The evolution of c-structure: prepositions and PPs from Indo-European to Romance[1]

NIGEL VINCENT

Department of Linguistics, University of Manchester, M13 9PL, UK

nigel.vincent@man.ac.uk

1. INTRODUCTION

Research in historical syntax sometimes seems in danger of polarization into two camps, each making diametrically opposed assumptions on a number of issues including those pertinent to the present paper, namely word order and constituency. On the one hand, generative historical syntax (cf recent conference collections such as Battye & Roberts 1995, van Kemenade & Vincent 1997) follows synchronic work in that framework in assuming configurational structure as a property of UG. Languages may differ in overt linear order in virtue of different movement possibilities and/or due to different settings of the directionality parameter (cf Kayne 1994 for a proposal to eliminate the latter in favour of the former and Roberts 1997a for an

---

[1]Parts of the material in this paper were presented at a workshop on grammaticalization held during LFG97 at UCSD in June 1997, and at a workshop on functional categories and syntactic change held during ICHL13 at Düsseldorf in August 1997. I am grateful to those who raised questions on those occasions, and to the organizers, respectively Farrell Ackerman and Ans van Kemenade, for inviting me to participate. I am also grateful to Ans again and to 'JC' Smith , who read and commented on early drafts. My thanks too to Andrew Spencer for his comments and notes on the development of prepositions in Slavic and Greek, and to Anna Morpurgo Davies, Jim Adams and David Langslow for conversations on Indo-European and Latin. The work reported here is part of a larger project funded through a Research Readership from the British Academy, whose support is gratefully acknowledged.

exploration of the diachronic consequences thereof), but all languages have some underlying order (indeed on Kayne's proposal they all have the same underlying order), and that order is in turn a reflection of configurational structure (cf Kayne's 'Linear Correspondence Axiom'). By contrast, most work on grammaticalization (Hopper & Traugott 1993: 50ff) seems to treat word order as just that — the linear arrangement of minimal syntactic elements — with no particular regard to the principles or mechanisms (i.e. the 'grammar') which determine such order beyond an assumption of simple concatenation driven either by syntactic or pragmatic considerations or both. To put it (perhaps too) simply, generative grammar seems to err in the direction of assuming too much in the way of linear order/constituency, while proponents of grammaticalization assume too little.

A third possibility lies in what Bresnan (1997) calls 'parallel correspondence theories'. Such approaches treat grammars as formal objects for exactly the same good reasons as inspired the original Chomskyan enterprise (Chomsky 1957: 5), but do not require that the representation of grammatical relations (subject, object, complement, specifier, etc) be couched in the same terms as the representation of overt linguistic structure (case marking, word order, etc). One such theory is LFG, and we will argue that it has much to offer in reconciling the goals for diachronic syntax of the two approaches sketched in the previous paragraph. On the one hand, it provides a formalism within which issues can be formulated and addressed with the precision and explicitness that we have rightly come to expect in morphosyntactic research. On the other hand, the formalism does not inherently privilege constituency and linear order, which are instead seen as one particular type of linguistic realization (cf Matthews 1981: Ch 12) of abstract grammatical relations rather than as the universal metalinguistic vehicle of those relations. The syntactically relevant properties of a given sentence — henceforth f(unctional)-structure — are instead expressed in a

feature-based notation subject to the basic principle of unification. This principle determines how smaller f-structures are combined into larger ones while remaining properly neutral on the question of how they are expressed. Linguistic expression is determined instead by the correspondences between f-structures and c-structures, where the latter may involve configurations, morphological patterns of case and agreement, or any combination thereof.

A key feature then of the architecture of LFG is the separation of f-structure and c-structure (Kaplan & Bresnan 1982).[2] For any given language or construction, the presence of constituency[3] and its frequent concomitant linear order can be seen for what it is: one of the ways — along with case/prepositions (dependent marking) and agreement (head marking) — that languages use to realize ('spell out') their underlying grammatical relations. Languages, and individual constructions within languages, may then differ in the extent of the evidence they offer for constituency just as they differ in the richness of their case and agreement systems. The diachronic consequence of this is clear: just as the histories of languages provide many examples of the rise and fall of case and agreement systems, we should also look for the possibility that constituency may be the product and/or victim of

---

[2]In this sense LFG enshrines the syntactic equivalent of what Aronoff (1994: 8ff) calls the 'separationist hypothesis' in morphology. For an elaboration of the latter from the perspective of LFG, see Börjars, Vincent & Chapman (1997) and Vincent & Börjars (1996).

[3] On a terminological note, I will use constituency to mean the hierarchical, configurational pattern familiar from X-bar theory. C-structure is the representation within LFG which encodes the same information. The difference is that within LFG c-structure may also be 'flat' and/or 'morphological' according to the dictates of the language under analysis, whereas X-bar theory is standardly conceived as a universal configurational template to which all morphosyntactic structures in all languages must conform.

the processes of grammatical change. To the extent that such situations arise, and can be adequately modelled, LFG can be seen to retain the benefits of a formal approach to syntax while still characterizing the types of change that were the original impulse behind work in grammaticalization. In the present paper, we shall examine one such instance in detail — the changing relations between cases and prepositions in a historical cross-section from Indo-European to modern Romance — before reflecting on the theoretical consequences of an LFG-based approach to the study of morphosyntactic change. Of particular interest in this connection are the recent moves to interpret LFG in the light of Optimality Theory (Bresnan 1997, 1998a, b).

2. THE PROBLEM OF PREPOSITIONS

Since the classic article by Jackendoff (1973), it has been standard practice within generative grammar to accord prepositions the status of a major category alongside nouns, verbs and adjectives. Yet the question is by no means clearcut. On the one hand the core prepositions of a language seem to form a small enumerable class of items much like determiners or complementizers; on the other hand if we include so-called complex prepositions like *in light of, by virtue of, in common with,* etc, then the border with NPs is no longer clear (cf also Lehmann 1998). The concept of grammaticalization offers a ready answer here. Original NPs may with the passage of time become compacted down into relation-marking elements. English *because of* is a classic instance, where the presence of the linking *of* betrays the earlier nominal source. Moreover, prepositions cross-linguistically provide good evidence for the principle of grammaticalization sometimes called 'Gabelentz's Law' (Vincent 1993: 145), viz.:

(1) Alle Afformativen waren ursprünglich selbständige Wörten
(All grammatical markers were in origin independent words)

Obedience to this 'law' is the one consistent respect in which prepositions behave like other functional categories, whose realizations are also not historically stable. In fact prepositions show a wide range of etymological sources, including verbs, nouns, participles, adjectives, adverbs and various types of particle (Kortmann & König 1992: 672; and cf Vincent 1997a: 212 for a compilation of representative examples from Italo-Romance). Consultation of the historical grammars of any language with sufficient recorded time depth will reveal that the modern prepositions are (almost) always the product of various grammaticalization processes. In this respect prepositions differ from nouns, verbs and adjectives which typically retain their categorial status over the full length of their recorded history even though they may undergo quite complex and many-layered semantic shifts. Put another way, prepositions are somewhat equivocal members of the class of major lexical categories. In some circumstances they surely do have independent lexical content — e.g. Eng *up/down, on/off, under/over, before/after* and equivalent items in many other languages — but on other occasions they seem to behave more like functional categories, serving to mark various kinds of structural case, e.g. Eng. *of, by* and kindred items in other languages. It is thus preferable to say that prepositions sit astride the lexical/functional divide rather than that they belong unambiguously to either type.

Even if we do accept that prepositions are — or better, can be — a major category, the usual X-bar definition of them as being [–N, –V] is unhelpful in at least one important respect, namely that prepositions in many languages typically govern cases just as verbs do. It seems perverse to make the prototypical verbal properties of taking an object and of determining the object's case depend on the [–N] feature rather than the [+V] one, but this is the only way to achieve the requisite generalization within the canonical X-

bar feature framework. An alternative feature characterization might therefore be the one proposed in Jackendoff (1977) where the defining properties are not categorial [±N, ±V] but relational [±subj, ±obj]. Verbs and prepositions would then share the property of having an object, but be distinguished in that verbs also have subjects and prepositions do not. Yet even this is not a completely satisfactory solution,[4] and the evidence of history suggests a more complex state of affairs. Whereas verbs can for the most part be assumed to govern cases as far back in history as we can go,[5] there is clear evidence that the nexus 'preposition + case-form' which is so much a part of the traditional grammar of Latin, German, Russian, etc is itself the product of a process of grammaticalization.

A second, and logically separate, property of prepositions in modern Romance languages is that they head prepositional phrases, characterized by strict linear order between the members of the PP and special co-occurrence restrictions between the prepositions and their specifiers. To see that these two properties — of case dependency and constituency — do not have to co-

---

[4] In all probability the difficulties of either type of feature system lie in the binarity, and the consequent problems of interpreting classes defined by negative coefficients. In phonology this has long since led to an assumption of unary rather than binary features and a reconsideration of the basics X-theory along similar lines is overdue. For an attractive and intriguing move towards such a theory of categories see Anderson (1997).

