Proper-name nominal compounds in Swedish between syntax and lexicon

Maria Koptjevskaja-Tamm

This paper focuses on Swedish nominal compounds with a personal proper name as their first component (PropN-compounds), e.g. *en Mozart+sonat* 'a Mozart sonata' or *Palme+mord-et* 'the Palme murder' ('Palme+murder-the'), and aims at showing that such compounds are truly situated between lexicon and syntax. Syntactic considerations are relevant for PropN-compounds for at least the following reasons: a) since proper names themselves are often phrases rather than single words, the first part of PropN-compounds can be a phrase, including pre- and postposed attributes of varying complexity, coordination and even definite suffixed articles on the head noun; b) PropN-compounds are often involved in competition with syntactic coding strategies, and c) the proper-names in PropN-compounds are occasionally accessible for external anaphora. Lexical considerations are also relevant for PropN-compounds: a) PropN-compounding provides a systematic way of creating new lexical items, with many PropN-compounds fully conventionalized or on their way to full conventionalization; b) various structural properties of PropN-compounds (e.g. recursion, the use of proper names with attributes or of conjoined proper names as the non-head component) are sensitive to the "lexicalization degree" of the involved elements, where lexicalization primarily refers to their recognition as expressions for unitary concepts. It is suggested that some of the theoretical claims on the morphology of compounding are not particularly relevant for PropN-compounds, since their defining property is that they are based on personal proper names or on *what is conceived of as a personal proper name*. These expressions can have quite different structural properties ranging from single simplex words to phrases, but all these structures (including certain types of close appositional structures) are allowed to function as the first-most component in compounds *as long as they are recognized as a name for a unique referent* – a person or a group of persons. PropN-compounds in Swedish are simply right-headed nominal compounds in which the first slot is occupied by any expression that is recognized as a personal proper name (with the marginal exception of expressions with a preposed syntactic determiner).

0. Introduction

Swedish, like other Germanic languages, frequently resorts to right-headed nominal+nominal compounds, e.g. *en student+lägenhet* 'a student appartment' or *ett student+mord* 'a student's murder'.[1] In such compounds the first part typically serves to qualify or classify the second one, the whole compound functioning more or less as a

Rivista di Linguistica 21.1 (2009), pp. 119-148 *(ricevuto nel marzo 2009)*

classificational label (a particular type of apartment or murder). The first nominal is normally non-referential – thus, in the examples above there is no reference to a particular student. Not infrequently, however, the first part in nominal compounds is a personal proper name, e.g. *en Mozart+sonat* 'a Mozart sonata' or *Palme+mord-et* 'the Palme murder' ('Palme+murder-the'). This is all the more surprising given that personal proper names are normally both referential and highly specific; in fact, both *Mozart* and *Palme* seem to refer to two particular and very well known persons even when used as the first part in the mentioned compounds.

Proper-name compounds (both in Swedish and otherwise) have received very little attention in the literature, as far as I am aware of. "The Swedish Academy Grammar", SAG (vol. 2:43, 45, 136-137) mentions them and lists a couple of their properties. Kajanus (2005) provides a first survey of personal proper-name compounds in Swedish, partly inspired by Dahl's (2003) conference paper. Warren (1978:43-44) devotes a couple of pages to the English variant of proper-name compounds.

There are several reasons for why proper-name compounds are relevant for the general theme of this volume, i.e. for the issues of lexicon(-morphology)-syntax interface in compounding. First of all, proper-name compounds are sometimes attested in examples like (1) in which the anaphoric pronoun 'his' refers to the first part of the compound 'Mozart year'.

(1) År 2006 är åter ett **Mozart+år**, nämligen **hans** 250:de födelsedag.
year 2006 is again a:N Mozart+year namely his 250:th birthday
'2006 is once again a **Mozart** year, more precisely **his** 250th birthday.'
(Marcus Boldemann in *Dagens Nyheter, Kultur*, January 22th 2005)

Such cases seem to contradict the Lexical Integrity Hypothesis, according to which "[t]he syntax neither manipulates not has access to the internal structure of words" (Anderson 1992:84). There is a long tradition of either dismissing authentic examples like (1) as "ungrammatical" or offering pragmatically anchored explanations for their occurrence (cf. Ward *et al.* 1991 for one exception).

Problems raised by examples like (1) will not be discussed in this paper. Here we will deal instead with the other, more structural aspects of proper-name compounds that pose problems for a strict separation between syntax and morphology or lexicon.

a) Proper names themselves are often phrases rather than single words. This does not prevent them from functioning as the first component of proper-name compounds, cf. ex. (2), and providing thus systematic counterexamples to the No Phrase Constraint whereby no phrases may appear within complex words (Botha 1981).

(2) [Emil i Lönneberga]+bus; e-n [Drottning Silvia]+brytning
Emil in Lönneberga + mischief.PL; a-COM Queen Silvia+accent
'Emil-in-Lönneberga mischief; a Queen Silvia accent'

b) Moreover, since proper names, in contrast to common names, have a unique reference and are therefore (inherently) definite, proper-name compounds seem to be problematic even for a more specific constraint on phrasal compounds, the "No DP!"-constraint, whereby DPs (or "full" NPs, i.e. determined NPs) are not allowed as non-heads in compounds (see the discussion in Meibauer 2007:236-238).

c) Finally, proper-name compounds (as compounds in general) are often involved in competition with syntactic coding strategies, where compounding is favoured by such factors as the degree of lexicalization and frequency of use.

In this paper I will limit myself to proper-name compounds with personal proper names as their first component, which will be referred to by the expression "proper-name (nominal) compound" and its abbreviation "PropN-compound", when not specified otherwise. The structure of the paper will be as follows. The first three sections provide the general information on proper names vs. common names in Swedish (Section 1), on Swedish nominal compounding (Section 2) and on semantics of proper-name nominal compounds (Section 3). Section 4 deals with the structure of nominal compounds with one-word proper names – both simple compounds (4.1.) and recursive compounds (4.2.), while Section 5 considers compounds with heavy (i.e., phrasal) proper names and their relation to lexicalized phrases and phrasal compounds. Section 6, finally, summarizes the main findings in the paper.

The paper is primarily based on authentic examples coming from the PAROLE corpus (http://spraakbanken.gu.se/parole/), largely due to Kajanus (2005), and from the Internet, occasionally complemented by solicited examples that have been provided (and checked) by several native speakers of Swedish.

1. Preliminaries: proper names vs. common names

In a recent important contribution Van Langendonck (2007:87) suggests a following definition for proper names that combines their pragmatic, semantic and syntactic properties:

> A proper name is a noun that denotes a unique entity at the level of established linguistic convention to make it psychosocially salient within a given basic level category [pragmatic]. The meaning of the name, if any, does not (or not any longer) determine its denotation [semantic]. An important formal reflex of this pragmatic-semantic characterization of proper names is their ability to appear in such close appositional constructions as *the poet Burns*, *Fido the dog*, *the River Thames*, or *the City of London* [syntactic].

The status and linguistic properties of proper names have been in the centre for much research coming from different traditions (cf. also Anderson 2004). Cross-linguistically proper names vary in how close they are to nouns in their grammatical behaviour. In addition (and partly independently of this variation) different linguists vary in whether they treat proper names in a particular language as a separate word class or as a subclass of nouns, highlighting or, conversely, downplaying the differences between the two. The most recent and complete reference grammar of Swedish, "The Swedish Academy Grammar" (SAG vol. 2:9,116-149), has a separate word class of proper names as opposed to that of nouns, in spite of the clear similarities between the two. The distinction is based on three types of criteria: a) morphologically, proper names do not inflect for number and definiteness; b) semantically, proper names refer to a unique entity, but lack descriptive meaning, and c) syntactically, proper names are heads in noun phrases, but show a restricted combinability with determiners and attributes as compared to common nouns. SAG's treatment of proper names as a distinct word class is fairly unusual within the Swedish grammatical tradition. SAG (ibid.:9) itself acknowledges that both semantically and syntactically proper names are very close to nouns in the definite form (and to the third person personal pronouns); there are also various processes whereby proper names become common nouns, and vice versa.

One of the shared grammatical properties for proper names and common nouns is their ability to function as the first part in nominal compounds.

