]
	GR.PL	nawi	[−additive +additive −atomic]
	PL	pe	[−additive −atomic]

It is analytically tempting to decompose the greatest plural into the concatenation of the other two morphemes, *pe-nawi* (as in Aikhenvald 1998). If we do this, however, then the inclusion relations are irreconcilable. Morphemically, GR.PL is a substring of GR.GR.PL, but featurally GR.GR.PL is a subset of GR.PL.

(47)	GR.PL		GR.GR.PL
	nawi	⊂	pe-nawi
	[−additive +additive −atomic]	⊃	[−additive −atomic]

There is no way of setting up the vocabulary items so as to yield this relationship.[32] The only option is to treat *penawi* as an unanalyzed whole. Although this may seem like a loss of insight, there is evidence to suggest that it is the correct solution.

An instructive case comes from Breton (and from several dialects of Arabic), where number morphemes may accrete. Corbett (2000:36), summarizing from multiple sources, gives the following examples (noting that the behavior is lexically restricted).

(48)	lagad	daou-lagad	daou-lagad-où
	eye	DU- eyes	DU- eye- PL
	'eye'	'(two) eyes'	'pairs of eyes'
(49)	bugel	bugal-e	bugal-e- où
	child	child- PL	child- PL-PL
	'child'	'children'	'groups of children'

What is odd about these examples is not their meaning: the third of each triple is just the plural of 'pair' or 'group'. What is odd is their form. 'Pair' and 'group' are not signaled by nouns. Instead, the dual of 'eye' and plural of 'child' have apparently been treated as the singular of new nominals, 'pair of eyes', 'group of children', which, like other nouns, can then be pluralized.

I suggest that Breton represents an intermediary stage by which greater plurals can arise. In its form, the greater plural of Fula is parallel to 'groups of children' in Breton. The GR.PL suffix in 4 (repeated in 43) is synchronically still a plural for other nouns (e.g. *faada-aji* 'palace-PL', *wannyo-oji* 'game-PL'; Anderson 1976:122). However, Fula

[32] Except by impoverishment-based hacking: *pe* realizes [−atomic], *nawi* realizes [+additive], but [−atomic] is deleted in the context of [−additive +additive].

differs from Breton semantically: *pucciiji* (and similar) do not mean 'groups of horses', but 'horses in great number'. It is easy to see how a plurality of pluralities would acquire this meaning (several herds of horses is a great number), but, crucially, Fula has reinterpreted what used to be only a plural marker as something that is plural ([−atomic]) for some nouns (e.g. 'palace') but greater plural ([+additive]) for others (e.g. 'horse').[33] Corbett (2000:37, following Appleyard 1987) observes that reinterpretation has also occurred in the Khamtanga (Chamir) dialect cluster of the Agaw Cushitic group, though with less semantically dramatic effect: where a century earlier the language exhibited, for instance, *iefír* 'children' and *iefírt* 'crowds of children', Appleyard records these simply as free variants of plural 'children'.

Returning to Warekena in this light, we find evidence of semantic reinterpretation there, too. Aikhenvald (1998:302) gives the meaning of *penawi* as 'very many … indeed, so many one cannot count them'. It occurs in another context with the sense of exhaustivity ('all the people, all the women, all the men', 1998:302). If we decompose this into *pe-nawi*, then we expect a different semantics, that of a plurality of pluralities. To capture its actual meaning, therefore, it appears necessary to not decompose it; and, without decomposition, the problem of the subset/substring mismatch in 47 does not arise. In other words, the feature system pushes us into an analysis for which there is independent evidence.[34]

In theoretical terms, we can understand the diachronic stages as follows. At the earlier stage, there would be two *n* projections, each dominated by a Number projection. As laid out in §2.3, each *n* defines a new nominal predicate. For Fula, for instance, the

[33] Fula preserves another sign of a stage prior to development of the greater plural. In Arabic, plural of plural forms can have a 'kinds of' reading (Ojeda 1992, Corbett 2000). Fula retains this for *nyaam-du-uji* 'kinds of food' (a different concatenation of number morphemes: food-SG-PL; Anderson 1976:123). In Miya, too, the same morpheme gives a greater plural for count nouns and a 'kinds of' reading for mass nouns. Greek offers a slight variation on the theme: the normal count plural, when applied to mass nouns, gives both 'kinds' and abundance readings (Tsoulas 2009).

A separate issue is whether the semantic range of pluralities of pluralities overlaps with that of plain plurals. I am not aware of any descriptions bearing on the issue, but I would not rule this out, since it seems similar to the overlap between greater plural and plain plural.

