

**INDEFINITE ARTICLES AND  
BEYOND**

Published by  
LOT  
Janskerhof 13  
3512 BL Utrecht  
The Netherlands

phone: +31 30 253 6006  
fax: +31 30 253 6406  
e-mail: [lot@uu.nl](mailto:lot@uu.nl)  
<http://www.lotschool.nl>

Cover illustration: fragment of the *Grammaire générale et raisonnée* by Lancelot & Arnauld

ISBN: 978-94-6093-022-5  
NUR 616

# **INDEFINITE ARTICLES AND BEYOND**

## **L'AU-DELÀ DES ARTICLES INDÉFINIS**

(avec un résumé en français)

## **ONBEPAAALDE LIDWOORDEN EN MEER VAN DIE DINGEN**

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor  
aan de Universiteit Utrecht  
op gezag van de rector magnificus, prof.dr. J.C. Stoof,  
ingevolge het besluit van het college voor promoties  
in het openbaar te verdedigen op woensdag 31 maart 2010  
des middags te 12.45 uur

door

Bert Simonne Walter Le Bruyn  
geboren op 30 september 1981 te Zoersel, België

Promotor: Prof.dr. H.E. de Swart

## TABLE OF CONTENTS

Acknowledgments	vii
<b>Indefinite articles and beyond</b>	<b>1</b>
1. Indefinite articles and what we know about them	1
2. Spanish <i>unos</i>	3
3. French <i>des</i>	6
4. Dutch <i>een</i>	8
5. Beyond this introduction	9
<b>Part I: <i>Unos</i> and article unite!</b>	<b>11</b>
General Introduction	11
Chapter 1: Bare plurals and <i>unos N</i>	13
1. The starting point	13
2. The semantics of <i>unos N</i> and the bare plural	14
3. Discourse salience	16
4. Discourse salience, <i>unos N</i> and the bare plural	19
5. Explaining the facts	22
6. Conclusion	25
Chapter 2: <i>Unos</i> as a plural indefinite article: synchronic evidence	27
1. Indefinite articles in the literature	27
2. <i>Unos</i> and the Blocking Principle	28
3. <i>Unos</i> and partitive readings	32
4. <i>Unos</i> and distributive readings	37
5. Further predictions	41
6. Conclusion	43
Chapter 3: <i>Unos</i> as a plural indefinite article: diachronic evidence	45
1. The grammaticalization of indefinite articles	45
2. The grammaticalization of <i>unos</i> : increase in frequency	46
3. The grammaticalization of <i>unos</i> : semantic bleaching	53
4. Simulating the evolution of indefinite articles	57
5. Conclusion	69
General Conclusion	71
<b>Part II: French <i>des</i> and articles: two worlds apart?</b>	<b>73</b>
General Introduction	73
Chapter 1: The semantics of <i>des N</i> : putting the pieces together	75
1. <i>Des</i> : a puzzling item	75
2. A bare partitive analysis for <i>des N</i>	77
3. The preference for narrow scope of <i>des N</i> and the anti-determiner line	87
4. The absence of bare plurals in French and the article line	89
5. Summary and outlook	90
Chapter 2: What about other bare partitives?	91
1. Recap	91
2. Beyond <i>des N</i>	92
3. Looking for bare partitives	93
4. An empirical investigation of bare partitives	99
5. The analysis of bare partitives	107
6. Conclusions and outlook	109

Chapter 3: Bare partitives and kinds _____	111
1. Introduction _____	111
2. The kind constraint _____	112
3. A note on <i>eat</i> -verbs _____	117
4. Summary and outlook _____	120
Chapter 4: Bare plurals and <i>des N</i> : a diachronic analysis _____	121
1. Introduction _____	121
2. The diachronic facts _____	121
3. Previous analyses _____	121
4. A novel way of looking at the evolution of <i>des N</i> _____	122
5. Simulating the evolution of <i>des N</i> : the software _____	124
6. Modelling the evolution of <i>des N</i> _____	131
7. Conclusion _____	136
General Conclusion _____	137
<b>Part III: An article gone astray _____</b>	<b>139</b>
General Introduction _____	139
Chapter 1: Article and Bare predication: A Dutch Case-Study _____	141
1. Introduction _____	141
2. Articles, bareness, capacities and kinds _____	141
3. De Swart, Winter & Zwarts (2007) _____	145
4. The coercion operators <i>kind</i> and <i>cap</i> _____	149
5. The link between article and kind predication _____	152
6. Taking stock _____	159
Chapter 2: Article and Bare predication: Extensions, limits and literature comparison _____	161
1. Introduction _____	161
2. Extensions _____	161
3. Limits _____	164
4. Literature comparison: Part I _____	166
5. Literature comparison: Part II _____	174
6. Summary and Conclusions _____	178
General Conclusion _____	179
<b>Concluding Summary _____</b>	<b>181</b>
1. Spanish <i>unos</i> _____	181
2. French <i>des</i> _____	183
3. Dutch <i>een</i> in predicate position _____	185
4. What we now know about indefinite articles _____	187
<b>Appendices _____</b>	<b>189</b>
1. Partitives and proper partitivity _____	189
2. The syntax of bare partitives _____	192
<b>References _____</b>	<b>195</b>
<b>Samenvatting in het Nederlands _____</b>	<b>203</b>
<b>Résumé en français _____</b>	<b>211</b>
<b>Curriculum Vitae _____</b>	

## ACKNOWLEDGEMENTS

I would like to start with a sincere thank you to a lot of people without whom this dissertation would simply not exist.

The first is my supervisor, Henriette de Swart. During the time I've known her she published / edited five volumes, guided at least 8 PhD students, raised 2 children and amazingly succeeded in always being there when I needed her. No matter where: KNG 29, *Orloff*, *De Poort*, Trans 8, conferences/workshops in Brussels, Paris, Oslo,... Next to being there when I needed her she also disappeared when I needed to think. And whenever I resurfaced she would then gladly comment on anything I came up with. No matter how unreal and no matter how often: the bare predication story you'll find in Part III of this dissertation is version 12.4bis. I thank her for her criticism, enthusiasm, support and trust. (do note that her name will pop up several times in the rest of these acknowledgements, this was just the research supervision part).

The second and the third (don't take these ordinals too seriously!) are Liliane Tasmowski and Brenda Laca who introduced me to DRT and GQT (or rather TRD et TQG) in Antwerp, supported me all along and always found time to listen to and shed new light on the issues I was struggling with. I often referred to them as my Belgian and my French supervisor and that's exactly the role they played.

I will stop using ordinals from now on, especially because the next people I want to thank form a group, *viz.* the Weak Referentiality group composed of Martin Everaert, Henriette de Swart (I told you she would pop up again!), Joost Zwarts, Ellen-Petra Kester, Min Que and Ana Aguilar. I would like to thank them for accepting me in their midst, sharing their thoughts with me on our bi-weekly meetings, commenting on draft versions of parts of this dissertation and allowing me to develop a broader view on linguistics. Special thanks also to the BA and MA students affiliated to the project: Marlies, Ingmar, Femke, Hanna, Tilly, Vera, Jolien, Martijn and Haifan.

Another group of people that I owe a huge debt of gratitude are all those working at the UIL-OTS. And when I say all, I mean all. From the cleaning lady with her beautiful smile that greets me every morning to Maaïke and the secretaries who are always there to help out PhD students that run into trouble. From all my fellow PhD students in the attic to all those trapped in Achter de Dom. Thanks to them all for being there, discussing, helping out, making life in Utrecht. A very special thanks to my paranimphs Gianluca and Sander.

For discussions and encouragement I gratefully mention my professors in Antwerp and Utrecht and the following people from everywhere: Jenny Doetjes, Radek Simik, Cristina Schmitt, Helen de Hoop, Mana Kobuchi, Evangelia Vlachou, Lucia Tovena, Donka Farkas, Jasper Roodenburg, Dorota Klimek, Albert Oosterhof, Celeste Augusto, Erik Schoorlemmer, Jan Schrotten, Fabienne Martin, Dany Jaspers, Guido Vanden Wyngaerd, Johan Rooryck, Lobke Aelbrecht, Karen De Clercq, Céline Signorini, Carmen Scarlat, Alexandru Mardale, Benjamin Fagard, Zlatka Guentchéva, Camelia Constantinescu and probably quite a lot of others that escape my attention for a second.

Part of my PhD life has also been about teaching. I enjoyed it a lot and this is due to the support I got from Henriette, all those involved in the French department in Utrecht, my co-teachers (Henriette (again!), Jan Schroten, Katell Lavéant, Marieke Schouwstra, Ana Aguilar and Joost Zwarts), Jenny Doetjes, Johan Rooryck and last but not least all my students. I thank them all for their trust, patience and encouragement.

Let me now go from my PhD life to my other life, that of a small boy born and raised in the Belgian Kempen in the beautiful village of Westmalle famous for its beer, cheese and cheerful people. I wish to thank my whole family for everything I owe them. A very special thanks to my parents: I cannot really calculate how often you made the trip Westmalle-Utrecht or Utrecht-Westmalle or Westmalle-train station or train station-Westmalle, how many hours you talked with me on the phone, how much support you gave me,... I do know that you used up a car on those trips, had quite impressive phone bills and were always there for me, no matter what. Thanks a lot! Another special thanks to my brother, sisters, their partners and children (especially my godson Stan): you're absolutely great and the last few months during which I had no time whatsoever to spend with you were hard. A very special place in my heart is also reserved for my godmother and her family who have always been there for me.

For the music in my life I want to thank all the members of the Ostinato choir. Thanks for making music with me and becoming my friends!

For the love in my life (the last but not least part) I owe a cute Chinese girl I met at the printer one day. Xiaolietje, thanks for everything!

## INDEFINITE ARTICLES AND BEYOND

### 1. Indefinite articles and what we know about them

What's there to be written about indefinite articles that we don't know yet? Well... a lot. To convince you of that let me first talk you through the things I think we know about them.

#### Indefinite articles are 'unmarked' argument markers

Indefinite articles are generally thought of as 'unmarked' argument markers in the sense that they mark arguments but don't add a lot of semantic content. For starters, compare the contribution of the indefinite *a* in (1) and that of the bare numeral *one* in (2):

- (1) A student came to see me.
- (2) One student came to see me.

Whereas (1) merely states that a student came to see me, (2) moreover suggests that some other student(s) didn't. Another indication of the unmarkedness of the indefinite article is that it's scopally free, i.e. it is free to take wide or narrow scope with respect to another scope bearing operator.

- (3) Mary didn't buy an apartment in San Francisco when she could have afforded it and now it's too late.
- (4) Mary didn't buy some apartment in San Francisco when she could have afforded it and now it's too late.

(3) has two readings: it either means that Mary didn't buy any apartment in San Francisco or that there is a specific apartment in San Francisco that she didn't buy. (4) only has the latter reading. What this shows is that the indefinite article differs from indefinite determiners like *some* in taking both narrow and wide scope with respect to operators like negation.

#### Indefinite articles 'block' bare nominal arguments

It is commonly assumed that a language that has articles doesn't allow for bare nominal arguments and vice versa. Think of a language like English and a language like Chinese: English has an indefinite article and doesn't allow for bare singular arguments (see (5)) whereas Chinese doesn't have an indefinite article but does allow for bare singular arguments (see (6)).

- (5) I saw \*(a) bear.
- (6) wò kànjiàn xióng le  
I see bear ASP  
I saw a | the bear

### Determiners don't get born as indefinite articles

Indefinite articles are generally thought to originate in the numeral *one* and to go through a number of stages. In the first stage they function as standard numerals, in the second they start to compete with bare nominals, marking pragmatically important arguments, and in the third and final stage they wipe out bare nominals from all argument positions. This is schematically represented in (7):

(7)

Stage 1	Determiner	Bare nominal
Stage 2	Determiner/Article	Bare nominal
Stage 3	Article	

I think the above three points summarize what we know about indefinite articles. Let me now return to my original question: is there more to be said about indefinite articles? The answer is affirmative: I think there are at least three more relevant questions to be asked about them. The first is what to do with determiners that are stage 2 of their 'article development' Are we supposed to analyze them as articles or are they just regular determiners? The second question is about the 'blocking principle'. This principle is supposed to account for a generalization, *viz.* that an article language does not allow for bare nominal arguments and vice versa. The question is whether this generalization actually holds. If the answer to the first question is that stage 2 determiners are articles, the generalization fails and the 'blocking principle' – in its current form – cannot be maintained. The third and final question is about the role of the indefinite article in predicates. Most people seem to assume that the role of the indefinite article is to mark arguments and one could therefore wonder why we find it so often in predicate position as well. These preceding questions are briefly summarized in (i) through (iii):

- |       |   |
|-------|---|
| (i)   | How should we define determiners that didn't evolve beyond stage 2? Are they articles or are they something else?                             |
| (ii)  | Does the 'blocking principle' hold? Is it the case that an article language does not allow for bare nominal arguments and <i>vice versa</i> ? |
| (iii) | People generally assume indefinite articles are argument markers. Why then do we find them in predicate position as well?                     |

These are the guiding questions this dissertation is about. It's about going beyond the neat cases and exploring the edges in order to get a better understanding of the notion *indefinite article*. Perhaps surprisingly, I'm not going to the end of the world to find interesting cases. I will stay rather close to home and look at Spanish *unos*,

French *des* and Dutch *een*. In this introduction I sketch why these cases are particularly relevant in the light of questions (i) through (iii). *Unos* will turn out to be a stage 2 determiner that will allow me to treat question (i) and part of question (ii). French *des* will turn out not to be an article but to block bare plural arguments anyway. This will be relevant for the *vice versa* part of question (ii). *Een* will turn out to be a prototypical indefinite article that shows a very clear semantic behaviour in predicate position and will therefore be a good case to study to tackle question (iii).

## 2. Spanish *unos*

For linguists working on diachrony *unos* is probably the most convincing plural indefinite article ever: morphologically it is the plural of the numeral *one* and both Lapesa (1975) and Pozas-Loyo (2009) have argued that it evolved in parallel to the singular indefinite article. Their most convincing argument is based on frequency data. The table in (8.A) shows the relative frequency of the singular indefinite article in the 13<sup>th</sup>, 15<sup>th</sup> and 17<sup>th</sup> centuries and the table in (8.B) shows the same for *unos* (data taken from Pozas-Loyo):

(8)

A.	
Century	Relative Frequency (per million words)
13 <sup>th</sup>	3280
15 <sup>th</sup>	3680
17 <sup>th</sup>	9813

B.	
Century	Relative Frequency (per million words)
13 <sup>th</sup>	226
15 <sup>th</sup>	253
17 <sup>th</sup>	680

As you can see the singular indefinite article and *unos* evolve in the same way: they are fairly stable up to the 15<sup>th</sup> century and then undergo an important increase in frequency. This can be taken as an argument in favour of the hypothesis that *unos* is the plural counterpart of the singular indefinite article.

Even though linguists working on diachrony might be convinced *unos* is an indefinite article, linguists working on synchrony are more reluctant. There are at least three reasons for this. The first is that *unos* doesn't seem to block bare plural arguments:

- (9)      María comió bizcochos.  
           María ate biscuits

(9) shows that bare plural arguments are grammatical in Spanish. In this sense *unos* would turn out to be a less prototypical indefinite article than its singular counterpart. The second reason *unos* is a less convincing indefinite article is that it cannot scope below negation:

- (10) A la reunión no asistieron unos profesores.  
at the meeting not assist UNOS professors

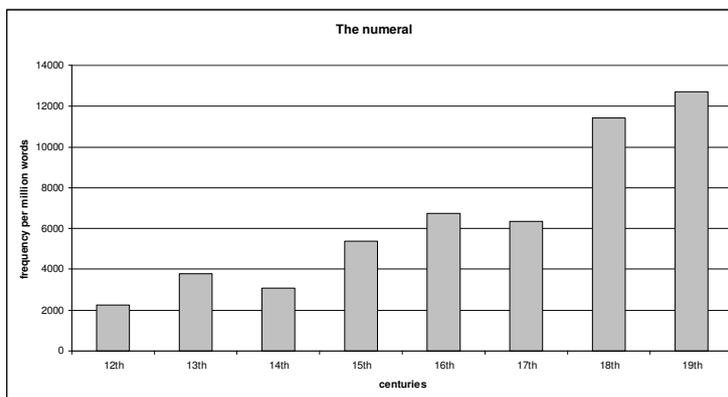
The only reading (10) can have is that there are some professors that didn't attend the meeting. The reading it crucially lacks is one in which no professors were present at the meeting. The third and final reason linguists working on synchrony might not be convinced *unos* is an indefinite article is that it has a clear preference for collective readings:

- (11) Unos hombres compraron un billete de lotería.  
UNOS men bought a ticket of lottery

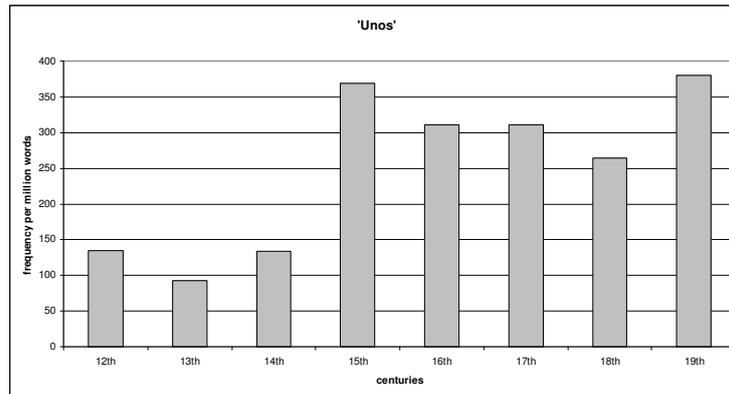
The strongly preferred reading of (11) is that one collective ticket was bought by some men. The distributive reading according to which each of the men bought a lottery ticket is strongly dispreferred. The point to be made is that if the indefinite article is semantically unmarked, one would expect *unos* not to have any preference. The fact that it does casts doubt on its articlehood.

What is interesting about *unos* is that there are clear arguments to say that it's not an indefinite article and that there are also clear arguments to say that it is. The question is whether we can tip the balance in one or the other direction. I believe we can. To give an idea of how I will do this I first have to give a fuller diachronic picture based on data drawn from the Corpus del Español:

- (12)



(13)



The graph in (12) shows the relative frequencies of the singular indefinite article throughout the whole history of Spanish, specifically adding data about the 18<sup>th</sup> and 19<sup>th</sup> centuries. The graph in (13) does the same for *unos*. What is remarkable is that the singular indefinite article underwent a second increase in frequency in the 18<sup>th</sup> century that has no counterpart for *unos*. This suggests that the singular indefinite article evolved further and that *unos* got stuck in the second stage of article development, *viz.* the one in which it is used to mark pragmatically important arguments.

With the new data in (12) and (13) in place I can explain how I will argue that *unos* should be analyzed as an indefinite article. The guiding questions will be the following:

- (14) Did *unos* undergo any semantic change from stage 1 to stage 2 that might indicate that it became an article?
- (15) Is there any semantic difference between *unos N* and the bare plural that cannot be explained in the pragmatics?
- (16) Is it possible that the development of indefinite articles stops at stage 2 for some determiners and at stage 3 for others?

As for (14), I will argue that *unos* did undergo a semantic change. More specifically, I will show that *unos* used to behave like other indefinite determiners in allowing a sentence like (17) to suggest that some students did not come, a suggestion it no longer has in present-day Spanish.

- (17) Unos estudiantes vinieron a verme.  
UNOS students came to see\_me

As for (15), I will argue that the two surprising properties exemplified in (10) and (11) – the fact that *unos* cannot scope below negation and has a strong preference for

collective readings – can be explained through its competition with bare plurals on the one hand and indefinite determiners on the other hand.

As for (16), I will show that the development of articles can be simulated in such a way that some articles go all the way to stage 3 and that others only go up to stage 2. What these answers amount to is that *unos* has a full article semantics and only differs from more prototypical indefinite articles because of its competition with bare plurals and other indefinite determiners. This is important for two of the general questions I intend to answer in this dissertation:

- (i) How should we define determiners that didn't evolve beyond stage 2? Are they articles or are they something else?
- (ii) Does the 'blocking principle' hold? Is it the case that an articleless language does not allow for bare nominal arguments?

As far as *unos* is concerned the answer to (i) is that stage 2 determiners are semantically full-fledged articles and should be analyzed as such. Their 'deviant' properties originate in their competition with bare plurals and other indefinite determiners. As for question (ii), the case of Spanish clearly shows that it makes no sense to assume that languages that have indefinite articles shouldn't allow for bare nominal arguments.

### 3. French *des*

For strong believers of the 'blocking principle' *des* is the prototypical plural indefinite article: French doesn't allow for bare plural arguments and forces the insertion of *des*.

- (18) \*Marie a mangé biscuits.  
Marie has eaten biscuits
- (19) Marie a mangé des biscuits.  
Marie has eaten DES biscuits  
Marie ate biscuits.

For those who take morphology seriously the articlehood of *des* is less straightforward: *des* is composed of *de* ('of') and *les* ('the') and is consequently more related to what is known as *bare partitives* than to indefinite articles. Bare partitives are expressions composed of the partitive preposition *of* and a definite DP that can appear in standard argument position. Examples are Dutch *van die N*, French *de ces N* and Italian *di quei N*, all of which can be glossed as *of* + demonstrative + N.

I will argue that *des N* originated as a bare partitive and never became a real indefinite article. If successful, this line of argumentation entails that there is no reason to assume that languages that don't allow for bare nominal arguments have indefinite articles. This complements the answer to question (ii) I formulated above: the 'blocking principle' does not hold in its current form, in whatever direction we interpret it.

In order to argue that *des* is more related to bare partitives than to indefinite articles, I will show that *des N* in general behaves in the same way as other bare partitives and that what we know about its diachronic behaviour can be explained without having to assume *des* became an indefinite article. I start by showing that it in general behaves like a bare partitive. To do so I will highlight some of the characteristics of bare partitives and then show how they carry over to *des N*.

What makes bare partitives special is that (a) their downstairs DP is necessarily interpreted as referring to a kind, (b) that their referent is necessarily plural and (c) that they can only take narrow scope. I will illustrate these properties for Dutch *van die N*.

The kind constraint on the downstairs DP of *van die N* can be illustrated on the basis of the following example:

(20)

A: Ik heb pas *De Slinger van Foucault* en *Baudolino* gelezen en ben nu bezig in *De Quincunx*.

B: Ik heb nog nooit [van die dikke boeken] gelezen, maar nu je het zegt, zou het eigenlijk nog een goed idee zijn om daarmee te beginnen.

A: I have just read *Foucault's Pendulum* and *Baudolino* and now I'm reading *The Quincunx*.

B: I have never read [of these thick books], but now that you come to mention it, it would be a good idea to do that.

If it were the case that *die dikke boeken* could refer to a specific set of books we would expect *van die dikke boeken* to preferably refer to a subset of the books introduced by A. This is however not the case; the only reading of B's utterance is that he might be interested in reading thick books. From this I conclude that the downstairs DP of *van die N* cannot refer to specific entities but only to kinds.

The plural interpretation of *van die N* can be demonstrated on the basis of agreement facts like the following:

(21) \*Er ligt [van die boeken] op de tafel.  
There is [of those books] on the table.

(22) Er liggen [van die boeken] op de tafel.  
There are [of those books] on the table.

The fact that a change in number of the verb in (21) makes the unacceptability disappear shows that *van die N* is necessarily interpreted as a plural.

The final relevant property of *van die N* is its narrow scope behaviour that I will demonstrate on the basis of (23):

(23) Marie heeft nog nooit van die ventjes gezien.  
Marie has yet never of those little\_guys seen  
<sup>ok</sup>Marie has never seen any of those little guys  
<sup>#</sup>There are some little guys that Marie hasn't seen

The only interpretation (23) allows for is that Marie has never seen any instantiations of the kind of little guys referred to by *die ventjes*. Crucially, (23) is incompatible with a situation in which Marie saw some but not all of the little guys referred to by *die ventjes*. This means that the bare partitive *van die N* cannot take scope over negation.

Let me now take a look at *des N* and see what predictions I make if I analyze it as a bare partitive. The first two predictions are that the downstairs definite DP of *des N* can only refer to kinds and that the referent of the whole bare partitive is necessarily plural. *Des enfants* ('of the children') would then refer to instantiations of the kind *children* which would be equivalent to what the bare plural *children* in English would mean. These predictions are borne out by the facts (see (19)). The third and final prediction is that *des N* should only take narrow scope. This prediction is correct to the extent that *des N* has a clear preference for narrow scope:

- (24) Il me faut des chemises.  
 There me must DES shirts  
 I need shirts.

The strongly preferred reading of (24) is that I need shirts, no matter which. This is – as predicted – the narrow scope reading.

From the preceding I conclude that *des N* can in general be analyzed as a bare partitive. Given its morphological make-up this is also its most likely origin. The next step in my argumentation is to show that what we know about the diachronic behaviour of *des N* can also be accounted for without having to assume that *des* became an indefinite article.

In order to argue that *des N* needn't have become an indefinite article I will show that its diachronic behaviour can be accounted for on the basis of a simulation that takes *des N* to be nothing more than a bare partitive. The simulation itself will be worked out in all its detail in the dissertation but the crucial insight is that what plays a role in the evolution of *des N* is not that *des* would have become an indefinite article but rather that *des N* can be truth-conditionally equivalent to the bare plural. The success of the simulation will be an important argument in favour of assuming that *des N* didn't become an indefinite article.

The conclusion I will draw on the basis of the discussion about *des N* is that *des* shouldn't be analyzed as an indefinite article. Given that French doesn't allow for bare plural arguments this allows me to conclude that languages that don't allow for bare plural arguments need not have an indefinite article. This crucially means that the 'blocking principle' – as it stands – cannot be maintained.

#### 4. Dutch *een*

Dutch *een* is a prototypical indefinite article: it is semantically unmarked, seems to block bare nominal arguments and has its origin in the numeral *one*. It is therefore one of the candidates to look at in order to answer question (iii):

- |   |
|---|
| (iii) People generally assume indefinite articles are argument markers. Why then do we find them in predicate position as well? |
|---|

The standard view on indefinite articles is that they are semantically empty in predicate position. This view has recently been challenged on the basis of contrasts like the following:

- (25) Henriette is manager.  
Henriette is manager  
(26) Henriette is een manager.  
Henriette is a manager

(25) has the unmarked interpretation that Henriette belongs to the set of appointed managers whereas (26) suggests that Henriette has the stereotypical properties of a manager (being organized, able to delegate, ...) without having to be one. To the contrast in (27) and (28) I add another one:

- (27) Jantje is een raaf.  
Little\_John is a raven  
(28) Jantje is raaf.  
Little\_John is raven

(27) has the unmarked interpretation that Jantje is a member of the kind *raven* whereas (28) is only interpretable if we assume Jantje is playing the part of *raven* in a game (e.g. the game *rats and ravens*).

When we try to make sense of the contrast in (25) and (26) on the one hand and (27) and (28) on the other hand it would seem that bare predication is used to refer to a position of an individual in a social setting, be it a professional or a game setting. Article predication on the other hand seems to be connected to more inherent properties of individuals: being appointed manager is less inherent than having the prototypical properties we associate with managers and playing the role of raven is also less inherent than belonging to the kind *raven*. The crucial point to be made on the basis of these minimal pairs is that article and bare predication seem to give rise to a systematic opposition between social and more inherent properties. This means that the indefinite article in predicate position cannot be semantically void and that our general unmarked semantics of indefinite articles has to be enriched in order to account for their behaviour in predicate position. In the third part of this dissertation I will do this for Dutch *een* and – as an extension – for English *a(n)*.

## 5. Beyond this introduction

In the preceding sections I have argued that Spanish *unos*, French *des* and Dutch *een* are particularly relevant in order to investigate the three questions that will guide me on my search for a more precise characterization of the indefinite article.

- |      |  |
|------|--|
| (i)  | How should we define determiners that didn't evolve beyond stage 2? Are they articles or are they something else?          |
| (ii) | Does the 'blocking principle' hold? Is it the case that an articleless language does not allow for bare nominal arguments? |

(iii) People generally assume indefinite articles are argument markers. Why then do we find them in predicate position as well?
---

I argued that *unos* is a determiner that evolved up to stage 2 and could therefore be studied to answer question (i). As for question (ii), I argued that both *unos* and *des* are relevant. If *unos* turns out to be an article, it makes no sense to assume that languages that have articles don't allow for bare nominal arguments. If *des* turns out not to be an article, it doesn't make any sense to assume that languages that don't allow for bare nominal arguments have an indefinite article. Put briefly: if *unos* is an article and *des* isn't, the blocking principle in its current form fails, no matter in which direction we interpret it. As for question (iii), I argued that a prototypical indefinite article like Dutch *een* is probably the best candidate to investigate.

The picture of indefinite articles that will emerge throughout this dissertation is that of a category of determiners that are poor in their semantics but rich in their pragmatic competition with bare nominals and other determiners. This picture is nothing more than a finer-grained version of the original 'blocking principle': once we accept that indefinite articles and bare nominals can co-exist – be it in argument or in predicate position – we open a whole new range of possible interactions that are worthwhile exploring.

## ***UNOS AND ARTICLES UNITE!***

The semantics and pragmatics of indefinite articles

### ***General Introduction***

The goal of this part of my dissertation is to probe the semantics and pragmatics of indefinite articles on the basis of a case-study of Spanish *unos* ( $\approx$  ‘some<sub>pl</sub>’) (Villalta 1994, Laca & Tasmowski 1994, 1996, Gutiérrez-Rexach 1999, 2001, 2003, Alonso-Ovalle & Menéndez-Benito 2002, Lopez-Palma 2007, Martí 2008, Pozas-Loyo *to appear*). On the basis of both diachronic and synchronic evidence I will argue that *unos* is best analyzed as the Spanish plural indefinite article.

In Chapter 1, I argue that *unos* should be analyzed as being in complementary distribution with the bare plural. In Chapter 2, this complementary distribution is exploited to argue that *unos* is the Spanish plural indefinite article and synchronic arguments in favour of this analysis are discussed. Chapter 3 presents diachronic arguments and concludes.



## Chapter 1: Bare plurals and *unos* N

### 1. The starting point

I start with two observations:

#### *Observation 1*

Spanish DPs headed by *unos* ( $\approx$  ‘some’) are felicitous in all argument positions except in the scope of negation. This is illustrated by the fact that (1) does not allow for a reading in which *unos profesores* scopes below negation:

- (1) A la reunión no asistieron unos profesores.  
At the meeting not attended some professors  
<sup>ok</sup> “There are some professors that didn’t attend the meeting.” (NEG< unos)  
# “No professors attended the meeting.” (NEG> unos)
- Laca (1996)

#### *Observation 2*

Spanish bare plurals are felicitous in argument position with one exception: the preverbal subject position. This is illustrated in (2) and (3):

- (2) María comió bizcochos.  
María ate biscuits
- (3) \*Políticos han ocupado el palacio.  
Politicians have occupied the palace
- Delfitto & Schrotten (1991)

I postpone the discussion of some exceptions to these observations to Chapter 2, section 2.

In principle one could propose separate accounts for the above observations. The lack of a narrow-scope reading for *unos* in (1) on the one hand could be accounted for if we assume that *unos* is a positive polarity item (PPI) and that PPI-behaviour is standardly built into indefinite determiners (cf. Martí 2008). The contrast between (2) and (3) on the other hand could be accounted for by assuming that Spanish can project null Ds but only in properly governed positions (cf. Longobardi 1994). I would like to argue though that the two phenomena are related. The gist of my proposal is given in (4):

- (4) Bare plurals and *unos* N are complementary: their truth-conditional contribution is the same but they differ in discourse salience. The salient *unos* N cannot be used in positions that are such that the newly introduced discourse referents they host cannot be picked up (such as the scope of negation) and the non-salient bare plural cannot be used in positions that require the introduction of a salient discourse referent (such as the preverbal subject position).

This proposal will be worked out in section 5. Before doing so I however have some explaining to do. First I have to define the truth-conditional contribution of bare plurals and *unos N* (section 2). Second I have to define the notion *discourse salience* (section 3) and third I have to show that *unos N* is more salient than the bare plural (section 4).

## 2. The semantics of *unos N* and the bare plural

In this section I present the semantics I assume for *unos N* and the bare plural. In 2.1. I look at *unos N*, in 2.2. at the bare plural and I sum up in 2.3.

### 2.1. The semantics of *unos N*

I assume the semantics of *unos N* is as follows:

$$(5) \quad \lambda Q \exists y (\text{plural}(y) \& N(y) \& Q(y))$$

According to (5) *unos N* is a standard plural indefinite. This analysis correctly predicts that (6) is true if and only if María has eaten at least two biscuits:

- (6)      María comió unos bizcochos.  
            María ate some biscuits

In the literature three types of additions to (5) have been proposed. In the remainder of this subsection I briefly present the facts that motivated each type of addition and refer the reader to the sections in which I treat them in detail. Crucially I will argue that none of the additions is necessary and that (5) suffices to account for all semantic facts concerning *unos N*.

The first type of addition was proposed to account for the PPI behaviour of *unos N* (see section 1). In the present chapter I argue that an addition of this type is not necessary given that the NPI behaviour of *unos N* can be derived from its competition with the bare plural.

The second type of addition was proposed to account for the fact that *unos* is infelicitous in examples like (7) (Gutiérrez-Rexach 1999, 2001, 2003):

- (7)      ??Unos estudiantes son abogados.  
            unos students are lawyers

Gutiérrez-Rexach (2001)

(7) and similar facts are discussed in section 3 of chapter 2. There I show that the semantics in (5) suffices to account for the infelicity of (7).

The third and final type of addition to (5) was proposed to account for the fact that *unos N* in examples like (8) doesn't easily allow for distributive readings but strongly prefers the collective interpretation according to which a group of men bought a single lottery ticket.

- (8) Unos hombres compraron un billete de lotería.  
unos men bought a ticket of lottery

Villalta (1994)

In section 4 of chapter 2 I argue that no addition is necessary to account for the preference for collective readings of *unos N* and that the additions that have been proposed might even lead to wrong predictions.

## 2.2. The semantics of the bare plural

The semantics I assume for the bare plural is the same as the one I proposed for *unos N* (see (5)). This assumption raises two questions. The first is related to economy: under the assumption that a language is economic one could wonder whether it is desirable to have two items that have the same semantics. The answer I propose is that it would be undesirable to have two items that are completely equivalent but that there is no reason to assume that a language couldn't have two items with the same semantics as long as the discourse properties of the two items are different. Given that I assume that the bare plural and *unos N* are to be distinguished in their discourse properties the economy view on language is not endangered.<sup>1</sup> The second question is about the correctness of the proposal: is it really the case that *unos N* and the bare plural have the same semantics? As far as standard declarative sentences are concerned the answer I propose is affirmative: there is no situation that would make (9) true and its counterpart with *unos N* in (6) false or vice versa.

- (9) María comió bizcochos.  
María ate biscuits

Declarative sentences such as (9) are however not the only type of sentences in which we find bare plurals and *unos N*. We also find them in questions and other scale reversal environments (see Fauconnier 1979) and the contrast in (10) and (11) seems to suggest that we do find a semantic difference between bare plurals and *unos N* in this type of sentences:

- (10) John: [Betty says that she saw children playing in the garden, but I don't think it's true.] ¿Tu viste *niños* jugando en el patio? (You saw children playing in the garden?)  
Mary: Yes, I saw one.

Martí (2008)

- (11) John: ¿Viste a *unos niños* jugando en el patio? (You saw *unos* children playing in the garden?)  
Mary: #Yes, I saw one.

Martí (2008)

---

<sup>1</sup> See chapter 2, section 2 for further discussion from the perspective of the Blocking Principle (Chierchia 1998).

The difference between (10) and (11) cannot straightforwardly be reduced to a difference in discourse properties and suggests there might be a semantic difference between the bare plural and *unos N*: (10) seems to indicate that the bare plural is semantically not really plural whereas (11) seems to indicate that *unos N* is. Probing this further would lead me beyond the scope of this dissertation, not only because there is no consensus in the literature on how to interpret the facts in (10) (compare Farkas & de Swart 2009, Sauerland et al. 2005 and Spector 2007) but also because recently new facts have been proposed that seem to suggest that the plurality of *unos N* is not the one we find in other plural DPs (see Alonso-Ovalle & Menéndez-Benito 2009). I will accordingly ignore possible variation in the plurality part of the semantics proposed in (5) in what follows.

### 2.3. Summing up

In this section I proposed the following semantics for both *unos N* and the bare plural:

$$(12) \quad \lambda Q \exists y (\text{plural}(y) \& N(y) \& Q(y))$$

I assume this semantics is adequate for both despite some subtle differences in their plural component. The crucial difference between the two, I claim, lies elsewhere, *viz.* in the fact that *unos N* is salient whereas the bare plural is not (see also Laca & Tasmowski 1994). In section 3 I zoom in on the notion of (discourse) salience.

## 3. Discourse salience

Discourse salience is a notion that is difficult to pin down. I will first present three examples that each illustrate a possible application of the notion and then narrow down its scope to its most easily verifiable version that I will define within Centering Theory (Walker, Joshi & Prince 1998).

### *Three examples*

For all three examples I will assume the author had the choice between a bare plural and *unos N* and selected *unos N* to introduce a salient discourse referent. Note that this is purely for expository reasons and that I do not dismiss the task of showing that *unos N* is indeed salient whereas the bare plural is not.

#### (13) **Example 1**

Por la mañana llegaron **unos milicianos** y nos sacaron del sótano. Se llevaron a todos los hombres. A las mujeres nos devolvieron a nuestras casas. Había siete soldados, sucios y malolientes, con ojos feroces y barba de muchos días. Tuve que cocinar para ellos. Así estuve cuatro meses.

‘In the morning **unos soldiers** arrived and took us from the basement. They took away all the men. They brought the women back to their houses. There were seven soldiers, filthy and with a bad smell, with fierce eyes and a beard of several days. I had to cook for them. This situation lasted for four months.’

data drawn from the CREA corpus<sup>2</sup>

In (13) the author arguably selected the salient *unos milicianos* instead of *milicianos* in order to signal that the referent would play a role in the following sentence.

(14) **Example 2**

Las clases no me parecieron difíciles. Tenía **unos profesores excelentes**. El trabajo era estimulante... [p.114]

‘The classes didn’t seem difficult to me. I had unos excellent professors. The work was stimulating.’

Mujeres de negro<sup>3</sup>

In (14) *unos profesores excelentes* is arguably preferred over the less salient *profesores excelentes* to underline the extraordinary quality of the professors.

(15) **Example 3**

Y recordé otra noche de Reyes, la del último año en Los Valles. Tampoco entonces podía dormir y oí **unos golpes fuertes** que resonaron en toda la casa. Mi padre bajó las escaleras y gritó: [p.46]

‘And I remembered another Kings’ eve, the one of the last year in Los Valles. I couldn’t sleep back then either and I heard unos strong blows that sounded throughout the whole house. My father went down and shouted.’

Mujeres de negro

In (15) it is not immediately clear why the author preferred *unos golpes fuertes* over *golpes fuertes*; the referent is not picked up and *being strong* is not an extraordinary property. It turns out though that these blows are highly relevant. Indeed, on the same page the author reports on another event:

(16) Dormí de un tirón y cuando desperté oí **golpes** en la puerta y mi madre dio un salto... [p.46]

‘I fell asleep immediately and when I woke up I heard blows on the door and my mother jumped...’

Mujeres de negro

<sup>2</sup> The CREA corpus is freely available online (<http://corpus.rae.es/creanet.html>). The part containing (mainland) Spanish texts is the largest corpus available of present-day (mainland) Spanish.

<sup>3</sup> *Mujeres de negro* is a novel by the mainland Spanish author J.R.Aldecoa and was written in 1995.

And a few lines later she makes the link between the two:

- (17) “Fue como aquella vez con la muñeca. Los golpes en la puerta y luego...”  
 [p.48]  
 ‘It was like the time with the doll. The blows at the door and then...’  
 Mujeres de negro

It thus turns out that the blows introduced in (15) are extremely relevant for the story because they are the associative link between two events the author clearly wants to connect. It can therefore be argued that the author wanted to use the salient *unos golpes fuertes* in order for the reader to remember these so that the link between the blows in (15) and (16) would be immediately apparent.

As should be clear from examples 1 through 3 the notion *discourse salience* can be interpreted in many ways. An important disadvantage of this multi-interpretability is that it is impossible to come up with clear and verifiable criteria that would allow one to argue that one type of NP/DP is salient whereas another is not. In order to avoid this problem I will use a very restricted version of discourse salience that I define within Centering Theory (Walker, Joshi & Prince 1998). I first briefly introduce the relevant notions of this framework.

### *Centering Theory*

Centering theory tries to model discourse coherence. Its input is a discourse chunked up in discourse segments. An example is given in (18), each line corresponding to a segment:

- (18) a. Jeff helped Dick wash the car.  
 b. He washed the windows as Dick waxed the car.  
 c. He soaped a pane.

Discourse segments evoke a number of discourse entities that can be grouped into sets which are referred to as *forward-looking centers*. For the discourse segments in (18) the forward-looking centers are given in (19):

- (19) a. [JEFF, DICK, CAR]  
 b. [JEFF, WINDOWS, DICK, CAR]  
 c. [JEFF, PANE]

The order of the elements in a forward-looking center is not random but crucially reflects the probability with which a discourse entity will reappear in the forward-looking center of the next discourse segment. The highest-ranked member is referred to as the *preferred center* and the member corresponding to the highest-ranked member of the previous discourse segment is referred to as the *backward-looking center*. JEFF is the preferred center of each of the discourse segments in (18) and the backward-looking center of (18.b) and (18.c). Note that (18.a) – being the first segment in a discourse – does not have a backward-looking center.

*Back to discourse salience*

I define discourse salience as the probability with which a discourse entity occurring in the forward-looking center of a discourse segment A will reappear in the forward-looking center of the discourse segment following it. I follow Roberts (1998) in assuming that discourse entities correspond to DRT discourse referents.

On the above definition an NP/DP *x* is salient if its referent is likely to be picked up in the discourse segment following it.

**4. Discourse salience, *unos N* and the bare plural**

In this section I will show on the basis of a corpus study that discourse referents introduced by *unos N* are likely to be picked in contrast to those introduced by bare plurals. On the definition of discourse salience presented in section 3 this means that *unos N* is salient whereas the bare plural is not. Before presenting the results of the corpus study I want to take a short look though at what has been said in the literature about Spanish bare plurals and their ability to introduce discourse referents.

**4.1. A note on Spanish bare plurals**

In the semantics literature on bare plurals in Spanish (see especially Laca 1996, 1999, McNally 2004) it has consistently been argued that they do not introduce discourse referents and can therefore not function as regular antecedents. If this holds *unos N* would be the default way to introduce discourse referents and they would automatically be not salient. In what follows I will go over Laca (1996, 1999) and McNally (2004) and argue that their complementary analyses do not account for all the facts.

*Laca (1996,1999)*

On the basis of examples like the one in (20) Laca (1996,1999) argues that bare plurals do not introduce discourse referents:

- (20) Al principio, Juan quería restaurar muebles de estilo, pero terminó vendiéndolos.  
 At\_the beginning, Juan wanted restore pieces\_of\_furniture of style, but ended\_up selling-them  
 ‘Initially Juan wanted to restore antiques but he ended up selling them.’  
 Laca (1996)

If the bare plural *muebles de estilo* were able to introduce a discourse referent we would expect (20) to be ambiguous between a reading in which the post-clitic pronoun *-los* refers to specific antiques and a reading in which it picks up the

descriptive content of the bare plural.<sup>4</sup> Given that only the latter reading is available – that of Juan being an antiques salesman – Laca concludes that bare plurals cannot introduce discourse referents.

*McNally (2004)*

Attested examples such as the ones in (21) through (23) show that something has to be added to Laca's analysis to account for contexts in which the referent of a bare plural is picked up:

- (21) Luego voy a darte un ramillete de **violetas** para tu madre, Juana. Las he cogido esta mañana, antes de que empezara a llover. [p.41]  
 Later I will give you a bunch of **violets** for your mother, Juana. I picked them this morning, before it started to rain.
- (22) Apenas encontramos coches, sólo alguna camioneta renqueante cargada de madera. Y **campesinos** en sus burros. Parecían dormidos; avanzaban despacio con la cabeza gacha, oculta por el amplio sombrero atravesando, como nosotros, la Sierra Madre del Sur. [p.89-90]  
 We hardly saw any cars, only some oldish van loaded with wood. And **farmers** on their donkeys. They seemed asleep; they progressed slowly with their head hanging, hidden by the large hat, crossing, like we were, the Sierra Madre del Sur.
- (23) Envié **peones** a caballo. Otros a pie con antorchas. [...] Hacia las ocho de la mañana llegaron los de los caballos con Octavio y Soledad montados en el caballo de Octavio. [p.127]  
 I sent **workers** on horseback. Others on foot with torches. [...] Around eight o'clock in the morning those of the horses came back with Octavio and Soledad both on Octavio's horse.

Mujeres de negro

McNally (2004) extends Laca's analysis to these cases suggesting that the underlined DPs in (21) through (23) do not pick up referents introduced by the bare plurals but rather accommodated referents that are facilitated by the descriptive content of the noun and the fact that the propositions in which the bare plurals occur entail the existence of referents corresponding to this descriptive content. The contrast between (21) through (23) on the one hand and (20) on the other hand follows given that *Juan quería restaurar muebles de estilo* ('Juan wanted to restore antiques') does not entail the existence of antiques whereas (21) through (23) do entail the existence of flowers, farmers and workers.

*Deciding between two approaches*

The Laca (1996,1999) and McNally (2004) type of analysis accounts in an elegant way for the bulk of the bare plural facts but does not offer a knock-down argument

---

<sup>4</sup> Cf. the token- and type-anaphora of Borthen (2003).

against analyses that take bare plurals to introduce non salient discourse referents. This raises the question whether there is a way to decide between the two types of approaches. I would like to suggest that there is: the availability of non-restrictive relative clauses.

What makes non-restrictive relative clauses interesting is that their head NP/DP is necessarily referential (*i.e.* is of type *e*, see Potts 2004 for a worked-out analysis). In DRT terms this means the head NP/DP of a non-restrictive relative clause has to introduce a discourse referent. Under McNally's assumption that accommodation plays at the clausal level Laca and McNally would predict that bare plurals cannot take a non-restrictive relative clause whereas the other type of approach would predict exactly the opposite. The attested example in (24) shows that it is the discourse referent approach that is on the right track.

- (24) Emily realizó un número significativo de pruebas con cada voluntario y sus resultados los sometió a sencillas estadísticas **que por lo demás no eran necesarias** [...]

Emily carried out a significant number of tests with each volunteer and submitted her results to **simple statistics** that for the rest were not necessary

data drawn from the CREA corpus

Note that I assume that *por lo demás* ( $\approx$  'by the way/for the rest') signals the non-restrictiveness of the relative clause.

#### 4.2. Discourse salience: a corpus study

In 4.1. I argued that bare plurals can introduce discourse referents. In this subsection I will present a corpus study that shows that the discourse referents introduced by bare plurals are less often picked up than those introduced by *unos N*.

##### *The corpus*

For the particular task at hand I selected the modern version of the Old Testament published by the World Bible Translation Center. There are two reasons to select this text. The first is that the Old Testament contains many characters which guarantees there are many discourse referents that can be introduced and picked up. The other reason to select this bible rather than e.g. the Biblia Reina Valera is that it was written to present the bible to a wide audience in a simple and up-to-date language.

##### *Selecting occurrences*

In order to find as many occurrences of bare plurals and *unos N* that might be picked up I restricted my search to argumental positions (subject, object, indirect object) in main clauses. For *unos N* I selected all occurrences in the Old Testament (n=49) and for the bare plural I selected all those that I manually found in the book Genesis (n=70).

In (25) through (30) I present some of the selected occurrences. (25)-(27) are occurrences of *unos* and (28)-(30) are occurrences of the bare plural. In (25), (26), (28) and (29) the underlined words are those picking up the referent introduced by *unos N* / the bare plural. Note that I did not distinguish between pronominal and full DP anaphora.

- (25) Jacob continuó su camino y se encontró con **unos ángeles** de Dios. Cuando los vio, les dijo: “¡Este es el campamento de Dios!”  
Jacob continued his way and he met **unos angels** of god. When he saw them, he said unto them: “This is the camp of God!”
- (26) Ellos le respondieron: – Tuvimos **unos sueños** y no hay nadie que nos los pueda explicar. Y José les dijo: – Sólo Dios puede interpretar los sueños. Cuéntenmelos.  
They answered him: – We had **unos dreams** and there is no-one who can explain them to us. And Joshua told them: – Only god can interpret dreams. Tell me them.
- (27) Vi entre **unos arrayanes** a un jinete en un caballo rojo. Detrás de él había otros caballos de color rojo, castaño y blanco.  
He saw a horseman and a red-haired horse between **unos myrtles**. Behind him there were other red, chestnut and white horses.
- (28) Vamos, hagamos **ladrillos** y pongámoslos en el fuego para endurecerlos.  
Let’s go, let’s make **tiles** and let’s put them in the fire to make them firm.
- (29) Najor también tuvo **hijos** con su concubina Reumá. Ellos fueron Tébj, Gaján, Tajás y Macá.  
Najor also had **children** with his concubine Reuma. They were Tebj, Gajan, Tajas and Maca.
- (30) También les dio **regalos muy caros** al hermano y a la mamá de ella. Luego él, y los hombres que estaban con él, comieron y pasaron ahí la noche. A la mañana siguiente se levantaron y dijeron:  
He also gave **very expensive presents** to her brother and mother. Afterwards he and his men ate and spent the night there. The following morning they got up and said:

#### *The results*

The results of my corpus study can be summarized as follows. Out of the 49 occurrences of *unos N* 30 were picked up (=61%) whereas out of the 70 occurrences of the bare plural only 10 were picked up (=15%). These results clearly show that – under the definition proposed in section 2 *unos N* is salient whereas the bare plural is not.

### 5. Explaining the facts

In section 1 I announced I would give a unified account of the unacceptability of bare plurals in preverbal subject position (see (31)) and the PPI behaviour of *unos N* (see (32)):

- (31) A la reunión no asistieron unos profesores. NEG< unos \*NEG> unos  
 At the meeting not attended some professors  
 “There are some professors that didn’t attend the meeting.”  
 Laca (1996)
- (32) \*Políticos han ocupado el palacio.  
 Politicians have occupied the palace  
 Delfitto & Schroten (1991)

I suggested that the basis for this unified account lies in the assumption that the bare plural and *unos N* are complementary: they have the same semantics but whereas the bare plural introduces non-salient discourse referents, *unos N* introduces salient ones. Now that I have defined discourse salience (section 3) and shown that *unos N* is indeed salient whereas the bare plural is not (section 4) I can spell-out how the complementarity of the bare plural and *unos N* is connected to the facts in (31) and (32).

Under the assumption that the bare plural is the non-salient member of a truth-conditional pair and *unos N* the salient member, one would predict that:

- (33) a. in a context that prohibits the picking up of discourse referents the bare plural is the preferred option  
 b. in a context that forces the picking up of discourse referents *unos N* is the preferred option

The rationale behind (33.a) is that a speaker will not introduce a salient discourse referent if s/he knows it cannot be picked up anyway and if s/he can choose between a salient and a non-salient discourse referent. The rationale behind (33.b) is that a speaker will not introduce a non-salient discourse referent if s/he knows that it will be picked up anyway and if s/he can choose between a salient and a non-salient discourse referent. In the following two subsections I will show how the predictions in (33.a) and (33.b) can be used to explain the facts observed in (31) and (32) respectively.