2. Swedish nominal compounds vs. syntactic adnominal modification

Most nominal compounds in Swedish, i.e. compounds based on nouns as head, are right-headed, subordinate and endocentric. Swedish compounds, both nominal and others, are normally easily recognizable as *words* rather than phrases on phonetic and morphosyntactic grounds. Thus, they are pronounced with the so-called "compounding accent" (Sw. *sammanfattningsaccent*), a variant of the Swedish lexical accent 2, cf. *²sko+₁kräm* 'shoe polish' vs. *²skola* 'school'.[2] The orthographic convention recognizes compounds as words; there is, however, a growing tendency to spell parts of compounds as different words. For all morphological and syntactic purposes, compounds behave like simplex nominals, i.e. like their last part. Among other things, nominal compounds show the usual opposition in definiteness in combining with preposed indefinite and suffixed definite articles (which agree with their head in gender and number), cf. *en student+lägenhet* 'a:COM student+apartment' vs. *student+lägenhet-en* 'student+apartment-DEF.COM', *en blå+klocka* 'a:COM blue+ bell' vs. *blå+klocka-n* 'blue+bell-DEF.COM', *en Mozart+sonat* 'a:COM Mozart+sonata' vs. *Mozart+sonat-en* 'Mozart+sonata-DEF.COM'.[3] The first part of nominal compounds can be common nouns, proper names, adjectives, verbs, numerals, pronouns, prepositions, adverbs and, occasionally, phrases.

Nominal compounds, which provide a morphological strategy for coding modifier–head relations within a word, co-exist and sometimes compete with several other, syntactic constructions used for adnominal modification, with the two most important for our purposes being as follows:

a) the "standard possessive noun phrase", in which the possessor (modifier), normally marked with *-s*, precedes the possessee (head, e.g. *Mozart-s sonat* 'Mozart-GEN sonat', *Lisa-s kind* 'Lisa-GEN cheek', *Peter-s syster* 'Peter-GEN sister'; and
b) constructions in which the head nominal is modified by a postposed prepositional phrase, e.g. *en sonart **av** Mozart* 'a sonata **of** Mozart', *kind-en **på** Lisa* 'lit. cheek-the **on** Lisa' ('Lisa's cheek'), *syster-n **till** Peter* 'lit. sister-the **to** Peter' ('Peter's sister').

Swedish possessive NPs with *-s*-possessors are, on the whole, fairly reminiscent of their English counterparts. Importantly, *-s*-possessors are (normally) incompatible with any articles pertaining to the possessive NP (or to its head), e.g. *en sonat* 'a sonata' vs. *sonat-en* 'sonata-the', but only *Mozart-s sonat* 'Mozart-GEN sonata' and not **en Mozart-s sonat* 'a Mozart-GEN sonata' or **Mozart-s sonat-en* 'Mozart-GEN

sonata-the'. This is one of the properties that distinguish possessive NPs from compounds (cf. *en Mozart+sonat* and *Mozart+sonat-en* given above).[4] In contrast to English, Swedish has no default preposition for postnominal modification comparable to *of*, but uses different spatial prepositions in many cases where English will use one and the same (cf. the examples above).

In addition to compounding and productive syntactic constructions, Swedish has a class of in-between expressions, so called *lexicalized phrases*, or *lexphrases*. These will be considered in Section 5.

3. Semantics of proper-name nominal compounds: general

Semantics of compounds has for a long time been a popular research topic for linguists coming from different research traditions. Similarly to compounds in general, PropN compounds show a whole spectrum of possibilities with respect to the relation between their meaning and the meanings of their components. In many cases, the meaning of a PropN-compound can easily be derived from the meanings of its parts, as the case often is for compounds with strongly relational heads or, more generally, with heads whose semantic frame is strongly associated with human participants. For instance, in *Kennedybegravningen* 'the Kennedy burial' the proper name refers to the Patient, or Undergoer of the corresponding deverbal noun. The first part in PropN-compounds that are based on nouns denoting musical pieces refers almost exclusively to their composers, e.g. *en Mozartsonat* 'a Mozart sonata', while *film* normally combines with the most prominent persons involved in its production – directors (*en Fellinifilm* 'a Fellini film'), film stars (*en Mastroiannifilm* 'a Mastroianni film') or protagonists (*en Bondfilm* 'a Bond film').

At the other end of this spectrum are the cases where the relation between the two parts of a PropN-compound is much less obvious or even completely idiosyncratic, e.g. *en Einstein-schäfer* 'an Einstein sheep dog' (a dog that is as clever as Einstein, *Göteborgsposten* 2001.51:0726) or *ett Barbiedagis* 'a Barbie kindergarten' (a kindergarten where the girls play with Barbie dolls and the boys play "real boys" games, *Göteborgsposten* 2002.4:0317). In addition, there are, of course, erstwhile PropN-compounds in which the first part is no longer conceived of as a proper name, e.g. *graham-s+mjöl* 'whole wheat flour' ('Graham-LNK+flour', named after an American 19th century's vegetarian S. Graham) or *dieselmotor* 'diesel motor'.

All these cases find more or less close parallels in other nominal compounds. In addition, there are commemorative PropN-compounds, "X named after person Y" without there being any "inherent" connection between the two, for which it is more difficult to find correspondences among ComN-compounds (Warren 1978:218, Kajanus 2005:44-45), e.g. *en Birgittabakelse* 'a Birgitta cake' (a cake "invented" for the 700 year anniversary of Holy Birgitta) or *Victoriaarenan* 'the Victoria arena' (named after the Swedish Crown Princess).

It is generally held that "[a] noun–noun compound such as *raincloud* is an ideal construction for providing a subcategorization" (Bauer 2006:722); in other words, a compound denotes a subtype of the kind of entity denoted by its head and behaves as a common noun (i.e., it inflects for number and attaches indefinite or definite articles following the pattern of its head). Also many PropN-compounds denote a class of possible referents, a subtype of the kind of entity denoted by their head, and function therefore as common names. A Montessori school is a type of school practicing the education method developed by Maria Montessori, a Barbie dress is a type of dress associated with Barbie fashion dolls, and in each of these cases, the class of possible referents is open and potentially endless.

On the other hand, there are many PropN-compounds that have unique reference and function as proper names. For instance, *Kristofferskolan* 'the Christopher school' is the name of a particular school in Stockholm, *Strindbergsgatan* 'the Strindberg street' is a particular street in the Eastern part of central Stockholm, and *Berwaldhallen* ('the Berwald hall', named after the Swedish composer Franz Berwald) refers to a particular concert hall. PropN-compounds constitute a regular means for coining names of particular buildings and institutions (schools, churches, theatres, hospitals), streets, etc. They differ from the most prototypical proper names in being obligatorily marked with the singular definite article, but, on the other hand, this is the only form in which they occur. Contrary to common names and similarly to other proper names they do not inflect for number and do not show contrasts in (in)definiteness.

Finally, there are compounds like *en Mozartopera* 'a Mozart opera' or *en Fellini film* 'a Fellini film'. Each of them denotes a class of possible referents, which, however, contain just a few members and are pretty much closed, since both Mozart and Fellini are dead and can therefore not create any new operas or films.

Proper names in PropN-compounds most often refer to persons, including fictive personalities, who are normally well known within the relevant speech community, ranging from fairly famous personali-

ties to the members of a family, the colleagues at the same working place etc. The "scope of popularity" often but not necessarily, correlates with the part of the name involved. Thus, *en Picassotavla* 'a Picasso painting', based on the last name, denotes a picture painted by the famous painter, while *en Kattitavla* 'a Katti painting', based on a Christian name, denotes a picture painted by my youngest daughter. On the other hand, many famous persons are commonly known by their first names, cf. such compounds as *en Britneylåt* 'a Britney (Spears) song' and *ett Zlatanmål* 'a Zlatan (Ibrahimovic) goal'.

PropN-compounds show a whole scale of entrenchment differences, with those completely conventionalized on the one pole and those created on the spot and being completely context dependent, on the other.

4. The structure of nominal compounds with one-word proper names

This section will be devoted to the structure of nominal compounds with one-word personal proper names – a first name (*Anna, Hassan*), a last name (*Levinsky, Palme*), a nickname (*Carro* for Caroline) or a fictional name (*Mumin, Tintin*). We will consider simple compounds in Section 4.1. and recursive compounds in Section 4.2.

