ON NOUN PHRASE ARCHITECTURE, REFERENTIALITY, AND ARTICLE SYSTEMS*

Øystein Alexander Vangsnes

Abstract. Some languages have articles, others have not. This article argues that semantically equivalent noun phrases have the same underlying phrase structure across languages, no matter whether articles exist or not. The core idea is that functional categories are abstract syntactic heads correlated with specific referential categories. Presence vs. absence of a functional category determines whether or not its particular referential properties are found on the noun phrase. The functional heads are found universally and are a part of Universal Grammar, but since they are abstract, they must be 'identified': *Identification* is accomplished when an element with the relevant morphological features is merged within the functional head projection. When the numeration contains several candidates for identification, the one with the highest number of relevant morphological features (with particular qualifications) is chosen. Crucially, articles are on principle available as potential identifiers, and languages that have articles must therefore use them if they are more appropriate identifiers than any of the elements in a given numeration. Having no semantic properties of their own, articles become specialized identifiers for particular functional categories. This theory, *Identification Theory*, captures important traits of the article systems of various languages, notably those of English, Finnish, French, Icelandic, Mainland Scandinavian, and Northern and Northeastern Swedish 'Österbotten' dialects.

1. Introduction

Can we model the referential properties of noun phrases in a manner which reflects their internal syntactic structure? And would such a theory of noun phrases provide us with valuable insights?

The aim of this article is to demonstrate that these questions should be answered in the affirmative. Such an objective may not appear controversial: any theory of grammar must include the insight represented by the *Principle of Compositionality*, i.e. the fact that the meaning of a linguistic expression is a function of the meaning of its part *plus* the way in which these parts are put together. Nevertheless, within generative grammar the more influential approaches to noun phrase interpretation and compositionality have focused on the interaction between noun phrases and clause structure, either on the positions noun phrases have at LF (May 1985, Diesing 1992) or on the relation between noun phrases and other constituents, notably Case assigners (Belletti 1988, De Hoop 1992).

Although there is little doubt that clausal relations are important for

* This paper is a revised and substantially reduced version of part I of Vangsnes (1999). For the present version I am grateful to Frank Drijkoeningen, Janne Bondi Johannessen, Pino Longobardi, Görel Sandström, Anna-Lena Wiklund, and an anonymous reviewer for discussion, comments, and criticism. Remaining shortcomings are my responsibility alone.

Studia Linguistica 55(3) 2001, pp. 249–299.
Published by Blackwell Publishers, 108 Cowley Road, Oxford OX4 1JF, UK, and
350 Main Street, Malden, MA 02148, USA

the understanding of the semantics of noun phrases, this paper will focus on the interaction between the internal composition of noun phrases and their interpretation and develop a theory where noun phrase internal structure directly reflects semantic properties.

Consider first the two English sentences in (1) which both have subjects containing the universal quantifier *all*.

(1) a. All linguists speak at least two languages.
b. All **the** linguists speak at least two languages.

These sentences differ in meaning. The sentence in (1a) is a generic statement about linguists and their ability to speak languages – the subject noun phrase refers to all existing linguists, i.e. the universal set of linguists. The sentence in (1b) on the other hand is a statement about a 'given' or 'familiar' group of linguists, a subset of the universal set, and the subject noun phrase refers to the whole of this set. If for instance the speaker is talking about a meeting at the Faculty of Arts at Discourse University and presenting general information about the people who attended, the most salient interpretation of (1b) would be that it is a statement about the totality of linguists who attended the meeting. The sentence in (1a) would however refer to the linguists at the meeting as well as all other linguists in the world.

The difference between the two sentences can be related to the presence or absence of the definite article whose core function is to relate the denotation of the noun phrase to a contextually given referent, and in fact the referential properties of the subject noun phrase in (1b) are quite similar if not identical, in the relevant respects at least, to those of the subject noun phrase in (2), the first one containing the definite article and the second one a demonstrative.

(2) a. **The** linguists speak at least two languages.
b. **These** linguists speak at least two languages.

In comparison with these sentences, in particular (2a), the presence of the universal quantifier in (1b) appears merely emphatic, stressing the 'totality' reading of the noun phrase. All the same, the semantic difference between the subject noun phrases in (1) can be related to an overtly expressed difference (the presence vs. absence of a 'word', i.e. the definite article).

Not all languages have a definite article. Finnish lacks this element, and the sentence in (3) may in fact correspond to both of the sentences in (1).

(3) Kaikki kielitietelijät puhuvat kahta kieltä. *Finnish*
all-NOM linguists-NOM speak two-PAR language-PAR
'All (the) linguists speak two languages.'

In this case Finnish has no overt correlate for the semantic distinction which may be expressed by the presence versus absence of the definite article in English. This illustrates the fairly trivial fact that languages

differ with respect to their grammatical inventory and with respect to which semantic oppositions are grammaticalized.

The core referential property of the definite article in English may however be expressed also in Finnish, for instance by adding a demonstrative to the subject noun phrase as in (4).

(4) Kaikki nämä kielitietelijät puhuvat kahta kieltä *Finnish*
all-NOM these-NOM linguists-NOM speak two-PAR language-PAR
'All these linguists speak two languages.'

The subject noun phrase in this example will necessarily denote (the totality of) a contextually given set of linguists and not the universal one.

The question then is whether the overtly expressed semantic distinction in (1) has a *covert* correlate in the Finnish sentence in (3) which in turn may have an overt reflex in some cases such as in (4). I will argue that this is the case, and a central task of the discussion to follow will be to explicate how.

Consider next a different issue. The noun phrases in (5) and (6) from Danish and English, respectively, correspond to each other in meaning.

(5) alle (*av) de tre lingvister *Danish*
all of the three linguists

(6) all (of) the three linguists *English*

Danish differs from English in that what we may call the 'partitive' preposition *of* may not intervene between the universal quantifier and the definite article. Danish is representative of Mainland Scandinavian in general in this respect. In English on the other hand, the partitive preposition may be present, and my impression is that its presence is preferred in American English.[1]

However, disregarding the presence of the partitive preposition, we may observe that the relative ordering of the three determiners in (5) and (6), i.e. the universal quantifier, the definite article, and the numeral, is quite fixed in both Danish and English. The examples in (7) and (8) illustrate that the ordering of constituents in (5) and (6) is the only possible one.

(7) a. *alle tre de lingvister *Danish*
b. *tre alle de lingvister
c. *tre de alle lingvister
d. *de tre alle lingvister
e. *de alle tre lingvister

(8) a. *all three the linguists *English*
*b. *three all the linguists*

[1] I thank Helge Dyvik for pointing this out to me.

c. **three the all linguists*
d. **the three all linguists*
e. **the all three linguists*

One may then ask whether the ordering of the constituents in (5) and (6) is coincidental. Data which suggest so do exist, and it is therefore not immediately clear that some deeper ordering principle should give us the constrained ordering facts witnessed by (5)–(8). For one thing, in English the structure in (9a) is acceptable.

(9) a. all three of the linguists
b. *all three the linguists

In (9a) the numeral precedes the definite article. However, the presence of the partitive preposition *of* is obligatory in this case, cf. (9b), and one might therefore argue that the construction is not immediately comparable to those in (6) and (8).

Moreover, in Icelandic a noun carrying a suffixed definite article may either intervene between the universal quantifier and the numeral or follow both. This is illustrated in (10).

(10) a. allir málvísindamennirnir þrír *Icelandic*
all linguists-DEF three
b. allir þrír málvísindamennirnir
all three linguists-DEF
Both: 'all the three linguists'

This fact about Icelandic can actually be taken to both parallel and not parallel the situation in Danish and English: in (10a) the element containing the definite article occupies the same position as in Danish and English, i.e. relative to the other constituents, whereas in (10b) the element containing the *noun* occupies the same position as in Danish and English – and, as we see, it is the same element in both cases.

Although examples like (9a) and (10) suggest that the fixed ordering of the constituents we witness in Danish and (varieties of) English does not reflect some universal underlying ordering principle, I will claim that the Danish/English pattern does reflect a particular universal ordering of functional categories and, in fact, that cases like (9a) and (10) can be related to this universal linearization of functional categories in a principled and non-exceptional way. We will return to the discussion of the functional categories involved shortly while a discussion of the 'exceptional cases' will have to wait until the end of the paper.

Once the syntactic domain of the definite article is defined, we may correlate its core semantic property with this domain. In turn, given that we want to argue that the overt reflex of the semantic distinction between the English sentences in (1) has a covert counterpart in the Finnish sentence in (3), one way of capturing this is to say that the Finnish subject

noun phrase in (3) is phrase structurally-ambiguous in a way which correlates with its referential ambiguity as well as with the overtly expressed difference in the English sentences in (1).

This constitutes the core idea to be explored here: functional categories correlate with referential properties and their presence or absence in the phrase structure determines the semantic properties of the noun phrase. Moreover, functional categories are abstract and since they are abstract they must be *identified* by some overt material: in languages with a definite article the definite article will be a specialized identifier for a particular functional category whereas this functional category will be identified by other means in languages that have no definite article. The theory to be presented may accordingly be termed *Identification Theory*.

In section 2 we will now first consider some aspects concerning noun phrase referentiality. Then, in section 3, the theoretical framework will be presented. In section 4 we subsequently turn to discuss various article systems in the light of the theory. In section 5 we return to the exceptional English and Icelandic examples in (9a) and (10) and certain other issues. Section 6 concludes the article.

2. Noun phrase architecture and referentiality

2.1. *Nominal functional categories*

Within contemporary generative syntactic theories, and in particular what we may call the 'Chomskyan' framework, noun phrases[2] are taken to have a domain of functional projections on a par with the domains projected from clausal functional categories such as C and I (AgrS, T, etc.). Szabolcsi (1983), Hellan (1986), and Abney (1987) were among the first to suggest that determiners head a projection above the NP-level. Since Abney (1987) this projection is generally referred to as 'DP'. Abney (1987:338f) moreover argued that quantifying determiners head a functional projection QP intervening between DP and NP (see also Löbel 1990). Since then other authors, e.g. Shlonsky (1991), Giusti (1991), Sigurðsson (1993), and Giusti and Dimitrova-Vulchanova (1996), have explored the idea that *universally* quantifying determiners (e.g. *all*, *every*, *each*) head a separate QP above the DP-level, leaving the lower QP for the existentially quantifying determiners. Moreover, Ritter (1991) proposed a functional projection, 'NumP', situated between DP and NP and associated with the morphological category number, and given that there is a clear affinity between existentially quantifying determiners and the category number, we sense a parallel

[2] Throughout the paper I will use the term 'noun phrase' to refer to what traditionally has been designated by the label 'NP' and within much of the contemporary generative literature by 'DP'. In other words, by 'noun phrase' I mean the phrasal units which consist of both the immediate projection of the noun as well as the functional projections associated with it. Later on this will be defined as the 'extended projection' of the noun, cf. Grimshaw 1991.

between NumP and the lower QP. Lamontagne and Travis (1986, 1987) suggested that there is a functional category related to Case above the DP-level. They labelled the projection of this functional category K(ase)P and viewed it as the nominal parallel of the clausal functional projection CP. This idea is developed further in Löbel (1994) and taken up by Sandström and Holmberg (1994) amongst others.

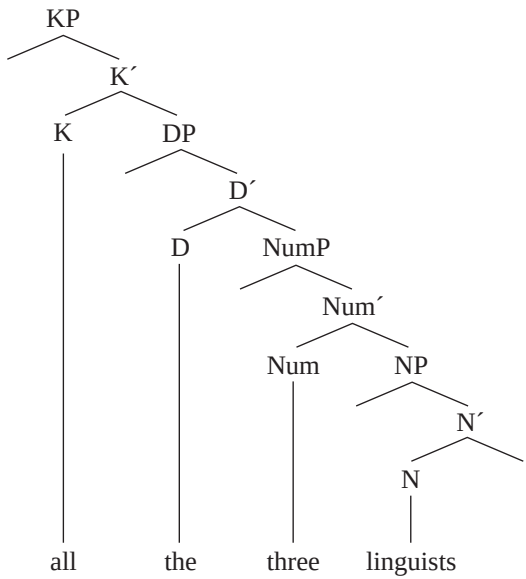
If we ignore the fact that additional functional projections have been proposed and also that the various proposals are not always compatible with each other, an amalgamation yields the structure in (11) as the maximal extended projection for noun phrases.

(11) $[KP/QP_{\forall} \ldots [DP \ldots [QP_{\exists}/NumP \ldots [NP \ldots .]]]]$

Let us dispense with the label QP but keep the idea of different domains for universally and existentially quantifying determiners in mind. Accordingly, the noun phrase in (12) could be analyzed as in (13).

(12) all the three linguists

(13)



Given this analysis, there is then a putative correlation between functional projections and a certain semantic classification of determiners: *three* belongs to the class of cardinal determiners, *the* to the class of definite determiners, and *all* to the class of universally quantifying determiners. In terms of referentiality cardinal determiners entail that the noun phrase they occur in is 'countable', definite determiners entail that the noun phrase is 'discourse anaphoric', whereas universally quantifying determiners entail that the noun phrase is 'uniquely referring'. We will return to these referential categories shortly.

The observation of the correlation between determiners and syntactic domains is central to the main ideas to be presented in this article: I will argue that the three functional categories K, D, and Num are correlated with the semantic properties that we can characterize and distinguish the three types of determiners by. This then gives us both the starting point and the goal for our investigation: I assume, and wish to show, that the ordering of determiners within the noun phrase follows certain syntactic principles, and moreover that these syntactic principles have a semantic basis.

Let us consider the semantic properties of the various determiners more closely.

2.2. *The strong/weak distinction*

As observed by Milsark (1974, 1977) noun phrases containing universally quantifying and definite determiners share the property of normally being excluded from occurring as the postverbal argument of existential sentences in English and a number of other languages. Noun phrases with (only) a cardinal determiner on the other hand are not excluded. This is illustrated in (14).

(14) a. There were three linguists/no linguists/*the linguists/*those linguists/*both linguists/ at the meeting.
b. There was a linguist/*every linguist/*each linguist at the meeting.
c. There is beer/much beer/a lot of beer/*all beer/*my beer in the fridge.

