Anders Holmberg
Department of Linguistics
University of Umeå

PROPERTIES OF NON-HEADS IN COMPOUNDS: A CASE STUDY

In this paper I will show that noun-noun compounds in Swedish conform to the following generalization:

(1) The head of a compound is a Word. A non-branching non-head of a compound is either a Root or a Case-marked word, a branching non-head is a Case-marked word.

This generalization will be explained in terms of a theory of word structure and Case. The following hypotheses are crucial in this theory:

(2) a. A Root is a maximal category.

b. Swedish nouns consist (minimally) of two morphemes, a Root and a Word Marker.

A Word Marker is a morpheme which may encode certain grammatical properties such as gender or number but whose essential grammatical function is to combine with a Root, and thereby permit the Root to take part in word- and phrase-syntactic processes. Various theoretical consequences of the theory will be discussed. I will also discuss some differences between Swedish and English word structure.

1. The structure of compounds in Swedish

There is a broad consensus that a compound in English is a word...
level category \( X^* \) made up of two or more words. Thus a noun-noun compound will have the structure (1):

\[
(1) \quad N \quad N
\]

See Selkirk (1982), Sciullo and Williams (1988). This is supported by paradigms such as the following:

(3) a. wallpaper  
b. paper wall  
c. wallpaper flower  
d. paper wall flower

The same constituent (wall, paper) can serve as head or as non-head in a compound without any morphological changes, and the compound may itself be a constituent of a complex compound (typically non-head, but also as head).! Thus the only structural distinction between head and non-head in English compounds appears to be linear order: the non-head precedes the head (the Right-hand Head Rule of Williams (1981)). Expressed in terms of government direction: \( X^* \) governs leftwards within \( X^* \). This distinguishes words from phrases, where characteristically head and non-head are usually distinguished by linear order but also by bar-level (heads are non-maximal, non-heads are maximal), and often by morphological marking, for instance case.

In Swedish the corresponding paradigm looks rather different. Consider first simple compounds ("+" separates head and non-head)

(4) a. skol+flicka  
school girl  
b. flick+skola  
girl school

With a large class of Swedish nouns (the 1st and 2nd declensions, so-called) the form of the non-head is systematically distinct from that of the head. These nouns consist of a Root (henceforth \( R \)) and a "stem vowel" (-a or -e).

(5) [\( N \) [\( R \) flick] a], [\( N \) [\( R \) poj] e] ("girl", "boy")

Only the \( R \) form is possible in non-head position, while only the \( N \)-form is possible in head position (in simple compounds). Compare (4a,b) and (6):

(6) *skola+flicka, *flicka+skola, *skol+flick, flick+skol

The non-head form of constituents in compounds is the same form that is found in combination with derivational affixes.

(7) skol+nin, *skola+nin  
school+ing  
flick+aXtig, *flicka+aXtig  
girl+ish/like

The generalization is: The head of a compound is a Word but the non-head is a Root. The structure of a simple noun-noun compound in Swedish is:

(8) \[ \begin{array}{c} \quad N \\ \quad R \\ \quad N \end{array} \]

Now consider complex compounds:

(9) a. skol+flick+s\( \text{d} \)om  
school girl dream  
b. flick+skola+s\( \text{d} \)om

When a compound is itself embedded as non-head of a compound the stem vowel (in the case of the relevant nouns) disappears and instead there is, obligatorily, a suffix -s.

(10) a. *skol+flick+s\( \text{d} \)om, *skol+flick+s\( \text{d} \)om

---

1 Incidentally, wallpaper is spelled as one word and paper wall as two, which also means that spelling disambiguates (cf). while constituents of English compounds are generally spelled as separate words, Swedish compounds, even very complex ones, are normally spelled as one word.
This is the case with binomophenic (i.e., branching) nominal non-heads in general, whether formed by compounding or derivational affixation; they are marked by -s.

(11) a. frihet, frihet+s-behov, *frihet + behov
freedom, freedom need

b. barn+dom, barn+dom+s +van, *barn+dom +van
childhood, childhood friend

In complex compounds with a branching head and a non-branching non-head there is no -s. Thus we get minimal pairs like (9a,b) and (12a,b):

(12) a. skol+flickfröms
   e.g. "a kind of dream that girls have at school as opposed to at home"

b. flickskoldrom
   e.g. "a dream about school typical of girls as opposed to boys"

The suffix -s is homophonous with the genitive suffix -s, and is without doubt historically related to the genitive -s found in phrasal possessive constructions such as (13):

(13) a. en skolflicka+s dröm
    a schoolgirl’s dream

b. min barn+dom+s stad
    my childhood’s town ("the town of my childhood")

In present day Swedish the genitive -s in possessive constructions has morpho-syntactic properties which are quite distinct from those of the -s of complex compounds. In particular, just like English genitive -s it is a "phrasal affix" (or clitic), which adjoins to NP, not to N (see Fiva (1985)). Whether we identify the -s found in complex compounds as a form of genitive or not (in synchronic grammar) is not crucial. What is crucial, though, is that the -s found in complex compounds is a Case morpheme.

Consequently the following generalization can be formulated:

(14) The head of a compound is a Word. A non-branching non-head of a compound is a Root. A branching nominal non-head is Case-marked.

One indication that it is Case is that it is found only on nominal non-heads.

