WHAT IS A CLITIC?

0. Introduction

Every language has elements with some of the properties characteristic of independent words and some characteristic of affixes, in particular, inflectional affixes, within words. Such elements act like single-word syntactic constituents in that they function as heads, arguments, or modifiers within phrases, but like affixes in that they are "dependent", in some way or another, on adjacent words.

A range of familiar examples will give some sense of the variety of the phenomena at issue here: (a) In Tagalog, a large assortment of forms can occur only in position immediately after the first word of a clause; in *Hindi ko siya nakita ngayon* 'I haven't seen her/him today' the 1 sg. agent pronominal *ko* and the 3 sg. topic pronominal *siya* cluster together immediately after the clause-initial word *hindi* 'not'. (b) In English, various forms of auxiliary verbs have "reduced" variants that are phonologically dependent on the word immediately preceding them, as in *Your friend from Chicago's going to arrive soon,* with a /z/ variant of *is* attached to *Chicago*. (c) In French, various pronominal words occur in "weak" versions that obligatorily precede the verb of their clause, though the corresponding "strong" phrases do not occur in that position; in *Il en remplit un verre* 'He fills a glass with it', the weak pronominal *en* immediately precedes the verb *remplit,* though a full PP

like *de ce vin* 'with this wine' would follow the direct object phrase. (d) In English, the possessive marker /z/ combines with an entire NP, though it is phonologically dependent on the last word of this NP, as in *your friend from Chicago's arrival*. (e) In English, unaccented object pronouns *her, him,* and *them* have "reduced" versions, consisting only of a syllabic sonorant, which are phonologically dependent on an immediately preceding verb or preposition, as in *We gave'em to 'er* 'We gave them to her'.

All of these elements, and many others similar to them, in hundreds of other languages, have been labeled clitics, because they "lean" on an adjacent word.

1. 'Clitic' as an Umbrella Term

It is argued here that *clitic*, understood in this broad fashion, is an umbrella term, not a genuine category in grammatical theory. Umbrella terms are names for problems, for phenomena that present "mixed" properties of some kind, not names of theoretical constructs. This is what Zwicky (1990) argued is the case for the notion of *serial verb* (serial verb constructions being those that present some characteristics of coordinate combinations and some of subordinate combinations), and the same could be argued for the notions of *compound* (compounds exhibiting some characteristics of single words and some of two-word combinations) and *incorporation* (incorporations exhibiting some characteristics of single words and some of clauses) as well as *clitic*. Each of these terms is ordinarily understood broadly, and on such an understanding none of them embraces a unified class of phenomena.

This conclusion does not mean that there are no constructs of theoretical interest under the umbrella of clitics (or serial verbs or whatever). In fact, there are at least two such constructs (with little in common) under the clitic umbrella: what can be called *bound words* (the Tagalog second-position items, for instance) and *phrasal affixes* (the English possessive suffix, for instance). One might choose to reserve the word *clitic* for one or another of these, but that would be a purely terminological decision, not a matter of substance at all; it would also be a rather confusing choice of terminology, given the widespread use of the word in its broad sense.

Faced with the observation that there is almost nothing shared by all of the things that have been referred to as clitics, many writers (most recently, Hopper & Traugott (1993:ch. 6)) have posited a scale, cline, squish, or continuum between the two poles of independent word and inflectional affix. Such proposals share the defects of many other cline hypotheses (e.g., a scale between derivational affix and compound element, verb and noun, parataxis and hypotaxis, or the thematic roles Agent and Patient): most notably, the two defects that there is no independently definable dimension for variation, so that the cline is an ad hoc creation, and that

there are many different 'paths' between the poles — not one dimension, but a number, along which different items can differ in different ways (see Dowty (1991) for an exemplary demonstration of multi-dimensionality in the case of Agent versus Patient).

2. Clitics and the Components of Grammar

In the case of clitics, in the broad sense, there are many paths between independent word and inflectional affix, indeed paths that wend their way through quite different components of a grammar: phonology, morphology, and syntax. The original notion of clitic is a phonological one, as is reflected in the *Oxford English Dictionary*'s (1st ed.) entry for *enclitic*:

> *adj* That 'leans its accent on the preceding word' (Liddell and Scott):
> in Greek grammar the distinctive epithet of those words which have
> no accent, and which (when phonetic laws permit) cause a secondary
> accent to be laid on the last syllable of the word which they follow.
> Hence applied to the analogous Latin particles *-que, -ve, -ne*, etc.,
> and in mod. use (with extension of sense) to those unemphatic words
> in other langs. that are treated in pronunciation as if forming part of
> the preceding word.