[5] Horrocks (1997) argues for the origins of Indo-European grammatical case in an earlier system of semantic relations. Even if this (controversial) view is accepted, however, the fact remains that the acquisition of case-marked dependents by prepositions is a much later development within the Indo-European family. This incontrovertible time lapse between the development of verbal and prepositional case assignment suffices for the purposes of our present argument.

occur, consider the historical development of the verb and the VP. The situation at three representative historical stages is set out in Table I:

	VP	PP
INDO-EUROPEAN	Verbs already govern cases (but cf Horrocks 1997 for a semantic view of (pre-) I-E case)	The ancestors of modern prepositions are independent clause level particles which do not take arguments
LATIN	Verbs clearly govern cases but exhibit no linear effects or other properties that require the postulation of a VP	Prepositions govern cases and, except in poetic style (Penney 1996), members of PP cohere in a single linear unit
ROMANCE	There is clear evidence for a VP according to the usual tests of adjacency, anaphora, and movement	There are strict constraints on order within the PP, and evidence for Spec PP
TABLE I		

Prepositional phrases, then, are of interest because they allow us to track the separate historical development of both government and constituency. Indeed, the historical facts (to be examined in more detail in the next section) present a clear challenge to the usual universalist X-bar theoretic

understanding of phrase-structure. On this view, preposition and PP are part of the major categorial core defined in terms of the features (whether [±N, ±V] or [±subj, ±obj]) and the bar levels. The language particular choice is whether there is a P or not. If there is, then PP, Spec of P, Complement of P come at once and 'for free' in virtue of the instantiation of the universal X-bar schema for the category P. The historical facts suggest, on the contrary, that the classic ingredients of a PP — the P itself, the case-governed complement of P and the category Spec of P — are the product of separate historical processes, and at given historical stages some of these ingredients may exist independently of others, a state of affairs which goes sharply against the usual view of X-theory as module of UG with values set by the parametric choice of head categories. Moreover, the change described here is unusual in that it involves the 'capture' of an argument by an item which previously had no argument structure. This is in contrast to the more commonly studied types of syntactic change which involve loss of argument or theta structure (Roberts 1985, 1997b; Barron 1997). It is also unusual — at least from the perspective of X-bar theory — in that the prepositional head is the grammaticalized result of something that was in origin a specifier. This last point is however consistent with the idea that when an item grammaticalizes, in generative terms it moves 'down' the tree (Roberts 1997b). All in all, prepositions are a complex and difficult category both synchronically and diachronically, and a closer look at the historical scenario is clearly called for. I will start from Latin and work back to Indo-European and forward to Romance.

3. CASE AND PREPOSITIONS FROM INDO-EUROPEAN TO ROMANCE

In this section we will document the stages whereby the full-fledged prepositional syntax of the modern Romance languages emerges. (The discussion inevitably involves a certain amount of detailed supportive argumentation to establish the historical data. Readers whose interest is

principally in the theoretical conslusions which we argue to follow from these patterns may prefer to skip ahead to the summary of this section contained in §3.8, and thence to the more directly theoretical §§4-6.)

3.1 *Particles, preverbs and prepositions*

In Latin we find a range of items which are both prepositions and also components of compound verb stems (henceforth 'preverbs'). Thus:

(2) a) *sub* 'under', *trans* 'across', *in* 'in, on', *ab* 'from', *ob* 'against', *cum* 'with', *ex* 'out of', *pro* 'for', etc.

b) *submittere transmittere, permittere, praemittere* etc
sufferre, transferre perferre, offerre, conferre, etc
subligare, obligare, colligare, deligare, etc.

A link between them is established in two frequently cited remarks by the grammarian Festus on the language of early prayers (cf most recently Cuzzolin 1995):

(3) **Sub vos placo**, in precibus fere cum dicitur, significat id, quod supplico, ut in legibus **transque dato** et **endoque plorato**
"When people say, mostly in prayers, 'sub vos placo', it means the same as 'supplico' and is like the expressions 'transque dato' and 'endoque plorato' in the laws"

(4) **ob vos sacro**, in quibusdam precationibus est, pro vos obsecro, ut sub vos placo, pro supplico
"'ob vos sacro' in certain prayers stands for 'vos obsecro', just as 'sub vos placo' stands for 'supplico'"

Here the classical forms *supplico, obsecro* 'I implore, beseech' are decomposed into their constituent parts *sub placo* and *ob sacro* and the object is placed between them. In the other cited examples the preverbs *endo* (an earlier form of *in*) and *trans* are separated from their verb stems (*dato* 'given', *plorato* 'begged') and act as hosts for the clitic conjunction -*que* 'and'. Other early examples also imply the independence of items that subsequently occur only in construction with a verb or a noun. Festus again quotes a passage from the famous legal text of the 4-5th cent BCE, the Twelve Tables:

(5) si calvitur pedemve struit, manum endo iacito
'if he plays tricks or runs off, lay hands on (him)'

Here *endo* occurs in isolation but in immediate pre-verbal position in a way that would not be acceptable in later Latin, but which has evident congeners elsewhere in Indo-European (Coleman 1991: 327).[6]

Items like *sub, endo* and *ob* were not in origin, therefore, prepositions but rather sentence particles or adverbs which could occupy a variety of positions in the sentence. The three that will be of particular interest in the present context are:

i) sentence-initial and separated from the verb (cf (3, 4))
ii) unstressed and adjacent to the verb
iii) before a local case form

[6] It is the absence of this construction in later Latin which argues against an analysis where *endo* is treated as a preposition with a *pro* object. It is not that Latin did not have object *pro*; it did but never with prepositions (van der Wurff 1993, Luraghi 1997), hence it is unlikely that an early example such as this is to be understood in this way.

For Latin, the alternation between the first two of these is only attested residually in the archaic examples adduced in (3) and (4) but is entirely systematic in Vedic (Renou 1933; Pinault 1995) and is also found in a number of other early Indo-European languages (the traditional name for this construction is *tmesis*[7]). A rather similar kind of alternation between the second and third possibilities is found with a word like *only* in English:

(6) a) Fred has only published two articles
b) Fred has published only two articles

Here too the interpretation of (6a) when pronounced with neutral intonation involves *only* having scope over *two articles* despite its adjacency to the verb and its separation from the NP over which it takes scope.

To sum up, then, in the ancestor language a deictic or orientational sentence particle could bear focal stress in its own right in absolute initial position or it could attach to a sentence constituent over which it had particular scope or it could default to the unstressed pre-verbal position (we assume an unmarked clause final position for the stressed verb[8]).

Given the foregoing, it would be wrong to see *sub vos placo* in (3) as involving a preposition at all (Cuzzolin 1995: 134). *Vos* is accusative as object of *placo* and *sub* is a sentence particle; as Coleman (1991: 324) notes, these particles could occur with intransitive verbs as well as transitive ones, and in

---

[7] As is often noted, the traditional label is something of a misnomer since it suggests a division of something that was formerly a unit, whereas at the point at which 'tmesis' abounds it is precisely because the full efects of so-called 'univerbation' have not yet set in.

[8] Unstressed main verbs migrate to second-position as already noted by Wackernagel: cf Anderson (1993).

the case of the transitive verb it is the particle which is omissible not the noun, a distributional property not consistent with these items being analysed as prepositions.[9]

In the subsequent history of Latin and Romance we find a number of changes which come together to create the class of prepositions and rigidly configurational PPs with which we are familiar from the modern languages.

3.2 *The acquisition of argument structure*

A clear property which distinguishes, cross-linguistically, adpositions from particles is that the former have their own argument structure, both in the semantic sense of selecting particular theta roles and in the syntactic sense of determining the case form of that argument. The passage from Indo-European is no exception and the former particles come to take a dependent argument in a fixed case. The source of this change lies in the use of the original particles

---

[9] This conclusion, which is no more than what is standardly agreed by all the manuals and commentators since the last century, is not threatened by evidence that there were other items which were already preverbs (and perhaps, though this is less clear, adpositions) in the Indo-European parent language (Pinault 1995). Nor are the objections of Pinkster (1972: Ch 9.2) especially persuasive. He is certainly right that throughout attested Latin there is a clear relation of government between preposition and case, but examples such as (3) and (4) above show us the residue of an earlier stage in which no such relation was required. The existence of such a stage is amply confirmed by the sister languages whose attestation is more ancient. The important point for our purposes is that the items which make up the bulk of the prepositions and preverbs in the daughter languages display signs of syntactic independence in texts such as those cited here for Latin and in Homeric Greek (8th cent BCE), Vedic and Hittite (2nd millennium BCE). Even the earliest of these is two or more millennia later than the latest datings for I-E unity.

as specifiers of the local cases (Instrumental, Ablative and Locative).[10] Thus a particle such as *sub* 'under' is not inherently directional or locational; it simply expresses an orientation. If it modifies a directional use of the accusative, the meaning will be '(move) under and to(wards)'; if it modifies a locative the meaning will be '(be) under and in/at'. It is only at a later stage that the particle and the case form are inseparable, and that we have the situation in which the preposition may be said to 'take' or govern a particular case form. When this happens, as we have noted, the change involved is therefore from a Spec-Head pattern to a Head-Comp, which is rather hard to reconcile with the usual constraints on X-bar. Note however that this origin fairly straightforwardly explains what many have taken to be a puzzle (Baldi 1979; Joseph 1991; Bauer 1995: Ch 5), namely the linear order whereby Latin has prepositions rather than postpositions. All the available evidence points to specifiers in Indo-European being on the left. There is no ordering paradox with a putative I-E OV order on this interpretation, a point that rather surprisingly seems to have been missed in the literature (cf further in 3.4 below).