### 5.1. Deriving the PPI behaviour of *unos*

There are three environments that are known to block the picking up of discourse referents: the scope of negation, that of a modal operator and that of a quantifier. This is illustrated in (34) through (36). Note that I use English bare plurals to illustrate because of their reluctance to take wide scope (cf. Carlson 1977).

- (34) John didn’t get presents<sub>i</sub>. #They<sub>i</sub> were great.  
 (35) John wants to meet scientists<sub>i</sub>. #They<sub>i</sub> are smart.  
 (36) Every boy kissed girls<sub>i</sub>. #They<sub>i</sub> were happy.

The picture in (34) through (36) is not complete though. Indeed, despite the fact that it is generally true that all three environments block the picking up of discourse referents two of them are known to be more permissive: the scope of a modal



Pedro tiene que escribir unas cartas hoy. Tiene que enviarlas por avión.  
 Pedro has to write unos letters today He\_has to send\_them by airmail

Before moving on to the explanation of the preverbal subject constraint on bare plurals it is important to stress that (33.a) nor (33.b) make any prediction about the (in)ability of bare plurals to take scope over negation. Indeed, given that the predictions in (33) are only about environments that necessarily block or necessarily force the picking up of discourse referents and given that discourse referents scoping above negation can but need not be picked up the predictions in (33) have no bearing on the possibility of any item to take wide scope with respect to negation.

## 5.2. Deriving the preverbal subject constraint on bare plurals

In order to show that the unacceptability of bare plurals in the preverbal subject position in (32) is connected to the prediction in (33.b) I first have to introduce the standard analysis of preverbal subjects in Spanish.

Zagona (2002) (following work by Contreras (1991) and Olarrea (1996)) claims that Spanish preverbal subjects are adjuncts of a silent left dislocated clitic. The main argument in favour of this analysis is that there need not be grammatical agreement between a preverbal subject and its verb:

- (43) Los estudiantes tenemos un alto concepto de nosotros mismos.  
 The students have-1st-pl. a high opinion of us- selves  
 “Students (we) have a high opinion of ourselves.”

The disagreement between *los estudiantes* and *tenemos* can be explained if we assume a silent first person plural clitic that picks up the referent introduced by *los estudiantes*. Under this assumption *los estudiantes* has to be picked up by the silent clitic in order to be interpreted as the subject of the sentence and – more generally – to be interpretable within the sentence at all. This crucially means that the preverbal subject position in Spanish has the property of forcing the picking-up of discourse referents introduced by the NPs or DPs occurring in it. To the best of my knowledge it is also the only position that has this property.

On the basis of what precedes the prediction in (33.b) can be restated as follows: given that the preverbal subject position is the only context forcing the picking up of discourse referents it is the only position in which Spanish bare plurals are expected not to be allowed. This correctly predicts the unacceptability of (32) and the general acceptability of bare plurals in all other argument positions.

## 6. Conclusion

In this chapter I have argued that the PPI behaviour of *unos N* and the unacceptability of bare plurals in preverbal subject position can be accounted for if –

next to adopting standard assumptions about the scope of negation and preverbal subjects – we assume that bare plurals and *unos N* have the same semantics but are complementary in discourse salience. The advantage of this approach over previous ones is that instead of proposing special restrictions for both the bare plural and *unos N* I make these restrictions follow from the independently established difference in salience between them (see section 4).

In the next chapter I will exploit the idea that the bare plural and *unos N* are complementary to argue that *unos N* should be analyzed as the Spanish plural indefinite article.

## Chapter 2: *Unos* as a plural indefinite article: synchronic evidence

In chapter 1 I argued that in order to capture constraints on both *unos N* and the bare plural it makes sense to analyze them as complementary expressions; the former introducing salient discourse referents, the latter introducing non-salient ones. The facts that gave rise to this analysis are the PPI behaviour of *unos* and the unacceptability of bare plurals in the preverbal subject position:

- (44) A la reunión no asistieron unos profesores. NEG< unos \*NEG> unos  
 At the meeting not attended some professors  
 “There are some professors that didn’t attend the meeting.”  
 Laca (1996)
- (45) \*Políticos han ocupado el palacio.  
 Politicians have occupied the palace  
 Delfitto & Schroten (1991)

In this chapter this complementarity will be exploited to argue in synchrony for the (indefinite) article status of *unos*. The theoretical basis for the argumentation in this chapter is presented in section 1.

### 1. Indefinite articles in the literature

Partee (1987) argued that there are certain ‘natural’ shifts between the basic types  $e$ ,  $\langle e, t \rangle$  and  $\langle \langle e, t \rangle, t \rangle$  that are such that every language allows them and that some languages might even have lexicalized them. These lexicalizations – by assumption – correspond to what is commonly referred to as *articles*. In what follows I will zoom in on what has been said in the literature about the shift corresponding to the indefinite article and the status of lexicalized shifts.

#### *The indefinite article*

In Partee’s analysis the indefinite article is taken to be the lexicalization of the existential type-shift as defined in (46) which takes an  $\langle e, t \rangle$  expression and turns it into an existential quantifier of type  $\langle \langle e, t \rangle, t \rangle$ . The distinction between singular and plural is reflected in the absence/presence of *plural* ( $x$ ).

- (46) *singular*:  $\lambda Q \exists x (P(x) \& Q(x))$   
*plural*:  $\lambda Q \exists x (\text{plural}(x) \& P(x) \& Q(x))$

Partee argued for the naturalness of this shift on the basis of the fact that it is the least marked inverse of the BE type-shift which takes a generalized quantifier and collects the elements of all its singletons into a set:

- (47)  $\lambda \wp [\lambda x [\{x\} \in \wp]]$

*The status of lexicalized shifts*

In the type-shifting literature following Partee (see esp. Chierchia 1998 and Dayal 2004) the following principle was adopted:

## (48) Blocking Principle

For any type shifting operation  $\tau$  and any  $X$ :

\* $\tau(X)$

if there is a determiner  $D$  such that for any set  $X$  in its domain,

$D(X) = \tau(X)$

Chierchia (1998)

What this principle does is to restrict covert applications of type-shifts; if a certain shift is lexicalized in a language it cannot apply covertly in that language. The relevant prediction this principle makes is that in languages that have an indefinite article nouns cannot appear bare in argument position with an indefinite interpretation. Insofar as Chinese does not have an indefinite article and English does this prediction is borne out:

## (49) wò kànjiàn xióng le

I see bear ASP

‘I saw a bear’

Chierchia (1998)

## (50) \*I saw bear

‘I saw a bear’

In the following sections I will argue on the basis of synchronic facts that *unos* should be analyzed as an indefinite article. In order to do this I will present four predictions Partee’s definition makes and show that they are borne out. This is done in sections 3 through 5. I however start with the most challenging feat; arguing that an adapted version of the Blocking Principle holds for Spanish, despite the existence of bare plural arguments and my claim that *unos* is the plural indefinite article.

## 2. *Unos* and the Blocking Principle

In this section I argue that the Blocking Principle cannot be maintained in its current form whether we argue that *unos* is an article or not. I furthermore propose an adapted version that completes the account of the competition between *unos N* and the bare plural given in Chapter 1.

### 2.1. The Blocking Principle and the Spanish plural paradigm

In order to get more grip on the Blocking Principle I first transform it into an OT analysis with two forms, two meanings and two (syntactic) constraints:

two forms:

- (51) X henceforth *BN*  
 (52)  $\tau(X)$  henceforth *ArtN*

two meanings:

- (53)  $\lambda x(P(x))$  henceforth *pred*  
 (54)  $\lambda Q\exists x(P(x)\&Q(x))$  henceforth *arg*

two constraints:

- (55) \*FORM  
 Avoid form.  
 (56) FAITHTYPESHIFT  
 Mark type-shifts.

The two possible rankings of the constraints reflect the two types of languages the Blocking Principle gives rise to: (i) *Article Languages*: those that have an indefinite article and do not allow bare nominals to acquire argument status through a covert type-shift (tableau 1) and (ii) *Bare Languages*: those that do not have an indefinite article and do allow bare nominals to type-shift (tableau 2). Note that both Tableau 1 and 2 are composed of two smaller OT syntax tableaux.

	FAITHTYPESHIFT	*FORM
<BN, arg>	*!	
<ArtN, arg> $\curvearrowright$		*
<BN, pred> $\curvearrowright$		
<ArtN, pred>		*!

Tableau 1: Article Languages

	*FORM	FAITHTYPESHIFT
<BN, pred> $\curvearrowright$		
<ArtN, pred>	*!	
<BN, arg> $\curvearrowright$		*
<ArtN, arg>	*!	

Tableau 2: Bare Languages

With Tableaus 1 and 2 in place I can turn to the question whether one of them can account for the Spanish plural facts. The answer is negative. Indeed, Spanish cannot be an Article Language given that it allows for bare plurals in argument position. It cannot be a Bare Language either though; if it were we would expect it to allow for bare nominals in any position (especially given that *Lift* and *Lower* seem to be freely available universally) and we would expect there to be no determiners that are sensitive to the availability of bare nominals. The facts discussed in Chapter 1

indicate that these expectations are not borne out. Spanish – in the plural paradigm – then seems to be somewhere in between a Bare Language and an Article Language.<sup>1</sup>

## 2.2. Adapting the Blocking Principle

In what follows I will present an enriched version of the analysis in Tableau 2 that accounts for the Spanish facts. In principle this is not too difficult: one can always add a constraint that requires faithfulness to a feature included in *unos*. This is unattractive though in two respects: on the one hand it would reduce the difference between *unos N* and the bare plural to a lexical one and on the other hand it would force me to put as much into the analysis as I hope to get out of it. The more attractive analysis I will present is one in which *unos N* and the bare plural are equally unmarked in their semantics and where it's the competition between the formally unmarked bare plural and the formally marked *unos N* that takes care of the marked interpretation of the latter. If we take salience to be an instantiation of markedness this analysis accounts for the facts. It furthermore makes three extra predictions connected respectively to the acceptability of *unos N* in the scope of negation, its acceptability in predicate position and the bare plural's acceptability in preverbal subject position. After the presentation of the analysis I take a closer look at these predictions.

### *The analysis*

The type of analysis I'm after is not formalizable in OT syntax but can be worked out in (weak) Bi-directional OT if we add the meanings and constraint in (57) through (59) to those used in Tableaus 1 and 2. Note that +*M* in (57) and (58) stands for 'marked interpretation':

*two extra meanings:*

(57) *pred + M*

(58) *arg + M*

*one extra constraint:*

(59) \**M-MEANINGS*  
Avoid marked meanings

---

<sup>1</sup> Note that I leave aside the singular paradigm here that can be analyzed along the lines of Tableau 1 with some exceptions (see e.g. Espinal & McNally 2007).

The worked-out tableau looks as follows:

	*FORM	FAITHTYPESHIFT	*MARKEDNESS
<BN, pred> $\emptyset$			
<BN, pred+M>			*
<BN, arg> $\emptyset$		*	
<BN, arg+M>		*	*
<ArtN, pred>	*		
<ArtN, pred+M> $\emptyset$	*		*
<ArtN, arg>	*		
<ArtN, arg+M> $\emptyset$	*		*

Tableau 3: between a Bare and an Article Language (general)

On the analysis in Tableau 3 the bare plural comes out as the optimal form and takes the unmarked meanings. *Unos N* comes out as the suboptimal form and takes the marked meanings.

*The predictions*

As I pointed out above the analysis in Tableau 3 not only accounts for the facts in Chapter 1 but also makes three further predictions. I briefly go over each of them.

1. In Chapter 1 I proposed that *unos N* – because of its salience – is not allowed in the scope of negation. Now suppose *salience* is not the only marked interpretation *unos N* can get through its competition with the bare plural. In that case the analysis in Tableau 3 predicts that if DPs can get an interpretation in the scope of negation that is independently argued to be marked *unos N* should be allowed to get it as well. Metalinguistic interpretations fit the description (see Horn 1985) and (60) shows we indeed find *unos* with this type of interpretation in the scope of negation:

- (60) No compré unos libros, sino que compré unas revistas.  
not I\_bought unos books but that I\_bought unos magazines  
“I didn’t buy some books but rather some magazines”  
Gutiérrez-Rexach (2001)

2. In Tableau 2 *unos N* was predicted not to be allowed in predicate position at all. Tableau 3 however predicts that in as far as predicates can get a marked interpretation they should allow for *unos N*. Under the assumption that markedness in predicates can be thought of in terms of metaphorical interpretations (61) shows that this prediction is borne out:

- (61) Todos los españoles somos unos conejillos mansos.  
All the Spanish we\_are some rabbits domesticated  
“We, the Spanish, are all domesticated rabbits.”  
Laca & Tasmowski (1996)

3. In Chapter 1 I proposed that the bare plural – because of its lack of salience – is not allowed to occur in the preverbal subject position. On the analysis in Tableau 3 this lack of salience is however merely connected to the bare plural's relative unmarkedness in form. This makes the indirect prediction that bare plurals should be allowed in the preverbal subject position if they are made 'heavier'. Under the assumption that stress and adjectives can be used to make bare plurals 'heavier' (62) and (63) show that this prediction is borne out:

(62) Políticos corruptos han ocupado el palacio.  
 politicians corrupt have occupied the palace  
 Delfitto & Schrotten (1991)

(63) RATONES van a salir de ese armario cuando te decidas a limpiarlo.  
 mice go to get\_out of that cupboard when you you\_decide to clean\_it  
 Laca (1999)

### 2.3. Conclusion and outlook

In this section I have argued that the Spanish facts discussed in Chapter 1 are compatible with an enriched version of the Blocking Principle. This is an important first step towards arguing in favour of the article status of *unos*.

In the following sections I will further strengthen my claim about the article status of *unos* on the basis of synchronic facts. I will present four crucial predictions Partee's analysis makes and show that they are borne out. In section 3 and 4 I discuss predictions that have not been addressed in the literature, section 5 deals with standard predictions.

## 3. *Unos* and partitive readings

In this section I discuss the first prediction the article analysis of *unos* makes: I argue that articles are expected to lack partitive readings (3.2) and that *unos* conforms to this prediction (3.3-3.5). I start however by introducing a specific background assumption I make about partitive readings of indefinites.

### 3.1. Partitive readings of indefinites

I assume most indefinites are ambiguous between a partitive and a non-partitive reading. In (64) I tentatively illustrate the partitive reading for the numeral *one*, the partitivity feature is in bold:

(64)  $\lambda P \lambda Q \exists x (one(x) \& P(x) \& Q(x) \& \exists y (P(y) \& \neg Q(x)))$

The reader is referred to Appendix 1 of this dissertation for a full motivation of this assumption.

### 3.2. The indefinite article and partitive readings

Partee defined the indefinite article as the least marked inverse of the BE type-shift. Taking this to its extreme one expects the indefinite article not to have semantic content that is ignored by this type-shift. As a consequence, we expect partitive readings to be unavailable for indefinite articles. To see this let's compare what happens to the partitive version of *one woman* and its non-partitive version for the model in (65):

(65) {Mary, Harriet, Amber} = [[woman]]

Given that BE ignores everything but singletons the result of applying it to the partitive (66) and the non-partitive (67) will be exactly the same, *viz.* the set in (68):

(66)  $\lambda P \lambda Q \exists x (one(x) \& woman(x) \& Q(x) \& \exists y (P(y) \& \neg Q(x)))$

(67)  $\lambda P \lambda Q \exists x (one(x) \& woman(x) \& Q(x))$

(68) {Mary, Harriet, Amber}

The fact that the application of BE neutralizes the difference between (66) and (67) shows that partitivity is ignored by BE. I consequently expect indefinite articles not to be able to induce partitive readings.

In the literature several tests have been proposed to check whether or not a determiner has partitive content. In 3.3. I will show for each test that *unos* comes out as non-partitive. In 3.4. I discuss an apparent counterexample and in 3.5. I give corpus data to further backup the claim that *unos* does not have partitive content. 3.6. concludes this section.

### 3.3. Partitivity Tests

In this subsection I will show on the basis of three tests that *unos* does not have partitive content. The judgements reported here are all agreed upon in the literature (Villalta 1994, Laca & Tasmowski 1994, 1996, Gutiérrez-Rexach 1999, 2001, 2003, Lopez-Palma 2007, Martí 2008).

The first test that has been used to check whether determiners and their corresponding DPs have partitive content is their acceptability in the subject position of an individual level predicate (*be intelligent, be black, be a lawyer...*). The rationale behind this test is that individual level predicates can only be felicitously predicated over a contextually available individual or part thereof.<sup>2</sup>

---

<sup>2</sup> The general consensus in the literature seems to be that individual-level predicates, unlike stage-level predicates, have no event argument that can be topicalized and therefore require

As expected, *unos N* is incompatible with an individual level predicate and patterns with indefinite articles and bare plurals on their non-generic readings:<sup>3</sup>

- (69) ??Unos estudiantes son abogados.  
unos students are lawyers  
Gutiérrez-Rexach (2001)
- (70) A student is a lawyer. [<sup>ok</sup>generic, \*non-generic]
- (71) Students are lawyers. [<sup>ok</sup>generic, \*non-generic]

The second test that has been proposed is the availability of noun ellipsis. Noun ellipsis is known to be felicitous only when reference is made to part of a contextually salient set. Determiners without partitive content are therefore expected to be incompatible with noun ellipsis.

As expected, *unos N* is incompatible with noun ellipsis and patterns with indefinite articles:

- (72) Los libros de matemáticas están en el cajón, los de física debajo de la cama y ?? unos de lingüística están sobre la mesa.  
The books about mathematics are in the drawer, those about physics under the bed and unos about linguistics on the table.  
Gutiérrez-Rexach (2001)
- (73) The books about mathematics are in the drawer, those about physics under the bed and \*an about linguistics on the table.

Note – for completeness – that the ungrammaticality of *unos de lingüística* in (73) is not due to the fact that *unos* has to be followed by a noun. This is clear from the fact that *unos* can be used as a pronoun:

- (74) Unos lo contaron anoche.  
Some told it yesterday evening  
Diccionario de la Lengua Española (2001)

The third and final test that has been used to check whether a determiner has partitive content is its acceptability in the upstairs determiner position of full partitives. Under the assumption that determiners in this position have to have partitive content we expect non-partitive determiners to be out.

*Unos* once more behaves as expected and patterns with the indefinite singular article:<sup>4</sup>

---

their subject to be the topic. In as far as topics cannot be discourse new indefinites would be out unless the set they quantify over is not discourse new and can serve as a topic. This seems to be what happens with partitive readings of indefinites (see de Swart 2001 and references therein).

<sup>3</sup> I abstract away from generic readings of *unos N* (see also Section 5)

- (75) ??unos de los familiares de Pedro  
 unos of the family members of Pedro  
 Gutiérrez-Rexach (2001)
- (76) \*a of the family members of Pedro

(75) and (76) show that *unos* patterns like an indefinite article in being proscribed in the upstairs D position of partitives.

### 3.4. A potential set of counterexamples

Despite the fact that – to my knowledge – nobody has ever questioned the facts in (69), (72) and (75), it is not generally accepted that *unos* does not have partitive content. The reason for this is that *unos N* sometimes does get a partitive reading. This is illustrated in (77) and (78):

- (77) Unos estudiantes son inteligentes, otros no.  
 unos students are intelligent, other not
- (78) De los libros que hablas hay unos sobre la mesa y otros en el cajón.  
 as for the books you're talking about, some are on the table and others in the drawer  
 Gutiérrez-Rexach (2001)

The reason why (77) and (78) are not necessarily counterexamples to the claim that *unos* does not have partitive content is that it is arguably *otros* and not *unos* that induces the partitive reading (see also Martí 2008). The same effect can be obtained with bare plurals which are in general left out of the debate on partitivity:

---

<sup>4</sup> Gutiérrez-Rexach (2001) argues that *unos* can appear in the upstairs D position of partitives on the basis of examples such as (i) which he claims has a partitive with a kind-referring downstairs DP:

(i) Éstos son unos de los mejores jugadores de la Historia.

These are unos of the best players in history

I think Gutiérrez-Rexach made a judgement error in assuming that *los mejores jugadores de la Historia* is a kind-referring downstairs DP. I have two reasons for supposing this. The first is that Spanish speakers in general reject partitives like the ones in (ii):

(ii) \*Unos de los leones / estos leones africanos escasean.

unos of the lions / these lions African are scarce

The second is that almost all attested examples I found on a corpus search looking for *unos* in the upstairs D position of partitives had a superlative downstairs DP. This is a clear indication that what is licensing *unos* is not the fact that *los mejores jugadores* is a kind but rather that it is a superlative. See also footnote 5.

- (79) [Por lo demás, el carácter pacífico y complaciente de Merceditas, su sensibilidad para captar los estados de ánimo de los demás, y su ausencia de susceptibilidad me invitaron a quererla sin reservas y a] compartir con ella periodos luminosos y otros sombríos de nuestras vidas. [p.88]  
share with her periods bright and others gloomy of our lives
- (80) Envié peones a caballo. Otros a pie con antorchas. [p.127]  
I sent workers on horse Others on foot with torches
- Mujeres de negro

Note also the parallel with the indefinite singular:

- (81) He picked up a toy and placed it with the others on the pile.  
data drawn from the BNC

(81) shows that *the others* can induce a partitive reading of the indefinite article.

### 3.5. Detecting partitive readings in a corpus

In 3.3. I presented facts taken from the literature. Given that none of them has been questioned it might seem somewhat redundant to doublecheck them on a corpus. In the next chapter I will however discuss older stages of Spanish for which corpus data are the only source of information. In order to be consistent in methodology, I should therefore also look at corpora for the stage of the language we have speaker's intuitions for.

Two preliminary remarks are in order. The first one is that the type of data I presented in 3.3. is negative. Corpus research for this type of data might at first sight seem useless given that corpora do not provide negative data. This is not entirely accurate though. Indeed, if we can show that a certain string is not attested in a sufficiently large corpus we have ground to believe that it is unacceptable in a language. Conversely, if we can show that a certain string is attested we have ground to believe that it is acceptable in a language. Corpus research can therefore function as an important touchstone for the representativity of made up examples and the reliability of speaker's judgements. The second remark is that there is one type of data in 3.3. that is easier to look for in a corpus than the other two, *viz.* the occurrence of *unos* in the upstairs D position of partitives. Indeed, for this type of data it suffices to look for strings of the type *unos de los* (or *unas de las*) whereas for the other two the search strings might be less straightforward. For convenience I will therefore restrict my attention to *unos* in the upstairs D position of partitives.

The corpus I consulted is CREA, a corpus of present-day Spanish funded and hosted by the Real Academia Española. I restricted the corpus to (mainland) Spanish texts (+85 million words) and expected to find no hits in which *unos de los* was part of a partitive. This expectation was borne out, the only significant type of

counterexample I found was that of partitives having a downstairs superlative DP (n=27):<sup>5</sup>

- (82) se sitúan sobre unos de los puntos más calientes del globo  
they are situated on some of the warmest spots of the globe

Given that this class of counterexamples is so consistent I conjecture that rather than being counterexamples to the claim that *unos* has partitive content there is something special about (plural) superlatives that needs further scrutiny. Coming up with a worked out proposal is beyond the scope of this dissertation, but for relevant discussion see Szabolczi 1986, Heim 1999, Farkas & Kiss 2000, Stateva 2004, Herdan & Sharvit 2006.

For present-day Spanish it appears then that corpus research confirms the intuitions about the data that I proposed in 3.3. to argue in favour of the fact that *unos* does not have partitive content.

### 3.6. Recap

I started out this section by arguing that Partee's analysis of indefinite articles predicts that they lack partitive content. I subsequently showed that this prediction was borne out for *unos* as well as for English *a* on the basis of speaker judgements and corpus data.

## 4. *Unos* and distributive readings

In this section I will discuss the second prediction the article analysis of *unos* makes: I will argue that articles are expected to lack distributive readings and that *unos* conforms to this prediction.

---

<sup>5</sup> The search string I used was *unos de los*. In total it yielded 116 hits, 47 of which were not partitives. Given that 29 hits of the remaining 69 were clearly typos (*unos* instead of *uno* 'one') I only retained those for which I had reason to believe that plural number was intended. This left me with 29 hits. 27 out of those contained a superlative downstairs DP. The two examples that did not contain superlatives are given in (i) and (ii):

- (i) [about the movie *Gladiator*, a movie set in Ancient Rome]  
en una plaza se pelean unos de los de antes  
in a square some of those of before fight
- (ii) [about an exposition of artists' shoes]  
sus zapatos son unos de los pocos nuevos que se exhiben  
his shoes are some of the few new ones on exhibition

#### 4.1. The indefinite article and distributive readings

Once more I take Partee's definition of the indefinite article as the inverse of the BE type-shift to its extreme and assume that it should not have semantic content that is ignored by this type-shift. Next to the prediction that the indefinite article should not have partitive content this also makes the prediction that it should not induce distributive dependencies. This is due to the fact that the BE type-shift only looks for singletons. There is an important *caveat* though: nothing prevents articles from getting a distributive reading through a covert or overt distributivity operator. In general the availability of this type of operator would make the prediction nearly unfalsifiable but Spanish might turn out to be an interesting exception. Indeed, given that *unos* has a near synonym – *algunos* ('some') – that readily allows for distributive readings it is not impossible that *unos* might still exhibit a strong preference for collective readings.

#### 4.2. The collectivity of *unos* in the literature

As expected *unos* has a strong preference for collective readings. This was first noted by Villalta (1994) who claimed that *unos* cannot induce distributive dependencies over some other indefinite in its scope. The strongly preferred interpretation of (83) e.g. is that a group of men bought one lottery ticket:

- (83) Unos hombres compraron un billete de lotería.  
unos men bought a ticket of lottery

Villalta (1994)

A stronger claim was made by Gutiérrez-Rexach (2001); according to him *unos* is not only unable to induce distributive dependencies over some other indefinite in its scope but actively blocks distributive dependencies:

- (84) Unos estudiantes se comieron una tarta \* cada uno.  
unos students clitic ate a cake each one

Gutiérrez-Rexach (2001)

(84) shows that *unos* cannot combine with an overt distributive operator like *cada uno* ('each') which suggests it actively blocks distributive dependencies.

The assumption that *unos* is the inverse of the BE type-shift is compatible with Villalta's claim that it cannot induce distributive dependencies. Gutiérrez-Rexach's claim goes farther though and gives *unos* extra semantic content which crucially endangers the article analysis I propose. The question that imposes itself then is whether or not Gutiérrez-Rexach's claim is tenable. I argue that it is not. My motivation comes from the fact that his claim makes the wrong prediction that there is no way in which DPs headed by *unos* will get a distributive interpretation or be

compatible with distributive operators like *cada uno* ('each'). (85) and (86) are counterexamples to this prediction:

- (85) Unos cuantos invitados se comieron un plato de jamón.  
unos many guests SE eat a plate of ham  
'Many guests (each) ate a plate of ham'  
Lopez-Palma (2007)
- (86) [...] quedaron sobre las armas unos treinta mil hombres, ganando cada uno un carlino  
stayed on the arms unos thirty thousand men, making each one a carline (=coin)  
data drawn from the Corpus del Español

What I did in (85) and (86) was to add quantifiers within the DP headed by *unos* that are in general able to induce distributive dependencies. If *unos* were indeed actively blocking distributive readings (strong claim) this should have no effect. If however *unos* is merely unable to induce distributive dependencies (weak claim) this should make a difference. The availability of a distributive interpretation of (85) and the acceptability of (86) strongly suggests it's the weaker claim that is on the right track. This leaves me with the unacceptability of (84) to which I turn presently.

#### 4.3. A note on *unos* and distributive operators

Gutiérrez-Rexach opposes the unacceptability of (84) to the acceptability of (87):

- (87) Algunos estudiantes se comieron una tarta cada uno.  
some students clitic ate a cake each one

When I asked my informant whether she agreed with the judgements of (84) and (87) she told me that neither of them sounded acceptable to her. This is relevant given that it suggests there is something weird going on with the example and not necessarily with the combination of *unos* and *cada uno*. Corpus data from the CREA corpus back this up:<sup>6</sup>

- (88) la proximidad de **unos grandes depósitos** de gasolina, con veinticuatro millones de litros de capacidad cada uno  
the proximity of unos big deposits of oil, with twenty-four million liters of capacity each ...  
'some big oil deposits with a capacity of twenty-four million liters each'

<sup>6</sup> I did a context search on the (mainland) Spanish part of the CREA corpus looking for *cada uno* in a 15 word window around *unos*. I got 209 hits most of which were not relevant. 3 of them involved *cada uno* as a distributive operator and 8 of them were of the following type:

(i) Isaias vio unos Serafines delante del trono de DIOS y cada uno tenía seis alas.

Isaias saw unos Seraphims in front of the trone of GOD and each one had six wings  
A similar search for *algunos* gave 55 hits, 3 of them relevant and each of the type in (i).

- (89) **unos espectadores** que [...] habían apoquinado sus buenas pesetas cada uno para ver un "espectáculo deportivo"...
- unos spectators that had paid their good pesetas each one to see a spectacle sports
- ‘Some spectators that had each spent their last penny to see a great sports event’

Two other ‘overtly distributive’ expressions have been claimed not to be compatible with *unos* and to be compatible with *algunos*: *uno a uno* (‘one by one’) (Martí 2008) and *sendos* (‘each one’) (Lopez-Palma 2007). I used the CREA and the CORDE corpus to check whether this is indeed the case.<sup>7</sup> As for *uno a uno* I did not find any example for *unos* nor for *algunos*. As for *sendos* I found two examples for *unos* and three for *algunos*:<sup>8</sup>

- (90) **unos medallones** antiguos con sendos relieves de pasta de marfil
- unos medals old with each\_one reliefs of pasta of marble
- ‘Some old medals each with its own marble paste relief’
- (91) Bertone ha preferido, contra las tendencias actuales, **unos faros** cubiertos por sendos cristales en lugar de un sistema escamoteable
- Bertone has preferred, against the tendencies present-day, unos headlights covered with each\_one windows in place of a system that\_can\_be\_stashed\_away
- ‘Bertone [a car designer] preferred, against present-day tendencies, headlights each covered by a window over a pop-up system’
- (92) Sacudiendo sendos abanicazos a **algunos hombres**.
- waving each\_one fans at some men
- ‘waving a fan to each of some men.’
- (93) Va abriendo **algunos paquetes** de piñones, almendras, etc. que sirve en sendos platitos
- goes opening some packages of pine\_tree\_pits, almonds, etc. that he\_serves on each\_one small\_plates
- ‘He opens some packages of pine tree pits, almonds, etc. that he serves each on small plates.’

<sup>7</sup> The CORDE corpus is freely available online (<http://corpus.rae.es/cordenet.html>). It is the historical twin of the CREA corpus and spans the whole history of the Spanish language.

<sup>8</sup> I did context searches on the (mainland) Spanish part of the CREA corpus and the (mainland) Spanish 20<sup>th</sup> century part of the CORDE corpus looking for *uno a uno* and *sendos* in a 15 word window around *unos* and *algunos*. The combination of *algunos* / *Algunos* and *uno a uno* gave 1 hit, which was irrelevant. The combination of *unos* / *Unos* with *uno a uno* gave 11 hits, all of them irrelevant. The combination of *algunos* / *Algunos* with *sendos* gave 4 hits, 3 of them relevant. The combination of *unos* / *Unos* with *sendos* gave 16 hits in total, two of them relevant.

- (94) **Algunos flecos** y sendos viajes del Piqué a Chile y Alemania impidieron que se firmara el acuerdo esta semana  
some fringes and each\_one journeys from el Piqué to Chili and Germany prevented the agreement from being signed this week  
'Some fringes and corresponding journeys from el Piqué to Chili and Germany prevented the agreement from being signed this week'

The examples presented in (90)-(84), together with the uncontested examples in (85) and (86) strongly suggest that the unacceptability of (84) should not be taken as evidence in favour of the claim that *unos* actively blocks distributive dependencies.

#### 4.4. Recap

In this section I argued that Partee's analysis of the indefinite article as the inverse of the BE type-shift makes the prediction that articles should not have a distributive semantics. I showed that this prediction was borne out for *unos* to the extent that it has a strong preference for collective readings but crucially does not block distributive dependencies.

#### 5. Further predictions

In the formal semantics literature two further tests have been proposed to distinguish articles from other semantically close determiners. Even though they were originally designed for the singular indefinite article they can be transposed to the plural. In this section I will briefly show that *unos* behaves as expected.

##### *Test 1*

According to Krifka (2004) the crucial difference between articles and numerals is that numerals are related to a number scale whereas this relation is absent in the case of indefinite articles. This can be established in a number of ways but one of the most compelling ones is that a DP headed by the indefinite article cannot be used as an answer to a *how many* question whereas a DP headed by a numeral can.

- (95) How many dogs did you see? I saw one dog.  
(96) How many dogs did you see? #I saw a dog.

A similar test shows that *unos* is not related to the number scale whereas its close relative – *algunos* ('some') – is:

- (97) ¿Cuántos libros tienes? ??Unos./ Algunos.<sup>9</sup>  
How\_many books you\_have Unos / Algunos

Gutiérrez-Rexach (2001)

A further indication that *unos* does not have any intrinsic number specification that would relate it to a number scale is that it can freely be added to NPs containing any quantity imaginable:

- (98) unos pocos                    ('unos few')  
      unos cinco                 ('unos five')  
      unos mil quinientos      ('unos one thousand five hundred')  
      unos cuantos              ('unos many')

data drawn from the CREA corpus

Note that even though *unos* is not directly related to a number scale it sometimes does seem to acquire a quantity specification via a pragmatic route. This is not unexpected; given that *unos* is the only indefinite D without a quantity specification it can serve in those contexts in which the speaker does not want to or cannot specify the exact quantity of N. In these cases the context will give some clues about the quantity involved. This type of analysis is supported by examples such as (99) in which the speaker first utters *unos días* ('unos days') and then specifies *pocos días* ('few days'):

- (99) Porque ella es más pequeña y también su papá la va a dejar por unos días,  
      pocos días, ya lo verás... [p. 70]  
      Because she is more small and also her daddy her goes to leave for unos  
      days, few days, already it you'll\_see

Mujeres de negro

#### Test 2

A second test to distinguish articles from other indefinites is that it should be able to appear in the scope of a generic operator (see Kratzer 1996, Chierchia 1998, Farkas 2002). This test was originally proposed to distinguish the English singular indefinite article from *some*:

- (100) \*Some seagull lays eggs in the sand.

- (101) A seagull lays eggs in the sand.

Farkas (2002)

The unacceptability of *some* in (100) is arguably due to the fact that it has an extra semantic specification that blocks it from appearing in that position (see Kratzer

---

<sup>9</sup> My informant points out that *unos* could be used as an answer to indicate that the speaker cannot or does not want to specify the quantity of books. This is in line with my line of argumentation.

1996 and Farkas 2002 for proposals). As shown in (102) Spanish *unos* behaves as an article on this test as well:<sup>10</sup>

- (102) [No puedo creer que esos dos sean millonarios] Unos millonarios no viajan en segunda clase.  
[I cannot believe that these two are millionaires] Unos millionaires do not travel second class.

Laca & Tasmowski (1996)

## 6. Conclusion

In this chapter I argued on the basis of synchronic facts that *unos* should be analyzed as the Spanish plural indefinite article. More specifically, I first argued that plural indefinite articles in a language that allows for bare plurals are expected to have the following characteristics:

- (i) they are more salient than the bare plural in argument position and, as a consequence, show PPI behaviour (section 2)
- (ii) they are marked in predicate position and are used for metaphorical predication (section 2)
- (iii) they do not have partitive content (section 3)
- (iv) they cannot induce distributive dependencies (section 4)
- (v) they are not related to a number scale (section 5)
- (vi) they do not have semantic content that blocks them from scoping below a generic operator (section 5)

I then showed that the characteristics (i) through (vi) correspond exactly to those of *unos* and accordingly concluded that *unos* is the Spanish plural indefinite article. In the next chapter I will argue that two diachronic predictions an article analysis makes are also borne out. This will complete my case in favour of an article analysis of *unos*.

---

<sup>10</sup> Note that the only point I try to make is that *unos N* can be embedded under a generic operator. I leave aside the discussion on why Spanish prefers definite plurals over indefinite plurals.



### Chapter 3: *Unos* as a plural indefinite article: diachronic evidence

In Chapter 2 I argued that *unos* behaves exactly as we would expect a plural indefinite article to behave in a language that allows for bare plurals. This chapter will make a similar case for *unos* but from a diachronic point of view. More specifically I will argue that there are two crucial predictions an article analysis of *unos* makes (section 1) and that these predictions are borne out (sections 2 and 3). I will furthermore show how we can simulate the evolution of articles in general which is important both to understand the mechanisms behind language evolution and to check whether language evolution can happen the way we think it does (section 4).

#### 1. The grammaticalization of indefinite articles

The grammaticalization of indefinite articles is assumed to be a three-stage process:

(103)

Stage I:

All indefinites are unmarked;

Stage II:

The new marker codes only pragmatically important indefinites;

Stage III:

The new marker gradually expands its scope to mark all indefinites.

adapted from Givón 1981

#### Stage I

In the first stage nouns can appear bare and the grammaticalizing item might have a semantics that makes it more marked than the bare form. For the indefinite article in English we know e.g. that it originated in the numeral *one* that can be distinguished from the indefinite article in allowing for partitive readings:

(104) A student is intelligent. [<sup>\*</sup>non-generic]

(105) One student is intelligent. [<sup>ok</sup>non-generic]

#### Stage II

In the second stage nouns can still appear bare but the grammaticalizing item gets a more general use marking pragmatically important indefinites. This second stage is generally accompanied by an increase in frequency and semantic bleaching – a process whereby the grammaticalizing item loses part of its semantic content. For

the indefinite article in English this process of semantic bleaching could e.g. be thought of as the loss of its partitive potential.

### Stage III

The third and final stage is characterized by the generalization of the use of the grammaticalizing item to all indefinites. This last stage is in principle accompanied by another increase in frequency of the grammaticalizing item.

On the basis of my claim that *unos* is an indefinite article but is crucially different from the singular indefinite article in not having superseded bare nominals I make the following prediction about their respective evolution:

(106) **Prediction 1**

*Unos* underwent one important increase in frequency (from Stage I to Stage II) whereas the indefinite singular article underwent two (from Stage I to Stage II and from Stage II to Stage III).

A further prediction my article claim about *unos* makes is formulated in (107):

(107) **Prediction 2**

If *unos* originally had a special semantics (e.g. the possibility to induce partitive readings) this semantics should disappear around the time *unos* underwent its increase in frequency.

In sections 2 and 3 I will show that these predictions are borne out.

Note that grammaticalization is also connected to other phenomena (like phonetic reduction, agglutination, morphological reduction, phonological reduction, reduction in syntactic combination possibilities, ...). Frequency increase and semantic bleaching are somewhat special though. Indeed, frequency increase is a necessary condition on grammaticalization (unlike most other phenomena) and the loss of semantic content seems particularly crucial in the transition from a standard, semantically marked determiner to an unmarked article.

## 2. The grammaticalization of *unos*: increase in frequency

In this section I will show that the following prediction an article grammaticalization analysis of *unos* makes is borne out:

**Prediction 1**

*Unos* underwent one important increase in frequency whereas the indefinite singular article underwent two.

I will show that the first increase in frequency of the numeral and *unos* coincide (section 2.1.) and that the numeral evolved further (section 2.2).

## 2.1. The joint increase in frequency of the numeral *one* and *unos*

In this subsection I will show that the numeral *one* and *unos* have undergone an important increase in frequency in the 16<sup>th</sup> century. I first look at the numeral.

### 2.1.1. The frequency of the numeral

I will present frequency data from four independent corpus studies: Lapesa (1975), Pozas-Loyo (*in preparation*) and two of my own.

#### 1. Lapesa (1975)

The first study presenting diachronic frequency data for the numeral is Lapesa (1975). The frequency data – relative to the number of nouns – are given in Table 1.<sup>1</sup>

		<b>Relative frequency</b>
Mio Cid (600 verses)	1140	1.26%
Quijote (Part I, Chapter 1)	1605	5.10%

Table 1: Frequency data from Lapesa (1975)

What these data suggest is that there was an important gain in frequency of the numeral from 1140 to 1605. Small as the corpus might seem (in total some 6800 words) it does show a tendency that will return in larger scale corpus searches.

#### 2. Pozas-Loyo (*in preparation*)

Pozas-Loyo (*in preparation*) looked at a 225.000 word corpus that was structured as follows: for the 13<sup>th</sup>, 15<sup>th</sup> and 17<sup>th</sup> century she took 5 texts each and from each text she selected a 15.000 word sample. This is schematically rendered in Table 2.

<b>Century</b>	<b># texts</b>	<b># words / text</b>	<b>total</b>
13 <sup>th</sup>	5	15.000	75.000
15 <sup>th</sup>	5	15.000	75.000
17 <sup>th</sup>	5	15.000	75.000

Table 2: Corpus of Pozas-Loyo (*in preparation*)

Table 3 presents Pozas-Loyo's frequency data. Note that the frequency in this table is relative to the number of words and not – as in Lapesa's study – relative to the number of nouns.

---

<sup>1</sup> I left out the data of the *Disputa* that – according to Lapesa – are not representative because of the shortness of the text (324 words). Note that Lapesa did not distinguish between the singular indefinite article and *unos*. Given that this has little influence on the results I will ignore this complication.

Century	Relative Frequency (per million words)
13 <sup>th</sup>	3280
15 <sup>th</sup>	3680
17 <sup>th</sup>	9813

Table 3: Frequency data from Pozas-Loyo (in preparation)

Pozas-Loyo's results confirm the tendency noted by Lapesa but narrow down the increase in frequency to the sixteenth and/or seventeenth century.

### 3. The CORDE corpus and the Corpus del Español

Given the interest of Pozas-Loyo in a qualitative study the size of her corpus is still fairly modest. If we however fully opt for a quantitative approach we can make use of some vast corpora that have recently become available: the CORDE corpus (235 million words) and the Corpus del Español (+100 million words). The number of words per century for each corpus is given in table 4.

Century	CORDE # words	Corpus del Español # words
13 <sup>th</sup>	8,9 million	7,9 million
14 <sup>th</sup>	8,8 million	3,0 million
15 <sup>th</sup>	22,2 million	9,7 million
16 <sup>th</sup>	50,6 million	19,7 million
17 <sup>th</sup>	36,4 million	14,8 million

Table 4: The CORDE corpus and the Corpus del Español

In table 5 I present the frequencies found for the numeral.<sup>2</sup> The relative frequencies in this table are also relative to the number of words.

Century	CORDE Relative Frequency (per million words)	Corpus del Español Relative Frequency (per million words)
13 <sup>th</sup>	3.718	2.278
14 <sup>th</sup>	4.745 (+28%)	3.198 (+40%)
15 <sup>th</sup>	4.911 (+4%)	3.097 (-3%)
16 <sup>th</sup>	7.036 (+43%)	5.373 (+73%)
17 <sup>th</sup>	8.328 (+18%)	6.745 (+26%)

Table 5: Frequency data from the CORDE corpus and the Corpus del Español

<sup>2</sup> Given that the CORDE corpus search engine is case-sensitive I looked for attestations of *un*, *Un*, *vn*, *Vn*, *una*, *Una* and *vna*, *Vna*. Note that this also includes all non-determiner uses of these forms. In the Corpus del Español which is tagged for parts of speech and whose search engine is not case-sensitive I looked for *un*, *vn*, *una* and *vna* followed by a singular noun.

Both separately and jointly these data confirm that the numeral has undergone an important increase in frequency and narrow down the period in which the most important increase took place to the 16<sup>th</sup> century.

#### 4. Recap

On the basis of 4 independent corpus studies I have shown that the numeral increases in frequency and that the most important century for this increase is the 16<sup>th</sup> century. If *unos* did indeed evolve in the same way we expect it to behave similarly. To show that this is the case is the goal of 2.1.2.

#### 2.1.2. The frequency of *unos*

I will present frequency data from three independent corpus studies to show that *unos* patterns with the numeral in undergoing an increase in frequency. I will also show that the most important century for this increase is the 16<sup>th</sup> century.

Before looking at the data I should point out that *unos* is expected to be far less frequent than the numeral. Indeed, given that the plural determiner paradigm is far more developed than the singular one we expect *unos* to have a smaller application domain.

##### 1. Pozas-Loyo (to appear)

For an overview of how Pozas-Loyo's corpus is composed see 2.1. The frequency data she obtained are presented in table 6:

Century	Relative Frequency (per million words)
13 <sup>th</sup>	226
15 <sup>th</sup>	253
17 <sup>th</sup>	680

Table 6: Frequency data from Pozas-Loyo (in preparation)

As was the case for the numeral in Pozas-Loyo's corpus we see an important increase in frequency from the 15<sup>th</sup> to the 17<sup>th</sup> century. This tightens the link between the numeral and *unos*. Note by the way that even the size of the increase is perfectly comparable; the indefinite singular article underwent an increase of 167%, *unos* an increase of 169%.

##### 2. The CORDE corpus and the Corpus del Español

For an overview of how the CORDE corpus and the Corpus del Español are composed see 2.1. The frequency data I obtained from both are presented in table 7:

Century	CORDE Relative Frequency (per million words)	Corpus del Español Relative Frequency (per million words)
13 <sup>th</sup>	456	135
14 <sup>th</sup>	331 (-27%)	93 (-30%)
15 <sup>th</sup>	434 (+31%)	134 (+44%)
16 <sup>th</sup>	679 (+56%)	368 (+175%)
17 <sup>th</sup>	707 (-16%)	311 (-16%)

Table 7: Frequency data from the CORDE corpus and the Corpus del Español

In order to interpret these data in the right way two remarks are in order. The first is that we expect a discrepancy between both corpora. This is due to the fact that the Corpus del Español is tagged for parts of speech whereas the CORDE corpus is not. The consequence of this is that the CORDE corpus did not allow me to automatically filter out the uses of *unos* in which it does not function as a determiner. Given that I did not try to do this manually either the frequencies from the Corpus del Español are expected to be lower. The second remark is that the overall frequency of *unos* is fairly low compared to that of the numeral. This makes the frequencies more dependent on the choice of texts and might lead to some unexpected frequency alterations. Taking these two remarks into account I think the data clearly show that *unos* underwent an important increase in frequency and that the most important century for this increase is the 16<sup>th</sup> century. This is in accordance with Pozas-Loyo's data and parallel to the data we obtained for the numeral.

### 3. Recap

On the basis of 3 corpus studies I have shown that *unos* behaves parallel to the numeral in undergoing an increase in frequency and most notably in the 16<sup>th</sup> century. This confirms the first part of prediction 1 according to which the numeral and *unos* are expected to undergo a joint increase in frequency. In 2.2. I will show that the second part of prediction 1 is also borne out, *viz.* that the numeral will undergo a second increase in frequency. Taken together the frequency data in this and the next subsection strongly suggest that both the numeral and *unos* underwent a process of grammaticalization even though *unos* did not grammaticalize as far as the numeral.

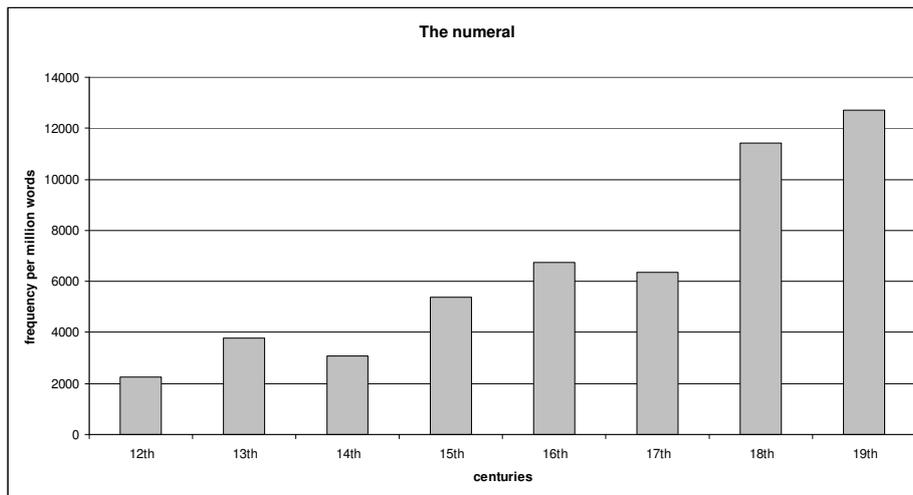
### 2.2. The extra increase in frequency of the numeral<sup>3</sup>

I present frequency data from two corpus studies to show that the numeral underwent a second increase in frequency that is absent in the case of *unos*. Given that neither Pozas-Loyo nor Lapesa looked into the frequencies of the numeral or

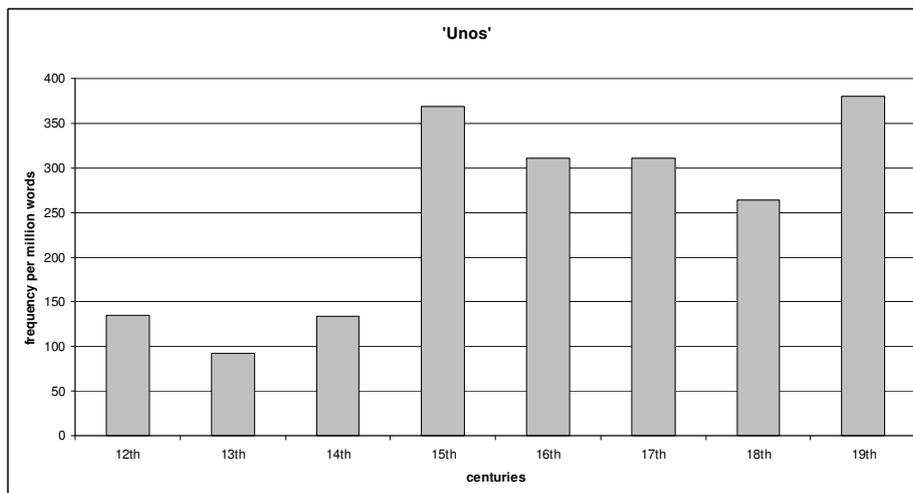
---

<sup>3</sup> Note that the data presented in this section are not only interesting for the study of the grammaticalization of indefinite articles but also more generally for the diachronic study of Spanish. Indeed, the fact that the indefinite article underwent a second important increase in frequency in the 19<sup>th</sup> century seems to support the recently proposed division of the history of Spanish into three periods instead of the standard two (medieval / modern vs. medieval / classic / modern). See Melis, Flores & Bogard (2003).