(15) a. röd+målad, *röd+farg
    red painted, red paint

b. mörkröd+målad, *mörkröd+farg
    dark red painted

   mörkröd+farg, *mörkröd+farg
    dark red paint

2 (The rule which assigns Case to branching non-heads is found in German as well (Fahrerlaubnis, Durchfahrerlaubnis; Halter, Kirchhofhalter). Indeed, the rule seems to have been imported into Swedish from German; it is not found in Old Scandinavian, as far as I know. However, I have the impression that the present day German rules of compounding are more complex than the Swedish ones, with more exceptions and sub-systems. In other words, it seems that the core rule system of Swedish compounding covers a larger part of word syntax than in German, where there is more periphery (in terms of the core-periphery distinction of Chomsky (1981: ch.1)).

3 There are two major classes of exceptions to the rule according to which a branching non-head must be Case-marked (see Wehnerberg (1983)): a. If the root is vocalic, there is usually no ( overt) Case:

   (i) [tåge-sklep] + klack, *[tåge+sklop] + klack
       lady shoe heel

   (ii) [förstör+rasse] + lust, *[förstörr+rases] + lust
       destruction desire

   I assume that these nouns take a vocalic Case suffix (like many nouns do: see text below) which is asssociated to the preceding vocal and hence inaudible at RP.

b. Nouns borrowed from French or formed according to the pattern of French loan words, which is to say that they have primary stress on the final syllable, tend not to take Case-marking.

   (iii) [data-program] + konstruktör,
       computer program constructor

   (iv) [skogs+promenade] + skater
       woods walk

b. Nouns borrowed from French or formed according to the pattern of French loan words, which is to say that they have primary stress on the final syllable, tend not to take Case-marking.

Clarity counterexamples to the rule can be found occasionally. For instance the official name of the Swedish national football association includes the complex compund fotboll+föreundet instead of the expected fotboll+föreundet.

I believe forms like this are written language constructs (in the case at hand probably influenced by the English term football association), not part of spontaneous spoken language word formation. Of course one could not

exclude the possibility that there is genuine diatopic variation concerning the rule in question.
(16) nedåtvand, uppåt+trend, nedåtvand, uppåt+trend
downward trend, upward trend

(17) a. skriv+maskin
write machine ("typewriter")
b. *snabbskriv+maskin, *snabbskriv+maskin
quick write machine

In (15) the non-head is an adjective, in (16) it is a locative
adverb, and in (17) a verbal root. As shown, a non-nominal
branching non-head cannot take -s. This is consistent with the
analysis of this suffix as a form of Case inflection. What matters
is the category of the non-head, not the category of the head. Cf.
(18), where the head is adjectival (or participial):5

(18) a. *brandbil+röd, brandbil+röd
fire engine red
b. *morot+formad, morots+formad
carrot shaped (N.B. morot is biformorphic)

Case-marking inside compounds is found in other languages, too.
For instance English and Finnish. In both languages the Case is
genitive.

(19) a. driver's seat, men's magazine, bird's nest
b. kuljetuajen+paikka, mieste+n+lehti, linna+n+pesä
driver GEN seat, men GEN magazine, bird GEN nest

4 In particular when the head is a noun, a branching verbal or adjectival
non-head is at best marginally acceptable (e.g., *snabbskrivmaskin). However,
the form with -s added is totally unacceptable. We can explain why branching
non-nominal non-heads are not good: Since they are branching they cannot be
roots, but since they are non-nominal they cannot be saved by Case. We
predict, in fact that they should be completely impossible. Possibly the
reason why mörkrödfärg and snabbskrivmaskin are marginally acceptable is that
they allow the analysis mörk-rödfärg and snabbs+skrivmaskin, an interpretation
which is not significantly different from the interpretation of mörk+rödfärg and
snabbs+skrivmaskin. In other words, the latter are indeed syntactically ill formed, but semantically well formed, and the semantic
interpretation can be recovered (marginally) from the alternative,
syntactically well formed analysis mörk+rödfärg and snabbs+skrivmaskin.

5 Since there are no verbal compounds (apart from participial ones, which
should perhaps be analysed as adjectival) we cannot test whether a nominal
branching non-head would have (the same) Case in a verb-headed compound.

In both languages the appearance of genitive on non-heads in
compounds is semantically conditioned, at least in part, and
furthermore very much in the same way, as it looks. According to
Warren (1978) the genitive is used in English compounds primarily
when the relation between head and non-head is a possession
relation (bird's nest) or what she refers to as a purpose relation
(men's magazine).6 It seems to me that very much the same is true
of the use of genitive in Finnish compounds. In neither English
nor Finnish is the appearance of case in compounds structurally
conditioned, the way it is in Swedish.

There are quite a few nouns which require Case-marking even
when they occur as non-branching non-heads of compounds. Some
examples are given in (20).

(20) stol+s+ben, skog+s+bryn, sätte+s+fält, minnes+förlust
chair leg forest edge grain field memory loss

Requiring Case in this way is an idiosyncratic lexical property
of some nouns, subject to dialectal/idiomatic variation. The nouns
which require Case in compounds seem to have no other properties
in common, except for nouns of the class to which minnes "memory"
belongs (bisyllabic, ending in -e, taking -n as a plural marker).
These nouns take -s obligatorily when they occur as non-heads in
compounds. Taking cases like (20) into account, the generalization
is (21):

(21) The head of a compound is a Word. A non-head is either a
Root or it is Case-marked.

This generalization entails that a non-head which is itself a
compound must have Case, as it cannot possibly be a Root.

In some cases there is a vowel, often called a "linking vowel",
marking the non-head of a compound, branching or non-branching.
For instance (22a), occurs as an alternative of (22b) (subject to
dialectal variation). (22c) contains a couple of examples of nouns
which take a linking vowel even as non-branching non-heads.

