HENNING ANDERSEN

The Plasticity of Universal Grammar

There is no denying the great diversity of human languages and the deep differences among them, not just in vocabulary – that practically goes without saying – but in phonology, inflection, and the structure of sentences. These are things anyone can observe by comparing a few examples of sentences in just a handful of languages from different continents.

Linguists have taken two radically different views of this diversity, perhaps reflecting different metatheoretical orientations. On the one hand, it has been claimed that languages can vary without limit. Obviously this claim cannot be demonstrated or falsified by looking at the 6,000 or so languages spoken today. It would not even help to have access to all the thousands of languages spoken by our species over the last perhaps 100,000 years. Logically this is not an empirical claim, but it might still be true. On the other hand, it has been claimed that "grammar is substantially one and the same in all languages, and that surface differences between them are merely accidental variations"; thus Francis Bacon [1]. This old idea of a universal grammar underlying all languages has dominated linguistic thinking periodically since the Middle Ages and is still current.

Professor Henning Andersen
Department of Slavic Languages and Literatures,
University of California, UCLA, USA
andersen@ucla.edu
CAS Fellow 2004/2005

These diametrically opposite views may seem irreconcilable, but they are actually not incompatible. The first is surely a useful reminder for the linguist who begins the investigation of a new language; he should be able to draw on previous experience without being constrained by it; he should keep an open mind and be prepared for the unexpected; we never do know what the next language will be like. The second view implies a serious research program, that of uncovering the principles of structure that can be hypothesized to underlie all human languages. Such a program was implicit in the theorizing of European structuralists in the last century, especially Hjelmslev and Jakobson, and has been advanced by their American successors, most forcefully, since the 1970s, by Noam Chomsky and his disciples.

Unlike his predecessors, Chomsky explicitly addressed the question of the nature of our innate capacity for language. In his view, every individual is born with a 'Universal Grammar', which is understood as a language matrix (he has called it a 'language organ' in the mind/brain) with a set of structural options for the syntax (and the phonology), and some basic principles of grammar formation, and, of course, plenty of storage space for a vocabulary (or lexicon). In this perspective, the 'task' of

the language learner who is exposed to a community language is to select options (or set parameters) and acquire the lexicon in accordance with the way that language is spoken.

This conception of Universal Grammar assumes that the categories of language are universal and correspond to our equally universal cognitive categories. But there is in fact one cognitive category that is well represented in some languages but practically or totally absent in others. What is even more remarkable, in languages where this category is represented it shows an extreme degree of syntactic idiosyncrasy. Indeed this category gives the impression that it does not fit well into the normal molds of languages. It appears to test the limits of Universal Grammar, or, if we look at the matter from another angle, it demonstrates the plasticity of Universal Grammar.

The cognitive category in question is that of numbers.

Here I will say very little about the fact that some languages have no number words, or numerals (section 1) and only a little more about two problems numerals appear to represent for Universal Grammar. One problem is what we can call nomination. This is the problem of naming higher numbers by means of lower numerals – in the larger perspective, of coping with the infinity of numbers by means of finite linguistic means (section 2). The other problem is that of fitting number expressions into clauses, which is to assign them to (one or more) syntactic classes. This will be the topic of section 3.

Languages without numerals

Neurological research has determined that infants 5 months of age have a clear sense of numbers [2; see also 3, 4]. From this one might expect that numbers are universal, and that all languages would have numerals to represent them. But in fact languages are known that make or made no use of numbers and have no native numeral expressions, in New Guinea, Australia, the Andaman Islands, Brazil. There are not many languages in this category, but from the ones that are known we can confidently infer that where numeral systems exist, they are a cultural attainment, that is, they have developed (or been borrowed from other languages) because they were culturally motivated. This means that languages entirely without numerals, such as Dani (Papua) [5] or Pirahã (Brazil) [6] represent the original state of affairs of human languages. Such languages have the universal quantifiers (one, some, all, none) and relative quantifiers (many, few), but no numerals or no precise numerals; cf. Nadëb in (1). Some languages have 'minimal' systems, such as Amanab (2). In others, a minimal system has been supplemented with words for 'hand', 'foot' as in Alamblak (3). They suggest how minimal systems of numerals may have been elaborated in the past and, in time, begun to reach toward infinity.[1]

(1) Nadëb (Brazil) 1: šæd, 2: pɔwɔp (also '3, a few'), 3: tamawɔb; and words for 'several', 'many', 'all' [7].

