

Essays on Nominal Determination

From morphology to discourse management

Edited by

Henrik Høeg Müller

Alex Klinge

Copenhagen Business School

John Benjamins Publishing Company
Amsterdam / Philadelphia

8TM The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Table of contents

Contributors	vii
The editors	ix
Introduction <i>Henrik Høeg Müller and Alex Klinge</i>	xi
Determiners and definiteness: Functional semantics and structural differentiation <i>Peter Harder</i>	1
Articles, definite and indefinite <i>Michael Herslund</i>	27
Typological correlations in nominal determination in Romance <i>Elisabeth Stark</i>	45
A stranger in the house: The French article <i>de</i> <i>Marc Wilmet</i>	65
Determination in endocentric and exocentric languages: With evidence primarily from Danish and Italian <i>Lørn Korzen</i>	79
Bare predicate nominals in Romance languages <i>Roberto Zamparelli</i>	101
Definiteness effect and the role of the coda in existential constructions <i>Manuel Leonetti</i>	131
Determination of N2 modifiers in Spanish nominal syntagmatic compounds <i>Henrik Høeg Müller</i>	163

Library of Congress Cataloging-in-Publication Data

Essays on nominal determination : from morphology to discourse management / edited by Henrik Høeg Müller, Alex Klinge.

p. cm. (Studies in Language Companion Series, ISSN 0165-7763 ; v. 99)

Includes bibliographical references and index.

1. Grammar, Comparative and general--Determiners. 2. Definiteness (Linguistics) I. Müller, Henrik Høeg. II. Klinge, Alex.

P299.D48E86 2008

415--dc22

2007052230

ISBN 978 90 272 3110 9 (Hb; alk. paper)

© 2008 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Reference to individuals, person, and the variety of mapping parameters <i>Giuseppe Longobardi</i>	189
English <i>th</i> -Forms <i>Judy B. Bernstein</i>	213
Stating the case for <i>p</i> -root and <i>hw</i> -root determiners <i>Alex Klinge</i>	233
On certain differences between noun phrases and clauses <i>Naoki Fukui and Mithoko Zushi</i>	265
Determination, nominalisation and conceptual processing <i>Helle Dam-Jensen</i>	287
The semantics and pragmatics of the possessive determiner <i>Georges Kleiber</i>	309
Reference, determiners and descriptive content <i>Thorstein Fretheim and Nana Aba Appiah Amfo</i>	337
Index	365

Typological correlations in nominal determination in Romance

Elisabeth Stark
University of Zurich

This paper discusses divergences and significant typological correlations found in the family of Romance languages, specifically French, Italian and Spanish. It proposes to reinterpret the complex system of *indefinite nominal determination* in two central Romance languages, viz. French and Italian, which both feature an indefinite article and a partitive article, as a device of *nominal classification* in a broad sense, marking the conceptually important distinction between a single, contoured referent and a non-contoured substance. It is argued that this classification system arose when nominal declension in Latin, which differentiated these two referentially highly relevant cognitive concepts via overt gender and number affixes, was partially or completely lost. In contrast to modern central Romance languages, which require obligatory (*indefinite*) determination in almost every argument position, modern peripheral Romance languages like Romanian or Spanish, possessing a simpler and more flexible system of determination, developed a system of *differential object marking* in order to unambiguously indicate contoured and highly individualized referents in direct object position.

1. The problem: Different systems of indefinite nominal determiners in Romance noun phrases

Despite some well-known and fruitful generalizations and hypotheses assuming homogenous semantic and syntactic systems of nominal determination for all Romance languages (e.g. Chierchia 1998, Longobardi 2001), the data in (1) demonstrates that there is considerable variation:

- (1) a. Sp.: *Has visto *(un) águila?*
Fr.: *As-tu vu *(un) aigle?*
It.: *Hai visto *(un) aquila?*
Rom.: *AI văzut (un) vultur?*
(Did you see an eagle?)

- b. Sp.: *Compro pan.*
 Fr.: *Jachète *(du) pain.*
 It.: *Compro (del) pane.*
 Rom.: *Cumpăr (niște) piine.*
 (I buy (some) bread).
 c. Sp.: *Me falta agua.*
 Fr.: *Il me faut *(de l')eau.*
 It.: *Mi occorre (dell')acqua.*
 Rom.: *Îmi trebuie (niște) apă.*
 (I need (some) water).
 d. Sp.: *Demostro paciencia en esta situación.*
 Fr.: *Elle montra *(de la) patience dans cette situation.*
 It.: *Dimostrò *(della) pazienza in questa situazione.*
 Rom.: *Demonstră răbdare în această situație.*
 (She showed patience in this situation)
 e. Sp.: *Veo (a unos) estudiantes en el edificio.*
 Fr.: *Je vois *(des) étudiants dans le bâtiment.*
 It.: *Vedo (degli) studenti nell'edificio.*
 Rom.: *Văd (niște) studenți în clădire.*
 (I see (some) students in the building)
 f. Sp.: *Salen estudiantes del edificio.*
 Fr.: *Il sortent *(des) étudiants du bâtiment.*
 It.: *Escono? (degli) studenti dall'edificio.*
 Rom.: *Îs studenți din clădire.*
 ((Some) students leave the building)