[34] There is, additionally, evidence of semantic variety in the interpretation of *pe* by itself (Aikhenvald 1998:301). It can express a unit group, that is, a plurality reconstrued as a singleton, in *ʃiani-pe* 'family' (alongside its straightforward plural reading in the pair *ʃiani* 'child', *ʃiani-pe* 'children'). And, it can have an exhaustive (hence possibly global) feel, in *ʃabine* 'family, household', *ʃabine-pe* 'all the members of a household', if the English translation is indicative. This adds support to the view that Warekena morphology has undergone some semantic reassignment.

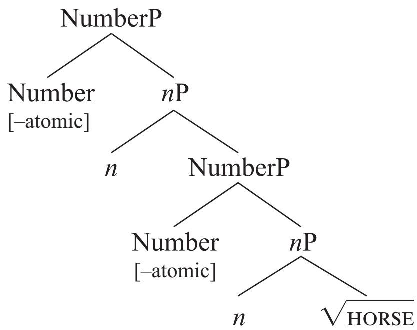
Furthermore, Koasati may present a similar case (Kimball 1991:447–49), though the numbers and diachronic sources are different from those in Warekena. Most Koasati nouns distinguish singular–plural; and plural ([−atomic]) is *ha*, as in *tayyí* 'woman', *tayyi-há* 'woman-PL'. Some nouns, however, built on the diminutive *(o)si*, distinguish singular–paucal–plural.

(i)	icofó:si	icofós-ki	icofós-kiha
	nephew.SG	nephew-PC	nephew-PL
	'nephew'	'a few nephews'	'nephews'

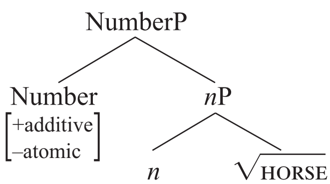
It would be tempting to regard *ki* as [−atomic] and *ha* as [+additive], as this yields *ki-ha* for plural [+additive −atomic], but then the language would have two different *ha*s and two allomorphs of [−atomic], namely *ha* of 'woman.PL' and *ki* of 'nephew.PC/PL'.

A neater system is one where *ki* realizes [−additive −atomic], *kiha* [+additive −atomic], and *ha* [−atomic]. Again, other language-internal facts seem to favor such reanalysis, in that, for some words, *ki* and *ha* do not follow the diminutive *os* in that order, but sandwich it in the reverse order, as in *tayyosí* 'girl', *tayyos-ki* 'girl-PC', *tayyi-haski* 'girl-PL': rather than posit radically different orders of identical morphemes (*os-ki-ha* versus *ha-s-ki*), one can simply regard *haski* as a lexically specific, undecomposed allomorph alongside *kiha*, and *ha*.

lower *n* creates a predicate true of horses; the upper *n* creates one true of (by convention, large) pluralities of horses, as in 50. (This ‘renominalization’ is similar to what occurs in [[[rest$_n$]-less$_a$]-ness$_n$], without the intervening adjectival detour.)

(50)


At a later stage, the upper [−atomic] was featurally reconstrued, permitting the upper *n* to be lost and the two Number heads to coalesce.

(51)


The nullness of *n* might, in part, drive this development, biasing the learner to posit featural identities for the two number morphemes that permit them to cooccur under a single head. (Recall that the axiom of extension prevents this while number morphemes realize identical features.)

The sketched analysis extends to Warekena and Khamtanga (Chamir) with minimal changes. For Warekena, the analogue of 50 would contain [+additive −atomic] on the lower Number head, and the analogue of 51 would contain [−additive +additive −atomic]. For Khamtanga, 50 would represent the earlier stage, but the analogue of 51 would lack [+additive]; rather, [−atomic] would have free allomorphic variants.

Paucal: purely semantic reconstrual. Cases like Warekena represent the opposite of those examined under the rubric of morphological compositionality. They present configurations of exponents that the current theory cannot capture with ease. The diachronic propensity of paucal to arise from nonapproximative numbers or numerals illustrates the same point. Examples are Sursurunga, in which the lesser paucal stems from the number ‘three’ and the greater paucal from ‘four’ (see also Lynch et al. 2002: 35, Wu 2003:333, Fried 2010:68), and Russian, where the paucal (governed by nominative and accusative numerals ending in ‘two’, ‘three’, or ‘four’)[35] derives from the