Milsark referred to this phenomenon as the 'Definiteness Restriction'. It has later also been called the 'Definiteness Effect' (cf. for instance Safir 1985, 1987; Belletti 1988), and there exists an enormous amount of literature on this topic which would lead too far to review here. (See e.g. Safir 1987 and Vangsnes 1994 for overviews.)

Based on whether or not a noun phrase is allowed to occur as the postverbal argument of an existential sentence, Milsark (1977) made a distinction between 'weak noun phrases', which are allowed, and 'strong noun phrases' which are disallowed. Based on this distinction Barwise and Cooper (1981) in turn termed the determiners that occur in the two types of noun phrases 'weak' and 'strong', respectively, and they developed a formal account of the semantic properties determiners yield. Thus universally quantifying determiners and definite determiners both came to belong to the class of 'strong determiners' whereas cardinal determiners belong to the class of 'weak determiners'.

In turn the class of weak determiners comprise a variety of determiners which strictly speaking do not entail cardinality in a narrow sense, i.e. as 'specifying a number'. Examples of such determiners are *some*, *many*, *few* etc.

By combining Milsark's test for noun phrase types and Barwise and

Cooper's terminology for determiners we arrive at the following classification of determiners.

(15) **Table 1** Classification of determiners as strong and weak

strong		weak
'∀'	'definite'	
all	*the*	*one, two three . . .*
each	*this*	*many*
every	*that*	*some*
both	*my*	*a*
[. . .]	[. . .]	*no*
		a lot of
		[. . .]

In the remaining subsections of section 2, I will discuss the three referential properties *uniqueness*, *discourse anaphoricity* (*specificity*) and *countability*, and how they relate to the various types of determiners in Table 1. Readers familiar with this issue may proceed directly to section 3.

2.3. *Uniqueness*

The core referential property of strong noun phrases is that they carry the presupposition that their referents exist. This presupposition is fairly indirect and not logically necessary, but the listener must at least take it that the speaker is referring to entities assumed to exist.[3] This presupposition of existence of course refers to either the real or some fictitious world, hence allowing noun phrases like *all unicorns* and *the present King of France* to have the same basic semantic properties as noun phrases where the noun denotes kinds of entities that no doubt exist.

One way to capture this core referential property of strong noun phrases is to say that they typically denote referents that are identifiable for both the speaker and the listener. Consider again the sentences in (1) and assume the same context as above (i.e. a faculty meeting at Discourse University).

[3] The use of a strong noun phrase to denote an empty set may come unexpected on the part of the listener and yield comic effects such as in the following example.

(i) All boys at the party were drunk, but there were no boys at the party!

An exception in this respect seems to involve the strong determiner 'free choice' *any* – since it may not occur in the postverbal argument of existential sentences, it is strong according to the Milsarkian test.

(ii) *There was any boy at the party.

Consider the following grammatical use of the determiner.

(iii) Any teacher will yell at a disobedient pupil like that.

Free choice *any* is arguably a universal quantifier in that it quantifies over universal sets, but it does not seem to entail a presupposition of existence in the same way as *every* and *all*. Rather it carries the information that *if* the referent exists it will have a certain property given a certain situation. See Sæbø (1999) and references cited there for discussion.

(1) a. All linguists speak at least two languages.
b. All **the** linguists speak at least two languages.

Upon hearing these sentences the listener can identify the referents of the subject noun phrases, i.e. the universal set of linguists in (1a) and the totality of the contextually given set of linguists in (1b). In that respect strong noun phrases can be said to be 'uniquely referring'.

It is an important point that the listener need not be able to identify the actual (real or possible world) referents of the strong noun phrases. Let us say that the sentence in (1b) were preceded by the one in (16).

(16) There are philosophers, historians, and linguists at the meeting, and it sure constitutes a collection of skilled people.

In that case the extension of *all the linguists*, which the listener is capable of identifying, is the one holding between that noun phrase and the set of linguists introduced by the noun phrase *linguists* in the preceding sentence. Whether or not the listener would be able to point these linguists out on a photo or on the street is irrelevant. What is relevant is that *all the linguists* can only refer to whichever linguists were at that particular meeting and not to any other set of linguists.

Another important point is that the extension of a strong noun phrase, which by definition then is identifiable for the listener, is not dependent on a discourse referent having been introduced in the linguistic context. The extension may also be *inferred* by the listeners by virtue of their knowledge and/or conventional conceptions of the world. For instance, consider a different text where the sentence in (1b) is not preceded by (16), but rather by (17), and when there has been no previous mention of linguists.

(17) There's a meeting going on at the Faculty of Arts and it sure constitutes a collection of skilled people.

In this case, upon hearing (1b) the listeners will infer from their knowledge that faculties of arts often comprise departments of linguistics, and hence linguists, that there were linguists at the meeting.[4]

[4] The relevance of 'inference' can be illustrated with a variety of examples. I will mention two additional ones. Take first the example in (i).

(i) I know a couple in Toronto. **The wife** is a teacher.

It is a (conventional) fact about the world that couples consist of a husband and a wife. The subject noun phrase in the second sentence therefore denotes a subpart of the discourse referent introduced by the weak noun phrase *a couple in Toronto*, and the listener can identify this extensional relation.

Consider next the sentence in (ii), which is not true about the real world.

(ii) The present King of France is bald.

However, this sentence could all the same be a part of a felicitous communication exchange, and the listener would then infer (a) that there exists a kingdom named France, and (b) that since kingdoms have only one king the subject noun phrase in (ii) denotes the one individual which happens to be the King of France. For the listener to identify a real world referent for the noun phrase would be a matter of possessing the appropriate knowledge.

Returning now to the sentence (1a), the reference of the subject noun phrase in this example will always be inferred. Its extension need not relate to a previously introduced discourse referent and actually not to an inferred discourse referent either. However, by its form the noun phrase conveys the information that the totality of a universal set is denoted, and the listener will know this. Accordingly, the listener will always be able to retrieve a referent for the noun phrase no matter what the context.

Stated more explicitly then, the difference between the two sentences in (1) is that the subject noun phrase in (1b) denotes a referent which is either contextually given or contextually inferred whereas the subject noun phrase in (1a) does not. Both noun phrases however have an extension which can be identified by the listener, and that is what we understand by 'unique reference'.

2.4. *Discourse anaphoricity and specificity*

Although noun phrases with a definite determiner typically have referents that can be identified by the listener, this does not always hold. A well-known case is the so-called 'indefinite-*this*' in English. As shown by the example in (18) indefinite-*this* may occur in the argument noun phrase of existential sentences.

(18) There was this linguist at the meeting.

The use of indefinite-*this* noun phrases is frequent in colloquial (American) English, and its typical textual function is to introduce new discourse referents (see Prince 1981 for discussion). The only formal difference between indefinite-*this* and 'definite-*this*' (i.e. the regular use of the demonstrative) is that indefinite-*this* is always unstressed. Unlike the standard use of the determiner the 'indefinite' use does not yield noun phrases with referents that can be identified by the listener. However, the speaker will necessarily have a particular referent in mind, and the noun phrase thus presupposes the existence of its referent. This becomes evident if we add negation to the sentence in (18).

(19) *There wasn't this linguist at the meeting.

The communicative function of existential sentences can be argued to be to predicate the existence of the referent denoted by the predicate internal noun phrase. Accordingly, negated existential sentences predicate the *non*-existence of the referent denoted by the predicate internal noun phrase. From this it follows that a noun phrase which presupposes the existence of its referent cannot occur in negated existential sentences.

This is what we see with indefinite-*this* noun phrases, and in contrast to the pair of examples in (18) and (19) we may consider noun phrases with weak determiners such as *a* (the indefinite article) and *many*. The

examples in (20) and (21) show that these determiners may occur in the predicate internal argument of both negated and non-negated existential sentences, suggesting that the noun phrases do not presuppose the existence of its referents.

(20) a. There was a linguist at the meeting.
b. There wasn't a linguist at the meeting.

(21) a. There were many linguists at the meeting.
b. There weren't many linguists at the meeting.

It is however generally agreed upon that the speaker *may* have a particular referent in mind when using a weak noun phrase: the predicate internal arguments in (20a) and (21a) may introduce a discourse referent which it later will be possible to refer back to with a uniquely referring noun phrase, say a noun phrase with a definite determiner or a pronoun. The examples in (22a) and (22b) could follow the examples in (20a) and (21a), respectively, in a text.

(22) a. [. . .] The linguist/he was enthusiastic about his new theory.
b. [. . .] The linguists/they were quarrelling about theoretical issues.

For the speaker to have a particular referent in mind is one way of defining the notion 'specificity'. This notion has been much debated, and I will not go into a discussion of it here, but simply state that I will adhere to the view just mentioned. This view is in line with Abbott (1992, 1993) and corresponds to the 'referential reading' of noun phrases discussed in Fodor and Sag (1981), but it goes against the view of specificity advocated in Enç (1991). (See Vangsnes 1994 for further discussion and references.)

Importantly, the view of specificity taken here provides us with a test for specificity: if the noun phrase may occur as the predicate internal argument of a non-negated existential sentence but not a negated one, it presupposes the existence of its referent and is therefore specific.

Although the exact understanding of the notion 'specificity' is controversial there is general agreement that noun phrases with adjectives like *certain* are specifically referring. This harmonizes with our test: the following pair of examples shows that the noun phrase *a certain linguist* is not possible in a negated existential.

(23) a. There was a certain linguist at the meeting.
b. *There wasn't a certain linguist at the meeting.

Moreover, we may now link the notions of 'specificity' and 'discourse anaphoricity': if a noun phrase is specific whenever the speaker has a particular individual (or set of individuals) in mind, such a noun phrase will necessarily be discourse anaphoric in the sense that the referent is a part of the context for the utterance. Returning to the sentence pair in (1),

the subject noun phrase in (1a) is non-specifically referring whereas the one in (1b) is specifically referring.

(1) a. All linguists speak at least two languages.
    b. All **the** linguists speak at least two languages.

In other words a universal set does not count as a 'particular individual'. Moreover, a noun phrase may be uniquely referring without being specific/discourse anaphoric and a noun phrase may be specific without being unique. Uniqueness and specificity are distinct semantic properties, and below I will argue that they are associated with distinct syntactic domains. More specifically, I will argue that uniqueness is associated with the functional projection KP and specificity with DP, cf. the structure in (13).

2.5. *Countability*

The third referential notion to be singled out is 'countability'. Above I classified *three* as a *cardinal* determiner. A noun phrase containing *three* will necessarily denote a countable referent, i.e. a set consisting of three members. However, it is not only numerals that yield noun phrases with countable referents. Any noun phrase that denotes a set is countable even if the exact cardinality of the set cannot be established: the important point is that the referent consists of distinct parts that *may* be counted. Accordingly, also other weak determiners like *many*, *some*, and the indefinite article *a* yield countable noun phrases, and the term 'cardinal determiner' should be understood in a broad sense to subsume determiners that yield noun phrases with countable referents.

In this sense the notion of countability distinguishes between noun phrases that denote sets and noun phrases that denote *masses*: masses do not consist of distinct, countable subparts. As an illustration consider the following examples involving the set denoting noun phrase *many linguists* and the mass denoting noun phrase *lemonade*.

(24) I met many linguists yesterday. In fact, fifteen!

(25) I drank lemonade at the party. #In fact, three!

The appositive statement in (24) is felicitous since *many linguists*, which the number *fifteen* relates to, denotes a set and thus a referent with a 'latent' cardinality. In (25) on the other hand the number relates to a noun phrase which under the most salient reading does not denote a set: in order for the example to be well-formed one must by pragmatic implicature understand *three* as meaning something like 'three glasses of __', i.e. where the mass is individuated by a container of some sort.

In the mass denoting noun phrase *lemonade* there is no determiner, but there do exist determiners which yield noun phrases that denote masses

rather than sets. Examples would be the weak determiners *much* and *a lot of*. The examples in (26) and (27) clearly show that these determiners do not yield noun phrases with countable referents, and are therefore 'non-cardinal'.[5]

(26) Much water has passed under this bridge since we first met.
*In fact, ten!

(27) The president has had a lot of trouble because of this affair.
*In fact, twenty!

Determiners like *much* and *a lot of* can therefore be labelled *mass determiners*.

In English and many other languages mass denoting noun phrases will often not contain any determiner but rather just a singular noun (cf. (14c) above). For such languages many linguists draw a distinction between *mass* and *count* nouns: nouns whose singular form by itself may constitute a mass denoting noun phrase (such as *lemonade* above) are considered *mass nouns* whereas nouns which lack this ability are considered *count nouns*. Although the distinction is convenient from a descriptive point of view, I will argue that it is not a part of the lexical information on nouns in the sense that each noun is specified as either mass or count.

I base this point of view on the fact that typical mass nouns may be used in a 'count noun way' and typical count nouns may be used as count nouns, sometimes by convention, sometimes not, but always with interpretational implications. Consider first the examples involving the noun *beer* in (28).

(28) a. There is beer in the fridge.
b. There is one beer in the fridge.
c. There are beers in the fridge.

Beer would normally be considered a mass noun, and the noun phrase *beer* in (28a), constituted by the singular form of the noun alone, does denote a mass. It is irrelevant whether the beer in the fridge is contained in one or more containers: the sentence is compatible with a state-of-affairs where there is only one bottle of beer in the fridge as well as one where there are ten bottles, and it would also be compatible with there being one or more kegs of beer in the fridge.

In (28b) on the other hand the singular form *beer* combines with the cardinal determiner *one*, and the denotation of this sentence is only compatible with a state-of-affairs where there is one and only one

[5] One may hold that abstract entities like 'trouble' do not involve masses, but often are referred to by expressions which show similar formal characteristics as those that denote concrete mass type entities like liquids. We may therefore treat such concrete and abstract entitites on a par.

container of beer in the fridge, by convention either a bottle or a can. In other words, by using the plural form of *beer* the existence of the beer in the fridge is conceptualized by its containers and the noun phrase denotes a set. This is then 'count noun' use of *beer*.