In Spanish, French, Italian and Romanian, indefinite nominals with existential reading show quite heterogeneous characteristics in argument position. Table 1 presents an overview over the three most frequent and grammaticalized indefinite determiners used with indefinite nominals in argument position:¹

Bare noun phrases in argument position occur in Spanish, Italian and Romanian under restricted grammatical conditions: in fact, only abstract nouns can appear freely in bare noun phrases in argument position even in the singular (cf. 1d), bare plurals surface postverbally in subject and object position independently of the lexical category of the noun (normally with non-specific interpretation of the nominal, cf. 1e and 1f). Bare singulars are also possible with "mass-denoting nouns" in postverbal subject and object position in Spanish, Italian and Romanian (see 1b and 1c), again with non-specific interpretation of the nominal. Conversely, the only Romance language which almost never permits bare noun phrases in argument position is French.

Table 1. Distribution of indefinite nominal determiners in four Romance languages

Spanish	French	Italian	Romanian
'Zero': abstract / "mass-denoting", rarely: "entity-denoting" nouns (non-specific); singular.	'Zero': rarely with abstract nouns (only in more or less idiomatic expressions).	'Zero': abstract / "mass-denoting" (non-specific); singular.	'Zero': abstract / "mass-denoting", "entity-denoting" (non-specific); singular.
Plural noun phrases (mostly postverbally).	No bare plurals.	Plural noun phrases (mostly postverbally).	Plural noun phrases (mostly postverbally).
<i>Uno</i> : singular countable noun phrases.	<i>Un</i> : singular countable noun phrases.	<i>Uno</i> : singular countable noun phrases.	<i>Un</i> : singular countable noun phrases.
No 'partitive article'.	<i>Du</i> : abstract / "mass-denoting" in non-countable singular noun phrases.	<i>Del</i> : "mass-denoting" in non-countable singular noun phrases.	No 'partitive article'.

As for overt indefinite nominal determination, every Romance language possesses a so-called indefinite article derived from the Latin numeral *unus*, 'one', which accompanies singular count noun phrases. 1a demonstrates that it seems possible for some Romanian speakers to accept even an "entity-denoting noun" like *vultur*, 'eagle', with a non-specific reading without *un*. Only two out of the four Romance languages discussed here, namely French and Italian, have a further indefinite determiner, the so-called 'partitive article', derived from the composition of Latin *de* and the definite article. It marks indefinite non-countable singular noun phrases², usually with "mass-denoting nouns", in pre- and postverbal subject and object position. In addition, it is obligatory with abstract nouns in French and optional in Italian (see examples in 1b, 1c and 1d).

In describing the facts in these terms, following Löbel (1993: 192ff) and with reference to Gil's (1987) typology, I assume a fundamental difference between the lexical categories "mass-denoting", "entity-denoting" and "abstract noun" (N), which derive from characteristics of the potential (extra-linguistic) referents (additivity, divisibility and so on) and which are based on denotational properties of the head noun, and the countability or non-countability of entire noun phrases. This last opposition is a grammatical category or a syntactic feature depending on the internal syntactic structure of the noun

1. Including 'zero' as a possible null determiner for the sake of a similar underlying syntactic structure (cf. Longobardi this volume), let us tentatively suppose, then, that mass/plurals, unlike singulars, can be introduced by an empty determiner.

2. I will not discuss the whole functional range and semantic properties of its morphological plural here, which seems to be the normal indefinite plural article and which is fully grammaticalized in French and optional in Italian.

phrase and it is characterized by the possibility of forming a morphological plural and/or to show compatibility with certain indefinite determiners (French/Italian: *uno* vs. *del* – NumP- or DP-level, see below). This assumption is justified by the fact that virtually any noun in Romance languages (like in any language with a grammaticalized countability distinction in this sense) can in principle appear in any kind of noun phrase:

- (2) a. Sp.: *Has comido águila?*
 Fr.: *As-tu mangé de laigle?*
 It.: *Hai mangiato (dell') aquila?*
 Rom.: *Ai mîncat (niște) vultur?*
 (Did you eat (some) eagle?)
 b. Sp.: *Compro un pan.*
 Fr.: *J'achète un pain.*
 It.: *Compro un pane.*
 Rom.: *Cumpăr un pîine.*
 (I buy one (a certain amount/piece of) bread)

Even if these examples seem semantically marked,³ due to prototypical affinities between “mass-denoting nouns” (like engl. *bread*) and non-countability, and between “entity-denoting nouns” (like engl. *eagle*) and countability (as already discussed for English in Allan 1980), they are grammatically well-formed and their ‘mass’ or ‘count’ readings derive exclusively from the pronominal indefinite determiners (‘zero’, ‘partitive’ or indefinite article).

2. An explanation proposal and its problems

In order to explain the striking differences between French and the other Romance languages concerning the possibility of permitting bare plurals or bare (‘mass’) singulars in argument position, the following correlation has often been observed (cf. e.g. Schrotten 2001): the loss of overt morphological number marking in nouns correlates with the necessity of number marking via determiners in spoken French.⁴

Le trait pertinent qui distingue [...] l'espagnol du français est la présence du nombre dans la prononciation du nom (Schrotten 2001: 196; similarly Wanner 2001: 1699).