[35] Dual-to-paucal development may be less common than trial to paucal. If the suggestion of a cognitive bias to perception of dyads over paucities is correct (§5.1), then this disparity might be expected, making the highest nonplural number more stable in {±atomic, ±minimal} systems than in {±atomic, ±minimal*}. Some possible examples of dual-to-paucal development, besides Slavic, include Foley’s (2005) reconstruction of *-ŋk-* for Sepik Ramu, the reflexes of which are dual in Ramu but paucal in Lower Sepik, and the Cook Islands Maori demonstrative *aua*, which is dual-cum-paucal despite containing what is, for pronouns, still the dual ending *-ua* (Buse 1995:86, 534). Corbett (2000:25) also cites Blanc (1970:45) for the development of paucal from dual in some Arabic dialects. (In a semantically, though possibly not featurally, related vein, consider English *a couple of* and German *ein Paar* ‘a pair (of)’, which have both spread from ‘two’ to ‘few’.)

ancestral dual.[36] In current terms, this development involves reinterpreting a number defined in terms of [±minimal] as one involving [±additive]. If we consider the evidence that leads the learner to posit one feature versus the other, it becomes clear that there is a disparity that creates an inherent instability in grammars without [±additive].

Consider a learner acquiring a number system of the form singular–dual–*something*–plural. This *something* must be the trial or paucal. Let us suppose, for the sake of simplicity, that *something* is acquired last. So, the child will have activated the features {±atomic, ±minimal}. Two parameters can generate a fourth number: feature recursion on [±minimal], or [±additive] as a third number feature. Let us consider how these two options fare when faced with the child's primary linguistic data.

One case is simple. If *something* is used in situations where there are no exact triads, but there are other paucities, then the child has positive data for [±additive] and against [±minimal]*, because the latter yields only triadic reference, given the other features that the child has activated. Presumably, therefore, the child settles on a unique I-grammar for number features.

In the other case, the evidence is less decisive. If *something* is only used to talk of triads, then the child has evidence for activation of [±minimal]*. However, the child has no evidence against activation of [±additive]: the trial is, after all, within the range of the paucal. It might be the case that the child cumulatively disfavors [±additive] because its fuller range is not exploited. But this is not the same as saying that the child activates only [±minimal]*. It may be that the child activates both but that the grammar incorporating [±additive] merely falls into disuse, rather than being ruled out.

It is interesting now to examine how these grammars differ with regard to their intergenerational stability (for a formalized approach to such issues, see Benz et al. 2006, Jäger 2008). The first case, involving the paucal, is presumably stable. Generation *n* has [±additive] and produces forms that refer, among other things, to tetrads and heptads. This gives generation $n + 1$ an unambiguous cue to activate [±additive], not [±minimal]*. So, the grammar replicates itself exactly from generation to generation.[37]

In the second case, generation *n* has [±minimal]* as the predominant grammar. Some speakers may also produce forms compatible only with [±additive], however, either as errors or else in virtue of their residual [±additive] grammar. Generation $n + 1$ might, in consequence, disfavor the [±additive] grammar slightly less than generation *n* does and might, in turn, produce slightly more forms using it (i.e. they might produce more non-trial paucals). Cumulatively, then, successive generations may come increasingly to favor [±additive], culminating at a tipping point where [±additive] becomes the only stable grammar.

We can also imagine the same scenario having occurred in Russian. If the dual began to decline in rates of use, there would have been less evidence to favor a parse of (+minimal(−atomic(P))) over (−additive(−atomic(P))), thus allowing a drift from dual to paucal.

Precisely what absolute and relative rates of production (of true duals/trials versus 'erroneous' paucals) are required to permit the drift from [±minimal]$^{(*)}$ to [±additive], and

[36] As explained in §3.2, I regard [±additive] as a selectional feature of the numerals two to four in Russian. Numerals with selectional [±additive] are attested also in Arabic and Byak, though there paucal can also occur without numerals. The retreat of the Russian paucal to so specific a grammatical niche is perhaps indicative of the decline that contributed to its reinterpretation from a dual. Its sensitivity to case ties in to the fact that it parasitizes case exponents.

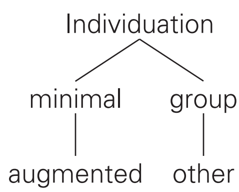
[37] Naturally, a community, or a part, might reconventionalize the bound that [±additive] induces to three. Later generations would then acquire [±minimal]*, not [±additive], as the extra parameter. See Corbett 2000:22 for some Australian examples.

so from dual/trial to paucal, and what social factors might encourage or restrain it, are questions beyond the scope of this article. However, the point of relevance to the theory of number features developed here is that two of the features overlap in the cardinalities they may pick out. This, in turn, provides a preponderant direction of development in number systems, which appears to be adequately attested in natural language.[38]

6. OTHER ACCOUNTS. Many of the topics touched on above have been addressed in other theoretical work. For reasons of space and locus, I consider only one other account of approximative number here, the feature-geometric approach of Harley & Ritter 2002 (see also Cysouw 2011 for an extensive evaluation of their approach). Besides some basic empirical differences, our accounts differ substantially in terms of their economy and the extent to which they serve the search for explanatory adequacy (Chomsky 1965, Fodor 1981).