4.1. The basic structure of Swedish proper-name and common-noun nominal compounds: simple compounds

In this section we will look at the basic and structurally simplest common-noun and proper-name nominal compounds – i.e., those mainly involving "simple" stems.

Common-name nominal compounds (ComN-compounds) fall into two major morphological types (the classification is due to Josefsson 1998). Non-mediated compounds consist of two nominal stems without any element between them and have two further subtypes depending on the form of their first part. Most nouns occur in their basic form when functioning as the first part in non-mediated compounds (3a), while a number of nouns have a special compound form. In particular, nouns ending in short unstressed vowel have sometimes their final vowel either truncated compared to their basic form (3b) or replaced with *-e-, -u-* or *-o-* (3c). Mediated compounds involve a linking element, typically *-s-*, in-between their two parts, akin to the German *Fugenmorphem*, the Greek *-o-*, the Russian *-o-/-e-* and the various Lithuanian intermorphemes. The linking marker *-s-* in Swedish (4) looks like the genitive marker and has developed

historically from a genitive inflection – a situation familiar from quite a few other languages.[5] Some nouns alternate between different forms as the first part in compounds (cf. *kyrktak* 'church roof' / *kyrkomusik* 'church music'). There are various factors governing the choice among the morphological types of compounds, but these remain tendencies rather than rules (in particular, the choice between non-mediated compounding and the linker *-s-*) (cf. SAG vol. 2:50-57).

(3) Non-mediated ComN-compounds
   a. basic form: *bil* 'car' + *tak* 'roof' => *biltak* 'car bonnet'
   b. truncated form: *kyrk**a*** 'church' + *tak* 'roof' => *kyrktak* 'church roof'
   c. vowel replacement:
      *kyrk**a*** 'church' + *musik* 'music' => *kyrk**o**musik* 'church music'
      *gat**a*** 'street' + *korsning* 'junction' => *gat**u**korsning* 'street junction'

(4) Mediated ComN-compounds
   stol 'chair' + *rygg* 'back' => *stol**s**rygg* 'chair back'

PropN-compounds have, on the whole, the same prosodical, inflectional and distributional properties as ComN-compounds. However, they show a restricted set of morphological types, as compared to ComN-compounds, occurring only in non-mediated compounds without truncation and, much less frequently, in mediated compounds with the linking *-s-*.

First, in contrast to common nouns, proper names get never truncated as the first part in compounds (SAG vol. 2:137). This is nicely demonstrated by the behaviour of two homonyms *lina* 'rope' vs. *Lina* 'Lina, a female name' – the first one appears in the truncated form in compounds, e.g. *lin+bana* 'a rope way' ('rope+road'), whereas the second one remains intact, e.g. *Lina+skämt* 'Lina joke'. The vowel replacement is likewise excluded from PropN-compounds. The marginal exceptions here are the old Latin genitive forms *Jesu-*, *Petri-* etc. that sometimes occur in PropN-compounds.

Moving to mediated PropN-compounds, in Kajanus' (2005) data, *-s-* seems to be restricted to very few proper names – mainly the last (family) names ending in *-berg*, *-borg*, *-beck* (erstwhile compounds, see Section 4.2. on the use of *-s-* in recursive compounds), certain names ending in *-man*, and the names of Swedish kings, e.g. *Gustav Adolf-s+dag-en* 'the Gustav Adolphus day' (Gustav Adolf-LNK+day-DEF.COM). The choice is not quite clear and the practice varies. Thus, of the 22 different compounds (i.e. compound types) with *Strindberg* as the first part in the PAROLE corpus (http://spraakbanken.gu.se/parole/),

15 have *-s-* and 7 lack it; the difference becomes even bigger when the tokens are considered, with 41 with the linking *-s-* and 8 without it (partly due to the highly frequent compound *Strindbergsmuseet* 'the Strindberg museum'). Some *Strindberg*-compounds are particularly prone to alternation: for instance, for 'a Strindberg play' Google has 107 hits with the non-mediated compound *en Strindbergpjäs* and 99 with the linker *-s-*, *en Strindbergspjäs* (checked on December 10th 2008). Once again, as already emphasized in Section 2, compatibility with articles distinguishes compounds from possessive NPs, with the compound *en Strindbergspjäs* being opposed to the regular possessive NP *Strindbergs pjäs*. SAG (ibid.:137) mentions also that *-s-* sometimes occurs in strongly lexicalized compounds, e.g. *Staffan-s+visan* 'the Stephan song' – a traditional Swedish Christmas carol.

Summarizing, compounding based on proper names tends to operate with a subset of the simpler morphological processes among those that are involved in ComN-compounding: the form of proper names is not influenced by compounding (they remain "morphologically inert"), and no extra morphological material is normally involved for combining the two parts.

4.2. Recursion

Either part of a compound can itself be a compound. Recursion within ComN-compounds can apply both to the first (5a-b) and to the second (5c) part. If the first part is complex, the linking *-s-* is almost obligatory.

(5) a. en [[barn+bok]-s+[klubb]]
a:COM child+book-LNK+club
'a club for children's books'

b. [svart+[peppar]]+[sås]
black+pepper+sauce
'black pepper sauce'

c. en [[barn]+[bok+klubb]]
a:COM child + book+club
'a book club for children'

Josefsson (1998:59) writes that "[t]he formation of three-segmented compounds is highly productive in Swedish, and the results are by no means marked. It is even possible to coin compounds of the proposed type [i.e., with recursive left adjunction, MKT] with four segments". Recursion in ComN-compounds is, however, far from

free, as I will briefly demonstrate in ex. (6-7) with compounds based on very common concrete nouns – *duk* 'cloth', *säck* 'bag', *mugg* 'mug', *tröja* 'sweater', and *klänning* 'dress'. The first component in the conventionalized and very frequent ComN-compounds in (6a) specifies a particular intended purpose, or use of the category denoted by the head, and the whole compound therefore denotes a particular subtype of this category (e.g., 'a coffee mug' ≈ 'a mug intended for coffee'). In (6b) the same heads are combined with left-hand components expressing material modification (e.g., *en bomullströja* 'a cotton sweater'). For material modification in general ComN-compounding is the most frequently used coding strategy, competing with regular syntactic attribution by postposed prepositional phrases (e.g., *en tröja i/av bomull* 'lit. a sweater in/of cotton'). The examples below are meant as illustrations rather than as a representative sample.

(6) a. hand+duk, rygg+säck, kaffe+mugg, dam+tröja, afton + klänning
hand+cloth, back+bag, coffee+mug, lady+sweater, evening + dress
'towel, rucksack, coffee mug, ladies' sweater, evening dress'

b. linne+duk, läder+säck, plast+mugg, bomull-s + tröja,
linen+cloth, leather+bag, plast + mug, cotton-LNK + sweater
silke-s+klänning
silk-LNK+dress
'linen cloth, leather bag, plastic cup, cotton sweater, silk dress'

The following examples show what happens if we try to combine the same heads with both purpose and material components within one and the same compound in an attempt to achieve recursive compounding.

(7) a. linne +[hand+duk], läder + [rygg+säck]
linen + hand+cloth, leather + back + bag
'linen towel, leather rucksack'

b. plast+[kaffe+mugg]
plast + coffee + mug
'plastic coffee cup'

c. ???bomull-s+[dam+tröja], ???dam+[bomull-s+tröja];
cotton-LNK+lady+sweater, lady + cotton-LNK+sweater
???silke-s+[afton+klänning], ???afton+[silke-s+klänning]
silk-LNK+evening+dress evening+silk-LNK+dress
'cotton ladies' sweater; silk evening dress'

As is clear from these examples, the felicity of recursion differs significantly across these cases (7a-c),[6] even though the intended compounds all denote very frequent objects. The reasons for these differences definitely deserve a study of their own (to be based on much more authentic data and on psycholinguistic experiments), but the examples above are suggestive of some factors at work here. Some of the best cases of what looks like recursive compounding involve either the compound stems that are very highly lexicalized (ex. (7a)), or conventionalized labels (ex. (7b)). To start with the former, the most felicitous results obtain for the most lexicalized compounds in the list, *hand+duk* 'towel' and *rygg+säck* 'ruggsack' (for *handduk* it can even be argued that it is a "frozen compound", i.e. synchronically perceived as simplex, since its meaning is no longer derivable from the meaning of its components). These can combine with diverse left-hand components denoting material, e.g. *bomull-s+handduk* 'cotton towel', *frotté+handduk* 'Turkish terry towel', or *nylon+ryggsäck* 'nylon+rucksack', *plast+ryggsäck* 'plast+rucksack', etc. An additional potential factor at work here is phonology of the compounded stems – both *handduk* and *ryggsäck* involve one-syllable long stems. Moving on, *plastkaffemugg* 'plastic coffee cup' is a conventionalized label for an everyday object (which also has a paper variant, *paperskaffemugg*), quite different from a normal coffee mug.