3. Cf. Behrens 1995: 47–50, Corbett 2000: 86f.; see also the sortal interpretation or “Artenplural” mentioned by Krifka 1991: 414f. for “mass-denoting nouns” in countable plural NPs and the unique meaning of the morphological plural in languages with grammaticalized countability: it is always understood as additive, “diskrete Gesamtheiten von Objekten derselben Art” (Link 1991: 418).

4. Cf. also Delfino/Schrotten (1991: 157): “...and bare nouns cannot be interpreted since there is no number affix which can be raised to the D-position at LF”.

However, even if this correlation explains the degree of how obligatory explicit nominal determination in argument position is in French, it is not precise enough to explain the considerably different behaviour of the other Romance languages in this respect, as they exhibit morphological number marking also in their spoken varieties. Italian seems to behave in a particularly striking way when compared to Spanish or Romanian: Italian has a quite restricted distribution of bare NPs (recall 1a to 1f) despite overt morphological number marking, and like French but unlike Spanish or Romanian, it possesses a ‘partitive article’. Consequently, it may be inferred that there may be more to an adequate explanation than just the problem of overt morphological number marking in Romance.

These facts give rise to the following questions:

1. From a diachronic perspective how can the loss of nominal inflection and the rise of obligatory nominal determination be explained in conjunction with the ‘countability distinction’?
2. Why are there different systems of overt indefinite determiners (singular) in central Romance (French, Italian) vs. peripheral Romance languages (Romanian, Spanish)?
3. Can we find a relation between the two major syntactic innovations in Romance languages as compared to Latin, i.e. the introduction of obligatory and explicit (indefinite) nominal determination in argument position and the phenomenon of “Differential Object Marking” (DOM) (cf. Bosson 1997)?

The remainder of this article will propose some tentative answers to these questions. Sections 3 and 4 will deal with Latin and Romance nominal morphology and its possible implications for the distribution of bare noun phrases in argument position (questions 1 and 2). Section 5 will present and discuss again well-known correlations between the existence of some special indefinite determiners and the DOM-phenomenon in Romance, and then develop a new interpretation of the function as well as the potential origin of DOM in Romance languages (question 3). I will try to show that both the Romance systems of indefinite determiners and the phenomenon of “Differential Object Marking” can be considered as devices of nominal classification in a broad sense, replacing the ancient Latin nominal inflection that was partially or completely lost.

3. A look at Latin and Romance nominal inflection

3.1 Latin

Classical Latin possessed a complex declension system divided into 5 classes, which required obligatory and overt marking of the morpho-grammatical categories case, of which there are 5, gender, of which there are 3, and number, of which there are 2. Even

though clear-cut correspondences between gender, declension class and 'semantic' or lexical noun class cannot be assumed (in contrast to the situation suggested for Proto-Indo-European in Ralli 2002), there was some 'classification potential' for nouns sharing the same lexical root but differing in gender and/or number:

- (3) a. *caseus*, 'one single (piece of) cheese',
olea, 'olive' / 'olive tree'
 b. *caseum*, 'cheese as a substance',
oleum, 'oil'
 c. *acinus/acinum*, 'berry',
actina, 'grape'
frumentum, 'wheat',
frumenta, 'corn's

(3a) shows lexical roots with masculine and feminine gender, resulting in 'entity-denoting nouns', whereas the nouns from the same root in (3b) with neuter gender are "mass-denoting nouns". In addition, (3c) shows the well-known 'collective' semantics of the Latin neuter plural ending in *-a* (cf. Schön 1971, Windisch 1973). Although these oppositions are not systematic, the Latin neuter and especially the Latin neuter plural in *-a* – both unambiguously marked in spoken and written varieties – can be re-interpreted as a partly generalised 'classification system' denoting mainly the opposition between 'single, contoured object' (e.g. one piece of cheese, one olive, one berry) and 'non-contoured substance' (e.g. cheese, oil) or 'collective' (grape).

That this important semantic opposition is as much related to gender as to number is shown by the fact that, unlike the plural in modern Indo-European languages, including the Romance languages, the Latin plural is neither automatically interpreted as additive (cf. Link 1991) nor restricted to "entity-denoting nouns":

- (4) a. *frigora caloresque*, 'an intense heat and cold': plural indicating intensification
 b. *acquae*, 'waters', *ceruae*, 'wax tablets': different appearances of a substance⁵

Although the Latin plural can have a sortal reading, bare plurals of abstract or "mass denoting-nouns" are not automatically re-categorized as for instance in modern Romance languages (compare Fr. *huile*, 'oil', *des huiles*, 'different sorts of oil'). Virtually any Latin noun can be pluralized, and in fact frequent occurrences of plurals of "mass-

denoting nouns" or abstract nouns, as in (4) above, are attested.⁷ This fact, together with the absence of compatibility restrictions for (optional) indefinite determiners with nouns (Lat. *quidam* or *aliquis* combine freely with abstract, "mass-denoting" and "entity-denoting nouns") indicates that Latin had no grammaticalized "countability distinction" at the level of noun phrases (cf. Löbel 1993).