Harley and Ritter's account proposes that person, number, and gender features are organized hierarchically. The features it posits are privative, not bivalent. For instance, the subtree for number is as in 52.

(52)



I have taken the liberty of naming a fourth feature, [other], and adding it to the tree. Harley and Ritter (2002:494) mention the necessity of a fourth feature in order to handle the five-value systems of Sursurunga and Mele-Fila and themselves suggest the locus shown. Although there are similarities of names, there is no simple correspondence between the features in 52 and those proposed above (see four paragraphs below). The different fonts serve to highlight this.

At first glance, both proposals are evenly matched. Three bivalent features plus feature recursion and a Number head balance against four privative features and an Individuation node. However, the accounts are neither conceptually nor empirically equal.

First, geometries are not obviously explanatory. For four features plus a root node, there are $(4 + 1)^{(4-1)} = 125$ different geometries, by Cayley's theorem. For 52 to be explanatorily adequate, it must be shown why none of the other 124 geometries is possible. This is not the same as showing that alternative geometries yield the wrong inventory of number systems. It involves showing why the alternative number systems could not have been the actual ones. It is not clear what this type of explanation would demand. A starting point, however, would be to determine which interface it is that forces features into geometric structures, a question that leads to the second problem.

Feature geometries are not syntactic structures, as they do not project according to phrase-structural rules. They are not phonological, because, on the nonlexicalist view of Harley and Ritter (which I share), phonological structures, unlike phi-features, do not interact with syntax. Besides, phonological geometries have, since Sagey 1986, been grounded in articulatory mechanics, which are surely irrelevant to number features. If

[38] Given that there is no featural quadral, the Sursurunga greater paucal, which derives from 'four' and related forms in other languages (Lynch et al. 2002:35), would have had to have moved immediately (in 'grammatical' time) from a numeral reading to a paucal reading. Alternatively, it may never have had a numeral reading when used as a grammatical affix, if it was formed by conscious analogy at a time after 'paucalization' of 'three': 'three' is less than 'four', so 'three-based paucal' is less than 'four-based paucal'.

geometries are sui generis morphological structures, then one must depart (again, contra Harley and Ritter) from the distributed morphology view of morphology simply as part of the mapping from syntax to phonology, not a special domain in its own right (a concept too close to the 'levels of representation' that syntactic MINIMALISM aims to do away with). So, if not syntactic, or phonological, or morphological, geometries must be semantic. This is, of course, more or less the line pursued here: constraints on features emerge from (the formal underpinnings of) the feature semantics. But this does not solve the question of what geometries are. It dissolves it. Geometries are not part of the formal apparatus above since there are no dependency relations between features.

Even putting aside the question of geometries, it is inaccurate to portray the two accounts as each positing four features/parameters. If [±additive], [±atomic], and [±minimal] have been previously motivated in treatments of masshood, aspect, and telicity, then the theory has not posited anything new. Furthermore, the parameter of feature recursion is also independently motivated: Harbour 2007, 2011e argue at length that 'inverse' number in Kiowa (the phenomenon whereby one and the same morpheme singularizes some nouns but pluralizes others) is the pronunciation of [+F −F]. Consequently, the parametric option permitting such specifications is required beyond the data examined above. Thus, all four parameters posited have independent motivation.

The independent motivation for the bivalent features does not obviously transfer to the privative ones. Despite some overlap in names, the two feature systems are nonisomorphic. For instance, [group] does double duty as English-style plural [−atomic] and Winnebago-style augmented [−minimal], and [minimal group augmented] does double duty as trial [+minimal −minimal −atomic] and Yimas-style paucal [−additive −minimal −atomic]. (Cysouw 2011 presents rather pointed criticism of this semantic underdetermination.)

Such double duty presents Harley and Ritter's account with an empirical problem mentioned in §4.2. If singular and minimal are featurally the same, and likewise for dual and unit augmented, then the operative semantic principle seems to be that each cardinally exact number, like singular or dual, corresponds to something in a minimal–augmented-type system. Consequently, one expects that, corresponding to trial, which is two more than singular, there should be dyad augmented, which is two more than minimal (see 25, §4.2). I claim that this number is featurally impossible, owing to the axiom of extension: it would involve multiple occurrences of [−minimal]. It is not clear what prevents this interpretation of 'trial' in Harley and Ritter's system.