6 It seems to me that possessor and purpose can be grouped together as
experiencers, taking this notion in a wide sense.
Nominal Word Markers in Swedish

What is the nature of the ‘stem vowel’ in nouns like **flicka**, **skola**, **pojke** etc.? What grammatical features, if any, do they encode? This is a controversial issue in Swedish morphology. According to one tradition, represented by Rieffer (1970), it is a stem forming affix, and as such encodes no features except a declension class feature.8 According to another tradition, represented by Dahlstedt (1965: 105) and more recently Ejerhed and Granley (1986), it is a number morpheme encoding [-PL] (and possibly also gender). See Linell (1972) for discussion.

A reason for analyzing the stem vowels -e/- or as number affixes is that they are each paired with a plural form: -a is paired with the plural suffix -or (or -er in many varieties of colloquial Swedish), while -e is paired with -ar.

(23) a. -e/or **flicka**, **flickor**
   SG PL girl girls
b. -e/ar **pojka**, **pojkor**
   SG PL boy boys

Formation of the plural would thus be straightforward. The affixes -e/or form a class of number affixes which select roots which are lexically marked for a certain declension class, call it class A. Thus roots like **flick-**, **skol-** etc. are marked with a feature [A]. The affix, -e is marked [-PL, A], while -or is marked [+PL, A], thus these affixes take roots like **flick-**, **skol-** etc. as complement. Correspondingly the affixes -e/ar select roots marked for a different declension class, call it class B. In this way forms like **flicka**, **flickor**, **pojka**, **pojkor** etc. are ruled out.

If -e/- is analyzed as part of the stem (unmarked for number), formation of the plural is complicated in that the stem vowel must be deleted (or not spelled out; see footnote 8) before addition of the plural. Alternatively, if the plural forms are analyzed as **flickor**, **pojkor** etc., you need a rule to change the

---

7 The following Icelandic paradigm is typical: The verb **fjóla** ‘increase’ takes a dative Theme argument, but the nominalized form fjólsun takes a genitive Theme argument (data provided by Halldór A. Sigurðsson, p.c.):

i. A. **Flugfelagi** fjólsun ferðunum
   The-airline-company increased the-travel+DAT
b. Ferðunum fjólsun.
   The-travel+DAT increased
c. fjólsun ferðunna
   Increase the-travel+GEN (*the increase of travel*)
d. *fjólsun ferðunum
   Increase travel+DAT

Icelandic has some constructions, however, where a predicative noun assigns lexical Case:

ii. Jóni er vorrinn.
   John is pitiful (*John is pitiable*)

---

8 In Rieffer’s theory, adopted and elaborated in Linell (1972), the roots in question are marked with a morphological feature, either [+a-Stem] or [-a-Stem]. The -e and -a affixes are spell-outs of these features.
quality of the stem vowel according to the declension class of the noun.

The two stem vowels have a connection with gender as well. Swedish has two grammatical genders, neuter and common gender. Nouns of the -a/-e class are all common gender, with no exceptions as far as I am aware (common gender has developed out of the earlier feminine and masculine genders, the -a class being historically feminine and the -e class historically masculine). Thus we may attribute to -a and -e the feature [-NEUT]]. On the other hand, as already mentioned the roots in question must be marked lexically as belonging to a particular declension class, in order to be selectable by the appropriate affix (to exclude flicke, flicker, etc.). Gender (in this case [-NEUT]] could be encoded on the declension class marker on the root (say, as a value of the declension class feature) instead of being encoded on the affix.

The stem vowels also have a connection with sex (natural gender). Nouns denoting females very often end in -a, and rarely, if ever, end in -e, while nouns denoting males rarely end in -a. Consider e.g. the following list of slang words (most of them obsolete) for "girl" and "boy", respectively:

(24) fjalla, böna, pingla, majsa, brutta ("girl");
    snubbe, kille, tjomme, jappe ("boy", "man")

With the exception of böna and pingla, which have a standard meaning as well ("bean" and "bail", respectively), these are all new coinages, and thus exemplify a productive rule of word formation. The rule is, informally, that words for "girl" and "boy" are either consonantal or end in -a for "girl" and -e for "boy". However, as the majority of words in the two noun classes denote sexless objects (matta "mat", hylla "shelf", gälge "hanger" etc.) it would clearly be misguided to attribute to these suffixes a feature for sex (say, [FEM]) even in the case of words like (24). Instead the connection between form and sex can be expressed as a lexical redundancy rule (cf. Harris's (1991) discussion of gender markers in Spanish).

In conclusion, I assume the Swedish nominal suffixes -a and -e encode [-PL].

Harris (1991) argues that the Spanish nominal suffixes -a/-o (muchacho/muchacha) do not encode gender features, contrary to the traditional view. Instead, he claims, -a/-o are exponents of declension class and nothing more. The members of the declension classes, he claims, have nothing in common (such as gender, etc.) except being members in the declension class. Thus Spanish -a/-o are "markers of pure form". Harris proposes to express existing relations between -a/-o and gender by means of redundancy rules. Harris refers to Spanish -a/-o as "word markers", where a word marker is an element which when added to a root or a stem forms a complete word.

If my analysis of Swedish -a/-e is correct, they are not "markers of pure form", since they encode (at least) number. I will, however, refer to Swedish -a/-e as word markers. The idea is that -a/-e belong to a class of morphemes the primary grammatical function of which is to make a Word out of a Root. These morphemes combine with a Root, and thereby allow the Root to take part in syntactic processes. In Swedish nominal word markers also encode number, incidentally as it were.

