

Stability and variation in article choice: generic and episodic contexts

July 2007

Donka F. Farkas and Henriëtte de Swart

Abstract

In Farkas and de Swart (2006), we proposed an analysis of the cross-linguistic variation in article choice in plural generics. In this paper, we develop an extended version of the analysis which accounts for cross-linguistic stability in article choice in episodic contexts and in singular generics in languages with definite and indefinite articles, such as English, Dutch, French, Spanish, Greek and Hungarian.

1. Introduction

The languages we are concerned with in this paper have a definite/indefinite contrast as well as a singular/plural distinction. For a wider typology that includes languages that lack plural morphology on nouns or lack articles we refer the reader to Hendriks et al. (2007, ch. 7).

Farkas and de Swart (2006) start by noting that the languages under consideration here pattern alike concerning definiteness and number in episodic (non-generic) environments. Definite forms are uniformly used with unique (or maximal) referents, such as (1) and familiar referents, whether singular (2) or plural (3). We exemplify with English (E), French (F) and Hungarian (H):

- | | | | |
|--------|---|---|---|
| (1) a. | The moon is round. | The stars are shining. | E |
| b. | La lune est ronde. | Les étoiles brillent. | F |
| | The moon is round. | The stars shine. | |
| c. | A hold kerek. | A csillagok csillognak. | H |
| | the moon round | the star.Pl shine.Pl | |
| | | | |
| (2) a. | A man and a child came in. The man was tall. | | E |
| b. | Un homme et un enfant entraient. L'homme était grand. | | F |
| | A man and a child came-in. The man was tall. | | |
| c. | Bejött egy férfi és egy gyerek. | A férfi magas volt. | H |
| | in.com.Past a man and a child | the man tall was | |
| | | | |
| (3) a. | Children and dogs were playing in the street. The children were noisy. | | E |
| b. | Des enfants et des chiens jouaient dans la rue. Les enfants faisaient du bruit. | | F |
| | Indef-pl children and indef-pl dogs played in the street. The children made indef-mass noise. | | |
| c. | Gyerekek és kutyák játszottak az utcán. | A gyerekek hangosak voltak. | H |
| | child.Pl and dog.Pl play.Past.Pl | the street.on the child.Pl noisy.Pl were.Pl | |

In Hungarian and English, bare plural forms are used with existential indefinite interpretations as in (3a,c) while in French the indefinite plural article *des* (3b) is used. Bare plurals and *des* NPs share many syntactic and semantic properties, as shown by Bosveld-de Smet (1998). For the purposes of this paper, they will be treated uniformly as indefinite plural forms. The

definiteness contrast these languages exhibit involves definite nominals (marked by a definite article) and indefinite ones (occurring with an indefinite article in the singular and with or without an article in the plural). We thus conclude that the languages under consideration do not differ significantly with respect to definiteness in episodic contexts.

The term ‘generic environment’ covers both sentences expressing direct kind reference and those involving generic generalizations. Both singular and plural forms are used in the languages under consideration in these environments. With singular NPs, the definite article is consistently used for direct kind reference, as in (4), and the indefinite article is used in generic generalizations, as in (5):

- | | |
|---|---|
| (4) a. The dinosaur is extinct. | E |
| b. Le dinosaure a disparu.
The dinosaur has disappeared. | F |
| c. A dinosaurusz kihalt.
The dinosaur die.out.Past | H |
| (5) a. A dog is dangerous when it is hungry. | E |
| b. Un chien est dangereux quand il a faim.
A dog is dangerous when it has hunger | F |
| c. Egy kutya veszélyes mikor éhes.
A dog dangerous when hungry | H |

Differences arise in the plural case, where our languages fall into two groups: (i) English-type languages, which use *indefinite (bare) plural* forms in generic environments, and (ii) Hungarian, Greek and Romance languages, which use *definite plurals*. We add a few more languages to show the consistency in the pattern for direct kind reference (6) and generic generalizations (7):