While it is not quite clear why none of the compounds in (7c) is felicitous, they seem to represent the more frequent case. In other words, when speakers need to express both the intended use of en entity and its material, they tend to avoid recursive compounds and resort to the alternative strategies for coding either the purpose or the material by a postposed prepositional phrase, e.g., *en dam+tröja i bomull* 'a ladies' sweater in cotton', *en bomull-s+tröja för damer* 'a cotton sweater for ladies'. Note that this alternative strategy is otherwise normally used only when the material or purpose expressions themselves are complex (i.e., when they contain an additional modifier), e.g., *en silver+ring* 'a:COM silver+ring' vs. *en ring i äkta silver* 'a:COM ring in genuine silver', en *dam+tröja* 'a lady+sweater' vs. *en tröja för äldre damer* 'a sweater for elder ladies'.

Recursion in PropN-compounds is also attested, both in their first (8a-b) and second (8c-d) parts. Strictly speaking, examples like (8a-b) do not demonstrate recursion within PropN-compounds, since their first part, while being a PropN-compound, is no longer a proper name,[7] and a similar terminological objection is valid for (8c-d) where the second part is a compound, but not a PropN-compound. The example below show rather than PropN-compounds can be involved in com-

pounds with complex stems – either in compounds with the structure [[Prop+Com]+Com], as in (8a-b), or [Prop+[Com+Com]], as in (8c-d).

(8) a. en [[Bond+film]-s + producent]
a:com Bond+ film-LNK + producer
'a Bond film producer' (i.e., 'a producer of Bond films')

b. [[Palme+mord]-s + utredning]-en
Palme+murder-LNK+investigation-DEF.SG.COM
'the Palme murder investigation' (i.e., 'the investigation of the Palme murder')

c. Armani+[sol+[glas+ögon]]
Armani + sun + glass+eyes
'Armani sun glasses'

d. en [Barbie+[rygg+säck]]
a:COM Barbie + back+bag
'a Barbie rucksack'

Now, the examples (8c-d) involve again highly lexicalized (and even frozen, in the case of 'eyeglasses') complex stems, the same that would also allow further compounding with material-denoting stems, e.g. *plast*+[*sol*+[*glas*+*ögon*]] 'plastic sun-glasses' and *plast*+[*rygg*+*säck*] 'plastic rucksack' (cf. above). PropN-compounds with a compounded second part are otherwise very much avoided. This is illustrated in ex. (9b), in which the stem denoting a musical piece cannot build a compound that would simultaneously specify both its composer and the intended musical instrument, even though simple compounds specifying either of those are perfectly fine (9a). For this purpose one of them (preferably the composer) would have to be coded by one of the alternative coding strategies – as a regular possessor in the standard possessive NP, or in a postnominal prepositional phrase (9c).

(9) a. en Mozart+sonat; en violin+sonat
'a Mozart sonata'; 'a violin sonata'

b. ???en Mozart+violin+sonat
'a Mozart violin sonata'

c. en violin+sonat av Mozart / Mozart-s violin+sonat
a:COM violin + sonata of Mozart / Mozart-GEN violin+sonata
'a violin sonata of Mozart / Mozart's violin sonata'

Judging from my data, PropN-compounds with a complex second part are extremely rare. At this stage it is difficult to estimate whether there are any significant differences in the acceptability of recursive PropN-compounding as compared to ComN-compounding, but my impression is that recursive PropN-compounds are proportionally less frequent than recursive ComN-compounds. As we have seen, acceptability of recursive compounding seems to be largely governed by such factors as conventionalization and, ultimately, frequency of use. Since PropN-compounds are, in general, significantly less common than ComN-compounds, it remains to be investigated whether the rarity of such complex compounds merely reflects the general differences in the frequencies of the PropN- and ComN-compounds or whether there are additional restrictions on the complexity of PropN-compounds, as compared to ComN-compounds. Judging by the Google searches in (10) (carried out on December 12th 2008), even though highly frequent compounds can be used as the second component in PropN-compounds, such complex compounds are still grossly dispreferred compared to the alternative phrasal strategy.

(10) a. Nobel+pris-et – 700 000 Google hits
Nobel + prize-DEF.SG.N
'the Nobel Prize'

b. Fred-s + pris-et – 59 800 Google hits
peace-LNK+prize-DEF.SG.N
'the Peace Prize'

c. Nobel+fred-s + pris-et – 58 Google hits
Nobel + peace-LNK + prize-DEF.SG.N

d. Nobel-s fred-s + pris – 226 000 Google hits
Nobel-GEN peace-LNK + prize
'the Nobel Peace Prize'

e. *Fred-s+nobel+pris-et
peace-LNK+Nobel +prize-DEF.SG.N

f. *Fred-s Nobel+pris
peace-GEN Nobel +prize

Interestingly, as witnessed by the English glosses in (9b) and (10d), the English correspondences to Swedish PropN-compounds seem to tolerate recursion to a higher degree.

Summarizing, recursion in compounds, although attested, is quite restricted. It is unclear whether the rarity of PropN-compounds with a complex second component is mainly related to such factors as conventionalization and frequency of use, as the case seems to be for recursion in ComN-compounds, or whether PropN-compounds are subject to additional restrictions on their complexity, as compared to ComN-compounds.

5. Lexicalized phrases, compounds with heavy proper names vs. phrasal compounds

5.1. Lexicalized phrases (lexphrases) and phrasal compounds in Swedish

In addition to compounding and well-behaved productive syntactic phrases, Swedish has a large number of so-called "lexicalized phrases" (*lexikaliserade fraser*, or "lexphrases", Anward & Linell 1975-1976, also called *lexikaliserade ordgrupper* 'lexicalized word groups' in SAG vol. 1:195) that are clearly intermediate between those. In a nutshell, lexicalized phrases are combinations of words with a special connective intonation pattern, whereby all the stressed syllables in the group, except for the last one, get deaccentuated, i.e. they either lose their stress completely or retain it in a weaker variant (in the examples this is symbolized by $_{(1)}$).[8] In the Swedish grammatical tradition this prosodical pattern is known as *sammanfattningsaccent* ('summarizing accentuation'), *idiomatiseringsaccent* ('idiomatizing accentuation') and *ordgruppsbetoning* ('word group stress'). In the absence of any accepted English terminology I will here adhere to the expression 'unit accentuation' used by Dahl (2004:252-259).

A prototypical lexicalized phrase has a syntactic structure that is more or less close to well-behaved syntactic phrases, consists of words that at least partly retain their inflection, but has a unitary meaning of its own, which can – but does not have to – be related to the meanings of its components, cf. examples in (11). Lexicalized phrases vary in their "superficial" similarity to normal syntactic phrases, in how much their components inflect, in how closely their meaning is related to those of their components, and in their degree of conventionalization.

