The place of locatives and the limits of the arbitrariness of signs

Sebastian Bican-Miclescu

August 15, 2012

Abstract

Locative forms are particularly wide-spread and salient in distributions that do not trivially come with a locative interpretation, such as the expletive subject in '**There** is a fiddler on the roof'. This phenomenon is central to issues such as the formalization of linguistic variation, the organization of the Lexicon with respect to poly-functional morphological forms and the syntactic analysis of specific constructions that come under its scope. Consequently, I begin by discussing the conceptual underpinning of such phenomena in order to make sense of this specific domain and inform research into the issues above (variation, the organization of the Lexicon, specific syntactic topics).

More concretely, I propose a way of dealing with the morphological 'locativity' of expletive subjects in some languages, in keeping with contemporary guidelines. Arguing against approaches that insist on introducing 'locativity' as a formal feature into the computation (and parametrizing its presence), I opt for identifying a salient 'subjecthood' property and hypothesize on how locatives may come to host it.

Other distributions of locatives, such as the Danish relative *der* or the Dutch function cluster, are also discussed in terms of how they could be integrated into a robust model of variation. The resulting pattern is also argued to suggest possible modifications to our model of the Lexicon. Thus, the emerging framework of Nanosyntax is put to the test as a technology that can capture the scalarity of functional polysemy. Dutch r-words (which are notoriously poly-functional locatives) are (re)analyzed as deriving from a principled morphological split in pronominal paradigms, partially reflected in dialectal variation. Their locative interpretation remains orthogonal to the pattern of distributions.

I reach the conclusion that we have a strong case for individually explaining the various distributions of locatives through mechanisms that render their primary interpretation as epiphenomenal to their functional interpretations. As an extension of this, I advance the hypothesis that the range of distributions of locatives (and other poly-functionals) can be captured in a scalar fashion that reflects a central design feature of natural language.

Acknowledgements

This thesis is dedicated to a great many people: I feel they are as much part of this effort as my own thoughts. Being part of humanity is an opportunity.

First of all, I need to thank my supervisor, Eric Reuland, for being an invaluable source of inspiration and support and making me believe in my ideas. He offered me the right hints and the freedom to explore them.

My second reader, Sjef Barbiers, deserves many thanks, not only for agreeing to be part of this endeavor but also for his inspiring teaching and advice during the program.

I am grateful for Maaike Schoorlemmer making things run smoothly for all of us over these two years, along with all our excellent professors. Sergey Avrutin's classes in particular have contributed to the thought process that lead to this thesis - I am grateful for the stimulating interaction.

Outside of Utrecht, I need to thank all the wonderful academics at CASTL (University of Tromso). My visit there last autumn was an opportunity to explore ideas that would later be significant for this research endeavor. Special thanks go to Michal Starke, Tarald Taraldsen, Marina Pantcheva, Pavel Caha, Tom McFadden and Sandhya Sundaresan. For providing me with companionship inside or outside of linguistics (as well as under an arctic sun or in day-long darkness), I thank Macek, Kati, other residents of Breiviklia 13, as well as my temporary office mates, Lucie and Sasha.

The wonderful group of colleagues in the Linguistics RMA at Utrecht deserves thanks for intellectual stimulation and genuine friendship. They have all left me with great memories, yet I will mention just a few: Anja (with her organized mind and open heart), Will (a real buddy), Heidi (always brilliant, always joyful), Masha (whoever has the most fun wins!). I am sorry I did not get to spend more time with the rest of you.

Very special thanks indeed go to my friends of Lombokstraat 76, for making a house (without a living-room) a home (with a living-room): Anne, Luca, Mircea, Martina, Juanita, Denis, Reinier, Randy, Mary Jane, Jarka & Stevo.

My stay in the Netherlands was also rendered spectacular by the presence of friends from back home. I give you 'The Romanian Mafia' - fine young intellectuals, all of them: Ileana, Maria, Dragos, Alex, Diana, Bogdan, Ela, Mircea and Brad.

Particularly affectionate thanks go to the warm people back home who were around and provided me with intellectual and emotional support during the finishing stages of this thesis: Mara, Andrei, Ioana, Maria and Adriana. Hot days and bad politics are more tolerable when you have love.

I have endless gratitude for all of my four grandparents' love, support and patience with

me being away. Many thanks go to my uncle and aunt, godparents and family friends for the part they have played in my life.

Two people in particular have made my path in life and academia possible. Thanks Dad for all the love, humor and the occasional push! I wish Mom could see this.

Contents

1 Introduction

This thesis sets out to explore the status of a peculiar class of words ('locatives', in a sense that will be explained below) in the model of natural language that modern grammarians have been striving to develop for about fifty years now, taking into consideration a time-frame that has Chomsky's revolutionary insights as its point of origin. Of course, the modern efforts to describe language in some systematic way are even older, and can be traced back (without jumping too much into the past) to Ferdinand de Saussure's insights about the relation between signs and meanings. In fact, both of these historical checkpoints are crucial for my current endeavor. On the one hand, the Chomskyan revolution has established a robust research program and resulted in the articulation of a powerful formal apparatus. On the other hand, the puzzles and solutions of this tradition largely elude the articulation of a robust model of sign-meaning relations. To be sure, de Saussure's key idea was that sign and meaning are divorced, and the generativist (i.e. Chomskyan) tradition has built on this, in the structuralist spirit of analyzing the relative distributions of signs. These developments, however, can also be viewed as only one theoretical branch starting from certain premises about signs and meanings. This leaves other options to be explored concerning sign-meaning associations - what I particularly take issue with regarding the 'underexploration' of this domain of possibilities is the fact that we lack an articulate model (or at least a coherent methodology) for dealing with words (signs) that have several distributions (or meanings, by extension) in the grammatical sense. How this kind of behavior should be regarded is a matter left largely open by the Saussurean thesis and the Chomskyan tradition, but of prime importance to both. In fact, at the interface between disciplines of study (like formal models and processing studies) we would expect to reach valuable insights precisely from such facts. Below, I sketch out a particular domain which requires an explicit analysis of the relation between certain words and the multitude of environments they play a role in.

Let us then take a look at this striking domain of words with a non-trivial range of distributions, labeled for convenience as 'locatives'. Consider the following examples:

(1)   The book I was looking for is right **there**.

(2)   **There** is a light that never goes out.

(3)   **There**!

The first of the three illustrates a core use of the word *there*, which is that of indicating a (non-proximal) physical location. In this sense, we label it as a 'locative'. Note that, in isolation (such as in (3)), the word will naturally be interpreted as such, so there is some indication of this being its core meaning in some sense. This meaning is not necessarily

carried over (in any trivial way) to the interpretation of *there* in (2).

The phenomenon which encompasses (1) and (2) is not just a peculiarity of English, neither is it trivial from a cross-linguistic perspective. The fact is that locatives such as *there* have other distributions as well, while the expletive subject position illustrated in (2) is filled by non-locative morphology in several other languages. This makes for a complex domain of linguistic variation. First, consider the following examples from Danish and Dutch[1]:

(4)   a.   Hun star lige **der**.

          She is standing just there.

     b.   **Der** sidder en fugl på min cykel.

          There's a bird sitting on my bike.

     c.   Kan du se den dreng **der** leger derhenne?

          Can you see the boy who is playing over there?

(5)   a.   Jan woont **er** nog maar kort.

          Jan has been living there briefly.

     b.   **Er** loopt een jongen in de tuin.

          There walks a boy in the garden.

     c.   Ik heb met hem **er**over gesproken.

          We have talked about it.

     d.   Ik heb **er** gisteren twee gekocht.

          I bought two yesterday.

It is clear that the first two examples from Danish parallel the locative use and expletive use in English (1, 2). The third example is however peculiar in the sense that it extends the range of use, compared to the English locative, by one function: the 'relative' (see Vikner 1991). The Dutch examples can also be viewed as an extension of the English data, while excluding the relative pronoun function we see in Danish. The function illustrated in (5c) can be referred to as 'prepositional', in the sense that it captures the morphological form a pronominal element takes in the context of a preposition. The function illustrated in (5d) can be referred to as 'quantitative', since it marks a similar effect on the shape of pronominal elements in the context of a numeral. More details on the Dutch function cluster will be given in Section 5.

Such examples make the domain of locative forms with different uses truly intriguing in its diversity and systematicity. Observing the pattern above, one might ask what the basis is for such a strong connection between expletives and locatives, for instance. This

---

[1]Bennis (1986)

would be an intriguing question in itself, but the addition of other distributions to the same range complexifies the data and the nature of questions to be posed. What is more, the link between expletive subjects (2, 4b, 5b) and locative morphology is not to be regarded as an absolute correlation. As the following examples illustrate, two closely related languages like Norwegian and Danish, or even a standard variety and a dialect (Standard English and Smith Island English)[2], can differ on such a minuscule but salient data point.

(6)   **Det** bor mange utlendinger i   Norge.   (Norwegian)
       it   live many   foreigners   in Norway
       There are many foreigners living in Norway.

(7)   **It** is poltergeists in this house. (Smith Island English, Parrott 2009)

Both examples illustrate the use of a 3rd person (neuter, singular) pronoun in the expletive distribution. Therefore, one needs to explain what makes locative morphology compatible with but not obligatory in certain distributions. The questions that one would pose are then, from the more fine-grained to the broader:

(8)   a. Can we explain the choice of locative morphology in particular syntactic distributions?

      b. Does 'locativity' as a formal property play a role in the syntax of these environments?

      c. Or can we explain the selection of that particular morphology on other grounds?

      d. What drives cross-linguistic variation in terms of the distribution of locatives?

      e. How are such items organized in the lexicon?

The issues addressed are critical at the present state of research into linguistic variation (from a formal perspective, within the Minimalist Program, particularly as synthesized in Barbiers 2009) and, I feel, they bring up what one may call "the dark side of the moon" in post-Saussurean linguistics: the status of words with multiple associated meanings and the limits of the arbitrariness of signs. What I mean by 'limits of arbitrariness' is that although the pairing of one sign and one meaning is never principled, once such an association has been performed, the range of other associations for the same sign will be to some extent restricted. In fact, in Section 6 I explore the hypothesis that the range of uses of a morpheme can be inferred if at least some of these uses are known. The technical consequences of addressing such issues are expected to impact the way syntactic computation and the Lexicon are modeled in formal linguistics and generative syntax in particular.

---

[2]Expletive *it* has also been noted in other dialects, such as Afro-American Vernacular English. See Setliff (1988) for a discussion of existential *it* in English dialects.

For these reasons, my thesis aims at tackling the questions above by starting with a conceptual and meta-theoretical discussion and then exploring some possibilities of modeling the observed data in a way that will be in line with contemporary theory and enlightening in terms of our understanding the place of locative words in natural language. Section 2 deals with some key conceptual aspects of my domain of inquiry, especially the idea of polysemy or homophony in the functional domain and also hints at the privileged place of lexical variation in contemporary theory. Two hypothesis are formulated regarding the place of poly-functional locative words in the Lexicon and their syntactic contribution. Section 3 presents the domain of expletive subjects empirically and theoretically, while formulating some objections to the way morphological variation in this domain is dealt with. Section 4 extends this case-study by presenting some facts that prompt a solution to the variation issue and discusses the resulting model in terms of the questions and hypotheses formulated under Sections 1 & 2. At the same time, I show that there is a link between expletive subjects and elements that front embedded clauses from where a subject has been extracted. Thus, the 'relative' function of Danish *der* is also explained. Section 5 discusses the multiple uses of Dutch *er*, as well as the morphological expression of these functions in non-standard dialects. In Section 6, I argue that novel tools might be needed in order to account for the complete picture, particularly the behavior of Dutch *er*. Finally, the conclusion draws on both the in-depth study of locative expletives (Sections 3 & 4) and the more tentative picture presented in Section 6.

2   Locatives in the Lexicon: some hypotheses

The present section establishes a conceptual framework in the sense of discussing some concerns and logical options that I take to be relevant in a non-derived, theory-neutral way. I deal with two different patterns of distribution for locative words, the status of poly-functional words in theory (and the mind) and the contemporary stance on linguistic variation (as stemming from morpho-syntactic variation), while finally establishing two working hypotheses regarding the representation of locative words and their interaction with syntactico-semantic computation, at the level of phrases and clauses.

2.1   The distribution of locatives

Before proceeding to more technical issues, I feel that an important descriptive point concerning locative words must be noted. For convenience, I will refer again to the examples initially given under (1) and (2):

(9)   The book I was looking for is right **there**.

(10)   **There** is a light that never goes out.

It is a trivial observation that the two distributions of *there* do not only differ in terms of their interpretation, but also in the way they respond to alternative fillers taken from the locative paradigm.

(11)   The book I was looking for is right **here**.

(12)   (*)**Here** is a light that never goes out.

The interpretation of (11) remains within the range of physical locations, yet (12) is not compatible with the existential interpretation of (10). If anything, the acceptability of (12) could be forced by a special context and stress pattern, but then it would simply be construed in terms of predicating location. This contrast can be formalized by saying that 'locative' forms can be chosen freely in unambiguously 'locative' environments (from a paradigm of spatial deictics), whereas in other environments only one locative form can occur. This is similar to saying that only one locative can be **grammaticalized** with a specific function. I will establish a terminological convention in calling environments like the expletive subject construction **'exclusive'** and environments of free alternation **'contrastive'**.

A clear pattern of such distributions can be observed in Dutch. Bennis (1986) explicitly states that quantitative *er* for instance has no 'r-variants', in the sense of locatives like *daar / hier*. This qualifies as an exclusive distribution. The same holds for expletive *er*, while the prepositional has the variants *er/daar/waar*. The locative flavour of *er* however has the free variants *daar / hier* and these depend not on the syntactic environment but rather on real-world conditions and speaker intention. This is characteristic of what I call a contrastive distribution. A point that is worth considering is that when we discuss exclusive distributions (i.e. functional uses of a morpheme) we are implicitly discussing patterns of grammaticalization.

Because the exclusive distributions block all locatives but one and in fact do not always require a locative cross-linguistically, it is doubtful that locativity is a key feature in those environments (i.e. that it plays a role in computation). It must rather be the case that in any one language there are certain formal conditions that an element in these particular distributions needs to meet and that locatives sometimes qualify for these formal conditions. In the following subsection I will show that contemporary ideas regarding linguistic variation make such an approach preferable. We therefore do not parametrize whether locativity is required or not (since languages are expected to behave uniformly in such issues of 'substance'), but rather parametrize the strategies that are made available for expressing certain

10

functions and explain how they target locatives precisely.

Finally, it must be noted that there seems to be an implicative relation between the occurrence of several exclusive distributions in one language for a particular morphological element. More details and an interpretation of this pattern will be given in Section 6.1.