Valence is another obvious difference between my account and Harley and Ritter's, and one, again, with empirical consequences. Harbour 2011e examines three varieties of privativity at length and concludes that privative number features cannot replicate the treatment of inverse number and associated phenomena in Kiowa-Tanoan. Rather, number features must permit the representation [+F −F] that only bivalence affords (and which it is hard to reconstitute in privative terms, since the very notion of value conflict is unexpected). Harbour 2013 presents a further range of semantic, morphological, and syntactic phenomena (including, respectively, feature recursion, polarity or alpha rules, and differential object marking in Kiowa-Tanoan) that require a three-way contrast between [+F], [−F], and absence of [±F]. Bivalence affords this set of distinctions, but privativity does not.

Finally, Harley and Ritter's target data also present several difficulties. For reasons of space, two examples will have to suffice. First, they claim that trial and paucal are different interpretations of 'the same geometric configuration' (2002:494) and so cannot cooccur in a single language. However, languages with both are reported (Lihir and

Marshallese: Corbett 2000; Mussau: Ross 2002, Brownie & Brownie 2007). Although one must approach such reports with caution (since some authors, like Palmer (2009: 68), use 'trial' to mean paucal), Corbett (2000:30, citing personal communication with B. Bender) is quite clear that the Marshallese 'trial' is indeed a trial, as is J. Brownie (p.c.) for Mussau.

Besides ruling out attested systems, Harley and Ritter's risks ruling in unattested ones. We noted the problem of dyad augmentation above. A similar issue arises from the fact that paucal can have a cardinally exact interpretation as trial. Parity of reasoning suggests that greater paucal should have an analogous interpretation as a cardinally exact quadral. Corbett fails to find this number, and Lynch and colleagues (2002:35) note that Oceanic pronouns based on the numeral 'four' are always paucal, not quadral (cf. §4.2).

I think it is unfair to Harley and Ritter to push the semantic arguments too far, as the feature dubbed [other] might avoid these difficulties, if properly defined. Even granted that latitude, however, the more substantial problems would remain: valence and empirical adequacy, geometries and explanatory adequacy, and the imparsimony of positing rather than recycling parameters.

7. CONCLUSION. The foregoing has argued for a feature inventory that provides a theoretically uniform treatment of the approximate and nonapproximate numbers, and in which the constraints on the features are general properties of the semantic or other cognitive systems. Besides predicting a range of typological, morphological, and (semantic) diachronic phenomena, the theory connects in potentially profound ways with domains of linguistic and nonlinguistic research that ought, at first glance, to lie far from the concerns of paucals, greater plurals, augmentation, duals, and trials. If there is a methodological moral to be gained from this, it is that our feature definitions should strive for a generality that moves them beyond such pretheoretic, descriptive confines as number, mass, aspect, and so on, and permits us to examine the primitives of our mental ontology at a more fundamental level. The theory of cognition that begins to emerge when we do this suggests that there is deep unity to be discovered.

REFERENCES

ACQUAVIVA, PAOLO. 2008. *Lexical plurals: A morphosemantic approach*. Oxford: Oxford University Press.

ADGER, DAVID. 2012. *A syntax of substance*. Cambridge, MA: MIT Press.

AIKHENVALD, ALEXANDRA Y. 1998. Warekena. *Handbook of Amazonian languages*, vol. 4, ed. by Desmond C. Derbyshire and Geoffrey K. Pullum, 225–439. Berlin: Mouton de Gruyter.

ANDERSON, STEPHEN R. 1976. On the description of consonant gradation in Fula. *Studies in African Linguistics* 7.93–136.

APPLEYARD, DAVID L. 1987. A grammatical sketch of Khamtanga—I. *Bulletin of the School of Oriental and African Studies, University of London* 50.241–66.

BENZ, ANTON; GERHARD JÄGER; and ROBERT VAN ROOIJ. 2006. An introduction to game theory for linguists. *Game theory and pragmatics*, ed. by Anton Benz, Gerhard Jäger, and Robert van Rooij, 1–82. Sydney: Palgrave MacMillan.

BITTNER, MARIA, and KENNETH HALE. 1995. Remarks on definiteness in Warlpiri. *Quantification in natural languages*, ed. by Emmon Bach, Eloise Jelinek, Angelika Kratzer, and Barbara H. Partee, 81–105. Dordrecht: Springer.

BLANC, HAIM. 1970. Dual and pseudo-dual in the Arabic dialects. *Language* 46.42–57.

BORER, HAGIT. 1983. *Parametric syntax: Case studies in Semitic and Romance languages*. Dordrecht: Foris.

BORER, HAGIT. 2005a. *Structuring sense, vol. 1: In name only*. Oxford: Oxford University Press.

