Explaining the bipositional head-complement order in adpositional systems

Martin Ehala
University of Cambridge

1. Introduction

In respect of the head-complement ordering, two orders are possible in adpositional phrases - either the head precedes its complement or follows it. On this basis languages are usually divided into prepositional and postpositional languages. However, as some (the Baltic-Finnic languages, for example) make use of both prepositions and postpositions, a total of three possible word order patterns can be distinguished in adpositional systems. The third one I will call the bipositional order.

It must be noted, however, that in bipositional languages, the two orders are never exploited to an equal extent - one position is always more frequent and one type of adpositions (either pre- or postpositions) always has significantly more items than the other. For example, Estonian has over 150 postpositions but only 30 prepositions. German has over 200 prepositions, but 18 postpositions (Fries 1991). Russian has 15 prepositions which have been attested also as postpositions, but the frequency of postpositional usages is very low - about 2% (Hill 1977). It seems also to be the case that the adpositions in the nondominant position have historically developed later than the majority which is used in the dominant position.

In this paper I attempt to specify the conditions on bipositional ordering in adpositional systems. For that task I will use the Finno-Ugric languages as examples.

2. Adpositional patterns in Finno-Ugric languages

The Finno-Ugric languages are mostly postpositional, but as already said, the Baltic-Finnic languages and Lappish make also use of prepositions. The Baltic-Finnic prepositions are commonly considered a late development in these languages, and their appearance is attributed both to the influence of Germanic and Balto-Slavonic

Proceedings of LP'94, edited by B.Palek ISSN 0567-8269

languages which make use of prepositions, and to the tendency of adverbs to turn into prepositions (Tauli 1966, Maitinskaya1982). This is undoubtedly so, but still it is not a fully satisfactory explanation as it does not explain why adverbs have not turned into prepositions in other Finno-Ugric languages or why there are no prepositions in Zyryan or Votyak (the two Permian languages of the Finno-Ugric family) though these languages have been for many centuries under heavy Russian influence and have a considerable amount of their lexicon borrowed from Russian (Lytkin et al. 1976).

These differences are usually accounted for by different sociolinguistic conditions or by the unlikeliness of borrowing or developing a new category or pattern if the language did not have it already (Maitinskaya 1982). As the Baltic-Finnic languages originally also had no prepositions, this explanation is not convincing without further argumentation - yet the strong resistance of Permian and Ob-Ugric adpositional systems to Russian influence suggests that there could be some structural constraints on bipositional ordering which are met in Baltic-Finnic languages and Lappish but not in the rest of the Finno-Ugric languages.

In fact, there are a few adpositions which Votjak (a Permian language) has borrowed from Russian: the first is '*pagu* 'for the sake of' from Russian '*pagu* 'for the sake of'; the second '*önpuч*'in addition' from Russian '*onpuч*'besides' (Maitinskaya 1982:50). These borrowings are a clear sign of Russian influence, but contrary to expectations the adpositions borrowed are exclusively used as postpositions. Furthermore, as in Russian '*pagu*' is used as a preposition most of the time and only infrequently as a postposition, and onpu is used either as a preposition or an adverb, in the latter case having the meaning 'apart', the fact that Votyak has borrowed from the Russian adpositions only these that were used also in the borrowing language as postpositions, strongly suggests that for the Votyak adpositional system, the bipositional ordering does not constitute a stable or consistent state.

If this is true, there should be quite remarkable differences between the adpositional systems of the Baltic-Finnic and the other Finno-Ugric languages. To characterise these differences better, I will give a short outline of the evolution of the Finno-Ugric adpositional systems.

The majority of Finno-Ugrists hold a view that the adpositional systems of the Finno-Ugric languages have not evolved from an equivalent in Proto Finno-Ugric, as there were no adpositions at this time, but have emerged independently in different

branches or separate languages (Lytkin et al. 1974). Most postpositions in all Finno-Ugric languages have developed from nouns via reanalysis of the head of the noun phrase as a postposition after a gradual loss of the semantic content and morphological transparency of this noun in the process of grammaticalisation. It should be noted that in the Proto Finno-Ugric language, beside the subject and object which were often not marked for case, the complements in the noun phrases (adjectives as well as nouns) also did not agree in case with their heads, but occurred in the basic nominative form. For that reason, differently from the contemporary Finno-Ugric languages, the word order in Proto Finno-Ugric must have had to be rigid (Lytkin et al. 1974).