(11) Lexicalized phrases (lexphrases)
a. (1)sun-t [1i]för[1i]nuft
sound-INDEF.SG.N reason
'common sense'

b. (1)lång [2i]klänning
'long dress' (the female correspondence to a tail-coat, i.e. an evening dress rather than any dress that happens to be long)

c. (1)frisk-a [1i]luft-en
fresh-DEF air-DEF.SG.COM
'out of doors' (used almost exclusively in the combination with *ute* 'out', *ute i friska luften*)

d. (1)röd-a [1i]hund
red-DEF dog
'Rubella, German measles'

(12) Regular definite NPs with adjectival attributes
de-n frisk-a luft-en
the-SG.COM fresh-DEF air-DEF.SG.COM
'the fresh air'

As the examples above show, lexphrases can contain both indefinite and definite adjectives (cf. *sunt, lång* vs. *friska, röda*), with the latter combining with either unmarked nouns or nouns with the definite suffix (cf. *hund* vs. *luften*); the pattern is, however, normally strictly fixed in each particular case. What seems to be avoided is the normal definite adjectival pattern with the preposed free determiner, the adjective in the definite form and the head noun with the suffixed article. Compare the lexphrase in (11c) with the regular phrase involving the same lexical items in (12).

Lexicalized phrases sometimes occur as the first part in a compound, these are called "word group compounds" (*ordgrupssammansättningar*) in the Swedish linguistic tradition (e.g. SAG vol. 2:43-44). For instance, whereas adjectives normally occur in the uninflecting form as the first part in nominal compounds, as in ex. (5b) (repeated here as (14a), in ex. (13) the adjective 'black' is inflected for definiteness within the compound, in accordance with its use within the corresponding lexphrase. However, there are restrictions on inflectional markers even within phrasal compounds: thus, the suffixed definiteness marker on the head (*börs)*, which is obligatory in the free standing lexicalized phrase, is likewise obligatorily dropped within

the compound in (13). This issue will be particularly relevant in the discussion of PropN-compounds in Section 5.3.

(13) svart-a börs-en => svart-a + börs + affär-er
black-DEF exchange-DEF.COM black-DEF+exchange+transaction-PL
'black market' => 'black-market transactions'

Phrasal compounds in Swedish "inherit" their prosody from lexphrases, which distinguishes them from recursive compounds, i.e. compounds with the compounded first part considered in Section 4.2. In both cases compounds are pronounced with the two-peak compounding accent (lexical accent 2), with the last stressed syllable of the compound constituting its second peak. The difference lies in the position of the first peak. In the Swedish spoken in Sweden (as opposed to the Finnish varieties), recursive compounds have basically the same overall prosody as simple compounds, with the first stressed syllable in the word being most prominent. In phrasal compounds, on the other hand, it is the last stressed syllable of the phrasal part that constitutes the first tonal peak, while all the preceding syllables are deaccentuated. This difference is, in fact, a good criterion for distinguishing recursive compounds from phrasal compounds even in the absence of other morphosyntactic clues, as in ex. (14a) vs. (14b), with the same unmarked adjective *svart* 'black' involved in both cases: in the recursive compound in (14a) it is stressed (*ˈsvart*), while in the phrasal compound it is deaccentuated (*$_{(ˈ)}$svart*).

(14) a. ^{2}[^{ˈ}svart+$_{(1)}$peppar]+$_{1}$sås
black + pepper + sauce
'black pepper sauce' (<^{2ˈ}*svart$_{ˈ}$ peppar* 'black pepper')

b. [$_{(1)}$svart+^{2ˈ}vinbär-s]+$_{1}$saft
black + currant-LNK+juice
'black-currant juice
(<$_{(1)}$*svart* ^{ˈ}*vinbär* 'black currant')

Lexphrases are, of course, related to such phenomena as formulaic expressions (e.g. Wray 2002) and idioms that have received a lot of attention in the recent years' theoretical discussions. The common theme in all these complex expressions is that they are "listed", or stored as a whole in the speaker's mental lexicon rather than being generated on the spot by productive syntactic rules. What is particularly interesting about Swedish lexicalized phrases is their distinct

prosody that can be productively used for creating new temporary expressions, something that seems to contradict the usual interpretations of lexicalization in any of its common uses (for an overview cf. Bakken 2006 and Svanlund 2002). Anward & Linell (1975-1976) talk about a "temporary lexicalization" that may, e.g., last for the rest of a conversation, an idea further developed by Wälchli (2005:114-116). "Local lexicalization" is also possible: in my family, for instance, we use the expression $_{(1)}$*fira* ˈ*lördag* 'celebrate Saturday' for referring to a particular kind of dinner that we normally eat on Saturdays (without any religious connotations at all) and that occasionally take place on other days of a week.

Lexphrases and the phenomena of temporal and local lexicalization are of primary relevance for the discussion of compounds based on complex proper names in the next section.

5.2. PropN-compounds with heavy proper names

We have so far considered PropN-compounds based on simple names (first names, like *Birgitta* or *Hassan*, last names like *Mozart*, nicknames like *Carro* for *Caroline*, and fictional names like *Barbie*). Quite often, however, the first part of a PropN-compound consists of what looks like a combination of several words (all referring to one and the same person). At least the following structural types of complex personal proper names occur in Swedish:

a) complex first names, e.g. *Carl Gustav*, *Gustav Adolf*
b) a name with a preposed attribute (epithet): *duktig-a Annika* 'efficient-DEF Annika', *Svart-e Petter* 'Black-DEF.M Petter' (approximately 'Old Maid' in card-game), *Magr-e Memet* 'Thin-DEF.M Memet', *Skön-a Helena* 'Beautiful-DEF Helena' ('la belle Hélène, Fair Helen');
c) a name with a postposed identificatory epithet, e.g. a royal name specified for the succession number, like *Karl den tolft-e* 'Charles the.COM.SG. twelfth-DEF' ('Charles the Twelfth'), or for other qualifications, like *Peter den stor-e* 'Peter the.COM.SG great-DEF.M. ('Peter the Great'), but also *Karlsson på tak-et* 'Karlsson on roof-DEF.SG.N' ('Karlsson on the Roof', the protagonist in a few books by Astrid Lindgren) and even *Jack the Ripper*.
   The two following classes are normally treated in the literature as close appositional structures (i.e. as syntactic constructions) rather than as "single" proper names (e.g., SAG vol. 2:134, 146) or Van Langendonck 2007:126-129):
d) a first name and a last name: *Astrid Lindgren*, *Selma Lagerlöf*, *Saddam Hussein*, *Florence Nightingale*, and

e) some very tight and institutionalized combinations of a title and a name, like *Drottning Silvia* 'Queen Silvia' or *Mr. Bean*.

In all these cases the complex proper name is pronounced with the prosody typical of lexphrases (cf. the preceding section), i.e. with the 'unit accentuation' whereby all the stressed syllables in the group, except for the last one, are deaccentuated.[9] All these types of complex proper names are attested in PropN-compounds and will be discussed below in a somewhat different order than in the list above.

The simplest and most frequent cases are illustrated in (15) – (16). Note that combinations of a first name and a last name are listed here, together with complex first names. SAG's (vol. 2:134) reason for not counting them among proper names is their great productivity due to which they are not perceived of as "lexicalized" proper names but rather as syntactic constructions of two proper names. Although this is in general true, such combinations do show different degrees of lexicalization (conventionalization, entrenchment). To take a couple of examples, the famous Swedish children's book author is normally referred to by *Astrid Lindgren* rather than by just *Lindgren* (which is a very frequent Swedish last name), another famous female writer is primarily known as *Selma Lagerlöf* (rather than *Lagerlöf*), while the legendary film director is very often mentioned as *Ingmar Bergman*, contrasting with the other Bergman celebrity, *Ingrid Bergman*. And in fact PropN-compounds based on such conventionalized 'first name – last name' combinations are particularly frequent among PropN-compounds of this type, but are definitely not the only ones here. In ex. (16b), the compound is based on a combination of a first and a last name that is in a sense "locally" lexicalized, i.e., fairly established within the community of professional linguists and students of linguistics (at least in Sweden). In principle, any combination of a first and a last name can get locally and temporarily "lexicalized" and therefore feed PropN-compounds; there is, however, a whole bunch of pragmatic and linguistic factors that influence the probability with which such compounds are prone to emerge in real language use.