- | | |
|---|----------|
| (6) a. Dinosaurs are extinct. | E |
| b. Dinosaurussen zijn uitgestorven.
Dinosaurians are extinct. | Dutch |
| c. Les dinosaures ont disparu.
The dinosaurs have disappeared. | F |
| d. A dinosauruszok kihaltak.
the dinosaur.Pl die.out.Past.Pl | H |
| e. Dinosaurii au dispărut.
Dinosaur-def has disappeared. | Romanian |
| f. Gli elefanti di colore bianco sono estinti.
The elephants of colour white are extinct. | Italian |
| g. Oi asproi elephantes echoun exaphanistei.
The white elephants are extinct. | Greek |
| (7) a. Dogs are dangerous when they are hungry. | E |
| b. Honden zijn gevaarlijk wanneer ze honger hebben.
Dogs are dangerous when they hunger have. | Dutch |
| c. Les chiens sont dangereux quand ils ont faim.
The dogs are dangerous when they have hunger. | F |
| d. A kutyák veszélyesek mikor éhesek.
the dog.Pl dangerous.Pl when hungry.Pl | H |
| e. Cîinii sînt intelegenti. | Romanian |

- Dog-def are intelligent.
- f. Gli ucelli di zone paludose sono intelligenti. Italian
The birds from marshlands are intelligent.
- g. Ta skillia einai eksipna. Greek
The dogs are intelligent.

The contrast between definite and indefinite generics is sometimes related to the contrast between Romance and Germanic languages, but the fact that typologically unrelated languages like Hungarian and Greek pattern with Romance indicates that we are dealing with a more widespread phenomenon. Since the cross-linguistic variation is restricted to plural generics, the correct account must locate the source of variation in the combination of plurality and genericity. The overall analysis of genericity, of the definite/indefinite contrast, and of the singular/plural distinction should be the same from one language to the next, so as to explain the cross-linguistic stability in article choice in episodic contexts and singular generics.

Farkas and de Swart (2006) focused on cross-linguistic variation in definiteness found in generic plural forms. Our aim here is to explain the cross-linguistic stability we find in the case of singular generics while maintaining the earlier account of the contrasts in the plural. To complete the account of cross-linguistic stability in article choice we extend the analysis of article choice in episodic contexts sketched in the earlier work.

The structure of the paper is as follows. In section 2, we summarize the analysis in Farkas and de Swart (2006) of the cross-linguistic contrast in article choice in plural generics. In section 3, we discuss episodic contexts, and account for cross-linguistic stability in article choice in these environments. In section 4, we pick up the line of genericity, focusing on singular, rather than plural generics, and show why the factors that give rise to cross linguistic variation in the plural case are irrelevant in the singular. Section 5 is a brief conclusion.

2. Article choice in plural generics

As mentioned above, we take the cross-linguistic variation in article choice in plural generics to arise from the combination of plurality and genericity. In section 2.1 through 2.3, we outline our background assumptions on number, the definite/indefinite contrast and genericity that allow us to explain what is so special about plural generics. Section 2.4 relates the semantics to the OT syntax, and develops an analysis of article choice in plural generics.

2.1 Background on number

As far as the semantics of number is concerned, we adopt the proposals made by Farkas and de Swart (2003) and Farkas (2006), who defend the view that singular nominals are semantically and morphologically unmarked for number, while plural nominals are morphologically marked by the feature [PI]. The effect of the presence of the feature [PI] is to introduce a presupposed discourse referent that gets the predicate Pl(ural) predicated of it. The canonical semantics of [PI] is to impose a group-level value on the discourse referent it is predicated of. DPs that lack this feature get an atomic interpretation in languages with morphological plural marking (treated in 2.4 as high ranked FPl (Faith Plurality)).

2.2 Background on the definite/indefinite contrast

For present purposes, we take an NP to be definite when it is preceded by the definite article and indefinite if preceded by an indefinite article or if it is a bare plural. We follow Farkas (2002) and references therein in taking the definite to be the marked member of the pair, and the indefinite the unmarked member. Just as in Farkas (2002), we characterize the semantic contribution of the definite article to be the marking of determined reference on the discourse referent it introduces. In the semantic representation we mark the contribution of the definite article by !. Determined reference (or dynamic uniqueness) concerns the question of choice of value for a referent at the point of update. A discourse referent x introduced by a DP has determined reference relative to an input function f and a model M iff for every f', f'' that are output functions satisfying the descriptive content of the DP, $f'(x) = f''(x)$. The special marking for determined reference is motivated by topicality or discourse prominence (de Hoop, this volume). Determined reference can be achieved in two ways: (i) uniqueness/maximality relative to M ; (ii) anaphoric dependency on a discourse referent in the input DRS. We can see determined reference at work in the examples in (1) and (2) from section 1. In example (1), world knowledge tells us there is a unique moon and a finite set of stars in our solar system, which are referred to by the definite *the moon* and *the stars*. In examples (2) and (3), the indefinites in the first sentence set up discourse referents that are picked up by means of the definite expressions *the man* and *the children* in the second sentence. The discourse referents are familiar, because they have already been introduced at the time the definites are evaluated. In this latter case the discourse referent has determined reference in virtue of its being familiar.