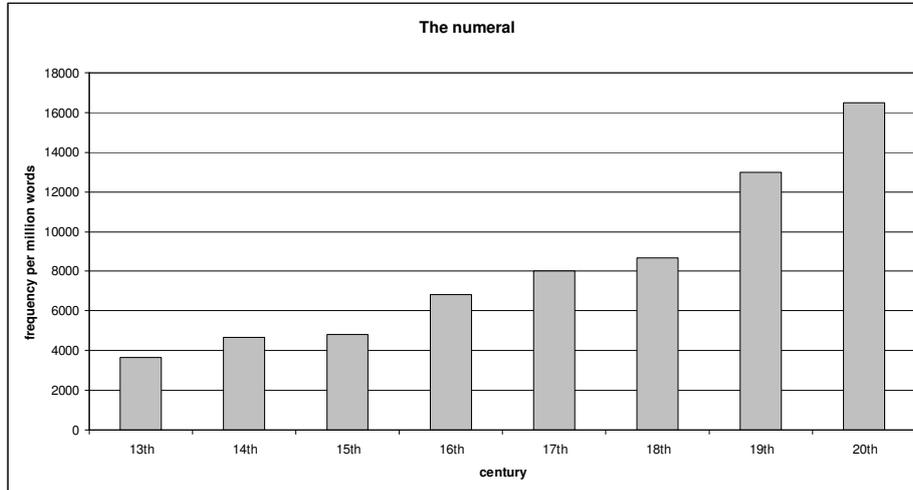
*unos* after the first increase I will restrict myself to the frequency data that can be drawn from the Corpus del Español and the CORDE corpus. For convenience I present them in the form of graphs. Graphs 1 and 2 show the respective evolutions in frequency of the numeral and *unos* in the Corpus del Español from the 12<sup>th</sup> to the 19<sup>th</sup> century and Graphs 3 and 4 do this for the CORDE corpus.



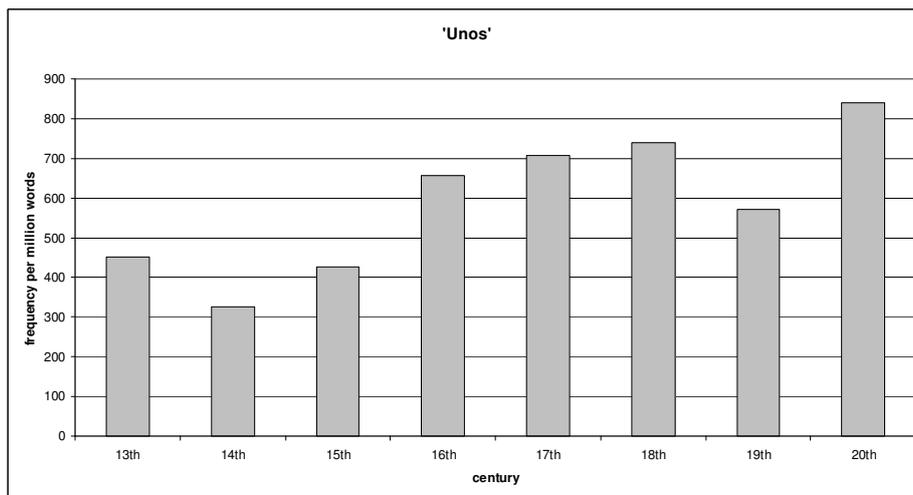
Graph 1



Graph 2



Graph 3



Graph 4

The data drawn from the Corpus del Español and the CORDE corpus show that *unos* did not undergo a second increase in frequency after the 16<sup>th</sup> century and that the numeral did. Together with the frequency data presented in 2.1. this shows that the first prediction a grammaticalization analysis of *unos* makes is borne out:

#### Prediction 1

*Unos* underwent one important increase in frequency whereas the indefinite singular article underwent two.

In the following section I will look at the second prediction that is concerned with the loss of semantic content.

### 3. The grammaticalization of *unos*: semantic bleaching

In this section I will show that the second prediction a grammaticalization analysis of *unos* makes is borne out:

#### Prediction 2

If *unos* originally had a special semantics (e.g. the possibility to induce partitive readings) this semantics should disappear around the time *unos* underwent its increase in frequency.

The special semantics I will be looking at is the one suggested, *viz.* the partitive potential of *unos*. The way to check whether the prediction is borne out was sketched in Chapter 2 section 3.5: performing a corpus search looking for attestations of *unos* in the upstairs D position of partitives.

The corpus I used for this search is the CORDE corpus.<sup>4</sup> As expected we do indeed find partitives headed by *unos*: the earliest one dates back to 1260 and the most recent one to 1623 (n=35).<sup>5</sup> This final date coincides more or less with the end of the period during which *unos* underwent its increase in frequency (see 2.1.2). In (108) through (111) I present four examples (the partitive is in bold, the element in the sentence indicating that a plural referent was intended is underlined):

- (108) como **unos de los nuestros** uinieron con armas batallasas contra el rey  
Laomedon  
how unos of the ours came with weapons fighting against the king  
Laomedon

Historia troyana (1376)

---

<sup>4</sup> The search string I used was *unos de los* which yielded 768 hits: 659 were not partitive, 13 were clearly typos. The remaining 96 can be divided into 4 groups. The first one contains those for which we have no way to check whether the intended number was plural or singular (n=10). The second group consists of partitives headed by *cada unos* ('each ones') but clearly refer to singular entities (n=37). I will not consider this group any further but do want to note that we are probably not dealing with a typo (we would expect the same kind of typo to occur across centuries but *cada unos* appears to be restricted to early stages of Spanish) (see also Pozas-Loyo, *in preparation*). The third group is the one containing superlatives which are also allowed in present-day Spanish and will therefore not count as evidence in favour of the possibility of *unos* to occur. The fourth group is that of partitives headed by *unos* for which there are clear indications that the intended number was plural (n=35).

<sup>5</sup> I made one important change in the dates: the CORDE corpus only contains the 1909 version of the Biblia Reina-Valera but given that this bible is known for its extreme conservativity and given that all occurrences of *unos de los* also occur in its 1569 version I think it's a justified move to count it as going back to 1569. Note that the 1569 version is not part of CORDE.

- (109) y todos ellos de edad madura y que parecían como **unos de los que imaginamos senadores de Roma**  
and all those of age ripe and that look like unos of the ones that we imagine senators of Rome  
Historia de las Indias (1527)
- (110) vio que **unos de los que a él avían robado ahorcavan** a otro de los que a él avían defendido  
he saw that unos of those that from him had stolen killed another of those that had defended him  
Libro áureo de Marco Aurelio (1528)
- (111) Entre los Canes y Sultanes que se escaparon de las manos del rey Xabas [...] fueron Mahamet Xerif Can y Sultán Mahamud y Azem Bec Golgachi, **unos de los que se escaparon**  
Between the Canes and Sultans that themselves escaped from the hands of the king Xabas were Mahamet Xerif Can and Sultan Mahamud and Azem Bec Golgachi, unos of those that themselves escaped  
Relaciones (1604)

Convincing as these examples might seem to some they might not be sufficient for the skeptical. One potential criticism is that they might all be accidental examples taken from different texts. In order to overrule this criticism I looked at whether there are texts that contain multiple examples of *unos* heading partitives. This is indeed the case: all four texts listed in Table 8 contain at least three occurrences of *unos* heading a partitive. Next to their approximate date, the number of occurrences of *unos* heading a partitive and their number of words I also indicate how many partitives they have headed by *algunos* ( $\approx$  ‘some’). Given that *algunos* is semantically very close to *unos* this is an indication of the potential domain of *unos*.

date	title	# words	# partitives <i>unos</i>	# partitives <i>algunos</i>
1260	<i>The Gospel according to manuscript I-j-6 of the Escorial</i> (Saint Matthew - Apocalypse)	172598	9	9
1275	<i>General Estoria. Primera parte</i>	556163	4	24
1280	<i>General Estoria. Cuarta parte.</i>	452515	3	19
1569	<i>The Reina-Valera bible</i>	733105	7	37

Table 8: Texts with multiple occurrences of partitives headed by *unos*

In (112)-(115) I give some more examples taken from these texts:

- (112) E ellas yendo se, fueron unos de los guardadores a la ciudad e dixieron a los obispos de los sacerdotes todo lo que acaeciera.  
And they leaving themselves went unos of the guards to the city and told to the helpers of the priests everything that had happened  
El Nuevo Testamento según el manuscrito escorialense I-j-6 (1260)

- (113) E fallamos segund lo retraye Plinio e el Libro de las provincias que **unos de los sabios** que se trabajaron de hablar de las particiones de la tierra que querién fazer  
 And we find as it portrays Plinius in the Book of the provinces that unos of the wise\_ones that themselves worked of talking of the parts of the earth that wanted make  
 General Historia. Primera Parte (1275)
- (114) **unos de los caualleros mancebos esforçados** tomaron sus espadas  
 unos of the knights young men strong took their swords  
 General Historia. Cuarta Parte (1280)
- (115) Y oyéndole **unos de los que estaban allí**, decían:  
 and hearing\_him unos of those that were there said  
 Biblia Reina-Valera (1569)

The skeptical might have one last objection, *viz.* that the two texts with most occurrences of *unos* – *The Gospel according to manuscript I-j-6 of the Escorial* and *The Reina-Valera bible* – are translations. The potential problem translations pose is that the original text might have influenced the translation. Indeed, the translator might have used *unos* in order to be faithful to the original even though it needn't have been completely acceptable in the target language. In order to argue that this is not the case I will show that identical constructions in the original text not only gave rise to translations involving *unos* but also to translations involving *algunos*. This shows that the translator had an alternative and did not hesitate to use it. The hypothesis that he would have used *unos* to be faithful to the original text is thus effectively discarded.

In table 9 I compare *The Gospel according to manuscript I-j-6 of the Escorial* to its original which is assumed to be the Vulgate (the latin translation by Saint Jerome). The table is organized in the following way: in the third column I list the partitive constructions of the original text giving rise to a partitive construction with *unos* or *algunos* in the translation. The fourth column indicates whether *unos* or *algunos* was used.<sup>6</sup>

Table 9 shows is that for none of the latin partitive constructions that are recurrent it is possible to predict how it will be translated; both *algunos* and *unos* appear in their translations. It furthermore shows that *unos* is as productive in the translation of partitives as *algunos*; both *unos* and *algunos* are chosen 50 percent of the time. This strongly suggests that the partitive potential of *unos* was comparable to that of *algunos*.

In table 10 I compare the new testament part of the Biblia Reina-Valera to its original which is assumed to be the Textus Receptus (a Greek translation of the bible by Stephanus dating back to 1550).<sup>7</sup> The table is organized in the same way as table

<sup>6</sup> I used the edition of the Vulgate that is freely available online via <http://www.biblegateway.com>.

<sup>7</sup> I used the edition of the Reina-Valera bible that is freely available online via <http://www.biblegateway.com>.

9. Table 10 shows is that for the only recurrent Greek construction it is not possible to predict how it will be translated; both *algunos* and *unos* appear in its translations.

	Reference	Vulgate	Gospel according to manuscript I-j-6 of the Escorial
1	Romans 11:17	<i>aliqui + ex + plural</i>	<i>algunos</i>
2	John 9:40	<i>ex + plural</i>	<i>algunos</i>
3	Matthew 9:3	<i>quidam + de + plural</i>	<i>algunos</i>
4	Matthew 12:38	<i>quidam + de + plural</i>	<i>unos</i>
5	Matthew 28:11	<i>quidam + de + plural</i>	<i>unos</i>
6	Mark 2:6	<i>quidam + de + plural</i>	<i>algunos</i>
7	Mark 15:35	<i>quidam + de + plural</i>	<i>algunos</i>
8	Luke 24:24	<i>quidam + ex + plural</i>	<i>unos</i>
9	John 9:16	<i>quidam + ex + plural</i>	<i>unos</i>
10	Acts 15:2	<i>quidam + ex + plural</i>	<i>algunos</i>
11	Acts 23:12	<i>quidam + ex + plural</i>	<i>unos</i>
12	Luke 6:2	<i>quidam + genitive plural</i>	<i>algunos</i>
13	Luke 13:31	<i>quidam + genitive plural</i>	<i>unos</i>
14	Luke 20:27	<i>quidam + genitive plural</i>	<i>unos</i>
15	Luke 20:39	<i>quidam + genitive plural</i>	<i>unos</i>
16	Acts 23:9	<i>quidam + genitive plural</i>	<i>unos</i>
17	Mark 12:13	<i>quosdam + ex + plural</i>	<i>algunos</i>

Table 9: Comparison between partitives in the Vulgate and the Gospel according to manuscript I-j-6 of the Escorial

	Reference	Textus Receptus	Reina-Valera
1	Mark 12:13	<i>tinas + genitive definite plural</i>	<i>algunos</i>
2	Acts 21:16	<i>genitive definite plural</i>	<i>algunos</i>
3	Acts 19:13	<i>tines + apo + definite plural</i>	<i>algunos</i>
4	John 7:25	<i>tines + ek + definite plural</i>	<i>unos</i>
5	Matthew 9:3	<i>tines + genitive definite plural</i>	<i>algunos</i>
6	Matthew 12:38	<i>tines + genitive definite plural</i>	<i>algunos</i>
7	Mark 2:6	<i>tines + genitive definite plural</i>	<i>algunos</i>
8	Mark 7:1	<i>tines + genitive definite plural</i>	<i>algunos</i>
9	Mark 11:5	<i>tines + genitive definite plural</i>	<i>unos</i>
10	Mark 15:35	<i>tines + genitive definite plural</i>	<i>unos</i>
11	Luke 6:2	<i>tines + genitive definite plural</i>	<i>algunos</i>
12	Luke 9:27	<i>tines + genitive definite plural</i>	<i>algunos</i>
13	Luke 19:39	<i>tines + genitive definite plural</i>	<i>algunos</i>
14	Luke 20:27	<i>tines + genitive definite plural</i>	<i>unos</i>
15	Luke 20:39	<i>tines + genitive definite plural</i>	<i>unos</i>

16	Luke 24:24	<i>tines</i> + genitive definite plural	algunos
17	Acts 10:23	<i>tines</i> + genitive definite plural	algunos
18	Acts 19:31	<i>tines</i> + genitive definite plural	algunos
19	Acts 23:12	<i>tines</i> + genitive definite plural	algunos
20	Romans 11:17	<i>tines</i> + genitive definite plural	algunos

Table 10: Comparison between the *Textus Receptus* and the *Reina-Valera* bible

From the above I conclude that *unos* used to allow for partitive readings which is an argument in favour of the fact that *unos* – like the numeral – underwent a process of semantic bleaching that ended around the same time *unos* underwent its increase in frequency. This means the second prediction a grammaticalization analysis for *unos* makes is borne out:

### Prediction 2

Insofar as *unos* and the numeral were standard indefinites having a partitive semantics they will lose part of their semantic content.

## 4. Simulating the evolution of indefinite articles

In sections 2 and 3 I argued that *unos* behaves in diachrony exactly in the way we expect an indefinite article to behave in a language that (in synchrony) allows for bare plurals. More specifically I argued that two crucial predictions a grammaticalization analysis makes are borne out.

This section is concerned with the question whether the evolution of indefinite articles can be simulated and if so which assumptions we should make it possible. The goal of evolutionary simulations is twofold: on the one hand they help us to understand which mechanisms are at play in language evolution and on the other hand they allow us to check whether the changes that occurred can have occurred for the reasons we think they did. I will be using EvolOT, a program that takes constraint reranking to be at the heart of evolution (Jäger 2002). The main challenge will be to come to a simulation in which we can incorporate the different stages of article grammaticalization (repeated below for convenience) and in which we can distinguish between articles that went all the way to stage III and articles that only went up till stage II.

(116)

Stage I:

All indefinites are unmarked;

Stage II:

The new marker codes only pragmatically important indefinites;

Stage III:

The new marker gradually expands its scope to mark all indefinites.

adapted from Givón 1981

In 4.1. I give a brief introduction to EvolOT. 4.2. discusses the problems a simulation of the grammaticalization of articles faces. In 4.3. I propose solutions to these problems and in 4.4. I conclude.

#### 4.1. EvolOT

EvolOT works in cycles, each cycle having two components: a learning component (4.1.1.) and a production component (4.1.2.). An iteration of the two gives rise to a model of evolution (4.1.3.).

##### 4.1.1. The learning component

Imagine we have a language with two forms A and B and two meanings A' and B'. Imagine furthermore that we have to learn how to use this language on the basis of a fully disambiguated corpus with the following frequencies of form-meaning pairs and a set of unranked constraints:

(117)	<i>form / meaning pairs</i>		<i>constraints</i>
	<A,A'> 0	1:	(o==A&i==A')
	<A,B'> 0	2:	(o==A&i==B')
	<B,A'> 100	3:	(o==B&i==A')
	<B,B'> 100	4:	(o==B&i==B')

The constraints given here are fairly simple; they each penalize one form/meaning pair (*o* stands for 'output' (or form) and *i* stands for 'input' (or meaning)). I will refer to this type of constraints as *bias constraints*.

Suppose now that we learn using the technique in (118):

(118) Bi-directional learning

We take a random form-meaning pair from the corpus and ...

I. ...make a prediction about what meaning would correspond to the form. If it turns out that our prediction is correct we do nothing but if it turns out that we are wrong we remember this and next time we have to interpret the same form we will favour the other interpretation a bit more. If we e.g. took an instantiation of <B,A'> and wrongly predicted B' to be the meaning we will favour the meaning A' a bit more. We do this by promoting the constraint that penalizes <B,B'> and by demoting the constraint that penalizes <B,A'> :

(i)  $\uparrow$  constraint 4 (ii)  $\downarrow$  constraint 3

II. ...make a prediction about what form would correspond to the meaning. If it turns out that our prediction is correct we do nothing but if it turns out that our prediction is wrong we remember this and next time we have to give a form for the same meaning we will favour the other form a bit more. This is done in the same way as illustrated in I.

We repeat the above process  $X$  times.

The effect of this learning technique and the input corpus will be that we will strongly favour  $\langle B, A' \rangle$  and  $\langle B, B' \rangle$  over  $\langle A, A' \rangle$  and  $\langle A, B' \rangle$ .

#### 4.1.2. The production component

After the completion of the learning process EvolOT starts its production component. This production component consists of an algorithm that randomly generates a set of form-meaning pairs. The frequency of each form-meaning pair depends on the preferences we obtained during the learning process. For our running example it would mean that  $\langle B, A' \rangle$  and  $\langle B, B' \rangle$  will be produced far more often than  $\langle A, A' \rangle$  and  $\langle A, B' \rangle$ .

The production component ends when for each meaning the number of times it was expressed in the original corpus is reached.

#### 4.1.3. Evolution

The output of the production component is a new corpus that can be used to start a new cycle. These cycles can be thought of as generations of speakers of a language, each learning the language and producing a corpus that serves as an input for the next generation. An iteration of cycles then corresponds to a simulation of language evolution. Figure 1 gives a pictorial representation of the cyclic set-up of EvolOT:

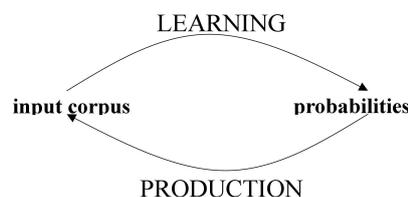


Figure 1: The cyclic set-up of EvolOT

#### 4.1.4. Recap

In this subsection I have briefly introduced EvolOT, a piece of software that models historic evolution and consists of two components that apply in a cyclic fashion: (i) a

learning algorithm that ranks predefined constraints in such a way that they make up an adequate grammar for an input corpus and (ii) a production component that uses this grammar to produce an output corpus. In 4.2. I will discuss two problems that pop up when we want to use EvolOT to simulate the grammaticalization of indefinite articles and in 4.3. I show how these problems can be overcome.

#### 4.2. Two potential problems

Ideally the constraints introduced in Chapter 2 to account for Article, Bare and Hybrid languages should be transposable to the historic simulation I am after now. I repeat them below for convenience:

- (119) \*FORM  
Avoid form.
- FAITHTYPESHIFT  
Mark type-shifts.
- \*M-MEANINGS  
Avoid marked meanings

If these constraints were transposable it would be sufficient to give form-meaning pairs and some input frequencies and let the system do its job. This approach runs into two problems that I will discuss in this subsection. In 4.3. I will show how they can be overcome.

##### 4.2.1. Problem 1: Language variation

The first problem can be illustrated on the basis of a learning simulation involving the input corpus and constraints in (120). Note that I'm using the evolution of a standard numeral like *one* to a standard indefinite article like *a* as a running example.

(120)  
*input corpus*

		BN (o==1)	Numeral (o==2)
pred	(i==1)	120	0
arg	(i==2)	5	0
'one'	(i==3)	-1	2

*constraints*

- 1: (o==1&i==2)  
2: (o==2)

The input corpus reflects the stage before the grammaticalization of the indefinite article in English: the numeral ‘one’ is only used as a numeral and the bare nominal can be used as an argument (though not very frequently) or as a predicate (in all DPs and in predicate position) but cannot semantically encode the meaning ‘one’ (this is indicated by ‘-1’). The constraints are those we had for Bare and Article Languages; constraint 1 states that a bare noun should not be used as an argument (FAITHSHIFT) and constraint 2 penalizes form (\*FORM).

The end result of a simulation of the learning process is given in (121):

(121)

*learned constraint ranking*

constraint 1: -20.63

constraint 2: 0.75

To understand this result it suffices to note that <1,1> and <1,2> are the most frequent form-meaning pairs and that the latter but not the former will frequently give rise to wrong predictions. Indeed, given that <1,2> but not <1,1> is penalized by a constraint we expect <1,2> often to lead to a wrong meaning prediction (1 instead of 2). This will result in a demotion of constraint 1.

The problem this simulation illustrates is that – as it stands – the system does not leave room for language variation. Indeed, any language displaying a situation similar to the one sketched in the input corpus in (120) is expected to evolve in the same way. This is clearly not what we want, especially in view of my claim that there are articles that do not supersede the corresponding bare nominal (cf. Chapter 2 Section 2). One way of circumventing this problem would be to assume that in some languages only constraint 1 is active whereas in others only constraint 2 is active. This would however go counter the basic idea in OT that constraints are universal and that language variation is a result of differences in constraint ranking.

#### **4.2.2. Problem 2: New meanings**

The second problem that pops when one tries to simulate the grammaticalization of the indefinite article is that EvolOT cannot create occurrences for new meanings. This is especially problematic when one tries to include Stage II (‘the new marker codes only pragmatically important indefinites’) in the analysis. Indeed, under the assumption that ‘pragmatic importance’ should be analyzed as a separate meaning the input corpus in (122) will never give rise to occurrences of this meaning in an output corpus.

(122)

*input corpus*

		BN (o==1)	Numeral (o==2)
pred	(i==1)	120	0
arg	(i==2)	5	0
'pragmatically important' argument (henceforth <i>arg</i> + <i>M</i> )	(i==3)	-1	0
'one'	(i==4)	-1	2

One solution would be to assume that this intermediate stage is simply hidden in the transition from a Bare to an Article Language but this would limit the scope of the simulation and it would make it impossible for there to be articles that do not end up superseding the corresponding bare nominal.

### 4.3. Overcoming the problems

In this subsection I will present a simulation of the grammaticalization of the indefinite article that overcomes the problems discussed in 4.2 and conclude with some remaining challenges.

#### 4.3.1. The basics

The basis for the simulation I propose is given in (123). In 4.3.2. and 4.3.3. I will suggest some crucial modifications:

(123)

*basic input corpus*

		BN (o==1)	Numeral (o==2)
pred	(i==1)	120	0
arg	(i==2)	5	0
arg + M	(i==3)	-1	0
'one'	(i==4)	-1	2

*basic constraints*

1:	(o==1&i==1)	5:	(o==2&i==1)
2:	(o==1&i==2)	6:	(o==2&i==2)
3:	(o==1&i==3)	7:	(o==2&i==3)
4:	(o==1&i==4)	8:	(o==2&i==4)

The input corpus is the one introduced in 4.2.2. and bias constraints (see also 4.1.1.), they each penalize one form/meaning pair.

#### 4.3.2. Getting occurrences for a new meaning

The first thing to be done is to solve the problem discussed in 4.2.2., *viz.* the fact that EvolOT cannot ‘create’ occurrences for a new meaning. This specific problem can be circumvented with a small trick: we can include a ‘ghost form’ in the input corpus that conveys the new meaning but is easy to be superseded in this meaning by another form. The way to do this is to let the ghost form convey an extra meaning – call it the ‘ghost meaning’ – that is far more frequent than the other one and can only be expressed by the ‘ghost form’. The implementation of this trick leads to the following input corpus:<sup>8</sup>

(124)  
*new input corpus*

		BN (o==1)	Numeral (o==2)	‘Ghost form’ (o==3)
pred	(i==1)	120	0	-1
arg	(i==2)	5	0	-1
arg + M	(i==3)	-1	0	3
‘one’	(i==4)	-1	2	-1
‘ghost meaning’	(i==5)	-1	-1	500

Intuitively ghost forms can be thought of as expressing the need of speakers to convey a certain meaning. The ghost meaning is simply there to ensure that the ghost form will not monopolize the meaning *arg + M*; without it *one* could never supersede the ghost form in expressing *arg+M*.

Note that the addition of a form and a meaning leads to extra constraints as well. Instead of 8 constraints (2 outputs times 4 inputs) we now have 15 constraints (3 outputs times 5 inputs).

#### 4.3.3. The grammaticalization of indefinite articles

I assume there are two grammaticalization paths, the first one leading all the way to Stage III the second one only going up to Stage II. In what follows I will show that both paths can be accounted for with the same set of constraints if we allow for constraint merger, an operation that will turn out to be compatible with the OT idea

<sup>8</sup> For those familiar with EvolOT I should note that this trick works well when one only uses bias constraints but does not necessarily work when one uses other types of constraints. Why this works well with bias constraints is explained in Part II of this dissertation.

that crosslinguistic variation is due to constraint reranking. This will allow me to overcome the problem presented in 4.2.1.

*Path 1*

The start of the development of an indefinite article that will go all the way to Stage III seems to reside in the desire to mark pragmatically important (argumental) indefinites. In a second step this desire to mark extends to all (argumental) indefinites. The way I propose to derive this is to merge the constraint penalizing <Numeral, Arg + M> (i.e. (o==2&i==3)) and the constraint penalizing <Numeral, Arg> (i.e. (o==2&i==2)). The result of this will be that these constraints will be promoted and demoted together. Given that the former will be (strongly) demoted this will automatically lead to a demotion of the latter. The result will be that the numeral starts conveying the meaning ‘arg + M’ and simultaneously (but at a slower pace) will start marking arguments in general.

The set of constraints then looks as follows:

(125)

*new constraints for Grammaticalization Path 1*

1:	(o==1&i==1)	8:	(o==2&i==4)
2:	(o==1&i==2)	9:	(o==2&i==5)
3:	(o==1&i==3)	10:	(o==3&i==1)
4:	(o==1&i==4)	11:	(o==3&i==2)
5:	(o==1&i==5)	12:	(o==3&i==3)
6:	(o==2&i==1)	13:	(o==3&i==4)
7:	(o==2&i==2)	14:	(o==3&i==5)
7:	(o==2&i==3)		

Note that we no longer have 15 constraints but only 14, the original number 7 and 8 being merged. Note also that the reduction of constraints is only apparent; the only thing the merger of constraints does is to explicitly state that the merged constraints are unranked with respect to each other. This is important given that this view on merged constraints allows me to maintain the OT idea of variation being due to differences in constraint ranking.

Output corpora showing the predicted evolution are given in (126) (first four generations of a random simulation)

(126)

*Random output corpora for Grammaticalization Path 1*

		BN (o==1)	Numeral (o==2)	'Ghost form' (o==3)
pred	(i==1)	120	1	-1
arg	(i==2)	4	1	-1
arg + M	(i==3)	-1	3	0
'one'	(i==4)	-1	2	-1
'ghost meaning'	(i==5)	-1	-1	500

*generation 1*

		BN (o==1)	Numeral (o==2)	'Ghost form' (o==3)
pred	(i==1)	119	1	-1
arg	(i==2)	2	3	-1
arg + M	(i==3)	-1	2	1
'one'	(i==4)	-1	2	-1
'ghost meaning'	(i==5)	-1	-1	500

*generation 2*

		BN (o==1)	Numeral (o==2)	'Ghost form' (o==3)
pred	(i==1)	120	0	-1
arg	(i==2)	3	2	-1
arg + M	(i==3)	-1	3	0
'one'	(i==4)	-1	2	-1
'ghost meaning'	(i==5)	-1	-1	500

*generation 3*

		BN (o==1)	Numeral (o==2)	'Ghost form' (o==3)
pred	(i==1)	120	0	-1
arg	(i==2)	0	5	-1
arg + M	(i==3)	-1	3	0
'one'	(i==4)	-1	2	-1
'ghost meaning'	(i==5)	-1	-1	500

*generation 4*

*Path 2*

As is the case with *unos* the grammaticalization of articles does not always go up to the stage at which the article supersedes the corresponding bare nominal (see also Wright & Givón 1987 on the numeral *yi* ‘one’ in Mandarin Chinese). What seems to happen in those cases is that arguments are not more marked than predicates except if they are also pragmatically marked. To formalize this we can merge the constraint penalizing <BN, Pred> (i.e. (o==1&i==1)) and the one penalizing <BN, Arg> (i.e. (o==1&i==2)). This merger will ensure that the numeral will never supersede the bare noun for unmarked arguments.

The set of constraints now looks as in (127). The same remarks apply as for the set of constraints for Grammaticalization Path 1:

(127)

*new constraints for Grammaticalization Path 2*

1:	(o==1&i==1)	8:	(o==2&i==4)
1:	(o==1&i==2)	9:	(o==2&i==5)
2:	(o==1&i==3)	10:	(o==3&i==1)
3:	(o==1&i==4)	11:	(o==3&i==2)
4:	(o==1&i==5)	12:	(o==3&i==3)
5:	(o==2&i==1)	13:	(o==3&i==4)
6:	(o==2&i==2)	14:	(o==3&i==5)
7:	(o==2&i==3)		

Output corpora showing the predicted evolution are given in (128) (generations 25, 50, 75 and 100 of a random simulation)

(128)

*Random output corpora for Grammaticalization Path 2*

		BN (o==1)	Numeral (o==2)	‘Ghost form’ (o==3)
pred	(i==1)	120	0	-1
arg	(i==2)	5	0	-1
arg + M	(i==3)	-1	2	1
‘one’	(i==4)	-1	2	-1
‘ghost meaning’	(i==5)	-1	-1	500

*generation 25*

		BN (o==1)	Numeral (o==2)	'Ghost form' (o==3)
pred	(i==1)	120	0	-1
arg	(i==2)	5	0	-1
arg + M	(i==3)	-1	3	0
'one'	(i==4)	-1	2	-1
'ghost meaning'	(i==5)	-1	-1	500

*generation 50*

		BN (o==1)	Numeral (o==2)	'Ghost form' (o==3)
pred	(i==1)	120	0	-1
arg	(i==2)	5	0	-1
arg + M	(i==3)	-1	3	0
'one'	(i==4)	-1	2	-1
'ghost meaning'	(i==5)	-1	-1	500

*generation 75*

		BN (o==1)	Numeral (o==2)	'Ghost form' (o==3)
pred	(i==1)	120	0	-1
arg	(i==2)	5	0	-1
arg + M	(i==3)	-1	3	3
'one'	(i==4)	-1	2	-1
'ghost meaning'	(i==5)	-1	-1	500

*generation 100*

#### 4.3.4. Remaining challenges

In 4.3.3. I showed how we can simulate two grammaticalization paths for the numeral *one*. Even though the simulation is important it cannot be the whole story though. Indeed, next to the fact that I have no account for phenomena that are known to be related to historic evolution (especially semantic bleaching) I have failed to give an account that predicts that the evolution from the numeral *one* to articles often spans more than one century. This can be due to a number of factors ranging from the specific parameters I set in EvolOT (next to the input corpus and the constraints) to sociolinguistic events that might slow down evolution. Another plausible reason why the evolution from the numeral *one* to articles might take longer than I predict is that the division of the evolution into three stages is too coarse grained. Assuming that there are multiple smaller evolutions than the simple '*one*' -> 'Arg + M' would immediately give rise to a longer evolution time. If this turns out to be the case – as e.g. claimed by Stark (2002) – the general mechanism of constraint merger will allow us to model this more fine-grained evolution.

#### 4.4. Summary

The goal of this section was double: checking whether the grammaticalization of articles can be simulated and if so, which assumptions we would be needing to do this. The starting point was the standard architecture of EvolOT and the fact that some articles go through three stages (I, II and III) of grammaticalization whereas others only go through two (I and II):

(129)

Stage I:

All indefinites are unmarked;

Stage II:

The new marker codes only pragmatically important indefinites;

Stage III:

The new marker gradually expands its scope to mark all indefinites.

adapted from Givón 1981

Next to the standard architecture of EvolOT and some representative input frequencies I used bias constraints, ghost forms / meanings and constraint merger. I will briefly go over each of them. Bias constraints are the least involved constraints that in general should give the best correspondence between an input and an output corpus. They can be taken to be the default choice of constraints. The use of ghost forms and meanings is somewhat more involved even though one cannot avoid using them. Indeed, a limitation of EvolOT is that an output corpus can never contain occurrences for a meaning that didn't have any occurrence in the input corpus. This means that if we want a form to be able to acquire a new meaning in an output corpus we have to make sure that there are already occurrences for that meaning in the input corpus. Ghost forms can take care of this: they are hypothetical forms that convey the meaning we want another form to be able to acquire. Intuitively they can be thought of as expressing the need of speakers to convey a certain meaning. Ghost meanings are there simply to make sure that ghost forms will not continue to mark the meanings we want other forms to be able to acquire. The last specific mechanism I used was that of constraint merger. This allowed me to simulate grammaticalization paths going from Stage I to Stage II or all the way from Stage I over Stage II to Stage III. What they formalize is the intuition that certain meanings should be expressed by the same form. For an article that goes all the way to Stage III the meanings that should be expressed by the same form are *arg+M* and *arg*. To obtain this I merged the constraints penalizing the form-meanings pairs involving the article and these meanings. For an article that only goes up to Stage II the meanings that should be expressed by the same form are *pred* and *arg*. To obtain this I merged the constraints penalizing the form-meaning pairs involving the bare nominal and these meanings.

Despite the remaining challenges raised in 4.3.4. the proposed simulations go a long way in giving a faithful view on what is going on in the grammaticalization of articles; they account for a stepwise evolution with different cut-off points thus distinguishing between for articles that go from Stage I to Stage II and those that go all the way to Stage III.

## **5. Conclusion**

In this chapter I have argued on the basis of diachronic facts that *unos* should be analyzed as the plural indefinite article in Spanish. More specifically I showed that two crucial predictions an article analysis makes are borne out: (i) like the indefinite singular article it underwent an important increase in frequency and (ii) it lost its partitive potential. Together with the synchronic evidence in chapter 2 this makes a convincing case in favour of the article analysis of *unos*. This chapter furthermore contains an attempt at simulating the historical evolution of articles. I succeeded in simulating it as a stepwise process with different cut-off points which seems to correspond to what we find in Spanish.



### ***General Conclusion***

In this part of my dissertation I probed the semantics and pragmatics of indefinite articles on the basis of a case-study of Spanish *unos* ( $\approx$  'some<sub>pl</sub>'). I compiled a list of synchronic and diachronic properties an indefinite article should have and argued that *unos* has all these properties.



## **FRENCH *DES* AND ARTICLES: TWO WORLDS APART?**

The semantics and pragmatics of bare partitives

### ***General Introduction***

In Part I of this dissertation I studied Spanish *unos* and the relation between indefinite articles and other indefinite determiners. The goal of the present part is to study the relation between indefinite articles and bare partitives on the basis of French *des N*.

An intuitive definition of bare partitives is that they are partitives lacking an upstairs determiner. (i) and (ii) show that French *des N* seems to qualify:

- (i) Jean a vu deux des filles.  
Jean has seen two of\_the girls  
Jean saw two of the girls.
- (ii) Jean a vu des filles.  
Jean has seen of\_the girls  
Jean saw girls.

The reason why bare partitives are interesting to study in relation to indefinite articles is that – despite the definite determiner they contain – they behave like indefinites with no other semantic content than the one contributed by the noun. This is clear from the translation of (ii).

The discussion in this part is organized along two dimensions. On the one hand, I present a fully worked-out semantics/pragmatics of bare partitives and oppose them to DPs headed by indefinite articles. On the other hand, I argue that, despite the fact that *des N* does not behave like a full-fledged bare partitive, most of its properties can be explained on the basis of a bare partitive analysis. Chapters 1 through 3 are concerned with a synchronic analysis of bare partitives and *des N* and Chapter 4 complements the picture with a diachronic analysis of the latter.



## Chapter 1: The semantics of *des N*: putting the pieces together

In this chapter I will argue that a semantically worked out bare partitive analysis of French *des N* has the merit of bringing together insights from various lines of analysis.<sup>1</sup>

### 1. *Des*: a puzzling item

This introduction is meant to briefly present the three main lines of thinking about *des*, each with its own motivation (1.1.-1.3.). At the end I will situate the analysis I will propose with respect to these three lines (1.4.).

#### 1.1. The article line

French does not allow for bare arguments and in particular for bare plural arguments.<sup>2</sup> This is illustrated in (1):

- (1) \*Marie a mangé biscuits.  
Marie has eaten biscuits  
Marie ate biscuits

This is an intriguing property given that in Italian and Spanish – its closest relatives – the counterparts of (1) are fine:

- (2) María comió bizcochos.  
María ate biscuits  
(3) Maria mangiò biscotti.  
Maria ate biscuits

In the cross-linguistic literature this property of French is often explained on the basis of two assumptions. The first is that articles block bare nominals from acquiring argument status. The second is that French has a plural indefinite article: *des*. When combined these assumptions derive the ungrammaticality of (1) on the basis of the availability of (4):

- (4) <sup>ok</sup>Marie a mangé des biscuits.  
Marie has eaten DES biscuits  
Marie ate biscuits.

It is because of the second assumption that I refer to this line of thinking as the *article line*.

---

<sup>1</sup> A brief discussion of the syntactic analysis I assume can be found in Appendix 2 of this dissertation.

<sup>2</sup> I ignore the coordinated cases like *livres et revues* ('books and magazines') (see Roodenburg 2004).

### 1.2. The anti-determiner line

Let us for the moment assume that *des* is an indefinite article or at least an indefinite determiner. Once we grant this we are confronted with at least one property that sets *des* apart from all other French indefinite determiners: its preference to scope below any other operator. This becomes clear in sentences like (5) where the strongly preferred reading is one in which *des chemises* scopes below the modal:

- (5) Il me faut des chemises.  
 There me must DES shirts  
 I need shirts.

In the literature it has not gone unnoticed that this property makes *des N* look more like a bare plural than like an indefinite DP. This observation has led semanticists and syntacticians alike to look for analyses that play down the determiner status of *des* in order to make *des N* resemble bare plurals.

This line is called the *anti-determiner line* to indicate that it favours analyses that take *des* **not** to be a determiner.

### 1.3. The bare partitive line

*Des* is known to be morphologically complex and to consist of the preposition *de* ('of') and the definite plural article *les*. Given that the preposition *de* is the same as the one used in partitive constructions it has been proposed that *des N* is a 'reduced' partitive. The thing that distinguishes 'reduced' from 'regular' partitive constructions is that the former lack a determiner in front of *de*. Compare in this respect the regular partitive in (6) to (7):

- (6) deux des femmes  
 two of\_the women  
 (7) des femmes  
 DES women

This line will be referred to as the *bare partitive line* to indicate that *des N* is analyzed as a partitive lacking a determiner in front of *de*.

### 1.4. Outlook

I presented three lines of thinking that can be found in the literature on *des*:

- i. The article line (see e.g. Chierchia 1998, de Swart & Zwarts 2008)
- ii. The anti-determiner line (see e.g. Dobrovie-Sorin & Laca 2003, Le Bruyn 2005)
- iii. The bare partitive line (see e.g. Kupferman 2001, Zribi-Hertz 2003, Roodenburg 2004)

The goal of this chapter is not to argue in favour of one line and against the others but to reconcile them and have each of them contribute a piece of the puzzle. This goal will be achieved in three steps:

### Step 1

Following the bare partitive line I will work out an explicit semantics for *des N* as if it were a partitive construction lacking a determiner in front of *de*.

### Step 2

I will then show that the analysis allows us to derive narrow scope for *des N* in a straightforward way.

### Step 3

As a final step I will show how we can use the proposed analysis to account for the main intuition behind the article line, *viz.* that French does not have bare plurals because of their competition with *des N*.

## 2. A bare partitive analysis for *des N*

In this section I will propose a semantic bare partitive analysis for *des N*, i.e. an analysis that takes *des N* to be a partitive construction lacking a determiner in front of *de* ('of'). I will do this step by step starting with the DP following *de*, henceforth called *downstairs DP*.

### 2.1. The semantics of the downstairs DP

#### *Two kinds of definite DPs*

*Des N* contains a definite downstairs DP. In French this type of DP can have two interpretations. I illustrate with (8.a-c):

- (8) a. les arbres  
the trees
- b. D'après ma voisine les arbustes se portent bien mais il faudra abattre les arbres.  
According to my neighbour the bushes themselves carry good but it must cut\_down the trees.  
According to my neighbour the bushes are doing fine but we'll have to cut down the trees.
- c. Les arbres jouent un rôle majeur dans le fonctionnement écologique terrestre.  
the trees play a role important in the functioning ecological of\_the\_earth  
Trees play an important for the earth's ecological system.

(8.a) can refer to the sum of all the trees in a (contextually restricted) model or to the kind *trees*. Put differently: it can have both a specific (see 8.b) and a kind interpretation (see 8.c). The question to be answered is which of the two can be used in *des N*. The strategy I will follow is simple: I will show that the definite downstairs

DP cannot have a specific interpretation and by elimination conclude that it only has a kind interpretation.

The simplest way to establish that the downstairs DP of *des N* cannot have a specific interpretation is to set up a context in which it could pick up a previously introduced discourse referent. This is done in (9.a):

- (9) a. Hier j'ai acheté cinq brouillons de Picasso et deux peintures de Modiano. J'ai mis **des brouillons** dans le living.  
 Yesterday I have bought three sketches by Picasso and two paintings by Modiano. I have put DES sketches in the livingroom.  
 Yesterday I bought three sketches by Picasso and two paintings by Modiano. I put sketches in the livingroom.
- b. Hier j'ai acheté cinq brouillons de Picasso et deux peintures de Modiano. J'ai mis **les brouillons** dans le living.  
 Yesterday I have bought three sketches by Picasso and two paintings by Modiano. I have put the sketches in the livingroom.  
 Yesterday I bought three sketches by Picasso and two paintings by Modiano. I put the sketches in the livingroom.

If the definite DP contained in *des N* had a specific interpretation we would expect it in (9.a) to have as a preferred referent some of the sketches by Picasso the I-person bought yesterday. This is not the case though: (9.a) contrasts with (9.b) in not allowing *les brouillons* included in *des brouillons* to refer back to the sketches introduced in the first sentence. I conclude that the downstairs DP of *des N* cannot have a specific interpretation and by elimination obtain that it has to have a kind interpretation.<sup>3</sup>

It is important to underline that establishing that the downstairs DP of *des N* is kind-referring does not mean I don't have to explain why the specific interpretation is not available. Given that it does not have any influence on the rest of the analysis I will however postpone this to Chapters 2 and 3. In the remainder of this subsection I will present a semantics for kind-referring DPs.

---

<sup>3</sup> Even though in general the downstairs DP in *des N* cannot have a specific interpretation there is one noteworthy exception: *des N* appearing in the object position of verbs like *to eat*.

(i) Hier j'ai acheté cinq pommes et trois poires. Aujourd'hui j'ai mangé des pommes.

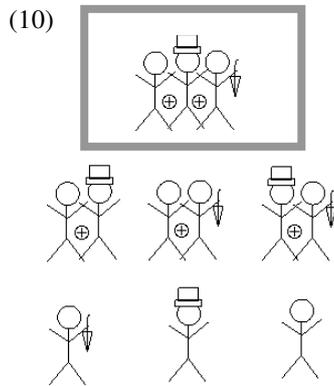
Yesterday I have bought five apples and three pears. Today I have eaten DES apples.

Yesterday I bought five apples and three pears. Today I ate apples / some of the apples.

In (i) *des pommes* can pick-up part of the five apples referred to before. The reading where it simply refers to apples is also available though. In the literature it is still under debate whether both readings should receive a bare partitive analysis (see Kupferman 2001 vs. Zribi-Hertz 2003). For the moment I will put aside these cases but I will come back to them (see Chapter 2).

*Semantics of plural kind-referring definite DPs*

A plural kind-referring definite DP is generally analyzed in much the same way as a specific plural definite DP. Both are assumed to refer to the top node of a lattice (see (10)). The lattice has to satisfy the following two conditions: (i) it has at least two atomic individuals and (ii) it contains all atomic individuals satisfying the descriptive content of the NP included in the DP and only those.



The only difference between kind-referring and specific plural definite DPs is that the lattice of the former includes atomic individuals from all possible worlds whereas the atomic individuals of the latter are restricted to the actual world.<sup>4</sup> The way I will formally distinguish between kind-referring and specific plural definite DPs is to use the  $\Sigma$  operator for the former and the *iota* operator for the latter:

- (11)  $\llbracket \text{es} \rrbracket_{\text{kind}}$   $\Sigma$   
 $\llbracket \text{arbres} \rrbracket$   $\lambda x(\text{plural}(x) \& \text{tree}(x))$   
 $\llbracket \text{es arbres} \rrbracket_{\text{kind}}$   $\Sigma \lambda x(\text{plural}(x) \& \text{tree}(x))$  (“the sum of all trees across all possible worlds”)
- (12)  $\llbracket \text{es} \rrbracket_{\text{pec}}$  *iota*  
 $\llbracket \text{arbres} \rrbracket$   $\lambda x(\text{plural}(x) \& \text{tree}(x))$   
 $\llbracket \text{es arbres} \rrbracket_{\text{spec}}$  *iota*  $\lambda x(\text{plural}(x) \& \text{tree}(x))$  (‘the sum of all trees in the actual world’)

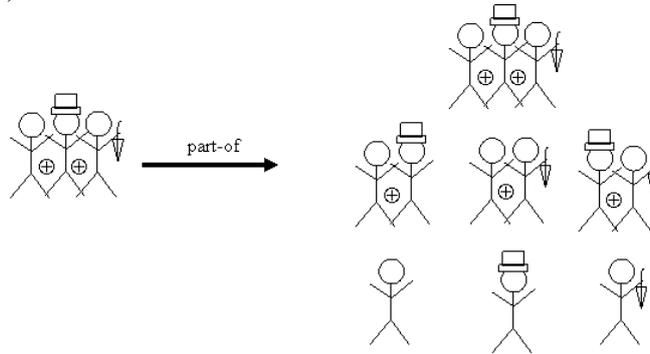
Now that I have a semantics for the downstairs DP of *des N* I can look into the semantic contribution of partitive *de* ‘of’. This is the topic of the next subsection.

<sup>4</sup> Note that this difference implies that in the world of reference specific plural definite DPs are presuppositional but kind-referring definite DPs are not. It might indeed be the case that there are no trees in the actual world, as long as there are at least two trees in some other worlds this would however not mean that the kind *trees* is undefined.

## 2.2. The semantics of partitive *de*

I follow the classical analysis of partitive *of* that is cast in terms of the part-of operator (see Keenan & Stavi 1986, Barwise & Cooper 1981) and assume it carries over to French *de*.<sup>5</sup> The part-of operator takes a sum individual and returns the set of all individuals making up the lattice of which that sum individual is the top node.

(13)



Formally:

$$(14) \quad \llbracket de \rrbracket = \lambda y \lambda x (\leq(x, y))$$

Even though this analysis was not designed to deal with kinds, the application to kind-referring plural DPs is straightforward. Combining the semantics we had for the downstairs DP of *des N* with that of *de* we obtain the following:

$$(15) \quad \begin{array}{ll} \llbracket \text{les arbres} \rrbracket_{\text{kind}} & \Sigma \text{tree} \\ \llbracket de \rrbracket & \lambda y \lambda x (\leq(x, y)) \\ \llbracket \text{des arbres} \rrbracket & \lambda x (\leq(x, \Sigma \text{tree})) \end{array}$$

In words: *des arbres* is the set of all (atomic and non-atomic) individuals that are part of the sum individual comprising all the trees in all possible worlds.

## 2.3. And now?

The attentive reader will have noticed that there is no part of *des N* I haven't given a semantics for and that the semantics I have now is that of an  $\langle e, t \rangle$  expression that is number neutral. This cannot be the whole story though. Indeed, even though *des N* can appear in predicate position it is in general incompatible with singular subjects (see (16)). It can furthermore also appear in argument position where it also has a plural interpretation (see (17)).

<sup>5</sup> See Appendix 1 for a worked-out motivation.

- (16) Marie et Jeanne sont des femmes.  
 Marie and Jeanne are DES women  
 \*Jean est des hommes.  
 John is DES men
- (17) J'ai vu des peintures. \*Elle était remarquable / Elles étaient remarquables.  
 I have seen DES paintings It was / They were remarkable  
 I saw paintings. It was remarkable / They were remarkable.

Should I conclude that the bare partitive analysis is flawed? I don't think so. In 2.4. I will show that the plurality of *des N* follows from more general considerations about number marking and in 2.5. I will show how standard type-shifting gets us from the  $\langle e, t \rangle$  type to an argumental type.

#### 2.4. The plurality of *des N*

In this subsection I will show that the plurality of *des N* follows from default reasoning. The analysis I will propose is inspired by number marking of nouns in e.g. (written) French.

##### 2.4.1. Number marking on nouns

In (written) French plural number is morphologically marked on nouns. This is illustrated in (18):

- (18) deux femmes  
 two woman\_pl

Singular number doesn't have a similar morphological marker. This is illustrated in (19):

- (19) une femme $\emptyset$   
 one woman

It appears then that the category *number* for French nouns has a marked and an unmarked value and that the unmarked value comes out by default reasoning of the following sort: if the speaker had wanted to mark that *femme* in (19) is plural (s)he would have done so; given that (s)he didn't *femme* is singular. See Farkas & de Swart (2009) for a worked out analysis.

##### 2.4.2. The gist of the proposal

In what follows I will argue that the kind of default reasoning used for the number of nouns can also be used for the number of partitives. The only difference is that partitives have a singular but no plural number marker. It's the plural then that comes out as the default. This is schematically represented as follows:

	singular number marker	plural number marker
nouns	∅ (default)	✓
partitives	✓	∅ (default)

Table 1

For this proposal to work I have to: (i) determine where number marking of partitives takes place, (ii) show that there is no plural marker, (iii) show that there is a singular marker. These are the steps I take from 2.4.3. to 2.4.6.

Note that the approach sketched here is strongly reminiscent of the one proposed by Sauerland & al. (2005) who use a variant of default reasoning – maximize presupposition (see Heim 1991) – to derive the plural reading of morphologically marked plural nouns.

### 2.4.3. Number in full partitives

On the basis of (20) and (21) I claim that the overall number of full partitives depends on the number of the determiner preceding *de*, henceforth *upstairs determiner*.

- (20) J'ai vu **un de ces hommes**. Il était amiable / \*Ils étaient aimables.  
I saw one of these men. He was amiable / They were amiable
- (21) J'ai vu **deux de ces hommes**. \*Il était amiable / Ils étaient aimables.  
I saw two of these men. He was amiable / They were amiable

What (20) shows is that using a singular upstairs determiner like the numeral *un* ('one') results in a singular semantics for the partitive. What (21) shows is that using a plural upstairs determiner like the numeral *deux* ('two') results in a plural semantics. I take (20) and (21) to be sufficient evidence in favour of the claim that the overall number of a full partitive depends on the number of its upstairs determiner.