When the adpositional systems started to evolve in the Ugric, Permian and Volga-Finnic branches, this syntactic pattern was taken over to the adpositional system. Thus, in the languages belonging to these branches, postpositions do not assign case to their complements, but these remain in the basic form or in the nominative case.

Despite the fact that all Finno-Ugric languages have developed an advanced case system by now, complements in noun phrases still do not agree with their heads in case and number in the Ugric, Permian and Volga-Finnic languages. So the structure of noun phrases in these languages is quite similar to the structure of noun phrases in English, except that the heads of noun phrases can have case endings assigned by the main verb in the sentence. If a noun phrase is a complement of a postposition the head of the noun phrase also stands in the basic form which is the same as in English, except that in English the adposition precedes its complement.

Though the postpositions in the Baltic-Finnic languages and in Lappish have also evolved mostly from nouns via reanalysis, the syntactic constructions serving as the base for the adpositional phrases in these languages were different. Unlike in Proto Finno-Ugric, in Early Proto Baltic-Finnic the adjective complements of noun phrases agreed in case and number with their heads, and nominal complements like ***brother's*** in ***brother's book*** were assigned genitive case.

As the postpositions evolved mostly from the heads of this latter construction, in the present-day Baltic-Finnic languages and Lappish they assign mostly genitive case to their complements. And differently from all other Finno-Ugric languages, in the contemporary Baltic-Finnic languages and Lappish heads of noun phrases still assign genitive to their nominal complements; and adjective complements agree with their heads in case and number.

On the basis of the structure of the adpositional systems, the Finno-Ugric languages can be divided into two groups: one group includes the Ugric, Permian and Volga-Finnic languages, the other group, the Baltic-Finnic languages and Lappish. As this division exactly matches the division of Finno-Ugric languages into exclusively postpositional languages and those that also make use of prepositions, it seems that the structural properties of adpositional systems of the first group are the cause which makes the bipositional order impossible in these languages.

3. Conditions on bipositional ordering

If we move away from the Finno-Ugric languages for a while we can test whether bipositionality and monopositionality are conditioned in other languages by similar structural properties as in the Finno-Ugric languages. From amongst the Indo-European languages at least German, Russian and Latin resemble Baltic-Finnic languages in making use of both pre- and postpositions. According to Fries (1991), there are at least 18 postpositions in German. To give a few examples, ***halber*** can occur only as a postposition as in *der Gesundheit halber* 'for the sake of health', ***gegenüber*** and ***nach*** can be used in both positions as in *gegenüber der Gefahr* or *der Gefahr gegenüber* 'in the event of danger' and *meiner Meinung nach* 'in my opinion' and *nach meinem Wünsch* 'by my wish'. In Russian, bipositionality is comparatively rare, the most common bipositional adpositions being '*ради*' for the sake of' and '*навстречу*' towards'. In Latin *tenus* 'by' is a postposition while *causa* and *gratia* 'for the sake of' can occur in both positions.

On the other hand, French and modern Greek seem to resemble the rest of the Finno-Ugric languages, permitting adpositions to occur exclusively in one position.

What is even more interesting is that in German, Russian, and Latin prepositions assign cases to their complements, and the case is marked on the surface by means of inflectional morphology. The complements of noun phrases also agree with their heads in case and number. In Swedish, French and Greek complements of prepositional phrases remain in the basic form (except pronouns), as do the complements of noun phrases.

To draw a conclusion from the resemblances between the adpositional systems of these languages, a constraint can be formulated specifying the conditions on bipositional ordering in adpositional systems:

(1) Bipositionality Condition (BC)

> No language may have both post- and prepositions unless the adpositions of this language assign surface case to their complements, and the complements of noun phrases agree in case and number with their heads.

This proposed universal is satisfied by the languages discussed so far, but there seem also to be some languages which do not obey it. For example, Dutch has mainly prepositions, but also several items which seem to be postpositions. This is a complex question, and I will discuss it in some detail.

