

Some core contested concepts

Noam Chomsky (Massachusetts Institute of Technology)
chomsky@mit.edu

The most fundamental question we can ask about language is: What is it? (Q). Curiously, a question rarely raised, usefully at least, in classic texts, or even recent work where it should be relevant.

On the basis of a tentative “best theory” T about Q we can proceed to ask further questions, among them how the concepts of T enter into acquisition, use, evolution, neural representation, historical change (not to be confused with evolution). Reciprocally, what is learned about these topics can lead to recasting of T; the logical hierarchy of questions does not of course determine order of research into them.

By *language* here I mean *I-language*. Another concept that appears in the technical literature is *E-language*, now conventionally used in ways that do not seem to me coherent.

I-languages are systems of discrete infinity, so their study falls within the theory of computation (Turing machine theory, the theory of recursive functions). An I-language can be taken to be a computational (generative) procedure GP, internal to the mind/brain, which yields structured expressions, each assigned an interpretation at two interfaces, sensorymotor (SM) and conceptual-intentional (CI). So regarded, I-language satisfies a classical concept of language as sound with meaning – more accurately, it seems, meaning with ancillary sound (or other externalization), a very different matter.

Recursion is often confused with center-embedding; these are very different notions. Sometimes recursion is incorrectly assumed to be necessarily infinite. A recursive function may yield a single output (or nothing). It is easy to construct a GP lacking resources to generate more than a fixed number of expressions – more precisely, of non-deviant expressions. Were such an impoverished language to exist, it would be a mere curiosity, with no implications for the general study of language, contrary to substantial recent literature.

Among other such devices, English and other I-languages freely generate center-embedding, though without external aids (time, memory), subjects of course cannot interpret such structures beyond some bound – about 7, as predicted by Miller’s famous magic number. Self-embedding, a narrower notion, is much more sharply restricted, a fact of interest for the study of perception/parsing. Discussion of these topics is sometimes obscured by failure to distinguish competence and performance; loosely, what we know and what we do. The distinction is contested, but needlessly. It is familiar in investigation of any organic system, cognitive or not. Where competence is unbounded, use is restricted by available memory (e.g., arithmetical knowledge). To put it differently, there is a crucial difference between a Turing Machine with all memory in the control unit, and one that relies on external memory, though with bounded resources the two cannot be distinguished by superficial experiment. Note that the distinction is not to be confused with Marr’s analysis of levels of abstraction for processing systems, similar in spirit but different conceptually.

Acquisition of language, like development of any organic system, involves several factors: (i) external data, (ii) genetic endowment, (iii) more general principles. Factor (ii) includes (a) language-specific UG, (b) other cognitive systems, (c) constraints imposed by structure of the brain, a potential source of such information. Existence of (iia) is contested, but apparently on the basis of misunderstanding. For a computational system like language, (iii) can be expected to include at least principles of computational complexity. Though this too is contested (mistakenly, I believe), virtually all of what has been learned about I-language derives from (iia) and (iii) (and of course (i)).

Uncontroversially, we seek the simplest theory of UG, eliminating stipulations and avoidable complexity. From the earliest origins of the modern study of these topics 60 years ago, research has been directed towards this goal. In recent years this research has sometimes been called *the minimalist program*, a notion that has been widely misunderstood. It is a seamless continuation of early inquiry, differing only in the suggestion of some new research programs, which have led to interesting and sometimes far-reaching empirical conclusions, and have some independent motivation in terms of the very little that can be said with any confidence about how language evolved.