3.3 *Separation of preverb and preposition*

The ability of the preverb to occur separated from the verb as in examples (3) and (4) is lost, due to reanalysis of a separated pre-verb as a preposition (see §3.4). When adjacent to the verb, the preverb quickly comes to form a lexical unit with it and to undergo semantic shifts which often obscure the etymological function of the preverb. Thus, *obligare* is rare its etymological

[10] Note that this origin explains the absence of prepositions governing the dative case, which did not in origin have local values. Where in a modern language we find a prepositional dative — e.g. Germanic, Greek — this is due to a later syncretism of the original dative with one or more of the local cases (Vincent 1997c).

sense of 'to bind around' (< ob+ligare) and the more usual and surviving sense of 'oblige, make liable' does not have a separable component of meaning that could be attributed to the *ob*-prefix. Similarly, *permittere* originally meant 'to let go through, to let loose' (< per 'through' + mittere 'to send, cause to go') but already by the time of Caesar and Cicero (1st cent B.C.) had acquired the meaning 'to allow' which is all that survives in French *permettre*, Italian *permettere*, etc. This kind of lexicalization is a familiar process in such circumstances and is described for the Brazilian language Nadëb by Weir (1986); see also the studies collected in Rousseau (1995). For more detailed studies of the Latin situation, see Lehmann (1983) and Rosén (1993).

3.4 *Word order*

Perhaps the most striking property of the newly emergent prepositional constructions is the rigidity of linear order they exhibit: the order is always P NP and the two items are very rarely separated by intervening material. In the larger Latin context, this has been seen as surprising because (a) Latin word order otherwise shows a fair degree of freedom, and (b) if there is a dominant tendency, it is for the verb to follow the object so we might have expected a similar linear order to hold between an adposition and its object (Hawkins 1983). How are we to explain this state of affairs?

Baldi (1979) suggests the following stages:

I attraction of unstressed particles to the preverbal slot, which would be consistent with the general verb-final typology reconstructible for I-E;

II the univerbation of pre-verb and verb to create compounds of the *praemettere, circumire* type (cf §3.3 above);

III the refinement of the case semantics of the argument of such verbs by the duplication of the preverb in a parallel pre-nominal position: *trans flumen transire/transducere* 'lit. to (cause to) cross across the river';

IV the extension of the pre-noun + noun groups thus formed to other contexts where the prenominal element is not supported by a preverbal element;

V the eventual loss of the transparent preverbal element once lexicalization of preverb+verb sets in: e.g. *inducere* comes to mean 'introduce, persuade' and the sense of 'lead in' is now expressed by *in*+NP *ducere*.

The problem with this proposal is that the key sentence type — *trans flumen transire* — has only a limited attestation, in particular in writers like Caesar, Cicero, Livy, etc (i.e 1st cent BCE and after; cf also Nocentini 1992: 228; Coleman 1991: 332). It is not found in early prose writers such as Cato (234-149 BCE). Even in those writers in which it is found, it is not always used and the verb is often treated as a simple transitive (even to the extent of allowing a passive: e.g.*Taurus transiri non potest* 'The Taurus mountains cannot be crossed' (Cic Att 5,21)). It is better therefore to regard this construction as an artefact of literary prose of the classical period, and thus something that came into existence long after prepositions had been established in the language, and perhaps even in virtue of the establishment of prepositions, rather than as the missing link between preverbal and prepositional syntax.

Joseph (1991: 188) also observes that the apparent anomaly of both the contiguity of preposition and noun and their relative order in the broader Latin context calls for some explanation. He too looks to the preverbal pattern that had established itself with the verb as a model, reinforced by what he claims is a conspiracy for Latin to have morphosyntactic clusters in which the inflected word is final. Thus, given the choice between *per fugam* 'through flight' and **fugam per*, he suggests only the former is consistent with the morphological 'target' of Latin. There is an inherent problem of teleology in this mode of explanation, but in any case there are empirical arguments

against Joseph's position. Not only are there are many instances of elements cliticising in Latin after the inflection — the already cited enclitic conjunction *-que* is a case in point as are the deictic and emphatic particles contained in items like *idem*, *hic*, etc (originally *is-**dem**, hi-**ce**) — but in Oscan we find precisely the development of a new inflectional sub-system out of agglutinated postpositions (Nocentini 1992).

Nocentini (1992: 227-230) in fact suggests that the origins of the matter lie earlier in time. The most common inherited patterns are NP-P-V and P-NP-V. The former yields the preverb construction, i.e. NP [P-V] and the latter the prepositional one, i.e. [P-NP] V. However, while this scenario narrows down the range of options, the only pattern it rules out absolutely (and correctly) is the possibility of a 'postverb', which would require the unattested *NP-V-P. Since NP-P-V does occur, we still need an explanation for why that doesn't yield a bracketing [NP-P] V, thus creating postpositions. Interestingly Craig & Hale (1988) propose just such a sequence as the pre-condition for the development of relational preverbs out of postpostions in the Chibchan language Rama and the Maku language Nadëb (Weir 1986). The fact that the change that must be ruled out in Latin has occurred elsewhere in the world's languages means that we cannot give up on the search for a special property of Latin which allowed P-NP-V to produce prepositions and yet blocked postpositions from developing out of NP-P-V.

The clue to the answer, I would suggest, lies in the fact that the preverbal/prepositional elements were inherently tonic when separated from the verb. At first sight this might appear paradoxical since we typically think of adpositions as weak and (en/pro-)clitic, but recall that in origin these elements were precisely not prepositions but particles which could focus or contrast some property of the sentence they were included in. Pinault (1995: 42) notes the accentability of the preverb in Vedic, especially when it is placed in initial position and separated from the verb (a pattern which corresponds

to the early Latin examples *ob vos sacro* and *sub vos placo* that we have already seen in (3) and (4)). It is the unaccented form which migrates to and cliticizes on the verb.[11] In this sense Baldi (1979) is right that the preverb pattern is older, and indeed on Pinault's account preverb+verb constructions must be postulated for the parent language anyway. What we must assume therefore is that when these particles were next to the verb they tended to cliticise onto it: the verb is the unmarked sentence head and the natural attractor of such sentence level particles in the unmarked — hence unstressed — case. Similarly, Joseph's discussion of the relative accent of P and NP (1991: 188-189) starts out on the right track but then makes the wrong assumption, viz that the P is proclitic on the NP. This may have been the later situation (cf the remark by Quintilian that he adduces), but we claim that the formation of prepositions could only come about if P and V were split by the NP and that in such a circumstance P bears the more prominent stress (cf the different stresses in English *only* and *even* in adnominal and adverbal position) just as in Vedic.

The stressed particle naturally occurs in initial position, so that unstressed elements — very commonly pronouns — will follow the particle by the normal Wackernagel effect. We thus get a very common pattern (bold means stressed and = marks a clitic boundary):

(7) **P**=pronoun … **V:** **sub**=vos **placo**; **ob**=vos **sacro**

---

[11] A clue to the different accentual status of the separated and unseparated preverb in Latin is found in the vowel alternations in *sub placo* vs *supplico*, *ob sacro* vs *obsecro*. In the separated form there is a stress on each part and hence the verb stem has the full vowel *a*. However, the preverb+verb cluster receives a single word stress (on the initial syllable) and hence the stem vowel reduces: cf *cano* 'I sing', *tango* 'I touch' beside *cecini* 'I sang', *tetigi* 'I touched' with stress on the reduplicated perfective prefix.

(cf **endo**=que **plorato**)

It is this pattern which provides the context for the reanalysis in which the (in origin pronominal) element attached to the particle becomes interpreted as an argument of the particle, thereby transforming the particle into a preposition.

A further attraction of this account is that it explains the virtually exceptionless nature of the process. Contrast a putative application to Indo-European of the reanalysis model postulated for Rama and Nadëb by Craig & Hale (1988) in which there is a postpositional phase prior to the reanalysis of the postposition as a preverb, something which happened to different degrees for different members of the class. This would predict a proportion of postpositions, with perhaps a gradual dominance of prepositions over time as the other word order changes take place, but this is not what we find. The striking thing about the P-NP order is, as Joseph (1991: 188) rightly notes and Bauer (1995: 128ff) endorses, that it is early and almost absolute, and runs counter to the prevailing OV typology.