2.2 Linguistic variation in the Lexicon

This subsection will establish a technical and conceptual framework for dealing with linguistic variation, which I show to be critical in dealing with the distribution of locative morphology. I intend to set the present study in a framework that assumes a principled (rather than case-by-case) approach to linguistic variation. The general stance on variation I subscribe to is that Syntax is uniform across languages (both operations, like Merge, and the building blocks that are operated on, like the cartographic component under some analyses) and that it is in the Lexicon that variation emerges. Parametrization thus amounts to specifying what features are associated with the words of a language. This is the general approach argued for in contemporary linguistics, at least since Chomsky's Minimalist Program (1995). The dawn of modern studies of linguistic variation can be identified in the Principles and Parameters (P&P) model of the 1980s, though this model accepted parameters in the sense of language-wide statements. Narrowing the domain of parametrization to lexical specifications seems to be a theoretical improvement.

In the Saussurean tradition it is assumed that there is no way of predicting the correspondence between an auditory unit and interpretive load. On the other hand, at least intuitively, we would expect to see some correlations between several apparently different distributions of one unique sign. It is undoubtedly a theoretically interesting notion to propose that we can infer the underlying properties of linguistic structure from the different positions one sign can occupy in the collection of phrases a language can generate. Extending this methodology to as many such items and as many languages as possible should give us a solid map of linguistic structure. Thus, the effort of capturing clausal templates is re-conceptualized as an effort of identifying patterns of morphological distribution. Section 6 should be understood as a strong proposal in this direction. In what follows I discuss the guidelines for studying linguistic variation and anchoring it in the morpho-lexical domain that are formulated in Barbiers (2009), whose study corresponds to the contemporary view I hint at in the previous paragraph, re-formulated as point (c) of the following list:

(13)   Three theoretical options for variation:

      a. Relatively free, possibly with some functional constraints (an option assumed by many typologists)

b. Constrained by some overarching innate principles (core UG, from the 50s to Minimalism, possibly implemented as P&P): morphological variation is relatively free (following the original Saussurean assumption)

c. Encoded morphologically (by formal features, not auditory form, obviously, as practiced in Minimalism): patterns of morphological distribution are also constrained.

Two dimensions for the formal study of natural language are established in Barbiers (2009): the dimension of universals (or Principles) and the dimension of variation (or Parameters). Syntactic principles are viewed as universal, and an extension of this perspective is that a Universal Base Hypothesis should be adopted. This restricts the range of syntactic operations and their domain of application (i.e. a uniform cartographic component). The dimension of parametrization is formally represented as the Lexicon and as the interfaces between core syntax, the Lexicon and other modules. In other words, variation can be reduced to the morphosyntactic properties of individual items (lexical entries) in a language or dialect. This approach makes space for describing fine-grained variation at the level of dialects or even idiolects.

Syntax, on the one hand, can restrict the legitimate combinations of lexical items, according to their featural make-up. At the same time, within the range of possible constructions, not all are instantiated in every language. These can be viewed as missing entries. At the interfaces, optionality can arise, i.e. when two configurations respect syntax and the range of available lexical entries, speakers are free to choose (this is argued to happen at the PF interface).

The core idea is that the locus of linguistic variation is the Lexicon (with a possible extension to processes of selection and competition that target lexical entries). There are however limits to how variation can manifest itself. That is, there are possible ranges of formal associations for a certain lexical entry within a linguistic system, as dictated by overarching properties of lexical organization and syntactic computation. It is precisely this idea that I attempt to explore in connection with the functions of locative words: how can a locative be used intra- and inter-linguistically? Is there an emerging pattern? And if so, what principles are responsible for such patterns? Naturally, I also assume that an explanation should be formalized at the level of the representation of such words, and not as language-level statements.

Finally, I need to draw the reader's attention to some empirical facts that have already featured in my introduction and that suggest that we should not conceptualize variation in terms of the larger syntax of the language, but rather in terms of more subtle differences in

the use of some forms. Smith Island English for instance is not a V2 or V-final language but its morphological choice of an expletive subject differs from the standard one, as illustrated in (7). Therefore, the two varieties do not seem to differ in terms of their clausal organization (available positions and patterns of attraction) but nevertheless one uses a locative as an expletive subject. Belfast English on the other hand accepts syntactic deviations from the standard (like transitive expletive constructions) but uses the same expletive as Standard English.

2.3 Formalizing polysemy

The flat Saussurean perspective allows two interpretations of locatives with exclusive distributions - either the same sign is associated to different meanings (formally translated as: different lexical entries with different properties) or somehow all of the exclusive distributions discussed are the same in terms of some unique property that constitutes the relevant and unique 'meaning' of the locative morpheme. Establishing the correct representation for such words is crucial at the present state of linguistic research, since, as argued above, even patterns of syntactic variation are being progressively reduced to lexical specifications. How the Lexicon organizes certain forms, such as poly-functional ones, is thus a key issue.

The first option that we have for describing words with multiple distributions is to appeal to the concept of homophony, i.e. a phonological coincidence with no semantic underpinning. In natural language it is quite a common situation for words to have several unrelated(!) meanings. Consider:

(14)   A **rose** smells just as sweet by any other name.

(15)   I **rose** up one morning.

Clearly, the examples above illustrate two completely different meanings that share a morphological form. But is this how we expect to analyze all such cases of form-meaning mappings? In fact, there is also a robust class of examples that suggest a certain degree of relatedness between the possible meanings of a word. Consider, then:

(16)   I used to have a pet **rabbit**.

(17)   Do you eat **rabbit**?

That such examples illustrate different meanings is not a pre-theoretical given. One can assume instead that *rabbit* in this case stands for a concept of 'rabbithood' that can be plugged into different contexts. This argument hinges on a question of **granularity**: at the level of natural language, is the distinction between [rabbit.animal] and [rabbit.food] a

relevant one? Since in this and other similar cases (like *duck*) natural language does not mark such a distinction in morphology, one could be tempted to reply negatively to the previous question. This would however be misguided, because in still other cases, the distinction between an animal and its edible flesh is marked in morphology:

(18)   I used to have a pet **pig/*pork**.

(19)   Do you eat **pork/*pig**?

Natural language is therefore sensitive to this level of semantic granularity, yet it remains optional for the distinction to be expressed morphologically. Not all meanings that can be contrasted in language are contrasted. This is a design feature (at least at the descriptive level) that should be kept in mind throughout this paper and when approaching linguistic topics in general.

Such examples also seem to be systematic and productive: some words behave like *rabbit*, others like *pig/pork*, and any new animal name could be assigned to one of the two classes. On the other hand, there is no systematicity behind the two uses of *rose* in the examples illustrating homophony - one would be hard-pressed to find another flower name that is also an inflected verb form, and the example is impossible to reduce to a productive template.

In short, this state of affairs partly motivates the traditional opposition between homophony and polysemy. With this opposition in mind we can proceed to ask the following question about the pair of examples in (1) and (2):

(20)   a. Is there a productive pattern that captures the two uses of *there* or are we dealing
          with an accident? This question can itself be decomposed into the following two:

       b. Is the contrast between the uses of there in (1) and (2) relevant at the level of
          semantic granularity natural language is sensitive to?

       c. If this contrast is relevant, can the two uses be translated into a productive
          template?

It will be obvious by this point that I am taking the idea of polysemy and homophony from the lexical domain (comprising words such as *rabbit, pig* etc.) and superimposing it on the functional domain (which is assumed to consist of words such as *there*). I will argue that this can be a fruitful way of reasoning about language and in particular about functional items.

In the domain of lexical polysemy, an approach that stands out is that of Pustejovsky (1995). By and large, his endeavor starts from observing productive templates such as the one I discuss for [animal - meat] and arguing that words are represented by a semantic grid that derives all of their possible uses by means of a limited set of operations. Note that this

level of representation is expected to hold for the conceptual domain. For the domain of functional words and their syntactic properties, there is no real consensus about a system of representation that can robustly derive the uses of one word or morpheme. Even the notion that there are such different uses for one functor is rarely argued for explicitly.

In syntactic literature, most research on apparent polysemy has been conducted with regard to verbs, especially those that seem to accept different argument structures. From Hale and Keyser (1993) to Reinhart and Siloni (2005) there have been many models that try to establish the representation of verbs and how their properties can be altered. I will not go into these proposals in detail. It suffices to say that the general topic of words with several functional identities has been under-explored in syntactic research and that the current paper is partly an attempt to cast some additional light on it. Locatives in particular have not enjoyed the preferential status of verbs, although, as I will show, their behavior is at least as strikingly diverse.

At this point, three options can be outlined for the status of *there* in the pair of examples I started from:

(21)   a. It might be one **atomic element** in what the syntactico-semantic system of language is concerned. I have already suggested that since other languages split the two roles (formally, locative adverb and expletive subject) in morphology, language is in fact sensitive to the contrast, hence does not treat both occurrences of there as the occurrence of the same object. This is therefore the least likely option, though I will hint to some syntactic analyses that seem to follow this assumption. However, I will argue that such analyses are usually not adequate for a number of reasons.

      b. It could also be the case that the two occurrences of *there* are **unrelated entries** in the mental Lexicon of a speaker. This is by no means an impossible state of affairs, since at least for lexical items homophony is easily observable. It would be interesting to establish whether homophony actually characterizes some or all functional items with many uses. For syntactic analysis, this is the least involved but also least rewarding option since, on the one hand, no relation needs to be established between the two contexts of use, allowing for completely different theoretical treatments, but on the other hand there are data which suggest a systematicity which is incompatible with the accidental nature of homophony.

      c. The third option is composite between the first two and can arguably overcome the problems that they pose: the two occurrences of there in (1) and (2) (and other similar cases), can be traced down to a **decomposable element**, which

corresponds to one finely structured entry in the mental Lexicon. This is reminiscent of Pustejovsky's (1995) treatment of lexical polysemy, who attempts to derive all meanings of polysemous words from an underlying semantic representation tracking basic properties of those words. In Section 6, I will suggest a way of doing this for functional words.

With regard to terminology, the observations above prompt me to refrain from using the term polysemy, since it partly entails the conceptual domain rather than the grammatical. Formalizing polysemy can be done successfully in Pustejovsky's model, but I do not expect it to be trivially applicable to the functional (grammatical) domain. [3] On the other hand, generative syntacticians tend to treat the class of functional items I deal with as homophones, at least implicitly. I argue, however, that more theoretical gains would be drawn from exploring the possibility that these items hint to a formal pattern awaiting to be uncovered. In this sense, they are closer to polysemic items and I adopt the label **'poly-functional'**.

If we are to check for hints of polysemic behavior in functional items (i.e. grammatical distributions) we will then expect to highlight the formal sensitivity of language to their different environments (recall the issue of granularity) and to observe some productive patterns (i.e. correlated distributions) intra- or inter-linguistically. I will show that these are in fact conspicuous properties of the different distributions of locatives.

2.4 Hypotheses on 'locativity' in syntax and the Lexicon

Locatives are a particular class of words that engenders questions about linguistic variation (i.e. when in apparently the same construction different languages use a locative and a non-locative) and also about functional polysemy (i.e. when a locative has several grammatical distributions). We have seen that syntactic variation is taken to stem from variation in the Lexicon (Section 2.2) and that functional words with multiple distributions pose a problem of representation in the Lexicon (Section 2.3). Consequently, it is natural to ask how locatives should be represented in the Lexicon, and also, how exactly they contribute to the syntactico-semantic computation of constructions they participate in. The status of the core interpretive property of locatives (the spatial dimension) in syntactic computation is an extremely sensitive point. Therefore, my paper is an attempt in the direction of answering the following questions:

(22)   What is the status of poly-functional items in the Lexicon of a natural language in
        terms of representation and compositional contribution at insertion? More impor-

---

[3] To be sure, there are interactions between conceptual specifications and patterns of distribution (in the sense of 'selection'), but I will not be concerning myself with this 'interface' in the present paper.

tantly, can we identify a design feature that is responsible for this type of organization in natural language?

(23)   Does a [+locative] feature participate in the computation of locative elements irrespective of the role these elements play in different syntactic environments? And in fact what are the relevant features of such entries (recall the issue of granularity).

At least for a specific locative word, similar concerns have been explicitly expressed in previous literature. Hans Bennis makes the following observation:

"The analysis of *er* as an adverbial pronoun is in fact the traditional view, as put forth in the Bech's (1952) article 'Ueber das niederlandische Adverbialpronomen *er*'.This analysis differs from most, if not all, recent analyses of *er*. These generally take the different syntactic functions of *er* as an indication or motivation for the existence of several distinct pronouns which may vary in their categorial status and inherent syntactic properties. A different view in a completely different framework is expressed in Kirsner (1979). Basing his account on the conception that one form should correspond to one function/meaning, he argues in favour of a unificatory analysis of *er* that is based on the idea that the meaning of *er* is constant and can be described as presentational. I agree with Kirsner that an optimal theory of *er* should relate all occurrences of *er*. I do not agree with him that such a unificatory approach should necessarily be based on a particular meaning of *er* that is present in all specific instances. It can be argued that *er* possesses inherent properties which allow *er* to appear in a variety of syntactic constructions. Differences between distinct occurrences of *er* should then be considered to follow from the distinct syntactic relations and not from the appearance of different pronouns which happen to have the same surface realization." (Bennis 1986: 171-172)

Then, there are two dimensions in the analysis of items such as *er*:

(24)   The dimension of identity: how many lexical entries correspond to the morpheme?

(25)   The dimension of formal properties: what is their categorial status and how does it fit into different environments?

I propose two hypotheses that can be explored Both assume that the different uses of locatives should not be explained in terms of how some unique feature (say, [+L]) satisfies multiple environments. Otherwise, parametrization would become problematic. The stronger hypothesis of the two however, also makes the claim that these different uses, although not related by one key property, entertain some kind of relationship that accounts for the range of scalar patterns observed.

(26) **Weak hypothesis**: it is not some feature [+L] that satisfies each environment where locatives appear. In each case (on a construction-by-construction basis) some other formal property satisfies the environment and the locative word is somehow the 'host' of this relevant property. There is no relation however between the features involved in the different environments.

(27) **Strong hypothesis**: the same assumption holds as above, but there is some implicational relation between the environments discussed, which can be expressed as a scale that derives from some principle of linguistic organization. [4]

---

[4] This scale is hard-wired in UG. It is possibly a relevant design feature of human language, with some specific effects on the nature of processing, in acquisition and production.