The structure of a noun of the -a/-e class, then, is as shown in (25) (where WM = word marker):

(25) R ———— WM
    ℓwick -a

---

9 A possibility worth considering is that the Word Marker encodes the lexical categorial feature(s) 3N, 4V. In this view a Root does not itself have categorial features, but can be made (part of) a noun, adjective, or verb by combining with the appropriate Word Marker. (I assume there are verbal and adjectival Word Markers as well as nominal ones. See section 7 below.) A Root itself would have only semantic features, including a theta grid. These semantic features can take part in semantic composition only by the mediation of syntactic categorial features provided by a Word Marker. Such a radical position is hard to maintain, though (for Swedish and cognate languages), in view of the fact that derivational affixes often select a root according to lexical categorial features, as it seems (e.g. in Swedish -ig selects roots which occur in nouns, -het selects roots which occur in adjectives, and so on).
3. Explaining generalization (21)

3.1 A Root is a maximal category

According to generalization (21), the non-head of a compound must be a (plain) Root or be Case-marked. It cannot be a Word without Case-marking. Why not? Why are the forms in (26) ungrammatical?

(26) *flickaskola, *skolisflicka, *pojkedrom

In other words, what do Roots and Casemarked Words have in common which separates them from Words without Case?

A non-head is by definition a non-projecting, i.e. maximal category, while a head is by definition a projecting, i.e. non-maximal category. But differently, if the (grammatical) features of a category A percolate to the node B immediately dominating A, then A is the head of B. If the grammatical features of A do not percolate to the node B immediately dominating A, then A is a non-head constituent of B. The fact that simple non-heads in Swedish compounds are typically Roots strongly suggests that Roots are maximal categories, as proposed by Muysken (1983). I assume (27):

(27) a. A Root is a maximal category.
   b. A Word is a non-maximal category.

(27b) should be entirely uncontroversial: Words (X’s) are heads in syntax, projecting their features onto X’ and XP. The grammatical features of a Root, by hypothesis, are not projected onto the immediately dominating node. Assuming (for the time being) that there are only two morphological X-bar levels, Roots and Words, then (27) explains the first clause of generalization (21) (the head of a compound is a Word). A Root, being a maximal category, cannot function as head. This rules out R R compounds (e.g. *skolisflick). (27) also explains the second clause of generalization (21) (the non-head of a compound is a Root). It also rules out W W compounds such as (26), the structure of which would be (28):

(28)

The structure has two heads, W2 and W3, projecting (all) their features onto W1, resulting in an uninterpretable construction.

What is the structure of a well-formed compound such as flickskola? The null hypothesis would seem to be (29). This structure is compatible with all of the assumptions made so far.

(29)

However, in this paper I will assume the more complex analysis in (30). (30a) is the D-structure, and (30b) is the S-structure, derived by word-internal head-movement.

(30) a. 

b. 

10 I have a notion, which will be substantiated only in part in this paper, that this analysis is preferable when a wider range of facts are taken into account, including other types of compounds and other structures involving inflection. A possible argument in favour of (29) is that selection of the root by the affix is more straightforward. In (30), the selection (strict subcategorization) feature of the affix is satisfied only in S-structure. In D-structure the sister of the affix is a featureless R1 compound.
Note that the category type labels are different in D-structure and S-structure: The Root skol moves and adjoins to the WM (or possibly substitutes into a position as complement of the WM). The index of the Root skol percolates to the W immediately dominating R and WM. This W thus heads the chain whose foot is the trace of the moved Root. This trace inherits W-status from the head of the chain. Therefore, at S-structure the compound embedded under the WM is a R W compound. That is to say, a R R compound is possible, in D-structure but not in S-structure. In S-structure the tree should be ready for semantic interpretation (abstracting away from effects of LF-movement, which are irrelevant for our present concerns), which is to say that the features of the various heads should be appropriately distributed in the tree, each node having features as determined by its head and general principles of feature percolation. Since a Root is a maximal category, a R R compound has no features, and therefore cannot be interpreted. The two Roots in (30a) can be unordered with respect to each other; the order is imposed by movement of one of the Roots to the WM.

The same mechanism of head-chain formation and inheritance of head-status is operative in connection with V-to-I movement. In languages such as the Germanic and the Romance languages the category which is moved in so-called V-to-I movement is not really V, but the verbal Root.11 We probably do not want to say that the trace of the moved verbal Root is itself a Root, though. In particular, if a Root is a maximal category as I have argued, we do not want to say that the trace is a Root, since in that case there would be no VP (or if you like, VP would be headless). Instead let us say, as is fairly standard, that the index of the verbal root percolates to I. Hence the category which locally binds the trace of the moved verbal root is I containing the verbal Root. This category is a head, i.e. a W (which is to say that I is a verbal WM). The trace inherits W-status from I containing the verbal Root, and consequently is able to function as head of VP. This is shown schematically in (31), where (31a) represents D-structure and (31b) S-structure (I assume the

---

11 This means that V-to-I is not strictly speaking head movement, but movement of a maximal category, although the chain created by V-to-I is a head chain, as discussed below.
maximal category by virtue of containing a governed Case-feature. Adopting a “checking approach” to Case, I will say that the Case feature is licensed by the governing head. This is all expressed in (34):

(34) a. Case must be licensed by a governing head.
   b. If a category A has a licensed Case feature, then A is maximal.

(34a) is a generalized version of the standard Case rules (nominative is assigned/licensed by I, genitive is assigned/licensed by N, and so on).