The noun phrase *beers* in (28c) also denotes a set. Unlike the sentence in (28a) but like the one in (28b) it is not irrelevant in what form the beer is contained in the fridge: there has to be more than one container of beer (and by convention these should be either bottles or cans). In this case the plural form of the noun *beer* is used with no determiner.

Consider next the following two passages.

(29) a. The bastards broke into the little cottage and ruined the furniture. First they chopped up **a table** and then they used **some chairs** to make a bonfire right on the middle of the floor.
b. The tree eating ants invaded the little cottage and ruined the furniture. It was as if they had **table** for dinner and **chair** for dessert.

Table and *chair* generally occur in noun phrases that denote sets. This is the case in (29a) where they combine with the cardinal determiners *a* (the indefinite article) and *some*. In (29b) on the other hand we find a more creative use of the two nouns. They occur in their singular forms and do not combine with any determiner, and importantly the two noun phrases do denote masses: the statement is compatible with the ants destroying one table or twenty, and there may be either dozens of chairs involved or there may be only one. Conceptually speaking there is a focus on the material 'table' and 'chair', not on the instances of the entities.

What facilitates this conceptualization is probably the likelihood of the ants destroying the furniture by eating it: it is then easier to regard it as a substance rather than a set of individual pieces. With regard to food terms we quite generally see an alternation between 'count' and 'mass' noun uses. Consider the following pair of examples.

(30) a. The linguists had turkey for dinner.
b. The linguists had a turkey for dinner.
c. The linguists had turkeys for dinner.

Again, for the sentence in (30a) it is of no importance how many turkeys were actually served at the dinner, whereas there could be only one at the event denoted by (30b). In turn there would have to be more than one turkey at the event denoted by (30c).

The mass/count distinction thus appears to be a matter of how we conceptualize the world, and the fact that certain nouns are more typical 'mass nouns' than others can be related to what kind of entity they denote, i.e. a mass noun will denote an entity that typically exists as a non-individualized substance.

Below I will argue that the mass/count distinction is a matter of whether or not the functional category Num is present or not in the phrase structure: if present the noun phrase will denote a set, if not present it will denote a mass. Moreover, only cardinal determiners, and not mass determiners, are associated with the functional projection NumP, cf. the structure in (13). Roughly speaking the abstract functional category assumed to head NumP is typically *identified* by a cardinal determiner or by a number affix.

Importantly, the property of yielding discourse anaphoric noun phrases cuts across the mass/count distinction. The noun phrase *the beer* in (31a) is ambiguous between a set and a mass reading: it could refer either to a bottle or can of beer or to a whole case full of bottles of beer.

(31) a. Put the beer in the fridge!
b. Put the beers in the fridge!

In (31b) where the noun is plural things are different: *the beers* can only refer to a set of containers of beer. These examples illustrate then that a noun phrase need not be countable in order to be discourse anaphoric. What is more a noun phrase need not be countable or discourse anaphoric in order to be uniquely referring. The example in (32) serves to illustrate this.

(32) All German beer is brewed according to das Deutsche Reinheitsgebot.

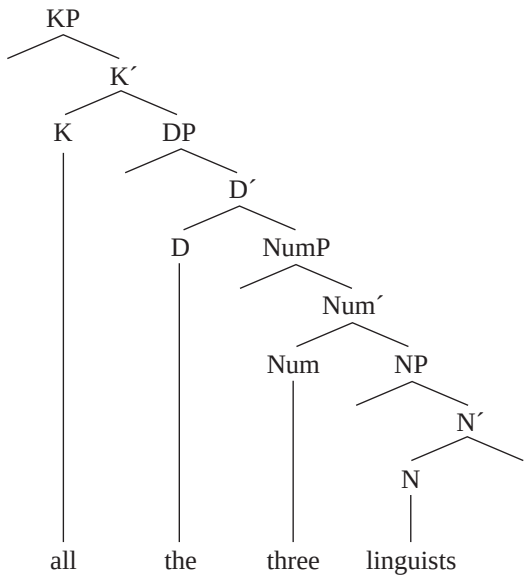
In other words, the properties uniquely referring, discourse anaphoric, and countable may be freely combined with each other. At this point we have set the stage for a presentation of Identification Theory.

3. Identification theory

3.1. *Functional categories and referential properties*

Returning to the structure in (13), repeated here, I will argue that there is a correlation between the three referential categories uniqueness, specificity/discourse anaphoricity, and countability and the three functional categories K, D, and Num, respectively.

(13)



The analysis will however be further qualified. First of all, I argue that the functional projections are not headed by the determiners but rather by abstract functional categories, listed in the lexicon, and it is these abstract functional heads that are correlated with the referential properties. More specifically NumP is headed by ν, and the presence of ν entails that the noun phrase is countable. DP is headed by δ, and the presence of δ entails that the noun phrase is discourse anaphoric. The presence of ν and δ, and hence NumP and DP, in a noun phrase is optional, but whether or not they are present will have consequences for the referential properties of the noun phrase: absence of ν entails absence of countability, and conversely absence of δ entails absence of discourse anaphoricity.

As for the KP-domain the story is slightly more complicated. First of all, I will assume that Case is assigned to this domain by Case assigners (e.g. verbs and prepositions), and that Case spreads from this domain to the rest of the noun phrase. An assumption related to this is that, unlike NumP and DP, the KP-domain is always present in argument noun phrases – if it is not present Case cannot be assigned.[6] In other words, KP may be regarded as the label for argument noun phrases and the projections that close them off (see Sandström and Holmberg 1994:85). In that respect there is a clear parallel between KP and CP at the clause

[6] The obligatoriness of the KP-domain thus represents a correlate of the Argument Rule of Delsing (1993:65) which says that all argument noun phrases must have a determiner position. See also Longobardi (1994). Moreover, it allows us to treat noun phrases as a uniform constituent type at the clause level, i.e. always as KPs and not as either NP, NumP, DP, or KP.

level, as originally proposed by Lamontagne and Travis (1986, 1987), cf. Löbel (1994). This idea also corresponds to the original 'KOMP' proposed by Szabolcsi (1983), i.e. the equivalent of COMP and the domain through which embedded constituents may move out of the noun phrase. Hence, the label K(P) should allude to the affinity between the projection and (i) 'komplementizer', (ii) 'kase', (iii) (universal) 'kwantifiers', and (iv) the 'klosure' of the noun phrase.

The head of KP is the functional category κ, and it exists in two varieties: $\kappa_{[+\text{unique}]}$ and $\kappa_{[-\text{unique}]}$. When KP is headed by the former, the noun phrase has unique reference, and when it is headed by the latter it does not. The reason we must assume that κ, unlike δ and ν, exists in two varieties, is the obligatory presence of KP – all argument noun phrases (i.e. noun phrases that receive Case) must contain this domain, but all of them are of course not uniquely referring.

A projection that is not headed by a functional category is headed by a 'substantive' category, and N is such a category along with the other big, open word classes (such as V and A). Substantive categories are thus not abstract. Moreover, semantically speaking the distinction between functional and substantive categories correlates with the distinction between extension and intension: the substantive categories contribute the core characteristic properties of the type of entity a phrase structural object denotes and the functional categories provide an extensional anchoring for it.[7]

Given the abstract functional categories and their syntactic and semantic definitions the system gives us a typology of 8 configurationally distinct kinds of argument noun phrases which are schematically given in (33) with their respective referential properties indicated in parentheses.

(33) a. $KP_{[+u]}$ DP NumP NP (+unique, +discourse anaphoric, +countable)
b. $KP_{[+u]}$ DP NP (+unique, +discourse anaphoric, –countable)
c. $KP_{[+u]}$ NumP NP (+unique, –discourse anaphoric, +countable)
d. $KP_{[+u]}$ NP (+unique, –discourse anaphoric, –countable)
e. $KP_{[-u]}$ DP NumP NP (–unique, +discourse anaphoric, +countable)
f. $KP_{[-u]}$ DP NP (–unique, +discourse anaphoric, –countable)
g. $KP_{[-u]}$ NumP NP (–unique, –discourse anaphoric, +countable)
h. $KP_{[-u]}$ NP (–unique, –discourse anaphoric, –countable)

These are then the possible 'extended projections' of a noun where 'extended projection' is understood in the sense of Grimshaw (1991) (i.e. as the projections from the lexical heads N and V and the functional categories associated with them).

[7] In the present paper we will not explore the syntax of the intensional domain: adjectives will be assumed to be heads in the extended projection from N (cf. Vangsnes 1999), merged lower than the functional projections (i.e. dominated by all of the functional projections of the same extended projection).

Examples from English of the different kinds of noun phrases are given in (34) – the list is of course not exhaustive.

(34)	a. the	chair	
	b. the	beer	
	c. all	chairs	
	d. all	beer	
	e. a certain	chair	
	f. a certain	beer	(i.e. 'a certain type/brand of beer')
	g. a/no	chair	
	h.	beer	

The construction of phrase structural objects otherwise observes the general principles of X′-theory. The order of the functional and substantive categories within an extended projection is fixed so that N is c-commanded by all the functional categories, ν is c-commanded by δ and κ but not by N, δ is only c-commanded by κ, whereas κ cannot be c-commanded by any of the other categories within the same extended projection.

This fixed order is semantically motivated: the category that yields a 'totality' reading and which determines the case for the noun phrase as a whole must c-command all the subparts; the category which yields a specific reading must 'specify' over both the cardinality and the intension of the expression; whereas the category which specifies the countability only need take scope over the intension of the expression.

In other words, we assume both that the inventory of functional categories is fixed and that their relative ordering is fixed, and that these facts both follow from semantic principles, presumably encoded in Universal Grammar.

The situation in the noun phrase also holds more generally in that I assume that *all* functional categories are headed by abstract heads, also those at the clausal level (i.e. C, AgrS, T etc.). It is beyond the scope of this article to consider the composition of clausal expressions, but some discussion can be found in Vangsnes (2001a, 2001b).

The presence of the abstract heads must be licensed through a process termed 'identification', essentially a visibility condition. Let us next turn to this notion.

3.2. *Identification and identifiers*

The principle of identification is stated as follows:

> **The Identification Principle**
> A functional category F must be identified in overt syntax.

The process of 'identification' is further defined as follows.

Identification$_{\text{def}}$
A functional category F in an extended projection P is identified iff a constituent of P that contains at least one feature relevant for F is merged in either the head or the specifier position of F.

Intuitively, then, 'words' and 'grammatical elements' serve to identify the abstract heads. Strictly speaking some version of the Identification Principle will apply to substantive categories also: a noun will identify N and an adjective A, but that is an issue peripheral to the following discussion.

There exist two kinds of features: lexical features and agreement features. This distinction resembles the [±interpretable] distinction in Chomsky (1995:277ff) and subsequent work: a *lexical* feature then corresponds to a [+interpretable] feature and an *agreement* feature to a [–interpretable] feature. As for the features relevant for identification of noun phrase internal functional categories we state the following:

The identification of noun phrase internal functional categories$_{\text{def}}$
(i) $\kappa_{[-\text{unique}]}$ must be identified by an element containing at least one of the following features: [MASS], [case], [N(OMINAL)],
(ii) $\kappa_{[+\text{unique}]}$ must be identified by an element containing at least one of the following features: [∀], [case], [N(OMINAL)],
(iii) δ must be identified by an element containing at least one of the following features: [gender], [DEIXIS], [N(OMINAL)],
(iv) ν must be identified by an element containing at least one of the following features: [num(ber)], [NUM(BER)], [N(OMINAL)].

For matters of exposition lexical features are rendered with small caps. As we see, then, [case] and [gender] are agreement features, whereas [N(OMINAL)], [∀], [MASS], and [DEIXIS] are lexical features. Number has a dual status with respect to this distinction: it may be both an agreement feature and a lexical feature. More specifically, certain determiners (e.g. numerals) carry a lexical number feature whereas other determiners (e.g. demonstratives and possessives) agree in number and their number feature therefore carries an agreement feature.

Agreement features may be considered attribute/value pairs in that they represent a property which has one of several possible specifications. The feature [case] will for instance be instantiated by one of the morphological cases which exist in the language in question, e.g. nominative or accusative depending on grammatical context, and likewise, [gender] will be instantiated by one of the gender categories found in the language, say masculine or feminine. Importantly, the specification of the features is not relevant for identification. Rather, it is the fact that they have a specification that matters. In other words, it is the presence of attributes that is important, not what values they have.

As for the lexical features, they are not attribute/value pairs. Rather, the feature carries a property as such.

The feature [∀] has a clear affinity to 'universal quantification'. However, in addition to being present in universal quantifiers such as *all*, *every* etc., I will argue that it also may be present in definite articles and demonstratives in the sense that they may yield also a 'totality' reading of noun phrases.[8]

Definite articles and demonstratives also contain the feature [DEIXIS] (henceforth '[DEIX]'), indicating that these are elements that 'point'. Importantly, there are demonstratives which do not carry the feature [∀]. In particular demonstratives like *such a* will be assumed to not carry the 'totality' feature whereas a demonstrative like *that* will. In section 4.6 we will also discuss the possibility that indefinite-*this* lacks the feature [∀].

As for the lexical feature [MASS] I will assume that it is found on mass determiners like *much* and *a lot of*. However, I will *not* assume that nouns that typically denote mass entities (such as *beer* and *lemonade*) are inherently specified for the feature (cf. the discussion in section 2.5).

The feature [N(OMINAL)] has a special status as it is the only one which is relevant for the identification of all the abstract heads. I will assume that [N] is present in nouns as well as all adnominal 'words', i.e. essentially in determiners and adjectives. It is not present in all adnominal constituents, and prepositional complements and relative clauses would be examples of modifying constituents that do not contain this feature. Hence, the feature is present only in those constituents of the noun phrase which are not themselves an extended projection, i.e. excluding phrases/extended projections such as PPs and relative clauses. Nouns, adjectives, and determiners are thus *in principle* capable of identifying any of the functional categories in the noun phrase.