3 Expletive subjects

The present section discusses one of the more salient distributions of locative elements - the expletive subject construction. First of all, I will present the key facts about expletives and the sentences where they surface, as discussed in the literature. For my present purpose, the central fact is the morphological split between languages which use a 'locative' expletive and languages which use a (3rd person) pronominal element instead. What I will attempt to do is formulate a model in which this morphological split can be parametrized in a way that is not anachronistic with other developments in generative linguistics. In order to derive this model, I will draw on contemporary ideas regarding linguistic variation (as presented for instance in Barbiers 2009). In the following section I will refer to some observations concerning the prototypical subject position and its interpretive effect and establish a model of variation by competition, where the properties of locative entries in the Lexicon make them the optimal fillers of the subject position, compared to 3rd person pronominals and other forms. The model will be extended to capture locative morphology in other clause-initial elements as well. This will be achieved without making use of parameters in the sense of language-specific statements and without forcing 'locativity' as a formal feature into the syntactico-semantic computation.

3.1 Describing expletives

Expletive subjects are a pan-Germanic phenomenon that varies along certain axes. On the one hand, there is the issue of variation in terms of the expletive element itself: what kind of morphology each language uses and where this element surfaces. In some cases the expletive can or must be dropped, while in others, one expletive form is replaced by another. On the other hand, the presence of expletives can also be discussed in terms of 'non-local' effects, like agreement patterns and the properties of co-occurring arguments or verb types. All of these domains allow for variation across the Germanic languages. For a standard reference I will point the reader to Vikner (1995), while a well-managed overview is available in Van Craenenbroeck (2011). The discussion that follows builds extensively on the organization of the latter.

First of all, we need to identify expletive constructions as those constructions where a lexical subject does not occupy the prototypical position and a non-thematic element is plugged in instead. Initially, this seems to largely correlate with the kinds of verbs used in a sentence. Verbs of existence usually go with expletive constructions, and this type of sentence is usually called an existential sentence. Existentials seem to be a core domain for the presence of expletives, and easily illustrated in English (see (28a) below). At the

same time, impersonal verbs (those which lack an identifiable subject) also require the use of an expletive, and even though they are not present in English, most Germanic languages exhibit the impersonal verb and expletive construction (see the Danish and Dutch examples, below). The widest range of constructions that allow the use of an expletive is signaled by transitive verbs occurring in sentences fronted by an expletive. Dutch, German and Icelandic prototypically exhibit this behavior, but not English or Mainland Scandinavian languages. The following examples illustrate this domain of incremental possibilities:

(28) a. There are many linguists who puzzle over such contrasts.

  b. * There may no smoking / was soccer played.

  c. * There read a book many men.

(29) Der ma ikke ryges. (Danish; Svenonius, 2002:5)
  there may not be.smoked

  Smoking is not allowed.

(30) a. (Fred denkt dat) er   twee koeien in de   tuin zijn. (Dutch; Reuland, 1988)
     Fred thinks that two cows in   the garden are
     (Fred thinks that) there are two cows in the garden.

  b. Er   wordt gevoetbald.   (Dutch; Bennis, 1986)
     there is   soccer.(play)ed
     Soccer is being played.

  c. (Fred zag   dat) er   veel   mensen een boek lazen. (Dutch; Reuland, 1988)
     Fred thinks that there many men   a   book read
     (Fred thinks that) many men are reading a book.

An apparently more trivial difference (but nonetheless, the crux of my interest in this domain) is that between the locative and non-locative morphology of the expletive in different languages. This morphological split seems to be a striking candidate for parametrization. Whether one claims or not that the morphology has anything to say about the syntax of the construction, it remains a well-known observation that it splits the Germanic languages into two groups. Locative expletives appear in English, Dutch, Danish and Frisian, whereas non-locative (3rd person) expletives appear in German, Swedish, Norwegian, Icelandic, Faroese and Yiddish. Dialects of the languages can differ in this respect, and sometimes one language uses both types of morphology in different contexts, as will be shown. Consider the following illustrations, taken from Van Craenebroeck (2011)[5]:

---

[5]The original examples appear in Vikner (1995).

(31) a. **There** arrived three men. (English)

      b. **Er**   is een jongen gekomen. (Dutch)
          there is a   boy   come

          A boy has come.

      c. **Der** er kommet en dreng. (Danish)
          there is come   a   boy

          A boy has come.

(32) a. **Det** har kommit en pojke. (Swedish)
          it   has come   a   boy

          A boy has come.

      b. **Es** ist ein Junge gekommen. (German)
          it is a   boy   come

          A boy has come.

      c. **Það** hefur komið strákur. (Icelandic)
          it   has   come a.boy

          A boy has come.

Another important dimension along which expletive constructions vary from language to language is the distribution of the expletive in terms of the structural position it occupies. In Mainland Scandinavian languages and English, it is clear that the expletive always needs to surface in specific kinds of clauses, regardless of phenomena such as inversion, etc. Dutch behaves the same in principle, but there are certain contexts where the expletive seems to be at least optional. In Icelandic and German though, the expletive is not compatible with inversion, while only in German is it also blocked by embedding. In this sense, German is at one end of the scale, with Dutch at the other, although in what concerns other phenomena, such as the range of verbal constructions, they pattern the same. In the middle of the scale, there are the issue of optionality and the issue of different expletives targeting different environments. Consider the following examples, taken from Van Craenebroeck (2011:2-3), with the exception of the Icelandic ones which can additionally be traced back to Thráinsson (2007:310,312,319):

(33) a. **Er**   staat een maan in de tuin.   (Dutch)
          there stands a   man in the garden

          There is a man standing in the garden.

      b. **Það** eru mýs í   baðkerinu. (Icelandic)
          it   are mice in bathtub.the

          There are mice in the bathtub.

  c. **Es** ist ein Junge gekommen. (German)

    it is a boy come

    A boy has come.

(34) a. ...dat **er** een man in de tuin staat. (Dutch)
    ...that there a man in the garden stands
    ...that a man is standing in the garden.

  b. ...að **Það** verði ball í skólanum á   morgun. (Icelandic)
    ...that it will.be dance in school.the tomorrow
    That there will be a dance in the school tomorrow.

  c. ...das (***es**) ein Junge gekommen ist. (German)
    ...that it a boy come   is
    ...that a boy has come.

(35) a. Staat **er** een man in de tuin?(Dutch)
    stands there a man in the garden
    Is there a man standing in the garden?

  b. Eru **(*Það)** mýs í baðkerinu? (Icelandic)
    are it   mice in bathtub.the
    Are there mice in the bathtub?

  c. Ist (*es) ein Junge gekommen? (German)
    is it a boy come
    Did a boy come?

Additionally, there are cases where the morphology of the expletive changes with the position it occupies. A specific case reported in the literature is that of Lapscheure Dutch (Belgium). The examples below are taken from Van Craenebroeck (2011:4-5), originally reported in Haegeman (1986)[6]. In this dialect, an expletive *t* surfaces in main declarative clauses, while under inversion and embedding we see the locative expletive *der*.

Such cases are interpreted as evidence for two different kinds of expletives: CP-expletives (high) and TP-expletives (low). The first class is never locative and has the particularity of being dropped under inversion, like in German, Icelandic, Lapscheure Dutch etc. Additionally, CP-expletives do not induce agreement on the verb, although they usually appear as 3rd person singular pronominals. TP-expletives do induce agreement (forced singular), but only when they are not locative. Verbs agree with the 'demoted-subject' DP whenever a CP-expletive or locative expletive appear. Empirically this results in striking differences, such as that between Standard English and Smith Island English (S.I.E.):

---

[6]qtd. in Van Craenebroeck (2011): Haegeman, L. 1986. *Er-sentences in West Flemish*. Ms. Université de Genève

(36) En.: There ... pl ... pl-DP

(37) S.I.E.: It ... sg ... pl-DP

There are additional correlations and interactions to be observed between these domains of variation. For instance, the type of expletive does not seem to correlate with the available verbal constructions: the Scandinavian group exhibits both types of expletives and the same range of verbs, while Dutch has the rather rare transitive expletive construction (TEC) and shares this with languages which do not have locative expletives, like Icelandic and German. Standard English shares the expletive with Belfast English, but unlike this dialect, it does not allow TECs.

A final point concerns the restrictions that are placed on the nominal associate of the expletive, the post-verbal DP. Starting with Milsark (1974), for English it was noted that, with the exception of some special pragmatic uses, definite nominals (and in fact a whole class of 'strongly quantified' nominals) cannot appear in expletive constructions. This was taken as a universal before further evidence was brought from languages such as Icelandic concerning the link between available syntactic positions and types of allowed nominals. Vangsnes (2002) for instance, shows that in Icelandic, although the definiteness restriction holds as expected in the post-verbal position, an intermediate position is also available for DPs in expletive construction, and this position allows strongly quantified nominals. There is then no direct influence of the presence of an expletive on the types of nominals allowed in a clause and doubt is cast on the idea of an expletive-associate chain.

3.2   Modeling expletives

Expletive constructions have received a generous amount of attention in the literature of generative linguistics. The particular domain of English existential sentences was perhaps the first to be approached. Given this state of affairs, the expletive construction was understood to be quite a restrictive environment, with idiosyncratic properties such as a definiteness restriction on the post-verbal nominal phrase (the associate) and an affinity for certain verb classes. Milsark (1974) is a classic reference that establishes this view of expletive constructions. With subsequent research, linguistic diversity has informed a more finely articulated perspective on the topic. In *The Representation of (In)Definiteness* (Reuland and Ter Meulen, 1987), despite its more general concern with the level(s) at which definiteness-related phenomena should be encoded, existentials occupy quite a central position, while the data and the approaches are clearly more diverse. This line of research however, still insists on the privileged position of the associate, and its connection to the expletive. A paradigmatic shift occurs with the exploration of expletive constructions in Germanic languages. Vikner

(1995) is the standard reference for this empirically diverse line of research, attempting to explain the syntax of expletives by referring to the syntax of the respective languages, in particular, what positions are available for the movement of verbs and arguments. In what concerns the contribution of the expletive itself, there are largely two kinds of approaches: the expletive originates within the VP and has a connection of some sort to the nominal associate or the expletive is inserted in its observable (high) position for strictly structural reasons. Neither of these lines of research attempts to explain the morphological nature of expletives.

First of all, I need to present a brief sketch of past syntactic models with regard to the status they attribute to the expletive element. Following the dichotomy used in Hartmann (2008), I oppose "in-core-predication" approaches to "insertion" approaches, as follows.

The first class refers to proposals in which the expletive originates within the VP, as a part of a predicational structure, such as a 'small clause' etc. In such systems, the expletive can function as a predicate (or subject of a predicate) and is assumed to make an essential contribution to the syntactico-semantic computation. To parametrize the locativity of the expletive under such a system amounts to parametrizing the articulation of argument structure and predication. Whether such a substantive notion as 'locativity' is expected to introduce cross-linguistic variation at this level is not an unproblematic question. Alternatively, one should find some other explanation for the way the dummy predicate or subject is selected morphologically and I am not aware of such attempts in the literature.

This line of research is primarily represented by Moro (1997 and earlier). He claims *there* is a dummy predicate occurring in a small clause (SC), along with the NP associate (the subject of this predicate). Its category is the same as an NP - just like *the picture on the wall* and *the cause of the riot* are two NPs that entertain a subject-predicate relation in a SC, in sentences such as: *The cause of the riot is a picture on the wall.* Hoekstra and Mulder (1990) adopt a similar approach, with the exception that the dummy predicate is considered to be a PP. This provides a link to the spatial PPs that sometimes occur in existential constructions. This could be taken as a hint towards the locativity of expletive *there*, but the issue of parametrization remains, like in all similar cases: how do we get expletive *it*? Alternatively, the expletive could be the subject of predication, either in SpecIP, as proposed by Williams (1994) (where the associate nominal is in the complement of the copula and functions as a predicate), or in SpecPredP, as proposed by Hazout (2004) (where the associate is the complement of the Pred head). Again, morphological variation in such a model is apparently unrestricted, and the empirical options, including the locative, cannot be explained in a principled way.

The second class of approaches (the "insertion" ones) at least implicitly assumes that

expletives are true 'dummies'. Thus, it is hypothesized that the reasons behind expletive insertion are structural and that these elements are semantically light. Compared to the modicum of semantic load in-core-predication expletives would have, these are even more minimal elements, perhaps reduced to a crucial syntactic feature. They are generally assumed to be base-generated in a high syntactic position, rather than originate in some other configuration. Consequently, parametrizing the morphological choice of the structural filler is quite a daunting puzzle. Whether implemented as EPP pure and simple, or some necessity of feature checking, the structural pressure on the presence of the expletive never has something to say about its morphology. What I will do in the following section is identify a common property of elements that satisfy this structural requirement and leave the issue of parametrization to the way morpho-lexical competition is resolved. In this sense, I opt for an 'insertion' approach of a more substantive nature.

3.3 Parametrizing expletives

So far, I have shown that there is barely an explicit proposal in the literature that could explain the locativity of expletive subjects in certain languages. Descriptively, some correlations between the expletive type and the syntax of the language have been observed. In what concerns syntactic models, very little research has been informed by or contributed to the issue of expletive morphology. On the one hand, there are approaches that vaguely find a place for the expletive in the way they conceptualize the syntactic representation of (existential) predication. On the other hand, a competing line of research simply assumes that expletive subjects are associated with a specific formal feature, quite independently of their morphology. Such approaches go around the morphological issue. This would not be a problem if the theoretical landscape lacked a robust theory of variation, as was the case at older stages of the theory. However, as I hope to have shown, there are currently solid proposals in the direction of a formal representation of variation (namely 'variation in the Lexicon', as captured in Barbiers (2009)) and adhering to such a model makes the issue of expletive morphology not only interesting but also unavoidable.

There is, then, a striking split in expletive morphology that most analyses so far ignore. A recent study by Jeroen Van Craenenbroeck (2011) attempts to also deal with this issue. The morphological split is therefore acknowledged and an attempt is made at encoding it. Van Craenenbroeck chooses to look in depth at Dutch dialectal variation regarding this issue and he adopts a proposal put forth in Ritter & Wiltschko (2009), that is assumed to explain locative forms arising in the vicinity of I(nflection), as an effect of agreement. Whether this locative kind of agreement manifests or not is supposedly fixed by a parameter. Furthermore,

Van Craenenbroeck sets his proposal in a 'Viknerian' framework: he explains (expletive) phenomena by the syntactic positions and patterns of movement available in the different dialects, and he plugs in the parameter proposed by Ritter & Wiltschko as a means of getting the locative morphology. Whether this morphology manifests itself (and in what positions!) is left to the mechanics of movement, landing sites and agreement patterns.