According to Riemsdijk (1978), besides prepositions, there are also postpositions in Dutch. However, he (Riemsdijk 1978) outlines certain properties of Dutch postpositions which, in my opinion, allow them to be classified as adverbs rather than postpositions.

First, Dutch postpositions can be incorporated into the verb like particles as the following examples show:

(2)

omdat hij de tandarts { probeerde op te bellen / op probeerde te bellen }

because he the dentist { cried up to call / up tried to call }

' because he tried to call the dentist'

omdat hij de boom { is in geklommen / in is geklommen }

because he the tree { has into climbed / into has climbed }

'Because he has climbed the tree'

(Riemsdijk 1978:97-98)

Secondly, Dutch postpositions may be separated from their objects by an adverbial phrase as in (3):

(3) omdat zij de boom op blote voeten in klommen
because they the tree on bare feet into climbed
'because they climbed the tree barefoot'
omdat zij het huis na een grote ruzie uit ging
because she the house after a big row out went
'because she went out of the house after a big row'

(Riemsdijk 1978:99-100)

As Riemsdijk (1978) points out, the separation of adpositions from their objects is not possible in the case of Dutch prepositions. These two properties make Dutch postpositions indeed quite distinct from ordinary adpositions. Furthermore, by analogy with Estonian, which has a similar kind of phenomenon, I suggest that strings like ***de berg op*** 'up the mountain', ***het huis uit*** 'out of the house' etc should not be analysed as postpositional phrases.

Estonian has both prepositions and postpositions. The properties illustrated in (2) and (3) are not characteristic either of prepositions or postpositions in Estonian:

(4) a) preposition
kuna ta püüab minna läbi metsa
because he tries to go through forest&G
'because he tries to go through the forest'
*kuna ta püüab läbi minna metsa
because he tries through to go forest&G
*kuna ta püüab minna läbi palja jalu metsa
because he tries to go through bare feet forest&G

b) postposition
kuna ta vale läbi püüab rikkaks saada
because he lie&G through tries rich to become
'because he tries to become rich by lying'
*kuna ta vale püüab läbi rikkaks saada
because he lie&G through tries rich to become
*kuna ta vale kiiresti läbi püüab rikkaks saada
because he lie&G quickly through tries rich to become

Yet the same adposition ***läbi*** 'through' as well as many other motional adpositions in Estonian show the patterns found in (2) and (3), when their object which is usually in the genitive in Estonian, is instead in a locative case:

(5) kuna ta püüdis metsast läbi minna
because he tried forest&EL through to go
'because he tried to go through the forest'
kuna ta metsast püüdis läbi minna
because he forest&EL tried through to go
'because he tried to go through the forest'
kuna ta püüdis metsast palja jalu läbi minna
because he tried forest&EL bare feet through to go
'because he tried to go barefoot through the forest'

As Estonian adpositions assign genitive case to their complements, except for a few items which assign the partitive, it is not reasonable to assume that motional adpositions can also optionally assign locative case. And if we take into account the fact

that, as a result of nominalisation of sentences with such adpositional phrases, we will have no adpositional phrase at all, but two noun phrases (see (6)) there seems to be clear evidence that the alleged adpositions really belong to V:

(6) ta läks metsast läbi
he went forest&EL through
'he went through the forest'
tema metsast läbiminemine
his forest&EL throughgoing
'his throughgoing of the forest'

Further, if the preposition has its object in the genitive, nominalisation never breaks up the adpositional phrase:

(7) ta läks läbi metsa
he went through forest&G
'he went through the forest'
tema läbi metsa minemine
his through forest&G going
'his going through the forest'

As we have seen so far, there can be strings of words in languages that look like adpositional phrases, but really are just strings of adverb followed or preceded by a noun phrase - i.e. not constituents at all. In some languages, like Estonian which has morphological case, these false adpositional phrases can easily be detected by the deviant case pattern, as the so called adpositional objects are not in the case that the adpositions assign. In languages such as Dutch, the lack of morphological case makes this distinction not so apparent at the first glance. However, as the Dutch postpositions also show the same behavior as Estonian adverbs, they should not be classified as adpositions.