If Latin inflectional affixes, which mark declension class, gender and number indicate oppositions between 'contoured single object', 'substance' and 'collective', they can be considered as **classification devices** in the following sense: they are part of the universal dimension of **nominal apprehension**, which is a central universal operation of establishing *reference*:

First of all, so it seems, one has to be able to express that something is a thing [= dimension of APPREHENSION, E.S.]. Only then can it be named: The dimension of NAMING [...]. Following that, it can be referenced: The dimension of DETERMINATION. (Seiler 1986: 9)

APPREHENSION is the universal operational dimension with corresponding subdimensions which explicate the grasping and representation of concepts corresponding to objects or things by means of language. (Seiler 1986: 145)

Consequently, apprehension concerns chiefly the classification of the denotation of the noun phrase as "an undifferentiated concept or as an individual" (Lehmann 1991: 206, see also Meisterfeld 2000: 328).

3.2 French and Italian

The main morphological changes in the nominal system from Latin to Romance are well-known and comprise the complete loss of morphological case, a reduction of gender (especially the loss of the **neuter gender**), contrasted with a solid **formal preservation of number**, but with now only the 'additive plural meaning' left.

In addition, Modern Standard French shows the complete loss of the declension classes (already in Old French, cf. Delfitto/Schroten 1991: 180f.). Gender and number are usually marked (in the phonetic code) only by pronominal determiners:

- (5) *un ami/une amie* – *des ami(e)s*
 [cenami/ynami – dezami]
 'a male friend'/'a female friend' 'male or female friends'

5. Cf. in detail Hofmann/Szantyr ((1997) [1965]: 7-10), Meisterfeld (1998: 56ff.) and for late Latin analogical neuter plurals following the same pattern cf. Morani (2000: 228).

6. Cf. Kühner/Siegmann ('1955: 69, 73), Hofmann/Szantyr (1997) [1965]: 18, 21).

7. Cf. Iurrioz Leza (1986: 295f.): "This individualization strategy [= pluralization of abstract nouns, E.S.] is more widespread in the classical languages (Greek, Latin) than in modern German or any other European language; thus it is often difficult to translate an abstract [plural, E.S.] NP without changing its number: [...] *Asperitates viarum et angustiae* [...] 'The roughness(es) and narrowness(es) of the ways'".

8. Cf. Delfitto/Schroten (1991: 177f.).

The French noun [ami] is thus not phonetically marked at all for gender or number.

By contrast, Modern Standard Italian is different from French in having preserved 3 main declension classes, 2 overtly marked genders as well as overt number marking. However, the declensional endings *-a* and *-e* are far from being unambiguous markers of singular or plural, as they can either indicate feminine singular, (rarely) masculine singular or feminine plural (*-a*) or masculine singular or feminine plural (*-e*). The morpheme *-o* unambiguously indicates singular, but both masculine and (rarely) feminine gender:

- (6) Sg.: *-o/-a/-e*;
Pl.: *-i/-a/-e*;
- | | | |
|------------------------------|----------------------------|------------------------------------|
| a. <i>libro - libri</i> (m.) | <i>mano - mani</i> (f.) | <i>braccio - bracci</i> (m. - f.), |
| 'book - 'books' | 'hand' - 'hands' | arm' - 'arms'; |
| b. <i>casa - case</i> (f.) | <i>poeta - poeti</i> (m.), | |
| 'house' - 'houses' | 'poet' - 'poets'; | |
| c. <i>cane - cani</i> (m.) | | |
| 'dog' - 'dogs' | | |

What is marked in bold characters in (6a) is a residue of the original Latin classification potential of the neuter plural in *-a*, as opposed to a regular plural form in *-i* (originating in Late Latin, cf. Hofmann/Szantyr (1997) [1965]: 21) and reanalysed as feminine (but still plural!), always indicating a collective or at least 'pair' reading. Some nouns ending in *-o* (masculine singular), usually denoting concrete objects like body parts (It.: *ginocchio* 'knee', *orecchio* 'ear' and so on, also *muro* 'wall' etc.), have a plural form in *-a* when denoting a plurality, body parts or a 'collective reading'. However, they form a plural in *-i* when used metaphorically to denote something similar in form, but without a collective denotation (e.g. It.: *le braccia* denotes both arms of an animate being, whereas *i bracci* denotes the arms of a river; It. *le mura* denotes the townwall, whereas *i muri* denotes the single walls of a building).