The empirical range this analysis explains is significant. However, I take issue with one assumption that is needed to derive locative morphology in his view: Inflection / SpecTP has a dimension of locative features that need to agree with an expletive. The presence or extent of this dimension however needs to be parametrized, and the way this is done, I argue, is entirely stipulative and goes against certain conceptual developments in the formal study of linguistic variation.

Therefore, Van Craenebroeck needs both INFL$_{Loc}$ and a parameter that dictates T-to-C movement. These are both problematic issues, conceptually and empirically. I will ignore the movement parameter since it does not directly influence my study of locatives and also since it is understandably a sensitive point in current syntactic research, usually approached without much explanatory ambition. It can be said however, that this is precisely the kind of parameter we expect to eliminate under a robust theory of lexically encoded variation. If we were to also adopt the INFL$_{Loc}$ parameter, it would have to be accepted in its fully-fledged form in order to explain the morphological contrast in Mainland Scandinavian languages: agreement of the locative type can be absent, in the sense of a negative setting of the parameter. Again, this goes against the idea of lexically encoded variation, delegating parametrization to a language-level statement.

Observe then, how Mainland Scandinavian languages either make the model under scrutiny crash or require the setting of a INFL$_{Loc}$ parameter on different values. Normally, Swedish and Norwegian are assumed to be the same as Danish in terms of T-to-C movement and so on. They differ, however, on the morphology of the expletive. In Van Craenebroeck's system, as I understand it, the expletive is by necessity locative when inserted in the low (TP) position for agreement reasons and always non-locative when inserted in the high (CP) position, since there is no (locative-)agreement requirement on it. What is more, these 'high' expletives are C-heads and generally exhibit reduced phonology. Applying this argumentation to Mainland Scandinavian languages, Danish locative *der* would have to be a C-head (with its locativity left unexplained), or Swedish and Norwegian *det* would have to surface under the pressure of locative-Agr - so where does their locativity go?

What is more, there is a conceptual imperfection in the system proposed by Ritter & Wiltschko (2009). They start from examples in Halkomelem (a Salish language), where in existential predication, the verbal particle agrees for location / deixis: it is morphologically

marked as proximal or distal. To take the "Agreement" label of this phenomenon and apply it to (Germanic) expletives, which are either proximal or distal for a given language (as I remark about exclusive distributions in Section 2.1), seems to be a rather forced strategy.

Additionally, in Dutch and its dialects there are other uses of *er*. Although Van Craenenbroeck works mainly on Dutch, he does not deal with these other uses. If the selection of this morpheme as an expletive subject instead of a *(he)t* form depends on 'locative agreement', then how do we fit this type of agreement in the other environments where *er* occurs? An agreement site similar to INFL should be proposed at some other syntactic position. This strategy already becomes very stipulatory and cartographically involved. Since Van Craenenbroeck already states that object expletives cannot be locative because there is no INFL node to generate locative agreement, Dutch *er* in the prepositional or quantitative distribution remains a morphological puzzle. After dealing with the expletive parametrization issue in the following section, in Section 6 I move on to sketch an analysis for the other uses of *er*, without resorting to locative agreement or some other model that assumes such a [+L] feature.

As a consequence of these objections, and especially keeping in mind the guidelines in Barbiers (2009) and most Minimalist syntax, I object to the parametrization of languages in terms of this locative-agreement phenomenon and formulate it mnemonically as: *don't parametrize the language to get the expletive; parametrize the locative to get its expletive use!*

4 Locatives first: competition & clause-initial elements

Building on the issues I have identified as problematic for the parametrization of expletive subjects, this section establishes a model of competition between lexical entries that can favor locatives for the expletive role, without targeting 'locativity' as a formal feature. This allows for variation being encoded in the lexical entries of each language, following the practice of the Minimalist Program, as presented in recent works such as Barbiers (2009). I point out a salient property of lexical subjects and even prepositional phrases in subject position that we expect to also hold for expletives, however this property may be encoded. The analysis is then extended to the 'relative' use of Danish *der*. The last subsection discusses the possibility of implementing a morphological competition model based on this salient 'subject' property and some consequences for the analysis of locative expletives.

4.1 Given a property of subjects...

This subsection deals with certain phenomena that I find to be central to the environment in which expletive subjects appear. I will not be pursuing a specific approach in terms of syntactic cartography, but rather intend to draw attention to a particularity of elements that are in complementary distribution with expletives. These constituents, whether DPs or locative PPs (as in the locative inversion phenomenon exhibited by English) seem to only occur with an interpretation as 'given'. The fact that subjects seem to act like discourse anchors, i.e. topic elements of some sort, has been noted previously, but not much syntactic argumentation is usually based on this (see Rizzi's (2006, 2007) treatment of subjects as 'criterial' elements similar to Topic / Focus / Wh-elements). I will be presenting data from English and Dutch that illustrate the relationship between expletive insertion and the interpretation of subject DPs and fronted PPs, whether reflected by morphology or not. What I argue for is that a 'giveness' requirement must be met at the relevant syntactic site and expletive insertion is one of the available strategies for doing so.

Immediately adjacent to the issue of expletives is, quite trivially, the nature of subjects in general. In fact, it is a well-known issue in the literature that the notion of subject in itself is rather hard to define. Falk (2006) for instance, has a thorough discussion of subjects, attempting to bridge the divide between typological description and formal explanation. He attempts a formalization in terms of LFG, associating several 'attributes' with prototypical subject functions. Depending on the language, a 'subject' will correspond to one or several of these attributes. In most of generative syntax, subjects are usually equated with a requirement on a specific clausal position. Most models therefore cluster around the idea of some position (the specifier of an inflectional phrase, in the most basic version) that must always

be filled by some constituent. The idea of an 'EPP' requirement is thus established, whether in the technical sense initially proposed in Chomsky (1982) or as a descriptive label for the phenomenon of subject obligatoriness. I will refer the reader to Svenonius (2002), which is a good overview of the issues involved in discussing subjects, expletives and the EPP. As will be detailed below, I attempt to identify a more substantive property that characterizes subjects and similar elements and that is involved in their obligatoriness (when this is the case).

Although I do not attempt to discuss more than the Germanic language space, there is still a wealth of subject-related phenomena that should be kept in mind. A rough description of this domain in Germanic can be set up along these lines: the languages under discussion all have the particularity that minimally a constituent must precede the verb. The agent argument of verbs would be a prototypical case of such a constituent. However: passive constructions allow a non-agent to function as subject, while in some languages a 'structural' subject can appear in a case other than the nominative (such as (experiencer) quirky subjects in Icelandic). What is more, there seems to be a difference between languages like English, which prefer noun phrases in the nominative as fillers for this position and languages like Dutch, which can fulfill the requirement by using a diverse array of phrases (subject or object noun phrases and prepositional phrases). As has already been discussed, when the subject position is filled by this latter strategy (discussed as V2 in the literature) or by inversion in an interrogative context, expletives either surface or not, depending on the language. This suggests that some languages need to fill-in one unique position, whether by a DP, PP or some wh-phrase, while other languages, independently of what occurs in the clause-initial position, also need an element in what is conventionally called the SpecIP.

The fact that English allows fronting locative PPs ('locative inversion') as a strategy that competes with lexical subjects and expletives seems to be crucial for explaining what exactly characterizes the obligatoriness of subjects. In English, therefore, we can have:

(38)   **Three clues** are in the living-room. (lexical subject)

(39)   **There** is a clue in the living room. (expletive subject)

(40)   **In the living room** is a clue. (locative PP)

However, other kinds of PPs and DPs interpreted as objects cannot be fronted similarly in English. These can only be fronted when an EPP/SpecIP filler-element (one of the three above) is already in place, giving rise to the following contrast:

(41)   In the castle lived a great wizard.

(42)   A great wizard lived in a castle.

(43)   * Spells did a great wizard.

(44)   Spells, a great wizard would have no problem with.

English locative inversion is not, therefore, analogous to a strategy like V2. It also casts some doubt on the idea / perception that the site of the traditional EPP has an exclusive affinity for (active or passive) subject DPs. An alternative way of discussing this position (along with the requirement that it be filled) is by formulating two alternative requirements: either a subject DP or a locative element must fill-in this position. The second case would refer to both the expletive (which in the case of English happens to have locative morphology) and locative PPs. It is clear however that such a formulation would be quite stipulative and not hold cross-linguistically. What I then argue is that all such elements have a common property, that they are inherently endowed with in the Lexicon. The following discussion aims at identifying this property.

It is important to note that the relation between expletive insertion and locative inversion (LI) seems to have been misrepresented in the literature, in the sense that the properties of the DP that is the complement of the preposition are what counts, and not the preposition in itself. A closer look at the properties of inverted PPs, suggests the environment is sensitive to definiteness / givenness / specificity (and for the purposes of this study I will use the three terms interchangeably since they do not have a direct effect on my analysis). We would expect expletives to satisfy this requirement in some sense. The following examples suggest that expletives are inserted in order to license clauses when the other fronted constituent does not satisfy a condition of giveness in SpecIP, and that a new / indefinite constituent (whether DP or PP) cannot generally satisfy this condition.

(45)   **In the castle** lived a great wizard.

(46)   In the castle **there** lived a great wizard.

(47)   In a castle **there** lived a great wizard.

(48)   ?* In a castle lived a great wizard.

Native speakers of English[7] seem to consistently find (48) worse than all the others. Therefore, LI is only an EPP strategy insofar as the fronted PP is interpreted as given, whereas expletive subjects satisfy the EPP by definition, so that when the available PP does not, an expletive is plugged in, such as in (47). The contrast between definite fronted PPs optionally co-occurring with expletives, and indefinite fronted PPs obligatorily co-occuring with expletives, is also found in the following Dutch examples:[8]

---

[7] I thank Heidi Klockmann and Will Schuerman among others for these judgments.
[8] Eric Reuland, p.c.

(49) Ik vraag me af   wie of (er)   in het kasteel woont.
     I ask   me V.PART who if (there) in the castle   lived

     I wonder who lived in the castle.

(50) Ik vraag me af   wie of *(er)   in een kasteel woont.
     I ask   me V.PART who if *(there) in a   castle lived

     I wonder who lived in a castle.

Here, an additional discussion is required: what would happen if an indefinite DP was fronted? For English, it is commonly argued that the [-specific/-given] post-verbal DP in a LI or expletive construction can be fronted and gain a possible [+specific] interpretation, while keeping the [-specific] one, thus:

(51) In the castle lived **a great wizard**. (-specif / * +specif)

(52) **A great wizard** lived in the castle. (+specif / -specif)

The [+specif] interpretation is to be understood in the sense that the individual is part of an already given set of such individuals. As we expect, given the definiteness restriction, such an interpretation is not possible in the post-verbal position. However, I take issue with the claim that a [-specif] interpretation is available when the DP is fronted. Note that in actual speech there would be several strategies (in different contexts of utterance) that make *a great wizard* [+specif], without necessarily referring to an already established set. It is sufficient for some aspect of the individual to be known, whereas the new or contrastive part will receive stress in speech:

(53) A *great* wizard lived in the castle.

(54) A great *wizard* lived in the castle.

As far as I can tell, these are all the available options for the interpretation of the respective DP and they are all [+specific / given] in some sense. For Dutch it has been noted that such morphologically indefinite DPs are ruled out from this position (i.e. without an interpretive repair strategy being available). Compare the following examples taken from Reuland (1988):

(55) Fred believes that **a cow** is in the garden.

(56) *Fred denkt dat **een koe** in   de tuin is.
     Fred   thinks that a   cow in the garden is

(57) Fred denkt dat **de koe** in   de tuin is.
     Fred thinks that the   cow in the garden is

These facts have received the interpretation that the Dutch indefinite article cannot receive the [+specific] interpretation of the English one, hence the whole construction is ruled

out. This might partly motivate the apparent difference in the strength of judgments between (48) and (50) Note that bare plurals fall under the same restriction, while numerals are protected by the same interpretive strategy as English *a*, forcing their [+specific] reading:

(58) *Fred denkt dat **koeien** in de tuin   zijn. [*+specif. / *-specif ; bare plural]
      Fred   thinks that cows   in the garden are

(59) Judy dacht   dat **zes mannen** arriveerden. [+specif. / *-specif.; numeral]
      Judy thought that six   men   arrived

As a native speaker of Romanian, my judgements on the interpretation of indefinite DPs are as follows. When the DP occurs in the prototypical subject position, a [+specific] interpretation is forced, whereas in the post-verbal distribution it receives a [-specific] interpretation, just like in English expletive constructions or LI sentences. Consider:

(60) **Un om** e în cameră. [+specific / given set]
      a.man   is in room

      A man is in the room.

(61) E **un om** în cameră. [-specific / unpredictable information 'existential']
      is a.man   in room

      There is a man in the room.

(62) În cameră e **un om**. [-specific / LI]
      in room   is a.man

      In the room is a man.

Given these facts I am fairly confident in assuming that the prototypical subject position (universally) forces a [+definite / given] interpretation and rules out constructions that cannot receive this interpretation (due to language-specific parameters). This is in symmetric distribution to the 'definiteness restriction' of the post-verbal position. A classic study of definiteness effects and their interaction with morphology, syntax and pragmatics is Reuland and Ter Meulen (1987). I will however not go into the proposals therein, but rather keep the basic hypothesis of these two symmetric effects and discuss some Dutch facts that complement the discussion of the interaction between [+/- specific/given] items and the presence of expletives. Consider the following, taken from Reuland (1988):

(63) Fred denkt dat *(er)   een koe in de tuin   is.
      Fred thinks that *(there) a   cow in the garden is

      Fred thinks that there is a cow in the garden.

(64) Fred denkt dat *(er)   koeien in de tuin   zijn.
      Fred thinks that *(there) cows   in the garden are

      Fred thinks that there are cows in the garden.

Clearly, constructions that were previously ruled out by their obligatorily [-specific] interpretation are licensed by the insertion of the expletive, much like in the case of the indefinite PP in (50). What is more, constructions with numerals are interpreted as [-specific] under this configuration. Observe the contrast:

(65) Fred denkt dat twee koeien in de tuin   zijn. [+specific]
      Fred thinks that two cows   in the garden are
      Fred thinks that two cows are in the garden

(66) Fred denkt dat **er**   twee koeien in de tuin   zijn. [-specific]
      Fred thinks that there two cows   in the garden are
      Fred thinks that there are two cows in the garden

Given the discussion so far I propose that expletives have the primary purpose of satisfying such a [+specific] requirement in the subject position, competing with [+specific] DPs and PPs.