Nonetheless, there are signs of apparent postpositional behaviour in early Latin texts and these need to be explained. It is commonly assumed that such cases — exemplified in (8) — require the postulation of an archaic Latin stage in which postpositions were normal (cf Bauer 1995: 130ff for a recent restatement of this view):

(8) mecum 'me + with'
quocum 'which + with'
ted endo 'you + in'
quibus ex 'which + out of'

In fact they fit in rather well with our story, and do not require the postulation of a separate postpositional phase of the language. To explain *mecum* and the

like we only have to assume that there could be competition for the initial stressed position between the particle and a stressed pronoun. If the latter wins, the particle is forced into second position, and as the relatively unstressed element itself cliticises: **me**=cum **venit**, typically written as a single word. In the other type of circumstance — exemplified by *quocum, quibus ex* — we have a similar kind of competititon for first position though here the motivation is different. The *qu*-words are anaphoric and therefore presumably reduced in stress; they are however required to obey the grammatical rule by which words of this class occupy first position, thus relegating the particle to second position. Although the existence of postposed forms in Sanskrit is widely cited in favour of the postpositional view, the detailed study of Renou (1933: esp 52ff) suggests that the determining principle here too is accentual and that in general postpositions are rare in Sanskrit. Nocentini (1992) similarly adduces a prosodic explanation for the origin of postpositions in Umbrian.[12]

In sum, we have two independent but converging patterns which generate prepositions in a language like Latin. On the one hand, since prepositions originate as specifiers and since specifiers precede the case form they modify, we predict the order 'P + Caseform' rather than the reverse. Second, separated preverbs would normally occur at the left edge of the clause and therefore necessarily to the right of any pronominal items which might later, by reanalysis, come to be interpreted as their arguments. Apparent instances of postposition (except perhaps Hittite — see note 12) would arise as the by-

---

[12] The one branch where postpositions are undeniable is Hittite, but it may not be legitimate to project the Hittite pattern directly back to Indo-European since the evidence for postpositions is stronger in Middle Hittite than in Old Hittite (cf Starke 1977). I am grateful to Anna Morpurgo Davies for bringing the Hittite facts to my attention and for explaining their significance in the broader context of reconstructing the syntax of I-E preverbs.

product of competing prosodic requirements and would not reflect an earlier stage of genuine postpositional syntax.

This principle of competing emphasis will also generalize very neatly to the other kind of reorganization of word order within the PP which we find, what we might call the *magna cum laude* effect. This pattern, now fossilized in many universities as labels for degree classes, typically affects quantificational adjectives such as *multus* 'much, many', *magnus* 'great', *maximus* 'greatest', *summus* 'top', *nimius* 'too great'. Such adjectives seem to be inherently emphatic and are natural candidates for the leftmost position in the constituent.[13] If they occupy this position, they too will trigger this kind of constituent (or perhaps colon) internal Wackernagel effect,[14] and force the preposition to be enclitic.

The data reviewed so far suggests a certain freedom in the word order of Latin prepositional phrases, though this freedom is not absolute. We have seen that a particular class of adjectives commonly precede the preposition. In poetic language, a wider range of adjectives and also genitives dependent on the argument of the preposition may precede (the following examples are taken from Penny (1996):

(9) a) Nestoris ad patriam (Ennius *Satur* 41V)
Nestor.GEN at homeland.ACC
'in Nestor's homeland'

b) Libyae vertuntur ad oras (Vergil *Aen* 3.18)
Libya.GENSG turn.3PSG to shore.ACCPL

---

[13] Similar emphasis must be attributed to the fronted numeral in *duabus de causis* 'for two reasons' (Caes *BG* 6.9).

[14] Though compare Nocentini (1992: 236) for the claim that this pattern is not to be attributed to the Wackernagel effect.

'they turn towards the shores of Libya'

As (9b) shows, such fronting may even move the genitive out of the prepositional phrase altogether, and these constructions seem to be part of poetic style often driven by metrical considerations. Penny (1996) for example notes that the only occurrences of this pattern to be found in the plain narrative prose text of Caesar's *De Bello Gallico* is *quorum per finis* 'through whose territory' where the fronted genitive *quorum* is a relative pronoun, which would have undergone this movement in any case. I have checked for comparison another and earlier prose text — Cato's agricultural handbook *De Agri Cultura* of the 2nd cent BC — and Penny's generalization holds true: instances such as *quo in agro* 'in which field' (VI.4) are found but otherwise both adjective or genitive and noun follow the preposition. Similarly poetic is the construction in which the noun precedes the preposition and its modifier follows (cf also Bauer 1995: 135-6):

(10) a) cava per calamorum (Lucr 5.1382)
hole.ACCPL through reedpipe.GENPL
'through the holes of the reedpipes'

b) litus harenosum ad Libyae (Vergil *Aen* 4.257)
shore.ACCSG sandy.ACCSG to Libya.GENSG
'to the sandy shore of Libya'

c) arbusta per alta (Ennius *Ann* 187)
timber-trees.ACC through tall.ACC
'through tall timber trees'

In this respect I agree entirely with the distinction Bauer (1995) draws between those patterns which are attributable to stylistic factors and those which properly feed into a diachronic account. I differ from her, however, in

that I see the changes in question as the gradual fixing of syntactic configurationality rather than as a shift from a head-final ('left-branching' in her terms) to a head-initial ('right-branching) PP. I see no difficulty in accepting the usual X-bar account of category structure for Modern French, Italian, etc, but I would hesitate to project it backwards onto Latin and Proto-Indo-European. One crucial area, therefore, in which our approaches differ concerns the role of specifier of PP, to which I now turn.

3.5 *Specifier of PP*

The only variant on P-NP order then that needs to be accounted for within the core syntax of (non-literary) Latin is that in which a quantificational or intensifying adjective immediately precedes the preposition. The fact of immediate linear precedence suggests that this 'displaced' adjective is still within the PP, a property which falls out naturally from the prosodic competition principle established above.[15] Moreover, note that there is often no semantic difference in such examples between interpreting the adjective as attached to the noun and as taking scope over the whole PP; in either case *summa cum laude*, for example, means 'at the highest point on the scale of praise'. This in turn suggests that we should consider the relation of this construction to other forms of prepositional modification.

---

[15] It remains true that Latin had fairly extensive patterns of topicalization or scrambling in which an element could 'move' to the front of its clause and some distance from its co-constituents (cf (9b)). It also seems likely that this was characteristic of all styles of the language since it leads in due course to the new patterns attested in Romance (cf Vincent 1998 for some discussion). I will asume however that these patterns are to be handled independently, an assumption supported by the fact that the grammatical distance involved and the classes of categories affected are quite different.

Adverbial modification of a prepositional construction is exemplified in the following:

(11) a) si ad te bene ante lucem venisset (Cic *Or* 2.64.259)
if to you well before light come.PLUPF.SUBJ.3SG
'if he had come to you well before daybreak'

b) longe ante alias specie insignem (Liv I,9)
long before other.FEM.ACC.PL beauty famed.ACCSG
'(a woman) famed for her beauty greatly above others'

c) paulo supra eum locum (Caes *BG* 6.9)
a little above that place
'not far beyond that place'

d) haud multum post mortem eius (Tac *Ann* 5.3)
not much after death her
'shortly after her death'

Such degree expressions were commonly but not necessarily placed before the the preposition, in other words in the same position as the adjective in *summa cum laude*. However, it seems clear that in earlier texts in particular the PP and its degree modifier could be separated. The following examples relate to the item *usque*, glossed in Lewis & Short (1879) as 'all the way to or from any limit of space or time. etc'. This item is of particular interest both because it later evolves from a fixed pre-prepositional position into part of an etymologically complex preposition in French *jusqu'à* (< de usque ad) 'until', and because it is the nearest we can come in Latin to an item that only modifies prepositional

and adverbial constructions and thus has the same diagnostic function that an item like *right* has in English:[16]

(12) a) ab imis unguibus
from bottom.ABLPL nail.ABLPL
usque ad verticem summum
right to top.ACCSG highest.ACCSG (Cic *Rosc Com* 7,20)
'from the top of his head to the tips of his toes'

b) usque a rubro mari (Nep. *Hann* 2.1)
right from red.ABLSG sea.ABLSG
'right from the red sea'

c) trans Alpes usque transfertur (Cic. *Quint* 3,12)
across Alps.ACC right carry.PRES.PASS.3SG
'he is carried right across the Alps'

d) ab Attica ad Thessaliam usque (Plin 4,12,21)
from Attica.ABL to Thessaly.ACC right
'from Attica right to Thessaly'

The examples in (12) show clearly that *usque* can both precede (12a,b) and follow (12c,d) the PP it modifies. It is only in later usage, presaging that which is found in Romance, that the specifier of a prepositional phrase has a fixed linear position. In sum, the fixing of the position of the specifier, whether adjectival as in the case of the *summus* class of items or adverbial as with *usque, longe,* etc, postdates by several centuries the fixing of the position of the