### 2.4.4. Plurality in full partitives

Suppose we want to mark that the referent of a full partitive is plural and nothing more. Interestingly, there is no determiner that we can use to do just that. Indeed, all plural determiners that can appear as upstairs determiners add semantic content beyond plurality. In (22) and (23) I demonstrate this for the most probable candidates: *plusieurs* ('several'), *certain* ('certain'), *différents* and *divers* ('diverse') (based on Bosveld-de Smet 1998, Jayez 2005, Corblin 2001, Laca & Tasmowski 2004, Comorovski 2004).<sup>6</sup>

<sup>6</sup> Note that *quelques* ('some') cannot appear as an upstairs determiner:  
(i) \*quelques de ces hommes 'some of these men'

- (22) ??Plusieurs gorilles de ce pays meurent par dizaines.  
 Several gorillas die by tens  
 “Several gorillas die by the tens.”
- (23) ??J’ai parcouru certains / divers / différents kilomètres.  
 I have travelled certain / diverse / different kilometers.

In (22) I show that *plusieurs N* objects are incompatible with quantitative modifiers. This suggests that *plusieurs* has a quantity component of its own that clashes with that of the modifier.

In (23) I show that *certains*, *différents* and *divers* are not compatible with nouns that refer to objects that shouldn’t be distinguishable one from the other (measure nouns like *kilometer*, *kilo*, *liter*, ...). The fact that *certains*, *différents* and *divers* are not compatible with these nouns indicates that these determiners impose an extra qualitative distinction on the individuals making up the referent of the DP they head. From the preceding I conclude that there is no clear way to mark plurality and only plurality in full partitives.

#### 2.4.5. Singularity in full partitives

Suppose now we want to mark that the referent of a full partitive is singular and nothing more. Unlike in the plural case there is a determiner that allows us to do exactly that: the numeral *un* (‘one’). Indeed, semantically the numeral *un* does not add anything beyond singular number. This is illustrated above in (20).

One possible question that could be raised is whether the indefinite singular article wouldn’t be a better candidate. The question is especially relevant for French where we cannot distinguish between the numeral and the article. The answer is negative though; languages in which the indefinite singular article can be distinguished from the numeral *one* (e.g. Dutch, English) show us that indefinite singular articles simply cannot appear as upstairs determiners in partitives (see (24)). By generalization I assume the same holds for a language like French.

- (24) \*een van de mannen  
 a of the men

From the preceding I conclude that the numeral *un* (‘one’) is a determiner that allows one to mark singularity and only singularity in full partitives.

#### 2.4.6. Number of bare partitives

In 2.4.3. to 2.4.5. I have argued that: (i) the number of partitives is marked by the upstairs determiner, (ii) there is no straightforward way to mark plurality and only plurality, (iii) there is a straightforward way to mark singularity and only singularity. If we take this one step further there are two important conclusions we can draw. The first is that partitives have a singular but no plural marker. The second is that in absence of the singular number marker partitives receive a plural interpretation by default:

	singular number marker	plural number marker
partitives	<i>un</i> ('one')	∅ (default)

Table 2

From the preceding it follows that bare partitives like *des N* – given the absence of overt singular marking – have a plural interpretation by default. This is exactly the kind of interpretation I wanted to arrive at.<sup>7</sup>

### 2.4.7. Conclusion

In this subsection I showed how default reasoning imposes a plural interpretation for bare partitives. This means that the interpretation we arrived at for *des arbres* ('DES trees') (see (25.a-c)) can be further completed (see (25.d)):

- (25) a.  $\llbracket \text{les arbres} \rrbracket_{\text{kind}} \quad \Sigma \text{tree}$   
 b.  $\llbracket \text{de} \rrbracket \quad \lambda y \lambda x (\leq(x,y))$   
 c.  $\llbracket \text{des arbres} \rrbracket \quad \lambda x (\leq(x, \Sigma \text{tree}))$   
 d.  $\llbracket \text{des arbres} \rrbracket + \text{default reasoning} \quad \lambda x (\text{plural}(x) \& \leq(x, \Sigma \text{tree}))$

In words: *des arbres* is the set of all non-atomic individuals that are part of the sum individual comprising all the trees in all possible worlds. For the use of *des N* in predicate position this gives us exactly the result we would want:<sup>8</sup>

- (26) Marie et Jeanne sont des femmes.  
 Marie and Jeanne are DES women

For *des N* in argument position, we still need a way to get from the  $\langle e, t \rangle$  type we have now to an argumental type. How exactly this can be done is the topic of 2.5.

<sup>7</sup> The attentive reader might object at this point and say that for this story to work for *des N* in French *un des hommes* ('one of the men') should be interpretable as referring to an instantiation of the kind of men and not only to a man who is part of a specific set of men (thanks to Brenda Laca for pointing this out to me). The objection is legitimate but – as will become clear in chapter 4 – I assume the plural interpretation of *des N* got fossilized in the 13<sup>th</sup> and 14<sup>th</sup> century where there was a slightly different kind of competition that led *des N* to a plural interpretation. In these centuries sentences like the following were fine (data drawn from the Middle French part of the Frantext corpus):

- (i) mon frere, qui est **un des preudoms** et un des plus vaillans roys du monde  
 my brother, who is one of\_the brave\_men and one of\_the most courageous kings of the world  
 my brother, who is a brave man and one of the most courageous kings of the world

The reason I keep the plurality part in this chapter is that it nicely carries over to other bare partitives that will be treated in chapter 2 and that don't present the problem of not having an attested singular counterpart of the type *one + bare partitive*.

<sup>8</sup> Given the pragmatic nature of the plurality of bare partitives I predict there might be some exceptions. This prediction is borne out as the following example shows:

- (ii) elle était **des femmes** heureuses de se laisser violenter (Zola, Au bonheur des dames)  
 she was of\_the women happy to themselves let screwed

## 2.5. Bare partitives and type-shifting

In this subsection I will show how the argumental interpretation of *des N* can be derived within standard type shifting theory. In order to do so I first briefly sketch the part of type shifting theory we need.

### *Background on type-shifting*

In the literature on type-shifting (Partee 1987, Chierchia 1998) it is assumed that there are three basic types ( $e$ ,  $\langle e, t \rangle$  and  $\langle \langle e, t \rangle, t \rangle$ ) and that there is a number of ‘natural’ shifts between types that may apply freely as long as there is no overt determiner denoting these shifts. In what follows I would like to take a closer look at one of these ‘natural’ shifts, *viz.* the existential shift:

$$(27) \quad P \rightarrow \lambda Q \exists x [P(x) \& Q(x)]$$

What this shift does is to take a set  $P$  and return the set of sets that have a non-empty intersection with  $P$ .

In a language like English there is a determiner that denotes this shift for the singular, *viz.* the indefinite singular article:

$$(28) \quad \llbracket a \rrbracket \quad \lambda P \lambda Q \exists x [P(x) \& Q(x)]$$

For the plural however English does not have a determiner with a semantics restricted to what is given in (28). The general assumptions underlying type-shifting and blocking then lead us to the prediction that there can be an existential use of the bare plural but not of the bare singular. This prediction is borne out:

$$(29) \quad * \text{There is book on the table.}$$

$$(30) \quad \text{ok} \text{There are books on the table.}$$

### *Type-shifting and bare partitives*

The semantics we ended up with in 2.4. for *des arbres* is as follows:

$$(31) \quad \begin{array}{ll} \text{a. } \llbracket \text{les arbres} \rrbracket_{\text{kind}} & \Sigma_{\text{tree}} \\ \text{b. } \llbracket \text{de} \rrbracket & \lambda y \lambda x (\leq(x, y)) \\ \text{c. } \llbracket \text{des arbres} \rrbracket & \lambda x (\leq(x, \Sigma_{\text{tree}})) \\ \text{d. } \llbracket \text{des arbres} \rrbracket + \text{default reasoning} & \lambda x (\text{plural}(x) \& \leq(x, \Sigma_{\text{tree}})) \end{array}$$

Under the assumption that there is no overt plural determiner whose denotation is restricted to what is given in (28) (see 2.4.4.) we expect the existential shift to be available covertly for (31.d). This leads to the semantics given in (31.e):

$$\text{e. } \llbracket \text{des arbres} \rrbracket + \text{type-shifting} \quad \lambda Q \exists x [\text{plural}(x) \& \leq(x, \Sigma_{\text{tree}}) \& Q(x)]$$

In words: the set of sets that have a non-empty intersection with the set of all non-atomic individuals that are part of the sum individual comprising all the trees in all possible worlds.

## 2.6. Summary and orientation

### Summary

In this section I presented a bare partitive analysis of *des N* tracing back its semantics to the sum of its parts (32.a-c), default number reasoning (32.d) and type-shifting (32.e):

(32)	a. $\llbracket \text{les arbres} \rrbracket_{\text{kind}}$	$\Sigma \text{tree}$
	b. $\llbracket \text{de} \rrbracket$	$\lambda y \lambda x (\leq(x,y))$
	c. $\llbracket \text{des arbres} \rrbracket$	$\lambda x (\leq(x, \Sigma \text{tree}))$
	d. $\llbracket \text{des arbres} \rrbracket$ + default reasoning	$\lambda x (\text{plural}(x) \& \leq(x, \Sigma \text{tree}))$
	e. $\llbracket \text{des arbres} \rrbracket$ + type-shifting	$\lambda Q \exists x [\text{plural}(x) \& \leq(x, \Sigma \text{tree})] \& Q(x)$

This semantics gives us the right truth conditions for (33) and (34):

- (33) Marie et Jeanne sont des femmes.  
Marie and Jeanne are DES women
- (34) J'ai vu des peintures.  
I saw DES paintings

(33) is true just in case Marie and Jeanne belong to the set of women. This is the interpretation we predict on the basis of (32.d). (34) is true just in case I saw more than one painting. This interpretation is the one we expect on the basis of (32.e). Note that the fact that the members of the set in (32.c) are in principle not restricted to the world of evaluation can be overcome by a straightforward contextual restriction.

### Recap

The goal of this chapter is to bring together three lines of thinking about *des N*:

- (i) the article line: *des* is a plural indefinite article and blocks the bare plural from appearing in French
- (ii) the anti-determiner line: *des* is not an indefinite determiner because it has a preference for narrow scope
- (iii) the bare partitive line: *des N* is a partitive lacking an upstairs determiner

In this section I have worked out one line in all its (semantic / pragmatic) detail, *viz.* the bare partitive one. In the next two sections I will show how the proposed

analysis can be reconciled with the anti-determiner line (section 3) and the article line (section 4).

### 3. The preference for narrow scope of *des N* and the anti-determiner line

In section 1 I pointed out that *des N* is notorious for its preference for narrow scope readings. I illustrate on the basis of (35.a-b):

- (35) a. Il me faut des chemises.  
           there me must DES shirts  
           “I need shirts (any shirts)” (*des chemises* scopes below *faut*)
- b. Il me faut une chemise.  
           there me must a shirt  
           “‘There is a shirt that is such that I need it.’” (*une chemise* scopes above *faut*)  
           “I need a shirt (any shirt)” (*une chemise* scopes below *faut*)

The strongly preferred reading of (35.a) is the one in which *des chemises* scopes below the modal. This contrasts with (35.b) that has both an interpretation in which *une chemise* scopes above the modal and one in which it scopes below.

In part of the literature on *des N* its narrow scope behaviour has been used to argue that *des N* behaves like a bare plural rather than like an indefinite DP and – consequently – that *des* is not a determiner. I dubbed this line of thinking about *des N* the anti-determiner line (see section 1). Interestingly, the bare partitive analysis I presented is perfectly compatible with the anti-determiner line. The reason for this is that I assume *des N* is similar to bare plurals in having to undergo a covert type-shift to appear in argument position. It’s the necessity of this covert type-shift that is generally assumed to be responsible for the narrow-scope behaviour of bare plurals. In the remainder of this section I will show how it works for *des N*.

The analysis of the narrow-scope behaviour of bare plurals I will extend to *des N* is the one proposed by Krifka (2004). The crucial assumption that derives narrow scope in his analysis is given in (36) (see also Chierchia 1998 and Dayal 2004):

- (36) type shifting occurs locally, i.e. when the mismatch between the NP and the verbal predicate becomes apparent

I will work out the analysis for the sentence in (35) but before doing so three remarks are in order. The first is that – for the sake of readability – I will use (38) as a short form for (37).

(37)  $\lambda x(\text{plural}(x) \& \leq(x, \Sigma \text{shirt}))$

(38)  $\lambda x(\text{shirts}(x))$

The second one is that I will use (39) as the semantics of *il l faut k* (‘I needs k’):

(39)  $\Box$  to\_have(k)(l)

The third and final remark is that – like Krifka – I will use type-neutral variables written  $\alpha, \beta, \gamma$ .

In (40.a-i) I give the semantic derivation up to the crucial step:

(40)  $\llbracket$ il l faut k $\rrbracket$   $\Box$  to\_have(k)(l) (a)  
 $\llbracket$ me $\rrbracket$  m (b)  
 $\llbracket$ il me faut k $\rrbracket$   
*lambda abstraction*  
 $\lambda\alpha(\Box$  to\_have (k)( $\alpha$ )) (c)  
*function application*  
 $\lambda\alpha(\Box$  to\_have (k)( $\alpha$ )) (m) (d)  
*lambda conversion*  
 $\Box$  to\_have (k)(m) (e)  
 $\llbracket$ des chemises $\rrbracket$   
 $\lambda x(\text{shirts}(x))$  (f)  
 $\llbracket$ il me faut des chemises $\rrbracket$   
*lambda abstraction*  
 $\lambda\alpha(\Box$  to\_have ( $\alpha$ )(m)) (g)  
*function application*  
 $\lambda\alpha(\Box$  to\_have( $\alpha$ )(m))( $\lambda x(\text{shirts}(x))$ ) (h)  
*lambda conversion*  
 $\Box$  to\_have ( $\lambda x(\text{shirts}(x))$ )(m) (i)

In (40.i) there is a type-clash that will have to be solved. One option is to type-shift to\_have, the other to type-shift  $\lambda x(\text{shirts}(x))$ . I choose the latter but the former is possible as well (see Van Geenhoven 1998):

*type shift*<sup>9</sup> + *function application*  
 $\Box \lambda P \lambda y \exists x(\text{chemises}(x) \& P(x)(y))(\text{to\_have})(m)$  (j)  
*lambda conversion*  
 $\Box \lambda y \exists x(\text{chemises}(x) \& \text{to\_have}(x)(y))(m)$  (k)  
*lambda conversion*  
 $\Box \exists x(\text{chemises}(x) \& \text{to\_have}(x)(m))$  (l)

The result of this derivation is that *des chemises* takes scope below the modal operator. The formal trick that allowed me to obtain this is the use of type-neutral variables; these make sure that type mismatches are detected and solved locally.

*Summary*

<sup>9</sup> The existential type-shift for combining with  $n$ -ary predicates is defined as follows:

(i)  $\lambda y P(y) \rightarrow \lambda Q \lambda x_0 \dots \lambda x_n \exists y (P(y) \& Q(y)(x_0) \dots (x_n))$

In this section I argued that the analysis I proposed for *des N* in section 2 is compatible with the anti-determiner line of thinking about *des N*. As such it reconciles the bare partitive and the anti-determiner line. Section 4 will sketch how these two lines can be reconciled with the article line.

#### 4. The absence of bare plurals in French and the article line

In section 1, I referred to French as an exceptional language in the Romance language paradigm given that it is the only one that does not allow plural nouns to occur bare in argument position. In the cross-linguistic literature, the explanation for the exceptional status of French goes back to what is commonly assumed to account for the absence of argumental bare singulars in languages that have articles: it is the availability of articles that blocks bare nominals from appearing in argument position. If we assume that *des* is the French plural indefinite article, the absence of French argumental bare plurals follows straightforwardly. I dubbed this line of thinking about *des N* the article line.

At first sight the article line is completely incompatible with the bare partitive and the anti-determiner line. I would however like to argue that this is only so to a certain extent. To see this, two things that are conflated in the article line have to be teased apart: the blocking account on the one hand and the article analysis of *des* on the other hand. The motivation for separating these two things out is very simple: there is nothing inherent to blocking that forces the blocking item to contain an article. Indeed, the only thing needed to set up a blocking account is two items that can express the same meaning.<sup>10</sup> Now that I have established this, the possibility of reconciliation becomes clear: if I can show that a blocking story can be set up without having to assume that *des* is an article I can reconcile at least one part of the article line and the two other lines. The remaining part – the article assumption – cannot be reconciled.

The complete blocking analysis I propose for *des N* and the bare plural has to be set up in diachrony. This will be worked out in detail in Chapter 4 but I will briefly sketch the main line of thinking in the remainder of this section.

In 2.4, I discussed the number oppositions in the French nominal and partitive domain. Number oppositions in the nominal domain are marked on the noun whereas in the partitive domain they are marked on the upstairs determiner. This is illustrated in (41) and (42):

- (41) livre            livres  
 (42) **un** des livres des livres

---

<sup>10</sup> Note that blocking here is seen as a general mechanism and is not restricted to the narrow definition given by Chierchia (1998).

In (41) the plural is marked by a suffixal –s and the singular comes out as the default. In (42) the singular is marked by the upstairs determiner and the plural comes out as the default.

Let us now take a look at diachrony. Around 1300, French final consonants start to undergo phonetic erosion. This has important consequences for number oppositions in the nominal domain but not in the partitive domain. Indeed, whereas phonetic erosion of final consonants affects the suffixal –s on nouns it does not directly affect the upstairs determiner in (42). As a result *livres* becomes indistinguishable in number from *livre* whereas *des livres* remains perfectly distinguishable from *un des livres*. It is in these particular circumstances that – in order to convey plurality – speakers will rather use a partitive than a nominal.

## 5. Summary and outlook

In this chapter I worked out a bare partitive analysis of *des N* and I argued that this analysis can be reconciled with two different lines of analysis:

- (i) the article line: *des* is a plural indefinite article and blocks the bare plural from appearing in French
- (ii) the anti-determiner line: *des* is not an indefinite determiner because it has a strong preference for narrow scope

A worked out blocking story of the bare plural by *des N* will be presented in Chapter 4 but there are some other questions the bare partitive analysis raises that should be addressed first. I briefly summarize them here and work them out in Chapters 2 and 3. The first question is why the downstairs DP in *des N* has to be kind-referring. In 2.1. I showed that this was a real constraint on *des N*, but I haven't yet given an explanation for it. The reason why I refrained from doing this is that I should first find out whether this constraint holds for bare partitives in general or whether it is specific to French *des N*. The second question is about scope. In section 3 I glossed over the availability of wide-scope readings of *des N* but they do exist. Whether or not this hybrid behaviour should be incorporated into the bare partitive analysis is also something that I can only decide on when I have had a look at some more bare partitives. The third and final question is about the validity of the pragmatic account I proposed to derive the plurality of *des N*. The question is whether this account derives a quirk of *des N* or whether it is as general as a pragmatic account should be and extends to other bare partitives as well.

The questions I just mentioned constitute serious challenges for the analysis I have proposed so far: the first can be restated as an overgeneration problem, the second as an undergeneration problem and the third can be seen as questioning the plausibility of the analysis. Chapters 2 and 3 will provide the answers to these challenges.

## Chapter 2: What about other bare partitives?

### 1. Recap

In Chapter 1 I argued that the puzzling facts that have driven linguists to explore different analyses for French *des* are all covered by an analysis in which *des N* is taken to be a partitive with a kind-referring DP and without an upstairs D. Let me briefly remind the reader of some core examples, the analysis and the facts it can derive.

#### *Core examples*

*Des N* can be used both in predicate and in argument position – see (43) and (44) respectively. The truth conditions it gives rise to are exactly the same as those of an English bare plural.

- (43) Marie et Jeanne sont des femmes.  
Marie and Jeanne are DES women  
Marie and Jeanne are women.
- (44) J'ai vu des peintures.  
I have seen DES paintings  
I saw paintings.

#### *The analysis*

In case *des N* is a partitive with a kind-referring DP and without an upstairs D its compositional semantics is the one in (45):

- (45)  $\llbracket \text{les arbres} \rrbracket_{\text{kind}}$        $\Sigma_{\text{tree}}$   
 $\llbracket \text{de} \rrbracket$                        $\lambda y \lambda x (\leq(x, y))$   
 $\llbracket \text{des arbres} \rrbracket$                $\lambda x (\leq(x, \Sigma_{\text{tree}}))$

I made two pragmatic additions to this semantics. The first is a plurality condition:

- (46)  $\llbracket \text{des arbres} \rrbracket$  +default reasoning     $\lambda x (\text{plural}(x) \& \leq(x, \Sigma_{\text{tree}}))$

This plurality condition follows from default reasoning if we assume that the number of partitives is marked by the upstairs determiner and that there is a determiner marking singularity – the numeral *one* – but no determiner marking plurality. In absence of a determiner plurality would then pop up as the default. The second pragmatic addition I made is the availability of the existential type-shift:

- (47)  $\llbracket \text{des arbres} \rrbracket$  +default reasoning     $\lambda Q \exists x (\text{plural}(x) \& \leq(x, \Sigma_{\text{tree}}) \& Q(x))$   
+ $\exists$  type-shift

The availability of the existential type-shift for type  $\langle e,t \rangle$  expressions in argument position is standard in type-shifting theory in as far as there is no overt determiner that encodes this shift. I argued that there is indeed no overt determiner that encodes the existential type-shift for plurals.

### *The puzzling facts*

There are three facts that are covered by the analysis of *des N* I just presented. The first is its morphological relatedness to full partitives (see Kupferman 2001, Zribi-Hertz 2003, Roodenburg 2004). The second is its narrow scope behaviour (see Dobrovie-Sorin & Laca 2003). This behaviour can be derived from the bare partitive analysis in the following way: given that *des N* and bare plurals are arguably of the same semantic type – type  $\langle e,t \rangle$  – and that both have a strong preference for narrow scope whatever mechanism derives the scope behaviour of the latter can also be assumed to apply in the former. The third fact is the absence of bare plurals in French (see Chierchia 1998, de Swart & Zwarts 2009). Even though I haven't yet given the specifics of a blocking account it should be clear that the truth-conditional equivalence of bare plurals and *des N* lends itself very well to a blocking account.

## **2. Beyond *des N***

The present chapter aims at providing a thorough description of a wider range of bare partitives. The motivation for this descriptive work comes from the fact that the analysis of *des N* I presented in Chapter 1 raises three questions that can only be addressed after a close investigation of other bare partitives. The first question is why *des N* can take wide scope even though the analysis now only covers narrow scope. This question can be answered in many ways but one thing that has to be established first is whether bare partitives in general can take wide scope or not. If it turns out that they can I have to adapt the analysis, if it turns out that they cannot the wide scope behaviour of *des N* should be explained in a different way. The second question raised by the analysis in Chapter 1 is what the status is of the assumption that the downstairs DP of *des N* is kind-referring. At the moment this assumption is nothing more than a stipulation but before I try to build this kind constraint into the analysis I need to know whether this constraint applies to other bare partitives as well. The third and final question raised by the bare partitive analysis of *des N* is about the validity of the pragmatic account I proposed to derive the plurality of *des N*. The question is whether this account derives a quirk of *des N* or whether it is as general as a pragmatic account should be. An investigation of other bare partitives will give me a straightforward answer to this question.

This chapter is organized as follows. In section 3 I make a list of bare partitives on the basis of criteria that are independent of the analysis presented in Chapter 1. In section 4 I present a thorough description of the bare partitives appearing in this list and in section 5 I address the three questions I raised above.

### 3. Looking for bare partitives

In this section I compile a list of bare partitives that will be the input for the descriptive work in section 4 and the discussion in section 5. The basis for this compilation is the set of criteria I present in 3.1. The compilation itself is presented in 3.2. Note that I restrict my search for bare partitives to languages closely related to French: English and Dutch as representatives of the Germanic side and French, Italian and Spanish as representatives of the Romance side.

#### 3.1. The criteria

In this subsection I present three criteria that have to be met in order to qualify as a bare partitive. Note that they are independent of the analysis I proposed in Chapter 1.

##### *The form criterion*

The first criterion for an expression to qualify as a bare partitive is to differ in form from a full partitive only in its lack of an upstairs determiner. In (48) I give a list of full partitives of the languages I will look at:

(48)	determiner	partitive ‘preposition’	determiner	noun	
	one	of	these	men	ENGLISH
	één	van	deze	mannen	DUTCH
	un	de	ces	hommes	FRENCH
	uno	di	questi	uomini	ITALIAN
	uno	de	estos	hombres	SPANISH

To be a bare partitive an expression then needs to have the following form:

(49) partitive ‘preposition’ + determiner + noun

An extra restriction I add is that the downstairs noun should be plural. Note that this is a restriction on the scope of the list I will compile and not on bare partitives in general.<sup>1</sup>

##### *The meaning criterion*

The second criterion for an expression to qualify as a bare partitive is that it can be of type  $\langle e, t \rangle$  and that *of* has the partitive semantics in (50):

---

<sup>1</sup> The existence of bare partitives with downstairs mass nouns is well-attested (see especially Bosveld-de Smet 1998). An explicit treatment of these is however beyond the scope of this dissertation as they raise a whole set of questions about the relation between mass and count domains.

(50)  $\lambda y \lambda x (\leq(x,y))$

The simple trick I will be using to identify partitive *of* is to check whether the individuals that are members of the set denoted by the candidate expression can be characterized by the downstairs noun. If yes, we are probably dealing with a partitive, if not, we are probably dealing with something else. To see this compare (51) and (52):

- (51) Jean et Pierre sont des hommes.  
 Jean and Pierre are DES men  
 Jean and Pierre are men.
- (52) Jean et Pierre sont de niveaux différents.  
 Jean and Pierre are of levels different  
 Jean and Pierre are of different levels.

On the basis of the most natural readings of (51) and (52) we can conclude that Jean and Pierre are men, but we cannot conclude that they are levels. It follows then that we are probably dealing with a partitive in (51) but not in (52).

*The use criterion*

The third and final criterion for an expression to qualify as a bare partitive is that it is able to occur in argument position. To this criterion I add one more restriction, *viz.* that I will not take into consideration the object position of verbs like *to eat*:

- (53) Van de koteletjes mag Marie eten, van de spareribs niet. DUTCH  
 Of the chops may Mary eat, of the spareribs not  
 Mary can have the chops but not the spareribs.

The reason for setting aside *van de koteletjes* and *van de spareribs* after verbs like *to eat* is that *de koteletjes* and *de spareribs* can be replaced by any NP, including bare plurals:

- (54) Van koteletjes mag Marie eten, van spareribs niet. DUTCH  
 Of chops may Mary eat, of spareribs not  
 Mary can eat chops but not spareribs.

Given that partitives – as shown in (55) – don't allow for bare plurals in their downstairs DP position (see Reed 1996, Abbott 1996 and de Hoop 1998) *van koteletjes* en *van spareribs* in (54) cannot be analyzed as bare partitives.

- (55) \*drie van koteletjes DUTCH  
 three of chops

I follow Kupferman (1979) in assuming that *van koteletjes* does not function as an NP/DP but rather as a prepositional complement of *to eat*. This is the reason why I set aside verbs like *to eat* for the moment. I will briefly come back to them in Chapter 3, section 3.

### 3.2. Bare partitives

The search for bare partitives in Dutch, English, French, Italian and Spanish on the basis of the criteria introduced in 3.1. yields the following list:

- |      |              |                    |        |         |
|------|--------------|--------------------|--------|---------|
| (56) | a. van       | + demonstrative    | + noun | DUTCH   |
|      | b. de ('of') | + demonstrative    | + noun | FRENCH  |
|      | c. di ('of') | + demonstrative    | + noun | ITALIAN |
|      | d. di ('of') | + definite article | + noun | ITALIAN |

In 3.2.1. I argue that the expressions in (56) meet the bare partitive criteria and in 3.2.2. I argue that Dutch does not have a bare partitive with the definite article and that English and Spanish lack bare partitives altogether. I summarize in 3.2.3.

#### 3.2.1. The bare partitives of Dutch, French and Italian

In what follows I will argue that the expressions in (56) meet the three bare partitive criteria.

##### *The form criterion*

On the basis of (58) I will argue that the expressions in (56) meet the form criterion:

- |      |                       |       |      |                             |                                 |
|------|-----------------------|-------|------|-----------------------------|---------------------------------|
| (58) | DUTCH                 |       |      |                             |                                 |
|      | <i>bare partitive</i> | van   | die  | rare ventjes                | <i>van die rare ventjes</i>     |
|      | <i>full partitive</i> | één   | van  | die rare ventjes            | <i>één van die rare ventjes</i> |
|      |                       | 'one' | 'of' | 'those' 'weird little_guys' |                                 |
|      | FRENCH                |       |      |                             |                                 |
|      | <i>bare partitive</i> | de    | ces  | pommes                      | <i>de ces pommes</i>            |
|      | <i>full partitive</i> | une   | de   | ces pommes                  | <i>une de ces pommes</i>        |
|      |                       | 'one' | 'of' | 'these' 'apples'            |                                 |
|      | ITALIAN               |       |      |                             |                                 |
|      | <i>bare partitive</i> | di    | quei | mascalzoni                  | <i>di quei mascalzoni</i>       |
|      | <i>full partitive</i> | uno   | di   | quei mascalzoni             | <i>uno di quei mascalzoni</i>   |
|      |                       | 'one' | 'of' | 'those' 'rascals'           |                                 |
|      | <i>bare partitive</i> | di    | i    | folletti                    | <i>dei folletti</i>             |
|      | <i>full partitive</i> | uno   | di   | i folletti                  | <i>uno dei folletti</i>         |
|      |                       | 'one' | 'of' | 'the' 'elves'               |                                 |

(58) shows that *van die N*, *de ces N*, *di quei N* and *dei N* differ from full partitives only in their lack of an upstairs determiner. This means they meet the form criterion.

*The meaning criterion*<sup>2</sup>

To argue that the expressions in (56) meet the meaning criterion I show in (59) that they can occur in predicate position and that the resulting sentences entail that their subjects belong to the sets denoted by the nouns included in the predicate. This entailment relation is indicated by the implication arrow.

(59) DUTCH

*van die N*

Jan en Piet zijn van die rare ventjes.

Jan and Piet are of those strange little\_guys

Jan and Piet belong to this type of strange little guys.

→ Jan and Piet are little guys

FRENCH

*de ces N*

Ce sont de ces révélations que le ciel ne donne pas deux fois à la terre.

This are of those revelations that the heaven NEG give NEG two times to the earth

These are revelations of the type that the heaven doesn't give twice to the earth.

→ These are revelations

ITALIAN

*di quei N*

Sono di quei mascalzoni.

They\_are of those rascals

They belong to this type of rascals.

→ They are rascals

*dei N*

Sono dei folletti.

They\_are DEI elves

They are elves.

→ They are elves

(59) shows that *van die N*, *de ces N*, *di quei N* and *dei N* felicitously occur in predicate position and that the resulting sentences entail that their subjects belong to the sets denoted by the nouns included in the predicate. This means these expressions meet the meaning criterion.

---

<sup>2</sup> Thanks to Gianluca Giorgolo and Tommaso Caselli for their help on Italian.

*The use criterion*

In (60) to (63) I show that the expressions in (56) can all occur in regular object position. This means they meet the use criterion of bare partitives.

DUTCH *van die N*

- (60) Ik heb van die rare ventjes gezien.  
I have of those strange little\_guys seen  
I saw (some) of those strange little guys.

FRENCH *de ces N*

- (61) J'ai vu de ces pommes.  
I have seen of these apples

ITALIAN *di quei N*

- (62) A quel punto, lui ha fatto di quelle smorfie...<sup>3</sup>  
at that point, he has made of those faces  
At that point, he made such (ugly/bizarre/...) faces...

ITALIAN *dei N*

- (63) Ho incontrato dei folletti.  
I\_have met of\_the elves  
I met elves.

The examples in (60) to (63) show that *van die N*, *de ces N*, *di quei N* and *dei N* all felicitously appear in object position which means that they meet the use criterion of bare partitives.

*Conclusion*

On the basis of the above discussion I conclude that the expressions in (56) all meet the necessary requirements of bare partitives.

**3.2.2. The partial or full absence of bare partitives in Dutch, English and Spanish**

In what follows I argue that Dutch does not have bare partitives with the definite article and that English and Spanish do not have bare partitives with the definite article nor with a demonstrative.

*Dutch*

On the basis of (64) I argue that Dutch does not have any bare partitives involving the definite article.

---

<sup>3</sup> *Quelle* is the feminine of *quei*.

- (64) \*Ik heb van de rare ventjes gezien.  
I have of the strange little\_guys seen

Given that the use criterion requires bare partitives to be able to appear in argument position the unacceptability of (64) shows that *van de rare ventjes* cannot be a bare partitive.

### English

Standard English does not seem to allow for bare partitives at all. The only examples that have been analyzed as bare partitives in the literature are the following (cf. Hoeksema 1996):

- (65) Again Tarzan came down into the village and renewed his supply of arrows and ate of the offering of food which the blacks had made to appease his wrath. (E. Rice Burroughs, *Tarzan of the Apes*)  
 (66) In the breast of his blouse he carries some coarse dark bread; he ate of this between whiles, and sat munching and drinking near Madame Defarge's counter. (Ch. Dickens, *A Tale of Two Cities*)

Both examples are outside the scope of what I defined as bare partitives, because both occur in the object position of verbs like *to eat*. Examples with other types of verbs are ungrammatical:

- (67) \*I saw of those strange guys.  
 (68) \*I saw of the strange guys.

(67) and (68) show that *of those N* and *of the N* cannot appear in regular argument position in English. On the basis of the use criterion I therefore conclude that *of those N* and *of the N* do not qualify as bare partitives.

### Spanish

Spanish does not seem to have bare partitives either. It does allow for an apparent partitive construction with verbs like *to eat* but not for bare partitives in the object position of any other verb:

- (69) ... Eva, representada en las Sagradas Escrituras como mujer-culpable por haber comido de la manzana prohibida.  
 ... Eva, represented in the Sacred Writings as woman-guilty of having eaten of the apple forbidden

data drawn from the CREA corpus

- (70) \*He visto de los hombres extraños.  
I\_have seen of the men strange

- (71) \*He visto de estos hombres extraños.  
 I\_have seen of those men strange

Given that the use criterion requires bare partitives to be able to appear in regular object position (70) and (71) show that *de los N* and *de estos N* cannot be analyzed as bare partitives.

*Conclusion*

On the basis of the data presented above I conclude that Dutch does not have bare partitives with the definite article and that English and Spanish lack bare partitives altogether.

**3.2.3. Overview**

In 3.2.1. I argued that *van die N*, *de ces N*, *dei N* and *di quei N* meet the necessary criteria I established for bare partitives in 3.1. In 3.2.2. I argued that Spanish and English do not have similar expressions and that Dutch does not have a variant of *van die N* with the definite article. These findings are summarized in the table in (72):

(72)

language	of + definite article + N	of + demonstrative + N
Dutch	✗	✓
English	✗	✗
French	✓	✓
Italian	✓	✓
Spanish	✗	✗

In section 4 I will use the existing cases listed in this table as the basis for my descriptive work on bare partitives.

**4. An empirical investigation of bare partitives**

In section 2 I raised three questions concerning the bare partitive analysis I proposed for *des N*. These questions were related to scope, the reference of the downstairs DP and number. I also pointed out that a thorough description of a wider range of bare partitives is needed to address these questions. The goal of this section is to present this description for the bare partitives I identified in section 3. The questions raised in section 2 will serve as a guideline.

#### 4.1. Scope

The bare partitive analysis I presented in Chapter 1 predicts that *des N* can only take narrow scope. Given that *des N* is also assumed to marginally allow for wide scope, this raises the question whether the analysis should be adapted. In order to answer this question, it is important to establish whether other bare partitives can take wide scope as well. If they can, the analysis should be adapted. If they cannot, I have to explain why *des N* – being a bare partitive – can take wide scope. In this subsection I argue that most other bare partitives cannot take wide scope.

The fact that Dutch *van die N* cannot take wide scope can be argued for on the basis of (73):

- (73) Marie heeft nog nooit van die ventjes gezien.  
 Marie has yet never of those little\_guys seen  
<sup>ok</sup>Marie has never seen any of those little guys  
<sup>#</sup>There are some little guys that Marie hasn't seen

The only interpretation this sentence allows for is that Marie has never seen any instantiations of the type of little guys referred to by *die ventjes*. Crucially, (73) is incompatible with a situation in which Marie saw some but not all of the little guys referred to by *die ventjes*. This means that the bare partitive *van die N* cannot take scope over negation.

French *de ces N* and Italian *di quei N* are similar to Dutch *van die N* as I will argue on the basis of (74) and (75):

- (74) Marie n'a jamais vu de ces enfants.  
 Marie NEG has NEG seen of those children  
<sup>ok</sup>Marie has never seen any of those children  
<sup>#</sup>There are some of those children that Marie hasn't seen
- (75) Non abbiamo di quei libri...  
 Not we\_have of those books  
<sup>ok</sup>We don't sell any of those books  
<sup>#</sup>There are some of those books we don't sell

(74) can only mean that Marie has never seen any of the children referred to by *ces enfants* and (75) can only mean that we do not sell any of the books referred to by *quei libri*. Crucially, (74) is incompatible with a situation in which Marie has seen some of the children and (75) is incompatible with a situation in which we sell some of the books. This means that the bare partitives *de ces N* and *di quei N* cannot scope over negation.

Italian *dei N* is different from the other bare partitives in freely allowing for wide scope. This can be shown on the basis of (76):

- (76) Non ho visto dei bambini.  
 Not I-have seen of the children  
<sup>ok</sup>*I haven't seen any children*  
<sup>ok</sup>*There are some children I haven't seen*

(76) allows for two interpretations: one according to which I haven't seen any children and one according to which there are some children that I haven't seen. This means (76) is compatible with a situation in which I saw some but not all of the children and that *dei N* can take narrow and wide scope with respect to negation. Crucial for the discussion in section 5 is also that *dei N* is assumed to freely allow for wide scope readings (see Chierchia 1997, Storto 2003 and Zamparelli 2008). In this respect it differs from *des N* that has a strong preference for narrow scope readings.

The table in (77) summarizes the above discussion:

(77)

<b>bare partitive</b>	<b>wide scope</b>
Dutch <i>van die N</i>	✗
French <i>de ces N</i>	✗
Italian <i>di quei N</i>	✗
Italian <i>dei N</i>	✓

Dutch *van die N*, French *de ces N* and Italian *di quei N* do not allow for wide scope readings whereas Italian *dei N* does. This table will be the starting point of the discussion about bare partitives and scope in section 5.

#### 4.2. The downstairs DP

In the analysis presented in Chapter 1 I assume that the downstairs DP of *des N* can only get a kind interpretation. At the moment this assumption is nothing more than a stipulation but in order to see where I should fit it into the analysis I need to know whether this constraint applies to other bare partitives as well. If it does, the constraint can be derived at a general level. If it does not, it should be derived as an idiosyncrasy of *des N*. In this subsection I argue that all bare partitives have a kind constraint on their downstairs DP.

##### 4.2.1. Dutch *van die N*

I will argue that the kind constraint holds for Dutch *van die N* on the basis of the following example (adapted from Oosterhof 2005):

- (78) A: Ik heb pas *De Slinger van Foucault* en *Baudolino* gelezen en ben nu bezig in *De Quincunx*.

B: Ik heb nog nooit [van die dikke boeken] gelezen, maar nu je het zegt, zou het eigenlijk nog een goed idee zijn om daarmee te beginnen.

A: I have just read *Foucault's Pendulum* and *Baudolino* and now I'm reading *De Quincunx*.

B: I have never read [of these thick books], but now that you come to mention it, it would be a good idea to do that.

If it were the case that *die dikke boeken* could refer to a specific set of books we would expect *van die dikke boeken* to preferably refer to a subset of the books introduced by A. This is however not the case; the only reading of B's utterance is that he might be interested in reading thick books. From this I conclude that the downstairs DP of *van die N* in Dutch cannot refer to specific entities but only to kinds.

#### 4.2.2. French *de ces N*

On the basis of (79) I will argue that the kind constraint holds for French *de ces N* (example adapted from Xavier Walter).<sup>4</sup>

(79) A: Là j'ai vu trois moutons qui étaient grands comme des boeufs et portent une longue toison très rêche.

B: J'ai vu [de ces bêtes] plus d'une fois!

A: There I saw three sheep that were tall as cows and had a long and very rough fleece.

B: I have seen [of these animals] more than once!

If it were the case that *ces bêtes* in *de ces bêtes* could refer to specific beasts we would expect it to have the sheep introduced by A as its preferred referent. Given that the preferred reading of B's utterance is one in which B has seen similar but not necessarily identical animals I conclude that the downstairs DP of *de ces N* in French cannot refer to specific entities but only to kinds.<sup>5</sup>

#### 4.2.3. Italian *di quei N*

Zamparelli (2008) observed that specific readings are available for DPs of the form in (80) but kind readings are not:

(80) demonstrative + numeral + noun

<sup>4</sup> X. Walter, 1997, *Avant les grandes découvertes: une image de la terre au XIVe siècle: le voyage de Mandeville*, Alban.

<sup>5</sup> This is in line with Kupferman (2001) and against Zribi-Hertz (2003). The analysis I will eventually propose of the kind constraint will allow for marginal counterexamples (see Chapter 3).

If it were possible to interpret the downstairs DP of *di quei N* as a specific DP there should consequently be no difference before and after the addition of a numeral. It turns out though that the bare partitive becomes unacceptable:

- (81) \*Ha fatto di quelle tre smorfie che conosci.  
 He\_has said of those three things that you know

The unacceptability of (81) shows that the downstairs DP in *di quei N* cannot get a specific reading. This strongly suggests that the downstairs DP of *di quei N* is necessarily kind-referring.

**4.2.4. Italian *dei N***

For Italian *dei N* Storto (2003) constructed a context in which the specific interpretation of the downstairs DP of *dei N* would be the only pragmatically felicitous one:

- (82) Ieri Gianni e Paolo sono stati attaccati da gruppi di cani; # sfortunatamente [dei cani] di Gianni non sono stati catturati.  
 Yesterday Gianni and Paolo were attacked by groups of dogs; unfortunately [of the dogs] of Gianni have not been captured.

The only way to make sense of the second part of (82) is to have *dei cani* pick up the dogs that attacked Gianni. Given that the second part of the sentence is odd we can conclude that this reading is not available and that the downstairs DP in *dei N* can only refer to kinds.

**4.2.5. Summarizing table**

The table in (83) summarizes the discussion in 4.2.1. through 4.2.4.

- (83)

bare partitive	kind-constraint
Dutch <i>van die N</i>	✓
French <i>de ces N</i>	✓
Italian <i>di quei N</i>	✓
Italian <i>dei N</i>	✓

I argued that all bare partitives I identified in section 3 are subject to the kind-constraint on downstairs DPs. The data in this table will constitute the starting point of the discussion about bare partitives and the kind constraint in section 5.

### 4.3. Plurality

In Chapter 1 I argued that bare partitives like *des N* are to be interpreted as referring to plural entities. This is illustrated by the fact that (84) can only be true if Jean saw at least two horses:

- (84) Jean a vu des chevaux.  
 Jean has seen of\_the horses  
 Jean saw horses.

The origin of this plurality feature is to be found in default reasoning. A schematic spell-out of the analysis is given in (85):

- (85) **Assumption 1**  
 The number of the referent of a partitive depends on the number of the upstairs determiner.  
**Assumption 2**  
 There is a determiner that marks singularity and only singularity – the numeral *one*. Partitives then have a singular marker.  
**Assumption 3**  
 There is no determiner that marks plurality and only plurality – all plural determiners have an extra quantity or quality component. Partitives then have no plural marker.  
**Default reasoning**  
 If the speaker wants to convey singularity he uses the numeral *one*. In absence of this numeral the hearer is entitled to conclude that the speaker did not want to convey singularity and is further entitled to conclude that the referent of the partitive is plural.

The question this account raises is whether it is as general as a pragmatic account should be. Deriving the plurality of any bare partitive in the pragmatics generates the prediction that all bare partitives have a plural interpretation. Whether this is indeed the case will be investigated in 4.3.1. through 4.3.4. for the bare partitives I identified in section 3.

#### 4.3.1. Dutch *van die N*

Even though it is never explicitly mentioned in the literature Dutch *van die N* only allows for a plural interpretation. This can be shown on the basis of the following minimal pair:

- (86) \*Er ligt [van die boeken] op de tafel.  
 There is [of those books] on the table.  
 (87) Er liggen [van die boeken] op de tafel.  
 There are [of those books] on the table.

The fact that a change in number of the verb in (86) makes the unacceptability disappear shows that *van die N* is necessarily interpreted as a plural.

#### 4.3.2. French *de ces N*

In French a test similar to the existential test in Dutch is not possible. The reason for this is that the existential construction in French does not distinguish between singulars and plurals:

	FRENCH		DUTCH
(88)	Il y a un homme dans la salle. He there has a man in the hall There is a man in the hall	vs.	Er is een man in de zaal. There is a man in the hall There is a man in the hall
(89)	Il y a deux hommes dans la salle. He there has a man in the hall There are two men in the hall	vs.	Er *is / zijn twee mannen in de zaal. There is are two men in the hall There are two men in the hall

Another way to test the plurality of *de ces N* would be to see whether *de ces N* can appear as a nominal predicate of a singular subject. This test is not fool proof though. Indeed, the singular subject might actually impose a singular interpretation (see Chapter 1, fn. 8, p. 84). The only test left is an interpretation test; asking native speakers whether sentences such as (90) can be true if Max saw only one of the horses referred to by *ces chevaux*.

- (90) Max a vu de ces chevaux.  
Max has seen of these horses

The judgements that have been reported in the literature all agree that (90) can only be true if Max saw at least two of the horses referred to by *ces chevaux* (see Morin 1976 for the most explicit statement).

#### 4.3.3. Italian *di quei N*

Italian – just like Dutch – has an existential construction that agrees in number with what follows. This means we can check the number of bare partitives on the basis of the number agreement in the existential construction. This is done for *di quei N* in (91) and (92):

- (91) Negli uffici ci sono di quelle discariche...!  
In\_the offices there are of these dumping sites  
(92) \*Negli uffici c'è di quelle discariche...!  
In\_the offices there is of these dumping sites

The contrast between (91) and (92) shows that *di quei N* forces plural agreement. This leads me to conclude that *di quei N* has a plural interpretation.

#### 4.3.4. Italian *dei N*

As for *di quei N* we can check the number of *dei N* by looking at the number agreement of the existential construction:

- (93) Ci sono dei folletti.  
There are of the elves
- (94) \*C'è dei folletti.  
There is of the elves

On the basis of the fact that *dei N* forces plural agreement I conclude that it has a plural interpretation.

#### 4.2.5. Summarizing table

The table in (95) summarizes the discussion in 4.2.1. through 4.2.4:

(95)

bare partitive	plural interpretation
Dutch <i>van die N</i>	✓
French <i>de ces N</i>	✓
Italian <i>di quei N</i>	✓
Italian <i>dei N</i>	✓

I argued that all bare partitives that were identified in section 3 have a plural interpretation. The table in (95) will be the starting point for the discussion about bare partitives and number in section 5.

#### 4.3. Summary

In the preceding subsections I have given a thorough description of the bare partitives I identified in section 3. The table in (96) gives a summary:

(96)

bare partitive	wide scope	kind constraint	plural interpretation
Dutch <i>van die N</i>	✗	✓	✓
French <i>de ces N</i>	✗	✓	✓
Italian <i>di quei N</i>	✗	✓	✓
Italian <i>dei N</i>	✓	✓	✓

All bare partitives are subject to the kind constraint and have a plural interpretation. The scope facts are less clearcut: *dei N* freely allows for wide scope whereas the other bare partitives only allow for narrow scope.

The empirical work done in this section is the basis for the discussion in section 5 concerning the questions I raised for the bare partitive analysis of *des N* in section 2.

## 5. The analysis of bare partitives

In section 2 I raised three questions for the bare partitive analysis of *des N* I proposed in Chapter 1. After the empirical work in sections 3 and 4 I am now in a position to address these questions. I treat each of them in a separate subsection.

### 5.1. Bare partitives and scope

The bare partitive analysis I presented in Chapter 1 predicts that *des N* can only take narrow scope. Given that *des N* is also assumed to marginally allow for wide scope, this raises the question whether the analysis should be adapted. In order to address this question I looked at a number of other bare partitives in section 4 and checked whether they allow for wide scope or not. The results are summarized in the table in (97):

(97)

bare partitive	wide scope
Dutch <i>van die N</i>	✗
French <i>de ces N</i>	✗
Italian <i>di quei N</i>	✗
Italian <i>dei N</i>	✓

What (97) shows is that most bare partitives do not allow for wide scope, the only exception being *dei N*. These data suggest that the analysis of bare partitives I presented in Chapter 1 is basically correct in deriving obligatory narrow scope for bare partitives. The data however also raise the question what the status is of *dei N* and *des N*. This is the question I focus on in this subsection.

In Chapter 1, I argued that DPs headed by indefinite determiners are different from bare partitives in being scopally free. Given that *dei N* freely takes narrow and wide scope I assume that *dei N* is not a bare partitive and that *dei* is an indefinite determiner (see Storto 2003 and Le Bruyn 2007). The status of *des N* is less clearcut: its preference for narrow scope prevents me from analyzing it as a DP and the fact that it allows for wide scope prevents me from analyzing it as a bare partitive. Given that the main goal of this part of my dissertation is to come to a clearcut distinction between bare partitives and DPs headed by an indefinite article, I will not blur the

distinction I argued for by weakening it to fit in *des N*. The consequence of this is that I have to accept that *des N* can be analyzed both as a bare partitive and as a DP headed by an indefinite determiner (see also Le Bruyn 2007). Note for completeness that there is at least one other option available for the analysis of *des*. In Le Bruyn (2005) I argued that *des N* is neither a bare partitive nor a DP but a NumP. *Des* would then occupy NumP and mark plurality in the same way as a plural morpheme. The exceptional wide-scope behaviour of *des N* could then be attributed to an exceptional raising of *des* to D. This analysis is elegant because it distinguishes *des N* from both *dei N* and bare partitives with a demonstrative D. It raises two questions though. The first is why a plural morpheme would ever optionally raise to D. The second is why *des N* is better at taking wide scope than a bare plural. Without a clear answer to these questions the analysis is as stipulative as the one that takes *des N* to be ambiguous between a bare partitive and a DP headed by an indefinite determiner. A further investigation of the scopal properties of bare plurals, bare partitives, *des N* and *dei N* is required to come to a better analysis of the status of *des N*. In joint work with Min Que and Femke Smits I have tried to make a contribution to this but I am fully aware of the fact that there is a lot more work to be done.<sup>6</sup>

## 5.2. Bare partitives and their downstairs DP

In the bare partitive analysis I proposed for *des N* I stipulated that the downstairs DP of *des N* is kind-referring. In order to turn this stipulation into an insight it was important for me to first establish at what level this kind-constraint plays. Crucially I needed to investigate whether or not other bare partitives are subject to the same constraint. This investigation was carried out in section 4 and its results are summarized in the table in (98). Note that I leave in *dei N* for completeness, despite the fact that I decided not to analyze it as a bare partitive.