BORER, HAGIT. 2005b. *Structuring sense, vol. 2: The normal course of events*. Oxford: Oxford University Press.

BORER, HAGIT. 2013. *Structuring sense, vol. 3: Taking form*. Oxford: Oxford University Press.

BOWERN, CLAIRE. 2004. *Bardi verb morphology in historical perspective*. Cambridge, MA: Harvard University dissertation.

BROADWELL, GEORGE AARON. 2006. *A Chocktaw reference grammar*. Lincoln: University of Nebraska Press.

BROWNIE, JOHN, and MARJO BROWNIE. 2007. *Mussau grammar essentials*. Ukarumpa: SIL-PNG Academic Publications. Online: www.sil.org/pacific/png/pubs/48552/MussauGrammarEssentials.pdf.

BUSE, JASPER, with RAUTUTI TARINGA. 1995. *Cook Islands Maori dictionary*, ed. by Bruce Biggs and Rangi Moeka'a. Canberra: ANU Printing Service.

CHAPPELL, HILARY. 1996. Inalienability and the personal domain in Mandarin Chinese discourse. *The grammar of inalienability: A typological perspective on body parts and the part–whole relation*, ed. by Hilary Chappell and William McGregor, 465–527. Berlin: Mouton de Gruyter.

CHENG, CHUNG-YING. 1973. Response to Moravcsik. *Approaches to natural language*, ed. by Jaakko Hintikka, Edith Moravcsik, and Patrick Suppes, 286–88. Dordrecht: Reidel.

CHIERCHIA, GENNARO. 2010. Mass nouns, vagueness and semantic variation. *Synthese* 174.99–149.

CHOMSKY, NOAM. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

CHOMSKY, NOAM. 2004. Beyond explanatory adequacy. *Structures and beyond: The cartography of syntactic structures, vol. 3*, ed. by Adriana Belletti, 104–31. Oxford: Oxford University Press.

CLARK, ROSS. 1975. Mele notes. *Auckland University Working Papers in Anthropology, Archaeology, Linguistics, and Maori Studies* 40.

CORBETT, GREVILLE G. 2000. *Number*. Cambridge: Cambridge University Press.

CORBETT, GREVILLE G., and RICHARD HAYWARD. 1987. Gender and number in Bayso. *Lingua* 73.1–28.

COWPER, ELIZABETH. 2005. A note on number. *Linguistic Inquiry* 36.441–55.

CYSOUW, MICHAEL. 2003. *The paradigmatic structure of person marking*. Oxford: Oxford University Press.

CYSOUW, MICHAEL. 2011. The expression of person and number: A typologist's perspective. *Morphology* 21.419–43.

DEIBLER, ELLIS W. 1976. *Semantic relationships of Gahuku verbs*. Norman, OK: Summer Institute of Linguistics.

DIXON, R. M. W. 1980. *The languages of Australia*. Cambridge: Cambridge University Press.

DIXON, R. M. W. 1988. *A grammar of Boumaa Fijian*. Chicago: University of Chicago Press.

DONOHUE, MARK. 1999. *A grammar of Tukang Besi*. Berlin: Mouton de Gruyter.

EVANS, NICHOLAS, and STEPHEN C. LEVINSON. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32.429–92.

EVERETT, DANIEL. 1986. Pirahã. *Handbook of Amazonian languages*, vol. 1, ed. by Desmond C. Derbyshire and Geoffrey K. Pullum, 200–325. Berlin: Mouton de Gruyter.

FODOR, JERRY. 1981. Some notes on what linguistics is about. *Readings in philosophy of psychology*, vol. 2, ed. by Ned Block, 197–207. Cambridge, MA: Harvard University Press.

FOLEY, WILLIAM. 1991. *The Yimas language of New Guinea*. Stanford, CA: CSLI Publications.

FOLEY, WILLIAM. 2005. Linguistic prehistory in the Sepik-Ramu basin. *Papuan pasts: Cultural, linguistic and biological histories of Papuan-speaking peoples* (Pacific linguistics 572), ed. by Andrew Pawley, Robert Attenborough, Jack Golson, and Robin Hide, 109–44. Canberra: Australian National University.

FRIED, ROBERT WAYNE. 2010. *A grammar of Bao'an Tu, a Mongolic language of northwest China*. Buffalo: State University of New York at Buffalo dissertation.

GÄRDENFORS, PETER. 2004. *Conceptual spaces: The geometry of thought*. Cambridge, MA: MIT Press.

GRÄTZER, GEORGE. 1978. *General lattice theory*. Basel: Birkhäuser.