We distinguish these two ways of achieving determined reference by means of the features $[\pm Max]$ and $[\pm Fam]$ on the discourse referent. A discourse referent marked as $[+Max]$ must have unique reference relative to the DRS if atomic and maximal reference if plural. A discourse referent characterized as $[+Fam]$ must be present in the input DRS or accommodatable therein. For convenience, we say an NP is maximal or familiar when the discourse referent the NP introduces is maximal or familiar respectively.

2.3 Background on genericity

As far as genericity is concerned, we distinguish between direct reference to kinds (examples in 4 and 6) and generic generalizations (examples in 5 and 7), along the lines of Carlson (1978), Krifka et al. (1995) and others. With respect to singular generics, we follow Carlson's (1977) original insight in treating the definite singular in (4) as referring to a special atom k in the model, one that is kind-level and unique. With respect to kind referring plural generics, we follow Ojeda (1993), Chierchia (1998) and Dayal (2004) in assuming that plural kinds involve the sum of their instantiations across different worlds and situations. So the definite singular in (4) gets the representation in (8a). Bare plurals and definite plurals in sentences expressing direct kind reference get the representations in (8b) and (c) respectively:

- (8) a. $\exists!k$ (Dinosaur(k) & Extinct(k)) definite sg kind reference
 b. $\exists K K = \sum_{x,w}$ (Dinosaur(x,w) & Plural(K) & Extinct(K)) bare pl kind reference
 c. $\exists!K K = \sum_{x,w}$ (Dinosaur(x,w) & Plural(K) & Extinct(K)) definite pl kind reference

\sum operates on individual level entities in different worlds and sums them up into their kind. The entity that results after application of \sum is the top node of the intensional lattice associated with the descriptive content of the NP. We denote it by K to distinguish it from

ordinary atomic kinds k (which can form their own hierarchical structures in taxonomic interpretations). Kind-level predicates like *to be extinct* are insensitive to the difference between atomic kinds (k) and derived kinds (K), as illustrated by the felicity of both (4) and (6). The atomic kind in (8a) is unique. In the singular, the fact that the discourse referent has determined reference is reflected in the use of the definite article, which induces the exclamation mark ! that follows the existential quantifier. The semantics of the sum operator confers maximality on the plural kinds in (8b) and (8c) as well. Romance and Hungarian reflect the determined reference of the discourse referent in the use of the definite article, which induces the exclamation mark ! in (8c). English and Dutch do not reflect the determined reference of the discourse referent, and use a bare plural, which does not induce an exclamation mark in (8b).

We follow de Swart (1991, 1996) in treating the indefinite singular in generic generalizations like (5) and (7) as an ordinary indefinite, introducing a discourse referent x that is indirectly bound by a silent generic operator Gen binding dog-sized situations. The semantics of generic generalizations involving singulars (as in 5) is given in (9a). The representation of bare plurals and definite plurals in generic generalizations is given in (9b) and (c) respectively:

- (9) a. $\text{Gen}_s ((\exists x \text{ Dog}(x,s) \ \& \ \text{Hungry}(x,s)), \text{Dangerous}(x,s))$ indef sg in GG
 b. $\text{Gen}_s ((\exists x \text{ Dog}(x,s) \ \& \ \text{Plural}(x) \ \& \ \text{Hungry}(x,s)), \text{Dangerous}(x,s))$ bare pl in GG
 c. $\text{Gen}_s ((\exists !x \text{ Dog}(x,s) \ \& \ \text{Plural}(x) \ \& \ \text{Hungry}(x,s)), \text{Dangerous}(x,s))$ definite pl in GG

Indirect binding arises when each situation s involves its own instance of the discourse referent x . In such a context, we have as many (typical) situations with a (typical) dog in them as we have (typical) dogs. As a result, indirect binding is indistinguishable from unselective binding (Farkas and Sugioka 1983, Krifka et al. 1995, and others), as far as truth conditions are concerned. In the plural case (9b, c), a hidden distribution operator guarantees that predication over the plural entity percolates down to the level of the atomic members of the plural individual. In order to keep the semantic representations as simple and transparent as possible, the particularities of this mechanism are not spelled out. The only difference between (9b) and (9c) is the visibility of determined reference in the semantics of the definite (Romance, Hungarian) generic plural in (9c), versus the invisibility of determined reference in the bare (English, Dutch) generic plural in (9b).