(98)

bare partitive	kind-constraint
Dutch <i>van die N</i>	✓
French <i>de ces N</i>	✓
Italian <i>di quei N</i>	✓
Italian <i>dei N</i>	✓

The table shows that all bare partitives are subject to the kind constraint. This constraint should consequently not be derived as an idiosyncrasy of *des N* but rather as something inherent to bare partitives. I will work out a proposal in Chapter 4.

<sup>6</sup> See Smits (2009) and De Vries (2009) on experiments that were developed by Min Que and further worked out by Min Que, Femke Smits, Hannah de Vries and me.

### 5.3. Bare partitives and plurality

The bare partitive analysis I presented in Chapter 1 derives the plurality of *des N* in the pragmatics. This makes the prediction that all bare partitives should have plural reference. In section 4 I checked this prediction and the results are summarized in the table in (99).

(99)

bare partitive	plural interpretation
Dutch <i>van die N</i>	✓
French <i>de ces N</i>	✓
Italian <i>di quei N</i>	✓
Italian <i>dei N</i>	✓

The results in this table show that all bare partitives have plural reference. This means plurality is not a weird quirk of *des N* and that a pragmatic account is feasible.

### 5.4. Summary

In this section I addressed the three questions that were raised by the bare partitive analysis I proposed for *des N*. First, I argued – on the basis of its scope behaviour – that *des N* cannot be unambiguously analyzed as a bare partitive. Second, I argued that the kind constraint on the downstairs DP of bare partitives should be derived as a general constraint and not as an idiosyncrasy of *des N*. Third, I argued that the prediction a pragmatic account of the plurality of *des N* makes is borne out: all bare partitives have plural reference. As an important aside I pointed out that Italian *dei* is probably to be analyzed as a determiner and not as a bare partitive.

## 6. Conclusions and outlook

This chapter aimed at presenting a description of a broader range of bare partitives that would allow me to address the questions that were raised by the analysis I proposed for *des N*. These questions concerned the scope behaviour of *des N*, the kind constraint on its downstairs DP and its number interpretation. The main conclusion I draw on the basis of the comparison with other bare partitives is that *des N* cannot unambiguously be analyzed as a bare partitive. Further conclusions concerning its bare partitive use are that the analysis I presented is on the right track as far as scope and number are concerned and that it should be supplemented with a general kind constraint on the downstairs DP of bare partitives. In Chapter 3 I show how this general constraint can be worked out.



## Chapter 3: Bare partitives and kinds

### 1. Introduction

In Chapter 1 I showed that the downstairs DP of *des N* has to be kind-referring. I argued for this constraint on the basis of the context in (100):

- (100) Hier j'ai acheté cinq brouillons de Picasso et deux peintures de Modiano. J'ai mis des brouillons dans le living.  
Yesterday I have bought three sketches by Picasso and two paintings by Modiano. I have put DES sketches in the livingroom.  
Yesterday I bought three sketches by Picasso and two paintings by Modiano. I put sketches in the livingroom.

If we assume that we always prefer to interpret definite descriptions as referring back to something that has been introduced before, we expect the downstairs DP of *des N* to have some of the sketches of Picasso as its preferred referent. Given that this is not the case it seems plausible to conclude that the downstairs DP of *des N* cannot refer to a specific referent. If we furthermore assume that definites can only refer to specific referents or to kinds, (100) allows us to conclude that the downstairs DP of *des N* necessarily refers to kinds. In chapter 2 I constructed similar arguments for other bare partitives. It appears then that the downstairs DP of bare partitives is necessarily kind referring. I will refer to this property of bare partitives as *the kind constraint*.

The goal of this chapter is to give an account of the kind constraint. In Section 2 I present the account itself. Section 3 picks up on apparent counter-examples such as the one in (101):

- (101) Hier j'ai acheté cinq pommes et trois poires. Aujourd'hui j'ai mangé des pommes.  
Yesterday I have bought five apples and three pears. Today I have eaten DES apples.  
Yesterday I bought five apples and three pears. Today I ate apples / some of the apples.

In Chapter 2 I argued that *des pommes* should not be analyzed as a bare partitive and assumed with Kupferman (1979) that it functions as a prepositional complement of the verb *to eat*. A question I left unanswered is why verbs like *to eat* can have two kinds of complements – a DP and a PP one. It is this question I will answer in Section 3.

## 2. The kind constraint

### 2.1. A specific semantics

The first thing to be done in order to give an account for the kind constraint is to work out the semantics of a bare partitive like *des N* in case it would have a specific downstairs DP. This is done in (102) for *des arbres* ('of\_the trees'):

- (102) a.  $\llbracket \text{les arbres} \rrbracket_{\text{specific}}$   $\iota \text{tree}$   
 b.  $\llbracket \text{de} \rrbracket$   $\lambda y \lambda x (\leq(x, y))$   
 c.  $\llbracket \text{des arbres} \rrbracket$   $\lambda x (\leq(x, \iota \text{tree}))$   
 d.  $\llbracket \text{des arbres} \rrbracket$  + default reasoning  $\lambda x (\text{plural}(x) \& \leq(x, \iota \text{tree}))$   
 e.  $\llbracket \text{des arbres} \rrbracket$  + type-shifting  $\lambda Q \exists x [\text{plural}(x) \& \leq(x, \iota \text{tree})] \& Q(x)$

In order to make the comparison between the “specific” and the “kind” version of *des N* I repeat below the analysis I proposed for the “kind” version:

- (103) a.  $\llbracket \text{les arbres} \rrbracket_{\text{kind}}$   $\Sigma \text{tree}$   
 b.  $\llbracket \text{de} \rrbracket$   $\lambda y \lambda x (\leq(x, y))$   
 c.  $\llbracket \text{des arbres} \rrbracket$   $\lambda x (\leq(x, \Sigma \text{tree}))$   
 d.  $\llbracket \text{des arbres} \rrbracket$  + default reasoning  $\lambda x (\text{plural}(x) \& \leq(x, \Sigma \text{tree}))$   
 e.  $\llbracket \text{des arbres} \rrbracket$  + type-shifting  $\lambda Q \exists x [\text{plural}(x) \& \leq(x, \Sigma \text{tree})] \& Q(x)$

The most important difference between (102) and (103) is that the former is presuppositional in the sense that it presupposes the existence of a referent in the actual world whereas the latter is not. This follows from the fact that the *iota* operator selects the supremum of an extensional lattice whereas the  $\Sigma$  operator selects the supremum of an intensional lattice. To be defined in a world the former but not the latter then requires the lattice to contain at least two atomic individuals in that world.

### 2.2. The competitor

In Chapter 1, section 4 I argued that “kind” *des Ns* are in competition with bare plurals and used this to explain the absence of bare plurals in French. A first attempt to explain the absence of “specific” *des Ns* would then be to argue that they too are blocked by “kind” *des Ns*. The semantics presented in 2.1. however prevents me from doing so. The reason for this lies in the assumption that an expression can only block another one if both are truth-conditionally equivalent. The fact that “specific” and “kind” *des Ns* are not can be shown on the basis of (104) and (105) (the †-sign indicates that the sentence is hypothetical):

- (104) † J'ai vu des femmes. SPECIFIC  
 (105) J'ai vu des femmes. KIND

In a model with no women the sentence in (104) is *undefined* because of a presupposition failure whereas the sentence in (105) is *false*.

The contrast between (104) and (105) not only tells me that my first attempt to explain the absence of “specific” *des Ns* cannot work but also suggests a direction for my second attempt: given that “specific” *des Ns* are presuppositional, whatever expression blocks them should also be presuppositional. In (106) through (110) I give a list of possible candidates:

- (106) plusieurs des femmes  
several DES women
- (107) certaines des femmes  
certain DES women
- (108) différentes des femmes  
different DES women
- (109) diverses des femmes  
diverse DES women
- (110) les femmes  
the women

Let me first look at (106) through (109). All of them are full partitives and presuppositional. Neither of them is truth-conditionally equivalent to the specific bare partitive though, all of them having an extra quantity or quality component (see Chapter 1, 2.4.4.). Let me now have a look at the candidate in (110) – the downstairs DP itself. It is clear that this DP is presuppositional. On the basis of (111) and (112) I will furthermore argue that it has no quantity nor quality component:

- (111) Les gorilles de ce pays meurent par centaines.  
The gorillas of this country die by hundreds  
“The gorillas of this country die by the hundreds.”
- (112) Les kilomètres qui nous séparent me rendent malheureux.  
“he kilometers that separate us make me unhappy”

The fact that DPs headed by the definite article are compatible with quantitative modifiers (see (111)) suggests that it has no quantitative component of its own. The fact that the definite article can combine with nouns that refer to objects that shouldn’t be distinguishable one from the other (measure nouns like *kilometer*, *kilo*, *liter*, ...) furthermore suggests that it has no quality component of its own.

Even though DPs headed by the demonstrative then seem to pass tests full partitives don’t pass it is far from clear that it could actually be seen as a competitor for “specific” *des N*. Indeed, whereas the definite DP refers to the supremum of a lattice the “specific” *des N* has the whole non-atomic part of the lattice as its denotation. Appearances can be deceiving though: definite DPs are known to allow for non-maximal readings quite easily.

- (113) The boys had a sandwich.

Sentence (113) is generally felt to be true even if not all the boys in the universe had a sandwich and even if not all the boys in a contextually salient subset of the universe had a sandwich. Linguistic support for this comes from the felicity of sentences such as the following (taken from Brisson 1998):

(114) The boys ate a sandwich, but Adam didn't.

If *the boys* only referred to all the boys in a contextually restricted subset of the universe, (114) would contain a contradiction and should therefore be infelicitous. The fact that it is not seems to suggest that definites should be given a non-maximal semantics. This can be done in two ways:

(115)  $\lambda Q \exists x(\text{plural}(x) \& \leq(x, \text{boys}) \& Q(x))$

(116)  $\iota \lambda x(\text{plural}(x) \& \text{boy}(x) \& D'(x))$

(115) introduces non-maximality explicitly through the use of the part-of relation whereas (116) gets it through a domain restriction indicated by the condition  $D'(x)$  where  $D'$  is a subset ( $\subseteq$ ) of the discourse domain  $D$ .<sup>1</sup> The crucial thing for my present purposes is that (115) and (116) are truth-conditionally equivalent and that (115) is identical to the analysis I proposed for bare partitives with a specific DP. This will be the basis of the blocking account I will present in 2.3.

### 2.3. Blocking of specific bare partitives

In this subsection I will present a blocking analysis for bare partitives cast in bi-directional OT (see Blutner 2000).

Before spelling out the details of the analysis it is important to give the intuitions that underlie it. The first intuition is that language is economic and that whenever we can, we will prefer shorter forms over longer ones. In the case of making a production choice between a definite DP and the bare partitive containing it this means that we would go for the definite DP. The second intuition is more subtle and is concerned with the kind / non-kind distinction : definite determiners can be used to refer to kinds but are not designed to do so. It is this intuition that was used by Farkas & de Swart (2005,2007) to account for the fact that in languages like English a DP headed by the definite article cannot be interpreted as referring to kinds. In the case of making an interpretation choice for a definite DP this means that – all other things being equal – we will prefer to give it a specific interpretation.

Let us now take a closer look at the ingredients of the analysis : there are two constraints, two forms and two meanings:

---

<sup>1</sup> The domain restriction could also be formalized with covers (see Brisson 1998).

(117) **two constraints:**

constraint 1: \*FORM: Avoid structure

constraint 2: \*DEF/KIND: Don't interpret definite DPs as referring to kinds.

(118) **two forms :**form 1 : *les femmes* ('the women')form 2 : *des femmes* ('of the women')(119) **two meanings :**meaning 1 :  $\lambda Q \exists x (\text{plural}(x) \& \leq (x, \iota \text{women}) \& Q(x))$ meaning 2 :  $\lambda Q \exists x (\text{plural}(x) \& \leq (x, \Sigma \text{women}) \& Q(x))$ *Constraints***\*FORM**

The first constraint is a standard economy constraint in OT and penalizes structure. Structure can be thought of in number of syntactic projections.

**\*DEF/KIND**

The second constraint is a constraint on meanings and penalizes generic readings of definite DPs.

*Forms*

The first form is the definite DP whereas the second form is its corresponding bare partitive. For the analysis it is relevant to be explicit about the structures I assume:

(120) *les femmes*: [<sub>DP</sub> *les* [<sub>NumP</sub> *pl* [<sub>N</sub> *femme*]]](121) *des femmes*: [<sub>PartP</sub> *de* [<sub>DP</sub> *les* [<sub>NumP</sub> *pl* [<sub>N</sub> *femme*]]]]]

Crucially *des femmes* comes out as having more projections than *les femmes*. When optimizing over forms *les femmes* will consequently come out as the optimal form.

*Meanings*

The first meaning (henceforth *specific*) translates the non-maximal reading of a specific definite DP or the meaning of its corresponding bare partitive. The second meaning (henceforth *generic*) translates the non-maximal reading of a kind-referring definite DP or the meaning of its corresponding bare partitive. When optimizing over meanings the specific meaning will come out as the optimal one.

*The bi-directional OT analysis*

The bi-directional OT tableau looks as follows:



## 2.4. Conclusion

In this subsection I presented a pragmatic account of the kind constraint within bi-directional OT. The gist of the analysis is that bare partitives on a specific interpretation of their downstairs DP are blocked by the downstairs DP on its non-maximal interpretation.

In the next section I turn to an old chestnut: the DP/PP alternation with *eat*-verbs that apparently allows for bare partitives with a specific interpretation.

## 3. A note on *eat*-verbs

In Chapter 2 I pointed out that in some cases bare partitives do seem to have downstairs DPs that can have a specific interpretation. This is the case for bare partitives appearing in the object position of verbs like *to eat*, *to drink*, ...:

- (127) Hier j'ai acheté cinq pommes et trois poires. Aujourd'hui j'ai mangé des pommes.  
 Yesterday I have bought five apples and three pears. Today I have eaten DES apples.  
 Yesterday I bought five apples and three pears. Today I ate apples / some of the apples.

In Chapter 2 I argued that *des pommes* in (127) cannot be analyzed as a bare partitive, and joined Kupferman in assuming that verbs like *to eat* allow for two kinds of complements: a (partitive) prepositional one and a DP one. Extra support for not analyzing *des pommes* in (127) as a bare partitive comes from the fact that languages that don't allow for bare partitives do allow for verbs like *to eat* to take partitive complementation (see Chapter 2). In (128) and (129) this is illustrated for English and Spanish:

- (128) Again Tarzan came down into the village and renewed his supply of arrows and ate of the offering of food which the blacks had made to appease his wrath. (E. Rice Burroughs, *Tarzan of the Apes*)  
 (129) ... Eva, representada en las Sagradas Escrituras como mujer-culpable por haber comido de la manzana prohibida.  
 ... Eva, represented in the Sacred Writings as woman-guilty of having eaten of the apple forbidden

data drawn from the CREA corpus

An analysis stipulating that verbs like *to eat* allow for two kinds of complementation only makes sense if there is something special about these verbs that makes them more 'partitive' than others (cf. Zribi-Hertz 2003). The answer – I would like to suggest – lies in the semantics of these verbs.

*The semantics of verbs like to eat*

Verbs like *to eat* have the following special semantic properties:

(130) *mapping to objects*

*R* obeys mapping to objects iff  $\forall e, e', x [R(e, x) \& e' \leq e \rightarrow \exists x' [x' \leq x \& R(e', x')]]$

*mapping to events*

*R* obeys mapping to events iff  $\forall e, x, x' [R(e, x) \& x' \leq x \rightarrow \exists e' [e' \leq e \& R(e', x')]]$

(see Krifka 1989, Kiparsky 1998)

If a verb obeys both *mapping to objects* and *mapping to events*, this means that there is a one-to-one correspondence between the events it denotes and the object parts that are subjected to these events. For a verb like *to eat* this means that every subevent of eating an apple involves the eating of a piece of the apple and vice versa.

Verbs that obey *mapping to objects* and *mapping to events* are arguably more sensitive to partitivity than others. The reason is that for these verbs every part of an object matters: the event of eating part X of an apple is different from the event of eating a part of part X and vice versa. This explains why these verbs in general allow for two kinds of complementation: DPs and PPs. However, this does not yet explain why we get these two kinds of complementation with definite DPs. Indeed, given that definite DPs can be interpreted non-maximally we would still expect verbs like *to eat* not to take PPs with definite DPs. Why they can is the question I turn to next.

*Eat-verbs and non-maximal definites*

From an economy point of view, the only way I can explain that *eat*-verbs can take PP complementation with a non-maximal definite is to show that there is a meaning difference between non-maximal definites and their corresponding PPs. This is what I will do: I will illustrate the difference and explain where it comes from. Note that it is crucial that the meaning difference is not truth-conditional: a truth-conditional difference would undermine the account I set up in section 2 to derive the kind constraint on bare partitives.

On the basis of (131) and (132) I argue that there is a non-truth-conditional difference between normal and partitive complementation with definite DPs:

(131) Yesterday I bought a bottle of wine and a bottle of brandy. This morning I found out that...

a. ... Marie a bu la bouteille de vin.

Marie drank the bottle of wine

b. ... Marie a bu de la bouteille de vin.

Marie drank of the bottle of wine

(132) a. ??Marie a bu la bouteille de vin et Jean aussi.

Marie drank the bottle of wine and John did too.

- b. Marie a bu de la bouteille de vin et Jean aussi.  
 Marie drank of the bottle of wine and Jean did too.

(131a) and (131b) are each compatible with Marie not drinking the whole bottle of wine. The contrast in (132a) and (132b) however suggests that – as far as the discourse is concerned – (131a) entails that the bottle of wine is empty whereas (131b) leaves open the possibility of there being some wine left.

The difference between DP and PP complementation illustrated in (131) and (132) can be linked to a subtle but crucial difference between non-maximal definites and their partitive version. I explain the difference between the two on the basis of (133) and (134):

(133)  $\lambda x(\text{wine}(x) \& D'(x))$

(134)  $\lambda P \exists x [\leq(x, \lambda y(\text{wine}(y) \& D'(y))) \& P(x)]$

I propose that (133) is the non-maximal semantics of *the bottle of wine* in (131a) and that (134) is the semantics of its partitive version in (131b). Note that  $D'(x)$  introduces a domain restriction. (133) refers to the maximum quantity of contextually relevant wine whereas (134) refers to part of the maximum quantity of contextually relevant wine. The difference between (133) and (134) is not truth-conditional: both would allow for Marie not to have finished all of the wine. For (133) this follows from the contextual restriction introduced by  $D'(x)$ , for (134) this follows from the same contextual restriction and – independently – from the part operator. The difference lies in the fact that (133) entails that the contextually relevant proportion of wine was finished whereas (134) allows a similar contextually relevant proportion of wine not to be finished. Formally, the crucial difference is that non-maximal definites are contextually maximal whereas their partitive version can be contextually non-maximal. This subtle difference explains the contrasts noted in (131) and (132). At a general level it also explains why verbs like *to eat* allow DP and PP complementation with non-maximal definite DPs.

### Conclusion

In the preceding discussion I argued that verbs like *to eat* are more ‘partitive’ than others because they obey *mapping to objects* and *mapping to events*. This explains why they in general combine with DPs and PPs. I furthermore argued that the DP/PP alternation with non-maximal definites is possible because of a subtle non-truth-conditional difference between the two.<sup>2</sup>

---

<sup>2</sup> In its strongest version, the analysis proposed here predicts that any verb that obeys mapping to objects and mapping to events would allow for partitive complementation. This would e.g. mean that a verb like *to build* should also be able to take partitive complements. The fact that *to build* does not easily take partitive complements indicates that the analysis should be further restricted. See Englebort (1996) and Zribi-Hertz (2003) for suggestions.

#### 4. Summary and outlook

In this short chapter I did two things. On the one hand I proposed a pragmatic account of the kind constraint on the downstairs DP of *des N*. On the other hand I showed how the semantics of *eat*-verbs makes them into plausible candidates for having both DP and PP complementation, even with non-maximal definites. The pragmatic account of the kind constraint complements the analysis I proposed in Chapter 1. In the next chapter I introduce another complement, *viz.* an analysis of the competition between *des N* and the bare plural worked out in diachrony.

## Chapter 4: Bare plurals and *des N*: a diachronic analysis

### 1. Introduction

In Chapter 1 I proposed a bare partitive analysis of *des N* and in Chapter 2 and 3 I defended and completed this analysis. The goal of the present chapter is to work out in detail one of the claims I made about the analysis, *viz.* that it can account for the fact that French does not allow for bare plurals. The gist of the proposal is simple: the bare partitive analysis of *des N* makes it truth-conditionally equivalent to the bare plural and – unless there were some pragmatic distinction between the two – we would expect only one of them to appear in French. There are however two questions that still need to be answered:

- (i) Did French ever allow for bare plurals?
- (ii) If yes, when and why did *des N* supersede them?

To answer these questions I have to look into diachrony. In section 2 I present the diachronic facts in an analysis-neutral way. In section 3 I evaluate the proposed analyses and in section 4 I present my own. Sections 5 and 6 are devoted to a historical simulation of my analysis and section 7 concludes.

Note that sections 5 and 6 can be read as appendices: they are not crucial for the understanding of the analysis but they do add to its plausibility and precision.

### 2. The diachronic facts

The most neutral description of the competition between *des N* and bare plurals is in terms of their availability and the possible interpretations of *des N*. Carlier (2007) distinguishes two stages. In the first stage, *des N* appears mainly with verbs like *to eat*, *to drink*, ... and only allows for a specific interpretation of its downstairs DP. At this stage, which lasted up to the 14<sup>th</sup> century, French productively allowed for bare plurals. In the second stage, *des N* is generalized to all verbs and no longer only allows for a specific interpretation of its downstairs DP but also for a kind interpretation. Note that the specific interpretation remained restricted *des N* in combination with verbs like *to eat*, *to drink*, ... From the second stage onwards, French no longer allowed for bare plurals.

### 3. Previous analyses

I distinguish between two analyses in the diachronic literature on *des N*: the Deleted Quantifier Hypothesis (Foulet 1965) and the Prepositional Object Hypothesis (based on Kupferman 1976, 1994, 1998 and worked out in Carlier 2007).<sup>1</sup> I will go over each of them and present a short evaluation.

---

<sup>1</sup> The labels are borrowed from Carlier (2007).

### *The Deleted Quantifier Hypothesis*

According to Foulet (1965) the origin of the present-day use of *des N* as an indefinite plural lies in the construction *assez de N* ('enough of the N') in Old French. The gist of his proposal is that *assez de N* evolved into *des N*, a process involving the subsequent addition of the definite article and the deletion of the upstairs quantifier. This process is schematized in (135):

(135) *assez de N* -> *assez des N* -> *des N*

### *The Prepositional Object Hypothesis*

The gist of the Prepositional Object Hypothesis worked out in Carlier (2007) is that *des N* was recruited as the plural indefinite article in French. The use that gave rise to its recruitment was the specific use of *des N* in constructions of the type *manger des N* ('eat of the N') where *de* still had its full prepositional use. The trigger for recruitment – according to Carlier – was the phonetic erosion of the final *-s* marking plurality which took place in the 14<sup>th</sup> century. This phonetic erosion led to the impossibility of distinguishing between singular and plural nouns and the necessity to reinstate this distinction through some other mean like a plural indefinite article.

### *Evaluation*

The main problem Foulet's and Carlier's analyses face is that they don't survive Occam's razor. Their starting point is that expressions of the type *of + determiner + N* should never appear in argument position and that the acceptability of *des N* has to be explained by some extra assumption: Foulet assumes *des N* was originally preceded by a quantifier and Carlier assumes *des N* was a prepositional object that started to lead a life of its own. The acceptability of other expressions of the type *of + determiner + N* suggests that the extra assumptions Foulet and Carlier make are unnecessary and consequently unwarranted. Carlier would probably object to this, claiming that if *des N* could be analyzed as a bare partitive, there would be no reason for it not to have been attested before the 15<sup>th</sup> century.<sup>2</sup> In the following sections I will argue that this claim is unsubstantiated.

## **4. A novel way of looking at the evolution of *des N***

In this section I informally present a new way of looking at the evolution of *des N* and work out the competition between *des N* and the bare plural. The crucial insight I will develop is that the reason the kind reading of *des N* was not attested before the 15<sup>th</sup> century is that this reading was blocked by the bare plural. *Kind* and *specific*

---

<sup>2</sup> Carlier develops an argument along these lines against Foulet on p.8 in the form of a rhetorical question: "Why is the generic interpretation impossible when the definite article is combined with *de* [...] while during the same period the generic interpretation is possible without *de* [...]"

*reading of des N* will be used as shorthand for *des N with a kind/specific reading of the downstairs DP*. I first look at the kind reading and then at the specific reading.

#### *The kind reading*

For the evolution of its kind reading I start from the assumption that *des N* and the bare plural are truth-conditionally equivalent and should therefore be mutually exclusive: at any stage of the evolution of French, only one of them should appear. As pointed out in section 2, this assumption is supported by the facts. The second assumption I make is that bare plurals – because of the fact that they are syntactically less complex – should be preferred over *des N*. This assumption derives the stage of French before the 15<sup>th</sup> century, in which *des N* did not appear. The last question to be answered is why *des N* did appear in the 15<sup>th</sup> century. The answer lies in the phonetic erosion of the final *-s* (cf. Carlier 2007 & Delfitto & Schroten 1991): the moment bare plurals are no longer distinguishable from bare singulars, they no longer form a competitor for *des N*. Put differently: the blocking of the kind use of *des N* by the bare plural stopped in the 15<sup>th</sup> century because bare plurals ceased to exist.

The analysis I just sketched is extremely simple and does not require *des N* to have undergone any semantic change. The only difference between the period before the 15<sup>th</sup> century and the period from the 15<sup>th</sup> century onwards is that the bare plural disappeared and stopped blocking *des N*.

#### *The specific use*

For the evolution of the specific use I have to make no further assumptions. As I pointed out in Chapter 3, the specific use of *des N* is always blocked by the non-maximal definite it contains, except with verbs like *to eat*. Given that there is no phonetic evolution that could have any effect on this blocking I predict this use of *des N* to be blocked throughout the history of French for all verbs except *eat* verbs. The data summarized in section 2 show that this prediction is borne out.

#### *Conclusion*

In this section I informally worked out a new analysis of the evolution of *des N* that boils down to saying that it did not evolve. The reason that the kind use of *des N* appeared in the 15<sup>th</sup> century has nothing to do with the semantic evolution of *des N* but only with the disappearance of the bare plural. This crucially means that it is not necessary to resort to the Deleted Quantifier Hypothesis nor to the Prepositional Object Hypothesis presented in section 3 to account for the evolution of *des N*.

The questions raised in section can now be answered:

- (i) Did French ever allow for bare plurals?
- (ii) If yes, when and why did *des N* supersede them?

Bare plurals were the first choice in French up to the moment they disappeared because of phonetic erosion. *Des N* was the second choice that got promoted once the first choice disappeared.

The analysis also generates new questions, the most important one of which is whether people can learn to use an item with a meaning it never manifested before. In the next two sections I will use historical simulations to argue that people can. Section 5 presents the software I will use for the simulation and section 6 presents the simulation itself.

## 5. Simulating the evolution of *des N*: the software

In this section I present a simplified introduction to EvolOT, the software I will be using in section 6 to argue that language learners can learn to use a form with a meaning it never manifested before and to simulate the evolution of *des N*. After the general introduction in 5.1. I zoom in on the predictions EvolOT makes about non-existing form/meaning pairs. These predictions will prove crucial for the discussion in section 6.

### 5.1. EvolOT

EvolOT works in cycles, each cycle having two components: a learning component (5.1.1.) and a production component (5.1.2.). An iteration of the two gives rise to a model of evolution (5.1.3.). The presentation here is slightly less complicated than in Part I of this dissertation. The reason for this is that the analysis does not require me to go to the level of the constraints.

#### 5.1.1. The learning component

Imagine we have a language with two forms A and B and two meanings A' and B'. Imagine furthermore that we have to learn how to use this language on the basis of a fully disambiguated corpus with the following frequencies of form-meaning pairs:

(136)

<A,A'>	50
<A,B'>	0
<B,A'>	50
<B,B'>	0

Suppose now that we learn using the technique in (137):

(137) Bi-directional learning

We take a random form-meaning pair from the corpus and ...

I. ...make a prediction about what meaning would correspond to the form. If it turns out that our prediction is correct we do nothing but if it turns out that we are wrong we remember this and next time we have to interpret the same form we will favour the other interpretation a bit more. If we e.g. took  $\langle A, A' \rangle$  and wrongly predicted  $B'$  to be the meaning we will favour the meaning  $A'$  a bit more. We do this by changing the probabilities we got from previous trials (on the very first trial there is no difference in probabilities) in the following way:<sup>3</sup>

(i)  $\uparrow \langle A, A' \rangle$  (ii)  $\downarrow \langle A, B' \rangle$

II. ...make a prediction about what form would correspond to the meaning. If it turns out that our prediction is correct we do nothing but if it turns out that our prediction is wrong we remember this and next time we have to give a form for the same meaning we will favour the other form a bit more. This is done in the same way as illustrated in I.

We repeat the above process  $X$  times.

The effect of this learning technique will be that we will strongly favour  $\langle A, A' \rangle$  and  $\langle B, A' \rangle$  over  $\langle A, B' \rangle$  and  $\langle B, B' \rangle$ .

### 5.1.2. The production component

After the completion of the learning process EvolOT starts its production component. This production component consists of an algorithm that randomly generates a set of form-meaning pairs. The frequency of each form-meaning pair depends on the preferences we obtained during the learning process. For our running example it would mean that  $\langle A, A' \rangle$  and  $\langle B, A' \rangle$  will be produced far more often than  $\langle A, B' \rangle$  and  $\langle B, B' \rangle$ .

The production component ends when for each meaning the number of times it was expressed in the original corpus is reached.

### 5.1.3. Evolution

The output of the production component is a new corpus that can be used to start a new cycle. These cycles can be thought of as generations of speakers of a language, each learning the language and producing a corpus that serves as an input for the

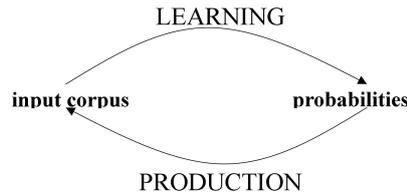
---

<sup>3</sup> The influence of wrong predictions on the probability passes in principle through a reranking of a set of constraints. Using bias constraints (Matusch 2007) we can minimize the influence of this intermediate stage. The bias constraints that will underlie the simulations to follow were defined as follows: 1:( $o==1 \& i==1$ ); 2:( $o==1 \& i==2$ ); 3:( $o==2 \& i==1$ ); 4:( $o==2 \& i==2$ ).

Constraint 1 penalizes the pair  $\langle A, A' \rangle$ , constraint 2  $\langle A, B' \rangle$ , constraint 3  $\langle B, A' \rangle$  and constraint 4  $\langle B, B' \rangle$ .

next generation. An iteration of cycles then corresponds to a simulation of language evolution. (138) gives a pictorial representation of the cyclic set-up of EvolOT:

(138)



The cyclic set-up of EvolOT

## 5.2. EvolOT and non-existing form/meaning pairs

In this section I zoom in on the predictions EvolOT makes with respect to unattested form/meaning pairs. I will focus on one concrete question, *viz.* why the input corpus in (139) will at some point in the evolution give rise to one of the output corpora in (140) despite the fact that (140) does not have any attested <B,A'> or <B,B'> form/meaning pairs:

(139)

form A	form B	
50	0	meaning A'
50	0	meaning B'

(140)

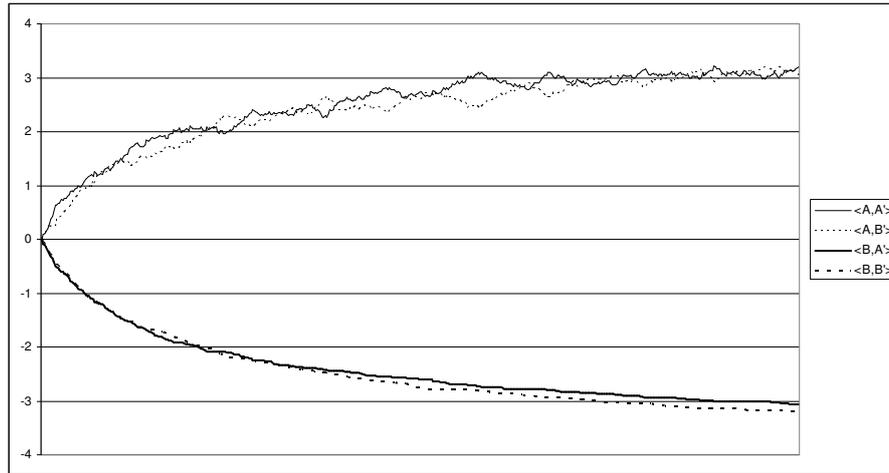
form A	form B	
50	0	meaning A'
0	50	meaning B'

form A	form B	
0	50	meaning A'
50	0	meaning B'

The corpus in (139) represents the state of a language in which both form A and form B are semantically ambiguous between A' and B' but in which form A is the only one that has attested occurrences of either of them. This might look as a weird language at first but in section 6 it will become clear that (139) is very close to the stage of French at the beginning of the 15<sup>th</sup> century: bare nominals had become fully ambiguous between a singular and a plural interpretation and *des N* was unattested with either of them even though its semantics made it compatible with both.

The course of the learning component is given in (141). Each line corresponds to the probability ranking of a form/meaning pair:<sup>4</sup>

(141)



What I would like to focus on are the lines corresponding to  $\langle A, A' \rangle$  and  $\langle A, B' \rangle$ : they constantly change their respective ranking even though the input corpus seems to suggest that they should have exactly the same one. This constant change in ranking is due to the fact that learning is item-based and that the ranking of  $\langle A, A' \rangle$  is not independent of the ranking of  $\langle A, B' \rangle$ . To see that the rankings are not independent it suffices to have a look at possible promotions and demotions in case the learning algorithm is fed with an instantiation of  $\langle A, A' \rangle$  or  $\langle A, B' \rangle$ :

(142) pair fed to the algorithm	possible promotion	possible demotion
$\langle A, A' \rangle$	$\langle A, A' \rangle$	$\langle B, A' \rangle, \langle A, B' \rangle$
$\langle A, B' \rangle$	$\langle A, B' \rangle$	$\langle A, A' \rangle, \langle B, B' \rangle$

Crucially, the promotion of  $\langle A, A' \rangle$  is not independent of the demotion of  $\langle A, B' \rangle$  and vice versa. This means that – no matter how long the learning process lasts –  $\langle A, A' \rangle$  and  $\langle A, B' \rangle$  cannot end up with the ranking they had in the input corpus. Given that the learning process cannot take forever the final ranking of  $\langle A, A' \rangle$  and

<sup>4</sup> For those familiar with EvolOT the graph might look as a mirror image of what it should be like. The reason for this is that I simplified the introduction to EvolOT (see fn. 3); the lines don't give the ranking of the constraints but the preferences that can be established on the basis of this ranking.

<A,B'> is left to chance in the sense that an extension of the learning process might lead to a different ranking.

Let me now take the result of the learning component to the production component. The important point to make here is that the probabilities in (141) will necessarily give an output corpus that presents form/meaning pairs that were not attested in the input corpus. The reason for this is that the ranking of <A,A'> and <A,B'> slightly differ and that one of them is therefore slightly less preferred over its competitor with the form B. This results in an output corpus like the one in (143):

(143)

form A	form B	
42	8	meaning A'
39	11	meaning B'

The step from the input corpus in (139) to the output corpus in (143) shows that form/meaning pairs that are not attested in an input corpus can be attested in an output corpus.

Suppose now we repeat the learning and production component. This would result in the following output corpora for the first eight generations:

(144)

**output generation 1**

form A	form B	
42	8	meaning A'
39	11	meaning B'

**output generation 2**

form A	form B	
35	15	meaning A'
30	20	meaning B'

**output generation 3**

form A	form B	
33	17	meaning A'
23	27	meaning B'

**output generation 4**

form A	form B	
36	14	meaning A'
21	29	meaning B'

**output generation 5**

form A	form B	
37	13	meaning A'
7	43	meaning B'

**output generation 6**

form A	form B	
47	3	meaning A'
1	49	meaning B'

**output generation 7**

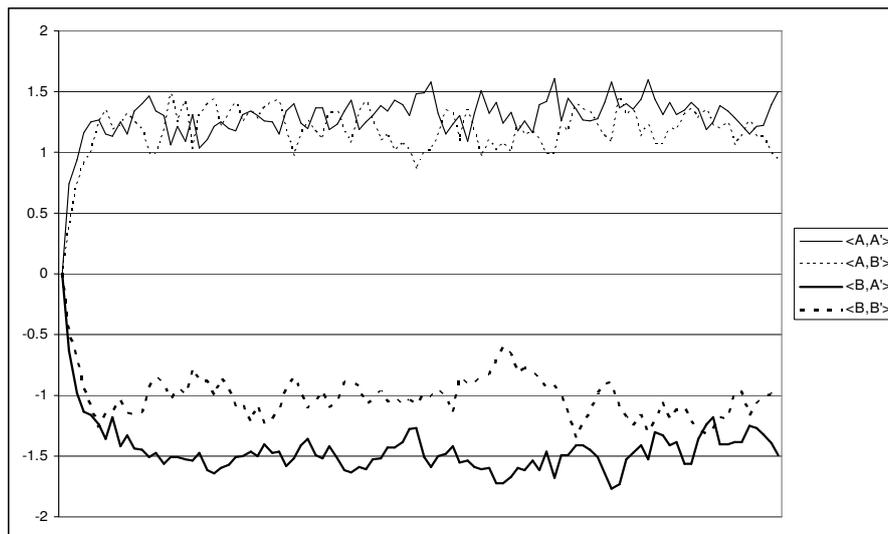
form A	form B	
49	1	meaning A'
0	50	meaning B'

**output generation 8**

form A	form B	
50	0	meaning A'
0	50	meaning B'

What happens is that the difference between  $\langle A, A' \rangle$  and  $\langle A, B' \rangle$  is progressively enlarged: in the output of generation 1 the difference is 3 and it increases in the following generations from 5 over 10, 15, 30, 46 and 49 to 50. There are two reasons for this increase. The first lies in the learning process: like before,  $\langle A, A' \rangle$  and  $\langle A, B' \rangle$  will constantly compete with each other leaving the final ranking to chance. The only difference is that it is more likely that  $\langle A, A' \rangle$  will be more and more preferred than  $\langle A, B' \rangle$ . This can be seen from the graph in (145) that represents the learning process on the basis of the output corpus of generation 1.

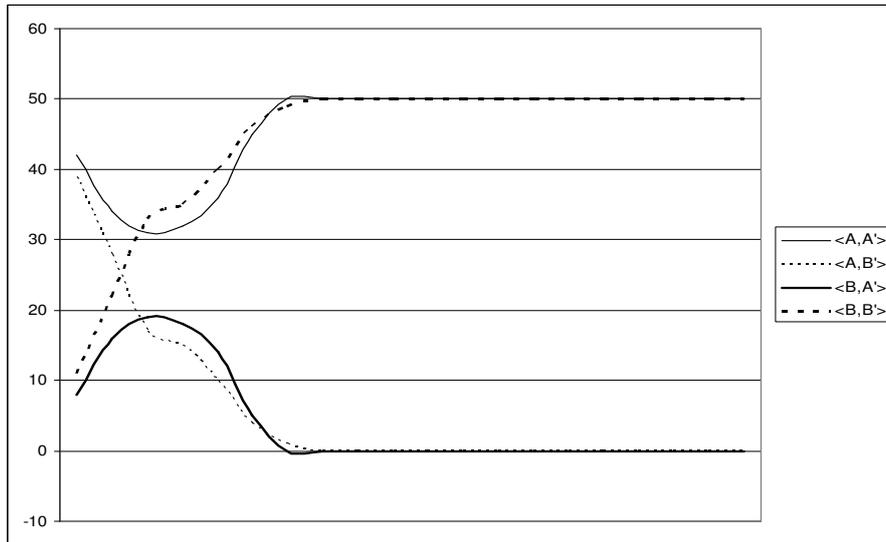
(145)



Even though  $\langle A, A' \rangle$  is dispreferred at some points in the learning process it is in general preferred over  $\langle A, B' \rangle$ . This is due to the difference in frequency in the input corpus. The second reason for the polarization in (145) is that  $\langle B, A' \rangle$  and  $\langle B, B' \rangle$  start to play a role. They too are in constant competition with each other and influence the rankings of  $\langle A, A' \rangle$  and  $\langle A, B' \rangle$ .

The last remaining question is what happens after generation 8. The following graph, representing the frequencies of all form/meaning pairs in the output corpora of the first twenty generations gives the answer:

(146)



Once  $\langle A,A' \rangle$  and  $\langle B,B' \rangle$  have taken over all instantiations from  $\langle A,B' \rangle$  and  $\langle B,A' \rangle$ , the situation stabilizes. The reason for this stabilization is that  $\langle A,A' \rangle$  and  $\langle B,B' \rangle$  do not compete with each other. This can be seen from the possible promotions and demotions in case the learning algorithm is fed with an instantiation of one of them:

(147)	possible promotion	possible demotion
$\langle A,A' \rangle$	$\langle A,A' \rangle$	$\langle B,A' \rangle, \langle A,B' \rangle$
$\langle B,B' \rangle$	$\langle B,B' \rangle$	$\langle B,A' \rangle, \langle A,B' \rangle$

Crucially,  $\langle A,A' \rangle$  cannot be demoted in case an instantiation of  $\langle B,B' \rangle$  is fed to the learning algorithm and  $\langle B,B' \rangle$  cannot be demoted in case an instantiation of  $\langle A,A' \rangle$  is fed into the learning algorithm. This makes the output corpus of generation 8 into a historically stable state.

### 5.3. Summary

In this section I briefly introduced EvolOT and I showed how it predicts that forms that do not appear with a certain meaning in an input corpus can appear in the output corpus. This will prove crucial in the simulation I will present for the evolution of *des N* in section 6.

## 6. Modelling the evolution of *des N*

In this section I simulate the evolution of *des N*. The crucial point I want to make on the basis of this simulation is that there need not have been attested cases of *des N* for it to start marking the indefinite plural in French.

The evolution I want to model is that of a language that – through a phonetic change – lost its ability to morphologically distinguish between singular and plural nouns and resorted to a constructional alternative. I will have two meanings and two forms:

### (148) meanings

‘singular’     $\lambda PP$   
‘plural’       $\lambda P\lambda x(\text{plural}(x)\&P(x))$

### forms

*N*  
*des N*

For the input corpus I assume bare nouns are fully ambiguous and extremely frequent whereas *des N* is unattested:

### (149) input corpus

<i>des N</i>	<i>N</i>	
42	8	‘singular’
39	11	‘plural’

A random simulation in EvolOT gives the following output corpora for the first 10 generations:

### (150)

#### output generation 1

<i>des N</i>	<i>N</i>	
23	77	‘singular’
28	72	‘plural’

#### output generation 2

<i>des N</i>	<i>N</i>	
27	73	‘singular’
47	53	‘plural’

#### output generation 3

<i>des N</i>	<i>N</i>	
28	72	‘singular’
61	39	‘plural’

#### output generation 4

<i>des N</i>	<i>N</i>	
29	71	‘singular’
65	35	‘plural’

**output generation 5**

<i>des N</i>	<i>N</i>	
18	82	'singular'
78	22	'plural'

**output generation 6**

<i>des N</i>	<i>N</i>	
9	91	'singular'
87	13	'plural'

**output generation 7**

<i>des N</i>	<i>N</i>	
4	96	'singular'
91	9	'plural'

**output generation 8**

<i>des N</i>	<i>N</i>	
1	99	'singular'
96	4	'plural'

**output generation 9**

<i>des N</i>	<i>N</i>	
0	100	'singular'
99	1	'plural'

**output generation 10**

<i>des N</i>	<i>N</i>	
0	100	'singular'
100	0	'plural'

What we see happening from the first generation onwards is that the plural meaning of the bare nominal quickly degrades and is being replaced by *des N* that completely takes over by the tenth generation. How this is possible was explained in detail in section 5.

In the remainder of this section I will show how we can further constrain the evolution of *des N* and avoid an unwarranted prediction that the setup I used up till now makes. I illustrate the prediction on the basis of the output corpora of another simulation:

(151)

**output generation 1**

<i>des N</i>	<i>N</i>	
15	85	'singular'
13	87	'plural'

**output generation 2**

<i>des N</i>	<i>N</i>	
31	69	'singular'
31	69	'plural'

**output generation 3**

<i>des N</i>	<i>N</i>	
55	45	'singular'
32	68	'plural'

**output generation 4**

<i>des N</i>	<i>N</i>	
71	29	'singular'
24	76	'plural'

**output generation 5**

<i>des N</i>	<i>N</i>	
78	22	'singular'
18	91	'plural'

**output generation 6**

<i>des N</i>	<i>N</i>	
86	14	'singular'
4	96	'plural'

**output generation 7**

<i>des N</i>	<i>N</i>	
97	3	'singular'
2	98	'plural'

**output generation 8**

<i>des N</i>	<i>N</i>	
99	1	'singular'
0	100	'plural'

**output generation 9**

<i>des N</i>	<i>N</i>	
100	0	'singular'
0	100	'plural'

**output generation 10**

<i>des N</i>	<i>N</i>	
100	0	'singular'
0	100	'plural'

The output corpora in (151) show that *des N* could have taken over the singular: from the ninth generation onwards *des N* unambiguously marks singular indefinites. Given that bare partitives are generally connected to plural interpretations this may seem like a result one would like to exclude. The crucial point I will make in what follows is that this result becomes impossible once we add the following data to the input corpus:

- (i) before the appearance of *des N*, 50% of the singular indefinites was already marked by the indefinite singular article (cf. Carlier 2007);
- (ii) in languages that distinguish between singulars and plural the former are more frequent than the latter;<sup>5</sup>
- (iii) there are some attested examples of *un des N* with a singular indefinite interpretation (see Chapter 1, section 2.4.6.)

The following input corpus reflects this fuller range of data:

**(152) input corpus**

<i>des N</i>	<i>N</i>	<i>un N</i>	<i>un des N</i>	
0	50	50	1	'singular'
0	50	-1	-1	'plural'

(i) is implemented by equally dividing the attested cases of <N, singular> over <N, singular> and <un N, singular>, (ii) is reflected in the reduction of the attested cases of <N, plural> and (iii) is reflected in the addition of one case of <un des N, singular>. '-1' for <un N, plural> and <un des N, plural> indicates that neither of these form/meaning pairs is expected to be compatible with a plural interpretation.

The most frequent outcome of a simulation that takes (152) as an input corpus is illustrated by the set of output corpora in (153):<sup>6</sup>

<sup>5</sup> This conclusion can easily be arrived at by taking a large corpus like the Corpus del Español ([www.corpusdelespanol.org](http://www.corpusdelespanol.org)) which contains on average 130000 singular nouns per million words and only 50000 plural nouns per million words.

(153)

**output generation 1**

<i>des N</i>	<i>N</i>	<i>un N</i>	<i>un des N</i>	
0	20	81	0	'singular'
11	39	-1	-1	'plural'

**output generation 2**

<i>des N</i>	<i>N</i>	<i>un N</i>	<i>un des N</i>	
0	4	96	1	'singular'
23	27	-1	-1	'plural'

**output generation 3**

<i>des N</i>	<i>N</i>	<i>un N</i>	<i>un des N</i>	
0	0	101	0	'singular'
32	18	-1	-1	'plural'

**output generation 4**

<i>des N</i>	<i>N</i>	<i>un N</i>	<i>un des N</i>	
0	0	101	0	'singular'
34	16	-1	-1	'plural'

**output generation 5**

<i>des N</i>	<i>N</i>	<i>un N</i>	<i>un des N</i>	
0	0	101	0	'singular'
34	16	-1	-1	'plural'

**output generation 6**

<i>des N</i>	<i>N</i>	<i>un N</i>	<i>un des N</i>	
0	0	101	0	'singular'
34	16	-1	-1	'plural'

---

<sup>6</sup> Note that the additions of form/meaning pairs also includes the addition of constraints. The full set of constraints is the following: 1:(o==1&i==1); 2:(o==1&i==2); 3:(o==2&i==1); 4:(o==2&i==2); 5:(o==3&i==1); 6:(o==3&i==2); 7:(o==4&i==1); 8: (o==4&i==2).

**output generation 7**

<i>des N</i>	<i>N</i>	<i>un N</i>	<i>un des N</i>	
0	0	101	0	'singular'
42	8	-1	-1	'plural'

**output generation 8**

<i>des N</i>	<i>N</i>	<i>un N</i>	<i>un des N</i>	
0	0	101	0	'singular'
46	4	-1	-1	'plural'

**output generation 9**

<i>des N</i>	<i>N</i>	<i>un N</i>	<i>un des N</i>	
0	0	101	0	'singular'
48	2	-1	-1	'plural'

**output generation 10**

<i>des N</i>	<i>N</i>	<i>un N</i>	<i>un des N</i>	
0	0	101	0	'singular'
46	4	-1	-1	'plural'

**output generation 11**

<i>des N</i>	<i>N</i>	<i>un N</i>	<i>un des N</i>	
0	0	101	0	'singular'
50	0	-1	-1	'plural'

The result of the simulation is that *des N* takes over the plural interpretation of the bare nominal from the eleventh generation onwards and that *un N* takes over the singular interpretation. The crucial point to be made is that out of the 20 simulations I ran in EvolOT, 6 had exactly the same end result as the simulation in (153) and none of them gave rise to *des N* marking singular indefinites. The typical variations in the simulations are represented by the following end results (after 100 generations):

**(154) output generation 100**

<i>des N</i>	<i>N</i>	<i>un N</i>	<i>un des N</i>	
0	0	100	1	'singular'
50	0	-1	-1	'plural'

## (155) output generation 100

<i>des N</i>	<i>N</i>	<i>un N</i>	<i>un des N</i>	
0	0	101	0	‘singular’
0	50	-1	-1	‘plural’

## (156) output generation 100

<i>des N</i>	<i>N</i>	<i>un N</i>	<i>un des N</i>	
0	4	97	0	‘singular’
50	0	-1	-1	‘plural’

In (154) *un des N* kept its singular indefinite reading, in (155) *des N* is blocked by the bare nominal and singularity is marked by *un N* and in (156) singular indefinites are marked both by the bare nominal and *un N*.

## 7. Conclusion

In this chapter I worked out the competition between the bare plural and *des N*. I argued that the bare plural blocked the kind use of *des N* up till the 14<sup>th</sup> century and then disappeared through phonetic erosion. The result of its disappearance was a gap in the singular/plural paradigm that was filled up by *des N*.

The main conclusion I draw on the basis of the analysis and simulation I proposed is that it is not necessary to assume that *des* became a plural indefinite article to explain its distribution in the different stages of the evolution of French. This is important given that this assumption played an important role in previous analyses like the one proposed by Carlier (2007) and is incompatible with the analysis I proposed for present-day *des N*.

### ***General conclusion***

In this part of my dissertation I did two things. On the one hand, I presented a fully worked-out semantics/pragmatics for bare partitives and opposed them to DPs headed by indefinite articles. On the other hand, I argued that even though *des N* cannot be analyzed as a full-fledged bare partitive, most of its properties can be explained on the basis of a bare partitive analysis.