(2) Amanab (Papua) ¹*mungu*, ²*sabaga*, 3: ²*sabaga* ¹*mungu*, 4: ²*sabaga* ²*sabaga*, 5: ²*sabaga* ²*sabaga* ¹*mungu* [5].

(3) Alamblak (Papua) ¹*rpa*, ²*hos*, 3: ²*hos-f-i-*¹*rpa-t*, 4: ²*hos-f-i-*²*hos-f*, 5: *tir* 'hand', 10: ⁵*tir* ²*hos-f*

15: ⁵*tir* ²*hos-f-i* ⁵*wura yoht-t* 'two hands and one whole foot'.

6:	*tir*	*yoht-t-i*	*anakor*	*tir-t-ho*	*rpa-t*
	hand	whole-3SG:F-CONJ	other=side	hand-3SG:F-POSS	one-3SG:F

'one whole hand and one from the other hand' [5].

The nomination of numbers

The basic device for naming a higher number by means of lower numerals is to dissolve the number into addends, as in (2). Addends may be joined with a word for 'and' or 'over', as in the nursery-rhyme's 24: ^{4}*four-and-*^{2}*twen-*^{10}*ty*, German 24: ^{4}*vier-und-*^{2}*zwan-*^{10}*zig* or Welsh 14: ^{4}*pedwar ar* ^{10}*ddeg*, or without one, as in Gm. 14: ^{4}*vier-*^{10}*zehn*, W. 15: ^{5}*pym-*^{10}*theg*, cf. W. 16: ^{1}*un ar* ^{5}*bym-*^{10}*theg*.

Addition may be supplemented with some use of subtraction (e.g., Latin 17: ^{7}*septen-*^{10}*decim*, but 18: ^{2}*duo-de-*^{20}*viginti* , 19: ^{1}*un-de-*^{20}*viginti*) or with 'overcounting', e.g., Tibetan 15: ^{10}*teri* ^{5}*pungu*, but 16: ^{20}*metsv maben* ^{6}*trok*, … 19: ^{20}*metsv maben* ^{9}*tvko*, 20: ^{20}*metsv*, where *maben* means 'before (i.e. on the way to)' [9]. Addends may be more or less implicit as in Finnish 15: ^{5}*viisi-toista* 'five of the second [ten]', as may other relational numbers, cf. Old Church Slavonic 25: ^{5}*pęti meždu* ^{10}*desętima* 'five between (two) tens'; and, with overcounting, Lithuanian 15: ^{5}*penkio-lika* 'five left behind [towards 20]'.

Where multiplication occurs, it implies addition. Factors can be small or large. Karen (4) uses the most elementary one, *swá* 'doubled'. Welsh and Breton are notable for their formations of 18: W. ^{2}*deu-*^{9}*nau*, Br. ^{3}*tri-*^{6}*wec'h*. Common factors are 10, as in English (^{2}*twen-*^{10}*ty*, ^{3}*thir-*^{10}*ty*, etc.), and 20, as in Dzongka (5) and Danish (6). These two also illustrate the use of fractions (divisions) in overcounting (e.g., Danish 'half the third' for 'two and a half'), Danish, an explicit word for 'times' as well.

(4) Karen (Tibeto-Burman) ^{1}*ɩ̃- * ^{2}*nↄ̃*, ^{3}*sō*, ^{4}*lwī*, ^{5}*ngɛ̃*, 6: ^{3}*sō swá*, 7: ^{3}*sō swá* ^{1}*ɩ̃-*, 8: ^{4}*lwī swá,* 9: ^{4}*lwī swá* ^{1}*ɩ̃-*, ^{10}*chↄ̃· swá* 'doubled' [9].