[16] The only exceptions here are the use of *usque* with certain locational nouns where it can modify the bare directional accusative: *terminos usque Libyae* 'to the borders of Libya' (Just 1,1,5). This is of course a parallel with the type of case modificational structure out of which prepositions emerged in the first place (cf §3.2).

preposition. As examples of this later usage with an item, *totum* 'whole', which did not have such a function in the classical language, consider the following from the fourth century CE text *Peregrinatio Aegeriae*:[17]

(13) a) totum per valle illa (I,2)
whole through valley that
'right through the valley'

b) totum per mediam vallem ipsam (II,3)
whole through middle valley the
'right through the middle of the valley'

3.6 *Morphosyntactic quivalence of PPs and case*

It is a commonplace of Romance historical syntax that the grammatical functions of the dative case — the marker of the indirect object — and the genitive case — the marker of a noun dependent on another noun — came to be replaced by grammaticalized uses of the prepositions *ad* 'to, towards' and *de* 'down, about' respectively. What is less commonly emphasised but is of particular importance in the context of the present discussion is that this development took place over many hundreds of years. Thus, Molinelli (1996: 83) notes the following pair of virtually equivalent examples from Plautus (254-184BCE):

(14) a) numquid aliud vis
INTERR other.NEUTSG.ACC want.2SG.PRES
patri nuntiari? (*Capt* 400)
father.DATSG tell.PASS.INF
'Is there anything else you wish to be told to your father?'

[17] I am grateful to David Langslow for pointing out to me the significance of these examples.

b) quae ad patrem vis
which.NEUTPL.ACC to father.ACCSG want.2SG.PRES
nuntiari
tell.PASS.INF (*Capt* 360)
'You wish these things to be told to your father'

The choice here may well have been dictated by metrical considerations since the dative case *patri* and the prepositional phrase *ad patrem* have different numbers of syllables, but even so it is hard to believe that the option would have been available in the first place if the expressions had not been semantically equivalent or almost so. This in turn requires that the grammar at this stage of the language be able to capture the fact that both a bare case form and a prepositional phrase can satisfy the indirect object requirement of the verb *nuntiare* 'to tell, report'.

There are similarly early examples of *de* + NP being used where classical norms would have dictated a genitive (again from Molinelli 1996, who provides ample documentation and vaulable discussion):

(15) a) memento ergo dimidium istinc mihi
remember.IMP therefore half.ACC thence me.DAT
de praeda dare
of booty.ABLSG give.INF (Plautus *Ps* 1164)
'Remember then to give me half of the booty from there'

b) faecem de vino bono (Cato *Agr* 96.1)
dregs.ACCSG of wine.ABLSG good.ABLSG
'dregs of good wine'

At the same time there are genitives and datives being used in similar contexts by writers over half a millennium later. Compare the example in (16) where there is both a genitive case (*loci ipsius*) and then a prepositional phrase *de pomis* 'of the fruit'. Note too the dative case *nobis* 'us' expressing the indirect object. The text is the fourth century AD *Peregrinatio Aetheriae*, which is notorious for its conversational and colloquial style.

(16) dederunt nobis presbyteri loci ipsius
give.PERF.3PL us.DAT priest.NOMPL place.GEN that.GEN
eulogias, id est de pomis …
praise.ACCPL that is of fruit.ABLPL (*Per Aeth* III,6)
'the priests of that place sang the praises of the fruit …'

At the same time there are many other contexts in the language in which *de* and *ad* will have independently definable semantic values. An adequate grammar will thus have to capture on the one hand the grammatical and the semantic uses of the same prepositions and on the other the morphosyntactic equivalence of a bare case form and a prepositional phrase. We will develop an LFG account of this in §5 below. Before that, however, let us briefly review a couple of other developments in the syntax of prepositions in modern Romance.

3.7 *Later Romance developments*

3.7.1 As is well known the eventual result of the breakdown and replacement of the Latin case system, part of which has been described in the last section, is the complete loss of morphological case as a grammatical mechanism. As the historical record shows, and as logic would dictate (cf Harris 1978), the rise of prepositions precedes the fall of case. Just as the fixed order of a noun adjacent to its governing preposition permits the case marking to be lost

without loss of content, so the fixed order of subject, verb and object allows for loss of grammatical case marking. The detail of these changes is complex and would take us too far afield. It is sufficient to note here that between the original, grammatically autonomous semantic cases and the complete loss of case morphology is a longlasting stage in which grammatical case survives beside an originally semantic use of prepositions, while the latter in turn become grammaticalized for the expression of non-core arguments as case is eliminated from the system.

3.7.2. Just as some prepositions pass from semantic to syntactic functions so a wide range of new prepositions are coined from many different sources (cf Vincent 1997a, b and Longobardi 1995). Thus:

(17) Ital *senza* 'without' < Lat *absentia(m)* 'absence'
Fren *chez* 'with, at the house of, in' < Lat *casa(m)* 'house'
Span *bajo* 'underneath' < Lat *bassu(m)* 'low'

3.7.3. One particularly interesting development in this latter connection is the emergence of prepositions whose etymology requires us to assume a combination of prepositions. Thus, for instance, Italian *dopo* 'after' can be shown to be the fusion of *po* < Lat *post* 'after'prefixed by *de*, the genitive marking preposition whose origin we reviewed in §3.6. Similarly French *avant* 'before' is from Latin *ab*'from' + *ante* 'before'. The most likely origin of such preposition + preposition sequences is from a stage in which a preposition could take another PP as complement (Vincent 1997a: 212-30). Thus the etymology of Italian *dopo cena* 'after dinner' would be [*de* [*post cenam*]], literally 'from after dinner'. This is important evidence for another piece in our puzzle. Classical Latin did not allow PPs to be the complements of other Ps, a pattern with which we are of course very familiar in English (Jackendoff 1973), and we can take their historical emergence in this function to be yet

another sign of the increasing internal structural complexity of the category PP.

3.8 *The changes summed up*

At this point it will be useful to summarise the sequence of changes that we have argued for in the preceding sections. There was undoubtedly a considerable degree of overlap in the instantiation of these changes (fuelled by the kinds of sociolinguistic factors discussed in Molinelli 1996, 1998), but by separating them out into discrete stages here we can better highlight the logical connections:

Stage I: Adverbial particles exist which do not have grammatically determined arguments and can occur in a variety of clausal positions.

Stage II: One particular usage of these particles causes them to precede and modify semantically independent local case forms of nouns. Another (independent) usage has them in stressed clause initial position where they may host second position clitic pronouns.

Stage III: Both the patterns identified in Stage II admit of a reanalysis in which the nominal or pronominal element may be interpreted as an argument of the adverbial particle, thereby converting the particle into a preposition of the which the case form is now a dependent, and creating the category complem.

Stage IV: The prepositional phrases thus created come in due course to be themselves subject to modification by items such as *usque, longe, totum*.

Stage V: Such secondary modifiers come to occupy a fixed pre-prepositional slot and thus to create a category Spec(ifier) of PP.

Stage VI: As prepositional syntax extends its range and case forms recede, both changes which are implemented over a number of centuries, two further changes occur:

i) PPs come to be complements of P;

ii) many new prepositions are created through the grammaticalization of a wide range of other categories.

Stage VII: Structures of the form [P [P NP]] which emerge in Stage VI (i) are reanalysed as [P-P [NP]] and eventually the compound prepositions so formed re-lexicalise as single Ps.

Stage VIII: The culmination of all these stages, which are crucially not only logically but also chronologically distinct, consists in the rigid, hierarchical, X-bar style, case-free PPs of the modern Romance languages.

What this sequence of events indicates is that, although the modern situation may be well described by X-bar theory, the stages which lead up to it are not. In the remainder of this paper we will try to draw out the theoretical consequences, in both the diachronic and the synchronic arenas, of such a state of affairs.

4. GRAMMATICALIZATION AND GRAMMATICAL THEORY

The case history (in both senses!) we have just sketched is exactly the kind of evidence that has been at the heart of the recent revival of interest in grammaticalization (cf Cuzzolin 1995). Items that in origin had independent lexical content gradually extend their semantic range and at the same time acquire a fixed subcategorization (Vincent 1980). This is perhaps most dramatically evident with the pan-Romance preposition *de*/*di* , which once had a sense of 'down, away from' (cf Latin *de-scendo* 'I come down', *de-sum* 'I am absent'), but now marks the purely grammatical relation of one noun on another (*la voiture de Paul* 'Paul's car', *la chûte du gouvernement* 'the fall of the government') or of one verb on another (*il a décidé de s'en aller* 'he decided to leave').