The complementary distribution of the -e suffix and the stem vowels -a/-ə is a decisive link in the analysis of these vowels as word markers, hence in a sense as inflectional elements rather than as “stem forming affixes”. The Case-affix clearly is not a stem forming affix: it is insensitive to declension class, and generally to lexical features of the root it is affixed to, except in cases like (20) and (22) where there seems to be a selection relation between the affix and the root. Except for these cases, the distribution of the Case affix is structurally conditioned: the Case-affix is obligatory on a branching non-head. Since the Case-affix is a word marker, and since it has complementary distribution with -a/-ə, it is natural to analyze -a/-ə as a word marker, too.12

I doubt whether Case-affixes in general can be analyzed as WMs. If not we can sum up the properties of word-internal Case in Swedish as follows: Like other instances of Case it closes (maximizes) a nominal projection when licensed by a governor. Unlike (perhaps) some other instances of Case, but like a least some instances of number and tense, it is a WM.

4. Case-marking and incorporation

Another context where we find a complementary distribution of roots and Case-marking is Noun Incorporation in Baker’s (1988a) sense. Consider (35a) (from Baker 1988a), exemplifying a pair of synonymous sentences, where in the first sentence the object has been incorporated into the finite verb, while in the second sentence the object remains in object position. Similar pairs can be constructed in many other languages which have Noun Incorporation.

(35) a. Poe waʔ+haʔ+hwist+ahuʔ+ʔaʔ. (Onondaga)
   b. Poe waʔ+haʔ+ahuʔ+ʔaʔ. neʔ aʔ+hwist+aʔ?
   Pat T+AGR+money-lose+CAUS+ASP
   Pat T+AGR+lose+CAUS+ASP the PRE-money-SUF

In (29a) the incorporated object is the noun root (see also Baker 1988b), while in (35b) it has a suffixed form which, as far as I can make out, is a Case-marked form. On the basis of such pairs, the following generalization can be formulated:

(36) A nominal argument either has Case or is incorporated.

That is to say, an incorporated nominal argument does not need Case. Absence of overt Case-marking does not necessarily imply absence of Case, given the existence of abstract Case, but Baker (1988: 105ff.) argues carefully, on the basis of other facts than just morphological ones, that the incorporated argument really lacks Case.13

Baker explains the generalization in terms of a theory of ‘PP-identification’. The idea is, basically, that arguments should be morphologically identified, in order for theta-role relations to be visible at PP. One way is for the argument to have Case assigned by the verb, another way is to incorporate the argument into the verb. In spite of the suggestive rhetoric this is not very satisfying. Why does incorporation, of all logically possible

12 The Case morpheme clearly encodes no other features than Case, i.e. no number or gender. Possibly this is an essential property of Case, which makes possible using Case as a form of default Word Marker of non-heads in this way.

13 Baker argues that incorporated nouns in some languages have Case (Greenlandic Eskimo is a case in point) (Baker 1988a: 124ff.). This is compatible with the present theory, of course.
syntactic processes have the same effect as Case-marking?

But given that

(a) one of the functions of Case is to close a nominal projection, thus turning it into a non-head;
(b) the incorporated nominal form is a Root, and
(c) Roots are maximal,
then it follows that incorporation can substitute for Case. In both (37a) and (37b) the nominal category is unambiguously maximal, hence a non-head.

(37) a.  R → V → V'  
      [+n]  
      [Case] 

b.  V → V' → NP 
      [Case] 

This is compatible with a Baker-style movement analysis of Noun Incorporation, if we assume that the trace of the incorporated Root can inherit the property of being maximal from the Root (following Huyssen (1983) I represent the property as a feature [+max]). In this way the trace, presumably a structureless empty NP, is maximal by virtue of incorporation.

(38)  R → V → V'  
      [+max]  

      NP  
      [+max]  

It seems to me that this theory offers a better analysis than Baker's of the possessor stranding construction. In this construction, found in several incorporating languages, a possessor argument of the object noun is stranded by incorporation, and is assigned objective Case in this position, rather than the possessive Case expected on account of its relation to the noun. The pattern is exemplified in (39) with English glosses (see Baker (1988: 96ff.) for authentic examples):

(39) a. Mary painted John's house.  
b. Mary painted John's.

The structure, under the present analysis will be (40):

(40)  

R → V → V'  
      NP  
      [+max]  

[Case]  

[Case]

In this case the index of the moved root (house) does not percolate to the immediately dominating node (house-painted), because house and painted are categorially distinct. (If it did percolate, the trace t1 would be a verb trace.) The antecedent of the trace t is, therefore, the Root house, a maximal category. Consequently the trace is a maximal category. We therefore do not expect it to be able to assign Case to the possessor. Formally the possessor in the possessor stranding construction is an adjunct to a (maximal) NP, and is thus in the government domain of the verb. This account seems preferable to Baker's. Baker proposes that traces of moved heads, as a general rule, are unable to assign Case.14

But as well known from studies of V-movement, the trace of V can assign Case.15 E.g. in (41), there is no other source of the Case of the object than the trace of the verb.