(5) Dzongka (Tib.-Burm.) 20: ^{20}*khe* ^{1}*ci:*, 30: ^{20}*khe* $^{1/2}$*pjhe–da-*^{2}*ňi:* 'twenty, half [a twenty] to two, 40: ^{20}*khe* ^{2}*ňi:*, 50: ^{20}*khe* $^{1/2}$*pjhe–da-*^{3}*sum* 'twenty, half [of twenty] to three', 55: ^{20}*khe* $^{3/4}$*ko* ^{3}*sum* 'twenty, three quarters [of twenty] to three', 60: ^{20}*khe* ^{3}*sum*, …, 100: ^{20}*khe* ^{5}*nga*, 200: ^{20}*khe* ^{10}*cu-thām* (*thām* 'complete'), 300: ^{20}*khe* ^{10}*ce-*^{5}*nga*, 400: ^{400}*ňiçu* ^{1}*ci:*, 600: ^{400}*ňiçu* $^{1/2}$*pjhe–da-*^{2}*ňi:* 'four hundred, half [of four hundred] to two', 1100: ^{400}*ňiçu* $^{3/4}$*ko* ^{3}*sum* 'four hundred, three quarters [of four hundred] to three' [9].

(6) Danish. 20: ^{10s}*tyve* [lit. 'tens'], 40: ^{4}*fyrre-*^{10s}*tyve* ['four-tens'], 50: $^{1/2}$*halv-*^{3rd}*tred-sinds-*^{20}*tyve* [lit.: half-the-third-times-twenty], 60: ^{3}*tre-sinds-*^{20}*tyve* ['three-times-twenty'].

(7) No. 10: *ti*, 10²: *hundre*, 10³: *tusen*, 10⁶: *million*, 10⁹: *milliard*, 10¹²: *billion*, 10¹⁵: *billiard*, ….

(8) Mayan 10: *lahun*, 20: *hun*, 20²: ^{400}*bak*, 20³: ^{8000}*pik*, 20⁴: ^{160,000}*calab* [10].

Among languages that utilize multiplication, some make use of exponentiation. Base-10 languages like Norwegian (7) and base-20 languages like Mayan (8) are illustrative.

Thus, when we compare systems of numerals we see a progression from no numerals to atomic numerals, to numerals formed through arithmetic operations, and among these, from addition (and subtraction) through multiplication (and division) to exponentiation. It is reasonable to suppose that this progression in arithmetic sophistication has been driven by advances in cultural complexity [10].

Although many languages make use of words for the arithmetic operations required to compute a number from a phrasal (or compound) numeral (e.g., 'and', 'times'), most commonly, addends and factors are simply juxtaposed; cf. (2), (4). This presupposes (i) that addition and multiplication are signified by distinct orders, as in Danish 104: ^{100}*hundrede* ^{4}*fire* and 400: ^{4}*fire* ^{100}*hundrede*; (ii) alternatively, if the same order is used for

addition and multiplication, then any numerals that serve both as addends and factors must have distinct shapes (or allomorphs) as in Eng. -^{10}*teen* and -^{10}*ty* (cf. 16: ^{6}*six*-^{10}*teen* and 60: ^{6}*six*-^{10}*ty*).

This seems perfectly reasonable, *ex post facto*. But it is in fact remarkable that these expression devices – allomorphy and element order – have been harnessed to those functions. Outside of phrasal (or compound) numerals, allomorphs are systematically synonymous and do not occur in the same environment. But these apparent allomorphs have been specialized to co-signify a number and an operation. On the other hand, alternative orders of identical lexemes are either excluded in a given language (cf. Eng. *(the) blue dress*, but not **(the) dress blue*), or they serve to indicate different information structure (cf. Russian *goluboe plat'e ~ plat'e goluboe* '(the) blue dress'), but not distinct ways of combining the referents of lexemes. In the construal of *(the) blue dress*, the reference potential of the phrase is the union of some individual dress and all things blue; there is no way the referents of the two constituents could be either added or multiplied.

What we have in the formation of phrasal (or compound) numerals such as 16: ^{6}*six*-^{10}*teen* and 60: ^{6}*six*-^{10}*ty* or Da. 104: ^{100}*hundrede* ^{4}*fire* and 400: ^{4}*fire* ^{100}*hundrede* are examples of linguistic means – allomorphy and element order – employed in functions that are alien to non-numeral language.