What the grammaticalization approach does not have — and indeed what some if not all proponents of grammaticalization would argue against in principle — is a theory whereby one can represent and compute grammatical structure. Thus, although they make fundamental use of the concept of 'reanalysis', as indeed all theories of syntactic change cannot fail to do (Harris & Campbell 1995: 000), they have no more than the traditional armoury of informal and intuitive grammatical categories with which to define the input to and output from diachronic reanalyses. The weakness of this position is perhaps not so obvious when the matter under discussion is semantic bleaching and the concomitant extension of the class of NPs associated with a given predicate. However, when we have a discrete shift from absence to presence of an argument, which is the key factor in the birth of a class of prepositions, then the inability to fall back on a theory of argument structure constitutes a serious difficulty.[18] In this respect, the acquisition of argument structure which is, as we have seen, inherent in the process of preposition formation is not simply the opposite of the loss of argument structure which is commonly, and reasonably, held to characterize auxiliarization (Roberts 1985, 1993). More precisely, Roberts (1985) argues that the properties of Modern English modals derive from the fact that they no longer assign theta-roles when compared to the etymological sources. Loss of theta structure is a clear formal analogue to the usual accounts of grammaticalization as involving 'semantic bleaching'. By the same logic one might expect acquisition of theta structure to be a kind of 'degrammaticalization' (on which

---

[18] Note that this is not intended as a plea for a particular theory of argument structure; there are many available in the current literature and the debates about which to choose would take us too far afield. The point is simply that the kind of change we are discussing here is hard to conceive of in the absence of some kind of explicit theoretical characterization of grammatical structure.

see Ramat 1992), yet the intuitive parallels break down. Rather, the fact of adverbs becoming prepositions, as already noted, is usually thought of as a kind of grammaticalization (Cuzzolin 1995) for the obvious reason that such items start to contract more than semantic relations of modification with the clauses to which they belong. What this in turn suggests is that it is wrong to think of the passage from lexical semantic autonomy to grammatical function as a cline or a diachronic continuum (cf Lightfoot (in press), Newmeyer (in press)); it is rather a more abrupt kind of leap from lexicon to grammar, understood as two independent but interacting components of language. It is the involvement in the mechanisms of grammar that characterises grammaticalization, and any change of argument structure whether involving loss or gain can have this effect. On this account, many of the valid intuitions and insights of the recent explosion of work in so-called grammaticalization 'theory' (cf Newmeyer (in press) Ch 5 on the misappropriation of the term 'theory' here) can be salvaged but only, and we would argue rightly, at the price of commitment to an explicit theory of grammatical structure and grammatical representation. Since we have seen in §3 that such a theory cannot be a configurational, position-based one, it behoves us to give at least a sketch of what an alternative account would look like and how it would handle the data discussed above. It is to this that we now turn our attention.

5. A LEXICAL-FUNCTIONAL MODEL OF CHANGE

Canonical X-bar theory can handle most if not all the patterns we have just observed, but only at the expense of postulating the existence of levels of structure prior to the emergence of the forms which would provide evidence for the structure. Thus, since the presence of a category P(reposition) in X-bar automatically entails both a complement and a specifier of P, the slot Spec of P will have to be assumed to exist as soon as the case forms come to be governed by the former independent adverbial particle. Indeed, on the

widespread view that such adverbial particles are simply intransitive prepositions (Jackendoff 1973, Emonds 1976: 172ff), the strongest form of X-bar theory might even be interpreted to include (unrealized) Spec and Comp of P at the Proto-Indo-European stage before case-governing prepositions in the traditional sense have emerged. Now there is indeed much to be said for the view that assimilates adverbial particles of this kind to the class of prepositions (cf already Jespersen 1924: 87ff for a similar idea), as long as we are not forced to associate with them unrealized and at some stages unrealizable arguments and modifiers. In other words, we need a theory in which categories may exist but in which projections of those categories are associated with them only at a historical stage at which there is overt evidence for such further layers of structure. On this view, and somewhat schematically, we could say that PIE has P, (Early) Latin has P and complement of P, and Late Latin/Romance has P, complement of P and specifier of P. Projection is then historically delayed until triggered by the changing distributional frequencies of certain items. The accretion of each new layer of structure can then be thought as corresponding to, and making precise, the intuitive idea of increased grammaticalization.

Compare now the usual X-theory treatment of morphological case (cf Bittner & Hale 1996). On this view, since the Latin ablative, say, corresponds to an English *with*-phrase, and since the latter is headed by the preposition *with*, then for Latin and other such languages we postulate a category KP (for Case Phrase) is postulated headed by a Case morpheme K, which will then contract the same structural relation to its associated nominal as the P does to its nominal. (For an instance of this form of argumentation applied in the historical domain, cf Weerman 1997.) Putting this latter idea toegether with the standard understanding of projection means that the postulated functional head K in turn automatically projects Spec of KP and complement of K. Once again there is a problem of unwanted structure. The essential

objection to this way of proceeding is that it creates a 'syntax' for the morphological form which is never realised in the phonology (cf Börjars, Vincent & Chapman 1997; Vincent & Börjars 1996). Such 'syntacticization' of morphology rests on the fallacious representational assumption that the only way to state morphosyntactic generalizations is in terms of X-bar trees. What is required rather is a means of representing morphological information on its own terms but ensuring that it can communicate within the grammar with syntactic information, also represented on its own terms. The way to achieve this is to extract the functional equivalences that the two patterns — inflectional case and prepositions — have, and to state this in a single common representation, an f(unctional)-structure. This f-structure can then be mapped onto the independently available and autonomously structured realizations in morphology and syntax. We will flesh out such an account for two types of situation, one involving a semantic use of case/prepositions and the other a grammatical use (Vincent & Börjars 1997).

Let us begin with the representation of the content of a semantic case such as an original Indo-European locative must have been (Kuryłowicz 1964: Ch 8). This can be expressed, following Simpson (1991) and Andrews (1995), by interpreting the semantic case as having its own argument. Intuitively, a locative case marker expresses a predicate of location, the argument of which is the (referent of the) noun to which the case marker is morphologically attached. Formally, this can be stated as in (18) for the form *urbe* 'in/at the city':

(18) $\begin{bmatrix} \text{PRED} & \text{'LOC (OBJ)'} \\ \text{OBJ} & \textit{urb} - \end{bmatrix}$

In the stage in which *in/endo* 'in, inside' has not yet become a preposition but is rather a specifier of this locative case, the representation of *endo urbe* 'inside the city' will be as in (19). (NB In what follows I adopt the fiction of

representing the item *in*/*endo* in its specifier function by its earlier shape *endo*, and in its prepositional function by its later shape *in*. I hope this will clarify the exposition, though it should be understood that the phonetic evolution from *endo* > *in* was independent of and parallel with the grammatical shift being analysed here.)

(19) $\begin{bmatrix} \text{SPEC} & \textit{endo} \\ \text{PRED} & \text{'LOC} < (\text{OBJ}) > \text{'} \\ \text{OBJ} & \textit{urb} - \end{bmatrix}$

After the grammaticalization of the preposition, we will have:

(20) $\begin{bmatrix} \text{PRED} & \text{'}\textit{in} < (\text{OBJ}) > \text{'} \\ \text{OBJ} & \textit{urb} - \end{bmatrix}$

A version of this last with its own specifier would be *longe in urbe* 'far inside the city' to which we can assign the following representation:

(21) $\begin{bmatrix} \text{SPEC} & \textit{longe} \\ \text{PRED} & \text{'}\textit{in} < (\text{OBJ}) > \text{'} \\ \text{OBJ} & \textit{urb} - \end{bmatrix}$

What these representations do not capture is the morphological forms of the nouns in question. We need to ensure that in (18) and (19) the LOC predicate is realized as the locative case, while in (20) and (21) the object of *in* is in the ablative case. In order to achieve this we first consider how to represent an item which has a grammatical case such as nominative or accusative (cf again Vincent & Börjars 1997 for discussion of the two types of case). In such circumstances the case will be a simple feature of the argument in question, but crucially will not have its own argument structure. Consider the sentence in (22a) and its associated representation in (22b):

(22) a) Exercitus urbem defendit
army.NOMSG city.ACCSG defend.3SG.PRES
'The army defends the city'

b)
$$\begin{bmatrix} \text{PRED} & \text{'}\textit{defendere} < (\text{SUBJ}), (\text{OBJ}) > \text{'} \\ \text{TENSE} & \text{PRES} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \textit{exercitu} - \\ \text{CASE} & \text{NOM} \\ \text{NUM} & \text{SG} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \textit{urb} - \\ \text{CASE} & \text{ACC} \\ \text{NUM} & \text{SG} \end{bmatrix} \end{bmatrix}$$

The association between the functions subject and object and the case markings nominative and accusative will be achieved through correspondence statements of the form:

(23) SUBJ ⇒ nom
OBJ ⇒ acc

The complex patterns of morphological realization of these case/number forms according to the declensional class of the noun in question will then be the responsibility of a separate morphological component, whose internal structure we will not further articulate here (cf Blake (1994: 20-23) for the arguments against mapping function straight into morphological form in a language like Latin). Note in particular that in a system of this kind there is no need to allocate the subject and object to configurationally determined slots in order for them to receive their case-marking. The kind of direct function-case mapping exploited in (23) fits very well with the fact that in fact the word order in Latin is relatively free, and that therefore any linear rearrangement of

the sentence in (22a) will retain the same f-structure values as long as the case forms are not changed. Conversely, if the case forms are changed, for instance by making *exercitus* 'army' accusative and *urbs* 'city' nominative, then the meaning and the f-structure of the sentence change even if word order is held constant.