3.3 Spanish and Romanian

Just like Italian, Modern (European) Standard Spanish has 3 main declension classes, 2 overtly marked genders and overt number marking. It is also "heterogeneous with respect to gender" (Harris 1992: 66ff.), but unambiguous with respect to number marking ("plurality is manifested consistently with the suffix *-s*", Harris 1992: 67):

- (7) Sg.: *-o/-a/-e*;
Pl.: *-s*;
- | | |
|-------------------------------|-----------------------------|
| a. <i>pas-o - pas-os</i> (m.) | <i>man-o - man-os</i> (f.), |
| 'step' - 'steps' | 'hand' - 'hands'; |
| b. <i>pas-a - pas-as</i> (f.) | <i>map-a - map-as</i> (m.), |
| 'raisin' - 'raisins' | 'map' - 'maps'; |

c. *jefe - jefes* (m.) *nube - nubes* (f.),
'chief' - 'chiefs' 'cloud' - 'clouds'

Unlike Italian however, there seems to be no 'classification potential' in nominal (declensional) endings in the morphological setup of Spanish nouns: the only slight 'classification potential' left in Spanish is a kind of 'neuter' (deriving from the Latin neuter singular) in the pronominal system, marking 'abstract antecedents', such as quotations, matters of fact, etc. Spanish personal pronouns and demonstratives show a threefold morphological opposition, with forms ending in *-e* for masculine singular, *-a* for feminine singular and *-o* for the so-called 'neuter' (e.g. span.: *este/esta/esto: lo que me interesa es esto...* 'what I am interested in is the following...' vs. *él que me interesa es este hombre* 'who I am interested in is this man').

Turning finally to Modern Standard Romanian, there exist 4 main declension classes, 2 overtly marked genders and, again, overt number marking. Almost like in Italian/Spanish, Romanian nominal endings are heterogeneous with respect to gender, but relatively unambiguous with respect to number marking and they seem to preserve a certain 'classification potential' within the so-called Romanian 'neuter' (cf. Windisch 1973, Herslund 1976):

- (8) Sg.: *-ul/-ă* [A]/ *-e/-K*;
Pl.: *-i/-e/-uri*;
- | | |
|----------------------------------|-------------------------------------|
| a. <i>membru - membri</i> (m.) | <i>teatr-u (m) - teatru-e</i> (f.), |
| 'member' - 'members' | 'theater' - 'theaters'; |
| b. <i>coleg-ă - coleg-e</i> (f.) | <i>săl-ă - săli</i> (f.), |
| 'colleague' - 'colleagues' | 'hall' - 'halls'; |
| c. <i>frat-e - frați</i> (m.) | <i>cart-e - cărți</i> (f.), |
| 'brother' - 'brothers' | 'book' - 'books'; |
| d. <i>coleg - coleg-i</i> (m.) | <i>caiet (m.) - caiete</i> (f.), |
| 'colleague' - 'colleagues' | 'booklet' - 'booklets' |
| | <i>tren (m.) - tren-uri</i> f. |
| | 'train' - 'trains' |

Singular nouns ending in *-u* (or consonant) with the respective plural in *-uri* (derived from the Latin neuter plural in *-ura*), and, less clearly, in *-e*, almost without exception indicate inanimate concrete objects or collectives, as opposed to nouns ending in *-i* (masculine or feminine plural).

4. Interesting correlations (1): Classification inside the noun phrase

How can we relate these morphological findings to the problem of the different indefinite determiner systems in the Romance languages? Let us summarize the main differences in the noun morphology of the four Romance languages investigated and look for possible correlations with the respective systems of indefinite determiners:

Modern standard French shows a complete reanalysis, a complete loss of the Latin neuter plural in *-a*: Lat. neuter plural *folia* becomes Fr. *la feuille*, feminine singular, just like Lat. feminine singular *femina* becomes Fr. *la femme*. This loss of the Latin 'classification system' via noun morphology is compensated for by the evolution of an obligatory 'classification system' via indefinite determiners (cf. Herslund 1998: 70ff.): 'zero' is practically excluded in argument position; the indefinite singular article, *un*, marks 'contour', 'individualized referent', and thus countability; the 'partitive article', *du*, marks 'substance', 'diffuse' (mass / abstract), and thus non-countability:

- (9) a. Lat.: *caseus*, Fr. *un fromage*: 'one single (piece of) cheese'
 b. Lat.: *caseum*, Fr.: *du fromage*: 'cheese as a substance'

Modern standard Italian shows some residue of the Latin neuter plural in *-a* with a certain 'classification potential', but it also has a French-like 'classification system' via indefinite determiners: 'zero' is partially permitted, but exclusively only for abstract / plural noun phrases; the indefinite article, *uno*, marks 'contour', 'individualized referent', and thus countability, just as in French; the 'partitive article', *del*, less grammaticalized than in French, marks 'substance' ('mass', as opposed to 'abstract') and thus non-countability.

The situation of these central Romance languages differs considerably from the situation in Modern Standard Spanish and Romanian. Here, we find partial preservation of the Latin neuter (in Spanish within the pronominal system, indicating 'abstract', in Romanian with the nominal 'neuter' meaning 'inanimate'/'collective'), but no clear-cut distinctions in the indefinite determination system. 'Zero' can mean 'abstract' / 'mass', even (rarely) 'entity-denoting', besides the additional possibility of marking non-specificity (cf. Laca 1999): the indefinite article, *un(o)*, less grammaticalized than in French or Italian, marks 'contour', 'individualized referent' and thus countability, but there is no explicit marking of non-countability and therefore no unambiguous simple obligatory classification system (cf. Herslund 1998: 70ff.).