I argued that bare partitives are nothing more than partitives without an upstairs determiner and are opposed to DPs headed by indefinite articles on the basis of their narrow scope behaviour. As for *des N*, I argued that except for its marginal wide scope behaviour, all of its synchronic and diachronic properties could be accounted for on the basis of a bare partitive analysis.



## AN ARTICLE GONE ASTRAY

The semantics of the indefinite article in predicate position

### *General Introduction*

In Parts I and II of this dissertation I defined the indefinite article as the determiner that realizes the existential type-shift. This definition is equipped to account for the appearance of the indefinite article in argument position but raises the question what role it plays in predicate position:

(1) John is a man.

The appearance of the indefinite article in (1) is not problematic in itself. Indeed, given that the BE type-shift is defined for DPs headed by the indefinite article nothing prevents these DPs from appearing in predicate position. The real problem is that the indefinite article is (in general) not optional:

(2) \*John is man.

(2) forms a minimal pair with (1) and shows that the indefinite article cannot be omitted, a fact that does not follow from the analysis of the indefinite article as the realization of the existential type-shift. The real question to be answered then is not why the indefinite article *can* appear in predicate position but why it (in general) *must*.

In order to investigate the presence of the indefinite article in predicate position I propose to compare the cases of predication in which it appears to the few ones in which it does not. The latter cases are generally referred to as *bare predication* and are exemplified in (3) and (4):

(3) Mary is chairwoman of the board of directors.

(4) John is CEO of Haver Foods.

My investigation into the semantics of the indefinite article in predicate position starts from Dutch that I treat as a case-study in Chapter 1. Taking Dutch as a starting point is motivated by the fact that the Dutch bare predication paradigm stands midway between the more restricted English one and the less restricted French and Italian ones that were intensively studied in the literature. The analysis I will propose for the indefinite article in predicate position in Dutch is closely related to the one proposed by de Swart, Winter & Zwarts (2007) even though there are some crucial differences that will be spelled out in detail. In Chapter 2 I present possible extensions of the analysis, including a semantics for the indefinite article in predicate position in English. I also present the limits of the analysis and compare it to other proposals that have been made in the recent literature.



## Chapter 1: Article and Bare predication: A Dutch Case-Study

### 1. Introduction

In this chapter I analyze the predication paradigm in Dutch focusing on the distinction between predication involving the indefinite article (henceforth *article predication*) and predication lacking an article (henceforth *bare predication*). The ultimate goal is to come to a semantics of the indefinite article in predicate position. In section 2 I introduce the notions *capacity* and *kind* and argue that they are crucial for a correct analysis of the Dutch predication facts. In section 3 I present the analysis of de Swart, Winter & Zwarts (2007) and point at the parts I will improve in sections 4 and 5. Section 6 summarizes.

### 2. Articles, bareness, capacities and kinds

#### 2.1. Bare predication and capacities

In this subsection I introduce the notion of *capacity* and argue that it is crucial for the interpretation of bare predication in Dutch (see also de Swart, Winter & Zwarts 2007).

#### *Capacities*

*Capacity* is a cover term for professions, nationalities and religions. The link licensing this cover term is that professions, nationalities and religions all stand for the position of an individual in society, be it at a professional, civil or religious level. They all specify what an individual stands for: a doctor is a professional we can turn to in case of illness, an American is a citizen of the US that has to defend his country and a muslim is an individual that believes in Allah and supports his fellow believers. These intuitions concerning the position of doctors, Americans and muslims are reflected in (5) through (7):

- (5) As a **doctor** I treat anyone who needs medical assistance.
- (6) As an **American**, I cannot stand silent as a truly evil man destroys the country I love.
- (7) As a **muslim** I must support my muslim brothers and sisters.

Crucial for capacities is that they are culturally defined. This has two consequences. The first is that there is no inherent property of an individual that makes it into a doctor, American or muslim. Indeed, any person could in principle have any of these capacities but could also have any other combination. The second consequence is that capacities exist by virtue of a cultural decision and need not reflect the position of any existing individual. One could e.g. establish the capacity *king of the US* independently of anyone ever holding this position.

Next to the broad categories of professions, nationalities and religions the notion *capacity* also covers smaller categories that are defined on smaller social settings. I give two examples. In the context of a shop *customer* can be thought of as a capacity: the customer is the person that buys goods. In the context of a game like the *werewolf game* one can also consider *wolf* a capacity: it stands for the position of a player in the game.

#### *Bare predication and capacities*

In order to argue that bare predication is linked to capacity predication I show that nouns referring to capacities appear in bare predication and that those that do not refer to capacities can only appear in bare predication if they undergo a capacity reinterpretation.

In (8) through (10) I show that nouns corresponding to the main categories covered by the term *capacity* can occur in bare predication:

- (8) Jan is dokter.  
Jan is doctor
- (9) Jan is Amerikaan.  
Jan is American
- (10) Jan is moslim.  
Jan is muslim

(8) through (10) show that professions, nationalities and religions felicitously occur in bare predication. The standard interpretation of (8) through (10) is that Jan works as a doctor, is an American citizen and believes in Allah.

On the basis of (11) I show that non-capacity nouns get a capacity interpretation when they occur in bare predication. The nouns I use are those that are generally considered not to be capacity nouns: those referring to kinds like *rat* and *raven* (see also 2.2.):

- (11) Eén kind van het tweetal is rat, de ander raaf.  
One child of the couple is rat, the other raven  
<http://www.groep3.mysites.nl/mypages/groep3/262230.html>  
[03/12/2009]

Crucially, the only felicitous way of interpreting (11) is to assume *rat* and *raven* refer to the positions of players in the game *rats and ravens*. This shows that a non-capacity noun gets a capacity reinterpretation when it occurs in bare predication.

In the preceding I have shown that any noun receives a capacity interpretation when it occurs in bare predication, independently of whether it qualifies as a capacity or not. This is strong support in favour of the assumption that bare and capacity predication are linked.

*Summary*

In this subsection I have defined the notion of *capacity* as referring to the position of an individual in society and I have argued that bare and capacity predication are closely intertwined. In the next subsection I define the notion of *kind* as the opposite of *capacity* and argue that kind predication is to be linked to article predication.

**2.2. Article predication and kinds**

In this subsection I oppose the notion of kind to that of capacity and argue that the former is crucial for the interpretation of article predication in Dutch (see also de Swart, Winter & Zwarts 2007).

*Kinds*

In 2.1. I pointed out that capacities are culturally defined. Kinds are exactly the opposite: they form natural classes of individuals. The term *natural* is opposed to *cultural* and reflects the intuition that kinds are different from capacities in two respects. The first is that kind membership is based on inherent properties: a wolf belongs to the kind *wolf* not by accident but because it has inherent properties that make it into a wolf. The second respect in which kinds are natural and capacities cultural is that kinds cannot be established by virtue of a cultural decision but exist by virtue of there being individuals that share the same inherent properties.

The cultural / natural distinction I established between capacities and kinds has two important consequences to which I will return later on. The first is that capacity-membership is less stable than kind-membership: a teacher can easily turn into a principal but a human cannot easily turn into a frog. This is related to the fact that kind-membership is based on inherent properties whereas capacity-membership is not. Apparent exceptions involve wizardry (a prince turning into a frog) and genetically programmed evolution (a caterpillar turning into a butterfly). The second consequence of the cultural / natural distinction is that a kind can only be defined if it has at least two instantiations whereas there is no such constraint on capacities. This is related to the fact that capacities can be established by virtue of a cultural decision whereas kinds exist by virtue of there being individuals that share the same inherent properties.

*Article predication and kinds*

In order to argue that article and kind-membership predication are linked I show that nouns referring to kinds appear in article predication and that those nouns that do not refer to kinds can only appear in article predication if they get a kind reinterpretation. In (12) through (14) I show that kind nouns occur in article predication:

- (12) Zoë is een wolf.  
Zoe is a wolf

- (13) Fifi is een hond.  
Fifi is a dog
- (14) Felix is een kat.  
Felix is a cat

(12) through (14) contain typical kind predicates and are standardly interpreted as stating that the subjects are members of the kind referred to by the predicate. On the basis of (15) I argue that non-kind nouns that appear in article predication get a kind reinterpretation:

- (15) Henriette is een manager.  
Henriette is a manager

The preferred interpretation of (15) is not one in which Henriette has been appointed manager but rather that she has the inherent properties we generally associate with managers: being organized, able to delegate, ... This reinterpretation can be seen as a kind reinterpretation: we look for inherent properties we associate with managers and predicate those of the subject.

In the preceding I have shown that nouns occurring in article predication get a kind interpretation whether they qualify as kind nouns or not. I take this to be strong support for the assumption that kind and article predication are linked.

### *Summary*

In the first part of this subsection I opposed kinds to capacities and defined the former as natural classes of individuals. The term *natural* captures the intuition that kind-membership is based on inherent properties and that kinds exist by virtue of there being individuals that share the same inherent properties. In the second part of this subsection I argued that kind and article predication are closely intertwined.

### **2.3. Conclusion**

The goal of this section was twofold. On the one hand I introduced the notions *capacity* and *kind*, the former being related to culture, the latter to nature. On the other hand I argued that capacity and kind predication are expressed in Dutch in two different ways: capacity predication is expressed through bare predication and kind predication is expressed through article predication. The crucial arguments in favour of this division of labour between bare and article predication are that kinds are reinterpreted as capacities in bare predication and that capacities are reinterpreted as kinds in article predication. In the next section I present the analysis of de Swart, Winter & Zwarts (2007) that derives these facts and point at the parts I will improve upon in sections 4 and 5.

### 3. De Swart, Winter & Zwarts (2007)

In this section I do two things. The first is to introduce the analysis de Swart, Winter & Zwarts (2007) proposed to account for the Dutch predication facts (see 3.1.). The second is to point at the parts of the analysis I will improve on in sections 4 and 5 (see 3.2.).

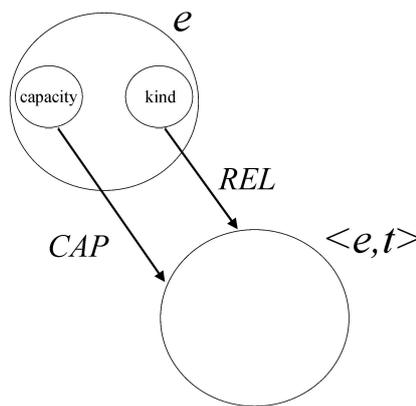
#### 3.1. The analysis

In this subsection I gradually build up the analysis of de Swart, Winter & Zwarts. First, I show how their analysis treats the distinction between kinds and capacities. Second, I focus on how they derive the links between article and kind predication and bare and capacity predication. Third, I zoom in on how they treat the capacity reinterpretation of kinds and the kind reinterpretation of capacities we saw in section 2.

##### *Kinds and capacities*

De Swart, Winter & Zwarts assume nouns are lexically stored as expressions of type  $e$  and can be divided into two subtypes: kinds and capacities. Given that kinds and capacities are of type  $e$  they have to undergo a type-shift to type  $\langle e,t \rangle$  in order to appear in predicate position. Kinds do this via *REL* and capacities via *CAP*. *REL* takes a kind and returns the set with its instantiations whereas *CAP* takes a capacity and returns the set with the individuals that have this capacity. This gives us the following (preliminary) sketch of what is going on in predicate position:

(16)



In the graph in (16) the big circles represent the  $e$  and  $\langle e,t \rangle$  domains and the smaller circles within the  $e$  domain represent the  $e$ -subtypes of capacities and kinds. The arrows stand for the type-shifts *CAP* and *REL*.

*The links between article and kind predication and bare and capacity predication*

The analysis as it stands now, gives us a lexical distinction but does not derive the fact that article and kind predication are linked nor the fact that bare and capacity predication are linked. De Swart, Winter & Zwarts derive these facts through two assumptions. The first is that NumP – and by extension the indefinite article – is the realization of *REL*. In type-shifting theory this means *REL* can only take place in the presence of NumP and that NumP can only appear if *REL* is applied. This creates a one-to-one correspondence between article and kind predication. The second assumption de Swart, Winter & Zwarts make is that *CAP* has no overt realization which means it can apply freely. Given that capacities are the complement of kinds this second assumption suffices to create a one-to-one correspondence between bare and capacity predication.

The analysis can now account for the fact that (17) has to be interpreted as stating that Zoë belongs to the kind of wolf and that (18) has to be interpreted as stating that Henriette is an appointed manager:

- (17) Zoë is een wolf.  
       Zoë is a wolf  
 (18) Henriette is manager.  
       Henriette is manager  
       Henriette is a manager.

Under the assumption that *wolf* is a kind noun, *een* in (17) takes the kind *wolf* and turns it into the set of its instantiations. The resulting semantics is:

- (19)  $z \in REL(wolf_k)$

In words: Zoë belongs to the set comprising all the individuals realizing the kind *wolf*.

Under the assumption that *manager* is a capacity noun (18) contains a covert application of the *CAP* shift. The resulting semantics of is:

- (20)  $h \in CAP(manager_c)$

In words: Henriette belongs to the set comprising the individuals that have the position of manager.

*Kind and capacity reinterpretations*

The analysis as it stands now, gives us a lexical distinction and derives the links between article and kind predication and bare and capacity predication. It is still underspecified though for the interpretation of examples such as (21) and (22):

- (21) Henriette is een manager.  
Henriette is a manager
- (22) Eén kind van het tweetal is rat, de ander raaf.  
One child of the couple is rat, the other raven

In (21) a capacity noun occurs in article predication and in (22) a kind noun occurs in bare predication. Even though we know from section 2 that these examples involve reinterpretation the analysis is not yet equipped to account for them. In order to accommodate (21) and (22) de Swart, Winter & Zwarts assume there is an operator coercing capacities into kinds in the presence of NumP and an operator coercing kinds into capacities in the absence of NumP. The former – called *kind* – was explicitly introduced in their paper, the latter – called *cap* – can be seen as a natural extension I add to their story. In section 3.2. I will get back to the exact semantics of these operators but for now it suffices to assume they exist to accommodate examples like (21) and (22). The interpretation of (21) on the proposed analysis would be the following:

- (23)  $h \in REL(kind(manager_c))$

In words: Henriette is a member of the set of individuals instantiating the kind obtained after kind coercion of the capacity *manager*. The interpretation of (22) in the proposed analysis would be the following:

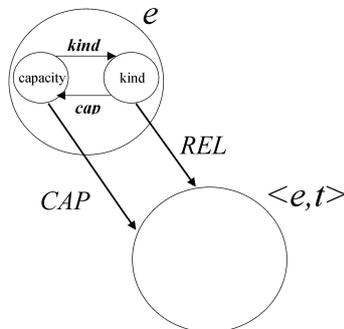
- (24)  $z \in CAP(cap(wolf_k))$

In words: Zoë is a member of the set of individuals performing the capacity obtained after cap coercion of the kind *wolf*.

### Summary

All elements of the analysis de Swart, Winter & Zwarts propose for Dutch predication are now in place. They are schematically represented in (25) that updates the earlier graph in (16):

- (25)



The two additions with respect to (16) are the horizontal arrows between capacities and kinds. They represent the coercion operators *kind* and *cap*.

### 3.2. Discussion

I assume the general setup of the analysis presented by de Swart, Winter & Zwarts is correct and fully equipped to account for the facts discussed in section 2. What I want to do in this subsection is to point at two parts of the analysis I will improve on in sections 4 and 5: the semantics of the operators *kind* and *cap* and the link between NumP and *REL*.

#### *The semantics of kind and cap*

In section 2 I discussed the interpretation of (26) and (27):

- (26)   Henriette is een manager.  
           Henriette is a manager
- (27)   Eén kind van het tweetal is rat, de ander raaf.  
           One child of the couple is rat, the other raven

(26) was said to be interpreted as stating that Henriette has the inherent properties we typically associate with managers and (27) was said to be interpreted as stating that one child plays the role of rat and the other the role of raven. These interpretations are very specific and should follow from the coercion of the capacity *manager* into the kind *manager* and from the coercion of the kinds *rat* and *raven* into the capacities *rat* and *raven*. Put differently: these interpretations should be obtained through the coercion operators *kind* and *cap*.

De Swart, Winter & Zwarts did not treat *cap* so I will leave it aside for the moment. As for *kind*, they give an indirect semantics based on the relation between the set obtained through the application of *CAP* to a capacity and the set obtained through the subsequent application of *kind* and *REL* to the same capacity:

- (28)   For all C, C being a capacity the following holds:  
            $CAP(C) \subseteq REL(kind(C))$   
           In words: the set obtained after application of *CAP* to a capacity is a subset of the set obtained after applying *kind* and *REL* to the same capacity

For a capacity like *manager* (28) states that the set obtained through the application of *CAP* is a subset of the set obtained through the subsequent application of *kind* and *REL*. In section 4 I will not improve on the indirectness of the semantics in (28); this would require a full ontology of capacities which is beyond the scope of this dissertation. What I will improve on is the concreteness and correctness of the proposal. (28) is inconcrete in the sense that it leaves a lot of room for interpretation, putting no upper bound on  $REL(kind(C))$ . As it stands now,  $REL(kind(C))$  could indeed be any superset of  $CAP(C)$  and (26) would consequently be trivially true. (28) is incorrect in the sense that it predicts (29) to contain a contradiction:

- (29) Jan is misschien wel manager maar hij is allesbehalve een manager.  
 Jan is maybe well manager but is he everything\_ but a manager  
 Jan may be a professional manager but he's everything but a manager.

Given that de Swart, Winter & Zwarts assume that the set corresponding to the kind *manager* is a superset of the set corresponding to the capacity *manager*, it is impossible for Jan to belong to the latter but not to the former. (29) should therefore contain a contradiction. This is however not the case.

In section 4 I will propose an indirect semantics for *kind* that puts an upper bound on  $REL(kind(C))$  and allows it not to be a superset of  $CAP(C)$ . I will furthermore present the beginning of an indirect semantics of *cap*.

#### *The link between NumP and REL*

The link between article and kind predication in the analysis of de Swart, Winter & Zwarts crucially depends on their assumption that NumP is the realization of *REL*. Even though this assumption derives the facts in a straightforward way one might wonder where it comes from: what makes NumP special and what links it to *REL* rather than to *CAP*? As far as I can see the link between NumP and *REL* is nothing more than a stipulation. The link between the indefinite article and *REL* can be motivated though. Linking the indefinite article to *REL* is the goal of section 5.

### **3.3. Summary and outlook**

In this section I introduced the analysis de Swart, Winter & Zwarts proposed for the Dutch predication facts discussed in section 2. I also pointed at the parts of the analysis I will improve on in the following sections. In section 4 I will present a semantics for *kind* and *cap* and in section 5 I will argue that there is a natural link between article and kind predication.

## **4. The coercion operators *kind* and *cap***

In this section I present an indirect semantics for the coercion operators *kind* and *cap*. In 4.1. I zoom in on *kind* and in 4.2. on *cap*.

### **4.1. The coercion operator *kind***

In section 3 I presented the indirect semantics de Swart, Winter & Zwarts propose for *kind*:

- (30) For all C, C being a capacity the following holds:  
 $CAP(C) \subseteq REL(kind(C))$   
 In words: the set obtained after application of *CAP* to a capacity is a subset of the set obtained after applying *kind* and *REL* to the same capacity

I furthermore pointed out that this semantics presents two problems. The first is that it is inconcrete: given that  $REL(kind(C))$  is only defined as a superset of  $CAP(C)$  sentences such as (31) would be trivially true:

- (31) Henriette is een manager.  
Henriette is a manager

The second problem I discussed was the incorrectness of (30): it makes the wrong prediction that a sentence like (32) would be contradictory:

- (32) Jan is misschien wel manager maar hij is allesbehalve een manager.  
Jan is maybe well manager but is he everything\_but a manager  
Jan may be a professional manager but he's everything but a manager.

Given that  $CAP(C)$  is a subset of  $REL(kind(C))$  it is impossible for Jan to belong to the latter but not to the former. (32) should therefore contain a contradiction. This is however not the case.

In what follows I work out a concrete semantics of *kind* on the basis of the interpretation of the predicate in (31) and the interpretation of its bare counterpart in (33). Interestingly this will not only solve the concreteness but also the correctness problem.

- (33) Henriette is manager.  
Henriette is manager

The predicate in (33) denotes the set of individuals that work as managers whereas the predicate in (31) denotes the set of individuals that have the inherent properties we typically associate with managers: being organized, being able to delegate, ... The semantics of the predicate in (33) is represented in (34.a), the semantics of the predicate in (32) in (34.b):

- (34) a.  $\lambda x(\text{professional\_manager}(x))$   
b.  $\lambda x \forall P(\forall y(\text{professional\_manager}(y) \rightarrow P(y)) \ \& \ \text{Inherent}(P) \rightarrow P(x))$

The set in (34.a) only contains professional managers. The set in (34.b) contains those individuals that have all the inherent properties we typically associate with professional managers. Note that “ $\rightarrow$ ” stands for “if...then normally” and is borrowed from the literature on non-monotonic inferencing (see e.g. Lascarides & Asher 1993). What *kind* seems to do then is to take a capacity and return the kind consisting of all the individuals that have the inherent properties that are typically associated with the individuals that have this capacity.

This semantics overcomes the problems posed by the one in (30). It correctly predicts (31) only to be felicitous if Henriette has all the inherent characteristics we typically associate with managers. It furthermore allows the set corresponding to the capacity *manager* not to be a subset of the set corresponding to the kind *manager*.

This is implemented in two ways. The first is that (34.b) only looks at inherent properties and not at accidental ones like being appointed as a manager. The second is that (34.b) is about properties that we merely expect professional managers to have; nothing prevents professional managers from not having them. On this semantics (32) comes out naturally: Jan is professionally employed as a manager but lacks the inherent properties we generally associate with managers.

#### 4.2. The coercion operator *cap*

In section 3 I pointed out that I assume the coercion operator *cap* to be at play in the interpretation of sentences such as (35):

- (35) Eén kind van het tweetal is rat, de ander raaf.  
One child of the couple is rat, the other raven

*Cap* takes the kinds *rat* and *raven* and coerces them into capacities. The question I want to answer here is what the semantics of this operator would look like. To get a preliminary feeling of what *cap* does I want to compare the interpretation of the predicate *rat* in (35) to that of *een rat* in (36):

- (36) Minnie is een rat.  
Minnie is a rat

Intuitively there need not be any link between the set denoted by *rat* and the set denoted by *een rat* beyond the fact that both are referred to with the noun *rat*: a child playing the role of rat in a game need not have any properties of a rat and rats in general are not expected to play the role of rat in a game. This is not surprising: capacities exist by virtue of a cultural decision and the name they carry is in principle arbitrary.

In (37) I formalize the above intuitions: (37.a) represents the set corresponding to *een rat* and (37.b) represents the set corresponding to *rat*:

- (37) a.  $\lambda x((\text{REL}(\text{rat}))(x))$   
b.  $\lambda x \exists C(\text{Capacity}(C) \& \text{referred\_to\_by\_the\_word\_rat}(C) \& C(x))$

What *cap* seems to do then is to take a kind and return a capacity with the same name. I of course do not exclude that capacities that were obtained via *cap* have some link with the kind they share their name with. What I am saying is that this is left open because of the arbitrariness of capacities.

#### 4.3. Summary

In this section I worked out the semantics of *kind* and *cap*. I defined *kind* as an operator that takes a capacity and returns the kind consisting of the individuals that have the inherent properties that are typically associated with the individuals that

have this capacity. I defined *cap* as an operator that takes a kind and returns a capacity with the same name. This concludes my first addition to the analysis of de Swart, Winter & Zwarts. In the following section I present the second one: a worked-out link between article and kind predication.

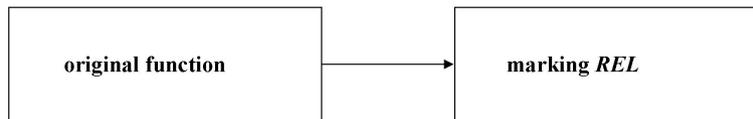
## 5. The link between article and kind predication

As I pointed out in section 3 the link between article and kind predication in the analysis of de Swart, Winter & Zwarts crucially depends on their assumption that NumP is the realization of *REL*. Even though this assumption derives the facts in a straightforward way one might wonder where it comes from: what makes NumP special and what links it to *REL* rather than to *CAP*? In section 3 I already pointed out that I do not see a way to make this link more insightful. What I argue for in this section is that there is a natural link between *REL* and one of the instantiations of NumP, *viz.* the indefinite article. This leads me to assume that it is not NumP that realizes *REL* but the indefinite article.

### 5.1. The gist of my proposal

In order to argue that there is a natural link between the indefinite article and *REL* I will argue that the indefinite article was the only expression that could have undergone the historical process of acquiring the marking of *REL* as a new function. This process is depicted in (38)

(38)



The box on the left represents the start of the historical process, the box on the right the end.

Historical processes like the one in (38) are interesting because they are constrained. Zeevat (2005) e.g. argues that it is impossible for an item to acquire a new function if its old function does not weakly entail the new function. I propose to exploit this and other constraints to argue that (i) the indefinite article was the only candidate that could have started marking *REL* for singular predicates and that (ii) it would be impossible for it to mark *CAP*. The argumentation is worked out in detail in 5.3. Before going there I however have to define the original function of the indefinite article in predicate position. I do this in 5.2.

### 5.2. The original function of the indefinite article in predicate position

In order to find out what the function of the indefinite article was before it started marking *REL* there are two options: the first is to look at corpora, the second to make a theory-based reconstruction. In this particular case I take the second option.

This choice is motivated by the fact that the indefinite article has a very weak semantics which makes it hard to come to any insightful conclusions on the basis of corpora.<sup>1</sup>

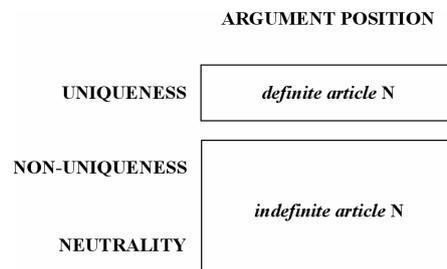
This subsection is organized as follows. I start by presenting the theoretical basis for my reconstruction. I then proceed to the reconstruction itself and I end with a short summary of the outcome.

#### *The basis of the reconstruction*

As the basis of my reconstruction I propose to take the well-studied function of the indefinite article in argument position. What we know about the indefinite article in argument position is that it introduces discourse referents that are non-familiar by implicature and that it carries a non-uniqueness implicature (Hawkins 1991, Farkas 2006). Given that predicates in general do not introduce discourse referents the only relevant part for my reconstruction is the non-uniqueness implicature. In what follows I will therefore ignore the discourse referential part.

The standard way of deriving the non-uniqueness implicature is to analyze the indefinite article as being neutral with respect to uniqueness and – in comparison to the definite article – as the least bad option for a speaker who wants to convey non-uniqueness. This division of labour is depicted in (39):

(39)



Under the assumption that hearers always go for the strongest interpretation except if they have evidence to the contrary the division of labour in (39) tells us that hearers will interpret the indefinite article in argument position as marking non-uniqueness except if they have reason to believe otherwise. This is exactly what it means to have a non-uniqueness implicature.

#### *The reconstruction*

Now that we have a clear idea of the function of the indefinite article in argument position we can reconstruct its original function in predicate position. The crucial difference between arguments and predicates is that there is no type requirement for

<sup>1</sup> Note furthermore that I have no knowledge of there being historical studies on the role of the indefinite article in predicate position in Dutch.

predicates that forces the appearance of an article to make a shift from type  $\langle e, t \rangle$  to type  $\langle \langle e, t \rangle, t \rangle$  or type  $e$ . As a consequence the two options that are available in argument position to mark uniqueness, non-uniqueness and neutrality are enriched with a third one. Indeed, in predicate position we not only have at our disposal DPs headed by the definite and the indefinite article but also bare NPs. As for the division of labour it seems plausible to assume that the definite article marks what it marks in argument position, *viz.* uniqueness. The question that imposes itself then is how the indefinite article and the bare form divide the two meanings that are not expressed by the definite article. Assuming that marked forms go with marked meanings and assuming the markedness scales in (40) we arrive at the conclusion that the indefinite article encodes non-uniqueness and the bare form neutrality with respect to uniqueness.

- (40) markedness in form:  
*indefinite article* N > bare form  
 (criterion : syntactic material)  
 markedness in meaning:  
 non-uniqueness > neutrality  
 (criterion : implication)

The analysis is formalized in a weak version of bi-directional Optimality Theory in (41):

- (41) forms: N  
*indefinite article* N  
 meanings:  $\lambda x(P(x))$  (= neutrality with respect to uniqueness)  
 $\lambda x(P(x) \& \exists y(x \neq y \& P(y)))$  (= non-uniqueness)
- |                | N |   | <i>indefinite article</i> N |
|----------------|---|---|-----------------------------|
| neutrality     | ☺ | ← | ●                           |
| non-uniqueness | ● | ← | ☺                           |

In this reconstruction the non-uniqueness marking of the indefinite article in predicate position is not seen as something inherent but rather as the result of a specific division of labour. In optimal communication the analysis however predicts a one-to-one correspondence between the indefinite article and non-uniqueness and would therefore be equivalent to an analysis that assumes the indefinite article marks non-uniqueness. This is reflected in the graph in (42):

(42)

<b>PREDICATE POSITION</b>	
<b>UNIQUENESS</b>	<i>definite article N</i>
<b>NON-UNIQUENESS</b>	<i>indefinite article N</i>
<b>NEUTRALITY</b>	<i>bare singular</i>

*Summary of the outcome*

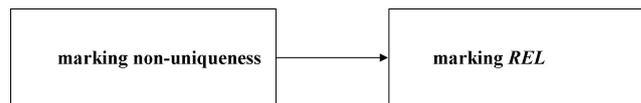
The outcome of the reconstruction of the original function of the indefinite article in predicate position is that it was used to mark non-uniqueness. I arrived at this outcome on the basis of standard pragmatic reasoning.

In the next subsection I will argue that the fact that the indefinite article originally marked non-uniqueness made it into the perfect candidate to mark *REL* for singular predicates and prevented it from marking *CAP*.

**5.3. Non-uniqueness, *REL* and *CAP***

In 5.1. I hypothesized that marking *REL* was not the original function of the indefinite article. Under this assumption there must have been a historical process through which the indefinite article acquired the marking of *REL* as a new function. With the outcome of the reconstruction in 5.2. this historical process can be schematically represented as in (43):

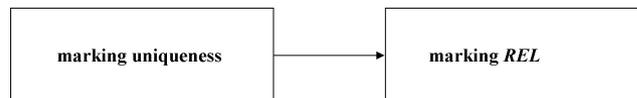
(43)



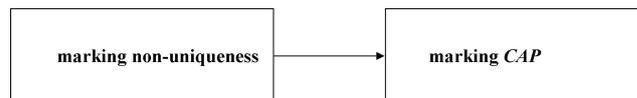
In this subsection I will reconstruct the process in (43) and argue that its most likely variants could not have taken place:

(44)

a.



b.



c.



The variant in (44.a) would have linked the definite article to *REL*, the variant in (44.b) the indefinite article to *CAP* and the variant in (44.c) the definite article to *CAP*.

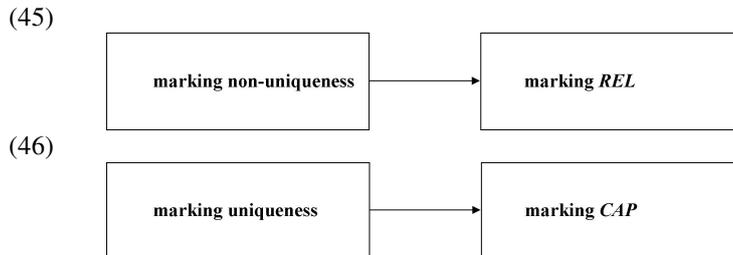
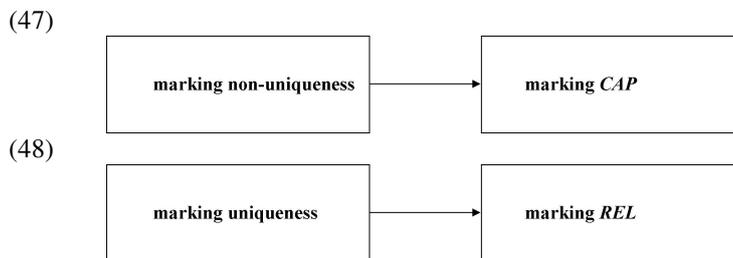
#### *Linking the indefinite article to REL*

The driving force behind many historical processes is the desire of speakers to formally mark a distinction. I assume this is also the case for the historical process leading the indefinite article to mark *REL*. The distinction in this case is the one between *CAP* and *REL*.

In order to formally marking the distinction between *CAP* and *REL* only one of them needs to be marked. Indeed, if one of them is explicitly marked the other can be inferred by default reasoning. This leads me to assume that marking the distinction could have led to two situations: one in which *REL* is explicitly marked and *CAP* is not and one in which *CAP* is explicitly marked and *REL* is not. With this assumption in place I can turn to the question which candidates could potentially have competed to mark the distinction. In principle the answer consists of a list containing all lexical items but I will narrow it down with three filters and end up with only one candidate: the indefinite article. I will furthermore argue that the indefinite article could only end up marking *REL*.

The first filter I use is one that is based on the fact that language evolution in general does not involve drastic changes. As a consequence I expect items that already co-occur with nouns in predicate position to be the most plausible candidates. These items are the indefinite and the definite article.

The second filter I use is based on the observation by Zeevat (2005) that it is impossible for an item to acquire a new function if its old function does not weakly entail the new function. Weak entailment of B by A is defined as A making B more likely than the negation of B. In the case of marking *CAP* or *REL* this means the old function of the indefinite or the definite article should make it more plausible that *CAP* was applied than that *REL* was applied or vice versa. To evaluate the effect of this second filter I have to go back to the distinction between capacities and kinds. In section 2 I pointed out that kinds are opposed to capacities in requiring their corresponding sets to contain at least two elements. This means that the elements of kind-sets necessarily obey non-uniqueness whereas this is not the case for the elements of capacity-sets. The consequence of this is that the indefinite article – as a marker of non-uniqueness – makes it more plausible that *REL* was applied and that the definite article – as a marker of uniqueness – makes it more plausible that *CAP* was applied. Even though this second filter does not eliminate any candidates from our list it does tell us that the processes in (45) and (46) are possible whereas those in (47) and (48) are not:

**possible processes****impossible processes**

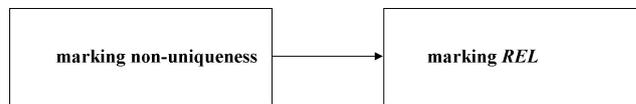
The processes in (45) and (46) would respectively lead the indefinite article to mark *REL* and the definite article to mark *CAP* whereas those in (47) and (48) would respectively lead the indefinite article to mark *CAP* and the definite article to mark *REL*.

The third and last filter I use is based on the observation by Zeevat (2005) that it is impossible for an item to acquire a new function if this new function does not give it a wider application domain than its old function. In order for the indefinite article to start marking *REL* this would require the application of *REL* to be more frequent than the explicit marking of non-uniqueness. In order for the definite article to start marking *CAP* this would require the application of *CAP* to be more frequent than the explicit marking of uniqueness. Present-day corpus data suggest that the requirement for the definite article to start marking *CAP* is not met.<sup>2</sup> Crucially *CAP* is unmarked in present-day Dutch and covers only 9 % of all cases of nominal predication whereas the definite article marks uniqueness and covers 28 %. This shows that marking *CAP* would not have extended the application domain of the definite article and that the definite article could not have started marking *CAP*. The requirement for the indefinite article to start marking *REL* is probably met though. In present-day

<sup>2</sup> The corpus data that will be referred to were established on the basis of a corpus in the Eindhoven Corpus (freely available online via [www.inl.nl](http://www.inl.nl)). The strings I looked for were “is\_[0-9]\* een\_[0-9]\* [a-z]\*\_000”, “is\_[0-9]\* de\_[0-9]\* [a-z]\*\_000” and “is\_[0-9]\* [a-z]\*\_000”. These strings gave me all attested examples of *is a N*, *is the N* and *is N*. I manually filtered out the predicates and ended up with 209 cases of indefinite article predication, 92 cases of definite article predication and 29 cases of bare predication.

Dutch *REL* is marked by the indefinite article and covers 63 % of all cases of nominal predication. Under the assumption that the indefinite article used to be infrequent in predicate position this shows that marking *REL* would (and did) extend the application domain of the indefinite article. The result of this final filter is that there is only one lexical item left of the two that could in principle have marked the distinction between *CAP* and *REL*: the indefinite article. This means that the only process that could have taken place to mark the distinction between *CAP* and *REL* is the one in (49):

(49)



This process is the one linking the indefinite article to *REL*.

Note for completeness that the simulations in Zeevat (2005) show that a recruited item can either keep or lose its original function. Given that I have no reason to believe the indefinite article still marks non-uniqueness in predicate position, I assume it lost this function.

#### *Summary*

In this subsection I reconstructed the process linking the indefinite article to *REL*. Crucial for this process were the indefinite article's original function in predicate position, the fact that kinds require their corresponding sets to have at least two elements and the assumption that the application of *REL* was more frequent than the explicit marking of non-uniqueness. I furthermore argued that the distinction between *CAP* and *REL* could not have been marked by any other lexical item. Most items were excluded on the basis of the assumption that only items that already co-occurred with nouns in predicate position were plausible candidates. The definite article was excluded on the basis of the interaction between two constraints due to Zeevat (2005): the weak entailment constraint would have linked the definite article to *CAP* rather than to *REL* and the application domain constraint blocked this process from taking place.

#### **5.4. Conclusion**

In this section I argued that there is a natural link between the indefinite article and *REL*. The argumentation contains the answers to the two questions I expect an analysis that relates *REL* to the indefinite article to answer: (i) what makes the indefinite article special and (ii) what links it to *REL* rather than to *CAP*. The answer to the first question is that the indefinite article used to mark non-uniqueness and the answer to the second is that non-uniqueness is closely linked to the application of *REL* because of the non-uniqueness constraint on the elements of kind-sets. I furthermore argued that the distinction between *CAP* and *REL* could only have been

marked by the indefinite article. Crucial for this part of the argumentation were the weak entailment and the application domain constraints of Zeevat (2005).

At a more general level I take the argumentation in this section to show that it is more elegant to link the indefinite article to *REL* than to link NumP to *REL*. The reason for this is that the former link can be motivated but the latter cannot. This leads me to the conclusion that the semantics of the indefinite article in predicate position in Dutch is that of the *REL*.

Note that I have only looked at singular predication in this section. This means that I leave open the possibility of having a plural realization of *REL*. I briefly come back to this in Chapter 2, section 2.

## 6. Taking stock

In this first chapter on the role of the indefinite article in predicate position I presented the Dutch article and bare predication facts as a case-study and proposed two improvements on the analysis presented in de Swart, Winter & Zwarts (2007). More specifically I proposed a concrete semantics for the operators coercing kinds into capacities and capacities into kinds. I furthermore argued that the realization of *REL* is not NumP but the indefinite article. Crucial in the argumentation was that there is a motivated link between the indefinite article and *REL* but not between NumP and *REL*.



## Chapter 2: Article and Bare predication: extensions, limits and literature comparison

### 1. Introduction

In Chapter 1 I presented an analysis of Dutch article and bare predication that can be summarized in two points. The first is that there is a link between article and kind-membership predication and that this link has its origin in the fact that the indefinite article is the realization of *REL*. The second is that there is a link between bare and capacity predication and that this link has its origin in the fact that capacity predication is the complement of kind-membership predication. In this chapter I extend this analysis and present its limits (sections 2 and 3). I furthermore show it complements and improves on other analyses in the recent literature (sections 4 and 5).

### 2. Extensions

In this section I discuss two extensions of the adapted analysis of de Swart, Winter & Zwarts I presented in Chapter 1. The first is concerned with the kind-capacity contrast in the plural paradigm in Dutch, the second with the semantics of the indefinite article in predicate position in English. The extensions crucially rely on the origin of the link between the indefinite article and *REL* worked out in Chapter 1, section 5.

#### 2.1. From singulars to plurals

De Swart, Winter & Zwarts (2007) note that the contrast between (52) and (53) is comparable to the contrast between (50) and (51):

- (50) Marie is manager.  
Marie is manager
- (51) Marie is een manager.  
Marie is a manager
  
- (52) Marie en Piet zijn manager.  
Marie and Piet are manager
- (53) Marie en Piet zijn managers.  
Marie and Piet are managers

According to de Swart, Winter & Zwarts (2007) both (50) and (52) receive a standard capacity reading of *manager* whereas (51) and (53) receive a kind reading. In their original analysis these facts follow from the stipulation that NumP as a whole is linked to *REL*. This link can potentially be motivated for plural marking in the same way as for the indefinite singular article. The gist of the proposal would be that plural marking is akin to the indefinite article in imposing non-uniqueness on

the set it applies to. For the singular indefinite article I argued that the marking of non-uniqueness followed from a specific division of labour. For plural marking this is hardwired into the semantics on the assumption that plural nouns necessarily have plural reference (see Farkas & de Swart 2009). Given that non-uniqueness was the crucial link between the indefinite article and *REL* it seems plausible to assume that the same link has been established between plural marking and *REL*, be it for the plural rather than for the singular paradigm.

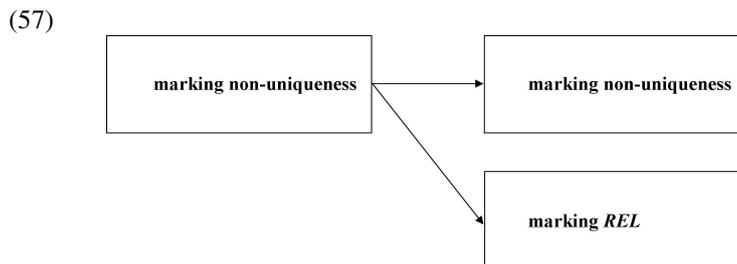
## 2.2. From Dutch to English

What we know about English is that it only allows for unique capacity nouns to occur bare. Unique capacity nouns are those that refer to functions that are necessarily unique within a certain social setting. In (54) through (56) I illustrate these facts.

- (54) \*John is man.  
 (55) \*Mary is doctor.  
 (56) Hyacinth is treasurer of the local Women's Institute.

(54) shows that kind nouns cannot appear bare in predicate position in English. (55) and (56) show that capacity nouns cannot either except if they refer to a function like *treasurer* that is necessarily unique within a certain social setting like the local Women's Institute.

These facts do not readily fit the analysis proposed in Chapter 1 but in what follows I will argue that they can be accommodated. Remember that I assume the indefinite article originally marked non-uniqueness and that this function made it into the perfect candidate to mark *REL*. Suppose now that in English the indefinite article got linked to *REL* in the same way as in Dutch but that it did not lose its original function of marking non-uniqueness. This historical process is depicted in (57):

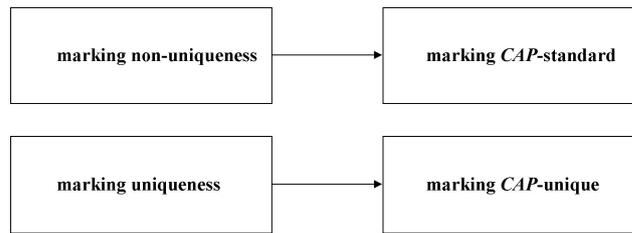


The result of this process is that there is a one-to-many mapping from the indefinite article to *REL* and to non-uniqueness. This means that the indefinite article was the default option for kind nouns and explicitly marked non-uniqueness for capacity nouns.

Suppose now that speakers have the desire to mark the distinction between unique and standard capacities and that – once more – the definite and the indefinite article

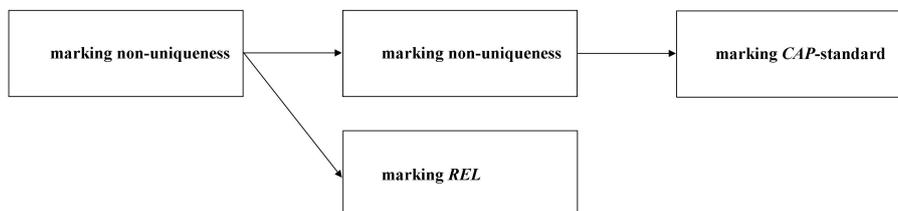
compete. The definite article would then go with unique capacities whereas the indefinite article would go with standard capacities. This follows from the weak entailment constraint of Zeevat (2005) according to which a form can only acquire a new function if its old function weakly entails the new function. We would then have one of the processes in (58):

(58)



The first process in (58) would lead the indefinite article to mark standard capacities and the second would lead the definite article to mark unique capacities. Which one of the two could have occurred can be decided on the basis of the application domain constraint of Zeevat (2005) according to which a form can only acquire a new function if the new function is applied more frequently than the old function. For the definite article I argued before that marking *CAP* would not lead to a widening of its application domain. The same holds by implication for the marking of a subtype of *CAP*. For the indefinite article however it is plausible to assume that the transition from explicitly marking non-uniqueness in some cases of capacity predication to obligatorily marking all standard capacities would lead to a widening of its application domain. What the application domain constraint tells us then is that the only process that might have taken place to mark the distinction between standard and unique capacities is the one leading the indefinite article to mark standard capacities. The complete historical evolution of the indefinite article in predicate position in English could then be represented as in (59):<sup>1</sup>

(59)



Interestingly the result of the process in (59) corresponds exactly to the situation in

<sup>1</sup> Note that the two steps in this reconstruction might not be distinguishable in corpora. This is due to the fine-grainedness of the distinctions and to the possibility of evolution to go quicker than historical corpora would allow us to record.

English where the indefinite article is used to mark *REL* and *CAP*-standard and where the bare form is used to mark *CAP*-unique. This shows that the same reasoning as the one I used to argue that the indefinite article in Dutch was linked to *REL* can be applied to derive the English predication facts. This leads me to the conclusion that the analysis proposed in Chapter 1 can accommodate the English predication facts.

### 2.3. Summary

In this section I argued that the adapted analysis of de Swart, Winter & Zwarts (2007) I presented in Chapter 1 can be extended in two directions. The first is the Dutch plural paradigm in which plural marking is linked to *REL*. The second direction is the English predication paradigm that differs from the Dutch one in not allowing standard capacity nouns to occur bare.

In the next section I shift my attention from extensions of the analysis to the limits it presents.

## 3. Limits

In this section I present three limits of the analysis I proposed in Chapter 1. The first one is related to the lexical ambiguity of the indefinite article, the second and the third are about data that require further scrutiny.

### 3.1. The lexical ambiguity of the indefinite article

In the original analysis of de Swart, Winter & Zwarts NumP was related to *REL* and this derived all the facts. My objection to this analysis was that the link between NumP and *REL* was a pure stipulation. I subsequently argued that the link with *REL* could be made insightful for two instantiations of NumP, *viz.* the indefinite article (see Chapter 1, section 5) and plural marking (see this Chapter, section 2). The disadvantage of the analysis I proposed is that I have to assume the indefinite article and the plural morpheme are ambiguous between their ‘standard’ semantics and marking *REL*. This raises the question which of the analyses is more attractive: the one that stipulates that a syntactic projection is related to *REL* or the one that posits that lexical items are ambiguous. The crucial reason for me to prefer my analysis over the original one proposed by de Swart, Winter & Zwarts is that the lexical ambiguities I posit are not stipulations but are based on motivated links between the different meanings of the lexical items. An analysis that posits no ambiguities would still be superior to mine though.

### 3.2. Bare predication and modification

In the analysis I propose the indefinite article is taken to be the realization of *REL* in predicate position. The question this raises is what to do in cases like (60) in which bare predication is impossible and article predication allows for a capacity reading of *advocaat*.

- (60) Jan is \*(een) goede advocaat. (Dutch)  
 John is a good lawyer  
 “John is a good lawyer”

The standard reading of (60) is that Jan is a lawyer and that he is good at his job. The answer to the question why this reading is available probably lies in the fact that bare predication does not allow for certain kinds of modification. As a result there would be no competition between article and bare predication and there would be no way to distinguish kind readings from capacity readings. The exact types of modification that are allowed with bare predication remain to be investigated though.

### 3.3. The *dat is* construction

The analysis I proposed predicts the availability of article and bare predicates to be independent of the subject of the sentence. The unacceptability of the bare version of (61) in the *dat is* construction consequently does not follow from the analysis:

- (61) Marie, dat is \*(een) advocate.  
 Marie, that is a lawyer

I leave a worked out account of the unacceptability of the bare version of (61) for future work but I do want to give an indication of what I think the account should look like. I first sketch the account proposed by Beyssade & Dobrovie-Sorin for the French counterpart of the *dat is* construction and argue that it cannot be on the right track. I then sketch an account that derives the unacceptability of the bare version of (61) from the fact that *dat* in principle refers to (non-human) objects.

For the French counterpart of the *dat is* construction Beyssade & Dobrovie-Sorin (2009) argue that a type constraint prevents it from taking bare predicates. Without getting into the details of their analysis the basic story they propose is that *advocate* is of type  $\langle e, t \rangle$ , *een advocate* of type  $e$  and that *dat is* is only compatible with the latter type. Crucially this analysis makes the prediction that  $\langle e, t \rangle$  expressions like adjectives should always be out in sentences like (61). As (62) shows this prediction is not borne out:

- (62) De hele nacht dansen, dat is geweldig.  
 The whole night dance, that is great  
 Dancing the whole night, that is great.

Under the assumption that adjectives like *geweldig* are of type  $\langle e, t \rangle$  (62) should be out on the analysis of Beyssade & Dobrovie-Sorin. Given that this is not the case I conclude that the type account they propose cannot be on the right track.

The account I will pursue in future work is connected to the observation that *dat* in principle refers to (non-human) objects. I hypothesize that this characteristic of *dat* prevents capacities in the *dat is* construction from shifting covertly from type  $e$  to type  $\langle e, t \rangle$ . The underlying intuition is that the shift from capacities to sets in ordinary bare predication sentences is triggered by the fact that it would make no

sense to equate an ordinary human individual with a capacity. A similar trigger is however unavailable if the subject refers to (non-human) objects and the shift would have to be explicitly marked – an extra function for the indefinite article. If this analysis is correct the bare version of (61) is not ungrammatical but pragmatically odd: Marie would end up being equated with the capacity *advocaat*. Crucial for this analysis to work is that *advocaat* on its *e*-type reading can appear in *dat is* sentences. (63) shows that this is indeed the case:

- (63) Het mooiste beroep ter wereld, dat is advocaat.  
The most beautiful profession in the world, that is lawyer.

In (63) *advocaat* appears in the predicate position of the *dat is* construction and refers to the profession *advocaat*. It is the availability of this *e*-type reading that I assume lies at the origin of the pragmatic oddness of the bare version of (61).

### 3.4. Summary

In this section I briefly pointed out some of the limits of the analysis I proposed. The first was related to the fact that I have to make the unattractive assumption that the indefinite article and the plural morpheme are lexically ambiguous. The second was connected to the fact that there is still work to be done on what kinds of modification are allowed with bare predication and why this is the case. The third and final limit was related to the unavailability of bare predicates in the *dat is* construction. In the following sections I will argue that – despite its limits – my analysis complements or improves on other ones.

## 4. Literature comparison: Part I

In this and the next section I compare my analysis to the ones that have been proposed in the recent literature. Given that most other analyses are not about Dutch I however first have to situate the predication facts of Dutch with respect to the facts of the other languages that have been treated in the literature. I do this in 4.1. focusing on bare predication and treat a first set of competing analyses in 4.2.