(15) Complex first names
  a. Gustav Adolf+dag-en
     Gustav Adolf + day-DEF.SG.COM
     'the Gustav Adolf-day (6th November)'
  b. Karl Johan-s+stil
     Karl Johan-LNK+style
     'the Karl Johan style, i.e. the Swedish empire'

(16) First name + last name
a. en Selma Lagerlöf+roman; en Ingmar Bergman+film
a:COM Selma Lagerlöf + novel; a:COM Ingmar Bergman + film
'a Selma Lagerlöf novel; an Ingmar Bergman film'

b. en Östen Dahl+föreläsning
a:COM Östen Dahl + lecture
'an Östen Dahl lecture' (Kajanus 2005:1)

What has been said about lexicalized combinations of a first name and a last name applies, mutandis mutandi, to institutionalized combinations of a title and a name. These behave basically like proper names and work fine in PropN-compounds (ex. 17). "Non-analyzability" is particularly evident with foreign titles, as in (17c).

(17) a. mi-n "Drottning-Silvia+klänning"
my-COM Queen-Silvia + dress
'my 'Queen Silvia dress'
(http://www.sr.se/cgi-bin/ostergotland/program/index.asp?ProgramID=2034
visited on December 16th 2008)

b. en drottning Silvia + brytning
a:COM queen Silvia + accent
'a Queen Silvia accent'
(https://www.flashback.info/archive/index.php/t-383060.html
visited on December 16th 2008)

c. ett Mr. Bean+-skämt; en Lord Byron+dikt
a:N Mr. Bean + joke a:COM Lord Byron + poem
'a Mr. Bean joke; a Lord Byron poem'

Names with preposed attributes (epithets) are likewise possible as first components in PropN-compounds, cf. ex. (18). At the risk of stating the obvious, it is worth underlying that the preposed adjectives in such examples cannot refer to accidental, transient states and properties of a person, but make up a part of the "label" whereby (s)he is identified.

(18) Names with preposed attributes
a. duktig-a Annika+syndrom-et
efficient-DEF Annika + syndrome-DEF.SG.N
'the efficient Annika syndrome' (i.e., the allegedly typical female habit of requiring too much of themselves)
https://www.flashback.info/archive/index.php/archive/t-147110-p-3.html
last visited on December 16th 2008

b. några Magr-e Memed+trick
some Thin-DEF.M Memed + trick.PL
'some Thin Memed tricks' (Tora Palm's Swedish translation of Yasar Kemal's *Ince Memed*, Sw. *Låt tistlarna brinna*. 1980:154)

The structurally most complex proper names are, finally, those with postposed identificatory epithets. These are also attested as the first part in PropN-compounds, exemplified in (19) by a few instances with *staty* 'statue' as the head:

(19) a. Peter de-n Stor-e + Staty-n
Peter the-COM Great-DEF.M+statue-DEF.SG.COM
'the Statue of Peter the Great'

b. Karl de-n tolfte + staty-n
Charles the-COM twelfth+statue-DEF.SG.COM
'the statue of Charles the Twelfth'

Compounds of this kind are, on the whole, not frequent. In general, they seem to be favoured by two frequency-related factors: a) a very high frequency and salience of the name with the postposed identificatory epithet, and b) a high salience (and sometimes even symbolic value) of the entity referred to by the whole compound.[10] For instance, Charles the Twelfth has a very special status in the history of Sweden, while his statue is a major landmark in Stockholm. Among other things, Charles the Twelfth has (unfortunately) become an important symbol for Swedish nationalists who celebrate the day of his death (*Karl XII+dag-en* 'Charles XII+day-DEF.COM', November 30th)[11] by demonstrations starting from the statue. In the same way, the special and symbolic value of the statue of Peter the Great in Saint Petersburg ("the Bronze Horseman") is, most probably, a factor that motivates the use of the compound in the following blog text:

(20) Och om Peter den store statyn[12] lämnar staden kommer hela St. Petersburg att gå under (yeah right, som om statyer kan röra på sig).
'And if Peter the great statue leaves the city, the whole of Saint Petersburg will be destroyed (yeah right, as if statues could move).'
(http://resedagbok.krokodil.se/index.php?elev=86)

Proper names with postposed identificatory epithets are, on the whole, quite rare in modern Swedish. In addition to royal names, this

class mainly contains several names referring to the protagonists in famous (children's) books and films such as Emil in Lönneberga and Karlsson on the roof, both coming from Astrid Lindgren's books, and other legendary persons, e.g. Jack the Ripper. The attested on the Internet compounds with these proper names all denote highly salient, outstanding attributes of their referents – Emil in Lönnerga's cap (the one he is always depicted with), Karlsson on the Roof's propeller (thanks to which he can fly) and Jack the Ripper's murders (basically the only thing we know about him), cf. ex. (21-22). It seems that these names in PropN-compounds are pretty much restricted to just one head (or a couple of heads, cf. also *Emil i Lönneberga+bus* 'Emil in Lönneberga mischief').[13]

(21) a. Han hade avklippta byxor, en randig murarskjorta och
[en blå [Emil i Lönneberga+keps]].
a:COM blue.SG.COM Emil in Lönneberga + cap
'He had cut-off trousers, a striped bricklayer's shirt and a blue Emil-in-Lönnerberga cap.'
(Värmlands Folkblad, http://blogg.vf.se/nojesbloggen/?p=47)

b. jag hade tänkt åka med tengil ut och kolla när han krashade
[si-n nya [karlsson-på-tak-et+propeller]]
REFL.POSS-SG.COM new-DEF Karlsson-on-roof-DEF.SG.N+propeller
'I had been thinking of accompanying Tengil and having a look when he would crash his new Karlsson-on-the-roof propeller'.
(http://www.sporthoj.com/forum/showthread.php?t=526&page=507)

(22) Vad kännetecknade samtliga [Jack the Ripper+mord]?
… all [Jack-the-Ripper+murder.PL]
'What characterized all the Jack-the-Ripper murders?'
(http://www.trivia.se/index.php?page=5&cat=55)

5.3. PropN-compounds with coordinated proper names

The first part of PN-compounds can also include reference to two or more people. Proper names can often be coordinated by means of the overt coordinator *och* 'and' (normally pronounced as [o] and sometimes spelled as "å"), cf. ex. (23).

(23) Romeo och Julia+förälskelse-n /+kör-en
Romeo and Julia+love-DEF.SG.COM/+choir-DEF.SG.COM
'the Romeo and Julia love; the Romeo and Julia choir'

Coordinated proper names are particularly felicitous as the first part of PropN-compounds when they refer to well-known couples, more or less conventionalized as such and having therefore properties of a unitary referent (a subtype of what Wälchli 2005:5-6 calls "natural coordination"). This conventionalization of couples can sometimes manifest itself in phonological reduction and in spelling. In ex. (24a), *Hasseåtage* is the established name for the legendary Swedish comedian couple – Hans (Hasse) Alfredsson and Tage Danielsson – pronounced with the typical prosodic pattern of lexphrases. Kain and Abel are so often mentioned together that their coordinated names function as the first part in the compound in ex. (24b) in spite of the fact that the two brothers have opposite roles in the event behind it (the example and the observation come from Kajanus 2005:44)!

(24) a. Hasseåtage (<[Hasse och Tage])+revy-n
Hasse and Tage + show-DEF.SG.COM
'the Hasse and Tage show'

b. Kain och Abel+mord-et
Kain and Abel+murder-DEF.SG.N
'the Kain and Abel murder' (examples from Kajanus 2005)

But even "less" conventionalised couples can occasionally be referred to by the first part of PropN-compounds if they act as a couple in what is denoted by the compound. For instance, a recording where the violinists Menuhin and Grappeli play in duo can be described as *min Menuhin och Grappeli+skiva* 'my Menuhin and Grappeli recording'.

In Swedish, conjunction is not restricted to PropN-compounds, but is possible in other types of compounds, e.g. *en tur-och-retur+biljett* 'lit. a trip-and-return ticket' ('a return ticket', derived from the lexphrase *tur och retur* 'there and back'), *en sol-och-vår+man* 'lit. a sun-and-spring+man ('a confidence trickster who obtains money from a woman by false promises of marriage'), *piano-och-flöjt+musik* 'piano and flute music', where the latter compound represents a fairly productive pattern. Productive conjunction in compounds has been discussed in the literature as one of the strong challenges for the *No phrase constraint* (and, further, for the *Lexical Integrity Hypothesis*), cf. Lieber & Scalise (2006).