What we can try to formulate now is an answer to the first question in section 2: The rise of indefinite determination in the Romance languages can be related to the loss of the complex Latin nominal morphology which indicated, among other things, the conceptually fundamental difference between a contoured and shaped individual, and diffuse substances/masses and collectives (the former Indo-European gender-based opposition between animate and inanimate entities had been obscured already in the Latin system). In Latin, the same lexical root could appear with different gender and number when denoting one or the other type of entities. The overall Romance indefinite article derived from the Latin numeral *unus*, 'one', originally seems to indicate 'nominal classification' in a broad sense, although at a higher level within the nominal's syntactic structure (probably NumP or PlP, cf. Delfitto/Schroten 1991 and especially Heycock & Zamparelli 2003), which indicates an (ongoing) 'countability grammaticalization' in Romance (see also the reduction of the different meanings of the Latin plural to an exclusively additive reading, cf. Meisterfeld 2000). Whereas Latin

nouns could have a phonologically expressed feature not only of (syntactic) PLURAL, just like most of the modern Romance languages, they could also have a phonologically expressed feature of 'semantic PLURAL' = COUNT, which covers mainly the difference between a countable and a non-countable interpretation of the NP. This second semantic feature has lost its overt phonological realization on the noun (N- or NP-level) via declensional endings, requiring thus an explicit indefinite determiner in a higher position than N in order to check the syntactic and semantic plural properties of the whole nominal.

Now we can also formulate an answer to the second question in section 2: (Un)ambiguous plural and thereby (non-)countability marking in Romance correlates with the presence or absence of a 'partitive classifier'.⁹ This is the main difference, for example, between the morphological set-up of Italian and Spanish nouns: Whereas the latter is marked by the overt, independent affix-like and unambiguous expression of [+PLUR], the former unambiguous plural-affix thus requiring explicit 'determination' via *uno* or *del* at least for the 'semantic plural-feature' [COUNT].

(Un)ambiguous plural and thereby (non-)countability marking does not correlate directly with obligatory determination. On the contrary, it is the complete loss of the Latin neuter that correlates with the development of an obligatorily explicit (indefinite) nominal determination (compare French with its necessity to mark both [PLUR] and [COUNT] or 'syntactic' and 'semantic plural' via determiners as against Italian, Romanian and Spanish).

With these findings in mind, we can now turn to the remaining question 3: How does 'Differential Object Marking', which exists in Spanish and Romanian, but not in French and Standard Italian, fit into this picture?

5. Interesting correlations (2): "Correlative typology" and classification outside the noun phrase

Körner (1987) observed a clear-cut correlation between the existence of a 'partitive article' and the existence of DOM in Romance. Whereas languages without "Differential Object Marking" like French, Occitan and Standard Italian possess a 'partitive article', the standard languages of Portuguese, Romanian and Spanish, together with several non-standard varieties and dialects, all of which show DOM, do not have any 'partitive article'.

9. Compare (Old) Spanish with (Old) French or Occitan, cf. Burdiant 2000: 73, 108, for Old French, Schultz-Gora 1973: 65ff., for Old Occitan, Zauner 1921: 56f., Penny 1993: 116-123, Lloyd 1987: 153f., for Old Spanish, see also certain Central-Southern varieties of Italian with an unambiguous morphological 'neuter' = mass declensional ending and without any partitive article, cf. Hall 1968, Delfitto/Schroten 1991: 167.

"Differential Object Marking" here means the phenomenon of selective prepositional case marking of the direct object according to certain lexical and / or semantic features of the intended referent or noun phrase:

- (10) *He visto (a) un hombre ingles con sombrero.*
(I saw an English man with a sombrero)

In (10), the insertion of a strongly favors a specific reading of *un hombre ingles*, whereas the noun phrase without *a* can only have a non-specific reading (cf. Leonetti 2003: 70–76, for a detailed discussion of *a* as a possible specificity marker in Spanish).

Now, based on this empirically valid observation that DOM correlates with the presence or absence of the partitive article, Körner (1987) interprets Sp. *a* or Rom. *pe* as markers of potential subject noun phrases (because of their lexical semantics or properties of their referents⁹) that have the function of direct objects in a concrete sentence, i.e. as a sort of structural or syntactic device of disambiguation (cf. Körner 1987: 42). And assuming a corresponding 'mirror function' to DOM, the (French) 'partitive article' would act as a marker of noun phrases that cannot be subjects. However, that this second generalization cannot be true is immediately shown by examples (11) and (12):

- (11) a. *Il y a de l'argent dans le portefeuille.*
b. **De l'argent est dans le portefeuille.*
(There is money in the wallet)
BUT:
c. *? Un franc est dans le portefeuille.*
(There is one franc in the wallet)
(12) *Du beurre était en train de fondre sur la table*
(Butter was melting all over the table)

(11c) demonstrates that the impossibility of putting *de l'argent* in preverbal subject position is not due to the determiner *du*, because *un franc* (with the French indefinite article *un*) is also odd in this position. Nonetheless, this restriction is not to be related to the kind of indefinite determiner in a subject NP, but to the kind of predication with a stative verb without any temporal specification or anchoring of the described event (compare (11c) to (12) which is perfectly fine, cf. Dobrovie-Sorin 1999: 173ff., Bosveld-de Smet 2000).