### 4.1. Bare predication beyond Dutch

The phenomenon of bare predication is not restricted to Dutch but also occurs in many other languages. Some examples are given in (64):<sup>2</sup>

- (64) a. Henry became treasurer. [English]  
b. Jean est médecin. [French]  
Jean is doctor  
Jean is a doctor.

<sup>2</sup> Example a. is taken from Huddleston & Pullum (2002), examples b.-d. are taken from Munn & Schmitt (2005) and example e. is based on an example in Giurgea (2008).

c. Gianni è dottore. Gianni is doctor Gianni is a doctor.	[Italian]
d. Juan es médico. Juan is doctor Juan is a doctor.	[Spanish]
e. Ion e medic. Ion is doctor Ion is a doctor.	[Romanian]

A striking resemblance between the languages is that they all allow capacity nouns to appear bare in predicate position. This should however not hide the fact that there is some variation. In what follows I will present this variation as a decrease of the restrictions on bare predication.

English appears to be the most restricted language allowing only unique capacity nouns to occur bare. Unique capacity nouns are those that refer to functions like *deputy leader of the party* that are unique within one social setting. Other examples are *treasurer, CEO, chairman of the board, president, ...* (Huddleston & Pullum 2002). Dutch is slightly more permissive allowing all capacity nouns to occur bare. Next to unique capacity nouns we therefore also find nouns like *bioloog* ('biologist'), *boerenwerkmán* ('farm labourer'), ...

From Dutch onwards languages become more and more permissive. In order to get some idea of the variation involved it is instructive to look at some quantitative data. The table in (65) compares the number of cases of bare predication to the number of cases of article predication. The Romance data are taken from a corpus study by Van Peteghem (see Van Peteghem 1993) and the Dutch ones are based on a study I carried out on the Eindhovencorpus.<sup>3</sup> Note that the data are restricted to unmodified nouns.

---

<sup>3</sup> The Eindhovencorpus is a +/- 720.000 word corpus, tagged for parts of speech and freely available via [www.inl.nl](http://www.inl.nl). The small corpus study had two parts. I first looked for article predication using the following (regular expression) string: "[a-zA-Z0-9\_]\*\_2[0-9]\* een\_450 [a-zA-Z0-9\_]\*\_0[0-1]0". This gave me all verbs followed by the indefinite article and a noun. I manually selected all non-modified count nouns and made a list of the predicative verbs. I then looked for bare predication taking the verbs on this list followed by a noun. Here too I manually selected all non-modified count nouns. The string I used was "VERB\_2[0-9]\* [a-z]\*\_0[02]0". The verb forms that were included are *ben, bent, bleek, blijft, is, leek, lijkt, vormen, waren, was, werd, wordt* and *zijn*.

(65)

	cases of bare predication	cases of article predication	total number	% of cases of bare predication	% of cases of article predication
Dutch	69	275	349	20	80
French	259	440	699	37	63
Italian	271	390	661	41	59
Spanish	356	412	768	46	54
Romanian	578	221	799	72.5	27.5

The table in (65) suggests that Romanian is at the extreme end of the scale at which everything can be bare. This is further illustrated by the examples in (66) that – on their most straightforward interpretation – would be out in the other languages of the table:

- (66) a. *Ăsta e scaun.*  
 That is chair  
 That is a chair.  
 b. *Mihaela e fată deșteaptă.*  
 Mihaela is girl smart  
 Mihaela is a smart girl.

data taken from Giurgea (2008)

As for French, Italian and Spanish, the table in (65) shows that they are clearly more permissive than Dutch but it is difficult to pin down where exactly lies the difference. Van Peteghem faced a similar problem in her attempt to explain the variation between French, Italian and Spanish (see Van Peteghem 1993). One of her suggestions was that at least part of the variation might be explained by the degree to which languages allow nouns to be reanalyzed as adjectives. If French is less permissive than Italian in allowing for an adjectival reanalysis of nouns and Italian in turn is less permissive than Spanish this would indeed account for part of the variation. The data in (67) through (70) seem to suggest that a similar story might work to distinguish Dutch from French, Italian and Spanish.

- (67) ??*Dat is heel jongen.* [Dutch]  
 That is very boy.  
 That is very boy-like.  
 (68) *C'est très garçon.* [French]  
 That is very boy.  
 That is very boy-like.  
 (69) *E molto uomo.* [Italian]  
 It is very man  
 It is very man-like.  
 (70) *Es muy hombre.* [Spanish]  
 It is very man  
 It is very man-like.

In (67) through (70) I tried to force an adjectival reanalysis of nouns by adding a gradable adverb to them that is in principle only compatible with adjectives. This works fine for French, Italian and Spanish but fails for Dutch which suggests that Dutch is indeed less permissive in allowing for an adjectival reanalysis of nouns than the other languages. In line with Van Peteghem (1993) one might then assume that part of the variation we see between Dutch, French, Italian and Spanish can be attributed to the degree to which these languages allow for an adjectival reanalysis of nouns. Note though that more research is needed to account for all of the variation.

On the basis of the preceding I end up with the following scale of languages allowing for bare predication:

(71) English – Dutch – French – Italian – Spanish – Romanian

The more the language occurs to the right of the scale the more easily it allows for bare predication. English is the most restricted one only allowing for unique capacity nouns. Dutch is slightly more permissive allowing for all capacity nouns. French, Italian and Spanish seem even more permissive than Dutch and this might be due to the fact that they allow nouns to be reanalyzed as adjectives more easily than Dutch. Romanian finally seems to be insensitive to the role – non-role distinction.

In my literature comparison I will focus on Dutch, French and Italian. Even though the slight differences between Dutch on the one hand and French and Italian on the other hand might sometimes blur the comparison I expect an analysis that accounts for the facts in one language to be able to account for the facts in the two other languages.

#### 4.2. Article and bare predication: *e* vs. $\langle e,t \rangle$ ?

The analyses I evaluate here are the ones proposed for French by Beyssade & Dobrovie-Sorin (2005, 2009) and Mari & Martin (2008, 2009). Two preliminary remarks are in order. The first is that I collapse the different analyses into one. This is motivated by the fact that they all share the same basic intuitions. The second preliminary remark is that I will illustrate the analysis on the basis of Dutch unless the French data are crucial for the argumentation.

This subsection is organized as follows. I first sketch how the *e* vs.  $\langle e,t \rangle$  distinction can be formalized and argued for. I then show that there are cases that cannot be analyzed in this way and I end with a short conclusion in which I suggest that the *e* vs.  $\langle e,t \rangle$  analysis and the analysis I proposed are complementary.

##### 4.2.1. Predicates and the *e* vs. $\langle e,t \rangle$ distinction

The analyses proposed by Beyssade & Dobrovie-Sorin (2005, 2009) and Mari & Martin (2008, 2009) start from the assumption that the article in predicate position has the same function as in argument position, *viz.* that of introducing new discourse

referents (Heim 1982, Kamp & Reyle 1993). *Jan is een man* ('Jan is a man') could then be analyzed along the following lines:

(72)

x,y
Jan (x) man (y) x=y

In words: there is an individual x called Jan, there is an individual y that is a man and x and y are identical.

In this line of analysis bare predicates would differ from article predicates in not introducing any discourse referents. *Jan is advocaat* ('Jan is lawyer') would then be analyzed along the following lines:

(73)

x
Jan (x) advocaat (x)

In words: there is an individual x called Jan and this individual is a lawyer.

For French two pieces of evidence have been proposed in favour of this analysis. The first comes from Beyssade & Dobrovie-Sorin (2005, 2009) who claim that article predicates are ungrammatical in small clauses whereas bare predicates are perfectly acceptable:

(74) \*J'imagine Romain Gary un écrivain.

I imagine Romain Gary a writer

(75) J'imagine Romain Gary écrivain.

I imagine Romain Gary writer

Under the assumption that small clauses can only host  $\langle e,t \rangle$  expressions this suggests that *un écrivain* necessarily has an argumental type whereas *écrivain* is of type  $\langle e,t \rangle$ . Evidence for the  $\langle e,t \rangle$  requirement of small clauses comes from the unacceptability of proper names (type  $e$ ) and the acceptability of adjectives (type  $\langle e,t \rangle$ ):

(76) \*I consider Romain Gary Emile Ajar

(77) I consider Romain Gary intelligent.

The second piece of evidence in favour of an analysis along the lines in (72) and (73) comes from Mari & Martin (2009) who claim that in a context in which it is established that Pierre is a lawyer it is impossible to use (78) to state this:

- (78) Pierre est un avocat.  
Pierre is a lawyer

Such a context can linguistically be forced through the addition of the presupposition-trigger *toujours* 'still':

- (79) \*Pierre est toujours un avocat.  
Pierre is still a lawyer

The unacceptability of (79) – according to Mari & Martin (2009) – would follow from the fact that the discourse referent introduced by *un avocat* is not new. Given that bare predicates don't introduce discourse referents the version with a bare predicate is correctly predicted to be fine:

- (80) Pierre est toujours avocat.  
Pierre is still lawyer

Given that Mari & Martin assume *avocat* in (80) does not introduce a discourse referent there is also no discourse novelty requirement that could clash with the presupposition triggered by *toujours*.

#### 4.2.2. The problem

The basic problem analyses like the one sketched in 4.2.1. face is to be found in the article predication part and more specifically in the assumption that all article predicates introduce discourse referents. (81) shows that this assumption cannot be correct:<sup>4</sup>

- (81) Dit boek is [een enorme openbaring]<sub>i</sub>. Ik vind \*haar<sub>i</sub> prachtig.  
This book is a huge revelation I find her magnificent  
This book is a huge revelation. I find it magnificent.

If it were the case that all article predicates introduce discourse referents we would expect *haar* to be able to pick up the feminine *een enorme openbaring*. This expectation is not borne out.

One might argue that (81) does not conclusively show that some article predicates do not introduce discourse referents. It could indeed be that discourse referents introduced by predicates are simply not salient enough to be picked up. The

---

<sup>4</sup> A note of caution: the pronominal system I use in this section is Belgian and not Dutch!

strongest argument against this comes from the fact that some article predicates do introduce discourse referents that are picked up. This is illustrated by (82):

- (82) Mijn geboorteplaats is [een dorpje in de Provincie Antwerpen]<sub>i</sub>. Het<sub>i</sub> ligt verscholen in de bossen.  
My place of birth is a village in the Province of Antwerp. It lies hidden in the forests.

In (82) *Het* picks up *een dorpje in de Provincie Antwerpen* which is the article predicate of the preceding sentence. To make sure that this is the case note that *mijn geboorteplaats* is feminine and would have to be picked up by the pronoun *ze*:

- (83) [Mijn geboorteplaats]<sub>i</sub> is een dorpje in de Provincie Antwerpen. Ik zoek \*het<sub>i</sub> / ze<sub>i</sub> nog vaak op.  
My place of birth is a village in the Province of Antwerp. I look it / her still often up.  
My place of birth is a village in the Province of Antwerp. I still often visit it.

I take (83) to show that some article predicates do introduce discourse referents and that (82) is conclusive evidence in favour of the claim that some article predicates do not. From this it follows that the *e* type analysis proposed for article predication cannot apply to all article predicates; at least some of them seem to pattern more with the  $\langle e, t \rangle$  type analysis proposed for bare predication.

Before concluding it is important to revisit the arguments that were proposed to defend the distinction between article and bare predication in terms of an *e* vs.  $\langle e, t \rangle$  distinction. The first was concerned with the fact that small clauses do not allow for article predication:

- (84) \*J' imagine Romain Gary un écrivain.  
I imagine Romain Gary a writer

Given that small clauses do not accept proper names in predicate position either but do accept adjectives, the unacceptability of (84) can be taken as an argument in favour of the assumption that article predicates are of type *e*. There are however several reasons to doubt that this argument is conclusive. The first is theoretical: assuming that article predicates and proper names are out in small clauses because of their type is only possible if we furthermore assume they cannot type-shift. The latter assumption is undesirable though given that it goes against standard assumptions in type-shifting theory about the type-shifting possibilities of DPs headed by the indefinite article and proper names. The second reason is that definite DPs are allowed in the predicate position of small clauses:

- (85) un homme que je croyais mon ami  
a man that I thought my friend

data taken from *Frantext*

- (86) je me croyais de nouveau le maître de ma vie  
 I me thought again the master of my life  
 I thought I had regained control over my life

data taken from *Frantext*

Given that definite DPs are in general analyzed as being of type *e* or type  $\langle\langle e,t\rangle,t\rangle$  their acceptability in small clauses is surprising if small clause predicates could only host expressions of type  $\langle e,t\rangle$ . The third and final reason to doubt that the unacceptability of (84) is due to a type-mismatch is that there are attested cases of DPs headed by the indefinite article in the predicate position of small clauses:

- (87) Mlle de Pen-Hoël [...] aurait cru Calyste un dissipateur [...]  
 Miss de Pen-Hoël would\_have believed Calyste a squanderer  
 Miss de Pen-Hoël would have believed Calyste to be a squanderer

data taken from *Frantext*

- (88) [...] Mme Bontemps croit Bloch un génie [...]  
 Mrs Bontemps thinks Bloch a genius  
 Mrs Bontemps thinks Bloch's a genius.

data taken from *Frantext*

Given that type-shifting is not sensitive to the type of nouns involved, an analysis that takes (84) to be out because of a type clash should predict (87) and (88) to be equally unacceptable. The fact that they are not indicates that the analysis is probably not on the right track.

The preceding discussion leads me to assume that the unacceptability of article predicates in small clauses cannot be attributed to their type *e* status. This entails that the first argument in favour of a generalized *e* type analysis of article predication does not hold.

The second argument in favour of an *e* type analysis of article predicates was the apparent discourse novelty requirement connected to article predication:

- (89) \*Pierre est toujours un avocat.  
 Pierre is still a lawyer

(89) would be out because of a clash between the presupposition triggered by *toujours* and the novelty requirement on the discourse referent introduced by *un avocat*. The problem this argument poses is that it is worked out on the basis of an example with a capacity noun. Given that capacity nouns prefer to occur bare anyway the example in (89) is expected to be degraded whether or not *toujours* appears. In order to make a proper evaluation of the argument we should therefore not look at capacity nouns but at kind nouns. The example in (90) shows that article predication with kind nouns is compatible with *toujours*. Note that I looked for an example in a fantasy world to make sure that *toujours* would be pragmatically felicitous with kind nouns.

- (90) S'il est toujours un loup à la fin de la bataille, il peut tenter une dernière fois de recouvrer sa forme humaine.  
 If he is still a wolf at the end of the battle, he can try a last time to recover his human form  
<http://fabien.jamme1.free.fr/Documents%20pour%20Mordheim/Pretre-Loup.pdf> [19/10/2009]

(90) shows that article predication can co-occur with *toujours*. This suggests that the unacceptability of (89) is not due to a clash between the presupposition triggered by *toujours* and some discourse novelty requirement. It furthermore suggests that the second argument in favour of a generalized *e* type analysis of article predication does not hold.

#### 4.2.3. Conclusion

In 2.4.1. I sketched an analysis that takes the difference between article and bare predication to lie in an *e* vs.  $\langle e, t \rangle$  distinction. This analysis was based on Beyssade & Dobrovie-Sorin (2005, 2009) and Mari & Martin (2008, 2009). In 2.4.2. I argued that this analysis might work to oppose bare predication to some cases of article predication but not to all. Crucially some article predicates seem to behave in the same way as bare predicates. This suggests that for these cases the *e* vs.  $\langle e, t \rangle$  distinction has to be supplemented with an extra distinction in the  $\langle e, t \rangle$  domain. I assume this distinction corresponds to the kind vs. capacity distinction and that the analysis presented in Chapter 1 complements the analyses proposed by Beyssade & Dobrovie-Sorin (2005, 2009) and Mari & Martin (2008, 2009). In the next section I evaluate a second set of analyses.

### 5. Literature comparison: Part II

In section 4 I evaluated analyses that take article predicates to be of type *e* and bare predication to be of type  $\langle e, t \rangle$ . In this section I evaluate three analyses that assume article and bare predication are of type  $\langle e, t \rangle$ : Roy (2006), Ebeling (2006) and Zamparelli (2008). After the evaluation of all three of them I present a short conclusion.

#### 5.1. Roy (2006)

Roy assumes all nouns come with an event argument that has to be bound. When bound by the indefinite article it is signalled that the predication holds for the maximal event around the 'time of utterance' (given by the Tense on the copula) and that this event cannot be split up into smaller intervals.<sup>5</sup> When bound by Tense it is

---

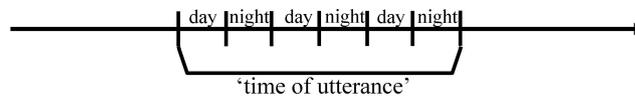
<sup>5</sup> 'Time of utterance' is Roy's term. Brenda Laca pointed out to me that the term 'time of reference' might be more appropriate.

signalled that the maximal event can be split up. The facts that led to this analysis are presented in (91) and (92):

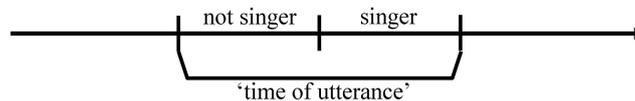
- (91) a. Jean est professeur le jour, danseur la nuit.  
[French]  
(French, *litt.* John is teacher by day, dancer by night)
- b. ?? Jean est un professeur le jour, un danseur la nuit.  
[French]  
(French, *litt.* John is a teacher by day, a dancer by night)
- (92) a. Paul est devenu chanteur.  
[French]  
(French, *litt.* Paul has become singer)
- b. ?? Paul est devenu un chanteur.  
[French]  
(French, *litt.* Paul has become a singer)

The reason why the b-variants are out on Roy's analysis is that adverbials like *le jour ... la nuit* ('by day... by night') and verbs like *devenir* ('become') split up the 'time of utterance'. This is depicted for the adverbials in (93) and for the verb in (94).

(93)



(94)



This analysis presents some problems. First of all it assumes all nouns are created equal which means that capacity readings of kind nouns should be as unmarked as their kind readings and vice versa for capacity nouns. The only way Roy can predict a contrast is through a pragmatic reasoning that makes capacities more likely to be split up. Note though that this pragmatic analysis would get into serious problems in English (see section 2.2). Indeed, assuming that capacities are more likely to be split up is one thing, assuming that only unique capacities are more likely to be split up another.

The second problem Roy faces is that her analysis is based on some strange intuitions. If we take her analysis seriously the default reading of (91.a) should be that John is not a teacher at night and not a dancer during the day. This is not the default reading though. The default reading is rather something along the lines of "At school John is a teacher but in the disco John is a dancer". Even though John's being at school and being in the disco are probably temporally disjoint what seems to play a role is not time but rather the social setting.

The third and most important problem Roy faces is that she makes wrong predictions. The empirical claim her whole analysis depends on is that temporal adverbials and transition verbs are incompatible with non-bare predication. As shown in (95) and (96) this is not correct.

- (95) In Lady Hawke is Rutger Hauer ‘s nachts een wolf en overdag een man.  
 In Lady Hawke is Rutger Hauer by night a wolf and by\_day a man  
 “In Lady Hawke Rutger Hauer is wolf by night and a man by day.”  
 [Dutch]
- (96) La chenille est devenue un papillon.  
 The caterpillar is become a butterfly  
 “The caterpillar has become a butterfly.”  
 [French]

One could wonder why (95) and (96) are in whereas in general kind predication is not. The answer goes back to the distinction between kinds and capacities: kind membership depends on inherent properties and changing these is in general not possible. The cases in (95) and (96) are the exceptions I mentioned in Chapter 1, section 2: whenever kind-membership is made dependent on wizardry or whenever there is genetically programmed evolution kind-membership can change.

## 5.2. Ebeling (2006)

Ebeling (2006) claims that the lack of an indefinite article in predicate position encodes [DST] of which the definition is given in (97):

- (97) [DST] “collection of distinctive properties that are acquired by their bearer  $x$  or of which the proportion of their presence in  $x$  is acquired, and of which the speaker presents the bearer  $x$  without paying attention to quasi-distinctive and accidental properties, in order to focus on the relation between  $x$  and  $x$ ’s surrounding reality insofar as this reality is a consequence of the distinctive properties”

Using the indefinite article would mean that one does not abstract away from quasi-distinctive and accidental properties. Note furthermore that Ebeling does not assume any lexical distinction between capacities and kinds.

As far as bare predication is concerned, Ebeling’s semantics has the same result as mine: bare predication according to him is about the relation between  $x$  and  $x$ ’s surrounding reality, which corresponds to what I would call the capacity of  $x$  in a social setting. His analysis of article predication is somewhat different though and as far as I can see counterintuitive. To see this, let me first present the minimal pair he proposed to back up his analysis:

- (98) Hij is boekhouder.  
 He is accountant
- (99) Hij is een boekhouder.  
 Hij is an accountant.

(98) is interpreted as stating that the *he*-person is an accountant by profession whereas (99) is interpreted as stating that the *he*-person has all the properties we generally associate with accountants without abstracting away from accidental properties linked to accountants. Given that Ebeling assumes there are no lexical restrictions on bare and non-bare predication we would expect (100) to also convey something accidental but as far as I can judge this expectation is not borne out.

- (100) White Fang is een wolf.  
WF is a wolf

Apart from the counterintuitive prediction concerning (100) Ebeling's analysis faces two major problems. The first is that his semantics of bare predication is purely stipulative in that it is not linked to any general semantics of bareness. The second is that the setup of his analysis is such that the indefinite article has (practically) no semantic content whereas the lack of the indefinite article does. From a compositionality point of view this is extremely odd.

### 5.3. Zamparelli (2008)

In Zamparelli (2008) capacity nouns come without a gender feature. Given that they have to get this feature and given that articles cannot provide it they have to enter into an agreement relation with their subject. Through this agreement they also get a value for semantic number. Other nouns have to get a value for semantic number from the indefinite article, the agreement relation being too costly to apply if other options are available.

This analysis derives the kind – capacity contrast in the lexicon. The insight behind it is that professions can be performed by men and women and that the gender of profession nouns is therefore expected to be underdetermined.

This analysis presents two problems. The first is that it predicts capacity nouns that are specified for gender like the Dutch *lerares* ('teacher<sub>fem</sub>') and *leraar* ('teacher<sub>masc</sub>') to take the indefinite article. This prediction is not borne out as is shown by (101):

- (101) Jan is leraar en Marie is lerares.  
Jan is teacher<sub>masc</sub> and Marie is teacher<sub>fem</sub>

The second problem Zamparelli faces is that his account cannot straightforwardly be extended to English. Indeed, if the indefinite article is nothing but a number marker the distinction between unique and standard capacities cannot be derived.

### 5.4. Conclusion

In this section I evaluated three analyses of the contrast between article and bare predication. The first is the one from Roy who proposes that the crucial difference between article and bare predication is to be located in their temporal properties. The main argument against this analysis is that it makes wrong empirical predictions.

The second analysis I evaluated was the one proposed by Ebeling. This analysis is very close to mine but has two unattractive features. The first is that it misses a lexical component. The consequence of this is that the analysis makes unintuitive predictions about standard kind-membership predication. The second unattractive feature Ebeling's analysis presents is that the proposed semantics is not linked to any deeper insight about articles or their absence.

The third and final analysis I evaluated was the one proposed by Zamparelli. This author argues that profession nouns lack a gender feature and that this lack can be related to their ability to appear bare in predicate position. He furthermore argues that the indefinite article functions merely as a number marker. The two arguments against this analysis are (i) that not all profession nouns that occur bare in predicate position lack a gender feature and that (ii) the English facts cannot be derived if the indefinite article is nothing more than a number marker.

## 6. Summary and Conclusions

In this chapter I argued that the analysis I presented in Chapter 1 can be extended to the Dutch plural paradigm and the English singular predication paradigm. I furthermore argued that – despite its limits – the analysis complements and improves on other analyses in the recent literature.

I draw two conclusions on the basis of the discussion in this chapter. The first is that the analysis I proposed in Chapter 1 is on the right track, not only for Dutch but also for languages like English, French and Italian. The second conclusion I draw is that the analysis I proposed has to be complemented with an analysis along the lines of those proposed by Beyssade & Dobrovie-Sorin (2005,2009) and Mari & Martin (2008, 2009). Crucially, the indefinite article in predicate position is not only the realization of *REL* but can also have its standard DRT interpretation. Indeed, allowing for the indefinite article to introduce discourse referents is the only way to derive the possibility of anaphoric pickup in examples like the following:

(102) Mijn geboorteplaats is [een dorpje in de Provincie Antwerpen]. Het<sub>i</sub> ligt verscholen in de bossen.

My place of birth is a village in the Province of Antwerp. It lies hidden in the forests.

If article predication were only of type  $\langle e,t \rangle$  the predicate in the first sentence of (102) would not be able to introduce a discourse referent. Given that it apparently does I have to assume that article predicates can also be of type  $e$ .

### ***General Conclusion***

In this part of my dissertation I investigated the role of the indefinite article in predicate position. I argued that this role is twofold. On the one hand the indefinite article introduces discourse referents in the same way as it does in argument position. On the other hand it is also the realization of *REL*, an operator that takes kinds and returns the sets of their instantiations. English is a bit special in the sense that the indefinite article is not only the realization of *REL* but also of *CAP-standard*, an operator that takes standard capacities and returns the sets of individuals that have these capacities.

The question I wanted to answer in this part of my dissertation was why the indefinite article – in its capacity of argument marker – appears almost obligatorily in predicate position. The answer I argued for is that at some point in history it was the right item in the right place. More specifically, I argued that the indefinite article was the only item that could end up marking the distinction between *REL* and *CAP* in the singular predication paradigm. Given that *REL* happens to be applied far more frequently than *CAP* the evolution linking the indefinite article to *REL* gave rise to the almost obligatory character of the indefinite article in predicate position.



## CONCLUDING SUMMARY

The main goal of this dissertation was to fine-tune our notion of indefinite article. It combined formal semantics and pragmatics, presented challenging new corpus data both from synchrony and diachrony and offered simulations of historical processes that – up till recently – could only be described. In this concluding summary I will briefly sketch the main lines of argumentation and the most important conclusions.

### 1. Spanish *unos*

In the first part of this dissertation I zoomed in on Spanish *unos* and I argued that it should be analyzed as a plural indefinite article. The argumentation was both synchronic and diachronic.

#### *Synchrony*

For the synchronic part I started from the standard assumption that indefinite articles are ‘default’ determiners in the sense that they have no semantic contribution beyond that of marking argumenthood. In order to argue that *unos* is an indefinite article I showed that it lacks several semantic ingredients that one might have expected to be standard for indefinite determiners: partitivity, the ability to induce distributive dependencies, a quantity specification, ... Examples (1) through (3) illustrate these properties (or rather their lack):

- (1) Unos estudiantes vinieron a verme.  
UNOS students came to see\_me  
<sup>ok</sup>*Some students came to see me.*  
*#Some students came to see me, other didn't.*
- (2) Unos hombres compraron un billete de lotería.  
UNOS men bought a ticket of lottery  
<sup>pref</sup>*Some men collectively bought a lottery ticket.*  
<sup>dispref</sup>*Some men bought a lottery ticket each.*
- (3) ¿Cuántos libros tienes? ??Unos.  
How\_many books you\_have UNOS

I also argued that *unos N* is to be distinguished from bare plurals only in its pragmatics: whereas the former is generally used to introduce salient discourse referents, the latter is used to introduce non-salient ones. I used this pragmatic difference – that I derive in Weak Bi-directional OT – to explain why indefinite articles can co-exist with argumental bare nominals, why *unos* can't occur in the scope of negation and why bare plurals can't occur in preverbal subject position. These facts are illustrated in (4) through (6):

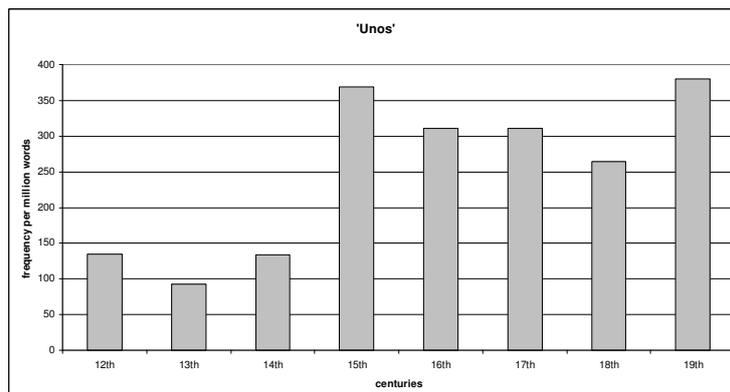
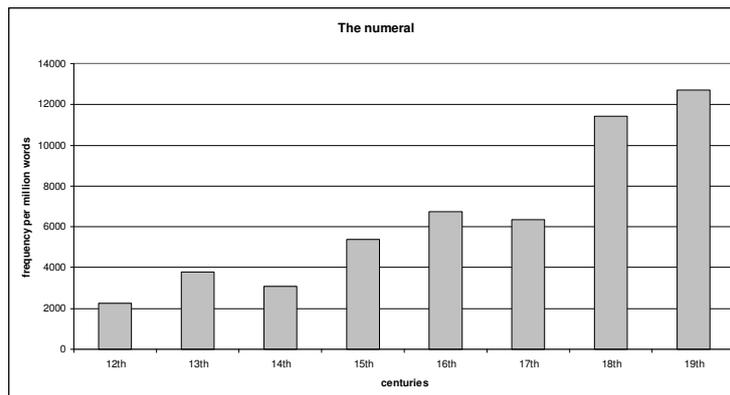
- (4) María comió bizcochos | unos bizcochos.  
María ate biscuits | UNOS biscuits

- (5) A la reunión no asistieron unos profesores.  
 At the meeting not attended some professors  
<sup>ok</sup>*There are some professors who didn't attend the meeting.*  
*#No professor attended the meeting.*
- (6) \*Políticos han ocupado el palacio.  
 Politicians have occupied the palace

### Diachrony

For the diachronic part I started from the observation in the grammaticalization literature that the development of indefinite articles generally goes through three stages: in the first stage argumental nouns occur bare, in the second indefinite articles start to be used to introduce salient discourse referents and in the third and final stage indefinite articles are generalized to all arguments. *Unos* is special in the sense that it didn't go all the way to stage 3. I showed this on the basis of the corpus data in (xx). The first graph shows the relative frequency of the singular indefinite article from the 12<sup>th</sup> to the 19<sup>th</sup> century. The second graph does the same for *unos*.

(7)



What the graphs show is that the singular indefinite article underwent two increases in frequency whereas *unos* only underwent one. This suggests that the singular indefinite article went through the three stages of article development whereas *unos* only went through the first two. The question that imposed itself then is whether stage 2 articles are real articles. The answer I proposed is affirmative. The first argument came from the fact that *unos* underwent a semantic change from stage 1 to stage 2: on the basis of qualitative corpus data I showed that *unos* did allow for partitive readings before it reached stage 2. The second argument came from the synchronic facts: there are no differences between bare plurals and *unos N* that cannot be explained on the basis of pragmatic competition. This indicates that *unos* is really an unmarked determiner. The third and final argument came from a simulation of the evolution of indefinite articles in EvolOT: I showed that their evolution can be simulated in such a way that it has different cutoff points (stage 2 for *unos* and stage 3 for indefinite articles in general). This means that the joint transition of the indefinite singular article and *unos* in the 15<sup>th</sup> century is a real transition from a standard determiner to an indefinite article.

### *Conclusions*

The discussion in Part I led to three conclusions. The first is that – in synchrony – DPs headed by indefinite articles are expected to lack certain properties other DPs have (partitivity, distributivity, quantity) and to differ from bare nominal arguments (if at all) only in their pragmatics. The second conclusion is that – in diachrony – articles are expected to exhibit a certain frequency pattern and to undergo semantic changes. The third and final conclusion is that *unos*, based on its synchronic and diachronic behaviour, should be analyzed as an indefinite article. This means that indefinite articles and bare nominal arguments can co-exist and that the existence of indefinite articles can never be used as a sufficient explanation for the ungrammaticality of bare nominal arguments.

## **2. French *des***

In the second part of this dissertation I zoomed in on French *des*. I started with the observation that *des N* can receive several analyses. The ones I focused on most were the DP analysis in which *des* functions as a plural indefinite article and the bare partitive analysis in which *des* is assumed to be composed of *de* ('of') and *les* ('the<sub>pl</sub>').

### *The plural indefinite article analysis*

Analyzing *des* as an indefinite article is very attractive in at least two respects. The first is that it straightforwardly accounts for the plural and the indefinite interpretation of *des N*. The second is that it nicely predicts the historical appearance of *des N* (on its plural indefinite interpretation) to coincide more or less with the generalization of the use of the singular indefinite article.

The problem an article analysis faces is that indefinite articles in general are free to take narrow or wide scope whereas *des N* has a clear preference for narrow scope:

- (8) Il me faut des chemises.  
 There me must DES shirts  
 I need shirts.

The strongly preferred reading of (8) is that I need shirts, no matter which. Under the assumption that articles are scopally free this strong preference is unexpected.

*The bare partitive analysis*

What makes the bare partitive analysis even more attractive than the indefinite article analysis is that – next to accounting for all the facts the latter can account for – it takes the complex morphology of *des* seriously. On the synchronic side I argued that a bare partitive analysis of *des N* can account for its plural and indefinite interpretation without making any *ad hoc* assumptions. The plural interpretation was derived through default reasoning, a variant of ‘maximize presupposition’ (see Heim 1991 and Sauerland et al. 2003). The indefinite interpretation was derived on the basis of a Weak Bi-directional OT analysis that made sure the definite DP in *des N* could only refer to kinds. I furthermore argued that the same pragmatic assumptions have to be made for bare partitives that cannot be analyzed as articles: Dutch *van die N*, French *de ces N* and Italian *di quei N* which are exemplified in (9) through (11).

DUTCH *van die N*

- (9) Ik heb van die rare ventjes gezien.  
 I have of those strange little\_guys seen  
 I saw (some) of those strange little guys.

FRENCH *de ces N*

- (10) J’ai vu de ces pommes.  
 I have seen of these apples

ITALIAN *di quei N*

- (11) A quel punto, lui ha fatto di quelle smorfie...  
 at that point, he has made of those faces  
 At that point, he made such (ugly/bizarre/...) faces...

On the diachronic side I showed – on the basis of a detailed historical simulation in EvolOT– that a bare partitive analysis would predict *des N* to appear at the same time an article analysis would.

As far as scope is concerned, the bare partitive analysis predicts *des N* to take narrow scope only. This fits the data in (8), so on that account the bare partitive analysis seems to fare better than the article analysis. However, wide scope readings for *des N* are not excluded, as illustrated in (12):

- (12) Il a refusé de collaborer avec des collègues de son équipe.  
 he has refused to collaborate with colleagues of his team  
<sup>ok</sup>*There are colleagues of his team he refused to work with.*

Given that *des N* marginally allows for wide scope, the bare partitive analysis faces the inverse problem of the article analysis

### *Conclusions*

The discussion in Part II of this dissertation led to two conclusions. The first is that DPs headed by indefinite articles and bare partitives are virtually indistinguishable. The only difference between the two arguably lies in their scopal behaviour: articles are expected to take scope freely whereas bare partitives are predicted to take narrow scope only. The second conclusion is that *des N* cannot unequivocally be analyzed as an indefinite article nor as a bare partitive. The reason for this is that it scopally behaves in a way that doesn't follow from either analysis. What this amounts to is that *des N* probably originated as a bare partitive, evolved in certain ways that brought it closer to an article DP, but never fully became one. I conclude that bare nominal arguments can be ungrammatical in a language without the language having indefinite articles and that the ungrammaticality of bare nominal arguments can never be used to argue in favour of the articlehood of a determiner.

### **3. Dutch *een* in predicate position**

In the third part of this dissertation I zoomed in on the Dutch indefinite article *een* and its use in predicate position. This study was motivated by the contradiction between the deeply rooted assumption that indefinite articles are used to mark arguments and the fact that indefinite articles are extremely frequent in predicate position as well. The analysis I proposed had three important ingredients: a lexical, a compositional and a diachronic one.

#### *The lexicon*

Following de Swart, Winter & Zwarts (2007) I assumed nouns are of type *e* and come in two subtypes: capacities and kinds. Capacities like *teacher*, *American* and *muslim* stand for the position of an individual in society, be it at a professional, civil or religious level. The crucial point is that they are culturally defined. Kinds like *wolf*, *rat* and *human* are exactly the opposite: they form natural classes of individuals. The term *natural* is opposed to *cultural* and reflects the intuition that kinds are different from capacities in two respects. The first is that kind membership is based on inherent properties: a wolf belongs to the kind *wolf* not by accident but because it has inherent properties that make it into a wolf. The second respect in which kinds are natural and capacities cultural is that kinds cannot be established by virtue of a cultural decision but exist by virtue of there being at least two individuals that share the same inherent properties.

#### *Compositional semantics*

Following de Swart, Winter & Zwarts (2007) I assumed there are two type-shifts turning expressions of type *e* into expressions of type  $\langle e, t \rangle$ . The first is *CAP*: it takes a capacity and returns the set of individuals that have this capacity. For a

capacity like *teacher* it would return the set of individuals that are appointed teachers. The second is *REL*: it takes a kind and returns the set of individuals that make up this kind. For a kind like *dog* it would return the set of all individuals that have the inherent properties that make it into dogs.

Under the assumption that *CAP* can apply covertly and that *REL* is linked to the appearance of the indefinite article, I predict that article predicates involve an application of *REL* and that bare predicates involve an application of *CAP*. I showed that this prediction – when supplemented with a worked out account of how kinds and capacities can be coerced one into the other – derives the interpretation of kind and capacity nouns in predicate position, whether they occur with the indefinite article or not. In what follows I briefly repeat the paradigmatic examples I analyzed.

- (13) Henriette is manager.  
Henriette is manager  
(14) Henriette is een manager.  
Henriette is a manager

(13) has the unmarked interpretation that Henriette belongs to the set of professional managers whereas (14) suggests that Henriette has the stereotypical properties we associate with the kind *manager* (being organized, able to delegate, ...) without having to be a professional manager.

- (15) Jantje is een raaf.  
Little\_John is a raven  
(16) Jantje is raaf.  
Little\_John is raven

(15) has the unmarked interpretation that Jantje is a member of the kind *raven* whereas (16) is only interpretable if we assume Jantje is playing the part of *raven* in a game (e.g. the game *rats and ravens*).

#### *Diachrony*

In the diachronic part I motivated the link between the indefinite article and *REL*. Crucial here were the assumption that the indefinite article in predicate position had a strong implicature of non-uniqueness and the assumption that sets corresponding to kinds necessarily have more than one element. When combined with general constraints on evolution these assumptions predict that the indefinite article was the only plausible candidate for marking *REL* in singular predicates.

#### *Extensions*

Even though the main goal of Part III was to investigate the Dutch indefinite article I also made specific suggestions about how the account can be extended e.g. to English *a*. The English case is slightly different from the Dutch one in the sense that *a* can only be omitted with capacities like *chairwoman* that refer to functions that are unique within a social setting:

(17) Hyacinth is chairwoman of the local Women's Institute.

### *Conclusions*

The discussion in Part III of this dissertation led to two conclusions. The first is that indefinite articles in predicate position may be recruited for new, related functions. The second is that the Dutch indefinite article was recruited as the realization of the type-shift *REL*.

### **4. What we now know about indefinite articles**

The picture of indefinite articles that emerged throughout this dissertation is that of a category of determiners that are poor in their semantics but rich in their pragmatic competition with bare nominals and other determiners. This picture is nothing more than a fine-grained version of the original *Blocking Principle* (Chierchia 1998): once we accept that indefinite articles and bare nominals can co-exist – be it in argument or in predicate position – we open a whole new range of possible interactions that are worthwhile exploring. The recently developed formal tools I exploited – OT, BiOT and EvolOT – turned out to be the perfect companions to make the exploration conjoin the insights of functionalism with the formal precision of the generative approach to linguistics.



## APPENDICES

### Appendix 1: Partitives and proper partitivity

Partitives are generally assumed to instantiate the proper part relation (see Barker 1998). If we look at partitives without an upstairs determiner this assumption cannot be maintained though (see e.g. Zamparelli 2008 on Italian *dei*, Roodenburg 2004 on French *des*, de Hoop et al. on Dutch *van die*). Indeed, the fact that the Martians in (1) are not necessarily telling a lie shows that partitives without an upstairs determiner do not instantiate proper (<) but standard partitivity (≤).

- (1) **Dei marziani** che sono atterrati nel mio giardino mi hanno detto che loro sono gli unici rappresentanti della loro specie.  
**of the Martians** that have landed in the my garden me have told that they are the only representatives of the their species  
 “Martians who landed in my backyard told me that they are the only representatives of their species.”

(based on Storto 2003)

The challenge then is to account for the difference between full and bare partitives while avoiding two unattractive assumptions: (i) bare partitives are radically different from full partitives, (ii) proper partitivity is encoded in the upstairs determiner and determiners such as *two* unambiguously have the following partitive semantics:

- (2)  $\lambda P \lambda Q \exists x (P(x) \& Q(x) \& \text{two}(x) \& \exists y (x \neq y \& P(y)))$

I will show that Zamparelli (1998, 2008) goes a long way in meeting the challenge but that his analysis cannot be the whole story yet. I then present my own analysis.

#### Zamparelli (1998, 2008)

Zamparelli’s analysis for a full partitive such as *two of the men* is that *two* combines with *men of the men* (with an upstairs copy of the downstairs noun) and that the latter part has the following subtraction semantics (in a model with three men):

- (3) 
$$\frac{\text{men} \quad \text{of the men} \quad \text{men of the men}}{\boxed{\{\{a\}, \{b\}, \{c\}, \{a,b\}, \{b,c\}, \{a,c\}, \{a,b,c\}\}} \quad \boxed{\{\{a,b,c\}\}} = \boxed{\{\{a\}, \{b\}, \{c\}, \{a,b\}, \{b,c\}, \{a,c\}\}}$$

Given that {a,b,c} contains all elements contained in all other subsets of *men* subtracting it gives us proper partitivity.

For bare partitives Zamparelli makes the standard assumption that the downstairs DP is kind-referring and proposes that the subtraction semantics is as follows:

$$(4) \quad \frac{\text{men} \quad \text{of the men}}{\boxed{\{\{a\},\{b\},\{c\},\{a,b\},\{b,c\},\{a,c\},\{a,b,c\}\}} \ominus \boxed{\{\{\mathbf{kindmen}\}\}} = \boxed{\{a\},\{b\},\{c\},\{a,b\},\{b,c\},\{a,c\},\{a,b,c\}}}$$

Given that the subtraction operation in (4) is a null operation Zamparelli correctly derives that bare partitives do not give rise to proper partitivity.

The problem Zamparelli faces is that his analysis predicts that partitives with kind-referring downstairs DPs never encode proper partitivity. This is however not the case as is shown by the unacceptability of (5):

- (5) \*the two of these lions ('the two instantiations of this kind of lions')

In (5) I use the standard diagnostics to check whether a full partitive instantiates proper partitivity or not. Adding an upstairs *the* should be fine in case of standard partitivity; there could be a unique plural individual consisting of two lions. In case there is proper partitivity adding *the* should be out because there can be no unique plural individual consisting of two lions.

### My proposal

My proposal consists of three parts. The first is that partitive *of* in both full and bare partitives denotes the standard part relation ( $\leq$ ). This goes back to earlier literature on partitives (Barwise & Cooper 1981, Keenan & Stavi 1986).

The second part is related to the semantics of those determiners that can appear in the upstairs D position of partitives. I propose they are ambiguous between a partitive and a non-partitive semantics. For *two* I already illustrated the partitive semantics in (2), its non-partitive counterpart is the same but without the part in bold. Note that this proposal can be motivated independently of partitives. Indeed, most indefinites are known to have at their disposal a non-partitive and a partitive semantics, the former being unmarked, the latter popping up in sentences such as (6) where the individual-level predicate forces a partitive reading:

- (6) Two students are intelligent.

With these first two parts in place I derive the impossibility of bare partitives to instantiate proper partitivity but I still need something more to explain why an upstairs determiner in a full partitive never has its unmarked non-partitive semantics. The answer lies in the pragmatics: if the upstairs determiner doesn't take its partitive semantics the referent of the whole partitive would necessarily be the same as the one of the downstairs DP. Indeed, *two of the men* would necessarily refer to the same plural individual as *the men*. Assuming that PartPs are dispreferred over DPs because of their syntactic complexity *the men* can then be taken to block *two of the men* on the non-partitive reading of *two*.

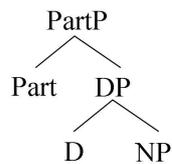
This pragmatic account is close to the one proposed by Ionin, Matushansky & Ruys (2006). The difference is that they assume the criterion for blocking is truth-conditional equivalence of a shorter expression whereas I assume the criterion is referential equivalence of an expression with a lower highest projection (DP and not PartP). My account cannot be distinguished from their's for standard partitives like *two of the men* but can be for full partitives with a downstairs kind-referring DP. Indeed, in a context in which one is talking about Berber lions the non-partitive reading of *two of these lions* cannot be blocked by *these two lions* which at first sight would be the shorter truth-conditionally equivalent expression. This is due to the fact that *these two lions* cannot pick-up the contextually introduced kind Berber lions and can therefore only refer to two contextually salient lions (cf. Zamparelli 2008 on the impossibility of definite + numeral + N to refer to kinds). My account however looks at referential equivalence and would correctly predict the non-partitive reading of *two of these lions* to be blocked by *these lions*.

With the above three parts in place I derive the impossibility of bare partitives to instantiate proper partitivity and the necessity of upstairs determiners to take on their partitive semantics. Crucial is that the proposed account meets the challenge: it accounts for the distinction between bare and full partitives without assuming that they are radically different and without assuming that the determiners that can appear in the upstairs D position of partitives are unambiguously partitive. An extra prediction the analysis makes and that plays an important role in this dissertation is that determiners that only have a non-partitive semantics cannot appear in the upstairs D position of partitives. Spanish *unos* is an example of this type of determiner.

### Appendix 2: The syntax of bare partitives

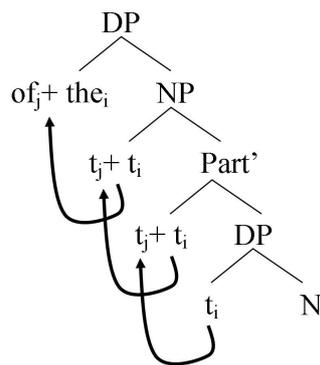
The syntax I assume for bare partitives is fairly simple and is represented in (1):

(1)

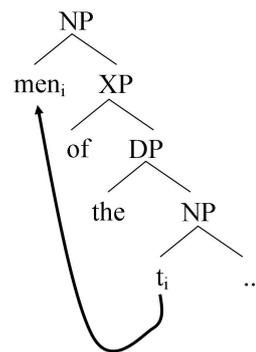


Two more complex bare partitive structures have been proposed in the literature. The first is the one by Chierchia (1997) who argues that Italian *dei* moves up to the upstairs D position as in (2). The problem for this kind of approach is that it is not clear why there is movement. Indeed, even if we grant that there is some clitic feature that licenses the movement from *the* to *of* it is not clear why *of\_the* would move any further. Chierchia claims this is because of ordinary  $\phi$  features of a determiner (gender and number) but fails to note that this presupposes *of\_the* is a determiner. The second bare partitive structure is the one of Zamparelli (2008) shown in (3). This structure is problematic in its own right. Indeed, if there were an upstairs silent copy of the downstairs noun as in (3), the upstairs noun would crucially be bare. For a language like Italian that allows for bare plurals this is fine but for a language like French that does not allow for bare plurals this is problematic. This is probably the reason why Zamparelli assumes *dei* moves up to D. This however raises the same question as in the case of Chierchia: why would there be movement to D?

(2)



(3)



Note that the structure Roodenburg (2004) proposes is not a real bare partitive structure: he takes the definite article to be the spell-out of number which makes his bare partitives more like pseudo-partitives.



## REFERENCES

- Abbott, B., 1996, 'Doing without a partitive constraint', in J. Hoeksema (ed.), *Partitives*, Berlin: Mouton de Gruyter, 25-56.
- Alonso-Ovalle, L. & P. Menéndez-Benito, 2002, 'Two types of weak quantifiers: Evidence from Spanish', in *Proceedings of the 1<sup>st</sup> North American Summer School in Logic, Language and Information*, Stanford, 11-19.
- Alonso-Ovalle, L. & P. Menéndez-Benito, 2009, handout of a talk given at Sinn und Bedeutung 14 'Distributed plurality in Spanish: *unos* vs. *algunos*'.
- Barker, C., 1998, 'Partitives, double genitives and anti-uniqueness', *Natural Language and Linguistic Theory* 16, 679-717.
- Barwise, J. & R. Cooper, 1981, 'Generalized quantifiers and natural language', *Linguistics and Philosophy* 4, 159-219.
- Beyssade, C. & C. Dobrovie-Sorin, 2005, 'A syntax-based analysis of predication', *Proceedings of Salt XV*.
- Beyssade, C. & C. Dobrovie-Sorin, 2009, *Predication and identity in copular sentences*, ms., Institut Jean Nicod & LLF, Paris 7.
- Blutner, R., 2000, 'Some aspect of optimality in natural language interpretation', *Journal of Semantics* 17, 189-216.
- Bosveld-de Smet, L., 1998, *On mass and plural quantification. The case of French des/du-NPs*, PhD Dissertation, University of Groningen.
- Brisson, C., 1998, *Distributivity, maximality, and floating quantifiers*, PhD Dissertation, Rutgers University.
- Carrier, A., 2007, 'From preposition to article', *Studies in Language* 31, 1-49.
- Carlson, G., 1977, *Reference to kinds in English*, PhD Dissertation, University of Massachusetts.
- Chierchia, G., 1997, 'Partitives, reference to kinds and semantic variation', *Proceedings of Salt VII*.
- Chierchia, G., 1998, 'Reference to kinds across languages', *Natural Language Semantics* 6, 339-405.
- Comorovski, 2004, 'Divers', in F. Corblin & H. de Swart (eds.), *Handbook of French Semantics*.
- Contreras, H., 1991, 'On the position of subjects', in S. Rothstein (ed.), *Perspectives on phrase structure*, New York: Academic Press, 63-79.
- Corblin, F., 2001, 'Où situer *certain*s dans une typologie sémantique des groupes nominaux?', in Kleiber et al. (eds.), *Typologie des groupes nominaux*, Rennes: Presses Universitaires de Rennes, 99-117.