Noun phrases may consist of more than one functional category (cf. (33)), and moreover they very often contain more than one constituent, for instance the noun plus one or more determiners. Since both nouns and determiners carry features that are relevant for the identification of functional categories, there is in principle a choice of identifier. In order for a choice to be made in a principled way we postulate the following preference algorithm:

Preferred identifier
When there are several candidate constituents for identifying the functional categories F^n of an extended projection P, the preferred identifier for F^n will be:

(i) the constituent containing the largest number of agreement features relevant for F^n; else, if there is no such constituent,
(ii) the constituent containing the largest number of lexical features relevant for F^n; else, if there is no such constituent,

[8] This may appear counterintuitive for readers familiar with formal logic, but it is in line with the view advocated by Milsark (1974, 1977) and Chomsky (1975) for capturing the linguistically relevant traits shared by universal quantifiers and definite determiners.

(iii) the constituent containing the smallest number of irrelevant lexical features for F^n; else, if there is no such constituent,
(iv) an X^0.

As we see, agreement features are more important than lexical features. The relevance of this point will not be demonstrated in any detail here, but as shown in Vangsnes (1999) it is arguably of great importance when one considers the Scandinavian noun phrase from a cross-dialectal point of view. On a general, conceptual note we may motivate the more prominent status of agreement features as follows: since the value of agreement features unlike lexical features is determined by some *other* constituent they constitute a redundant part of the system *a priori*, but if they do play a prominent and important role in the mechanics of syntax they would not be redundant after all.

3.3. *Identification in an articleless language*

We may then briefly consider how the identification of the functional categories takes place. Let us take some examples from Finnish, which is a language which lacks articles, both indefinite and definite ones. In this language a bare noun can be the single constituent of both an indefinite and a definite noun phrase, i.e. Finnish 'N' may correspond both to English 'a N' and English 'the N'. Accordingly, in isolation the sentence in (35) is ambiguous in that the car denoted by the object noun phrase can be a car which is either contextually given or not. The object noun phrase must carry partitive case since it is under the scope of negation,[9] and on the 'indefinite' reading it cannot be specifically referring according to the view of specificity advocated in section 2.4. ('ELA' stands for *elative* case.)

(35) En ostanut autoa Ruotsista. *Finnish*
NEG-1PL buy-AUX car-PAR Sweden-ELA
'I didn't buy the/a(ny) car in Sweden.'

The referent of the noun phrase will in both cases be countable, and under the present approach, the ambiguity of the object noun phrase is reflected in its phrase structure: on the 'definite' reading, the noun phrase contains the functional category δ (discourse anaphoric) and moreover $\kappa_{[+\text{unique}]}$ (uniquely referring), whereas it on the 'indefinite' reading contains $\kappa_{[-\text{unique}]}$, and *no* δ. Accordingly, the object noun phrase has either of the structures in (36).

(36) a. [KP $\kappa_{[+\text{unique}]}$ [DP δ [NumP ν [NP]]]] 'definite'
b. [KP $\kappa_{[-\text{unique}]}$ [NumP ν [NP]]] 'indefinite'

[9] This is a fact of Finnish object case marking. If there were no negation an alternation between accusative and partitive could be used to express telicity: an accusative object would entail a telic event whereas a partitive object would entail an atelic event. See Karlsson (1983) or Kiparsky (1998) for further discussion and references.

The noun *autoa* exhibits both case and number marking. Case is assigned by a noun phrase external Case assigner, and the actual case marking on the noun is therefore an instance of agreement. In turn the (partitive) case affix carries information about number, and I therefore take it to carry the features [case] and [NUM].

This means that the noun as a syntactic unit carries the lexical features [N] and [NUM], and these features are both relevant for the identification of ν. In both of the cases in (36) the derivation will start by insertion/merging of the noun under N. From this position it will head-move leftwards and adjoin to the functional categories one by one, and since it carries relevant features for each of them, the identification requirement on functional categories will be met. For the sake of completeness the derivation of the noun phrases can be represented as in (37). (Henceforth '[± unique]' will be abbreviated as '[± u]' whenever convenient.)

(37) a. [KP autoa$_i$ – $\kappa_{[+u]}$ [DP t_i – δ [NumP t_i – ν [NP t_i]]]]
b. [KP autoa$_i$ – $\kappa_{[-u]}$ [NumP t_i – ν [NP t_i]]]

Consider next the sentence in (38) where the object noun phrase contains a demonstrative.

(38) En ostanut tätä autoa Ruotsista. *Finnish*
NEG-1PL buy-AUX this-PAR car-PAR Sweden-ELA
'I didn't buy this car in Sweden.'

In this sentence the object noun phrase cannot have an 'indefinite' reading in the sense that the referent is not identifiable for the listener. (No indefinite-*this* reading is possible.) Still, when the sentence is interpreted in isolation this noun phrase is also ambiguous between two readings: the most salient reading is that the noun phrase refers to a particular, physically manifested car which is identifiable for the listener, but it may also refer to a particular *kind* of car. In other words we have a type/token distinction, and the same ambiguity is present in the English translation.

It nevertheless seems unnecessary to give the type/token distinction a syntactic correlate, and we therefore assume that the noun phrase has the structure in (39) on both readings.

(39) [KP $\kappa_{[+u]}$ [DP δ [NumP ν [NP]]]]

The demonstrative will be the identifier of the functional categories δ and $\kappa_{[+u]}$, i.e. preferred over the noun since it carries the lexical features [DEIX] and [$\forall$] in addition to the feature [N].

As for ν the noun is preferred over the demonstrative even though both contain the relevant features [num] and [N]. This is so since the noun carries fewer irrelevant features for this particular functional category. In

other words, the noun moves vacuously to ν as before, and we basically have the structure in (40) for the object noun phrase in (38).

(40) [KP tätä$_i$ – $\kappa_{[+u]}$ [DP t_i – δ [NumP autoa$_j$– ν [NP t_j]]]]

3.4. *The availability of elements*

The assumption that the feature [N] is relevant for identification of all of the noun phrase internal functional categories in principle yields a highly unconstrained system, but the discussion of Finnish has illustrated that the preference algorithm for identifiers represents an important constraint in that one will always have to consider which of the elements in a numeration it is that carries most features relevant for the identification.

Nevertheless, the system needs to be further constrained. An important question is why a demonstrative is not always merged in Finnish when the noun phrase in question is discourse anaphoric and uniquely referring: the demonstrative will always be a better identifier for δ and κ than the noun. Conversely, one may ask why an English noun by itself may never constitute the single constituent of a discourse anaphoric and uniquely referring noun phrase, i.e. why *car* can never mean 'the car'.

In order to achieve a better understanding of this let us first introduce a distinction between substantive and functional *elements*. This notion should not be confused with the distinction between substantive and functional *categories*. By a 'substantive element' we understand an element which contributes substantial denotative information of its own, whereas by a 'functional element' we understand an element which does not contribute substantial denotative information of its own.

As a first illustration N and A are substantive categories whereas for instance *cat* and *furry*, i.e. a noun and an adjective, respectively, are substantive elements which may instantiate the substantive categories N and A. Conversely, κ, δ, and ν are functional categories whereas for instance a definite article is a functional element, in fact the core instantiation of the functional category δ. In other words substantive/functional elements pertain to the lexical inventory whereas substantive/functional *categories* pertain to phrase structure.

Although demonstratives may instantiate a functional category on a par with a definite article they are *substantive* rather than functional elements. By hypothesis both demonstratives and definite articles carry the lexical feature [DEIX], but unlike an article, a demonstrative also carries information about the locational orientation of the speaker with respect to the referent of the noun phrase it is contained in. The article, on the other hand, entails something like a default deictic reading: it points

to a given referent, but rather than picking out one which is located in certain ways spatially, it merely picks out the most salient one in the discourse. Hence, we may say that a demonstrative contributes more denotative content than a definite article, and on the basis of that classify it as a substantive element.

Somewhat similarly we may say that the difference between the (singular) indefinite article and the numeral 1 is that whereas both express 'number', only the numeral expresses 'cardinality', and if we take only the latter to be a substantial denotative property the indefinite article will be a functional element whereas the numeral will be a substantive element.

The distinction between substantive and functional elements as conceived of here may appear more scalar than strictly dichotomous, and it seems difficult to give a clear definition of where on the scale the dividing line should be drawn, i.e. what it means to contribute sufficiently substantial denotative information to be considered a substantive element. Nevertheless, I do believe that languages make this distinction, and hopefully the intuition behind it will become clearer as we proceed.

With reference to the distinction between the two sorts of *elements* we now state the following principle, henceforth referred to as the 'Availability Principle'.

(41) **The Availability of Elements**
  (i) No substantive element can be merged post-lexically.
  (ii) Functional elements may be merged post-lexically and they are available for evaluation of economy throughout the derivation.

In order to understand this principle we must next define the notion 'post-lexical(ly)'. Intuitively it simply means 'after the lexical component has fed the syntactic component': we assume a lexicon and a syntactic component that consists of the generalized transformations 'move' and 'merge' which build up phrase structural objects according to X′-theory and other general and standard assumptions. Before the construal of the phrase structural object starts, i.e. before the derivation starts, all *substantive* elements are chosen from the lexical inventory.

Once the derivation has started, new substantive elements cannot be part of the derivation – the lexicon is shut off for such elements. Functional elements on the other hand may be picked from the lexicon after the derivation of the phrase structural object has started.

The essence of the Availability Principle is that once elements are chosen from the lexicon and the derivation of a phrase structural object has started, its semantic properties, both the extensional and intensional ones, are fixed and cannot be altered. The merging of a functional element will however not alter the semantic properties of the phrase structural

object since a functional element does not contribute any denotative information of its own.

Accordingly, whereas the first conjunct of the principle can be said to be both semantically and syntactically based, the second conjunct is a strictly syntactic one which captures the fact that languages may have elements which function as mere identifiers of functional categories and that these elements must be used whenever that yields the best numeration.

Returning to the question raised at the beginning of this subsection, the answer to why a demonstrative is not merged in Finnish whenever a noun phrase is discourse anaphoric and uniquely referring is that the demonstrative is a substantive element and therefore only present in the numeration if picked from the lexicon in the first place, i.e. if the phrase structural object should have a deictic/demonstrative reading.

On the other hand, the reason why a bare noun can never constitute the single constituent of a discourse anaphoric and uniquely referring noun phrase in English is that one in principle may always merge the definite article. And since such a numeration will contain a more ideal identifier for the functional categories involved, merging of the definite article will always take place. In other words, $\{car, \kappa_{[+u]}, \delta, \nu\}$ must be compared with $\{car, the, \kappa_{[+u]}, \delta, \nu\}$, and the latter is the best numeration.

Another important fact about the English article system which follows from the Availability Principle is that the definite article does not occur in a non-specific indefinite. This is so because a non-specific indefinite does not contain the functional categories δ and $\kappa_{[+unique]}$: for identification of ν and $\kappa_{[-unique]}$ the indefinite article is more appropriate than the definite article.

Articles are the most typical functional elements of the nominal system, and below we will discuss articles and article systems in more detail.

As for most other determiners they provide substantial enough denotative information of their own to be considered substantive elements. The universal quantifier *all* contributes the information that the referent of the noun phrase comprises a total entity. Possessives are of course substantive elements since they themselves have a referent that is distinct from that of the matrix noun phrase but nevertheless bears a certain relation to the latter. Numerals are substantive elements since they specify the cardinality of the set referred to by the noun phrase, arguably a denotative property.

Importantly then, most determiners cannot be merged postlexically – they must be part of the original numeration. Moreover, given that *substantive determiners* will often carry many of the same features as *functional determiners*, there may be cases where merging of the latter is rendered unnecessary. This is an important point to keep in mind, among other things in order to understand the interaction between demonstratives and definite articles.

3.5. *The complementary distribution of the definite and the indefinite articles*

An interesting question that arises is why the definite and the indefinite articles cannot cooccur in for example English. The examples in (42) illustrate this fact.

(42) a. *the a chair
b. *a the chair

Intuitively one would perhaps say that the two articles cannot co-occur since they have incompatible semantic properties (say 'definiteness' versus 'indefiniteness'), but under the present account they do not have semantic properties in and of themselves – rather, it is the functional categories which they identify that have the semantic properties. And consider then the fact that the English definite article does not show any variation in number.

(43) a. the car
b. the cars

(43) a. a car
b. cars

A natural conclusion to draw from this observation is that the definite article does not contain the feature [NUM]. The indefinite article on the other hand we would assume does, and the feature specifications of the articles would then be something like those given in (44).

(44) a. *the*: [DEIX]
[∀]
[N]

b. *a*: [NUM]
[N]

If we next evaluate the articles with respect to the Preference Principle, the indefinite article will be the preferred identifier for ν whereas the definite article will be the preferred identifier for δ. On that account, we could expect that both were merged, the indefinite article in NumP and the definite article in DP. However, as we know that does not happen: the two articles cannot co-occur (cf. (42) above).

On the other hand, the definite article does contain a feature relevant for the identification of ν, namely [N], and the solution to the problem is to capitalize on this fact: in a given numeration the various available functional elements are evaluated for all of the functional categories to be identified within the numeration. The question then is why the definite article is merged when a definite noun phrase is derived and the indefinite article when an indefinite one is derived; in other words when the structures in (45) are derived (assuming for the sake of the argument that the indefinite is one which is non-specifically referring).

(45) a. [KP $\kappa_{[+u]}$ [DP δ [NumP ν [NP]]]] (definite)
b. [KP $\kappa_{[-u]}$ [NumP ν [NP]]] ((non-specific) indefinite)

What we then find is that whereas the definite article is the preferred identifier for δ and $\kappa_{[+\text{unique}]}$, the indefinite article is the preferred identifier for ν and $\kappa_{[-\text{unique}]}$. As for $\kappa_{[-\text{unique}]}$ both articles contain one relevant feature ([N]), but since the definite article contains more irrelevant features, the indefinite article is the preferred one.

Summing up, the definite article is the preferred identifier for most functional categories in (45a) whereas the indefinite article is the preferred identifier for most functional categories in (45b), and this matches the empirical facts the structures are intended to cover.

It is also worth mentioning that under the present theory there is nothing peculiar about the fact that the definite article and the numeral 1 can co-occur, i.e. as for example in *the one car*. The numeral is quite close to the indefinite article with respect to the semantic properties its presence entails, but since it is not a functional element (it has denotative content of its own, i.e. 'cardinality') the Availability Principle does not bar it from being merged along with the definite article. On the other hand, the fact that the numeral 1 and the indefinite article cannot co-occur naturally follows from the fact that the indefinite article is not a better identifier than the numeral for any functional category.