The following examples illustrate the potential selectional restrictions and/or the triggering factors for DOM in Modern Spanish:

10. Due to identical properties as to animacy etc. of 'I' and 'an English man' in example (10), both noun phrases, *pro* in subject position and *un hombre ingles*.. in direct object position could be subjects of a verb like Sp. *ver*, 'to see'.

- (13) a. *Busco a un camarero* (preferably specific)
b. *Busco un camarero*
c. *Busco (*a) camarero* (preferably non-specific) (cf. Leonetti 2003: 71)
(I am looking for a (new) waiter)
- (14) a. *Busco (*a) coche* (non-specific)
(I am looking for a car)
b. *Busco (*a) agua*
(I am looking for water)
- (15) a. *Está buscando a alguien*
(He is looking for somebody)
b. *No está buscando a nadie*
(He is not looking for anybody)
c. *No está buscando (*a) nada*
(He is not looking for anything)
(Leonetti 2003: 73)
- (16) *Un adjetivo califica a un sustantivo*
(An adjective modifies a noun)
(Torrego Salcedo 1999: 1788)
- (17) a. **No conozco a candidatos*
(I do not know any candidates)
b. *No conozco a candidatos con esas características*
(I do not know any candidates with these characteristics)
- (18) *Este profesor admite (*a) ofensas pequeñas*
(This teacher accepts small offences)

The examples in (13) show specificity effects of DOM in Spanish – in these sentences, the use of *a* is optional and marks or strongly favours a specific reading of *un camarero* ('a certain waiter'). (13c) and (14) – see also (17a) – show that *a* is impossible whenever an indefinite noun phrase is bare and not even modified by adjectives, relative clauses etc. – bare noun phrases like this are admitted in certain intensional contexts in Spanish, e.g. after the verb *buscar*, 'to look for', for "entity-denoting nouns" and for "mass-denoting nouns". This could indicate that *a* would be a specificity marker, but the examples in (15) and (17b) clearly show that (pro)nominals with a non-specific reading (e.g. *nadie*, 'nobody') or nominals without a clear-cut indication of (non-) specificity like *candidatos con esas características* in (17b) can, or even must, be marked by *a* – whenever they denote animate (human) beings. The fact that adjectival attribution does not in general save the construction is illustrated in (18): abstract nouns occurring as bare indefinites in direct object position can never be marked by *a*. Cases like (16) present serious difficulties for all theories that consider 'animacy' as the prime feature triggering DOM in Spanish; in fact, they seem to corroborate Körner's disambiguation hypothesis (see above): *adjetivo* and *sustantivo* denote both referents with

identical positions on any presumed 'animacy scale'¹¹, and *a* in front of *un sustantivo* seems to mark the direct object (recall the relatively free word order in Spanish).

Without going into detail (for recent accounts of DOM in Spanish and other languages see Torrego Salcedo 1999: 1784ff., 1790ff., Aissen 2003, Leonetti 2003, Naess 2004), it looks like "differential object marking" is always related to the relative degree of affectedness and control of the nominal arguments in a sentence (cf. Naess 2004). This is indicated by examples (13) to (18) above, and furthermore by several other factors governing DOM: DOM can in fact be obligatory with (animate) direct objects after certain verb classes (as with Sp. *atacar/insultar*, 'to attack', 'to insult' vs. optional DOM with Sp. *encontrar/ver*, 'to meet', 'to see') and it sometimes reflects specificity and/or topicality of the respective referents in direct object position (cf. Leonetti 2003: 76ff., Sora 2002: 360ff., and especially Farkas / von Heusinger 2003 for Romanian):

What is at stake here [...] is the emphasis on the individualization of the referent triggered by *a*, compared to the emphasis on quantity or descriptive content that predominates in unmarked objects. (Leonetti 2003: 80; similarly Torrego Salcedo 1999: 1789+1793ff.).

- (19) a. *Estaba dibujando a una niña*
(He was portraying a girl)
 - b. *Estaba dibujando una niña*
(He was drawing a girl)
- (Leonetti 2003: 80)

(19) presents a 'minimal pair' which clearly shows the semantic contribution of *a* to the sentence. Only when marked with *a*, does the direct object NP refer to a single, autonomous entity affected by the action that is denoted by the verbal predicate, whereas the omission of *a* in (19b) licenses a weak reading relating to something similar to 'semantic incorporation' (indefinite NPs with predicative readings, unspecified for animacy etc.; cf. Van Greenhoven 1998, Leonetti 2003; for similar remarks see also cf. Torrego Salcedo 1999: 1800, similar remarks concerning Romanian *pe* can be found in Sora 2002: 359+362f.). Moreover, the diachronic evolution of DOM in Spanish (cf. von Heusinger / Kaiser 2005) seems to confirm this interpretation: DOM starts with highly individualized referents denoted by personal pronouns or proper names, and subsequently (from the 12th century on) spreads towards definite topical NPs, before also marking indefinite specific or topical referents in the sense described above at a later stage. However, the current state of affairs is that DOM never marks non-specific, i.e. not clearly individualized or contoured referents (this also holds true for some modern varieties of American Spanish, which seem to allow DOM also with indefinite NPs referring to inanimate objects – they all appear at least in clearly countable NPs, cf. von Heusinger / Kaiser 2005). But even if DOM in Spanish maybe did not start out

as a marker of contoured referents in the first place, its actual distribution in indefinite object NPs could now be interpreted as a metaphoric shift from the original meaning with definite NPs denoting 'highly affected referent + maximally identifiable' (i.e. high on the definiteness scale) towards 'individualized, contoured referent' with indefinite NPs, in which classification is crucial, because in contrast to definite NPs the identification of the intended referent is impossible from the hearer's point of view.