- Dayal, V., 2004, 'Number marking and (in)definiteness in kind terms', *Linguistics and Philosophy* 27, 393-450.
- Delfitto, D. & J. Schroten, 1991, 'Bare plurals and the number affix in DP', *Probus* 3, 155-185.
- Dobrovie-Sorin C. & B. Laca, 2003, 'Noms sans déterminants', in D. Godard (ed.), *Les langues romanes, problèmes de la phrase simple*, Paris : CNRS Editions.
- Dobrovie-Sorin, C., 2004, 'Generic plural indefinites and (in)direct binding', in F. Corblin & H. de Swart (eds.), *Handbook of French semantics*, Stanford: CSLI Publications.
- Ebeling, C.L., 2006, *Semiotaxis*, Amsterdam: Amsterdam University Press.
- Englebert, A., 1996, *Le 'petit mot' de: étude de sémantique historique*, Genève: Droz.
- Espinal, M.T. & L. McNally, 2007, 'Bare singular nominals and incorporating verbs', in G. Kaiser & M. Leonetti (eds.), *Proceedings of the III NEREUS International Workshop*, Universität Konstanz *Arbeitspapier* 122: 45-62.
- Farkas, D. & H. de Swart, 2005, 'Généricité et indéfinitude', in C. Dobrovie-Sorin (ed), *Noms nus et généricité*, Paris: Presses Universitaires de Vincennes, 98-126.
- Farkas, D. & H. de Swart, 2007, 'Article choice in plural generics', *Lingua* 117, 1657-1676.
- Farkas, D. & H. de Swart, to appear, 'The semantics and pragmatics of plurals', to appear in *Semantics and Pragmatics*.
- Farkas, D. & K. Kiss, 2000, 'On the comparative and absolute readings of superlatives', *Natural Language & Linguistic Theory* 18: 417-455.
- Farkas, D., 2002, 'Varieties of indefinites', in *Proceedings of Salt XII*.
- Fauconnier, G., 1979, 'Implication reversal in natural language', in F. Guenther & S.J. Schmidt (eds.), *Formal semantics and pragmatics for natural languages*, Dordrecht: Reidel.
- Foulet, L., 1965, *Petite syntaxe de l'ancien français*, Paris: Champion.
- Geurts, B. & R. van der Sandt, 'Domain restriction', in P. Bosch & R. van der Sandt (eds.), *Focus: linguistic, cognitive and computational perspectives*, Cambridge: Cambridge University Press, 268-292.
- Giurgea, I., 2008, handout of a talk given at the Workshop on bare singulars, argument structure and their interpretation 'On Romanian predicative bare nouns'.
- Givón, T., 1981, 'On the development of the numeral 'one' as an indefinite marker', *Folia Linguistica Historica* 2, 35-53.

- Gutiérrez-Rexach, J., 1999, 'Spanish indefinites and type-driven interpretation', in J.M. Authier, B. Bullock & L. Reed (eds.), *Formal perspectives on romance linguistics*, Amsterdam: John Benjamins, 151-166.
- Gutiérrez-Rexach, J., 2001, 'The semantics of Spanish plural existential determiners and the dynamics of judgment types', *Probus* 13, 113-154.
- Gutiérrez-Rexach, J., 2003, *La semántica de los indefinidos*. Madrid: Visor Libros.
- Heim, I., 1982, *The Semantics of Definite and Indefinite Noun Phrases*, PhD Dissertation, University of Massachusetts, Amherst.
- Heim, I., 1991, 'Artikel und Definitheit', in A. von Stechow & D. Wunderlich (eds.), *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*, Berlin : de Gruyter, 487-535.
- Heim, I., 1999, *Notes on superlatives*, Ms., MIT.
- Herdan, S. & Y. Sharvit, 2006, 'Definite and nondefinite superlatives and NPI licensing', *Syntax* 9, 1-31.
- Heyd, S., 2003. *L'interprétation des syntagmes nominaux en "des" et "de" en position sujet et objet - Généricité, habitualité et incorporation sémantique*, PhD Dissertation, Université Marc Bloch.
- Hoop, H. de, 1998, 'Partitivity', *Glott International* 3, 3-10.
- Horn, L.R., 1985, 'Metalinguistic negation and pragmatic ambiguity', *Language* 61, 121-74.
- Huddleston, R. & G.K. Pullum, 2002, *The Cambridge grammar of the English language*, Cambridge: Cambridge University Press.
- J. Hoeksema (ed.), *Partitives*, Berlin: Mouton de Gruyter.
- Jäger, G., 2002, 'EvolOT: software for simulating language evolution using Stochastic Optimality Theory', ms., University of Potsdam.
- Jayez, J., 2006, 'How many are 'several'? Argumentation, Meaning and Layers', *Belgian Journal of Linguistics* 19, 187-209.
- Kamp, H. & U. Reyle, 1993, *From Discourse to Logic*, Dordrecht: Kluwer.
- Keenan, E. & J. Stavi, 1986, "A semantic characterization of natural language determiners." *Linguistics and Philosophy* 9, 253-326.
- Kiparsky, P., 1998, 'Partitive case and aspect', in M. Butt & W. Geuder (eds.), *The projection of arguments*, Stanford: CSLI, 265-307.
- Kratzer, A., 1996, *Scope or Pseudoscope? Are There Wide Scope Indefinites?*, ms., University of Massachusetts.
- Krifka, M., 1989, 'Nominal reference, temporal constitution, and quantification in event semantics.', in R. Bartsch, J. van Benthem & P. van Emde Boas (eds.), *Semantics and contextual expressions*. Foris: Dordrecht.

- Krifka, M., 2004, "Bare NPs: kind-referring, indefinites, both, or neither?", in O. Bonami & P. Cabredo Hofherr (eds.), *Empirical Issues in Formal Syntax and Semantics* 5, available on [www.cssp.cnrs.fr](http://www.cssp.cnrs.fr).
- Kupferman, L., 1976, *Études sur l'article*, PhD Dissertation, University of Paris VIII.
- Kupferman, L., 1979, 'L'article partitif existe-t-il?', *Le français moderne*, p. 1-16.
- Kupferman, L., 1994, 'Du: un ature indéfini?', *Faits de langue* 4, 195-203.
- Kupferman, L., 1998, 'Des: pluriel de du?', in M. Bilger, K. Van den Eynde & F. Gadet (eds.), *Analyse linguistique et approches de l'oral*, Leuven: Peeters.
- Kupferman, L., 2001, 'Quantification et détermination dans les groupes nominaux', in X. Blanco et al. (eds.), *Détermination et Formalisation*, Amsterdam/Philadelphia: Benjamins, 219-234.
- Laca, B & L. Tasmowski, 2004, 'Différents', in F. Corblin & H. de Swart (eds.), *Handbook of French Semantics*.
- Laca, B. & L. Tasmowski, 1993 (1994). 'Le pluriel indéfini de l'attribut métaphorique', *Linguisticae Investigationes* 18, 27-48.
- Laca, B., 1996, 'Acerca de la semántica de los plurales escuetos del español', in I. Bosque (ed.), *El Sustantivo sin Determinación*, Madrid: Visor Libros, 241-268.
- Lapesa, R. 1975, 'Dos estudios sobre la actualización del sustantivo en español', en *Boletín de la Comisión Permanente de la Real Academia Española* 21, 3-31.
- Lascarides, A. & N. Asher, 1993, 'Temporal interpretation, discourse relations, and common sense entailment', *Linguistics and Philosophy* 16, 437-493.
- Le Bruyn, B., 2005, *Looking for the bare necessities*, MA thesis, Utrecht University.
- Le Bruyn, B., 2007, 'The partitive article dei in Italian', *Linguistics in the Netherlands* 24, 138-149.
- Longobardi, G., 1994, 'Reference and proper names', *Linguistic Inquiry* 25, 609-665.
- Lopez-Palma, H., 2007, 'Plural indefinite descriptions with *unos* and the interpretation of number', *Probus* 19, 235-266.
- Mari, A. & F. Martin, 2008, 'Bare and indefinite NPs in predicative position in French', Working Papers of the SFB 732 *Incremental Specification in Context* 1, 119-144.
- Mari, A. & F. Martin, 2008, talk given at Going Romance 2008 'Generic indefinite *des*: groups, sums and classes'.
- Mari, A. & F. Martin, 2009, talk given at the workshop *Genericity: Interpretation and Uses* 'Generic (in)definites and the analytic/synthetic distinction'.

- Mari, A. & F. Martin, 2009, *To be (an) idiot in French. Systematic polysemy in copular sentences* (abstract).
- Martí, L., 2008, 'The semantics of plural indefinite noun phrases in Spanish and Portuguese', *Natural Language Semantics* 16, 1-37.
- Mattausch, J., 2007, 'Optimality, bidirectionality and the evolution of binding phenomena', *Research on language and computation* 5, 103-131.
- McNally, L., 2004, 'Bare plurals in Spanish are interpreted as properties', *Catalan Journal of Linguistics* 3, 115-133.
- Melis, C., M. Flores & S. Bogard, 2003, 'La historia del español: propuesta de un tercer período evolutivo', *Nueva Revista de Filología Hispánica* 51, 1-56.
- Morin, Y.-C., 1976, 'De ces: the lexicalization of a syntactic expression', *Linguistic Inquiry* 7, 706-707.
- Munn, A. & C. Schmitt, 2005, 'Number and indefinites', *Lingua* 115, 821-855.
- Olarrea, A., 1996, *Pre- and Postverbal Subject Positions in Spanish*, PhD dissertation, University of Washington.
- Oosterhof, A., 2005, 'Dutch bare plurals, faded partitives and subject-object asymmetry', *Belgian Journal of Linguistics* 19, 59-91.
- Partee, B., 1987, 'Noun phrase interpretation and type-shifting principles', in J. Groenendijk, D. de Jongh & M. Stokhof (eds.) *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, Dordrecht: Foris, 115-143.
- Poesio, M. & A. Zucchi, 'On telescoping', in *Proceedings of Salt II*.
- Potts, C., 2004, 'Conventional implicatures, a distinguished class of meanings', in G. Ramchand & C. Reiss, *The Oxford Handbook of Linguistic Interfaces*, Oxford: Oxford University Press.
- Pozas-Loyo, J., to appear, *The development of the indefinite article in medieval and golden-age Spanish*, PhD Dissertation, Queen Mary, University of London.
- Reed, A., 1996, 'Partitives, existentials, and partitive determiners', in J. Hoeksema (ed.), *Partitives*, Berlin: Mouton de Gruyter, 143-178.
- Roberts, C., 1989, 'Modal subordination and pronominal anaphora in discourse', *Linguistics and Philosophy* 12, 683-721.
- Roodenburg, J., 2004, *Pour une approche scalaire de la déficience nominale: la position du français dans une théorie des "noms nus"*, PhD Dissertation, Universities of Amsterdam and Paris 8.
- Roy, I., 2006, *Non-verbal predications. A syntactic analysis of predicational copular sentences*, PhD Dissertation, University of Southern California.
- Sauerland, U. et al., 2005, 'The plural is semantically unmarked', in S. Kepser & M. Reis (eds.), *Linguistic Evidence*, Berlin: Mouton-De Gruyter.

- Smits, F., 2009, *La portée dans le domaine du possible*, MA thesis, Utrecht University.
- Spector, B., 2007, 'Aspects of the pragmatics of plural morphology: On higher-order implicatures', in U. Sauerland & P. Stateva (eds.), *Presuppositions and implicatures in compositional semantics*, Palgrave-Macmillan, 243-281.
- Stark, E., 2002, 'Indefiniteness and specificity in Old Italian texts', *Journal of Semantics* 19, 315-332.
- Stateva, P., 2004, *Presuppositions in superlatives*. Ms., ZAS Berlin.
- Storto, G., 2003, 'On the status of the partitive determiner in Italian', in *Selected Papers from Going Romance 2001*, Amsterdam: John Benjamins, 315-330.
- Swart, H. de & J. Zwarts, 'Article use across languages: an OT typology', in A. Grønn (ed.), *Proceedings of Sinn und Bedeutung 12*, 628-644.
- Swart, H. de, 2001, 'Weak readings of indefinites: type-shifting and closure', *The Linguistic Review* 18: 69-96.
- Swart, H. de, J. Zwarts & Y. Winter, 2007, 'Bare nominals and reference to capacities'. *Natural Language and Linguistic Theory* 25.
- Szabolcsi, A., 1986, 'Comparative superlatives', in N. Fukui, T.R. Rapoport & E. Sagey (eds.), *Papers in Theoretical Linguistics*, Cambridge: MITWPL, 245-265.
- Van Geenhoven, V., 1998, *Semantic incorporation and indefinite descriptions*, Stanford: CSLI Publications.
- Van Peteghem, M., 2003, *La détermination de l'attribut nominal. Étude comparative de quatre langues romanes (français, espagnol, italien, roumain)*. Brussel: Koninklijke Academie voor Wetenschappen, Letteren en Schone Kunsten van België.
- Villalta, E., 1994, *Plural indefinites in Spanish and distributivity* [unpublished manuscript].
- Vries, H. de, 2009, *Scopal behaviour of English bare plurals*, Mphil Internship report, Utrecht University.
- Walker, M., A. Joshi & E. Prince (eds.), 1998, *Centering Theory in discourse*, Oxford: Clarendon Press.
- Wright, S. & T. Givón, 1987, 'The pragmatics of indefinite reference: Quantified text-based studies', *Studies in Language* 11, 1-33.
- Zagona, K., 2002, *The syntax of Spanish*. Cambridge: Cambridge University Press.
- Zamparelli, R., 1998, 'A theory of kinds, partitives and OF/Z possessives', in A. Alexiadou & C. Wilder (eds.), *Possessors, predicates and movement in the determiner phrase*. Amsterdam : John Benjamins.

- Zamparelli, R., 2008, 'Bare predicate nominals in Romance languages', in H. Müller & Klinge, *Essays on nominal determination*, John Benjamins.
- Zamparelli, R., 2008, 'Dei ex machina', *Studia Linguistica* 62, 301-327.
- Zeevat, H., 2005, *Simulating recruitment in evolution*, ms., University of Amsterdam.
- Zribi-Hertz, A., 2003, 'Pour une analyse unitaire de DE partitif', in L. Kupferman, et al. (eds.), *Indéfinis et prédictions en français*, Paris: Presses Universitaires de la Sorbonne.



## SAMENVATTING IN HET NEDERLANDS

Het doel van deze dissertatie is tot een fijnmazigere definitie te komen van de notie *indefiniët lidwoord*. De dissertatie combineert formele semantiek en pragmatiek, presenteert uitdagende nieuwe corpus data – zowel diachroon als synchroon – en stelt simulaties voor van historische processen die – tot voor kort – enkel beschreven konden worden. In deze samenvatting schets ik de grote argumentatielijnen en de belangrijkste conclusies.

### 1. Het Spaanse *unos*

In het eerste deel van deze dissertatie zoom ik in op het Spaanse *unos* en beargumenteer ik dat het geanalyseerd moet worden als een meervoudig indefiniët lidwoord. De argumentatie is zowel synchroon als diachroon.

#### *Synchroon*

Voor het synchrone gedeelte start ik van de standaard aanname dat indefiniëte lidwoorden ‘default’ determinatoren zijn in de zin dat ze geen semantische bijdrage leveren op het markeren van argumenten na. Om te beargumenteren dat *unos* een indefiniët lidwoord is toon ik aan dat het verschillende semantische ingrediënten mist die vrij standaard zijn voor andere indefiniëte determinatoren: partitiviteit, de mogelijkheid om distributiviteit uit te drukken, een hoeveelheidsspecificatie, ... Voorbeelden (1) tot (3) illustreren deze eigenschappen (of althans het gebrek daaraan):

- (1) Unos estudiantes vinieron a verme.  
UNOS studenten kwamen om zien\_me  
<sup>ok</sup>*Enkele studenten zijn naar mij gekomen.*  
*#Enkele studenten zijn naar mij gekomen, anderen niet..*
- (2) Unos hombres compraron un billete de lotería.  
UNOS mannen kochten een biljet van loterij  
<sup>pref</sup>*Enkele mannen kochten tesamen een loterijbiljet.*  
<sup>dispref</sup>*Enkele mannen kochten elk een loterijbiljet.*
- (3) ¿Cuántos libros tienes? ??Unos.  
Hoeveel boeken heb\_jij UNOS

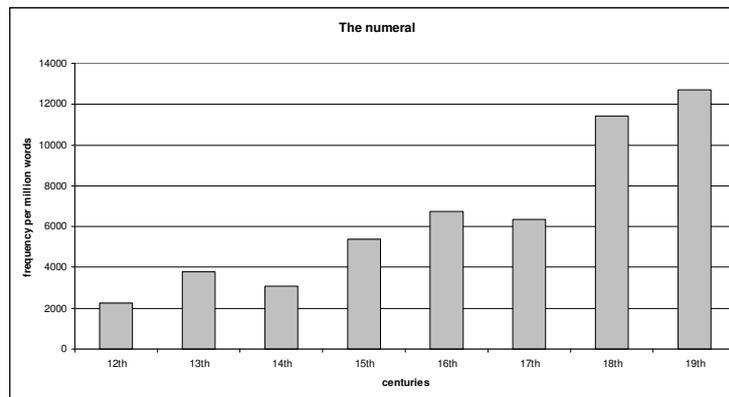
Ik beargumenteer ook dat *unos N* zich enkel in de pragmatiek van kale meervouden onderscheidt: de eerstgenoemde wordt meestal gebruikt om saillante discourse referenten te introduceren terwijl de laatstgenoemde gebruikt wordt voor niet-saillante discourse referenten. Ik gebruik dit pragmatisch verschil – dat ik analyseer in Weak Bi-directional Optimality Theory – om uit te leggen waarom indefiniëte lidwoorden kunnen bestaan naast kale naamwoord argumenten, waarom *unos* niet in het bereik van negatie kan vallen en waarom kale meervouden niet kunnen voorkomen in preverbale onderwerpspositie. (4) tot (6) illustreren deze feiten:

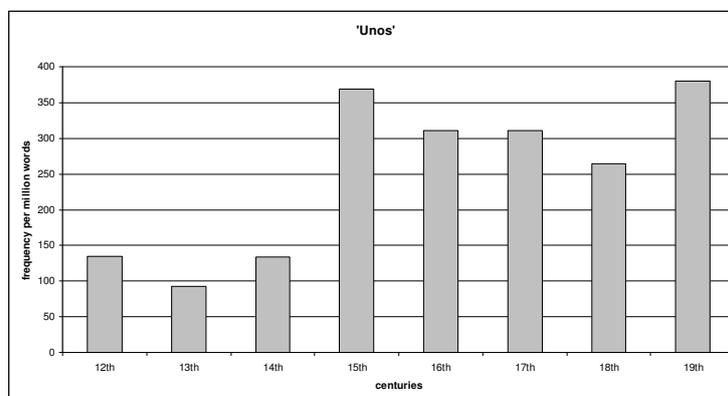
- (4) María comió bizcochos | unos bizcochos.  
María at biscuits | UNOS biscuits
- (5) A la reunión no asistieron unos profesores.  
Op de vergadering niet bijwonen unos professors  
<sup>ok</sup>*Er zijn enkele professoren die de vergadering niet bijwoonden.*  
*#Geen enkele professor woonde de vergadering bij.*
- (6) \*Políticos han ocupado el palacio.  
Politici hebben bezet het paleis

### Diachroon

Voor het diachroon gedeelte start ik van de observatie in de grammaticalisatie literatuur dat de ontwikkeling van indefiniete lidwoorden in het algemeen door drie fasen gaat: in de eerste fase komen argumentsnaamwoorden kaal voor, in de tweede fase worden indefiniete lidwoorden gebruikt om saillante discourse referenten te introduceren en in de derde en laatste fase wordt het gebruik van indefiniete lidwoorden gegeneraliseerd naar alle argumenten. *Unos* is een beetje een uitzondering in de zin dat het niet tot de derde fase is doorgedrongen. Ik laat dit zien op basis van de corpusdata in (7). De eerste grafiek toont de relatieve frekwentie van het indefiniet lidwoord van de 12de tot de 19de eeuw. De tweede grafiek toont dezelfde data maar dan voor *unos*.

(7)





Wat de grafieken duidelijk laten zien is dat het enkelvoudige indefiniete lidwoord tweemaal een belangrijke stijging in frequentie heeft gekend terwijl dit voor *unos* maar éénmaal het geval is. Dit suggereert dat het enkelvoudig indefiniet lidwoord de drie fasen van lidwoordontwikkeling heeft doorlopen terwijl *unos* is blijven steken in de tweede. De vraag die dan gesteld moet worden is of tweede fase determinatoren al echte lidwoorden zijn. Het antwoord dat ik voorstel is bevestigend. Het eerste argument voor dit antwoord komt van het feit dat *unos* in de overgang van fase 1 naar fase 2 een semantische verandering ondergaan heeft: op basis van kwalitatieve corpus data toon ik aan dat *unos* partitieve lezingen toeliet voor het fase 2 bereikte. Het tweede argument komt van de synchrone feiten: er zijn geen verschillen tussen kale meervouden en *unos* *N* die niet kunnen verklaard worden op basis van pragmatische competitie. Dit is een indicatie dat *unos* een echte niet-gemarkeerde determinator is. Het derde en laatste argument komt van een simulatie van de evolutie van indefiniete lidwoorden in EvolOT: ik laat zien dat hun evolutie kan gesimuleerd op zo'n manier dat de evolutie verschillende eindfasen toelaat (fase 2 voor *unos* en fase 3 voor indefiniete lidwoorden in het algemeen). Dit betekent dat de gezamenlijke overgang van het enkelvoudig indefiniet lidwoord en *unos* in de 15de eeuw een echte transitie is van een standaard determinator naar een indefiniet lidwoord.

### Conclusies

De discussie in het eerste deel van deze dissertatie leidt tot drie conclusies. De eerste is dat – in synchronie – DPs ingeleid door indefiniete lidwoorden bepaalde eigenschappen missen die andere DPs wel hebben (partitiviteit, distributiviteit, hoeveelheid) en zich enkel in de pragmatiek van kale naamwoord argumenten onderscheiden. De tweede conclusie is dat – in diachronie – lidwoorden verwacht worden een bepaald frequentiepatroon te hebben en bepaalde semantische veranderingen te ondergaan. De derde en laatste conclusie is dat *unos*, op basis van zijn synchrone en diachrone eigenschappen, geanalyseerd moet worden als een indefiniet lidwoord. Dit betekent dat indefiniete lidwoorden en kale naamwoord argumenten naast elkaar kunnen bestaan en dat het bestaan van indefiniete

lidwoorden nooit kan gebruikt worden als een voldoende verklaring voor de ongrammaticaliteit van kale naamwoordargumenten.

## 2. Het Franse *des*

In het tweede deel van deze dissertatie zoom ik in op het Franse *des*. Ik start met de observatie dat *des N* op verschillende wijze geanalyseerd kan worden. De analyses waar ik het meeste aandacht aan besteed zijn de DP analyse waarin *des* als een indefiniet lidwoord functioneert en de kale partitief analyse waarin *des* wordt opgesplitst in *de* ('van') en *les* ('de<sub>meervoud</sub>')

### *De meervoudig indefiniet lidwoord analyse*

*Des* analyseren als een meervoudig indefiniet lidwoord is zeer aantrekkelijk in twee opzichten. Het eerste is dat deze analyse op een eenvoudige manier de meervoudige en indefiniete interpretatie van *des N* voorspelt. Het tweede is dat de analyse ook mooi voorspelt dat het historische verschijnen van *des N* in het Frans min of meer samenvalt met de generalisatie van het gebruik van het enkelvoudige indefiniete lidwoord.

Het belangrijkste probleem voor een lidwoord analyse van *des N* is dat indefiniete lidwoorden in principe zowel klein als wijd bereik kunnen nemen en *des N* een duidelijke voorkeur heeft voor klein bereik:

- (8) Il me faut des chemises.  
Er mij moet DES hemden  
I heb hemden nodig.

De sterk geprefereerde lezing van (8) is dat ik hemden nodig heb, eender welke. Onder de aanname dat lidwoorden vrij zijn in bereik is deze sterke voorkeur onverwacht.

### *De kale partitief analyse*

Wat de kale partitief analyse nog aantrekkelijker maakt dan de lidwoord analyse is dat – naast het feit dat het dezelfde feiten kan verklaren – het de complexe morfologie van *des* serieus neemt. Aan de synchrone kant beargumenteer ik dat een kale partitief analyse de meervoudige en indefiniete interpretatie van *des N* afleidt zonder *ad hoc* aannames. Ik leid de meervoudige interpretatie af op basis van default reasoning, een variant van 'maximize presupposition' (zie Heim 1991 en Sauerland et al. 2003). De indefiniete interpretatie leid ik af op basis van een Weak Bi-directional OT analyse die ervoor zorgt dat de definitieve DP in *des N* noodzakelijk verwijst naar soorten. Voorts laat ik ook zien dat de zelfde pragmatische aannames nodig zijn om andere kale partitieven te analyseren: het Nederlandse *van die N*, het Franse *de ces N* en het Italiaanse *di quei N*. Cruciaal hierbij is dat geen van deze uitdrukkingen als lidwoord DPs kunnen geanalyseerd worden.

DUTCH *van die N*

(9) Ik heb van die rare ventjes gezien.

FRENCH *de ces N*

(10) J'ai vu de ces pommes.

I heb gezien van die appels

ITALIAN *di quei N*

(11) A quel punto, lui ha fatto di quelle smorfie...

op dat moment, hij heeft gemaakt van die gezichten

Op dat moment trok hij van die gezichten...

Aan de diachrone kant toon ik – op basis van een gedetailleerde historische simulatie in EvolOT – dat een kale partitief analyse voorspelt dat *des N* in het Frans verschijnt op hetzelfde moment als een lidwoord analyse zou voorspellen.

Wat bereik betreft, voorspelt een kale partitief analyse dat *des N* enkel klein bereiklezingen toelaat. Dit komt overeen met wat we zien in (8) en het lijkt dan dat de kale partitief analyse op dit punt beter is dan de lidwoord analyse. Echter, wijd bereik lezingen van *des N* zijn wel degelijk mogelijk:

(12) Il a refusé de collaborer avec des collègues de son équipe.

Hij heeft geweigerd van samenwerken met DES collega's van zijn team

<sup>ok</sup>*Er zijn collega's van zijn team met wie hij weigerde samen te werken.*

Aangezien *des N* uitzonderlijk ook wijd bereiklezingen toelaat heeft de kale partitief analyse het omgekeerde probleem van de lidwoord analyse.

*Conclusies*

De discussie in het tweede deel van deze dissertatie leidt tot twee conclusies. De eerste is dat DPs ingeleid door een meervoudig indefiniet lidwoord vrijwel ononderscheidbaar zijn van kale partitieven. Het enige verschil tussen de twee ligt waarschijnlijk in hun bereikeigenschappen: lidwoorden zijn vrij in het nemen van klein en wijd bereik en kale partitieven laten enkel klein bereik toe. De tweede conclusie is dat *des N* niet éénduidig als een lidwoord of een kaal partitief kan geanalyseerd worden. De reden hiervoor is dat zijn bereikeigenschappen uit geen enkele analyse volgen. Wat dit suggereert is dat *des N* ontstaan is als een kale partitief, op een bepaalde manier zich heeft ontwikkeld in de richting van een lidwoord DP, maar er zelf nooit één geworden is. De conclusie die ik trek is dat kale naamwoord argumenten ongrammaticaal kunnen zijn in een taal zonder dat deze taal indefiniete lidwoorden moet hebben en dat de ongrammaticaliteit van kale naamwoord argumenten nooit als een doorslaggevend argument gebruikt kan worden voor het bepalen van de lidwoordstatus van een determinator.

**3. Het Nederlandse *een* in predikaatspositie**

In het derde deel van deze dissertatie zoom ik in op het Nederlandse indefiniete lidwoord *een* en zijn gebruik in predikaatspositie. De motivatie voor deze studie komt van de contradictie tussen de diepewortelde aanname dat indefiniete

lidwoorden argumentsmarkeerders zijn en het feit dat indefiniete lidwoord ook extreem frequent zijn in predikaatspositie. De analyse die ik voorstel bestaat uit drie belangrijke ingrediënten: een lexicaal, een compositioneel en een diachroon.

#### *Het lexicon*

In navolging van de Swart, Winter & Zwarts (2007) neem ik aan dat naamwoorden van het type *e* zijn en kunnen onderverdeeld worden in twee subtypes: capaciteiten en soorten. Capaciteiten zoals *leraar*, *Amerikaan* en *moslim* staan voor de positie van een individu in de samenleving, zij het op professioneel, burgerlijk of religieus niveau. Het cruciale punt is dat capaciteiten cultureel gedefinieerd zijn. Soorten als *wolf*, *rat* en *mens* zijn exact het tegenovergestelde: zij vormen natuurlijk klassen van individuen. De term *natuurlijk* moet tegenover de term *cultureel* geplaatst worden en is een uitdrukking van de intuïtie dat capaciteiten en soorten in twee opzichten verschillen. Het eerste is dat het behoren tot een soort is gebaseerd op inherente eigenschappen: een wolf behoort tot de soort *wolf* omdat hij de inherente eigenschappen heeft van een wolf en niet omdat iemand beslist heeft dat het een wolf zou zijn. Het tweede opzicht waarin capaciteiten cultureel zijn en soorten natuurlijk is dat soorten niet bestaan omwille van een culturele beslissing maar wel omwille van het feit dat er meerdere individuen zijn die dezelfde inherente eigenschappen hebben.

#### *Compositionele semantiek*

In navolging van de Swart, Winter & Zwarts (2007) neem ik aan dat er twee type-shifts zijn die uitdrukkingen van het type *e* omzetten in uitdrukkingen van het type  $\langle e, t \rangle$ . De eerste is *CAP*: deze neemt een capaciteit en zet deze om in de verzameling van individuen die deze capaciteit hebben. Voor een capaciteit als *leraar* levert dit de verzameling van professionele leraren op. De tweede is *REL*: deze neemt een soort en zet deze om in de verzameling van individuen die behoren tot deze soort. Voor een soort als *hond* levert de verzameling van alle honden op. Onder de aanname dat *CAP* onzichtbaar kan toegepast worden en dat de toepassing van *REL* vasthangt aan het verschijnen van het indefiniete lidwoord, voorspel ik dat lidwoord predikaten een toepassing van *REL* insluiten en dat kale predikaten een toepassing van *CAP* insluiten. Ik toon dat deze voorspelling – wanneer ze wordt verrijkt met een uitgewerkt verhaal van hoe capaciteiten en soorten ook onderling kunnen shiften – de correcte interpretatie afleidt voor soort en capaciteit naamwoorden in predikaatspositie, of deze nu met of zonder lidwoord voorkomen. In wat volgt zet ik de paradigmatische voorbeelden nog even op een rijtje.

- (13) Henriette is manager.  
 (14) Henriette is een manager.

(13) heeft de ongemarkeerde interpretatie dat Henriette behoort tot de verzameling van professionele managers. (14) daarentegen suggereert dat Henriette de stereotiepe eigenschappen van een manager heeft (georganiseerd zijn, in staat te delegeren, ...) zonder daar een professionele manager voor te moeten zijn.

- (15) Jantje is een raaf.  
 (16) Jantje is raaf.

(15) heeft de ongemarkeerde interpretatie dat Jantje behoort tot de soort raven terwijl (16) enkel interpreteerbaar is als we aannemen dat Jantje de rol speelt van *raaf* in een spel (zoals het *ratten en raven spel*).

### *Diachronie*

In het diachroon gedeelte motiveer ik de link tussen het indefiniete lidwoord en *REL*. Cruciaal hierbij zijn de aanname dat indefiniete lidwoorden in predikaatpositie een sterk niet-uniciteitsimplicatuur hebben en de aanname dat verzamelingen die overeenkomen met soorten meer dan één element hebben. Als we deze twee aannames combineren met standaard restricties op evolutie voorspellen we dat het indefiniete lidwoord de enige mogelijke kandidaat was voor het markeren van *REL* in predikaatpositie.

### *Uitbreidingen*

Ondanks het feit dat deel III vooral gaat over het Nederlandse indefiniete lidwoord maak ik ook enkele specifieke suggesties over hoe de analyse kan uitgebreid worden naar bijv. het Engelse indefiniete lidwoord. Het Engelse *a* verschilt licht van het Nederlandse *een* in de zin dat het enkel kan weggelaten worden voor unieke capaciteiten zoals *voorzitster* die verwijzen naar functies die uniek zijn in een bepaalde sociale setting:

- (17) Hyacinth is chairwoman of the local Women's Institute.  
 Hyacinth is voorzitster van het lokale Women's Institute

### *Conclusies*

De discussie in het derde deel van deze dissertatie leidt tot twee conclusies. De eerste is dat indefiniete lidwoorden in predikaatpositie gerekruteerd kunnen worden voor nieuwe, gerelateerde functies. De tweede is dat het Nederlandse indefiniete lidwoord gerekruteerd werd als realisatie van de *REL* type-shift.

## **4. Wat we nu weten over indefiniete lidwoorden**

Het beeld van indefiniete lidwoorden dat in deze dissertatie wordt gepresenteerd is dat van een klasse determinatoren die arm zijn in hun semantiek maar rijk in hun pragmatische competitie met kale naamwoorden en andere determinatoren. Dit beeld is niets meer dan een fijnmazigere versie van het oorspronkelijk *Blocking Principle* (Chierchia 1998): wanneer we aanvaarden dat indefiniete lidwoorden en kale naamwoorden naast elkaar kunnen bestaan – in argument of in predikaatpositie – openen we een heel nieuw domein van mogelijke interacties die het waard zijn om ontdekt te worden. De recent ontwikkelde formele tools die ik heb gebruikt – OT, BiOT en EvolOT – blijken op deze ontdekkingsstocht de perfecte metgezellen te zijn

om de inzichten van het functionalisme te koppelen aan de formele precisie van de generatieve taalkunde.

## RÉSUMÉ EN FRANÇAIS

Cette thèse présente trois études de cas qui ont comme but principal d'approfondir notre conception de la notion d'article indéfini. Elle combine la sémantique et la pragmatique formelle, présente de nouvelles données synchroniques et diachroniques tirées de corpus et fournit des simulations de processus historiques qui – jusqu'à maintenant – ne pouvaient être que décrites. Dans ce résumé je présente les grandes lignes de l'argumentation et les conclusions les plus importantes.

### 1. *Unos* espagnol

Dans la première partie de cette thèse je propose d'analyser *unos* comme un article indéfini. L'argumentation est tant synchronique que diachronique.

#### *Synchronie*

Pour la partie synchronique je pars de l'hypothèse que les articles indéfinis sont des déterminants 'par défaut' dans le sens qu'ils n'ont pas de contribution sémantique au-delà du marquage d'arguments. Pour soutenir que *unos* est un article indéfini je montre que plusieurs ingrédients sémantiques lui manque qu'on pourrait croire standards pour les déterminants indéfinis: la partitivité, la capacité d'induire des dépendances distributives, une spécification de quantité, ... Les exemples (1) à (3) illustrent ces caractéristiques (ou plutôt leur manque)

- (1) *Unos* estudiantes vinieron a verme.  
UNOS étudiants vinrent à voir\_me  
<sup>ok</sup>*Quelques étudiants vinrent me voir.*  
*#Quelques étudiants vinrent me voir, d'autres non.*
- (2) *Unos* hombres compraron un billete de lotería.  
UNOS hommes achetèrent un billet de loterie  
<sup>pref</sup>*Quelques hommes ont collectivement acheté un billet de loterie.*  
<sup>dispref</sup>*Quelques hommes ont chacun acheté un billet de loterie.*
- (3) ¿Cuántos libros tienes? ??Unos.  
Combien livres as-tu UNOS

Je défends également que le pluriel nu et *unos* *N* ne se distinguent qu'au niveau de la pragmatique: le premier s'emploie en général pour introduire des référents de discours non-saillants tandis que le second s'emploie pour introduire des référents saillants. J'exploite cette différence pragmatique – que je formalise en Weak Bi-directional Optimality Theory – pour expliquer comment il se fait que les articles indéfinis peuvent co-exister avec les noms nus argumentaux, pourquoi *unos* ne peut pas prendre portée sous la négation et pourquoi les pluriels nus ne peuvent pas apparaître en position de sujet préverbal. Ces faits sont illustrés de (4) à (6):

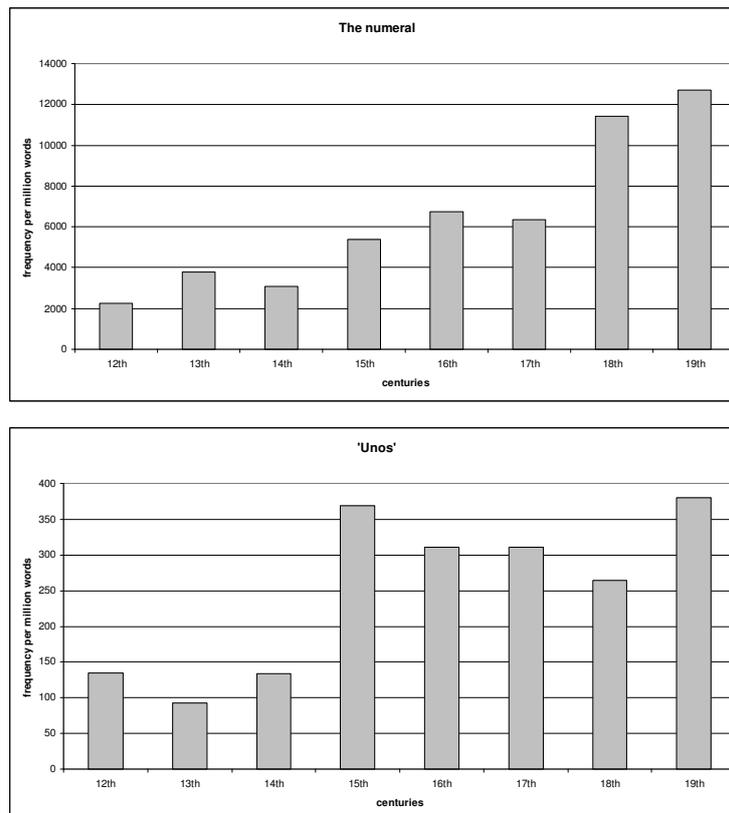
- (4) María comió bizcochos | *unos* bizcochos.  
María mangea biscuits | UNOS biscuits

- (5) A la reunión no asistieron unos profesores.  
 À la réunion NEG assistèrent UNOS professeurs  
<sup>ok</sup>*Il y a quelques profs qui n'ont pas assisté à la réunion.*  
 #Aucun prof n'a assisté à la réunion.
- (6) \*Políticos han ocupado el palacio.  
 Politiciens ont occupé le palais

### Diachronie

Pour la partie diachronique je pars de l'observation que le développement des articles indéfinis se divise en trois étapes: dans la première les noms argumentaux peuvent apparaître nus, dans la deuxième les articles indéfinis commencent à s'employer pour introduire les référents de discours saillants et dans la troisième et dernière étape les articles indéfinis se généralisent à tous les arguments. *Unos* est un peu particulier parce qu'il n'est jamais allé au-delà de la deuxième étape. J'ai illustré cette particularité de *unos* sur la base de données de corpus reprises en (7). Le premier graphique présente la fréquence relative de l'article indéfini singulier du 12<sup>ème</sup> au 19<sup>ème</sup> siècle. Le second graphique présente les mêmes données pour *unos*.

(7)



Les graphiques montrent que l'article indéfini singulier a subi deux montées en fréquence tandis que *unos* n'en a subi qu'une seule. Ceci suggère que l'article indéfini singulier a parcouru les trois étapes du développement des articles tandis que *unos* n'en a parcouru que deux. La question qui s'impose alors est de savoir si les déterminants dont le développement s'arrête à la deuxième étape sont de vrais articles ou non. La réponse que je propose est affirmative. Le premier argument en faveur de cette réponse est que *unos* a subi un changement sémantique lors de sa transition de la première à la deuxième étape: sur la base de données de corpus je montre que *unos* donnait lieu à des lectures partitives avant d'atteindre l'étape 2. Le deuxième argument vient des faits synchroniques: il n'y aucune différence entre les pluriels nus et *unos N* qui ne puisse être expliquée sur la base d'une compétition pragmatique. Ceci indique que *unos N* est un véritable déterminant non-marqué. Le troisième et dernier argument vient d'une simulation de l'évolution des articles indéfinis: je montre sur la base d'une simulation en EvolOT que l'évolution des articles indéfinis peut être conçue comme une évolution qui donne lieu à des points terminaux différents (étape 2 pour *unos* et étape 3 pour les articles indéfinis en général). J'en conclus que la transition collective de l'article indéfini singulier et de *unos* est une vraie transition de déterminants standards à des articles indéfinis.

### *Conclusions*

La discussion dans la première partie de cette thèse mène à trois conclusions. La première est que – en synchronie – on s'attend à ce que les SNs introduits par un article indéfini ne se distinguent des pluriels nus qu'au niveau pragmatique et que certaines caractéristiques que possèdent d'autres SNs leur manquent (partitivité, distributivité, quantité). La deuxième conclusion est que – en diachronie – on s'attend à ce que les articles manifestent une certaine évolution tant dans leur fréquence que dans leurs propriétés sémantiques. La troisième et dernière conclusion est que *unos*, sur la base de son comportement synchronique et diachronique, doit être analysé comme un article indéfini. La conséquence en est que les articles indéfinis et les noms nus argumentaux peuvent co-exister et que l'existence d'articles indéfinis ne peut jamais être une explication suffisante pour l'agrammaticalité de noms nus argumentaux.

## **2. Des français**

Dans la deuxième partie de cette thèse je traite *des*. Mon point de départ est le constat que *des N* peut recevoir plusieurs analyses. Celles que je traite le plus en détail sont – d'une part – celle qui considère *des* comme un article indéfini pluriel et – d'autre part – celle qui suppose que *des N* est un partitif nu, composé de *de*, *les* et un nom.

### *L'analyse en termes d'article*

Une analyse en termes d'article indéfini est attrayante pour deux raisons. La première est qu'elle permet facilement de rendre compte de l'interprétation plurielle et indéfinie de *des N*. La seconde est qu'elle prédit que son apparition historique (en

ce qui concerne son interprétation plurielle indéfinie) coïncide plus ou moins avec la généralisation de l'emploi de l'article indéfini singulier.

Le problème que pose une analyse en termes d'article est qu'en général les articles sont libres à prendre portée étroite ou large. Or, *des N* a une préférence claire pour la portée étroite:

(8) Il me faut des chemises.

La lecture fortement préférée de (8) est que j'ai besoin de chemises, n'importe lesquelles. Dans l'hypothèse où les articles sont libres en matière de portée cette préférence est inattendue.

*L'analyse en termes de partitif nu*

Ce qui rend l'analyse en termes de partitif nu encore plus attrayante que l'analyse en termes d'article est que – à part le fait qu'elle rend compte de tous les faits dont rend compte le dernier – elle prend au sérieux la morphologie complexe de *des*. Au niveau synchronique je défends qu'une analyse en termes de partitif nu rend compte de l'interprétation plurielle et indéfinie de *des* sans avoir recours à des suppositions *ad hoc*. J'obtiens l'interprétation plurielle sur la base du raisonnement par défaut, une variante de 'maximize presupposition' (voir Heim 1991 et Sauerland et al. 2003). L'interprétation indéfinie est obtenue sur la base d'une analyse en Weak Bidirectional Optimality Theory qui garantit que le SN défini inclus dans *des N* réfère à une espèce. Je défends en plus que ces mêmes procédés pragmatiques doivent être appliqués pour rendre compte de l'interprétation d'autres partitifs nus qui ne peuvent être analysés comme articles : *van die N* néerlandais, *de ces N* français et *di quei N* italien :

*van die N* néerlandais

(9) Ik heb van die rare ventjes gezien.  
J'ai de ces bizarres bonshommes vu  
J'ai vu de ces bonshommes bizarres.

*de ces N* français

(10) J'ai vu de ces pommes.

*di quei N* italien

(11) A quel punto, lui ha fatto di quelle smorfie...  
à ce point, lui a fait de ces visages  
À ce point, il a fait de ces visages (laid, bizarres, ...)

Au niveau diachronique je montre – sur la base d'une simulation historique détaillée en EvolOT – qu'une analyse en termes de partitif nu prédit que *des N* apparaît au même moment que le ferait une analyse en termes d'article.

En ce qui concerne la portée, une analyse en termes de partitif nu prédit que *des N* ne peut que prendre portée étroite. Ceci rend compte des données en (8), ce qui semble suggérer que – sur ce point – l'analyse en termes de partitif nu est supérieure à l'analyse en termes d'article. Or, (12) montre que *des N* n'exclut pas les lectures à portée large:

- (12) Il a refusé de collaborer avec des collègues de son équipe.  
<sup>ok</sup>*Il y a des collègues avec lesquels il a refusé de collaborer.*

Étant donné que *des N* admet exceptionnellement la portée large, l'analyse en termes de partitif nu semble poser le problème inverse de l'analyse en termes d'article.

### *Conclusions*

La discussion dans la deuxième partie de cette thèse mène à deux conclusions. La première est que les SNs introduits par un article indéfini et les partitifs nus ne se distinguent presque pas les uns des autres. La seule différence se situe probablement au niveau de leur portée: on s'attend à ce que les article n'aient aucune préférence et que les partitifs nus ne prennent que la portée étroite. La seconde conclusion est que *des N* ne peut être analysé de façon univoque ni en termes d'article ni en termes de partitif nu. La raison en est que son comportement au niveau de la portée ne suit ni de l'une ni de l'autre analyse. Ceci semble indiquer que *des N* était un partitif nu à l'origine, a évolué et s'est rapproché d'un SN introduit par un article sans pour autant en devenir un. Je conclus que les noms nus argumentaux peuvent être agrammaticaux dans une langue sans que cette langue dispose d'articles indéfinis et que l'agrammaticalité de noms nus argumentaux ne peut jamais être employée comme un argument décisif en faveur d'une analyse d'un déterminant en termes d'article.

### **3. Een néerlandais en position de prédicat**

Dans la troisième partie de cette thèse je traite l'article indéfini néerlandais *een* et son emploi en position de prédicat. La motivation derrière cette étude vient de la contradiction entre la supposition fortement enracinée que les articles indéfinis s'emploient pour marquer les arguments et le simple fait que les articles sont extrêmement fréquents – voire obligatoires – en position de prédicat. L'analyse que je propose a trois ingrédients importants: le premier est lexical, le deuxième compositionnel et le dernier diachronique.

#### *Le lexique*

Avec de Swart, Winter & Zwarts (2007) je suppose que les noms dénotent dans le domaine *e* et peuvent être subdivisés en deux sous-types: les capacités et les espèces. Les capacités telles que *professeur*, *Américain* et *musulman* réfèrent à la position d'un individu dans la société, que ce soit au niveau professionnel, civil ou religieux. Le point crucial est qu'elles sont définies de façon culturelle. Les espèces telles que *loup*, *rat* et *humain* sont exactement le contraire: elles forment des classes naturelles d'individus. Le terme *naturel* est opposé à *culturel* et reflète l'intuition que les espèces se distinguent des capacités à deux niveaux. Le premier est que l'appartenance à une espèce est basée sur des propriétés inhérentes: un loup appartient à l'espèce *loup* non par accident mais parce qu'il a les propriétés inhérentes qui en font un loup. Le second est que les espèces n'existent pas grâce à

une décision culturelle sinon grâce au fait qu'il y a au moins deux individus qui partagent les mêmes propriétés inhérentes.

### *Sémantique compositionnelle*

Avec de Swart, Winter & Zwarts (2007) je suppose qu'il y a deux changements de type qui convertissent les expressions de type  $e$  en expressions de type  $\langle e, t \rangle$ . Le premier est *CAP* qui prend une capacité et retourne l'ensemble d'individus qui ont cette capacité. Pour la capacité *professeur* il retourne l'ensemble des professeurs professionnels. Le second est *REL* qui prend une espèce et retourne l'ensemble des individus qui la composent. Pour l'espèce *chien* il retourne l'ensemble d'individus qui ont les propriétés inhérentes qui en font des chiens.

Supposant que *CAP* s'applique de façon invisible et que l'application de *REL* est liée à l'apparition de l'article indéfini, je prédis que les prédicats avec un article comporte une application de *REL* et que les prédicats nus comportent une application de *CAP*. Je montre que cette prédiction – enrichie d'une analyse traitant la coercion d'espèces en capacités et *vice versa* – rend compte de l'interprétation des noms qui réfèrent à des espèces et des noms qui réfèrent à des capacités en position de prédicat, qu'ils soient accompagnés ou non d'un article.

(13) Henriette is manager.

Henriette est manager

(14) Henriette is een manager.

Henriette est un manager

(13) a l'interprétation non-marquée qu'Henriette appartient à l'ensemble des managers professionnels tandis que (14) suggère qu'Henriette a les caractéristiques stéréotypées qu'on associe à l'espèce *manager* (être organisé, capable de déléguer,...) sans pour autant être un manager professionnel.

(15) Jantje is een raaf.

Petit\_Jean est un corbeau

(16) Jantje is raaf.

Petit\_Jean est corbeau

(15) a l'interprétation non-marquée que Jantje est un membre de l'espèce *corbeau* tandis que (16) n'est interprétable que si on suppose que Jantje joue le rôle de corbeau dans un jeu (comme le jeu *rats et corbeaux*).

### *Diachronie*

Dans la partie diachronique je motive le lien entre l'article indéfini et *REL*. La supposition que l'article indéfini en position de prédicat porte une implicature forte de non-unicité et la supposition que les ensembles correspondant à des espèces contiennent plus d'un élément jouent un rôle crucial dans cette motivation. Si on combine ces suppositions avec des contraintes générales sur l'évolution linguistique

on prédit que l'article indéfini était le seul candidat pour marquer *REL* dans les prédicats singuliers.

#### *Extensions*

Même si le but principal de la troisième partie est d'étudier l'article indéfini singulier néerlandais je fais également des suggestions spécifiques pour étendre l'analyse, p.ex. à l'article indéfini singulier anglais. Ce dernier article se comporte un peu différemment: il ne peut être omis que devant les capacités comme *présidente* qui réfèrent à des fonctions qui sont uniques dans un cercle social:

- (17) Hyacinth is chairwoman of the local Women's Institute.  
Hyacinth est président de le local Women's Institute

#### *Conclusions*

La discussion dans la troisième partie de cette thèse mène à deux conclusions. La première est que les articles indéfinis en position de prédicat peuvent être recrutés pour des fonctions apparentées mais nouvelles. La seconde est que l'article indéfini singulier néerlandais a été recruté comme la réalisation du changement de type *REL*.

#### **4. Ce que nous savons maintenant sur les articles indéfinis**

L'image des articles indéfinis qui émerge tout au long de cette thèse est celle d'une catégorie de déterminants qui sont pauvres en sémantique mais riches dans leur compétition pragmatique avec les noms nus et les autres déterminants. Cette image n'est rien d'autre qu'une version élaborée du *Principe de blocage* (Chierchia 1998): dès qu'on accepte que les articles indéfinis et les noms nus peuvent co-exister – que ce soit en position d'argument ou de prédicat – on ouvre un réseau tout à fait nouveau d'interactions possibles qui méritent d'être explorées. Les outils formels récemment développés que j'ai exploités – OT, BiOT et EvolOT – se présentent comme les compagnons parfaits pour cette exploration. Ils permettent de joindre les mérites du fonctionnalisme à la précision formelle de l'approche générative.



## **CURRICULUM VITAE**

Bert Le Bruyn was born in Zoersel (Belgium) on September 30<sup>th</sup> 1981. In 1999 he started his studies in Romance Languages at the University of Antwerp where he obtained a master in 2003. After a short period of teaching in his old highschool he joined the Mphil Programme in Linguistics at the University of Utrecht from which he graduated in 2005. From November 2005 till August 2008 he was enrolled in the International PhD programme of the UiL-OTS and from September 2008 onwards he has been working as a PhD student on the NWO project 'Weak referentiality: bare nominals at the lexicon-syntax-semantics interface'. During his PhD period he also worked as a teacher in the French departments of Utrecht and Leyden.