Greenberg, Joseph H. 1966. *Language universals, with special reference to feature hierarchies*. The Hague: Mouton.
Grimshaw, Jane. 2005. Extended projection. *Words and structure*, ed. by Jane Grimshaw, 1–74. Stanford, CA: CSLI Publications.
Hale, Kenneth. 1973. Person marking in Walbiri. *A festschrift for Morris Halle*, ed. by Stephen R. Anderson and Paul Kiparsky, 308–44. New York: Holt, Rinehart, and Winston.
Hale, Kenneth. 1997. Some observations on the contribution of local languages to linguistic science. *Lingua* 100.71–89.
Hale, Kenneth; LaVerne Masayesva Jeanne; and Paula M. Pranka. 1990. On suppletion, selection, and agreement. *Interdisciplinary approaches to languages*, ed. by Carol Georgopoulos and Roberta Ishinara, 255–70. Dordrecht: Kluwer.
Harbour, Daniel. 2007. *Morphosemantic number: From Kiowa noun classes to UG number features*. Dordrecht: Springer.
Harbour, Daniel. 2010. In a word (Letter). *New Scientist* 206.2766.29–30.
Harbour, Daniel. 2011a. An arithmetic calculus for presuppositions. London: Queen Mary University of London, ms.
Harbour, Daniel. 2011b. Descriptive and explanatory markedness. *Morphology* 21.223–40.
Harbour, Daniel. 2011c. The generative syntax of the natural numbers. London: Queen Mary University of London, ms.
Harbour, Daniel. 2011d. Mythomania? Methods and morals from 'The myth of language universals'. *Lingua* 121.1820–30.
Harbour, Daniel. 2011e. Valence and atomic number. *Linguistic Inquiry* 42.561–94.
Harbour, Daniel. 2012. The architecture of features and the derivation of person. London: Queen Mary University of London, ms.
Harbour, Daniel. 2013. 'Not plus' isn't 'not there': Bivalence in person, number, and gender. *Distributed morphology today*, ed. by Ora Matushansky and Alec Marantz, 135–50. Cambridge, MA: MIT Press.
Hardman, Martha J. 2001. *Aymara*. Munich: LINCOM Europa.
Harley, Heidi, and Elizabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78.482–526.
Harrison, Sheldon. 1976. *Mokilese reference grammar*. Honolulu: University Press of Hawaii.
Hutchisson, Don. 1985. Sursurunga morphophonemics. Ukarumpa: Summer Institute of Linguistics, ms. Online: http://www-01.sil.org/pacific/png/pubs/928474543575/Sursurunga_Counting_Systems.pdf.
Hutchisson, Don. 1986. Sursurunga pronouns and the special uses of quadral number. *Pronominal systems*, ed. by Ursula Wiesemann, 1–20. Tübingen: Gunter Narr.
Jäger, Gerhard. 2008. Applications of game theory in linguistics. *Language and Linguistics Compass* 2.406–21.
Kaufman, E. L.; M. W. Lord; T. W. Reese; and J. Volkmann. 1949. The discrimination of visual number. *The American Journal of Psychology* 62.498–525.
Kimball, Geoffrey D. 1991. *Koasati grammar*. Lincoln: University of Nebraska Press.
Kitcher, Philip. 1984. *The nature of mathematical knowledge*. Oxford: Oxford University Press.
Krifka, Manfred. 1992. Thematic relations as links between nominal reference and temporal constitution. *Lexical matters*, ed. by Ivan Sag and Anna Szabolcsi, 29–53. Stanford, CA: CSLI Publications.
Krifka, Manfred. 2007. A note on the pronoun system and the predicate marker in Tok Pisin. *Form, structure, and grammar: A festschrift presented to Günther Grewendorf on the occasion of his 60th birthday*, ed. by Patrick Brandt and Eric Fuss, 79–92. Berlin: Akademie.
Lasersohn, Peter. 1995. *Plurality, conjunction and events*. Dordrecht: Kluwer.
Le Corre, Mathieu, and Susan Carey. 2007. One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition* 105.395–438.
Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. *Meaning, use and interpretation of language*, ed. by Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow, 302–23. Berlin: Walter de Gruyter.

LIPKIND, WILLIAM. 1945. *Winnebago grammar*. New York: King's Crown.

LYDALL, JEAN. 1976. Hamer. *The non-Semitic languages of Ethiopia*, ed. by Lionel Bender, 393–438. East Lansing: African Studies Center, Michigan State University.

LYNCH, JOHN; MALCOLM ROSS; and TERRY CROWLEY. 2002. Typological overview. *The Oceanic languages*, ed. by John Lynch, Malcolm Ross, and Terry Crowley, 34–53. Surrey: Curzon.