Let us return now to the question of assigning case forms in the context of the grammatical change from specifier to preposition itemised in (18) through (21) above. At the original stage represented in (18) the locative value of the locative case can be expressed as follows:

(24) LOC ⇒ loc

The prepositional case instead requires the following statement, which must be construed as part of the lexical entry for *in*:[19]

(25) OBJ_{in} ⇒ abl

In other words the shift from specifier to preposition involves *in*/*endo* acquiring both an OBJ argument in its PRED feature and a case-assigning feature which determines the morphological shape of that argument. Inspection of the above f-structures reveals that all the necessary equivalences and differences are represented. When the original locative case is serving a semantic function then it is functionally equivalent to a preposition, and this is expressed in the similarity of their respective PRED features, viz:

---

[19] I ignore here the additional complexities of the alternation between ablative and accusative after certain prepositions in Latin. For discussion see Vincent (1997c) and Vincent & Börjars (1997).

(26) a) loc ‘LOC <(OBJ)>’ (as in (18) and (19))

b) *in* ‘*in* <(OBJ)>’ (as in (20) and (21))

When the erstwhile specifier aquires prepositional force, then it takes over the PRED feature and the case is automatically downgraded to grammatical status and is assigned in virtue of the f-structure value of the item to which it attaches, just as happens for clausal subjects and objects:

(27) SUBJ ⇒ nom

OBJ ⇒ acc

OBJ*in* ⇒ abl

Finally in this section we see how an extension of the apparatus just established will provide us with a natural account of the equivalences between case forms and PPs reviewed in §3.6 above. I will simplify the discussion somewhat by extracting the core propositional structure out of the Plautine examples cited in (14), and repeating it as (28):

(28) a) ea patri nuntias

that.NEUT.ACCPL father.DATSG tell.PRES.2SG

‘you tell those things to your father’

b) ea ad patrem nuntias

that.NEUT.ACCPL to father.DATSG tell.PRES.2SG

‘you tell those things to your father’

We can assume a function-case equivalence for dative inherited from Indo-European as follows:

(29) IND OBJ ⇒ dat

This will allow us to state the f-structure of (28a) as:

(30)
$$\begin{bmatrix} \text{PRED} & \text{'}\textit{nuntiare}\ <(\text{SUBJ}),\ (\text{OBJ}),\ (\text{IND OBJ})>\text{'} \\ \text{TENSE} & \text{PRES} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'addressee'} \\ \text{CASE} & \text{NOM} \\ \text{NUM} & \text{SG} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \textit{is} \\ \text{CASE} & \text{ACC} \\ \text{NUM} & \text{PL} \end{bmatrix} \\ \text{IND OBJ} & \begin{bmatrix} \text{PRED} & \textit{patr} - \\ \text{CASE} & \text{DAT} \\ \text{NUM} & \text{SG} \end{bmatrix} \end{bmatrix}$$

To complete the story we now need it to be possible for the f-structure of *ad patrem* in (28b) to substitute for and be equivalent to the f-structure of *patri* in (28a). In other words, we need to treat the PP here not as headed by P with an independent PRED feature as in our treatment of *in* above, but as constituting a periphrastic case form. To achieve this, I will exploit the type of analysis developed in Börjars, Vincent & Chapman (1997). The function-case equivalence statement required is:

(31) IND OBJ ⇒ *ad* + NP

Now this will cause a failure of unification if we try to plug in the preposition *ad* with its usual semantic value 'direction', for which we assume a lexical entry as in (32) along the same lines as that given for *in* above, since the PRED feature of lexical *ad* will be superfluous to the requirements of the realization statement in (31).

(32) PRED '*ad* <(OBJ)>'
OBJ*ad* ⇒ acc

The only way to avoid the incompatibility is for *ad* in this function to lose its PRED feature just as the locative case lost its PRED value when it came to be governed by a preposition. Note that in losing its PRED feature it does not give up its syntactic subcategorization, which is still accusative, and hence the correct case is asigned to *patrem*. An exactly parallel account can be developed for the grammaticalization of *de* 'down, about' into the marker of nominal dependent, first in competition with and finally as a replacement for genitive case. The consistent generalization across both the diachronic scenarios discussed in this paper and those in Börjars, Vincent & Chapman (1997) is that grammaticalization can be neatly modelled in LFG terms as loss of the PRED feature, with the remainder of the changes following from the independently required architecture of the model.

6. TOWARDS AN OPTIMALITY THEORY ACCOUNT OF SYNTACTIC CHANGE

In the previous section I have sought to show that a parallel correspondence model such as LFG provides a more natural account of the data outlined in §3. In recent work, Bresnan (1997,1998a,b) has argued further that the LFG framework is best construed according in Optimality Theory terms. On this view, the relation between f-structure and c-structure is determined by the interaction of re-rankable constraints rather than in absolute terms. Although we will not work out a full account here, two observations indicate the attractiveness of the OT approach also in the present instance.

The first concerns the constraint proposed in Bresnan (1997: 20) under the name of the Principle of Economy of Expression:

(33) DON'T PROJECT: All non-preterminal c-structure nodes are optional and not used unless required by independent principles

As she notes this constraint 'privileges lexical over phrasal expression wherever possible'. We should note too that for these purposes 'lexical' includes fully inflected items according to the principle of lexical integrity. The consequence of this principle applied to a case-form of a language like Latin or Sanskrit (and hence, for the purposes of the present argument, to the relevant reconstructible stage of the parent language) is that unless other factors intervene, such as the requirements of other items in the f-structure (or members of the 'numeration' in Minimalist terms), no further structure is required. In particular, there are no empty Ds or Ps required. Nor, even if a lexical category node is required, will it be necessary to project higher level structure from that node. In the context of the data studied here, this means that at the earliest Indo-European stage the morphologically complete lexical entries will not need to be li censed by empty prepositions or the like. They will simply express the categories of f-structure directly. As each successive historical reanalysis takes place, it will introduce new levels of structure via a re-ranking of DON'T PROJECT vis-à-vis the endocentricity constraints which ensure overt expression of grammatical content.

The second observation which leads naturally in the OT direction concerns the notion of morphosyntactic competition (Bresnan 1998a, b). In these papers Bresnan develops the idea that there is a hierarchical ranking of different forms of expression for a given grammatical content. Just as she proposes a scale for the expression of pronominal content ranging from zero to full syntactically independent pronouns, we can think of a similar scale for the expression of case and prepositional type meanings. On the not unreasonable assumption that independent prepositions are higher on such a hierarchy than bound case morphology just as pronouns are higher than

agreement morphology, we should expect the development of the former to take precedence over the historically inherited latter. It is difficult to be more precise at this point in the research since a full case–preposition scale remains to be worked out. However, for a more detailed investigation of the logic of this type of argumentation in a diachronic projection of Bresnan's pronoun hypothesis, see Vincent (in prep).

REFERENCES

Anderson, Stephen (1993) Wackernagel's revenge: clitics, morphology and the syntax of second position. *Language* 69: 68-98.

Anderson, Stephen (1996) How to put your clitics in their place, or why the best account of second-position phenomena may be something like the optimal one. *The Linguistic Review* 13: 165-191.

Anderson, John (1997) *A notional theory of syntactic categories*. Cambridge: Cambridge University Press.

Andrews, Avery (1995) Semantic case-stacking and inside-out unification. *Australian Journal of Linguistics* 15: 127-181.

Aronoff (1994) *Morphology by itself*. Cambridge, Mass: MIT Press.

Baldi, Philip (1979) Typology and the Indo-European prepositions. *Indogermanische Forschungen* 84: 49-61.

Barron, Julia (1997) LFG and the history of raising verbs. Paper presented at LFG97, University of California at San Diego, June 19-21 1997.

Battye, Adrian & Ian Roberts (eds) (1995) *Clause structure and language change*. Oxford: OUP.

Bittner, Maria & Ken Hale (1996) The structural determination of case and agreement. *Linguistic Inquiry* 27: 1-68.

Blake, Barry (1994) *Case*. Cambridge: Cambridge University Press.

Bauer, Brigitte (1995) *The emergence and development of SVO patterning in Latin and French*. Oxford & New York: Oxford University Press.

Börjars, Kersti, Nigel Vincent & Carol Chapman (1997) Paradigms, periphrases and pronominal inflection: a feature based account. *Yearbook of Morphology 1996*: 155-180.

Bresnan, Joan (1997) Optimal syntax. To appear in J. Dekkers *et al* (eds) *The pointing finger. Conceptual studies in optimality theory*. Oxford: Oxford University Press.