In contrast, it was the morphological peculiarities of clitics — in particular, their being subject to template conditions on their combinability and ordering — that brought them to the attention of theoretical linguists in the 1960s (notably via Perlmutter (1971)). The Tagalog clitics, for instance, are subject to the condition that monosyllabic pronominals must precede other clitics, and that nonpronominal clitics must precede disyllabic pronominals. A concern with clitics as morphological objects continues in such works as Simpson & Withgott (1986) and Zwicky (1992).

Most recent theoretical work, though, has tended to emphasize the syntactic peculiarities of clitics, including their occurrence in the "Wackernagel" position (second within the clause), their connection to other syntactic phenomena (as in "clitic climbing"), and their interaction with agreement (as in "clitic doubling"); see Spencer (1991:ch. 9) for several compact case studies.

There are two main threads in current approaches to clitics by theorists, each of which treats clitics — or at least the interesting ones — as having some special syntax plus some further grammatical peculiarity, in either phonology or morphology. In one thread, exemplified by Klavans (1982, 1985) and Anderson (1992:ch. 8), the single nonsyntactic peculiarity of a clitic is its being phonologically dependent, and so obligatorily adjoined to some adjacent material in

a prosodic domain. In the other, as in Nevis (1988c) and the introduction to Borer (1986), what characterizes clitics is their playing a part in both syntactic and morphological structures.

The two threads are combined in some recent work — Inkelas (1989) and Halpern (1992), in particular — that allows for the possibility of elements with special syntax, morphology, *and* phonology.

3. A Few Common Types of "Clitics"

A variety of phenomena have appeared under the clitic umbrella but merely have marked properties in one or more components of a grammar. The claim advanced here is that these marked properties are in principle independent of one another (even if they tend to occur in combination). They are simply marked options that must be made available, one by one, in the grammars of the world's languages, and no special place need be made for them in linguistic theory.

3.1. Lexicon, Morphology, and Syntax

The first such items to be surveyed here have marked properties in the lexicon, morphology, or syntax.

3.1.1. Grammatical Category Lexemes

These are lexical items — for example the English agentive preposition *BY*, the comparative determiner *MORE*, the possessive preposition *OF*, and the future modal verb *WILL* — that are affix-like in their semantics. Like affixes (and unlike ordinary lexical items), they express grammatical categories — case, degree, tense, and the like — rather than more concrete meanings.

Every language has an inventory of such grammatical category lexemes, variously called "function words", "grammatical words", "particles", "markers", or (because the inventory expands very slowly indeed) "closed category words". Most lexical items in a language have more concrete meanings, but grammatical category lexemes are entirely natural, though marked.

3.1.2. Marginal Category Members

These are lexical items — among them the English articles *A* and *THE*, infinitival *TO*, complementizers *FOR* and *THAT*, negator *NOT*, and the "particles" in verb-particle combinations — that are affix-like in that they are not easily assigned to a syntactic category. The temptation is to treat each as *sui generis*.

There is now a fair amount of literature in which it is argued that particular instances of such lexemes are in fact members of major syntactic categories, though members with special (sometimes idiosyncratic) properties in addition to the properties they share with other members of their category. The arguments of Pullum (1982) that English infinitival *TO* is an (exclusively nonfinite) auxiliary verb and of Emonds (1985) that "particles" of verb-particle combinations are (intransitive) prepositions are paradigms here. Following in this vein, Zwicky (1985a) advanced the hypothesis that there are no acategorial lexical items, that upon investigation it will turn out that each lexeme is assignable to a syntactic category, albeit a somewhat idiosyncratic subcategory of it. Marginal category members would then be not really analogous to inflectional affixes; they would merely be marked, unusual, members of one of the standard categories.

3.1.3. Invariables

These are lexical items, like the English modal *MUST* and infinitival *TO*, that are affix-like in not showing overt inflectional morphology.

