

On clitics and canons

Canons are uncontroversial sets of descriptive criteria which generally converge on a single linguistic construct ('agreement', 'suppletive form', ...) and define a logical space in terms of sets of polar oppositions. We propose canonical properties for 'word', 'clitic', and 'affix'. A word is canonically a monomorphemic noun/verb lexeme heading a phrase and distributed according to principles of syntax. An affix is a bound element attached to the lexical head being marked for that property. Affixes take narrow scope over coordinated phrases. A clitic is a prosodically non-prominent element dependent on a full-word host. Clitics attach promiscuously to the relevant NP/clause ('phrasal affix' property), distributed according to clitic-specific principles, canonically in Wackernagel position. They take wide scope over coordinated phrases.

We show that our fairly uncontroversial set of clitics canons fails to define a coherent notion. This is because the clitic is a 'mixed construct'. We show that for every putative canonical clitic property it makes more sense to regard that property as a deviation from either the canons for (function) words or for affixes. As a result, there is no more reason to believe in a canonical clitic than a canonical suspendable affix (e.g. in Turkish) or conversely, a canonical function word obligatorily repeated on conjuncts (e.g. monosyllabic articles/prepositions in some Romance languages), or even a canonical 'non-projecting word' (Toivonen 2003).

Our results are compatible with the misgivings often expressed about the notion 'clitic' and the way it is (mis)used in descriptive/typological discussion. Whatever the heuristic, informal value of a clitic category in a descriptive grammar, we generally find that it has to be hedged by specifying exactly what clitic properties its instances do or don't possess. But given these hedges there is no role in the grammar for an attribute 'clitic', moreover, these hedges are *all* definable with respect to word or affix canons. This makes the notion practically useless for typological analysis. Moreover, this is typical for 'mixed constructs' due to partial grammaticalization (e.g. we wouldn't expect to be able to identify a coherent notion of, say, 'verbal gerund' lying somewhere between a canonical noun and canonical verb).

We argue that the canons-based approach allows us to understand the logical determinants of these misgivings: if the clitic notion cannot be coherently defined in terms of canonical properties then that is because it is not a coherent notion. Rather, it represents a family of deviations from other, canonically definable constructs. We argue that the canonical approach allows us to see the logical structure of this descriptive problem with greater clarity than other descriptive/analytic approaches.