If my argumentation above is correct, Dutch does not have postpositions and is therefore not exceptional to BC.

Chinese is another problematic language for BC. As it does not have morphological case, it should not have adpositions in both positions. Chinese has prepositions, yet there is a small class of words (like *shang* 'on', *li* 'in', *xia* 'under', *wai* 'outside' etc)

which some linguists (Ernst 1988) claim to be postpositions. The problem of Chinese postpositions has a long history, but no general agreement on their classification has been reached yet. Three possible classifications have been suggested for this word class: either they are nouns (Li 1985, McCawley 1992), postpositions (Ernst 1988) or affixes (Starosta 1985).

My impression of this debate is that the noun analysis is the hardest to defend at least for these few items (above) which are most discussed in the literature. Ernst (1988) points out five syntactic properties of these words in which they differ from ordinary nouns. For example, these words may not occur alone, without a modifier; the preceding modifier must be a noun; usually, they have a neutral tone, which is very uncommon for head categories in Chinese; and they cannot occur in argument positions either.

These properties, while very unusual for nouns, could indeed be characteristic both of postpositions as well as affixes. Ernst (1988) who defends the postpositional approach to these words, admits the possibility of an affixal analysis, but offers no serious arguments against it. Starosta (1985) argues that the problematic words are, in fact, not words, but bound morphemes. The process of derivation where they combine with nouns he calls "pseudocompounding". He further argues that two of these affixes - *shang* and *li* - are extremely productive, and can combine with almost every noun, in which they are different from other such suffixes that occur only in certain fixed combinations. They also contrast with *shang* and *li* that occur in pseudocompounds. Two of his examples are given in (8):

(8) a) chéng.lǐ 'downtown'
qiáng de shang.tou '(the space) above the wall'
b) chéng.lǐ 'in the city'
qiáng.sihàng(.tou) 'on the wall'

As can be seen, the affixes *li* and *shang* in (8a) have their own tone contour and the words containing them are semantically less transparent than their counterparts in (8b) where *shang* and *li* do not have their own tone. On the basis of this and further evidence, Starosta (1985) concludes that these two morphemes should be classified as inflectional case suffixes.

Above I have outlined some arguments presented in Chinese linguistics in favour of three possible classifications for a class of linguistic units. On the basis of the

evidence available to me, it seems that the noun analysis is the most unlikely, whereas the decision in favour of either postpositional or affixal analysis is very hard to make. I conclude that Chinese may or may not be exceptional to BC.

Unlike Dutch and Chinese which have been claimed to have both types of adpositions - pre- and postpositions, there are languages which, according to BC, may have adpositions only in one position, but also have one or two highly restricted adpositions which occur in nondominant position. Swedish and English can serve as examples.

In Swedish, *runt* 'around' is a postposition (as for example in *jorden runt* 'around the world'), and in English, *ago* is claimed to be a postposition and *notwithstanding* to be able to both precede and follow its nominal complement. While the Swedish *runt* appears to be a genuine postposition, there seems to be a way of questioning postpositional analysis for English *ago* and *notwithstanding*. Kortmann & König (1992) point out that *three years* in *three years ago* could be regarded as specifier rather than complement of *ago*, as there is also the possibility of saying *three years ago today*, in which case *today* serves as a complement, omitted in the first example. They also note that *notwithstanding* if following a NP could be an adverb. Be this as it may, as I will show in the next section, single counterexamples to BC can be easily accounted for.

Though the apparent counterexamples to BC known to me do not invalidate it, at least conclusively, it is still possible that there exist languages which are exceptional for BC. In this case BC should considered a statistical tendency rather than an absolute universal. I leave this question open and proceed to the problem of why bipositionality is sensitive for the presence of inflectional morphology.

4. Explaining bipositionality

There are two possible solutions to this problem: either bipositionality is a consequence of the formal structure of grammar, or it is caused by some functional properties of language usage. In the first case BC has a formal explanation, in the second case a functional one. Let us consider the first possibility.

