

**ANALYSING SMALL DATA: LESSONS FROM THE TYPOLOGY OF NUMERAL SYSTEMS**

Bernard Comrie, UCSB <comrie@linguistics.ucsb.edu>

(Abstract for Summer School session)

The typology of numeral systems might seem like a small and arcane sub-sub-field of linguistics, one to which few linguists have paid any serious attention. Nonetheless, it presents in a microcosm a number of interesting phenomena that are of general linguistic interest, including from a methodological perspective. Some of the issues that will be discussed – with appropriate audience participation (no hand-held calculating devices allowed!) – will include the following.

While all humans have the ability to acquire any natural language numeral system, many languages restrict their numeral systems considerably, for instance having numerals only up to ‘three’. What psychological processes might underpin such a limitation? Does the extension of numeral systems to higher numbers represent cultural evolution? Are there natural language numeral systems with no upper limit?

A general issue in linguistics is whether the attested cross-linguistic variation is representative of the human potential for cross-linguistic variation. For languages as a whole, the answer may be unknowable. For numeral systems, it is almost certainly “no”: numeral systems are frequently more endangered than the languages they form part of; what is currently accessible to us is probably only a fraction of what once existed, let alone of what is possible.

Different numeral systems have different bases, e.g. decimal (base 10), vigesimal (base 20). Are there constraints on possible bases? If so, are they somatic, i.e. reflecting properties of the human body (“If God had wanted us to use the decimal system, He’d have given us ten fingers”) or commercial (it’s easier to pack things in twelves or twenties than in tens), and what is the relation between numeral bases and these other factors?

What arithmetic operations are used in natural language numeral systems, and what are the constraints on them? Might a language express 57 as ‘sixty-two less five’? If not, why not?

The near-universal representation of numbers by means of figures uses a position-based decimal system ordering powers of ten from higher to lower, e.g. ‘hundreds, tens, units’. The numeral systems of nearly all (if not all) languages depart from this to varying degrees. Does this create processing and/or production problems for the interface between natural language numerals and representations by means of figures? How do speakers / writers solve this problem? Are they even aware of the problem?