This suggests that whatever is in the subject position is always interpreted as given, discourse-wise. When no argument qualifies, some filler is selected. These facts seem to be absolute . What we need to explain is the basis on which these fillers are selected. One option would be that they have a certain feature which licenses them. However, no substantive stipulations should be made, since, in the other cases of subjects/fronted PPs etc. we have seen it is simply some 'given' interpretation (sometimes coerced even when the morphology seems to be incompatible) that licenses them. So, expletives should satisfy the same requirement. The question is how individual languages decide on the better suited element from their inventory. Note that it is not the case that several functional (!) elements are compatible with this position in one language. What we therefore need to establish is a competition model that determines what unique element from a language's inventory will function as an expletive: is a locative element or a personal pronoun 'given' in the required sense? A sketch for such a model will be presented in the last subsection of the present section.

4.2   An extension to relative pronouns

In the previous subsection I have suggested that locatives can be selected as expletives by a competition model that targets [+given] elements. In Danish however, if we regard morphology as meaningful, the locative appearance of 'relative' *der* remains unexplained, in spite of my proposal for the competition-based explanation of the locative expletive. Below I sketch out a model that accounts for the range of use of Danish *der*, building on my previous observations and Taraldsen (2002).

First, I summarize the relevant facts and part of the analysis used in Taraldsen (2002), who starts from the following alternation in French:

(67)   Quel livre crois-tu   que/*qui les filles vont acheter?
       which book think-you that   the girls will buy

(68)   Quelles filles crois-tu   *que/qui vont acheter ce   livre-la?
       which   girls think-you that   will buy   that book-there

We see that *qui* correlates with an extracted subject, while *que* is the standard complementizer of French, as in (68) or:

(69)   Je sais que tu as raison.

Taraldsen notes that this property would be uncharacteristic of West Germanic agreeing complementizers, contrary to what Rizzi (1990)[9] claims. In fact, the pattern is more similar to one attested in the Rhaeto-Romance variety of the Engadine: Vallader (Taraldsen 2002:30). Consider:

(70)   Qual cudesch crajast   cha/*chi las mattas cumpraran?
       which book   think.you that   the girls   will.buy

(71)   Qualas mattas crajast   chi/*cha cumpraran quel cudesch?
       which girls think.you that   will.buy   that book

The same subject versus object extraction facts hold as for French. Vallader however, exhibits two additional patterns that seem to be connected to the same alternation. Firstly, *chi* can introduce sentences from which no subject has been extracted, as long as the subject is not in its prototypical position, such as *quels temps docts* in:[10]

(72)   ... la   spranza chi/*cha   turnaran quels temps   docts.
       the hope that   will.return those   times learned

When the subject occupies such a non-prototypical position in a non-embedded clause, a free expletive *i* surfaces instead:

(73)   I turnaran quei   temps docts.
       it will,   return those times learned

Taken together, these facts suggest that the *i* in *chi* correlates with the absence of a subject from its prototypical position. Then, Taraldsen suggests, *chi* is only *cha+i*, and the analysis can be extended to French, where *qui* should be understood as *que+i*. Note that the French subject expletive *il* is frequently pronounced as *i*, at least in informal speech

---

[9]qtd. in Taraldsen (2002): Rizzi, L. 1990. *Relativized Minimality*. MIT Press, Cambridge, Mass.

[10]The reader is lead to understand that in a parallel example where the nominal phrase sits between the complementizer and the verb, *cha* is the required form.

and non-standard varieties, and although postponed subjects do not correlate with the *qui* complementizer, an apparently expletive *i* also appears in another context: complex subject inversion. Taraldsen (2002:32) gives the following example:

(74)   Pourquoi tu   dois-ti partir?
         why   you must-ti leave

The *i* in *ti* is then identified as the expletive, which has been plugged in the position that is normally reserved for subjects in interrogative (inverted) contexts. Since the pronominal subject is higher under complex inversion, the expletive must fill-in the empty slot.

(75)   ... [dois-tu] ...

(76)   ... * tu [dois ‖] ...

(77)   ... tu [dois-(t)i] ...

This analysis suggests that expletive subjects (or rather an expletive condition on the empty subject position) can have a morphological effect on the complementizer. I use this line of reasoning to account for Danish *der*, which functions both as an expletive and as a particle that introduces subject relative clauses ('relative *der*' as Vikner (1991) calls it). Before tackling the Danish facts, some very well-known English facts should be considered:

(78)   The man that I saw. (OBJ)

(79)   The man NULL I saw. (OBJ)

(80)   The man that saw me. (SBJ)

(81)   * The man NULL saw me. (SBJ)

The asymmetry between the second and fourth examples (extracted object vs extracted subject) is coherent with the idea that extracted subjects require the obligatory presence of some morphological material (ignoring the concept of silent expletives). English obviously satisfies this requirement in embedded contexts in a disjoint manner from the way it satisfies the standard expletive requirement. Vallader, as we have seen, resorts to one overarching strategy. Danish, I suggest, is able to do what Vallader does, with the exception that the expletive completely takes over the role of the complementizer, instead of just coalescing with it.

In the literature that deals with the status of *der* in Danish, there are two main lines of reasoning. Vikner (1991) can be taken to represent an approach where the two occurrences of *der* (relative and expletive) are formally different: the relative is a head which occurs in C^o, while the expletive is an XP which occurs in SpecIP. This state of affairs however seems to resist any explanation in terms of the status of this morpheme in the Lexicon, and at

the same time does not make clear what drives the difference cross-linguistically. Mikkelsen (2002) on the other hand, assumes that we are dealing with the same expletive *der* in both cases.

Of the two, based on the observations made in Taraldsen (2002), I believe Mikkelsen must be on the right track. If we are able to explain the need for an expletive as I do in the previous sub-section, then most issues are solved. What is not clear however, is why exactly a complementizer does not surface alongside the obligatory subject expletive. I will not delve into this technical issue but rather point out a more immediate empirical concern. In a North Frisian dialect (Fohr-Amrum to be precise), an element *diar* does the same as *der* in Danish, only it can also be used when an object has been extracted from the embedded clause.[11] Such cases make it hard to explain the locative morphology if this morphology is exclusively characteristic of subject fillers in the left-wise clausal domain. I integrate such facts in Section 6 and propose a solution, in the guise of a framework that is suited to deal with massively poly-functional morphemes. In what remains of this section I draw some temporary conclusions, but not before attempting to show how a competition model would single out locatives as the best-suited '+given' elements in some languages' inventories, but not in others'.

4.3   Half-time: some conclusions

The previous two sections have been dedicated to elucidating the reasons behind the locative morphology that characterizes expletive subjects in some Germanic languages. In Section 3 I have suggested that current approaches are not satisfactory in terms of giving a principled explanation for the split between languages which choose a locative for the expletive function and those that do not. I claim that this split should be derived by a competition model, articulated as follows:

(82)   a.   The prototypical subject position needs to be occupied by an element that is interpreted as [+given].

   b.   Competition establishes which lexical entry from a language's inventory fits better.

In this model, locative words are better fits for the [+given] requirement in some languages. It must therefore be explained why English *there* out-competes *it*, but Norwegian *der* for instance loses to *det*. Following the guidelines for parametrizing linguistic variation I present in Section 2.2, such differences should not be stipulated (at the level of the language)

---

[11] Hoekstra(2001)

but integrated with the inherent specifications of such words in the language-specific Lexicon. Such differences are expected to be seen in other distributions of the respective words. It can be shown that English and Norwegian, for instance, exhibit a distributive contrast between (adverbial) locatives. This can be taken to correlate with different formal specifications for those lexical entries, hence with a difference at the level of how the expletive-competition is resolved. Consider the following construction from Norwegian:

(83) en sånn **her/der** bil
      a   such here/there car
      a car like this

I believe that we can treat this distribution as adjectival, whereas the standard usage of locatives (such as in English) is pronominal. The pronominal distribution will correlate with a [+given] interpretation, while the adjectival one with no such specification. Therefore, English would only have a [+given] version of the locative, whereas Norwegian would have both possibilities. If our competition model for selecting expletives were to trivially look at the available specifications of competitors, the Norwegian expletive should do the same as the English one. I will therefore propose that the competition model looks for an element that is unambiguously [+given]. In a sense, English locatives can be regarded as 'strong' (exclusively pronominal) and Norwegian locatives as 'weak' (both pronominal and adjectival). We see that when the locative is 'strong', the pronominal (*it* in English) loses. This might be because the best competitor would be the one that not only satisfies the requirement but that also brings a minimal amount of features into the computation. Pronominals automatically come with a specification for number and this needs to be resolved at the level of agreement.

Two further issues arise. The first one is: what would motivate such a competition model, especially the preference for forms that have minimal feature endowment. The second one: does the English vs Norwegian pattern apply to other (Germanic) languages?

First of all, the principles that dictate how competition behaves in these cases seem to come from different modules of the linguistic machinery. For a detailed account we would have to characterize these modules and their interfaces. In short, it looks like the preference for minimal feature specifications (especially phi-features) has to do with a syntactic preference for resolving as few agreement-based dependencies as possible. The preference for 'strong' (unambiguously [+given]) elements in the prototypical subject position could be delegated to the module that mediates between syntax and information structure (IS). We have seen that elements in the prototypical subject position are obligatory and have a [+given] interpretation. In this sense, they function as topic-elements that anchor utterances in discourse. I tentatively propose that a clause has a corresponding 'information contour':

arguments derive their place in discourse in a contrastive manner. A [+given] anchor is needed in order to establish the status of other arguments. Thus, an element that does not always arise in [+given] (pronominal) distributions is 'weak' (like Norwegian locatives) and cannot establish the obligatory discourse-oriented contour. The contrastive patterning of arguments at the level of IS bears some resemblances to the idea of an accessibility scale for the argument to discourse-function mapping. Since it is only a suggestion for a future implementation, I stop here with this side-note concerning competition, the interfaces and discourse. In fact, in the following paragraph I will argue that we cannot rely exclusively on the English-Norwegian contrast I describe above, hence attempts at establishing a robust model will have to wait for more robust data as well.

We are therefore left with the question of cross-linguistic applicability for the English - Norwegian contrast. Trivially, Swedish exhibits the same 'adjectival locative' as Norwegian, as well as the pronominal expletive. Dutch on the other hand, lacks such a construction and has a locative expletive, exactly like English. Yet, two languages resist integration so far. Firstly, there is Smith Island English: it has already been mentioned that in this dialect the expletive is *it*. We would expect, considering my model of competition, that the dialect also exhibits adjectival locatives, like Norwegian and Swedish. I do not have access to such data but I believe the possibility is rather slim that the dialect will differ from Standard English on this count. Secondly, Danish arises as a strong counter-example: although it has a locative expletive like English, it exhibits the same type of construction as Norwegian and Swedish, paralleling (83). For these reasons, while maintaining my idea that we need a competition model that makes use of contrasts such as the one discussed above, I will refrain from claiming that I have a found a definitive solution for the parametrization problem. To some extent, I believe to have stumbled over the form of the solution, while its substantive content still eludes me.

In what concerns the status of locative expletives in the Lexicon, I believe the model that I sketch can work with a minimal representation of locatives. The expletive and locative do not need to be separate entries. The locative is chosen for the expletive function because of an inherent feature it has, yet it is not targeted because of its locativity. Thus, we do not need to postulate a role for locativity in the syntax of expletives, or work with the notion of two disconnected but homophonous entries. However: even if this model can also integrate the relative use of Danish *der*, the extended range of usage that we see in North Frisian or Dutch suggests that a stronger system is needed. As a consequence, Section 6 will show how a 'decompositional' view of locatives would derive more empirical facts than the 'atomic' view I just established.

5 The Dutch functional cluster

After discussing the place of locatives in expletive constructions and their connection to Danish 'relatives', I will proceed to another intriguing range of distributions for locatives. Dutch *er* is such a locative, and it comes in two additional flavors, on top of the expletive use and its core locative interpretation. Note that compared to *daar* ('there'), it does not take part in a proximal-distal opposition, but is regularly translated as 'there' in English. In Standard Dutch, it is this element that plays the role of expletive, but in certain dialects forms that are closer to *daar* are used. In the following subsection I present the distribution of *er* in the context of prepositions and it will be seen that *daar* participates in a similar alternation: while *er* replaces the non-human personal pronoun, *daar* replaces the non-human demonstrative. Therefore, descriptively, Dutch has a full paradigm of locatives corresponding to each relevant type of pronominal, while English only has the form *there* that can be seen to occasionally (but more rarely than Dutch) compete with a pronominal.

The distributions of Standard Dutch *er* will be discussed in the first subsection, while the second subsection presents the *(e)r / (he)t* split in dialects, i.e. how the functions of standard *er* can be divided between different morphemes.

5.1 The r-word conundrum

Van Riemsdijk (1978) presents an intriguing paradigm of words that are in a certain sense locative but are in fact used to replace non-human pronominals in the context of a preposition. Morphologically, they all contain a salient *r*, which is characteristic of locative terms in Dutch especially, but also, as we have seen, a good indicator of locatives in English (*here, there, where*), Danish (*der*, see Section 4.2) etc. This observation motivates the terminology *r-pronoun* or *r-word*. An intriguing property of such forms is that they do not only replace a regular pronominal form, but also surface to the left of the preposition that is apparently involved in the phenomenon. What is more, unlike the regular (animate) complements of prepositions, these are free to raise, stranding the P(reposition). Van Riemsdijk's (1978:37) pattern follows:

(84)   a. op hem ('on him' [+human])

 b. * op het ('on it' [-human])

 c. * op er ('on there' [-human / locative])

 d. er op ('there on' [-human / locative])

(85)   a. op die ('on that' [+human])

 b. * op dat ('on that' [-human])

    c. * op daar ('on there' [-human / locative])

    d. daar op ('there on' [-human / locative])

(86)   a. op wie ('on whom' [+human])

    b. * op wat ('on what' [-human])

    c. * op waar ('on where' [-human / locative])

    d. waar op ('where on' [-human / locative])

The (a) examples illustrate the relative distribution of [+human] pronouns and their dependent prepositions, while the cases under (b) show that this pattern is unavailable for [-human] pronouns. As is apparent from the translation, English does not present such an asymmetry. Then, it is the pattern under (d) that is grammatical in Dutch: [r-word P]. Note that the reverse order would be unavailable, as suggested by the (c) examples. The three groups of examples illustrate the fact that this shift in morphology and distribution is triggered by prepositions for several pronominal classes: personal, demonstrative, interrogative, each with its own specific morphology that carries over to the locative (d-forms for demonstratives, w(h)-forms for interrogatives). Within these classes, the *er / het* alternation is the one I will focus on. *Er* is particularly interesting in that it has not only this 'prepositional' distribution, but also a quantitative one (similar to partitive clitics in French or Italian), as well as the expletive subject function, paralleling English *there*. In what follows, I will look at this extended range of functions from the perspective I outline in Sections 1 and 2.