Following X-bar theory, Universal Grammar (UG) is assumed to contain a very general principle which defines possible phrase structures:

(9) XP → Spec; X'
X' → X'; YP
X' → X; YP

As the order of constituents with respect to the head of a construction is subject to variation across languages, the exact order of specifiers, heads, adjuncts and complements is not specified in UG. Attempts have been made to relate the basic word order to the interaction of parameters which specify the directions of theta-role and case assignments (Li 1985), but this is not very reliable as there is a number of languages in which verbs assign case in one direction and adpositions in the other. It is reasonable then to assume that the relative order of constituents is fixed by the linguistic environment. Thus, if a language has prepositions, a child growing up in this language environment will specify the universal rule X' → X; YP for adpositional phrases as (10a). If the language has postpositions, the rule will be specified as (10b), and if both pre- and postpositions are used, both (10a) and (10b) will be learned.

(10) a) P' → P - NP
b) P' → NP - P

This generates the appropriate results, but it does not explain why the grammars of certain languages historically are unable to develop both rules. If such an ability is dependent on the presence of morphological case, it could be desirable to make the setting for the rules in (10) dependent on the setting for morphological case.

Let us assume that in the learning process, the child, following UG, is unbiased as to the ordering of adpositions and their complements. So, the child is satisfied with the rule in (11):

(11) P' → P; NP

If, however, the parameter for morphological case has been set to 0, the child is forced to specify (11) further, to allow only one possibility (either (10a) or (10b)).

This analysis has two weak points. First, the assumed parameter for morphological case is not as straightforward as it seems. It is not a question of whether a language has or has not got morphological case. As specified by BC, the co-occurrence of pre- and postpositions is sensitive only to surface case marking of the complements of PPs and NPs. Thus, a language may have morphological case, and still not allow both pre- and postpositions as far as the complements of PPs and NPs do not decline. A good

case is Hungarian which has a number of cases, but the complements of PPs and NPs always stand in the basic form. Following BC, Hungarian has only postpositions. Votyak, also a Finno-Ugric language which has morphological case, but nondeclinable complements in PPs and NPs, has not developed a single preposition, despite being under heavy Russian influence for centuries.

Now, if we believe that there is a very specific parameter which handles morphological case marking for complements of [-V] heads, and the value for this parameter triggers the parameter value for head-complement order in adpositional phrases, then we will have to assume that children have to master almost the whole morphology before they can rule out Adv;NP sequences as impossible PPs. Maybe this is the case, and Hungarian children really misanalyse Adv-NP as prepositional phrases in their early stages. A further study is needed to test this hypothesis.

The second counter-argument derives from the basics of the generative theory. According to the Subset Principle (Berwick 1985), learning from positive evidence can be successful only if the rules specified are consistent with the narrowest language that it is possible to infer from the available data. Thus, a child acquiring rules for adpositional phrases of a language consistent with BC, but accidentally not having developed bipositionality, cannot have (11) in its grammar, despite the dispersion allowed by BC, since this would violate the Subset Principle, and must opt either for (10a) or (10b), whichever is appropriate. This means that we could not have different formal representations for potentially bipositional languages and monopositional languages.

And finally, even if we find a way to avoid the Subset Principle, in which sense we could consider the account similar to that, presented above as an explanation? By stating BC in formal terms we can only shift the problem from the level of the structure of language to the level of the structure of grammar, but this itself does not answer the question of why there should be such a correlation.

Now, as the formal explanation for BC suggested above leaves us with some questionable points and actually does not explain anything, let us turn to the possibility of a functional explanation.

If a language does not have morphological case, the syntactic relationship between an adposition and its complement is expressed by means of word order. Therefore, if both pre- and postpositions were possible, the parsing of strings like NP-P-NP would

create processing difficulties. This also explains the fact that BC is sensitive to morphological case in complements of NPs and PPs only: as they are the most likely governees for Ps, something has to signal the government.

If a language has morphological case, government is expressed both by means of word order and case assignment, and therefore the ambiguity of NP-P-NP strings would be eliminated, or at least reduced.

This explanation certainly saves us from the rigorous mechanism where the parameter for head-complement order in adpositional phrases is set by a morphological parameter. Furthermore, the functional explanation also allows us to account for counterexamples of BC.