The Availability Principle is thus of great importance for understanding the article system of a language. Let us next look at article systems in more detail.

4. Article systems

4.1. *'Function words' vs. 'functional inflection'*

Finnish is an articleless language with neither a definite nor an indefinite article, and we have given an account of how such a grammatical system is possible: if the language for some reason does not have elements with the special purpose of identifying particular functional categories, other constituents may very well do the job.

English is a language which has specialized identifiers, i.e. both a definite and an indefinite article. Icelandic in turn has a grammar different from both English and Finnish. This language has a definite article but no *indefinite* article. Moreover, the definite article appears as a suffix on the noun. The following example serves to illustrate these facts as it contains a definite subject noun phrase and an indefinite object noun phrase.

(46) Strákurinn kyssti konu útá götunni. *Icelandic*
boy-DEF kissed woman out-on street-DEF
'The boy kissed a woman in the street.'

The indefinite article in the English translation is obligatory, thus showing the absence of an obligatory indefinite article in Icelandic.

Furthermore, because of the absence of an indefinite article there are cases where a noun phrase consisting of just a singular indefinite noun will be ambiguous between a mass and a count reading. This is the case with the object in (47): it may be translated into English as either *beer* or *a beer*.

(47) Ég er með bjór í ísskápnum. *Icelandic*
I am with beer in fridge-DEF
'I have (a) beer in the fridge.'

The lack of an indefinite article sets Icelandic apart from the other Scandinavian languages and dialects, and in Vangsnes (1999:129f) I argue that the reason for this is that Icelandic unlike the other varieties[10] has morphological case marking on nouns. In addition to marking case the affix carries information about number: in other words it carries the lexical feature [NUM] as argued above for Finnish. In turn this renders Icelandic singular indefinite nouns just as suitable for the identification of ν as a potential indefinite article like the one found in languages like English (cf. the feature specification in (44b) above).

For this to make sense we must argue that singular indefinite nouns in other Scandinavian varieties do not contain the feature [NUM]. This is not entirely straightforward. Although the majority of nouns have singular indefinite forms that are identical to the inflectional stem, there are quite a few that do not. As an illustration consider the inflectional paradigm for two feminine and two neuter (Nynorsk) Norwegian nouns in (48) and (49), respectively.

(48) a.	*elv*	'river'	b.	*jent-e*	'girl'
	elv-a	'the river'		*jent-a*	'the girl'
	elv-ar	'rivers'		*jent-er*	'girls'
	elv-ane	'the rivers'		*jent-ene*	'the girls'
(49) a.	*hus*	'house'	b.	*gjerd-e*	'fence'
	hus-e	'the house'		*gjerd-e*	'the fence'
	hus	'houses'		*gjerd-e*	'fences'
	hus-a	'houses'		*gjerd-a*	'the fences'

By treating the vocalic endings in the b.-examples as an affix we get what most plausibly is the stem in the inflectional paradigm. One could argue that these vocalic endings are number affixes, the realization of *singular indefinite*. However, most nouns are like the ones in (48a) and (49a) in having no affix in the singular indefinite, and along the lines pursued in Vangsnes (1999:162) we will say that since the majority of nouns in the non-Icelandic varieties of Scandinavian does not have an affix in the singular indefinite *all* singular indefinite nouns behave as if there was no

[10] Faroese does have case inflection as well as an obligatory singular indefinite article. For discussion of and a solution to that problem see Vangsnes (1999:167ff).

marking of number, and in turn no [NUM] feature present. In other words the dominant morphological properties of the word class as a whole determine how all its members behave at the syntactic level.

In the same way we can account for the potential problem that several Icelandic indefinite singular nouns do not have an overt case affix. Consider the indefinite forms of the masculine noun *hestur* 'horse' and the neuter noun *hús* 'house' given in (50) where forms with no overt affix are rendered bold face.

(50)	a.			b.		
	hest-ur	'horse, NOM, SG'		***hús***	'house, NOM, SG'
	hest	'horse, ACC, SG'		***hús***	'house, ACC, SG'
	hest-i	'horse, DAT, SG'		*hús-i*	'house, DAT, SG'
	hest-s	'horse, GEN, SG'		*hús-s*	'house, GEN, SG'
	hest-ar	'horse, NOM, PL'		***hús***	'house, NOM, PL'
	hest-a	'horse, ACC, PL'		***hús***	'house, ACC, PL'
	hest-um	'horse, DAT, PL'		*hús-um*	'house, DAT, PL'
	hest-a	'horse, GEN, PL'		*hús-a*	'house, GEN, PL'

Although there exist forms of nouns in Icelandic that do not have an overt case marker, the great majority of noun forms do have such a marker which in addition to marking case also carries information about number. For this reason singular indefinite forms act uniformly as if they all had an affix – they are all specified as carrying the feature [NUM].

In turn this means that the Icelandic case affix serves as a functional element corresponding to the indefinite article in other languages. That does not seem far-fetched, especially not when we consider this in comparison with the Scandinavian *definite* article which, as we have seen, is an affix on the noun. We have already encountered instances of this element in the examples above, and for further illustration the full paradigm of the definite forms of the Icelandic nouns in (50) is given in (51).

(51)	a.			b.		
	hesturinn	'the horse, NOM, SG'		*húsið*	'the house, NOM, SG'
	hestinn	'the horse, ACC, SG'		*húsið*	'the house, ACC, SG'
	hestinum	'the horse, DAT, SG'		*húsinu*	'the house, DAT, SG'
	hestsins	'the horse, GEN, SG'		*hússins*	'the house, GEN, SG'
	hestarnir	'the horse, NOM, PL'		*húsin*	'the house, NOM, PL'
	hestana	'the horse, ACC, PL'		*húsin*	'the house, ACC, PL'
	hestunum	'the horse, DAT, PL'		*húsunum*	'the house, DAT, PL'
	hestanna	'the horse, GEN, PL'		*húsanna*	'the house, GEN, PL'

There is little doubt that we would treat the Scandinavian suffixed definite article as a functional element on a par with for example the English definite article. In other words, functional elements can exist in a language either as a separate syntactic unit (as a *function word*) or as an affix on a given syntactic unit, i.e. as a part of the inflectional system of that unit (as *functional inflection*).

The Availability Principle does not distinguish between these two forms

of existence for functional elements. This means that in the case of Scandinavian noun phrases that are discourse anaphoric and uniquely referring a numeration consisting of a definite noun will always be evaluated, and preferred, over a numeration consisting of the corresponding indefinite form of the noun. Technically speaking, this means that the principle does not count the indefinite and definite forms as different substantive elements: the principle refers to *lexemes* and *morphemes* rather than word forms and affixes (i.e. to types rather than tokens).

In summary, let us now compare the three article systems represented by Finnish (no articles), Icelandic (definite article, but no indefinite article), and English/Norwegian (both indefinite and definite article). (The Icelandic and Finnish examples are in the nominative.)

(52)	a. tuoli chair 'a chair'	b. tuoli chair 'the chair'	*Finnish*
(53)	a. stóll chair 'a chair'	b. stóllinn chair-DEF 'the chair'	*Icelandic*
(54)	a. *(en) stol a chair 'a chair'	b. stolen chair-DEF 'the chair'	*Norwegian*

By hypothesis the noun phrases in the a.- and b.-examples, respectively, have the same phrase structure across the three languages, i.e. the ones given above in (45) (still assuming that we have indefinite noun phrases with a *non-specific* reading).

In all cases just a noun would in principle suffice for identifying the functional categories since a noun carries the feature [N]. Just a noun is what we find in the Finnish examples, and it is the feature [N] that makes it suitable as identifier of δ, and [N] and [NUM] (the number marking on the case affix) as identifier of ν. (We will however revise this analysis of Finnish slightly in section 4.6.)

Just a common noun is also what we find in the Icelandic indefinite noun phrase: the noun itself carries the feature [N] and the case affix the feature [NUM]. In the case of the uniquely referring, discourse anaphoric noun phrase (i.e. the b.-examples) the definite form of the noun must be chosen from the numeration since the definite affix carries the feature [DEIX] which thereby renders the definite form a better identifier for δ than the indefinite.

Lastly, in the Norwegian pair of noun phrases in (54) functional elements must be generated in both cases: a definite form of the noun is preferred over an indefinite form for the same reason as in Icelandic, and since singular indefinite forms do not carry the feature [NUM] but the

indefinite article does, the indefinite article is a better identifier for ν than a bare noun.

The general conclusion from this comparison of Norwegian, Icelandic, and Finnish is then that if a language has articles it must use them when possible. The question of why some languages have articles whereas others do not is no doubt a diachronic one. In that respect, the present theory offers an account of why definite articles typically have developed from demonstratives and indefinite articles from the numeral 1, respectively. The etymological origins of the articles are ideal identifiers for certain functional categories (with a particular semantic content), and at some point in the evolution of a language they have been interpreted as 'mere' identifiers with no denotative content of their own. At that point they are no longer optional with respect to possible numerations and the evaluation of numerations – they have become functional *elements* which must be used whenever that gives the best numeration.

The general line of reasoning in this respect should not be very controversial. Many factors, ranging from language contact to 'accidental' phonological innovations, seem to be involved when we consider grammaticalization processes, and it is beyond the scope of the present article to discuss the evolution of article systems from a diachronic point of view. We will nevertheless touch on the issue again in later sections, albeit only in terms of speculative notes rather than in-depth studies.

4.2. *Partitive, singular and plural indefinite articles*

The examples in (55) illustrate the fact that both Norwegian and English have an indefinite article which is obligatory in a singular noun phrase which denotes a non-unique countable referent.

(55) a. Det er *(en) hund i hagen. *Norwegian*
EXPL is (a) dog in garden-DEF
b. There is *(a) dog in the garden. *English*

The obligatoriness of the English indefinite article is of course accounted for in the same way as for Norwegian: English singular nouns do generally not carry any affix, but are identical to the inflectional stem. Therefore they do not carry the feature [NUM].

Still, in plural indefinite noun phrases we do not find any obligatory indefinite article in either of the two languages.

(56) a. Det er hunder i hagen. *Norwegian*
EXPL is dogs in garden-DEF
b. There are dogs in the garden. *English*

The answer to why there is no obligatory indefinite article in the plural is straightforward: the great majority of English and Norwegian plural

nouns do carry an overt number marker, and we may therefore take it that plural nouns, unlike singular ones, *do* carry the feature [NUM]. That makes the plural nouns equally well suited as identifiers of ν as the indefinite article since both a plural noun and the indefinite article will carry the relevant features [N] and [NUM].

A language which requires an indefinite article both in the singular and the plural is French. Consider the following examples which are the French counterparts of the Norwegian and English examples in (55) and (56).

(57)	a.	Il y a	*(un)	chien	dans	le	jardin.	*French*
		There's	a	dog	in	the	garden	
	b.	Il y a	*(des)	chiens	dans	le	jardin.	
		There's	PL	dogs	in	the	garden	

As we see French has an obligatory singular indefinite article just like Norwegian and English, but unlike the latter two languages a determiner is also required in a plural noun phrase.

However, a crucial difference between French on the one hand and Norwegian and English on the other is that the distinction between singular and plural forms of nouns is seldom audible – there is no distinction in the pronunciation of *chien* and *chiens* in the examples in (57), and apart from a few idiosyncratic nouns (cf. e.g. Togeby 1965: 27f; Doetjes 1997:20f and references cited there), this reflects the general situation in French.[11]

On that basis, and given that it is the morphological properties of a class which are syntactically relevant (cf. above), we can argue that French nouns do not carry the features [NUM] – there is no plural affix which contributes this feature.

In turn, given the existence of an indefinite article where the [NUM] feature is specified 'plural', this will be the preferred identifier for ν in a noun phrase denoting a non-unique (non-specific) countable referent in French, and the ungrammaticality of the version of (57b) without the plural indefinite article is thus accounted for.

Unlike in Norwegian and English, French singular nouns cannot be the single constituent of a mass denoting noun phrase. This difference is illustrated by the examples in (58) and (59).

[11] A plural *-s* does however occur in so-called 'liaison contexts', i.e. where the plural noun is followed by an adnominal constituent (e.g. an attributive adjective) with an initial vowel, and one could therefore argue that there is an underlying plural affix which only surfaces in certain phonological contexts. In turn, however, one could argue that the affix surfaces so seldom that it is not syntactically relevant. This line of argument would then be on a par with the arguments for saying that instances of nouns with a singular affix in Norwegian (cf. section 4.1) are too few to mark singular nouns as carrying the feature [NUM]. Still, as pointed out to me by Frank Drijkoeningen (p.c.), the phenomenon of liaison is much more general and may in fact turn out to be determined by strictly phonological rather than morpho-syntactic factors.

(58) a. Il y a *(du) vin dans le réfrigérateur. *French*
There's ART wine in the fridge
b. Det er vin i kjøleskapet. *Norwegian*
EXPL is wine in fridge-DEF
c. There's wine in the fridge *English*

(59) a. Je veux *(du) vin. *French*
I want ART wine
b. Jeg vil ha vin. *Norwegian*
I want have wine
c. I want wine. *English*

The element whose presence is obligatory in the French examples is often referred to as the 'partitive' article, and etymologically (and perhaps synchronically) it consists of the partitive preposition *de* 'of' and the definite article: *du* is the masculine form which is derived from *de*+*le*, and the feminine form is *de la* which is etymologically transparent. Interestingly, the plural indefinite article also has a similar etymology: whereas the singular indefinite article is derived from the numeral 1, the plural indefinite article (*des*) is derived from *de* plus the plural definite article *les*.

I will argue that the synchronic feature composition of the partitive and the plural indefinite article in French reflects their common etymological origin. The preposition part of the articles, i.e. *de*, has contributed the lexical feature [MASS], a feature relevant for identifying $\kappa_{[-\text{unique}]}$. In this respect a parallel can be drawn to the English mass determiner *a lot of*: etymologically speaking this determiner consists of a set denoting noun phrase plus the equivalent of the French preposition *de*, i.e. *of*, and in most varieties of English this etymologically complex determiner has replaced the simplex mass determiner *much*. To say that it is *of* that contributed the mass feature of the complex determiner should be straightforward.