A thorough discussion of the functions of *er* is featured in Bennis (1986). He illustrates each function and discusses some salient properties thereof. Another striking aspect of his presentation is the discussion of interactions between these multiple functions. Additionally, his reflections on the status of this type of morpheme are insightful. Recall the lines I quote in my own conceptual preamble (Section 2), especially the following stance:

> " I agree with Kirsner that an optimal theory of *er* should relate all occurrences
> of *er*. I do not agree with him that such a unificatory approach should necessarily
> be based on a particular meaning of *er* that is present in all specific instances."
> (Bennis 1986:171-172)

In the next section I will explore this statement in its formal aspect and argue for a contemporary framework in which it can be captured. Before that, I will go through Bennis' descriptive exploration of the domain of *er*.

First of all, one needs to discuss the expletive subject function of *er*. Bennis (1986:173-175 etc.) talks about such 'dummy elements' and claims that they stem from 'independently

motivated principles' (and not some EPP requirement). Expletives seem to correlate with de-
moted subjects, however this does not mean they are syntactically linked to the co-occurring
nominal argument, especially since in languages like Dutch expletives can front sentences
that lack a nominal argument ('logical subject'), such as passives, as illustrated in (88). The
other core properties of expletive constructions, as discussed in Section 3, also hold in Dutch.
Consider the following illustrations from Bennis' discussion:

(87) Er   loopt een jongen in de tuin.
      There walks a   boy   in the garden.

(88) Er   wordt gevoetbald.
      there is   soccer.(play)ed

      Soccer is being played.

The following function of *er* discussed by Bennis (and a central one in the literature,
at least after Van Riemsdijk's 1978 analysis) is the 'prepositional' one, as discussed and
illustrated in the beginning of this subsection. It is assumed that the adverbial pronoun
functions as an argument of P, but one generated to its left, though inside the PP. This
is argued to allow the stranding of the preposition. An interesting fact is that not any
preposition can participate in this construction. Prepositions like *over, met, van, tegen, in*
and *op* prototypically require r-pronouns, whereas *zonder* does not, as indicated by Van
Riemsdijk (quoted in Bennis, 1986:191). Again, Bennis' (1986: 176) illustrations:

(89) Ik heb met hem [er   over] gesproken.
      i   have with him there about talked

      We have talked about it.

(90) Ik heb *(er) met hem [-   over] gesproken.
      i   have there with him talked about

The two sentences are equivalent, only it is to be noted that in the second example the r-
pronoun can be separated from its preposition, as long as it does not disappear completely, as
indicated by the *(). Bennis treats such (left of P) PP-internal gaps as traces, but I will not
use such a specific notation. As already mentioned, these are striking facts, in the sense that
they introduce a [+human]/[-human] asymmetry in terms of both the morphology and the
position of the element that accompanies the preposition. The fact that such elements have
a locative reading is in itself problematic. Next, Bennis (1986) discusses the quantitative
function of *er*. This refers to distributions in the context of a "quantified NP containing an
empty head or N' " (Bennis 1986:177). The phenomenon is comparable to the use of French
en and Italian *ne* (with an important difference though). Bennis also notes that there are
no "r-variants" for this distribution. Whereas prepositional *er* has the demonstrative and

interrogative variants, and the locative can be replaced with an explicitly proximal or distal r-word, the quantitative function can only be expressed as *er*. Consider the following example (Bennis 1986:177):

(91) Ik heb er   gisteren [twee-] gekocht.
      I   have there yesterday two   bought
      I bought two yesterday.

An important restriction on the distribution of quantitative *er*, which sets it apart from its French or Italian counterparts, is that the nominal 'gap' must refer to a countable noun, reflected or not in the presence of the numeral. Therefore, Dutch can use such constructions when referring to bottles, but not wine, whereas French has no such restriction. Compare:

(92) Biertjes, ik heb er   gisteren ([twee-]) dronken.
      beers   I have there yesterday (two)   drank
      Beers, I drank some (/two) yesterday. [+count]

(93) *   Wijn, ik heb er   gisteren dronken.
      wine I   have there yesterday drank
      Wine, I drank some yesterday. [-count]

(94) Bieres, j'en   ai (trois) bu.   (French)
      beers, I.PartCL have (three) drank
      Beers, I drank three.

(95) Vin, j'en   ai bu.   (French)
      wine, I.PartCL have drank
      Wine, I drank some.

Note how in the English translations, either a numeral or *some* must appear, where *some* can receive either a countable or a mass interpretation. Thus, *some* and *en* pattern together, as opposed to Dutch *er*. As a pattern of variation, this is reminiscent of the rigidly [-specific] interpretation of Dutch indefinite articles and bare plurals, as discussed in Section 4.1.

In addition to the descriptive facts, it must be noted that Bennis indicates the presence of an empty category *e* in the gap, unlike the trace used for the prepositional environment.

The functions of *er* are exhausted once the locative one has also been discussed. Bennis (1986:177-178) makes an interesting distinction between environments where a locative *er* is obligatory since the verb requires a locative argument, and environments where *er* is optional, paralleling the addition of optional PPs in a clause. In both of these environments *er* functions as a minimal filler (with a spatial interpretation), that can be replaced by r-variants and other locative constituents (PPs). Clearly, under the terminological system I

proposed in Section 2.1, this is a contrastive distribution. Bennis' examples follow, as an illustration:

(96) Jan koopt (er)   een boek.
      Jan buys (there) a   book
      Jan buys a book (there).

(97) Jan woont *(er) nog maar kort.
      Jan lives   there only briefly
      Jan has been living there briefly.

To summarize the picture of distributions so far, I will group the prepositional, expletive and quantitative as exclusive distributions, where *er* is obligatory and does not have contextual r-variants, while the locative distribution is contrastive, in allowing *er* and its variants. The exclusive distributions are rigidly determined as contexts where *er* receives the particular interpretation, whereas the locative flavour of *er* can conceivably appear in any clause.

Another interesting phenomenon extensively discussed by Bennis (1986) is the co-occurrence of multiple functions of *er*. He notes that these can surface as one form, although formal diagnosis suggests that all of the functions are present even if not individuated morphologically. At least seven such combinations are possible. First, I will illustrate the co-occurrence of expletive and prepositional *er* with one realization and present a diagnosis for the construction, as conceived by Bennis (1986:178).

(98) ...dat (er)   twee jongens [-op] zaten.
      ...that there two boys   on   sat

   a. *...dat er   deze jongens [-op] zaten.
      ...that there these boys   on   sat

   b. *...dat deze jongens [-op] zaten.
      ...that these boys   on   sat

The version under (b) is ruled out since the definite nominal (interpreted as an active subject) is incompatible with the expletive, as discussed previously. This suggests that *er* cannot drop the expletive interpretation in this context. The example under (c) is not improved by simply dropping the *er*, since the preposition requires its presence in the prepositional flavor. Thus, (a) is the only felicitous version, and it has been shown that despite the unique realization of *er*, both the prepositional and expletive function are represented (at some level of representation). Bennis also notes examples of the expletive co-occurring with the quantitative and the locative:

(99)   ...dat er   niemand [twee-] gekocht heeft.
       ...that there nobody two   bought has

(100) ...dat er   niemand meer   woonde.
       ...that there nobody anymore lived

It is clear how in both the indefinite *niemand* allows the expletive, just like *twee jongens* above. In (99) the *er* must obligatorily also function as a quantitative, while in (100) it needs to function as a locative, since a locative argument is obligatory in the context. A prepositional and a locative can also co-occur, as illustrated in:

(101) ...dat Jan er   [twee-] [-over] gelezen heeft.
       ...that Jan there two   about read   has

More instances of *er* (at least up to three kinds) are also possible in the same clause, as exemplified in:

(102) ...dat er   niemand [drie-] [-over] geschreven heeft.
       ...that there nobody two   about written   has

(103) ...dat er   nog maar [twee-] -   woonden.
       ...that there yet only two   lived

In the first example, by applying the diagnoses used previously, one can highlight the presence of expletive, quantitative and pronominal *er*. In the second one, the expletive also hosts a quantitative and a locative.

An additional phenomenon mentioned by Bennis is the co-occurrence of two instances of the same kind of gap, requiring two differently linked (indexed) abstract instances of *er*. In combination with the patterns above, this can result in clauses where one realization of *er* satisfies several requirements (perhaps four or more). Nevertheless, it would seem *er* can be realized twice. In the first example to follow, one position is the expletive and the other one is lower, standing for a quantitative (and we expect the same to happen for the prepositional and locative. The second one (taken from Barbiers 2009b) illustrates the presence of two quantitative instances of *er*, while the next one shows that they can be conflated in the standard position to the left of the leftmost numeral.[12]

(104) Er   lopen er   [twee-].
       there walk there two

(105) Ik heb er   [twee-] er   toen [een-] gegeven.
       I have there two   then one given

---

[12]It would be interesting to see whether such a conflation requires that both quantities refer to the same reference set of items, though it would be conceptually difficult to find an example where they do not.

(106) Ik heb er   twee toen een gegeven.
       I   have there two then one given

There are therefore three issues to be explained concerning *er*: number of functions, *het /
er* alternation (induced by some prepositions), r-movement and r-conflation. The following
subsection presents the dialectal picture of the morphological split between the functions
of *er*. Then, Section 6 attempts to derive a picture that captures the facts in terms of
representation and computation, suggesting a possible way of lexical organization for poly-
functional words. The problematic issues identified by Van Riemsdijk, Bennis and others in
the syntax of *er* are also accommodated in the model I adopt.

5.2   Morphological splits in Dutch dialects

After having presented the poly-functional status of Dutch *er*, I turn to dialects where these
functions are distributed between several morphemes (a 'split' similar to the inter-linguistic
one concerning expletive subject morphology). I show that there are apparently limits to
this kind of variation, suggesting the fact that the morphological patterns might reflect some
principles of linguistic organization. I will make reference to both some previous litera-
ture and also independent research I conducted through the DynaSAND[13] and DIDDD[14]
databases. The better-known facts from previous literature regard dialects where the exple-
tive subject has a *(he)t*-form. I have already presented such a pattern in Section 3. However,
the prepositional, quantitative and locative are not expected to use the same morphology in
the respective dialects, which results in a Expletive vs. Prep/Quant/Loc split. Note how this
state of affairs already starts casting some doubt on theories that would have all instances of
er reflect the same central property, as Bennis also remarks. In addition to this, I also look
at dialects where the quantitative does not take an *r*-form, and even present cases where
based on the gender of the 'elided' noun, the quantitative can take either an *r* or a *t*-form.
Morphology therefore seems to support a formal split between the features that play a role
in the distributions of Standard Dutch *er*.

Formally, the morphological splits in dialects suggest that we are not dealing with a
grammatical functor [X] which is associated with a unique morpheme /Y/ (possibly differing
across dialects), but rather with a series of functions [A-B-...-*n*] which can be associated to an
expected maximum of *n* morphemes. The Standard Dutch case would qualify as complete
morphological conflation. The relation between functional granularity and morphological
opacity will be formally explored in the following section, where a specific framework is

---

[13]Barbiers, S. et al. 2006. *Dynamische Syntactische Atlas van de Nederlandse Dialecten (DynaSAND)*. Amsterdam, Meertens
Instituut.

[14]Corver, N. et al. 2011. *Database Diversity in Dutch DP Design (DDiDD)*. Utrecht, UiL-OTS.

introduced in order to capture this dimension of variation. At this point, we must establish possible restrictions on how the series of functions can be mapped onto morphology, i.e. whether the four-way distinction is ever instantiated in morphology, what functions can be expressed by what morpheme and so on. Following in the Greenbergian tradition at least in spirit, I attempt to identify some generalizations in this domain, whether absolutes or just tendencies. The focus is on contrasting the distribution of *t*-forms with *r*-forms.

As already mentioned, it is relatively common for the expletive subject to be expressed as *(he)t* in Dutch dialects, especially in the southern half of the language area. Looking at the informant responses for SAND test sentence 412 gives a clear picture of the distribution of such forms. The test sentence *Er waren veel mensen op het feest.* ('There were many men at the party.') has been chosen as a standard environment for expletive subjects (existential). Note that for all of the sampled areas, some informants use *r*-forms. What is crucial is that in other areas such *t*-forms are ruled out. Three representative examples follow, and it must be noted that they stand for a very common situation and are not some local anomaly.

(107)   **'t** woaren vele mensen op die feeste (West-Flanders: Bredene, H017p)

(108)   **'t** waren do veel mense. (East-Flanders: Gent, I241p)

(109)   **Ut** wore veul luuj op ut fes. (Dutch Limburg: Simpelveld, Q116p)

For the quantitative function, dialectal morphology does not deviate as strikingly from Standard Dutch *er*. In very few cases is there a *t*-form used, and this tends to correlate with an expletive *t*-form. I do not claim that the correlation is absolute, but I have not found clear counterexamples. Two such instances follow:

(110)   a. Van al disse boek'n bint die dik'n   't   mooiste, mar ik heb **'t** nog gien ene
             of   all these books are the fattest the nicest   but I have T yet no   one
             ekoch. (Central-East Holland: Vaassen, F129p)
             bought

             Of all these book, the fat ones are the best, but I have not bought any.

        b. Van voetbalbillekes gekalt,   van al men kamerote is **to** eine wo   ze alemwol
             of   football-cards speaking of all my   friends   is T one who he all
             hiet. (Belgian Limburg: Tongeren, Q162p)
             has

             Speaking of football cards, of all my friends there is one that has them all.

It must be noted that the case under (a) is an absolute out-lier in terms of geography and also compared to neighboring locations or other informants from the same location. Additionally, in this area expletive subjects never occur as a *t*-form. This suggests further

research is needed to elucidate the case, but in any case, it is worth reporting as one of the few examples that seem to illustrate quantitative *(he)t.*

The dialect of Tongeren (b) seems to be another case where the quantitative function is not expressed as an r-form. The expected correlation is that in this dialect the expletive subject is also expressed as *(he)t.* This is possible (compare with the Limburgish dialect exemplified earlier), though not attested in any of my sources. Whatever the case might be for the dialects I mentioned, there is older literature that records and discusses similar facts.