(41) a3 Jón keypiti1 [yp ekki [yp t1 bókins]] (Icelandic)  
    that John bought not the-book(ACC)

We can also explain why the trace of the verbal Root here can assign Case to the object, while the trace of the moved nominal Root in possessor stranding constructions cannot. This is because the moved verbal Root is incorporated in a WM, a verbal inflection (say \( \theta \)). The index of the verbal root percolates to the node dominating R and T. The local binder of the trace is thus a W-

---

14 Baker's formulation is 'If A is a trace, A cannot assign a Case index to R' (Baker 1988a: 121).
15 See for instance Holmberg (1986: 175f.).
level category, and the trace is thereby a W-level category, i.e. a head, capable of Case-assignment. The incorporated nominal Root, on the other hand, is incorporated in a verb. The index of the nominal Root does not percolate to V, since the two are categorially distinct. The local binder of the trace is therefore the Root, a maximal category, and hence the trace is a maximal category, incapable of Case-assignment.

5. Generalizing the analysis of Swedish noun structure

Not all Swedish nouns belong to the -a/-e class. A great many nouns do not have a vocalic, overt Word Marker. For instance the ones listed in (42) do not.

(42) bild, knapp, stol, minne, uniform, frihet, ungdom
picture, button, chair, memory, uniform, freedom, youth

If the analysis of number marking with nouns of the -a/-e class is generalized to other nouns as far as possible, the structure of e.g. bild would be (43a) and knapp (43b):

(42) a. \[ R \overset{\text{bild}}{\rightarrow} W \]  
   b. \[ R \overset{\text{knapp}}{\rightarrow} W \]

The WM \( \text{-er} \) again encodes number: \( \emptyset \) (null) is the singular form, -er the plural form. This WM selects a class of roots to which bild belongs (which means that bild must be lexically specified as belonging to this class), while \( \text{-er} \) selects the class to which knapp belongs.\(^{16}\)

bild and knapp take the Case suffix -e when they occur as head of the nonhead of a complex compound.

(44) a. pärmbild på idé cover picture idea (“idea for a cover illustration”)

The analysis is the same as for the -a/-e nouns: cf. (30).

The rest of the nouns in (42) take Case even as nonheads of a simple compound.

(45) stol+rystg. minnes+förlust, frihets+långtan, ungdoms+kärlek chair back memory loss freedom longing youth love

uniform, frihet, and ungdom take Case because they are branching (bimorphemic). Stol takes Case as an idiosyncratic lexical property (subject to dialectal variation). Minne belongs to a particular declension class: neutres with e-stem. They take -n as a plural marker, and they all require Case-marking when occurring as non-head of a compound. I assume nouns of the minne class consist of two parts (two morphemes) like the -a/-e nouns.\(^{17}\)

However, unlike the -a/-e nouns the vocalic second morpheme of minne is not a NW, and does not encode number. Instead I assume it combines with a Root and forms a Stem (i.e. it heads a Stem).\(^{18}\)

That is to say, I assume word structure contains a category level between R and W, which I call Stem (broadly following traditional morphology in this respect). Like W it is a non-maximal category. Unlike W it does not contain number or Case. It follows that it requires Case when occurring as non-head of a compound.

(46)

\[ R \overset{\text{minne}}{\rightarrow} \overset{\text{e}}{\rightarrow} \begin{cases} \emptyset / n \end{cases} \]

\(^{17}\) One indication of this is their pitch accent: Like the -a/-e nouns, and like bimorphemic words in general (derived words and compounds) they take Accent 2, so-called. It should be noted, though, that there exist (apparently) monomorphemic, bisyllabic roots which take Accent 2, e.g. himmel “heaven”, spegel “mirror”.

\(^{18}\) In a sense it corresponds more closely to Harris’s (1991) Spanish word markers than the -a/-e suffixes do.
(0/fn is the number WM, -s is the Case-marked WM form.)

As for nouns like stol, which require Case when occurring as non-heads of compounds, I assume that they have no Root form. They are inherent Stems. The only way they can meet the maximality requirement on non-heads is by means of Case.

\[
\begin{array}{c}
\text{stol} \\
\text{\{0/-ar\}} \\
\text{-s}
\end{array}
\]

As for derivational (nominal) affixes like -het, -dom, -ning, -skap, -else, they have the following pertinent properties:

(a) they are subcategorized to take a complement with (often) specific categorial properties,
(b) they take overt Case when heading a nonhead of a compound (see (45)).
(c) they take specific plural forms (-het+er, -dom+ar, etc.)
(d) they do not take overt Case when they occur as non-head (complement) of a derivational affix.

Property (d) is exemplified in

\[
\begin{array}{c}
\text{frihet}+lig, *\text{frihets}+lig, \\
\text{frihet}+lig, *\text{frihets}+lig, \\
\text{freedom}+ly, (**anarchistic**) \\
\text{ungdom}+lig, *\text{ungdoms}+lig \\
\text{youth}+ly, (**youthful, young-looking**) \\
\end{array}
\]

I propose the derivational affixes are inherent Stems, subcategorized for a complement, and taking a WM specifier. The structure of e.g. frihet is

\[
\begin{array}{c}
\text{FRIHET} \\
\text{WIN} \\
\text{WM} \\
\text{stol} \\
\end{array}
\]

When occurring as nonhead of a compound, a derived word has to be Case-marked, i.e. the WM has to contain Case, to satisfy the maximality requirement. The reason why the derived word does not need to have overt Case when occurring as non-head of a derivational affix (ungdomlig) I assume is because the derivational affix licenses abstract (phonetically empty) Case on its complement. The structure of ungdomlig is

\[
\begin{array}{c}
\text{ung} \\
\text{dom} \\
\text{lig} \\
\text{0}
\end{array}
\]