The definite article part of the French partitive and plural indefinite articles on the other hand contributed the features [N], [DEIX], and [∀], and we thus have the following feature specifications for the two articles:

(60) *du*: [N] *des*: [N]
[NUM] [NUM]
[DEIX] [DEIX]
[∀] [∀]
[MASS] [MASS]

We leave out the agreement feature [gen(der)] which the articles also contain, but which will not be of relevance to the present discussion.

To complete the picture the definite and the singular indefinite articles on the other hand have more or less the same feature specifications as their English and Norwegian counterparts, i.e. as shown in (61), again leaving out the agreement feature [gen].

(61) *le*: [N] [NUM] [DEIX] [∀] *les*: [N] [NUM] [DEIX] [∀] *un*: [N] [NUM]

Cutting the discussion short with respect to when the various articles are used, somewhat simplified we can say that the definite articles are used in the same contexts as the English definite articles, i.e. whenever a noun phrase is uniquely referring and discourse anaphoric. The singular indefinite article is used in the same contexts as the English singular indefinite article, i.e. when the noun phrase refers to a non-unique, non-specific set with one member. The plural indefinite article is used in contexts where bare plural nouns are used in English, and finally the partitive article is used where an English singular noun by itself constitutes a mass denoting noun phrase. Examples illustrating these facts are given in (62)–(64).

(62) a. Je veux **le** chien.
I want the dog
'I want the dog.'
b. Je veux **les** chiens.
I want the.PL dog(s)
'I want the dogs.'

(63) a. Je veux **un** chien.
I want a dog
b. Je veux **des** chiens.
I want PL dog(s)
'I want dogs.'

(64) a. Je veux **de la** bière.
I want MASS beer
'I want beer.'
b. Je veux **une** bière
I want a beer
'I want a beer.'
Je veux **la** bière.
I want the beer
'I want the beer.'

Notice that whereas the object noun phrase in (64a) must denote a mass and the one in (64b) a set, the object noun phrase in (64c) is ambiguous between either a mass or a set denotation: it can be the definite counterpart of either of the two former.

Recall now the setup in (33) for various types of noun phrases, the statement of which features are relevant for the identification of noun phrase internal functional categories, and the preference algorithm for identifiers, all repeated here for the sake of exposition.

(33) a. $KP_{[+u]}$ DP NumP NP (+unique, +discourse anaphoric, +countable)
b. $KP_{[+u]}$ DP NP (+unique, +discourse anaphoric, –countable)
c. $KP_{[+u]}$ NumP NP (+unique, –discourse anaphoric, +countable)
d. $KP_{[+u]}$ NP (+unique, –discourse anaphoric, –countable)
e. $KP_{[-u]}$ DP NumP NP (–unique, +discourse anaphoric, +countable)
f. $KP_{[-u]}$ DP NP (–unique, +discourse anaphoric, –countable)
g. $KP_{[-u]}$ NumP NP (–unique, –discourse anaphoric, +countable)
h. $KP_{[-u]}$ NP (–unique, –discourse anaphoric, –countable)

The identification of noun phrase internal functional categories$_{\mathbf{def}}$

(i) $\kappa_{[-\text{unique}]}$ must be identified by an element containing at least one of the following features: [MASS], [case], [N(OMINAL)],
(ii) $\kappa_{[+\text{unique}]}$ must be identified by an element containing at least one of the following features: [$\forall$], [case], [N(OMINAL)],
(iii) δ must be identified by an element containing at least one of the following features: [gender], [DEIXIS], [N(OMINAL)],
(iv) ν must be identified by an element containing at least one of the following features: [num(ber)], [NUM(BER)], [N(OMINAL)].

Preferred identifier

When there are several candidate constituents for identifying the functional categories F^n of an extended projection P, the preferred identifier for F^n will be:

(i) the constituent containing the largest number of agreement features relevant for F^n; else, if there is no such constituent,
(ii) the constituent containing the largest number of lexical features relevant for F^n; else, if there is no such constituent,
(iii) the constituent containing the smallest number of irrelevant lexical features for F^n; else, if there is no such constituent,
(iv) an X^0.

The only functional category present in an indefinite mass denoting noun phrase is $\kappa_{[-u]}$ (cf. (33h)), and in French a numeration containing the partitive article *du*/*de-la* in addition to a noun will be better than one containing just a noun because the partitive article is a functional element containing one more feature relevant for the identification of $\kappa_{[-u]}$ than the noun, namely [MASS].

Similarly, the plural indefinite article (*des*) is a better identifier for ν than a French noun: the plural indefinite article contains the feature [NUM] which is absent from French nouns. Moreover, the plural indefinite article also carries the feature [MASS], relevant for the identification of $\kappa_{[-u]}$. Accordingly, in French plural non-unique, non-specific set denoting noun phrases (cf. (33g)) there can never be just a noun: a functional element better suited for the identification of ν and $\kappa_{[-u]}$ is available, and must be merged.

A puzzling issue that arises from the present approach to the French

article system is the fact that the plural definite article (*les*) is a better identifier for ν than the plural indefinite article (*des*): the former carries equally many relevant features as the latter, but crucially *les* carries fewer *irrelevant* features than *des*. (The difference pertains to the lexical feature [MASS] which is present in *des* but absent in *les*.)

Still, the reason why *les* is not merged in a plural non-unique, non-specific, countable noun phrase (cf. (33g)) is that the plural indefinite article (*des*) is better suited for the identification of $\kappa_{[-u]}$ exactly because of its [MASS] feature. For this line of reasoning to hold we must argue that the presence of relevant features is more important than the presence of irrelevant features: the feature [MASS], which constitutes the only difference between the *les* and *des*, is irrelevant for identification of ν but relevant for the identification of $\kappa_{[-u]}$, and our conjecture is that relevance is more important than irrelevance.

On the other hand the reason why the definite singular article (*le*/*la*), and not the partitive one (*du*/*de-la*), is merged when a noun phrase is a uniquely referring, mass denoting, and discourse anaphoric noun phrase (cf. (33b) and the object noun phrase in (64c)), is that $\kappa_{[-u]}$ is not present, i.e. the functional category which the partitive article is especially well-equipped to identify.[12]

The French article system could certainly be discussed in greater detail, and it is worth pointing out that some of the issues discussed above would have been more straightforward had we not assumed that the etymological origin of the partitive and indefinite plural articles is fully reflected in their feature composition. That assumption is strictly speaking not necessary from a synchronic point of view, and if we for example took these articles to not contain the features [∀] and [deix], the discussion above would have proceeded differently – on that account the plural indefinite article (*des*), and not the plural *definite* article, would have been the preferred identifier for ν, and the solution to why *des* is the functional element chosen for a plural non-unique, non-specific countable noun phrase would be quite unproblematic.

Other issues would also have been more straightforward: for instance, the account of why the singular definite article is preferred over the partitive article as identifier of $\kappa_{[+u]}$ and δ would be less convoluted since the former (*le*) then would contain more relevant features than the latter (*du*) (i.e., it would not contain the features [DEIX] and [∀]).

Moreover, Helge Dyvik (p.c.) points out that it is quite plausible also from a diachronic point of view to say that only the feature [MASS] is 'carried over' to the new elements (i.e. the partitive and indefinite plural articles) since that then would amount to saying that only the feature of

[12] Moreover, although the partitive article carries features relevant for the identification of δ (e.g. [deix]), it also carries a higher number of irrelevant features for the identification of δ than the definite article.

the head of the original structure, i.e. the preposition *de*, is the one that 'survived'. In conclusion, it seems that an equally natural pre-theoretical assumption about the partitive and plural indefinite articles would have been that the only feature reflecting their etymology is the feature [MASS].

A highly interesting question is why and how French has developed its article system. It will lead too far here to discuss that issue, but it seems quite plausible that an important factor has been the loss of overt plural marking on nouns: once the plural marking was lost the singular and plural forms of count nouns would be homophonous, and the strategy of using the plural form of the noun in order to get an uncountable noun phrase was no longer available, so instead a different strategy, involving the use of a function word, emerged.

In the next subsection we will consider another language which also has a functional element similar to the French partitive article, but where the partitive article has a slightly different etymological origin, and synchronic status, than in French: it is in fact homophonous with the definite article.

4.3. *The Northern and Northeastern Swedish suffixed article*

Like Scandinavian dialects in general, Northern Swedish dialects and the Swedish dialects spoken in Österbotten on the western coast of Finland (henceforth 'North-Northeastern Swedish', abbreviated 'NN-Swedish') have an element suffixed to nouns which is present in noun phrases with a unique, discourse anaphoric referent, in other words a suffixed definite article. In for instance the examples in (65) and (66) the suffixed article is used in the same way in NN-Swedish and Standard Swedish, i.e. on a par with the English definite article in the translation.[13] The examples in (65) are from NN-Swedish and those in (66) from Standard Swedish.[14] We gloss the suffixed article as ART in the following discussion.

(65) a. Om **vårn** då **isn** jick bårt sä fann däm **flicka**.
about spring-ART when ice-ART went away so found they girl-ART
'In the spring when the ice disappeared they found the girl.'

b. **Bjärn'n** tågga **armen** hans nästan oppvä **axla**.
bear-ART bit arm-ART his almost up-with shoulder-ART
'The bear bit his arm almost by the shoulder.'

[13] The Standard Swedish article system parallels the Norwegian system in the relevant aspects, and what we have said about Norwegian earlier thus can be carried over to Standard Swedish.

[14] The NN-Swedish examples in (96) are excerpted from texts in the Skellefteå dialect in Dahlstedt and Ågren (1980), and the orthography is slightly altered.

(66) a. Om **våren** då **isen** gick hittade de **flickan.**
about spring-ART when ice-ART went found they girl-ART
'In the spring when the ice disappeared they found the girl.'
b. **Björnen** bet honom i **armen** nästan vid **axeln**.
bear-ART bit him in arm-ART almost by shoulder-ART
'The bear bit him in the arm almost by the shoulder.'

However, a fact which is fairly well known within traditional Swedish dialectology is that in NN-Swedish we also find uses of the suffixed article in noun phrases that are neither uniquely referring nor discourse anaphoric. As Delsing (1993:49ff) shows, in this use the suffix may appear on a singular noun which by itself constitutes a mass denoting (non-unique, non-specific) noun phrase, and the suffix may also appear on a plural noun which by itself constitutes a non-unique, non-specific, countable (plural) noun phrase. The phenomenon, which is not found in other Scandinavian dialects, can be illustrated by the NN-Swedish existential sentences in (67), in part excerpted from Delsing (1993). For comparison the Standard Swedish counterparts are given in (68).

(67) a. He jer mjölka/*mjölk i kylskåpe. *NN-Swedish*
EXPL is milk-ART/milk in fridge-ART
'There is milk in the fridge.'
b. He jer moröttren/*morötter i kylskåpe.
EXPL is carrots-ART/carrots in fridge-ART
'There are carrots in the fridge.'
c. He sitt en katt/*katta/*katt dera trappa.
EXPL sits a cat/cat-ART/cat on stairs-ART
'There is a cat sitting on the stairs.'

(68) a. Det finns mjölk/*mjölken i kylskåpet. *Standard Swedish*
EXPL exists milk/milk-ART in fridge-ART
'There is milk in the fridge.'
b. Det finns morötter/*morötterna i kylskåpet.
EXPL exists carrots/carrots-ART in fridge-ART
'There are carrots in the fridge.'
c. Det sitter en katt/*katten/*katt på trappan
EXPL sits a cat/cat-ART/cat on stairs-ART
'There is a cat sitting on the stairs.'

As we see from the c.-examples NN-Swedish and Standard Swedish behave alike with respect to disallowing the suffixed definite article on the noun in a singular set denoting noun phrase appearing in the postverbal argument of an existential sentence, but the a.- and b.-examples show that they differ with respect to singular mass and plural set denoting noun phrases.

Delsing (1993) considers the suffixed article in examples like these an element homophonous to, but lexically distinct from, the suffixed definite

article. The present theory however supports treating the suffix as just one element: a functional element which carries the feature [MASS] in addition to [DEIX] and [∀].

If we assume that the difference between NN-Swedish and the other Scandinavian dialects is the presence versus absence of the [MASS] feature on the suffixed definite article, the syntactic difference illustrated above will follow. In NN-Swedish a noun which carries the article will be a more appropriate identifier for $\kappa_{[-u]}$ than a noun which does not carry it, and accordingly merging just a noun without the suffixed article is never possible, neither in the singular nor in the plural.

As for Mainland Scandinavian in general (cf. the discussion of Norwegian in section 4.1.) we argue that NN-Swedish singular nouns do not carry the feature [NUM], and given the existence of an indefinite article in this dialect this is the reason why the indefinite article must be merged in singular non-unique, non-specific, countable noun phrases in NN-Swedish just as in Standard Swedish (cf. the examples in (67c) and (68c)).

The feature specifications for the NN-Swedish functional elements, i.e. the (lexical) indefinite article (cf. 67c) and the suffixed article, can then be given as follows:[15]

(69) *en*: [N] [NUM] *-a*: [N] [NUM] [DEIX] [∀] [MASS]

In other words, in the singular the NN-Swedish suffixed article corresponds to both the definite and partitive (lexical) article in French, and in the plural it corresponds to both the definite and the indefinite French article.

And somewhat similarly as in French, the singular suffixed article is not merged in a singular non-unique, non-specific, countable noun phrase since there exists an element that is more appropriate for the identification of ν, namely the indefinite article. Such a noun phrase will contain the following functional categories.

(70) [KP $\kappa_{[-u]}$ [NumP ν . . .

Although the singular suffixed article is a better identifier for $\kappa_{[-u]}$ it carries several features that are irrelevant for the identification of ν, whereas the indefinite article carries no irrelevant features for either of the two functional categories. Again we take it that relevance is more important than irrelevance, and hence the indefinite article will be merged in order to identify both ν and $\kappa_{[-u]}$.

[15] Again, as for French, we leave out the feature [gen(der)] since it is of no importance to the present discussion. This feature is otherwise present on both of the functional elements.