In many PropN-compounds the first part contains simply two or several proper names listed one after another without any overt coordinators in-between them, e.g. *Baader-Meinhofligan* 'the Baader-

Meinhof liga', *Clinton-Levinskyaffären* 'the Clinton-Levinsky affair', *Holmes-Watsonforskningen* 'the Holmes-Watson research', *Istomin-Sterne-Rose-trion* 'the Istomin-Sterne-Rose trio', etc. (examples from Kajanus 2005:19).

5.4. PropN-compounds with proper names containing articles and plural markers

Although prototypical proper names in Swedish are not marked with any definiteness markers, some proper names have arisen through lexicalization of definite common names or definite noun phrases and carry therefore definite articles. Of interest for us here are proper names referring to one person, as e.g. *Stål+mann+en* 'lit. Steel+man-DEF.SG.COM' ('Superman'), or to a group of persons (i.e., a musical band), e.g. *Imperi-et* 'Empire-DEF.SG.N' ('the Empire', the name of a pop group). Most remarkably, the suffixed definite article is retained when such proper names are used as the first part in PropN-compounds, as in ex. (25). Number markers, occasionally present in proper names (such as the names of musical bands), are likewise retained in PropN-compounds, cf. ex. (25c). Note that number markers, although normally absent in the first part of ComN-compounds, can sometimes occur there as well, e.g. *fem+veck-or-s+semester* 'lit. five+week-PL-LNK+vacations', 'five weeks long vacations' (but *fem+våning-s+hus* 'lit. five+floor-LNK+house', 'house of five floors').

(25) a. [Stål+mann-**en**] + dräkt-en
steel + man-**DEF.SG.COM** + dress-DEF.SG.COM
'the Superman dress' (107 Google hits, December 13th 2008)

b. en Imperie-**t** + låt
a:COM Empire-**DEF.SG.N**+melody
'a melody coming from the Empire pop group'

c. en Viking-**ar-na** + låt
a:COM Viking-**PL-DEF.PL**+melody
'a melody coming from the Vikings pop group'

Personal proper names can occasionally have both the free syntactic determiner *den/det/de* and the suffixed definite article. As mentioned in in Section 5.1 in connection with ex. (12), the free syntactic determiner *den/det/de* is obligatorily used in Swedish definite noun phrases with adjectival attributes in addition to the definite suffix on their nominal head. This pattern is sometimes – although not very

commonly – attested in personal proper names, e.g., *den Vandrande vålnaden* 'the Wandering ghost' (i.e. the cartoon figure Phantom), *den Flygande Köttbullen* 'the Flying Meatball' (the nickname of a Swedish handball player) and *den flygande holländaren* 'the flying Dutchman'. All my informants agree that the syntactic determiner *den / det / de* has to be dropped when such names occur in PropN-compounds, in contrast to the suffixed definite article on their head, cf. ex. (26).

(26) a. de-n Vandrande vålnad-en
the-SG.COM Wandering ghost-DEF.SG.COM
'the Wandering Ghost'

b. e-tt typisk-t [Vandrande vålnad-en] + äventyr
a-SG.N typical-SG.N Wandering ghost-DEF.SG.COM+adventure
'a typical Wandering Ghost adventure'

These examples show that the morphologically bound definite article and the free syntactic definite element *den / det / de* in Swedish differ in their impact on various morphosyntactic phenomena, with *den / det / de* being closer to what has been termed "determiner" in languages like English, French and German (cf. Börjars 1998 for a careful analysis) and with the suffixed definite article being involved in and sensitive to morphological processes. However, retention of the suffixed definite article in personal proper names used as the first part in compounds is in fact an amazing property of PropN-compounds, one that singles them out from the other kinds of compounds:

a) ComN-compounds are formed with uninflected nominal stems as their first component, which includes, among other things, the absence of definiteness markers;
b) the suffixed definite article on the head of a lexphrase is obligatorily dropped when the whole lexphrase is used as left-most component in a compound, as discussed in Section 5.1. (recall the contrast between *svarta börsen* 'the black market' and *svartabörsaffärer* 'black-market transactions' in ex. (13));
c) finally – and most surprisingly – the suffixed definite article is normally lost in geographical proper names when these appear in compounds, cf. (26) below[14]:

(27) a. *Hö+torg-**et*** => *en* [*Hö+torg*]-*s+skrapa*
Hay+Place-DEF.SG.N a:COM Hay.Place-LNK+skyscraper
'Hötorget (lit. 'the Hay Place'); a skyscraper in Hötorget'

b. *Kung-s+holm-**en*** => *en* [*Kung-s+holm*]*-s+gata*
King-LNK+Island-DEF.SG.COM a:COM King-LNK+Island-LNK +street
'Kungsholmen (lit. 'the King Island'); a street on Kungsholmen'

According to the "No DP!"-constraint (cf. Lawrenz 1996), no DPs (or "full" NPs, i.e. NPs with determiners) are allowed as a non-head in compounds. While this constraint seems to make good predictions for ComN-compounds and "regular" phrasal compounds in various languages (including Swedish), it does not appear adequate for compounds with personal proper names as their non-head component (even though the latter appear to show restrictions on syntactic determiners within them).

6. Concluding remarks

As I hope to have shown in this paper, Swedish PropN-compounds are truly situated between lexicon and syntax.

Syntactic considerations are relevant for PropN-compounds for at least the following reasons:

a) since proper names themselves are often phrases rather than single words, the first part of PropN-compounds can be a phrase, including pre- and postposed attributes of varying complexity, coordination and even definite suffixed articles on the head noun;
b) PropN-compounds are often involved in competition with syntactic coding strategies, and
c) the proper-names in PropN-compounds are occasionally accessible for external anaphora.

Lexical considerations are also relevant for PropN-compounds:

a) PropN-compounding provides a systematic way of creating new lexical items, with many PropN-compounds fully conventionalized or on their way to full conventionalization;
b) various structural properties of PropN-compounds (e.g. recursion, the use of proper names with attributes or of conjoined proper names as the non-head component) are sensitive to the "lexicalization degree" of the involved elements, where lexicalization primarily refers to their recognition as expressions for unitary concepts.

These issues are, of course, not new in the linguistic literature (even though they have hardly ever been discussed with regard to PropN-compounds), and there are ample proposals on how to account for the access of word formation to syntax. For instance, Lieber & Scalise (2006)'s suggestion operates within a modular and constraint-based grammar, where the morphological module specifies well-formedness constraints on complex words, including kinds of phrase that may appear within them, whereas the well-formedness of these word-internal phrases is defined by the syntactic module.

In my opinion, in the case of PropN-compounds this proposal misses the point. The defining property of these compounds is that they are based on personal proper names or on *what is conceived of as a personal proper name*. These expressions can have quite different structural properties ranging from single simplex words to phrases, but all these structures (including certain types of close appositional structures) are allowed to function as the first-most component in compounds *as long as they are recognized as a name for a unique referent* – a person or a group of persons. The only exception is constituted by complex names with a preposed syntactic determiner that has to be dropped in compounding.

Dahl (2004:257) discusses the idea that utterances in spoken language are delivered in units that can be called "packets", which are highly integrated units with constraints on their internal complexity and on what can be done with their elements. According to him, Swedish lexphrases are a good example of a "packet". A proper name, or a structure that is used as a proper name, would be a related example of a "packet". PropN-compounds are simply right-headed nominal compounds in which the first slot is occupied by any expression that is recognized as a personal proper name (with the marginal exception of expressions with a preposed syntactic determiner). Booij's (2009) theory of Constructional Morphology appears to be a promising framework for dealing with these kinds of phenomena.

Acknowledgements

This paper has a long prehistory. My initial curiosity about proper-name compounds was aroused by Östen Dahl's elegant study of the Palme-murder related compounds presented in a talk at the First Swedish Linguist Meeting (SLING) in Uppsala in the Spring of 2003. This interest grew much stronger while I was acting as a supervisor for Eliza Kajanus' BA thesis on proper-name compounds.