The reason why a derivational affix licenses abstract Case when lexical items do not (ungdomlig, but ungdoms+kallek) is that derivational affixes are subcategorized for a complement. There is, informally speaking, a stronger government relation between a derivational affix and its complement than between a lexical head of a complement and its complement in a compounds. The weaker form of government is not sufficient to license phonetically empty Case.\footnote{There are at least two derivational affixes which do require Case on a branching complement: the adjectival affixes -aktiv and -näsig: flicka-aktiv "girlish", skol-flicka-aktiv, skol-näsig "school-like", flickaskola-näsig. Intuitively these are more "word-like" than affixes like -lig, -het, etc. (Though possibly this is just because they are bisyllabic) When comparing the adjectival affix -lig and in particular -aktiv, it may be significant that while the former strictly takes nominal roots as complement, -aktiv can take nominal, adjectival and even verbal roots as complement: blå-aktiv "blueish", lär-aktiv "apt to learn". Thus the selection relation (hence the government relation) between -lig and its complement is stricter than between -aktiv and its complement, which accords with the hypothesis that a stronger government}
6. Some Remarks on English

As mentioned, there is a broad consensus that both head and non-head in English compounds are W-level categories. *Schoolgirl* has the structure (51a) and *schoolgirl uniform* (51b) (the numbers indicate head vs. non-head).

(51) a. N2 N1 N1  
  school girl  
  b. N2 N1 N1  
  school girl uniform

According to our theory the non-head of a compound should either be a Root or be Case-marked (or be marked as a maximal category in some other way). English does not appear to support this theory. One might analyze *school* in *schoolgirl* as a Root, but there is ample evidence that the non-head of a (simple) compound can be an unmistakable word in English. In particular, the non-head of a compound can be plural (see Selkirk (1983: 52), Warren (1978)). Furthermore, this seems to be at least to some extent a productive mechanism in English word syntax.

(52) parks commissioner, programs coordinator, arms merchant, private schools catalogue

A plural non-head is out of the question in Swedish.

(53) a. *parker+forvaltare, park+forvaltare  
parks commissioner

b. *privatskolor+katalog, privatskola/privatskole+katalog  
private schools catalogue

I assume English nouns do not consist of a Root or Stem plus a WM. They are quite simply inherent W-level items. If so, the non-head

relation obviates the need for morphologically realized Case.

21 On the variation between -s and -e on the non-head form of *skola*, see (22).

of e.g. *schoolgirl* is indeed N. And of course, in *schoolgirl uniform* the non-head *schoolgirl* must be an N. Given the present theory we would expect non-heads in English compounds always to have Case-marking. But only some English compounds require genitive -s on the non-head (*driver's seat, men's magazine, pastor's cap*). Usually there is no kind of overt Case-marking of the non-head of compounds in English. Apparently, then, linear order is, indeed, sufficient to mark the first W in a compound as non-head in English.22 Given the theory of Case assumed here, we would say that N can license abstract Case on a nominal category which governs the left, in English. If abstract Case is represented as an empty head, the structure of e.g. *schoolgirl* will be (54):

(54)  

The phonetically empty Case on N2 is licensed by N3, apparently simply by virtue of N3 governing N2 leftwards.

The possibility of plural non-heads in English as opposed to Swedish can be derived from the absence vs. presence of WMs. If nouns in English are inherent Ws, the plural marker is simply adjoined to N.

(55)  

Since *schools* is the same category type as *school* there are no structural reasons not to allow a plural non-head in English, if this is somehow semantically motivated. In Swedish, on the other

---

22 As mentioned, the occurrence of genitive -s in English compounds is semantically conditioned, at least in part.

23 Note, however, that since English nouns, by assumption, have no WM, there is no position corresponding to the WM position in Swedish for word-internal Case in English. This may be part of an explanation why overt Case in compounds is such a restricted phenomenon in English.
hand, the number affix is a WM, i.e., it takes a Root or a Stem as complement. This difference is reflected in the fact that English has only one (productive) plural affix, while Swedish has a variety of plural forms. In Swedish but not in English there can be a selection relation between the WM and the Root. In English, where the plural morpheme is an adjunct, there is no selection. Consider the structure of a Swedish simple compound such as skolakatalog(er) "school catalogue(s)."

![Diagram 56](56)

\[ \begin{align*}
\text{R} & \quad \text{W} \\
\text{W} & \quad \text{W} \\
\text{skol} & \quad \text{katalog} \quad \emptyset/er
\end{align*} \]

The non-head is a Root. A plural non-head skolak would be a Word. As such it would require Case. But since word-internal Case is a WM, it is incompatible with the plural form. Hence both *skolakatalog(er) and *skolakerskatalog(er) are ill-formed.

In English, the structure of schools catalogue is

![Diagram 57](57)

\[ \begin{align*}
\text{N} & \quad \text{N} \\
\text{Case} & \quad \text{Case} \\
\text{school} & \quad \text{pl} \\
\text{L} & \quad \text{case} \\
\text{catalogue}
\end{align*} \]

Combining a non-head with Pl is no problem, from a structural point of view. Combining a plural N with (abstract) Case is also no problem from a structural point of view, since Case is also an adjunct to N.24

---

24 This raises the question why plural morphology inside compounds is so restricted in English. Why are "paperbacks" writer, "letterhead", "dressmaker", etc., ungrammatical? If I am right, the ungrammaticality is semantic in nature. I have no definite ideas of how the relevant restrictions should be expressed formally. It should be noted that plural nonheads in compounds occur frequently in German and Icelandic.