In my view treating the NN-Swedish suffixed article as one element with two identifiable uses is superior to acknowledging two separate but homophonous elements, say as on the one hand a definite article and on the other a partitive/plural indefinite article.

An interesting question is of course how the partitive/plural indefinite use of the suffixed article has arisen, but again that is beyond the scope of this article. It should however be noted that there is widespread homophony between indefinite and definite plural forms of nouns in NN-Swedish. The degree of homonymy ranges from being total (e.g. in the Karleby dialect, cf. Hagfors 1891:93) to being found in only a certain class of nouns (e.g. in the Västerbotten dialects, cf. Larsson 1929:113ff). One possible scenario is that the generalized plural affix first has been recognized as a suitable identifier for both δ and $\kappa_{[-u]}$, and that it therefore has been reanalyzed as carrying the feature [MASS] in addition to [DEIX] and [$\forall$]. In turn, the singular affix (i.e. the singular suffixed article) has been given the same feature specification as the plural affix. At that point the singular affix would be a better identifier for $\kappa_{[-u]}$ than a singular noun in a mass denoting, non-unique, non-specific noun phrase, and hence a noun carrying it would be merged rather than a noun not carrying it.

These speculations however require further studies of the NN-Swedish article system, and I leave that for future research.

4.4. *Finnish case marking*

Finnish, which lacks articles, may in some syntactic contexts mark the distinction between a mass and a set denotation of a singular noun phrase by case marking. When a predicate denotes a non-negated, telic state-of-affairs the postverbal argument (i.e. the direct object) shows an alternation between on the one hand partitive case and on the other accusative or nominative: if the noun phrase is uncountable it will carry partitive, otherwise it will carry accusative (if there is a nominative subject) or nominative (if there is no nominative subject, e.g. in passives, imperatives, unaccusative constructions, and constructions with an oblique subject). If the predicate is atelic or negated the postverbal argument always carries partitive case, and partitive case thus has both a 'VP-related function' and an 'NP-related function' – see Kiparsky (1998) and references cited there for details. The NP-related function is illustrated in (71) and the VP-related function is illustrated in (72).

(71) Ostin olutta / oluen. *Finnish*
bought-1SG beer-PAR / beer-ACC
'I bought beer/a beer (+ the beer).'

(72) a. Kirjoitin kirjettä. *Finnish*
wrote-1SG letter-PAR
'I was writing a/the letter.'

b. Kirjoitin kirjeen.
wrote-1SG letter-ACC
'I wrote a/the letter.'

In all of the translations I have included the possibility of a 'definite' reading, and this possibility cuts across the partitive/accusative alternation: in (71) both the mass denoting expression (i.e. *olutta*) and the set denoting expression (i.e. *oluen*) may have a unique, discourse anaphoric referent.

If we focus on the NP-related function it is interesting to observe that in the plural the distinction between partitive and accusative-nominative is sensitive to specificity/discourse anaphoricity in a more subtle way. In fact, whether or not a non-unique noun phrase is specific can be marked by case alternation: if it is specific it will carry accusative/nominative case, and if it is non-specific it will carry partitive. Consider the following examples.

(73) a. Ostin kirjeitä.
bought-1SG letters-PAR
'I bought letters.'
b. Ostin kirjeet.
bought-1SG letters-ACC
'I bought certain letters/the letters.'

On the telic reading of (73a) (i.e. dispensing with the non-salient atelic reading *I was buying (the) letters (over and over again)*), the object noun phrase can only be interpreted as having a non-unique, non-specific, countable referent. The object noun phrase in (73b) on the other hand is ambiguous between a non-unique, specific set reading and a unique specific set reading. Importantly then, the noun phrase *must* be discourse anaphoric (specific). If we then compare partitive singular and partitive plural of the NP-related function of partitive case, we see that its use is similar to the partitive and plural indefinite article.

Given the present theory an important difference between the NP-related use of partitive singular and partitive plural in Finnish is that partitive singular is used in noun phrases without the functional category ν, whereas partitive plural is used in noun phrases *with* this functional category. Let us now first conjecture that the singular partitive case suffix does not carry the feature [NUM] whereas the plural partitive case affix does so. This runs counter to what we said in section 3.3, but we will return to that problem below. Moreover, let us take it that both the singular and plural partitive affixes carry the feature [MASS], making them ideal identifiers for $\kappa_{[-u]}$. As for the accusative/nominative affixes we take them to carry [NUM], but otherwise nothing more. That would give us the following feature specification for accusative singular, accusative plural,

partitive singular, and partitive plural, respectively, for the Finnish noun *olut* 'beer' – we take the noun itself to contribute the feature [N].

(74)

oluen:	[N]		*oluet*:	[N]
(ACC.SG)	[NUM]		(ACC.PL)	[NUM]
olutta:	[N]		*olueita*:	[N]
(PAR.SG)	[MASS]		(PAR.PL)	[MASS]
				[NUM]

We can now calculate the NP-related use of case marking in Finnish in the light of the present theory. The various types of noun phrases we should consider are the following, where the case used is given in parentheses, and remember that we are only taking into account objective case marking in cases with non-negated, telic predicates.

(75)

a.	[KP	$\kappa_{[-u]}$	[DP	δ	[NumP	ν	. . . (ACC.SG or ACC.PL)
b.	[KP	$\kappa_{[-u]}$			[NumP	ν	. . . (ACC.SG or PAR.PL)
c.	[KP	$\kappa_{[-u]}$					. . . (PAR.SG)
d.	[KP	$\kappa_{[+u]}$	[DP	δ	[NumP	ν	. . . (ACC.SG or ACC.PL)
e.	[KP	$\kappa_{[+u]}$					. . . (PAR.SG)

The difference with respect to uniqueness does not have any consequences for the present discussion.

The reason why accusative is chosen in both the singular and the plural of discourse anaphoric (specific) noun phrases (75a, 75d) is that the partitive affixes carry a feature that is irrelevant for the identification of δ ([MASS]) whereas the accusative affixes do not.

The reason accusative singular is preferred over partitive singular in non-specific, countable noun phrases (75b) is that the singular accusative affix carries the feature [NUM], relevant for the identification of ν, whereas the singular partitive affix does not, and the fact that the singular partitive affix is a better identifier for $\kappa_{[-u]}$ (because of its [MASS] feature) does not affect the choice. Relevance is more important than irrelevance, and since the singular accusative affix is a better identifier for ν than the partitive, and the partitive better than the accusative for $\kappa_{[-u]}$, it is the irrelevance of the [MASS] feature that becomes the deciding factor.

However, in the plural things are different. The partitive plural affix carries the feature [NUM] and therefore has equally many features relevant for the identification of ν as the plural accusative affix, and the plural partitive affix is therefore better equipped to compete with the plural accusative affix: all in all the plural partitive is the preferred identifier since it carries more features relevant for $\kappa_{[-u]}$ than the plural accusative (because of the feature [MASS]) and since the two carry equally many features relevant for the identification of ν.

In mass denoting noun phrases partitive singular is of course chosen

because it carries the feature [MASS] and since all other forms will include some irrelevant feature for identification of $\kappa_{[-u]}$.

The question then is why partitive case, both singular and plural, can be related to other senses in the VP-related use. Why is it for example that an object under the scope of negation or in an atelic predicate must be partitive and not accusative no matter whether the noun phrase has a mass or set denotation?

The answer to this is that in such cases partitive case is an external requirement imposed on the noun phrase, and although it surely would be interesting to investigate this requirement further, the important point for the present discussion is that although partitive singular does not carry the feature [NUM], a partitive singular noun does carry a feature relevant for the identification of both ν and δ, as well as $\kappa_{[-u]}$, namely [N]. So in contexts where the use of accusative/nominative case is blocked for NP-external reasons, identification will thus nevertheless be ensured.

A broader investigation of the interaction between the NP-related and VP-related use of partitive case in Finnish lies beyond the scope of the present paper, but it surely constitutes an interesting object for further studies.

4.5. *A 'specific' article*

In section 2.4 we encountered a use of the English demonstrative *this* where the noun phrases in question is specifically but not uniquely referring. Let us consider the possibility that in certain varieties *this* is a functional rather than a substantive element with a feature specification different from both the indefinite and the definite article.

We may then compare the three elements, starting with the most straightforward case: (indefinite) *this* versus the indefinite article. If *this* carries the feature [DEIX] we immediately predict that it will be preferred over the indefinite article for identification of δ, and hence that it will be the chosen identifier in a non-unique, specifically referring noun phrase.

As for the difference between (indefinite) *this* and the definite article, consider the possibility that the former lacks the feature [$\forall$]. If so the definite article will be the preferred identifier in a noun phrase containing δ and $\kappa_{[+u]}$ whereas indefinite *this* will be the preferred identifier in a noun phrase containing δ and $\kappa_{[-u]}$, i.e. a non-unique specific noun phrase.

In turn, the indefinite article will be the preferred identifier in a (countable) noun phrase lacking the functional category δ. Summarizing, the three articles will have the following feature specifications.[16]

[16] We take *this* to carry the feature [num] since it agrees with the noun (*this* versus *these*). The definite article on the other hand is invariant with respect to number in English.

(76)

the:	*this*:	*a*:
[N]	[N]	[N]
[DEIX]	[DEIX]	[NUM]
[∀]		[num]

If such a system exists it sheds some light on the phenomenon of demonstrative reinforcement in some American English dialects discussed by Bernstein (1997) where a deictic adverb co-occurs with demonstrative. One of Bernstein's examples is given in (77).

(77) this here guy *colloquial American English*

If *this* has become a functional element in these dialects the complex, reinforced demonstrative may simply be a new substantial element replacing the original substantive *this*.

In fact, as shown by Haspelmath (1993:282f) the etymology of the English (and more generally West Germanic) proximal demonstrative *this* is an exact parallel to the (potential) grammaticalization process witnessed in the modern American English dialects: it emerged from the combination of the simple demonstrative *sa* (masc. sg. nom.) (with an initial dental fricative in many of the other case and gender forms) and the deictic particle *se*/*si*. The original demonstrative on the other hand became the Old English definite article, and later the present-day Modern English 'the'.[17]

5. Quantification, functional categories, and word order

In the introduction we noted the existence of English and Icelandic noun phrases where the order of constituents did not follow the general pattern on which we have developed the present theory. The relevant examples are repeated here.

(9) a. all three of the linguists *English*
b. *all three the linguists

(10) a. allir málvísindamennirnir þrír *Icelandic*
all linguists-DEF three
b. allir þrír málvísindamennirnir
all three linguists-DEF
Both: 'all the three linguists'

In (9a) the cardinal determiner precedes the definite article, and in (10a) the noun plus the suffixed definite article precedes the cardinal determiner, and hence these orderings do not conform to the general pattern. That of course does not necessarily mean that there is no underlying general pattern.

[17] The existence of complex demonstratives consisting of a determiner and a deictic adverb or clitic is moreover well known from other languages, and for discussion of such demonstratives in contemporary Scandinavian dialects see Vangsnes (1999:140ff).

An important point regarding the English example is that the presence of the 'partitive' preposition *of* is required. That could indicate that we are dealing with another construction, i.e. another kind of noun phrase than we have been discussing, perhaps a 'partitive construction' where *of* is a preposition taking a noun phrase complement, and where (9a) then would be comparable with noun phrases like the ones in (78).

(78) a. many of the linguists
b. some of the linguists
c. three of the linguists

On the other hand, and especially in the light of the discussion of the French article system above, it might be that both (9a) and the noun phrases in (78) have the same phrase structure as other noun phrases we have been discussing, more specifically where *of* is merged in K^0, adjoined to κ, and where the determiners have raised from Num^0 and adjoined to *of*. The structure for (78a) would then roughly be as in (79).

(79) $[_{KP}\,[_{K^0}\,many_i\text{–}of\text{–}\kappa\,[_{DP}\,[_{D^0}\,t_i\text{–}the\text{–}\delta\,[_{NumP}\,[_{Num^0}\,t_i\text{–}\nu\,[_{NP}\,linguists]]]]]]]$

There are certain attractive sides to such a line of reasoning. KP is by hypothesis associated with Case, and *of* being a preposition (the least marked one in English) surely could be regarded as a partitive Case marker, i.e. yielding a subset interpretation, not to be confused with morphological partitive case in Finnish. Moreover, the affinity between this partitive preposition and the partitive article in French, which is an identifier of κ and which is etymologically speaking derived from the French equivalent of *of* (i.e. *de*) is highly suggestive.

The reason why the cardinal determiner raises in this construction is then presumably scope – it must take scope over the whole of the KP in order to quantify over the denoted set.

Another phenomenon which involves the preposition *of* in English concerns the count/mass distinction. With a noun like *beer* we may easily get a set denotation by combining it with a simple determiner such as the indefinite article or a cardinal determiner. With other nouns that typically denote masses this is not equally straightforward, and a more complex expression is then needed to yield a noun phrase with a set denotation. Consider the following examples involving the noun *water*.

(80) a. I want (?*a) water.
b. I want (a glass of) water.

This case may seem unrelated to the ones in (9a) and (78), but under the present theory that need not be so. In (80b) there are two functional categories that must be identified, κ and ν, and a not altogether

implausible analysis is that the preposition serves as the identifier of κ, whereas the element yielding the set denotation, *a glass*, is a phrasal category merged in Spec-NumP where it serves as the identifier of ν. From Spec-NumP this 'piece maker' (or 'individualizer') subsequently raises to Spec-KP to take scope over the entity that it quantifies over. In other words it would roughly have the structure in (81).

(81) $[_{KP}$ [a glass]$_i$ $[_{K^0}$ of–κ $[_{NumP}$ t_i $[_{Num^0}$ ν $[_{NP}$ water]]]]]

At this point consider the possibility that also simple cardinal determiners originate as XPs in Spec-NumP on a par with the structure in (81), so that instead of (79) we have (79′).