Nevertheless, upon more careful inspection these items can be seen as merely having a "zero" realization of any relevant grammatical categories, much as the English nouns *SHEEP* and *MOOSE* have "zero" realizations of the grammatical category of plural number. For at least certain syntactic categories in some languages, having no overt mark is a morphological peculiarity, though a widespread peculiarity in the world's languages.

Marginal category membership and invariance often go together. Elements that analysts are inclined to assign to some ad hoc class, or to refer to by their phonological content, are marginal category members or invariables or both.

3.1.4. Loners

These are syntactic words that are affix-like by virtue of not allowing phrasal modifiers or arguments. English loners include the inverted auxiliaries (*Really must you go?* alongside *You really must go*), "particles" contiguous to their transitive verbs (*We gave right up the fight* alongside *We gave the fight right up*), and the degree modifier *ENOUGH* (*This is small much enough to lift* alongside examples with the degree modifier *TOO*, like *This is much too big to lift*).

It is undoubtedly marked to have syntactic words that must stand alone like this. Nevertheless, they occur with modest frequency in the world's languages.

3.1.5. Syntactic Dependents

These are syntactic constituents that are affix-like in that they are distributed with respect to some companion element (a word or phrase), as are clitics in the Wackernagel position. In English, for example, direct objects are located with respect to — in particular, immediately following — their head verbs (*We passed quickly the other racers* alongside *We passed the other racers quickly*).

Again, this is a marked characteristic of syntactic constituents, which are usually "equal participants" in combinations.

This characteristic can occur along with the previous one, giving a syntactically dependent loner — the type of element which, to adapt Wackernagel's own term, can be labelled a "quasi-clitic" (like the Latin adverb *IGITUR* 'therefore'). The combination of two independent, but both marked, somewhat affix-like characteristics is certainly very marked. But nothing makes the result anything less than a syntactic word, at least so far as the synchronic grammar of the language is concerned.

3.2. Phonologically Special Items

Next in this survey are various ways in which syntactic words can have marked phonological properties, in particular, properties that make them seem to some degree affix-like.

3.2.1. Idiosyncratic Sandhi Targets

Like many inflectional affixes, some lexemes are subject to special (morphological) alternations in *shape* (in the sense of Zwicky 1992), triggered by the phonology, syntactic category, and/or syntactic structure of adjacent words within an expression. For instance, the English indefinite article *A* exhibits an idiosyncratic alternation *a/an*, triggered by the phonology of the following word, and the Spanish definite article *LA* exhibits an idiosyncratic alternation between the default *la* and the special shape *el,* conditioned by (at least) the phonology and syntactic category of the following word.

Such shape alternations are undoubtedly marked, and they usually (perhaps always) affect grammatical category lexemes. But there is no doubt that the affected elements are in fact lexemes.

3.2.2. Obligatory Leaners

Again like many inflectional affixes, some lexemes have no free-standing shapes; they cannot form phonological domains on their own and must lean

phonologically on an adjacent word. In English, such leaners include the articles *A* and *THE*, and infinitival *TO*.

Obligatory leaning is undoubtedly marked, and it is usually (perhaps always) confined to grammatical category lexemes. But again there is no doubt that the dependent elements are in fact lexemes.

3.2.3. Accentless Words

Many obligatory leaners — the English articles *A* and *THE*, for instance — are such because they are prosodically defective, unable to receive accent (except in the sort of contrastive utterance where even meaningless syllables can be accented). There are, however, obligatory leaners that are accentable, for instance, the English infinitival *TO*: accentable in *TO invite them would be insulting*, but not available when it has no material it can lean on, as in **To would be insulting* (Zwicky (1982)).

Some lexemes — the English personal pronouns, for instance — have been classified as "clitics" merely because they have accentless (and, as a result, phonologically dependent, and often reduced) variants. But again, being able to occur accentless, with all that implies for prosodic organization and the applicability of automatic phonological processes, is merely a marked property of certain lexemes.

3.2.4. Phonologically Located Words

Words within a sentence can seem affix-like because their distribution is accounted for in part by virtue of their having to satisfy some condition on prosodic organization. This is the case for Hausa emphatic *FA* (Zec & Inkelas (1990)), which must occur immediately after a phonological phrase (in addition to having to satisfy some modest conditions on syntactic structure).