Bresnan, Joan (1998a) Explaining morphosyntactic competition. To appear in Mark Baltin & Chris Collins (eds) *Handbook of contemporary syntactic theory*. Oxford: Blackwell.

Bresnan, Joan (1998b) The emergence of the unmarked pronoun II. MS, Stanford University.

Chomsky, Noam (1957) *Syntactic structures*. The Hague: Mouton.

Chomsky, Noam (1995) *The Minimalist Program* Cambridge, Mass: MIT Press.

Coleman, Robert (1991) Latin prepositional syntax in Indo-European perspective. In R. Coleman (ed.) *New studies in Latin linguistics* Amsterdam: Benjamins, pp. 323-338.

Craig, Colette & Ken Hale (1988) Relational preverbs in some languages of the Americas: typological and historical perspectives. *Language* 64: 312-344.

Cuzzolin, Pierluigi (1995) A proposito di *sub vos placo* e della grammaticalizzazione delle adposizioni. *Archivio Glottologico Italiano* 80: 122-143.

Emonds, Joseph (1976) *A transformational approach to English syntax*. new York & London: Academic Press.

Harris, Alice & Lyle Campbell (1995) *Historical syntax in cross-linguistic perspective*. Cambridge: CUP.

Harris, Martin (1978) The inter-relationship between phonological and grammatical change. In J. Fisiak (ed.) *Recent developments in historical phonology* The Hague: Mouton, pp. 159-172.

Hawkins, John (1983) *Word order universals*. New York: Academic Press.

Hopper, Paul & Elizabeth Traugott (1993) *Grammaticalization*. Cambridge: CUP.

Horrocks, Geoff (1997) Greek voice, Indo-European echoes. Paper delivered to the Philological Society, Oxford, 7 June 1997.

Jackendoff, Ray (1973) The base rules for prepositional phrases. In S.R. Anderson & P. Kiparsky (eds) *A festschrift for Morris Halle*. New York: Holt, Rinehart & Winston, pp. 345-356.

Jackendoff, Ray (1977) *X-bar syntax*. Cambridge, Mass: MIT Press.

Jespersen, Otto (1924) *The philosophy of grammar*. London: Allen & Unwin. (Reprinted University of Chicago Press, 1992)

Joseph, John (1991) Latin prepositions and Romance syntax. In D. Wanner & D.A. Kibbee (eds) *New analyses in Romance linguistics* Amsterdam: Benjamins, pp. 187-199.

Kaplan, Ronald M. & Joan Bresnan (1982) Lexical-Functional Grammar: a formal system for grammatical representation. In J. Bresnan (ed) *The mental representation of grammatical relations*, Cambridge, Mass: MIT Press, pp. 173-281. Reprinted in M. Dalrymple *et al* (eds) *Formal issues in Lexical-Functional Grammar*, Stanford: CSLI Publications, 1995, pp. 29-130.

Kayne, Richard (1994) *The antisymmetry of syntax* Cambridge, Mass: MIT Press.

van Kemenade, Ans & Nigel Vincent (eds) (1997) *Parameters of morphosyntactic change*. Cambridge: CUP.

Kortmann, Bernd & Ekkehard König (1992) Categorial reanalysis: the case of deverbal prepositions. *Linguistics* 30: 671-697.

Kuryłowicz, Jerzy (1964) *The inflectional categories of Indo-European*. Heidelberg: Winter.

Lehmann, Christian (1983) Latin preverbs and cases. In H. Pinkster (eds) *Latin linguistics and linguistic theory*. Amsterdam: Benjamins, pp. 145-165.

Lehmann, Christian (1998) German abstract prepositional phrases as reduced subordinate clauses. Paper delivered at the Workshop on Clause Combining, Copenhagen Business School, 16-17 March 1998. To appear in the Proceedings.

Lewis, Charlton & Charles Short (1879) *A Latin dictionary*. Oxford: Oxford University Press.

Lightfoot, David (in press) *The dynamics of language* Oxford: Blackwell.

Longobardi, Giuseppe (1995) A case of construct state in Romance. In R. Ajello & S. Sani (eds) *Scritti linguistici e filologici in onore di Tristano Bolelli*. Pisa: Pacini, pp. 293-329.

Luraghi, Silvia (1997) Omission of the direct object in Latin. *Indogermanische Forschungen* 102: 239-257.

Matthews, P.H. (1981) *Syntax*. Cambridge: CUP.

Molinelli, Piera (1996) Casi e preposizioni in latino: lo sviluppo del genitivo e del dativo. Una prospettiva (anche) sociolinguistica. *Linguistica e Filologia* 3:73-125.

Molinelli, Piera (1998) Dai casi alle preposizioni in latino: analisi sociolinguistica e spiegazione tipologica. In Ramat & Roma (eds), pp. 147-166.

Newmeyer, Frederick J. (in press) Cambridge, Mass: MIT Press.

Nocentini, Alberto (1992) Preposizioni e posposizioni in oscoumbro. *Archivio Glottologico Italiano* 77: 196-242.

Penney, John (1996) Archaism and innovation in Latin poetic syntax. To appear in J. N. Adams & R.J. Mayer (eds) *Aspects of the language of Latin poetry*. Oxford: Oxford University Press.

Pinault, Georges-Jean (1995) Le problème du préverbe en indo-européen. In A. Rousseau (ed.), pp. 35-59.

Pinkster, Harm (1972) *On Latin adverbs* Amsterdam: North Holland Publishing Company.

Ramat, Paolo (1992) On degrammaticalization. *Linguistics* 30: 549-560.

Ramat, Paolo & Elisa Roma (eds) *Sintassi storica.* Rome: Bulzoni.

Renou, Louis (1933) La séparation du préverbe et du verbe en védique. *BSL* 34: 49-96.

Roberts, Ian (1985) Agreement patterns and the development of English modal auxiliaries. *Natural Language and Linguistic Theory* 3: 21-58.

Roberts, Ian (1993) A formal account of grammaticalization in the history of Romance futures. *Folia Linguistica Historica* 13: 219-258.

Roberts, Ian (1997a) Directionality and word order change in the history of English. In van Kemenade & Vincent (eds), pp. 397-426.

Roberts, Ian (1997b) A formal account of grammaticalization. Paper presented at Workshop on Functional Categories and Morphosyntactic Change, ICHL13, Düsseldorf, 12 August 1997.

Rosén, Haim (1992) Die Komposita mit *co(n)*- in funktioneller und vergleichender Sicht. In O. Panagl & T. Krisch (eds) *Latein und Indogermanisch*. Innsbruck, pp. 357-367.

Rousseau, André (ed.) (1995) *Les préverbes dans les langues d'Europe: introduction à l'étude de la préverbation*. Lille: Presses Universitaires du Septentrion.

Simpson, Jane (1991) *Warlpiri morphosyntax* Dordrecht: Kluwer.

Starke, Frank (1977) *Die Funktionen der dimensionalen Kasus und Adverbien im Althethitischen.*. Wiesbaden:

Vincent, Nigel (1980) Iconic and symbolic aspects of syntax: prospects for reconstruction. In P. Ramat (ed.) *Linguistic reconstruction and Indo-European syntax*. Amsterdam: John Benjamins, pp. 47-68.

Vincent, Nigel (1993) Head vs dependent marking: the case of the clause. In G. Corbett *et al* (eds) *Heads in grammatical theory*. Cambridge: CUP, pp. 140-163.

Vincent, Nigel (1997a) Prepositions. In M. Maiden & M. Parry *The dialects of Italy*. London: Routledge, pp. 208-213.

Vincent, Nigel (1997b) Lexical and historical grammar: the case of *chez*. Paper delivered at Conference on the Boundaries of Morphology and Syntax, Rome, 23-25 October 1997. To appear in Proceedings.

Vincent, Nigel (1997c) On the origin of prepositions. Paper delivered to the Philological Society, Manchester, 21 November 1997.

Vincent, Nigel (1998) Tra grammatica e grammaticalizzazione: articoli e clitici nelle lingue (italo-)romanze. In Ramat & Roma (eds), pp. 411-440.

Vincent, Nigel (in prep) Competition and correspondence in syntactic change: null arguments in Latin and Romance. Paper to be delivered to DIGS5, York, 30 May-1 June 1998.

Vincent, Nigel & Kersti Börjars (1996) Suppletion and syntactic theory. M. Butt & T. King (eds) *Papers from the 1st LFG Colloquium*, Grenoble, August 1996, pp. 448-462.

Vincent, Nigel & Kersti Börjars (1997) Stating your case. Paper delivered at the First Mediterranean Morphology Meeting, Mytilini, September 1997.

Weerman, Fred (1997) On the relation between morphological and syntactic case. In van Kemenade & Vincent (eds), pp. 420-459.

Weir, Ellen M.H. (1986) Footprints of yesterday's syntax: diachronic development of certain verb prefixes in an OSV language (Nadëb). *Lingua* 68: 291-316.