As reported in Coppen(1991), the morphology of the quantitative (pronoun) can track number (Spaubeek) and gender (Weert). This presents as a much more complex picture than that available for Standard Dutch. What I focus on mainly is the morphological split in the Weert dialect. Consider:

(111)   a. Ik heb d'r ejn/eine   gezeen. (Weert)
           I have R one:[fem/masc.] seen

        b. Ik heb 't ein   gezeen. (Weert)
           I have T one:[neuter] seen

(112)   a. Ich hub d'r eine. (Spaubeek)
           I   have R one

        b. Ich hub d'rs   acht. (Spaubeek)
           I   have R(pl.) eight

Both dialects are from Limburg. In the case of Weert we notice a common versus neuter split (feminine *ejn* and masculine *eine* vs. neuter *ein*), that parallels a split in morphology (*d*-form vs *t*-form). This is an extremely interesting case for the present survey since an *r* and a *t*-form alternate within the dialect. Note how this gender contrast parallels the [-human] restriction on prepositional *er*. The Spaubeek dialect on the other hand marks plurality. Tentatively we can suggest that this is an indication of the status of *er* as true pronominals, rather than some other (functional) category. Concerning the issue of variation with respect to locative morphology, I take such facts as indicators of a more fine grained pattern. How could it be that one 'locativity' parameter would take scope over the quantitative pronoun in only one gender? Clearly, a robust system of representation is needed to capture such subtle differences without wildly stipulative parameters. Note that the phenomenon in Lapscheure Dutch (the availability of both a *t* and an *r* expletive) is analyzed as a contrast between different syntactic positions, of which only the TP (low) expletive can be locative. Therefore, the unitary 'expletive in SpecIP' function is not split morphologically.

In what concerns the prepositional function of *er*, dialects do not seem to allow *(he)t* forms, even in those cases where they do use expletives or quantitatives of this type. This is another

indicator of the fact that there is a formal difference between the different distributions of *er*. For the locative, *er* or *der* is the expected default.

So far I have treated the morphology-to-functions mapping as an orderly affair: the expletive is more freely expressed as *(he)t*, the quantitative in at most a subset of those dialects that have a *t* expletive, and the prepositional is always an r-pronoun. It is not absolutely certain that this is true, but it is a good working hypothesis in the absence of counter-evidence, qualifying as an implicative scale.

Therefore, some dialects suggest there is granularity in the functional range of *er*, while others introduce complete opacity in this respect. We see a scale of non-locative forms and mixed locative and non-locative forms for apparently the same phenomenon, such as the quantitative that tracks gender (Weert). If these were all expressions of a standard filler of some sort, there would be no formal reason for their morphological separation. The overwhelming impression is one of 'different packaging'.

Just like the presence of a morphological split argues against the monolithic nature of the grammatical functor expressed as *er* in Standard Dutch, the scalar aspect of variation in this domain seems to suggest that the 'homophony' answer is not adequate either. That is, we expect some link to exist between the different instances of *er*. The question becomes: how do we parametrize this? The idea of 'different packaging' gives a possible answer to the issue of parametrizing morphological splits in an apparently incremental way. A concrete proposal in this direction is made in Starke (2011). The following section capitalizes on these incremental splits and introduces the core theory that Starke (2011) departs from, while attempting to model the Dutch facts regarding r-pronouns within that model.

6 Incremental variation and competition

After having established a model that tentatively explains variation in the morphological expression of expletive subjects (and Danish relative *der*), I attempt to integrate more problematic distributions of locatives in the picture of linguistic variation I set up. As suggested at the end of Section 4.2-3, the coherence of the model that integrates the expletive and relative functions of Danish is threatened by data from a Frisian dialect that uses *diar* in both subject and object relatives. Section 5 presents additional data from the domain of Dutch r-words (locatives) that suggest the functional range of locatives is even greater. Below, I show that incremental morphological splits can be captured in a model of the Lexicon that also derives the complexities of domains such as r-words.

6.1 Incremental morphological splits

Based on the issues discussed in Section 5 (and as already hinted by the examples in the introduction), we can organize the distribution of locatives in English, Danish and North Frisian (Fohr-Amrum) in a table that shows their incremental nature:

LANGUAGE	Locative	Expletive	S. Relative	O. Relative
Smith Island English	**THERE**	IT	-	-
Standard English	**THERE**	**THERE**	-	-
Danish	**DER**	**DER**	**DER**	-
North Frisian	**DIAR**	**DIAR**	**DIAR**	**DIAR**

Smith Island English does not use the adverbial locative in any grammatical distribution. Standard English on the other hand, uses the locative as an expletive subject. Danish adds the 'relative' function to this scale (introducing clauses from which a subject has been extracted), while the North Frisian variety of Fohr-Amrum generalizes the use of its locative to the front position of clauses from which a subject or object have been extracted. It is interesting to note that none of the languages that come under the scope of my study introduce any gaps in this sequence: if a position in the table is realized as the locative, then all positions to its left are also realized as the locative. This looks like a significant typological generalization. I will then formulate the generalization as:

(113)   a.   Within a language, if a position on the following scale is realized as the locative, then all positions to its left are as well:

   b.   L(ocative) - E(xpletive) - S(ubject relative) - O(bject relative)

Recall how in the previous section I showed that in dialects where the functions of Standard Dutch *er* are split between *r*-forms and *t*-forms, the *t*-forms target the expletive primarily, and only in a subset of those cases the quantitative as well. In some cases the quantitative has both a *t* and a *r* form, depending on gender. The split is then incremental, as can be shown in the following table:

	E	Q	P
a	er	er	er
b	het	er	er
c	het	het	er
d*	het	het	het
e*	er	er	het
f*	er	het	er
g*	het	er	het

The first row corresponds to the pattern we see in Standard Dutch. Row (b) refers to dialects with a *t*-form expletive. Row (c) refers to dialects where the quantitative is also expressed by a *t*-form. There are apparently no dialects where the P(repositional) is expressed as a *t*-form and, just like in the previous table, gaps seem to be ruled out (as suggested by the last two lines). Once more, we have a scale over which a generalization can be expressed, namely the one in (113a):

(114)   L(ocative) - P(repositional) - Q(uantitative) - E(xpletive)

What the generalizations above show is that when a (locative) word expresses multiple functions, those positions must be contiguous in a universal scale. In other words, when a morpheme X targets functions [A] and [C] in a scale [A-B-C], then [B] is also necessarily expressed as X. This type of observation has been used in several recent works in the **nanosyntactic framework**. It is usually referred to as the *ABA principle, adapting terminology used by Bobaljik (2007) in a study of suppletion. Within nanosyntax, patterns that conform to the *ABA are interpreted in the sense that syncretic morphological forms reflect a sub-morphemic level of organization. In his dissertation on morphological case, Caha (2009:17) notes Jakobson's (1962)[15] observation that "syncretism points to the existence of a hidden level of linguistic organization inside an apparently indivisible unit: the morpheme" and proceeds to investigate the possibility of syntactic organization within morphological

---

[15]qtd. in Caha (2009): Jakobson, R. 1962. *Beitrag zur allgemeinen Kasuslehre: Gesamtbedeutungen der russischen Kasus.* In "Selected writings", vol. 2, pp. 23âĂŞ71. Mouton, The Hague.

units, developing the nanosyntactic technology originally proposed in Starke (2009). This line of research will play a role in the analysis I sketch in the following sub-section.

Recall the issue of functional polysemy discussed in Section 2.3. Integrating those concerns with the generalizations formulated in this chapter, the emerging picture is that of locatives being precisely such morphemes that have a rich internal organization. I will explore this hypothesis in the following section by testing it against the behavior of Dutch *er*.

6.2   Nano-locatives

This subsection is dedicated to a brief sketch of the nanosyntactic picture of morphological case under the analysis of Caha (2009) and an adapted version of this analysis that I use to elucidate the behavior of Dutch *er*.

Caha formulates the following main generalization: "Non-accidental case syncretism targets continuous regions in a sequence invariant across languages." (Caha 2009:49). Below I give the most elaborate case sequence that is used in his study, as established by comparing syncretic cases across languages and arranging them so that the *ABA is not offended:

(115)   Nom - Acc - Loc1 - Gen - Part - Loc2 - Dat - Loc3 - Ins (Caha 2009:130)

As already mentioned, he appeals to Nanosyntax in order to explain the systematicity of these patterns. One of the basic tenets of Nanosyntax, which makes it very appealing for dealing with such scales, is the assumption of sub-morphemic organization. That is, morphemes correspond to a well-articulated representation that dictates their range of use and implicitly patterns of competition. They are not therefore associated to unordered bundles of features, but rather to asymmetric trees of terminals (in the sense of functional heads). A case morpheme X can therefore be associated to a syntactic tree that includes terminals corresponding to several adjacent (!) positions in the case scale. It is assumed that there is no operation of 'fusion' in syntax (like in Distributed Morphology) and no NULL morphemes are postulated. Instead, every syntactic structure is lexicalized by using morphological entries that exhaust the range of needed functional heads. These functional heads are assumed to follow a universal hierarchy (in the cartographic sense) and gaps in this sequence are assumed not to exist. This relates to the way semantic computation is mapped onto syntactic structure: the semantics of a structure [X-Y] are derived compositionally by applying [X] to [Y], where [X] is a functional head and [Y] corresponds to a complex structure of heads. The process of selecting lexical entries and lexicalizing the target syntactic structure consists of matching the representational trees corresponding to morphemes to the target syntactic structures. Here, several principles apply that dictate the way competition is resolved. Pertaining to this mechanism of matching and lexicalization is the phenomenon of

Spell-Out driven movement. This phenomenon is responsible for movement without semantic consequences, that is, a surface reshuffling of the order of morphemes, relative to the universal order of their underlying spans of terminals. The motivation for evacuating pieces of structure (with the effect of surface movement) resides in the representations of morphemes competing for the position: certain structures can only be lexicalized once movement has been performed. I will refer to these technical tools in the analysis that follows.

A natural consequence of Nanosyntax, apart from its explanation of syncretism scales, is the model it suggests for linguistic variation. As formulated in Starke (2011), variation is reduced to the size of syntactic trees that correspond to individual morphemes across languages. This would motivate scales of variation in the distribution of a morpheme, such as the ones I point out in the previous subsection. More precisely: a morpheme can take on as many functions as its representation contains different terminals.[16] As these terminals are distributed between different morphemes, different patterns of use can arise.

Let us now look at how such a picture would look in the case of poly-functional locatives. Say A1, A2 and A3 are morphemes in languages L1, L2 and L3. All three have a locative use. If this type of semantics is reflected in syntactic representation (as is commonly assumed), we expect it to be implemented as a terminal, call it X. If A2 also has an additional syntactic use compared to A1, it must be that it corresponds to a sequence [Y-X]. Then, if A3 has a pattern of distribution that points to an additional function, it will correspond to [Y-Z-X]. Consider the case of Dutch *er*: this morpheme seems to have several different functions. We will suggest these correspond to a sequence of terminals [X-Y-Z]. In a dialect where a function of *er* has been delegated to a *t* form, there will be a morphological split between what can express the sequence [X-Y] and what can express the remaining terminal [Z]. A similar story would go for the morphological split between expletive subjects in different languages. In English, Dutch and Danish, the terminal corresponding to the relevant feature of expletive subjects will be a part of the representation of the morpheme that also has locative semantics. The Danish locative (and this has become a label that glosses over the other readings of the morpheme) will have a representation that includes some terminal responsible for the 'relative' function. The problematic North Frisian *diar* will now simply correspond to a syntactic span that includes a terminal that also licenses the 'object relative' use.

Thus, a picture of principled morphological variation emerges, and with it, a tool for syntactic analysis. I will then pursue the options provided by this technology in order to analyze the behavior of Dutch *er*. What we therefore need to capture is the multitude of distributions it has, including the fact that in some contexts it seems to replace *het*, surfacing

---

[16]I refer to 'functions' that are genuinely contrasted at the syntactic level of granularity

in a non-prototypical position. From the nanosyntactic sketch I present above, I will take the following observations to be significant to my domain:

(116) a. There is a scalar ordering of cases (and prepositions): morphemes can compete for continuous stretches of this structure.

b. Displacement of morphemes is a diagnosis for a repair strategy that operates at lexicalization, as a result of the representation of a certain morpheme.

From the empirical perspective, it seems that *er* competes with *het* but occupies an exceptional syntactic position. It has several interpretations and co-occurs with prepositions. This is reminiscent of Caha's case scale and illustrates a predicted effect thereof: as the structure is lexicalized bottom-up (starting from Nominative), once prepositions begin occurring (instead of postpositional marking), all the case positions left will be lexicalized with the help of prepositions. Therefore I will make the following proposals in what concerns the representation of Dutch *er*:

(117) a. It is associated to a syntactic tree that bottoms out in whatever categorial terminal is needed to instantiate (pro)nominal elements, perhaps a head N.

b. The associated tree stretches up to a position in the case scale after which the prepositions start (*van, op etc.*).

c. The associated tree also includes a terminal responsible for the locative reading, one for the quantitative and one for the expletive.

Note that we are dealing with a functional field that starts out at the nominal root N, includes terminals responsible for the instantiation of phi-features, quantification and reference and also includes all of the positions in Caha's case sequence. Clearly, some 'cases' have to do more with theta-roles, others with spatial semantics and others still might be involved in quantification. I will not go into the conceptual and technical issues involved in opting for this representation, but simply attempt to identify the functions of *er* as particular positions in Caha's case scale and show how the other properties are expected to derive. Recall the sequence of cases:

(118) Nom - Acc - Loc1 - Gen - Part - Loc2 - Dat - Loc3 - Ins

I believe that the crucial region of this scale is the [Loc1-Gen-Part]. The Loc1 terminal would be responsible for the locative reading of *er*. The Partitive will allow the Quantitative reading: I have already mentioned how the Dutch Q-*er* is similar to French partitive *en*, except for its [+count] restriction. The Genitive will be associated with the preposition *van* (which is similar in its polysemy to English *of*, but just like *of* is also characterized by the

semantics of possession, traditionally associated to the Genitive). Simply adopting Caha's case scale would give us the following picture:

(119)   (rootN-Phi-...)- Nom - Acc - Loc1 - Gen - Part - Loc2 - Dat - Loc3 - Ins

This structure is the target of lexicalization by items such as *er, het, hem*. First of all, we can assume that the N(ominal) root comes with specifications of the type [gender/+human/-human etc.] Each different class will instantiate a parallel functional span that starts out at the root and can stretch up to the highest case terminal (and beyond) but will be lexicalized by different morphemes. For masculine human animates, the basic morpheme will be *hij* which can lexicalize the functional sequence or span, up to the Nom(inative) position. After this, the object pronoun *hem* is required. The object pronoun's representation includes that of the subject pronoun (since the sequence up to Acc needs to be lexicalized exhaustively even in the absence of *hem*). It is by competition (the traditional Elsewhere principle) that *hij* wins for the low part of the sequence. For the higher cases prepositions will be required: the object pronoun lexicalizes as much as it can, while a prepositional morpheme is used for higher terminals. We would expect this model to carry over for the [-human] pronouns: *het* would lexicalize an extensive range of terminals, with the exception of high prepositions. We see however that this is not the case. When prepositions are used, *het* is ruled out and *er* is plugged in instead, but at an unexpected position. Building on traditional observations on the distribution of *er* I propose that it is simply associated to a structural superset of the representation of *het*, say starting with Loc1 in the case scale. As such, it will be able to get a locative reading as the default. If we equate its quantitative use with the Partitive case and the preposition *van* with the Genitive there will be an ordering conflict: the Partitive / Quantitative would be non-prepositional in the middle of a prepositional field. I propose that in fact inverting the Gen and Part positions in Caha's scale is a possibility that would solve this ordering issue. In the initial table of syncretism patterns (Caha 2009:130), the Partitive is never syncretic with Loc2 and always syncretic with Gen when Gen is syncretic with Loc1. Changing [Loc1-Gen-Part] to [Loc1-Part-Gen] would not pose problems to the initial model and empirical data presented by Caha. Thus, *er* could be associated to a sequence of terminals that includes maximally [root- ... - Loc1 - Part]. The bare morpheme thus comes with the quantitative and locative readings, as expected. The expletive reading might be derived by a model such as the one I propose in Section 4 or simply associated to a case position in this area. Again, cases in this picture should not be exclusively equated with roles of some sort: it can very well be that some discourse properties such as 'given' are mapped onto terminals in this region.