Let us summarize: DOM in Spanish marks almost exclusively animate, more precisely non-abstract or non-mas-like referents in definite and / or specific (but not exclusively) and topical direct objects. Moreover, the higher a verb's transitivity ('other directed', cf. König 1999, Naess 2004: 1191), the more DOM becomes grammaticalized. In addition, whenever its use is optional, DOM marks single, individualized and autonomous referents with stable referentiality, i.e. referents that have to be considered at least as contoured, shaped entities (recall 16 and 19). These generalizations are also valid to almost the same extent for direct objects in Romanian (marked with *pe*), albeit exhibiting some syntactic differences (more obligatory clitic-doubling with DOM) as well as some special semantic triggering conditions like the denotation of N: With certain Ns denoting either individuals or social roles, DOM is disallowed with coinciding role-denotation:

- (20) *In America, data inchirietzi un apartament și ai vreo problema, trebuie să contactezi (pe) proprietarul.*
(In America, whenever you rent an apartment and have any problem, you have to contact the owner)

This also supports the interpretation of DOM in modern Romance languages as a classification device in a broad sense, indicating a 'contoured object'.

With these generalizations in mind, we can now try to give an answer to question 3 (recall section 2): DOM in peripheral Romance languages seems to be functionally parallel to the complex system of indefinite determiners in central Romance languages, at least in direct object position, which is the most important position for the development of determiners or nominal determination (cf. Leiss 2000). In addition, DOM might be understood – just like the opposition between 'zero', 'partitive' and 'indefinite article', e.g. in Italian¹² – in terms of 'nominal classification' and surfaces or specializes in this direction in exactly those Romance languages lacking a sufficient complex indefinite determiner system. DOM nowadays marks explicitly individualized referents; in Modern Spanish and Modern Romanian, *a* / *pe* encodes the "instruction to process

12. Compare the following examples from Delfino/Schroten (1991: 160), one without the 'partitive plural' and one with the 'partitive plural' in the direct object, clearly indicating shaped, individualized objects:

Gianni ha venduto libri solo per cinque minuti (with a possible meaning: 'Gianni has been a bookseller only for five minutes', even if he did not sell a single book) vs. *Gianni ha venduto dei libri per cinque minuti* meaning only 'Gianni has been selling some books for five minutes'.

11. Cf. von Heusinger / Kaiser 2005 for a discussion and presentation of possible 'animacy scales' for Spanish.

the object DP as a [...] prominent and referentially autonomous argument" (Leonetti 2003, 84) – via *classification* and/or determination.

6. Conclusion

The beginning of this article raised two points concerning current structural differences within the family of Romance languages despite their common historical development from Latin: Firstly, why is there considerable variation between the systems of indefinite determiners among Romance languages, and secondly, why is "differential object marking" found only in peripheral Romance languages? Since previous approaches based exclusively on overt number marking have to be considered insufficient, this paper presents a different, more comprehensive proposal. At first we reconsidered Latin nominal morphology and discovered a defective, but still functionally valuable system of 'nominal classification' via declensional affixes that make particular use of oppositions in gender and number (masculine/feminine vs. neuter, neuter plural in *-a* vs. other plural endings). The subsequent loss of this 'classification system' had different results in the Romance languages investigated: It led either to various complex systems of indefinite nominal determiners (where the Latin neuter has been almost completely lost, e.g. in French, and to a smaller extent, Standard Italian), or to the development of a device to differentially mark direct objects as autonomous, shaped entities vs. abstract, mass-like entities with special emphasis on their descriptive content (in languages that preserve overt number and even gender marking to some degree, e.g. Spanish and Romanian). What remains to be done now is a detailed diachronic description of the different stages of grammatical change from Latin to Romance from this new perspective, re-evaluating data from older stages of French, Italian, Spanish and Romanian, ultimately discovering possible grammaticalization paths related to the conceptually basic dimension of **apprehension**, or to put it more precisely, **nominal classification**.

References

- Aissen, J. 2003. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory* 21: 435–483.
- Allan, K. 1980. Nouns and countability. *Language* 56(3): 541–567.
- Behrens, L. 1995. Categorizing between lexicon and grammar: The MASS/COUNT distinction in a cross-linguistic perspective. *Lexicology* 1(1): 1–112.
- Bossong, G. 1997. Le marquage différentiel de l'objet dans les langues d'Europe. In *Actance et Valence dans les Langues d'Europe*, Jack Feuillet (ed.), 193–258. Berlin: Mouton de Gruyter.