Being phonologically located, being accentless, and obligatorily leaning are all ways in which material can be subject simultaneously to syntactic conditions and to conditions on prosodic organization. Although the distribution of most lexemes can be sufficiently described in terms of their syntactic properties (and their semantics and pragmatic values), there are some that are "prosodically subcategorized" as well.

4. Items that Are Special in Both Syntax and Morphology

Next are two classes of items that appear not only to be special in both syntax and morphology, but also to require an analysis that links the two peculiarities.

4.1. Bound Words

These are items that seem clearly to be words in syntax but also function as parts of a special type of composite lexeme in morphology. They are clitics *par excellence*.

The second-position elements of Tagalog (for details, see Schachter & Otanes (1972)), for instance, have a number of word-like characteristics; some of them are in complementary distribution with undoubted words, several of them are bundles of grammatical categories (like grammatical category lexemes), and there is even a certain amount of freedom in their ordering with respect to one another. Nevertheless, their syntax is special (they have to occur in the Wackernagel position), and this syntactic peculiarity is linked to a kind of morphological peculiarity (their template requirements, which are unlike conditions on syntactic words otherwise, but similar in a number of ways to template requirements on inflectional affixes, in languages with multiple slots for such affixes). These "clitic groups" are in some ways like compound lexemes (they are made up of lexemes, to start with), but they do not behave syntactically like words. And they are in some ways like inflected forms, though some of the kinds of ordering conditions in clitic templates (by number of syllables, for instance) are quite unlike attested schemes for the ordering of slots for inflectional affixes.

In any case, the syntactic peculiarity of these bound words predicts their morphological peculiarity, and vice versa. It is only the second-position elements that are subject to the template requirements, and all of them are.

The classification of an element as a bound word is not always sure, especially when dependent elements do not pile up as they do in Tagalog. The English reduced auxiliaries, often treated as bound word clitics, might be analyzed instead merely as idiosyncratic sandhi targets.

4.2. Phrasal Affixes

Bound words behave like syntactic words, that is, formatives that combine with other syntactic formatives. There are also special elements that behave like inflectional features, like properties of entire syntactic phrases. When these elements are also morphologically special — in particular, when they are located in an extra, outer, layer of inflection in morphology — then we have the profile of a *phrasal affix*, like the English possessive suffix in *the person with that cat's judgment*.

Again, there is a special link between a syntactic peculiarity (functioning as an inflectional feature rather than a syntactic formative) and a morphological peculiarity (location in an outer layer of inflection, or in other words, combining with an already inflected word, as in *the person who played's judgments*).

5. Conclusion

Every language has a number of elements that are in some way "dependent" (syntactically, morphologically, morphonologically) on adjacent material. For the most part, these elements are words, and their special characteristics are simply marked options in the grammars of languages. Marked options are, of course, a worthy object of study, and much of the clitic literature investigates them.

In some cases, dependent elements turn out not to be words at all, but rather inflectional affixes. This seems to be the right analysis for pronominal "clitics" in some colloquial varieties of modern Romance languages; see Miller 1991, in particular, for arguments of this sort for modern spoken French.

There remain at least two types of words with linked special properties in both syntax and morphology (and phonology as well): bound words and phrasal affixes. Though these surely do not together constitute a uniform class of phenomena, each of them merits study on its own.

6. References Not in Bibliographic Entries A – Z

Anderson, Stephen R. 1992. *A-Morphous Morphology*. Cambridge: Cambridge Univ. Press.

Dowty, David R. 1991. "Thematic Proto-roles and Argument Selection". *Language* 67.3:547–619.

Emonds, Joseph E. 1985. *A Unified Theory of Syntactic Categories*. Dordrecht: Foris.

Halpern, Aaron L. 1992. *Topics in the Placement and Morphology of Clitics*. Ph.D. dissertation, Stanford Univ.

Hopper, Paul J. & Elizabeth Closs Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge Univ. Press.

Zwicky, Arnold M. 1990. "What Are We Talking about When We Talk about Serial Verbs?". *OSU WPL* 39:1–13.

_____. 1992. "Some Choices in the Theory of Morphology". *Formal Grammar: Theory and Implementation*. Ed. by Robert D. Levine, 327–71. Oxford: Oxford Univ. Press.

A.M.Z.