George Rebisch (p.c.) has suggested that the plural nonheads in (47) and similar examples are genitive plurals. This is compatible with my description of English compounds. However, if it is true that the distribution of genitive in English compounds is semantically conditioned, as

---

7. Speculations concerning verbal WMs

I have already suggested that L, or some subpart of L such as T is a verbal WM. It combines with the verbal root and forms a verb. In this case joining of Root and WM is effected by movement in the syntax (i.e., across phrasal boundaries), either V-floating to T or V-moving to T as in English. The relevant verbal WM may be T or it may be some other, more abstract functional category closer to VP (perhaps even internal to VP). If it is T, then English has at least one kind of WM (since English has T-morphology), and English verbs, just like Swedish, Icelandic, French, etc., verbs would consist of a Root (or Stem) plus a WM. If it is not T but some other functional category, then the description of English nouns above may be generalized to verbs.

A striking difference between English on the one hand and the other Germanic languages, and the Romance languages, on the other hand, is that English has no infinitive marker. In e.g., Swedish an infinitival verb is usually marked by a suffix -a. The structure of e.g., hänga "hang" (infinitive) is

![Diagram 58](58)

\[ \begin{align*}
\text{R} & \quad \text{N} \\
\text{hang} & \quad \text{a}
\end{align*} \]

I assume the infinitive marker is a WM, possibly a functionally empty WM signifying absence of Tense. The fact that the infinitive a-suffix disappears when the verb combines with tense-infections (häng-x "hang-3RS", häng-gets "hang") indicates that the inflection is syntactic, not lexical in the sense that number inflection (e.g., the WM a/- or on nouns is lexical in Swedish.25 There are various ways one may analyze the infinitive. It may be inserted as a default WM in the absence of finite inflection. It may be a functional head, distinct from finite T, in at least some

---

25 The number affix -a of e.g., flikas is retained regardless which affixes are added to the noun in the syntax: of. flickas (girl-CMF-POS3)
languages, located closer to the head of VP than finite T (to account for why the verb moves a shorter distance in infinitival than in finite sentences; cf. Pollock (1989)).

The absence of infinitival inflection may indicate that verbs are, indeed, inherent words in English. It is tempting to try to relate the absence of V-to-T movement in English (except for be and have) to this, assumed, property of English verbs (cf. Pollock (1989)). Assume that I-lowering is possible only if I can adjon to a a Word (i.e., a head). Then I-lowering to V is possible in English (creating a structure [y V I], corresponding to the structure of e.g. plural nouns) where verbs are inherent Word-level categories. In, say, French I-lowering will not be possible, since the lowered I would have to adjon to a verbal root, yielding the (plausibly) impossible configuration [y R INF]. Alternatively it would have to adjon to [y R-ON] where the infinitival WM has features which are incompatible with finite I.

The situation is complicated by the fact that Swedish and the other Mainland Scandinavian languages have I-lowering (see Holmberg and Platek (1991)), in spite of having infinitival inflection. Possibly the nature of the infinitival WM is such that it does not preclude I-adjunction to the infinitival [y R-ON].

8. Conclusions

I have argued that there is a class of functional heads which I call Word Markers. These heads may encode various features. I have argued that nominal WM’s in Swedish encode number or Case. Verbal WM’s may encode tense, and possibly other features such as mood. We may also have to recognize the existence of functionally empty WM’s. The primary function of WM’s is to combine with Roots (which are maximal, non-projecting categories), thus making it possible for the Roots to take part in various syntactic processes (e.g. selection of complements, theta role assignment, etc.). In metaphorical terms, I assume the WM provides a gate by which the features of the Root can get out of the maximal category and percolate to higher nodes.27

WM’s belong to the functional categories which most people would call inflectional (as opposed to derivational). A widely held view is that inflection is a syntactic phenomenon: it is something which happens to words when they enter syntax and are engaged as carriers of syntactic categories such as Tense, Agreement, Definiteness, Case, etc. I hope to have shown that some forms of inflection (those which involve WM’s) can be strictly word internal. The case in point was WM-marking in Swedish compounds. That is to say, inflection is something which takes place inside as well as outside W, i.e. by word internal processes (word internal movement, government, percolation) as well as by word external processes (movement, government, etc. across phrasal boundaries, as in the case of V-to-T).

From the point of view of the syntax-morphology relation, perhaps the most interesting finding discussed in this paper is that a typically syntactic category such as Case is involved in a productive, systematic, and structurally determined fashion within words, namely Swedish compounds. This paper can thus be seen as a contribution to the theory of syntax/morphology represented by Mark Baker, who has argued that Roots, traditionally a strictly morphological category, are involved in syntactic processes.

---

27 I have not provided the formal rules or rules required to achieve this. There are several possible ways it can be done. It remains to figure out which way has the most favourable consequences.
(undergoing syntactic movement), which entails that some forms of word formation take place in the syntax.

I have not discussed the possibility that Swedish compounding is a syntactic process (like Noun Incorporation), as opposed to a lexical process. I leave open the possibility that some cases of compounding are truly syntactic. However, clearly all compounding conforming to the rules and principles outlined in this paper cannot be syntactic, as the compounds exhibit varying degrees of lexicalization.

References


WORKING PAPERS
IN
SCANDINAVIAN
SYNTAX

Volume 49

Halldór Ármann Sigurðsson: The Case of Quirky Subjects 1
Anders Holmberg: Properties of Non-heads in Compounds: A Case Study 27
Gunlög Jósefsson: Object Shift and Weak Pronominals in Swedish 59
Peter Svenonius: The Extended Projection of N: Identifying the Head of the Noun Phrase 95

June 1992

Working Papers in Scandinavian Syntax
Department of Scandinavian Linguistics
Belgornäsbcken 14
S-223 62 Lund
Sweden