The external syntax of numerals

Unlike any other semantic class, numerals occur in different languages variously as verbs, as adjectives, and as nouns.

One specialist has inferred from this fact that numerals must basically be adjectives [11]. But one might as well draw the inference that Universal Grammar has no pre-ordained part of speech for numerals. In fact, a close look at numeral systems reveals that although numerals form a well-defined semantic class, they do not fit any part of speech particularly well and are commonly distributed among several.

Yurok is a language in which numerals are verbs, that is, 1–4 are verbs, whereas higher numerals 5 and 10 are nouns, constructed with a verb meaning 'it is a collection'. The Yurok verb includes a classifier that shows what kind of thing the sentence is about; likewise also when there is a multiple of 10. The expression for 'thirty-one logs' breaks up into 'there are three tens and there is one log'; see (9).

In other languages, numerals are adjectives or nouns. In Latin, all are adjectives except the plural of 'thousand' (10); in Lithuanian, 1–9 are adjectives, the rest are nouns (11); whereas in Old Church Slavonic, only 1–4 are adjectives and the rest are nouns (12). But note how in Latin, many numerals deviate from normal adjectives by having no inflection; in Lithuanian, it is 10 and its multiples that are uninflected. This morphological coyness of numerals fits in well with the fact that even though they sometimes look like adjectives (e.g., *the old books* like *the three books*), they mostly do not behave like adjectives. Books can be *old and dusty*, but not **three and eight*; they can be *very dusty* or *dustier*, but not **very eight* or **eighter*. Numerals are not a class of adjectives.

(9) Yurok *ceyk-ok's* it is narrow' (*-ok's-* CL 'flat thing'), ^{1}*koht-ok's* it is one (flat thing)', ^{2}*no'-ok's* 'it is two (flat thing)', ^{3}*nahks-ok's* 'it is three (flat thing)', ^{1}*koht-e'r* 'it is one' (*-e'r-* CL 'stick-like thing'); 1: ^{1}*koht-e'r tektoh* 'it is one log', 2: ^{2}*na'-a'r tektoh*, 3: ^{3}*nahks-e'r*

tektoh, 5: [5]*meruh tom-u'r tektoh* (*tom-* 'be a collection'), 10: [10]*wehlowaa tom-u'r tektoh*, 13: [10]*wehlowaa nima* [3]*nahks-e'r tektoh*, 15: [10]*wehlowaa nima* [5]*meruh tom-u'r tektoh*, 20: [2]*na'-ami* (*-ami-* CL 'numeral'), [10]*wehlowaa tom-u'r tektoh*, 31: [3]*nahks-emi* [10]*wehlowaa nima* [1]*koht-e'r tektoh* 'there are thirty-one logs' [12].

(10) Latin [10].

Adjectives	Simple	Complex
Declined	[1]*unus*, [2]*duo*, [3]*tres*	[200]*ducenti* [300]*trecenti* … [900]*nongenti*
Indeclinable	[4]*quattuor* … [9]*novem* [10]*decem*, [100]*centum*, [1000]*mille*	[11]*undecim* … [17]*septendecim* [20]*viginti*, [30]*triginta* … [90]*nonaginta*
Nouns	[1000]*milia*, nt.pl.	

(11) Lithuanian

	Simple	Complex
Adjectives	[1]*vienas*, [2]*du* … [9]*devyni* (pl.t.)	
Nouns	[100]*šimtas*, [1000]*tūkstantis, milijonas, ...*	[11]*vienuolika*, … [19]*devyniolika* (sg.t.)
Indeclinable	[10]*dešimt*	[20]*dvi-dešimt*, … [90]*devynias-dešimt*

(12) Old Church Slavonic.   Adjectives: [1]*jedinŭ*, [2]*dŭva*, [3]*trĭje*, [4]*četyre*:

Nouns: [5]*pętĭ* … [10]*desętĭ*, [100]*sŭto*, [1000]*tysǫšta*, [10,000]*tĭma.*

What numerals really are is something different. Semantically they differ radically from nouns, adjectives and verbs. These parts of speech are used to describe what sorts of things, properties, or situations sentences are about; but numerals have no descriptive content. While the descriptive content of, say, a noun can help us identify its referent, a numeral can only tell us how many there are of whatever the noun refers to; thus the noun and the numeral provide us with truly complementary kinds of information, descriptive and quantitative [14]. It is significant that in some of the simplest numeral systems, such as Nadëb, the words for 'two' and 'three' have overlapping meaning; see (1) again. Here the kinship between numerals and the relative quantifiers 'few', 'many' is obvious. It is only through their later development that numerals have become distinct from the relative quantifiers by their absolute numerical value. Maybe such precise values arose only with the introduction of addition, as in Amanab (2).

It appears that throughout their development, numerals have been a liminal category in languages. We cannot say that numeral systems can vary without limit, but they stand out by their diversity of syntactic properties, often mixed and varying from language to language, evidence of makeshift extensions of grammatical devices that are basically used for other purposes. By their historical dependence on the development of cultures, numerals suggest an evolutionary perspective on Universal Grammar. It is apparent that during the time that numerals have existed in languages, our innate capacity for language has not evolved to accommodate this cultural attainment as a basic lexical/syntactic category. One can imagine that the existence of numerals will favor such an evolutionary step some time in the future. Looking back, one can wonder what sequence of prior developments in human culture have favored the gradual evolution of existing grammatical parameters and principles of grammar formation.[1]

End-notes

1. I am grateful to members of the 2004–2005 CAS team for comments on the oral version of this paper. Special thanks go to Dr. Rosemary Tannock for several useful references.

References

[1] Robins, R. H. 1967. A Short History of Linguistics. London: Longmans.

[2] Kobayashi, Tessei, Kazuo Hiraki, Ryoko Mugitani, Toshikazu Hasegawa. 2003. "Baby arithmetic: one object plus one tone". *Cognition* 91.B23–B34.

[3] Dehaene, Stanislas, E. Spelke, P. Pinel, R. Stanescu, S. Tsivkin. 21999. "Sources of mathematical thinking: behavioral and brain-imaging evidence". *Science* 284.970–974.

[4] Dehaene, Stanislas, Nicolas Molko, Laurent Cohen and Anna J. Wilson. 2004. "Arithmetic and the brain". *Current Opinion in Neurobiology* 14.218–224.

[5] Claudia Gerstner-Link. MS. "Typology of numeral systems in Papua languages". Paper presented in a Workshop on Numerals in the World's Languages, Max Planck Institute for Evolutionary Anthropology, March 2004.

[6] Daniel L. Everett. MS. "On the absence of number and numerals in Pirahã". Paper presented in a Workshop on Numerals in the World's Languages, Max Planck Institute for Evolutionary Anthropology, March 2004.

[7] Pattie Epps. MS. "Tracing the development of numerals in the Guaviareå-Japuraå (Maku) family". Paper presented in a Workshop on Numerals in the World's Languages, Max Planck Institute for Evolutionary Anthropology, March 2004.

[8] Jadranka Gvozdanović. 1992. *Indo-European Numerals*. Berlin–New York: Mouton de Gruyter.

[10] Martine Mazaudon. 2002. "Les principes de construction du nombre dans les langues tibéto-birmanes". *Mémoires de la Sociéteé de linguistique de Paris* (Nouvelle série) 12.91–119.

[11] Karl Menninger. 1958. *Zahlwort und Ziffer. Eine Kulturgeschichte der Zahl*. (2. neubearbeitete und erweiterte Auflage.) Göttingen: Vandenhoeck und Ruprecht.

[12] James R. Hurford. *Language and Number. The Emergence of a Cognitive System*. Oxford: Basil Blackwell.

[13] Juliette Blevins. MS. The morphology of Yurok numerals". Paper presented in a Workshop on Numerals in the World's Languages, Max Planck Institute for Evolutionary Anthropology, March 2004.

[14] Henning Andersen. 1998. "Does the past have a future. Reflections on the Jakobson Heritage", *The Roman Jakobson Centenary Symposium, Copenhagen 10–12 October 1996* ed. by Per Aage Brandt et al., *Acta Linguistica Hafniensia* 29.149–177.