MARANTZ, ALEC. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics (Proceedings of the 21st annual Penn Linguistics Colloquium)* 4.2.201–25. Online: http://www.ling.upenn.edu/papers/v4.2-contents.html.

MCKAY, GRAHAM. 1978. Pronominal person and number categories in Rembarrnga and Djeebbana. *Oceanic Linguistics* 17.27–37.

MCKAY, GRAHAM. 1979. Gender and the category *unit augmented*. *Oceanic Linguistics* 18.203–10.

NAKANISHI, KIMIKO, and SATOSHI TOMIOKA. 2004. Japanese plurals are exceptional. *Journal of East Asian Linguistics* 13.141–79.

NEVINS, ANDREW. 2011. Marked targets versus marked triggers and impoverishment of the dual. *Linguistic Inquiry* 42.413–44.

NORMAN, JERRY. 1988. *Chinese*. Cambridge: Cambridge University Press.

NOYER, ROLF. 1992. *Features, positions and affixes in autonomous morphological structure*. Cambridge, MA: MIT dissertation.

OJEDA, ALMERINDO. 1992. The semantics of number in Arabic. *Proceedings of Semantics and Linguistic Theory (SALT)* 2.303–25.

OJEDA, ALMERINDO. 1998. The semantics of collectives and distributives in Papago. *Natural Language Semantics* 6.245–70.

ONISHI, MASAYUKI. 1994. *A grammar of Motuna (Bougainville, Papua New Guinea)*. Canberra: Australian National University.

OUWAYDA, SARAH. 2013. Oils and waters: On Lebanese Arabic mass plurals. Paper presented at the annual meeting of the Linguistic Society of America, Boston.

PALMER, BILL. 2009. *Kokota grammar*. (Oceanic linguistics special publication 35.) Honolulu: University of Hawai'i Press.

QUINE, WILLARD VAN ORMAN. 1960. *Word and object*. Cambridge, MA: MIT Press.

RAMCHAND, GILLIAN. 1997. *Aspect and predication: The semantics of argument structure*. Oxford: Oxford University Press.

REGIER, TERRY; PAUL KAY; and NAVEEN KHETARPAL. 2007. Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences* 104.1436–41.

ROSS, MALCOLM. 2002. Mussau. *The Oceanic languages*, ed. by John Lynch, Malcolm Ross, and Terry Crowley, 148–66. Surrey: Curzon.

RUSSELL, BERTRAND. 1919. *Introduction to mathematical philosophy*. London: Allen & Unwin.

SAGEY, ELIZABETH. 1986. *The representation of features and relations in non-linear phonology*. Cambridge, MA: MIT dissertation.

SAUVAGEOT, SERGE. 1967. Note sur la classification nominale en baïnouk. *La classification nominale dans les langues négro-africaines*, ed. by Gabriel Manessy, 225–36. Paris: CNRS.

SCHUH, RUSSELL G. 1998. *A grammar of Miya*. Berkeley: University of California Press.

SILVA, AMÉLIA REIS. 2003. Pronomes, ordem e ergatividade em Mebengokre (Kayapó). Campinas: Universidade Estadual de Campinas master's thesis.

SILVERSTEIN, MICHAEL. 1976. Hierarchy of features and ergativity. *Grammatical categories in Australian languages*, ed. by R. M. W. Dixon, 112–71. Canberra: Australian Institute of Aboriginal Studies.

THOMAS, DAVID. 1955. Three analyses of the Ilocano pronoun system. *Word* 11.204–8.

TSOULAS, GEORGE. 2009. On the grammar of number and mass terms in Greek. *MIT Working Papers in Linguistics (Proceedings of the 2007 Workshop in Greek Syntax and Semantics at MIT)* 56.333–48.

TUITE, KEVIN. 1995. Svan and its speakers. Montreal: Université de Montréal, MS. Online: http://mapageweb.umontreal.ca/tuitekj/publications/Tuite-1997-SvanSpeakers.pdf.

van den Heuvel, Wilco. 2006. *Biak: Description of an Austronesian language of Papua*. Utrecht: LOT.
Wu, Hugjiltu. 2003. Bonan. *The Mongolic languages*, ed. by Juha Janhunen, 325–45. New York: Routledge.
Wynn, Karen. 1990. Children's understanding of counting. *Cognition* 36.155–93.
Zwarts, Joost. 1995. The semantics of relative location. Utrecht: OTS, Utrecht University, ms.

Department of Linguistics (SLLF)
Queen Mary University of London
Mile End Road
London E1 4NS, United Kingdom
[d.harbour@qmul.ac.uk]

[Received 10 September 2011;
accepted pending revisions 18 June 2012;
revision received 20 March 2013;
accepted 24 June 2013]