For example, in Estonian some noun classes have the same case form for the nominative, genitive and partitive case. This means that in the case of such words (as subjects are mostly in nominative, objects in partitive, and adpositional complements in genitive case), ambiguous NP-P-NP strings may arise. This property of Estonian constitutes a local violation to BC. Similarly, the Swedish *runt* and English *ago* and *notwithstanding* constitute local violations to BC. Within the formal framework these cases can only be treated by formulating the exceptional properties of these words in their lexical entries in the lexicon. This again does not explain why these exceptions are allowed in this language. Within the functionalist approach, however, it is possible to account for these exceptions by hypothesising that such violations are too slight to create serious problems for comprehension. For example, it might be argued that Swedish *runt* and English **ago** and *notwithstanding* are tolerated since they usually occur in a narrow and well-defined environment where the ambiguity does not arise. Similarly the defectiveness of case marking in some Estonian word classes does not prohibit bipositionality as the chance of it creating ambiguity is too low to make the state intolerable.

Of course, it might be argued that this hypothesis is unsatisfactory since it allows us to explain away any possible counterexamples by simply saying that they do not disrupt comprehension. This is true, but the hypothesis can be made more precise and restricted by a systematic study of such violations in different languages. For example, if the ambiguity explanation for BC is correct we would expect that there are fewer exceptions to BC in SOV and VSO languages, and languages with free word order since in these languages NPs are often adjacent to each other. Similarly, we would

presume that adpositions exceptional to BC have a well-defined meaning, allowing only a narrow class of semantically restricted complements which again would reduce ambiguity. Finding out these restrictions is a complicated task which requires precise data from a large set of languages, but I feel that it can be far more illuminating than simply stating the regularities and exceptions in formal terms and appealing to the UG as the ultimate explanation.

5. Conclusion

In this paper I have tried to explicise the conditions on bipositional ordering in adpositional systems and provide a plausible explanation to this phenomenon. The results, though somewhat preliminary, suggest that bipositionality in adpositional systems is sensitive to the presence of inflectional morphology, and that this is due to language processing considerations - if the government between the head and its complement is expressed by word order only, as in languages with no inflectional morphology, bipositionality can increase ambiguity. If government is expressed both by means of word order and case assignment this possibility is eliminated.

References

Berwick, R.C. (1985) *The acquisition of syntactic knowledge*. The MIT Press: Cambridge, Massachusetts.

Ernst, T. (1988) Chinese postpositions? - again. *Journal of Chinese Linguistics*, Vol 16, 2: 219-245.

Fries, N. (1991) Prepositions and prepositional phrases: a contrastive analysis. In: Rauh (1991): 53-76.

Hill, S.P. (1977) *The* n*-factor and Russian prepositions*. Mouton Publishers: The Hague.

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

Li, Y. (1985) *Abstract case in Chinese*. PhD dissertation, USC, Los Angeles.

Lytkin, V.I. et al. (1974) *Osnovy finno-ugorskovo yazykoznaniya. Voprosy proishozdeniya i razvitiya finno-ugorskikh yazykov*. Nauka: Moscow.

Lytkin, V.I. et al. (1976) *Osnovy finno-ugorskovo yazykoznanya. Mariiskiy, permskie i ugorskie yazyki*. Nauka: Moscow.

Maitinskaya, K.E. (1982) *Sluzhebnye slova v finno-ugorskikh yazykakh*. Nauka: Moscow.

McCawley, J.D. (1992) Justifying part-of-speech assignments in Mandarin Chinese. *Journal of Chinese Linguistics*, Vol. 20, 2: 211-245.

Rauh, G. (ed.) (1991) *Approaches to prepositions*. Gunter Narr: Tübingen.

Riemsdijk, H. van (1978) *A case study in Syntactic markedness*. Foris Publications: Dordrecht, Holland.

Starosta, S. (1985) Mandarin case marking: a localistic Lexicase analysis. *Journal of Chinese Linguistics*, Vol. 13, 2: 216-266.

Tauli, V. (1966) *Structural tendencies in Uralic languages*. Indiana University Publications. Uralic and Altaic Series. Vol. 17. Mouton: The Hague. .