(79′) $[_{KP}$ many$_i$ $[_{K^0}$ of–κ $[_{DP}$ t_i $[_{D^0}$ the–δ $[_{NumP}$ t_i $[_{Num^0}$ ν $[_{NP}$ linguists]]]]]]]

Continuing this line of reasoning I would then suggest that the analysis of (9a) is one where the universal quantifier and the cardinal determiner constitute a phrase initially merged in Spec-NumP, and that this complex phrase moves to Spec-KP thereby giving the word order where both of the determiner precede *of*.

(82) $[_{KP}$ [all three]$_i$ $[_{K^0}$ of–κ $[_{DP}$ $[_{D^0}$ the–δ $[_{NumP}$ t_i $[_{Num^0}$ ν $[_{NP}$ linguists]]]]]]]

In turn, given this analysis, the existence of the word order in (83a) can be given a straightforward account, namely the structure in (83b) where just the universal quantifier has moved to Spec-KP.

(83) a. all of the three linguists
b. $[_{KP}$ all$_i$ $[_{K^0}$ of–κ $[_{DP}$ $[_{D^0}$ the–δ $[_{NumP}$ [t_i-three] $[_{Num^0}$ ν $[_{NP}$ linguists]]]]]]]

Coming now to the Icelandic example in (10a) it is quite clear that the construction involves noun phrase internal movement of an XP which contains the noun and the suffixed definite article. Consider the examples in (84) – the example in (84b) is based on an example (without the possessive) given in Sigurðsson (1993:194).

(84) a. hinar þrjár frægu bækur **mínar** *Icelandic*
the three famous books my
b. frægu bækurnar **mínar** þrjár
famous books-DEF my three
c. *frægu bækurnar þrjár **mínar**
famous books-DEF three my
d. *bækurnar þrjár frægu mínar
books-DEF three famous my
e. þessar þrjár frægu bækur mínar
these three famous books my

These examples clearly show that the phenomenon involves phrasal movement, and moreover that the movement is obligatory in definites containing no lexical definite determiner. In the latter respect it should however be noted that the examples in (10) show that the movement is optional in noun phrases containing a universally quantifying determiner but no lexical definite determiner.

It would lead too far here to consider the phenomenon in detail, but as argued in Vangsnes (1999:144ff, 2001c) the XP-movement witnessed in Icelandic definites containing a numeral can be taken to represent a strategy whereby the identification of δ is met by phrasal movement to Spec-DP of a suitable identifier in cases where the numeral blocks head movement of the identifier. The obligatoriness of the movement in definite noun phrases with no universally quantifier and no lexical definite determiner suggests that this is the case, and moreover its optionality when there is a universal quantifier suggests that the quantifier may meet the identification requirement on δ: along with the analyses suggested for English above the universal quantifier is merged in Spec-NumP, and from this position it then moves via Spec-DP to Spec-KP, first identifying δ, then $\kappa_{[+u]}$.[18]

A phenomenon related to this, pointed out to me by Janne Bondi Johannessen (p.c.), concerns the fact that the Norwegian lexical definite article which normally must precede numerals in discourse anaphoric (and unique) noun phrases, is optional if there is a universal quantifier.

(85)	a.	*(de)		tre	lingvistene
		the		three	linguists-DEF
	b.	alle	(de)	tre	lingvistene
		all	the	three	linguists-DEF

Given that the universal quantifier originates in the NumP-domain and that it may identify δ on its way to the initial position, that may account for the optionality of the lexical definite article.

The question is of course what it is that makes the universal quantifier not only able to identify δ but also to yield a numeration that is equally good as one containing the definite article. In Vangsnes (2001c) I capitalize on the fact that the Scandinavian universal quantifier is inflected for gender and therefore carries an agreement feature relevant for the identification of δ ([gen]). Along with a more technical argument

[18] Þorsteinn Indriðason (p.c.) informs me that in universally quantifying noun phrases (with no lexical definite determiner) the 'heaviness' of the moved XP plays a certain role – movement seems to be dispreferred if the XP contains an adjective, but preferred if it does not, hence as exemplified in (i).

(i)	a.	allir	þrír	?(frægu)	málvísindamennirnir		*Icelandic*
		all	three	famous	linguists-DEF		
	b.	allir		(?frægu)	málvísindamennirnir	þrír	
		all		famous	linguists-DEF	three	

whereby δ is claimed to recover the feature [DEIX] of the suffixed article through c-command, an account of why the lexical definite article need not be merged is developed.

What is important for the present investigation is that the differences across languages with respect to the ordering of determiners do not constitute evidence against there being a universal ordering of functional categories. The surface ordering may be a result of crosslinguistic lexical differences as to what is the preferred identifier for a given abstract functional category in a given numeration.

6. Conclusion

The present article has advocated a theory of noun phrase structure in which referential properties are taken to be directly reflected in the phrase structure in the form of abstract functional categories whose presence imply the referential properties in question. I have discerned three such referential properties, and thus limited the inventory of noun phrase internal functional categories to three, albeit one of these (κ) exists in two varieties so as to allow us to preserve various formal syntactic requirements pertaining to the relation between noun phrases and the rest of the clause.

I have furthermore postulated a formal licensing requirement on the abstract functional categories, namely that they must be 'identified' by an element which carries relevant features, and the investigation of a variety of phenomena has given us an important understanding of the principles that govern the construction of noun phrases.

Such semantically founded principles presumably exist for the construction of phrase structural objects in general, and thus also for the clause. I have not investigated the clause in the light of Identification Theory here, but some attempts can be found in Vangsnes (2001a, 2001b). In fact, I believe that we find exact parallels to the nominal functional categories at the clausal level: $\kappa_{[\pm\text{unique}]}$ parallels [±finite] C ($\kappa_{[\pm\text{finite}]}$), δ (discourse anaphoricity) parallels event interpretations for the clause, located to AgrS, and countability has its parallel in tense anchoring (T).

Moreover, I have not fully investigated in what way noun phrases interact with other constituents at the clause level, notably the predicate, with respect to semantic properties. That has been a conscious choice: it is well known that the clausal status of noun phrases in many cases has important consequences for their semantic interpretation, but the present investigation has focused on the internal composition of noun phrases and how it follows certain principles irrespective of which clausal status the noun phrase has.

Nevertheless, the goal presented at the outset has been reached: we can model the referential properties of noun phrases in a manner which

reflects their internal syntactic structure, and such a theory of noun phrases does provide us with valuable insights. Importantly, we have developed an understanding of article systems whereby articles do not have semantic properties in and of themselves. Rather, they are licensers for abstract functional categories with semantic properties.

The present understanding thus lays the foundation not only for understanding how articles interact with other nominal constituents within a language, but also for how languages can differ with respect to the inventory of articles while at the same time following the same universal principles for the composition of noun phrase architecture.

References

ABBOTT, B. 1992. Some remarks on the notion 'specificity, ms., Michigan State U.

ABBOTT, B. 1994. Referentiality, specificity, strength & individual concepts, *Proceedings from the West Coast Conference of Formal Linguistics 12* 473–484.

ABBOTT, B. 1995. Some remarks on specificity, *Linguistic Inquiry* 26, 341–347.

ABNEY, P. S. 1987. *The English noun phrase in its sentential aspect*. PhD dissertation, MIT.

BARWISE, J. & COOPER, R. C. 1981. Generalized quantifiers and natural language, *Linguistics and Philosophy* 4, 159–219.

BELLETTI, A. 1988. The case of unaccusatives, *Linguistic Inquiry* 19, 1–34.

BERNSTEIN, J. 1997. Demonstratives and reinforcers in Romance and Germanic languages, *Lingua* 102, 87–113.

CHOMSKY, N. 1975. Questions of form and interpretation, *Linguistic Analysis* 1, 75–109.

CHOMSKY, N. 1995. *The Minimalist Program*, Cambridge, MA: MIT Press.

DAHLSTEDT, K.-H. & ÅGREN, P.-U. 1980. *Övre Norrlands bygdemål*, Cewe, Umeå: Skrifter utgivna av Johan Norrlander-sällskapet 2.

DELFITTO, D. & SCHROTEN, J. 1991. Bare plurals and the number affix in DP, *Probus* 2, 155–186.

DELSING, L.-O. 1993. *The internal structure of noun phrases in the Scandinavian languages*, PhD dissertation, U. of Lund.

DIESING, M. 1992. *Indefinites*, Cambridge, MA: MIT Press.

DOETJES, J. 1997. *Quantifiers and selection: On the distribution of quantifying expressions in French, Dutch, and English*. HIL dissertation 32, The Hague: Holland Academic Graphics.

ENÇ, M. 1991. The semantics of specificity, *Linguistic Inquiry* 22, 1–25.

FODOR, J. & SAG, I. A. 1982. Referential and Quantificational Indefinites, *Linguistics and Philosophy* 5, 335–398.

GIUSTI, G. 1991. The categorial status of quantified nominals, *Linguistische Berichte* 136, 438–454.

GIUSTI, G. & DIMITROVA-VULCHANOVA, M. 1996. Quantified noun phrase structure in Bulgarian, ed. J. Toman, *Formal approaches to Slavic linguistics: The College Park meeting*, 123–144, Ann Arbor, MI: Michigan Slavic Publications.

GRIMSHAW, J. 1991. Extended projections, ms., Brandeis U.

HAGFORS, K. J. 1891. *Gamlakarlebymålet: Ljud- ock formlära samt språkprov*, Stockholm: P. A. Nordstedt & Söner.

HASPELMATH, M. 1993. The diachronic externalization of inflection, *Linguistics* 31, 279–309.

Hellan, L. 1986. The headedness of NPs in Norwegian, eds. P. Muysken & H. van Riemsdijk, *Features and projections*, Dordrecht: Foris, 89–122.
Holmberg, A. & Platzack. C. 1988. On the role of inflection in Scandinavian syntax, *Working Papers in Scandinavian Syntax* 42, 25–42.
Hoop, H. de. 1992. *Case configuration and noun phrase interpretation*, PhD dissertation, U. of Groningen.
Karlsson, F. 1983. *Finnish grammar*, Helsinki: WSOY.
Kiparsky, P. 1998. Partitive case and aspect, ms., Stanford U., to appear in eds. M. Butt and W. Geuder, *Projecting from the lexicon*, Stanford: CSLI.
Lamontagne, G. & L. Travis. 1986. The Case Filter and the ECP. *McGill Working Papers in Linguistics*, 3, 51–75.
Lamontagne, G. & L. Travis. 1987. The syntax of adjacency, ms. McGill U. paper presented at West Coast Conference of Formal Linguistics 6.
Larsson, S. 1929. *Substantivböjningen i Västerbottens folkmål jämte en exkurs till ljudläran*, Appelbergs: Uppsala.
Löbel, E. 1990. D und Q als funktionale Kategorien in der Nominalphrase, *Linguistische Berichte* 127, 232–264.
Löbel, E. 1994. KP/DP-syntax: Interaction of case-marking with referential and nominal features, *Theoretical Linguistics* 20, 37–70.
Longobardi, G. 1994. Reference and proper names: a theory of N-movement in syntax and Logical Form, *Linguistic Inquiry* 25, 609–665.
May, R. 1985. *Logical Form*. Cambridge, MA: MIT Press.
Milsark, G. 1974. *Existential sentences in English*. PhD dissertation. MIT, publ. 1979. New York/London: Garland.
Mil[s]ark, G. 1977. Toward an explanation of certain peculiarities of the existential construction in English, *Linguistic Analysis* 3, 1–29.
Prince, E. 1981. On the inferencing of indefinite-*this* NPs, eds. A. Joshi, I. A. Sag, & B. L. Webber, *Elements of discourse understanding,* Cambridge U. Press, 231–250.
Ritter, E. 1991. Two functional categories in noun phrases: Evidence from Modern Hebrew, ed. S. Rothstein, *Syntax and Semantics* 26, San Diego: Academic Press, 37–62.
Sandström, G. & Holmberg, A. 1994. Adjective incorporation and the syntax of the Scandinavian noun phrase, *Reports from the Department of General Linguistics, University of Umeå* 35, 81–97.
Sæbø, K. J. 1999. *Free choice items in Scandinavian*, Göteborg U., NORDSEM Report 4.
Safir, K. J. 1985. *Syntactic chains*, Cambridge U. Press.
Safir, K. J. 1987. What explains the definiteness effect?, eds. E. Reuland and A. ter Meulen, *The representation of (in)definiteness*, Cambridge, MA: MIT Press, 71–97.
Shlonsky, U. 1991. Quantifiers as functional heads: a study of quantifier float in Hebrew, *Lingua* 84, 159–180.
Sigursson, H. Á. 1993. The structure of the Icelandic NP, *Studia Linguistica* 47, 177–197.
Szabolcsi, A. 1983. The possessor that ran away from home, *The Linguistic Review* 3, 89–102.
Togeby, K. 1965. *Fransk grammatik*, Copenhagen: Gyldendal.
Vangsnes, Ø. A. 1994. *The syntactic source and semantic reflexes of the definiteness effect*. Cand. philol. thesis, *Skriftserie fra Institutt for fonetikk og lingvistikk, Universitetet i Bergen* 48.
Vangsnes, Ø. A. 1995. Referentiality and argument positions in Icelandic, *Working Papers in Scandinavian Syntax* 55, 89–109.

VANGSNES, Ø. A. 1999. The identification of functional architecture, PhD dissertation, U. of Bergen.
VANGSNES, Ø. A. 2001a. Icelandic expletive constructions and the distribution of subject types, to appear ed. Peter Svenonius, *Subjects, expletives, and the Extended Projection Principle*, 43–69, Oxford: Oxford U. Press.
VANGSNES, Ø. A. 2001b, A verb's gotta do what a verb's gotta do! On Scandinavian infinitivals and the AGR parameter, to appear. eds. J. W. Zwart & W. Abraham, *Studies in comparative Germanic syntax*, Amsterdam: John Benjamins.
VANGSNES, Ø. A. 2001c. Passing by numbers: On phrasal movement in the Icelandic noun phrase (and related issues), ms., U. of Bergen/U. of Tromsø. Talk presented at the 16th Comparative Germanic Syntax Workshop in Montreal, 5 May 2001.

Received October 6, 1999
Accepted August 27, 2001

Øystein A. Vangsnes
Dept. of Linguistics
University of Tromsø
N-9037 Tromsø, Norway
oystein.vangsnes@hum.uit.no