Then, what of prepositions and prepositional *er*? First of all, we will map *van* to Gen

and *met* to Ins(trumental & Comitative), while spatial prepositions such as *op, in, tegen* will correspond to the different locative cases and their flavors (see Pantcheva (2011) for a nanosyntactic exploration of the domain of directional prepositions). Note that the position of the Dat(ive) is a problem for my analysis in several ways since it is part of what I treat as the prepositional field but is expressed by bare object pronouns. I will go around this matter here and show how the system would work in its absence.

At this point I can also sketch a proposal for the position of *er* relative to prepositions. I propose that the mechanism of 'spell-out driven movement' can be held responsible for the *er-hem* asymmetry. This mechanism (which I shorthand as SODM) is explored in more detail in Caha (2010), following on an original idea of Michal Starke. SODM relies on the idea that re-ordering of morphological material relative to the underlying (universal) structure it maps out is a result of strategies that mediate the lexicalization process. Depending on what entries (terminal trees paired with morphemes) are available in a language, it is not always possible for lexicalization to proceed in the expected order. Some structures need to be evacuated (to the left, in an LCA perspective) so that lexicalization takes place. In this way, movement that is not correlated with semantic differences (respective to base configurations) is restricted to the only level at which it would not influence computation (PF) and is parametrized without resorting to language-level parameters. I believe this is what happens in the case of Dutch *er*: because of its own representation or that of higher prepositions (and Nanosyntax makes available several possibilities which I will not detail here), *er* is evacuated to a position above the preposition when lexicalization occurs. If this is the case, then my model holds without stipulating a different category for *er*: it is syntactically parallel to other pronominals but its position is influenced by a representational conflict. Why *hem* escapes this phenomenon might have to do with its possibly not going as far up as the [Loc1-Part] point. This would look like a gap in the syntax of [+human] pronominal sequences. The prepositions above would spell-out the needed terminals without clashing with the representation of *hem*, hence no evacuation would be needed. These suggestions are however only a sketch for capturing the pattern of r-movement. Making the analysis work under the proposed framework is not at all trivial. In addition, I have proceeded from the assumption that this pattern of movement is simply a lexicalization/PF phenomenon. Note that this seems to be indeed the case for expressions such as English *thereon*, in which the morphemes appear in an unexpected order and are frozen together, as it were. However, Dutch *er* crucially allows the stranding of the preposition, so the freezing we would expect to occur once lexicalization has taken place is somehow circumvented. These issues motivate further research in this direction, research which falls outside the scope of this paper.

There are two remaining issues that I propose to solve by assuming that minor differences

in the lower part of the sequence of terminals can influence the semantics of the maximal sequence and instantiate parallel spans. I mentioned that the Quantitative in Dutch is dissimilar to Romance Partitive clitics because it only works for [+count] entities. I believe this has to do with the fact that in Dutch the terminal responsible for these Partitive semantics appears in the functional sequence of pronominals, which are by default [+count].[17] Recall the 'parallel' spans I propose for different classes of pronominals [+/-human]. Dutch *er* and French *en* for instance, will be parallel spans, but the Dutch one includes a [+count] specification. An additional issue that I mentioned previously is that *zonder* ('without') does not trigger r-replacement and r-movement. I believe this is because its semantics are derived by reversing the semantics of *met* ('with'). This 'anti-instrumental' or 'anti-comitative' is probably built as a parallel span to the regular instrumental/comitative: in the lower part of its structure it has a specification for absence (akin to the semantics of negative quantifiers). Then, the structure [root(-human)...Neg...Ins] will be lexicalized by *het* and the dedicated preposition. It has also been noted that other, more complex, prepositions share this behavior. Future exploration into this domain could indicate a different, but more inclusive, solution.

Additionally, *er*-conflation would work under the set-superset nano-architecture. Recall that multiple functions of *er* can appear as a single morphological instance. Recall, also, the view of lexical items in the current model: functions/features are not simply derived by the context of insertion and associated loosely with the morpheme - they are organized asymmetrically in the very representation of the item. Thus, all readings are accessible at once at insertion. Therefore, an instance of *er* can be read in several ways at once, since its representation in the Lexicon contains all of the relevant features.

To summarize: I propose a modified case sequence in which *er* corresponds to a tree including the Locative and Partitive cases maximally and overrides *het* (Nominative/Accusative) in these cases. Starting with the Genitive position, that sits above Partitive in my modified scale, prepositions appear (such as *van*). *er* is forced to evacuate as an effect of its own representation and that of prepositions (a lexicalization phenomenon). This does not happen for *hem* under prepositions, since there is something like a gap for a relevant sequence of cases: it is difficult to conceptualize an animate used as a Locative or Partitive. As a general point, the reader should note that Caha (2009) does not deal with pronouns in different cases but rather with explicit case marking on nouns. Also, we cannot be sure that the different functions of *er* are different cases, but the larger picture seems to be on the right track. Thus, irrespective of how well my facts map onto the model of representation used in Caha

---

[17]I take countability to be a property of entities or sets of entities, with substances being of a different type. When a pronoun such as *it* refers to a mass noun, the substance will have been recategorized as an entity.

(2009) and related work, I believe there is an isomorphism between these domains and that exploring it should point the way to a solution.

7 Conclusions

There is in a sense a feedback-loop between the organization of the Lexicon, items with multiple distributions and the issue of parametrizing linguistic variation. Locatives provide a good case study for this cluster of phenomena, while putting some specific issues on the table. At the descriptive level I provided a survey of the distributions of locatives. I have shown that the range of distributions, within and across languages, behaves like a scale, suggesting some implicational relation between the several uses. This suggests a certain model of representation in the Lexicon for poly-functional words / morphemes.

Looking at expletives, I raised the question of why such a function should be available for locatives in some but not all Germanic languages. I suggested that a language-level parameter is not needed to explain the presence of locatives in this position, but rather a model of competition between such elements and other pronominals. Thus, there is no need to set apart languages which have locative expletives in terms of their syntactic configuration. The requirement for expletives is a constant, but what elements can satisfy this requirement depends from one language to the other. Under this approach there is no need to say that locative expletives and adverbials represent different lexical entries. We can then parametrize the use of locative expletives without appealing to locativity as a formal feature, hence we get 'clean' variation (i.e. without substantive language-specific statements). This is an illustration of the **weak hypothesis** I establish in Section 2.4.

The extension of my study to Dutch r-words (*er* in particular) presents the following challenges: how can we account for the number of uses such locatives have, the positions they occupy relative to other pronominals and the way they behave across dialects? The dialectal range of uses of *er* vs *het* becomes an issue for parametrization. I proposed to deal with all of these issues by adopting a recent theoretical framework: Nanosyntax. This makes available the perspective that there is sub-morphemic organization. These structures are asymmetric in nature, hence variation in the available uses for one morpheme naturally receives a scalar aspect. A locative such as *er* can in fact correspond to a rich structure that derives its uses in a principled way. This is a possible implementation for what I advance as my **strong hypothesis** in Section 2.4: the different functions a (locative) word can have are related - under this implementation they are structurally related.

An interesting and possibly far-reaching idea to consider is why it should be that lexical entries have the richly structured representation I proposed in connection to my strong hypothesis. To establish that indeed natural language relies on such a design feature would add an entirely new dimension to the concrete pursuit of linguistic exploration, while giving research into language evolution and its neurological basis a more solid foothold than ever.

While this thesis draws to an end, linguistic research is only now gathering real momentum, with more novel perspectives and more solid principles being established everyday. I can only hope that the perspective I tentatively adopt here has the capacity to fully develop as research continues.

References

[1] Barbiers, S. 2009. "Locus and Limits of Syntactic Microvariation". In *Lingua*: 119 (2009), 1607-1623.

[2] Barbiers, S. 2009b. *Quantitative 'er', doubling and the structure of noun phrases*. Workshop on Agreement, doubling and DP. Leiden, September 4, 2009.

[3] Bennis, H. 1986. *Gaps and dummies*. Foris, Dordrecht.

[4] Bobaljik, Jonathan. 2007. *On comparative suppletion*. Ms., University of Connecticut.

[5] Caha, P. 2009. *The Nanosyntax of Case*. PhD diss., University of Tromso.

[6] Caha, P. 2010. "The parameters of case marking and spell out driven movement". In: Van Craenenbroeck, J. (ed.), *Linguistic Variation Yearbook 2010*. 32-77.

[7] Chomsky, N. 1982. *Some concepts and consequences of the theory of government and binding*. MIT Press, Cambridge, Ma.

[8] Chomsky, N. 1995. *The Minimalist Program*. MIT Press, Cambridge, Ma.

[9] Falk, Y. 2006. *Subjects and universal grammar : an explanatory theory*. Cambridge: CUP.

[10] Hale, K. & Keyser, J. 1993."On Argument Structure and the Lexical Expression of Syntactic Relations". In: Hale, K. and Keyser, J. (eds.): *The View From Building 20*. Cambridge: The MIT Press. 53-109.

[11] Hartmann, J. 2008. *Expletives in existentials*. LOT dissertation.

[12] Hoekstra, J. 2001. "Comparative Aspects of Frisian Morphology and Syntax". In: Munske, Horst H. (ed.) *Handbook of Frisian studies*. TÃijbingen: Max Niemeyer Verlag, 775-786.

[13] Mikkelsen, L. 2002 ."Expletive subjects in subject relative clauses". In: Zwart, J-W. and Abraham, W. (eds.) *Studies in Comparative Germanic Syntax: Proceedings from the 15th Workshop on Comparative Germanic Syntax (Groningen, May 26âĂŞ27, 2000)* . 69-91.

[14] Milsark, G. 1974. *Existential sentences in English*. New York / London: Garland.

[15] Moro, A. 1997. *The raising of predicates. Predicative noun phrases and the theory of clause structure*. Cambridge Studies in Linguistics, CUP, Cambridge, England.

[16] Pantcheva, M. 2011. *Decomposing Path*. PhD diss., University of Tromso.

[17] Parrott, J. 2009. "A Minimalist, Distributed Morphology approach to intra-individual variation: expletives and agreement in an insular English variety". In Grohmann, K. et al. (eds.), Interface Issues in Linguistic Minimalism. Cambridge: Cambridge Scholars Press.

[18] Pustejovsky, J. 1995. *The Generative Lexicon*. MIT Press, Cambridge, MA.

[19] Reinhart, T. and Siloni, T. 2005. "The Lexicon-Syntax Parameter: Reflexivization and Other Arity Operations". In *Linguistic Inquiry* 36: 3, pp. 389-436.

[20] Reuland, E. 1988. "Indefinite Subjects". Proceedings of NELS 18, 375-94. GLSA, Amherst, MA.

[21] Reuland, E. & Ter Meulen, A. 1987. *The Representation of (In)Definiteness*. MIT Press.

[22] Ritter, E. & Wiltschko, M. 2009. "Varieties of INFL: TENSE, LOCATION, and PERSON". In: H. Broekhuis., J. Craenenbroeck, H. van Riemsdijk (eds.). *Alternatives to Cartography*. Berlin, New York: Mouton de Gruyter.

[23] Setliff, R. 1988. *Sociolinguistic Variation in Smith Island English: Existential It*. PhD Ms, Georgetown University.

[24] Starke, M. 2009. "Nanosyntax: A Short Primer to a New Approach to Language". In ... *Nordlyd Special Issue on Nanosyntax*, Vol 36, No 1 (2009).

[25] Starke, M. 2011. *Towards elegant parameters: Language variation reduces to the size of lexically stored trees* . Available at: lingbuzz.auf.net/001183.

[26] Svenonius, P. 2002. "Introduction". In: Svenonius, P. (ed.) *Subjects, Expletives, and the EPP*. Oxford: OUP. 1-25.

[27] Rizzi, L. 2006. On the form of chains: Criterial positions and ECP effects. In Wh-movement: Moving on, edited by L. Cheng and N. Corver, 97-134. Cambridge: MIT Press.

[28] Rizzi, L. 2007. On some properties of Criterial Freezing. In CISCL Working Papers on Language and Cognition,edited by V. Moscati, 145-158. Interdepartmental Centre for Cognitive Studies of Language, University of Sienna.

[29] Taraldsen, T. 2002. "The Que/Qui Alternation and the Distribution of Expletives.". In: Svenonius, P. (ed.) *Subjects, Expletives, and the EPP*. Oxford: OUP. 29-42.

[30] Thráinsson, H. 2007. *The syntax of Icelandic.* Cambridge: Cambridge University Press.

[31] Van Craenenbroeck, J. 2011. *Germanic expletives revisited.* Comparative Germanic Syntax Workshop 26. Amsterdam. June 23-24, 2011.

[32] Van Riemsdijk, H. 1978. *A Case Study in Syntactic Markedness: The Binding Nature of Prepositional Phrases.* Foris, Dordrecht.

[33] Vangsnes, O. 2002. "Icelandic Expletive Constructions and the Distribution of Subject Types". In: Svenonius, P. (ed.) *Subjects, Expletives, and the EPP.* Oxford: OUP. 43-70.

[34] Vikner, S. 1991. "Relative 'der' and other C Elements in Danish". Lingua 84.2/3, 109-136.

[35] Vikner, S. 1995. *Verb movement and expletive subjects in the Germanic languages.* Oxford:. Oxford University Press.

[36] Williams, E. 1994. *Thematic structure in syntax.* MIT Press.