I am greatly grateful to both Östen and Eliza for all our inspiring discussions and collaboration – in a way, this paper is a fruit of our collective thinking! I would also like to thank several other colleagues whose wise and inspiring advise has been indispensable – Livio Gaeta, John Payne, Peter Pagin, Anette Rosenbach, Tomas Riad and Jan Svanlund – and the members of my family and my numerous colleagues for their judgements regarding acceptability and interpretation of Swedish proper-name compounds. The earlier versions of this paper have been presented in several places – at the 3rd Swedish Linguist Meeting (SLING) in Lund (April 2004), at the Yerevan Typological School (September 2004), during a guest lecture at the Konstanz University (December 2004) and at the symposium "Nouns cross-linguistically" at the Molise University in June 2007. I am grateful to the audience at all these occasions and to the two anonymous reviewers for their questions and comments.

Address of the Author:

Dept. of Linguistics, Stockholm University, 106 91 Stockholm, Sweden
<tamm@ling.su.se>

Notes

[1] I have chosen to use "+" for separating the compounded stems in the examples.

[2] Swedish distinguishes between two lexical accents – accent 1 and accent 2, the realization of which varies across the different varieties. In Central Swedish varieties, accent 2 is pronounced with two peaks.

[3] Swedish nouns fall into two genders – Common (COM in the glosses) and Neutral (N in the glosses); articles, demonstratives and adjectives agree with nouns in number and, for the singular, also in gender. In addition, adjectives distinguish between the indefinite and definite form (the latter is not normally sensitive to the number and gender distinctions). Nominals referring to singular male persons can occasionally combine with a special (but optional) definite adjectival form ending in *-e*, see e.g. ex. (18b) and (19).

[4] There are also cases in-between well-behaved compounds and possessive, such as "measure expressions" like 'a three days long journey' – *en tredagars resa*, or *en tredagarsresa*, or *en tre dagars resa* (cf. Koptjevskaja-Tamm 2003).

[5] A few consonant-ending stems can be followed by the vowels *-e-* or *-a-* as linking markers, e.g. *jord-e-liv* 'earth-LNK-live' ('our life on earth'), or *barn-a-vård* 'child-LNK-care' ('child care'). The vowel alternation in the compound forms of (ex. 3c) also goes ultimately back to the old genitive inflection.

[6] The data behind these claims come from Google checks for the occurrences of these and similar compounds on the Internet and from replies to my queries directed to a number of native speakers including the members of my family and several colleagues.

[7] Female complex first names, like *Ing-Marie* or *Anna-Lisa*, that are normally pronounced with the compounding prosody and behave like any other first names for purposes of further PropN-compounding, could in principle be used for illustrating the possibility of genuine recursive PropN-compounding.

[8] SAG uses another notation ($_0$) to mark complete deaccentuation. The notation adopted here reflects the idea that the stress of the syllable is in fact retained, but receives no tonal realization – a solution for which there are theoretical arguments (I am grateful to Thomas Riad for the help with this issue).

[9] The only exception here are female complex first names that are normally pronounced with the compounding prosody, cf. fn. 7.

[10] A nice indication that the use of complex proper names with postposed identificatory epithets in PropN-compounds is restricted, comes from the cake-and-tart domain. In contrast to the PropN-compounds denoting Gustav Adolf cakes and Napoleon tarts, *Oscar II:s tårta* 'Oscar the 2nd's tart' with the possessive NP is much more frequent than the compound *Oscar II-tårta-n* 'the Oscar II-tart' (compound), at least in written sources (including the Internet).

[11] The succession number in royal names appearing in PropN-compounds is normally given in Roman numerals rather than spelled as words, cf. *Karl XII dagen* (72 Google hits) vs. *Karl den tolfte dagen (*no Google hits), *Karl XII statyn* (422 Google hits) vs. *Karl den tolfte statyn* (2 Google hits), checked on December 16th 2008.

[12] The spelling in this example, as well as in the examples immediately above, reflects the growing tendency in Swedish to spell parts of a compound as separate words.

[13] In addition, there are compounds like *en Karlsson på taket+film* 'a Karlsson on the roof film' and *Karlsson på taket+sången* 'the Karlsson on the roof song', the first part of which is ambiguous between being a proper name and a film, book or song title.

[14] There are, however, examples of geographical proper names that do retain their suffixed definite article in compounds, e.g. *en Gaml-a Sta-n+krog* 'an Old-DEF Town-DEF.SG.COM+restaurant', where *Gamla Stan* ('the Old Town') refers to the oldest part of Stockholm.

Bibliographical References

ANDERSON Stephen 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.

ANDERSON John M. 2004. On the grammatical status of names. *Language* 80. 435-474.

ANWARD Jan & Per LINELL 1975-1976. Om lexikaliserade fraser i svenskan. *Nysvenska studier* 55-56. 77-119.

BAKKEN Kristin 2006. Lexicalization. In Brown (ed.). 106-108.

BAUER Laurie 2006. Compounds. In Brown (ed.). 719-726.

BOOIJ Geert 2009. Compounding and construction morphology. In LIEBER Rochelle & Pavol STEKAUER. *The Oxford Handbook of Compounding*. Oxford: Oxford University Press. 201-216.

BÖRJARS Kersti 1998. *Feature Distribution in Swedish Noun Phrases*. Oxford: Blackwell.

BOTHA Rudolph P. 1981. A base rule theory of Africaans synthetic compounding. In MOORTGAT Michael, Harry VAN DER HULST & Teun HOEKSTRA (eds.). *The scope of lexical rules*. Dordrecht: Foris. 1-77.

BROWN Keith (ed.) 2006². *Encyclopedia of Language and Linguistics*. Amsterdam / London: Elsevier.

DAHL Östen 2003. *Palmesyndromet*. A talk given at the First Swedish Linguistic Meeting, March 27-28 2003, Uppsala University.

DAHL Östen 2004. *The growth and maintenance of linguistic complexity*. Amsterdam / Philadelphia: John Benjamins Publishing Company.

JOSEFSSON Gunlög 1998. *Minimal words in a minimal syntax. Word formation in Swedish*. Amsterdam / Philadelphia: John Benjamins Publishing Company.

KAJANUS Eliza 2005. Barbiedagis, Hitlerfylla, Åsa-Nissemarxist – *en studie av svenska sammansättningar med personnamn som förled*. Stockholm: Dept. of linguistics, Stockholm University. BA thesis.

KOPTJEVSKAJA-TAMM Maria 2003. *A woman of sin, a man of duty, and a hell of a mess*: Non-determiner genitives in Swedish. In PLANK Frans (ed.). *Noun phrase structure in the languages of Europe*. Berlin: Mouton de Gruyter. 515-558.

KRIPKE Saul 1978. Naming and necessity. In DAVIDSON Donald & Gilbert HARMAN (eds.). *Semantics of natural language*. Boston: Dordrecht. 253-355.

MEIBAUER Jörg 2007. How marginal are phrasal compounds? Generalized insertion, expressivity, and I/Q-interaction. *Morphology* 17. 233-259.

LAWRENZ Birgit 1996. *Der Zwische-den-Mahlzeiten-Imbiß* und *der Herren-der-Welt-Größenwahn*: Aspekte der Struktur und Bildungsweisen von Phrasenkomposita im Deutschen. *Zeitschrift für Germanistische Linguistik* 24. 1-15.

LIEBER Rochelle & Sergio SCALISE 2006. The Lexical Integrity Hypothesis in a new theoretical universe. *Lingue e Linguaggio* 6. 7-32.

SAG = TELEMAN Ulf, Staffan HELLBERG & Erik ANDERSSON (eds.) 1999. *Svenska Akademiens Grammatik*. Vol. 1-4. Stockholm: Norstedts Ordbok.

SVANLUND Jan 2002. Lexicalization. *Språk & Stil. Tidskrift för svensk språkforskning* 12. 7-45.

VAN LANGENDOCK Willy 2007. *Theory and typology of proper names*. Berlin / New York: Mouton de Gruyter.

WARD Gregory, Richard SPROAT & Gail MCKOON 1991. A pragmatic analysis of so-called anaphoric islands. *Language* 67. 439-474.

WARREN Beatrice 1978. *Semantic Patterns of Noun-Noun Compounds*. Gothenburg: Acta Universitatis Gothoburgensis.

WÄLCHLI Bernhard 2005. *Co-compounds and natural coordination*. Oxford: Oxford University Press.

WRAY Alison 2002. *The transition to language*. Oxford: Oxford University Press.