2.4 OT analysis of article use in plural generics

In Farkas and de Swart (2006), we argue that the general definite/indefinite contrast cannot be made responsible for article choice in plural generics, because English/Dutch and Romance/Hungarian behave very similarly as far as the use of definite and indefinite articles in episodic contexts is concerned. We can embed this in the typological theory of definiteness and number developed by de Swart and Zwarts (2006). In this analysis, the interaction of three faithfulness constraints and two markedness constraints determines which distinctions get expressed in a language. The constraints used are the following:

- ◆ FPl: parse plurality by means of a NumP.
- ◆ FDef: parse determined reference.
- ◆ Fdr: parse a discourse referent by means of a functional layer above NP.
- ◆ *Art: avoid article
- ◆ *FunctN: avoid functional structure in the nominal domain

The constraint interaction is used as a generation model in OT syntax. If all the faithfulness constraints are ranked below the markedness constraints *Art and *FunctN, no number and definiteness distinctions are expressed. If one or more faithfulness constraints are ranked above the markedness constraints, the OT syntax produces nominals with functional layers (NumP and/or DP). In this paper, we are interested in languages that have a full definiteness/indefiniteness contrast, as well as a singular plural distinction. In such languages, all the relevant faithfulness constraints are ranked above the two markedness constraints, so we have the ranking {FDef, Fdr, FPl} >> {*Art, *FunctN}.

According to Farkas and de Swart (2006), the constraints in de Swart and Zwarts (2006) are not sufficient to determine article use in plural generics. The similarities between Romance, Hungarian, Greek on the one hand and English, Dutch on the other hand in episodic contexts (examples 1-3 in section 1) and in singular generics (examples 4 and 5) can be accounted for under the assumption that the grammar of all these languages involves the ranking {FDef, Fdr, FPl} >> {*Art, *FunctN}. But how would this grammar account for the divergence in article use in generic plurals (examples 6 and 7)? We could attempt to relate the use of the definite article in Romance/Hungarian/Greek bare plurals to the constraint FDef. FDef is a faithfulness constraint that requires the output syntactic form to reflect the determined reference nature of the discourse referent. We observe that the discourse referent in plural generics has determined reference because of the analysis of genericity we assumed in section 2.3 above. Both the sum operator (Σ) and the Gen operator confer maximality upon the plural discourse referent, so the discourse referent has determined reference. Because of the high ranking of FDef in Romance/Hungarian/Greek as well as English/Dutch, we should see a reflection of this in the syntactic form. However, if we take determined reference to trigger the use of the definite article in plural generics in Romance/Hungarian/Greek (6c-g, 7c-g), we are left with no explanation for the bare generic plurals in English/Dutch (6a-b, 7a-b). If we assume that FDef is somehow independently satisfied, and has nothing to do with article choice in plural generics, we might appeal to *Art to explain why generic plurals remain bare in English/Dutch (6a-b, 7a-b). However, this means that the use of the definite article in Romance/Hungarian/Greek (6c-g, 7c-g) is not accounted for. We conclude that we need the constraint setting advanced in de Swart and Zwarts (2006) to account for the similarities in article selection in episodic contexts and singular generics, but we need an additional mechanism to account for the divergences of article selection in plural generics.

The key insight formulated by Farkas and de Swart (2006) is that FDef is independently satisfied in plural generic contexts, and the cross-linguistic variation arises by exploitation of the maximality and familiarity components of the definite article. We take it that the genericity of sentences involving Σ or Gen is marked in the syntax, and parses determined reference. Given that either Σ or Gen is always present in plural generic contexts, according to the analysis of genericity outlined in section 2.3, FDef is automatically satisfied in examples like (6) and (7) in all the languages under consideration. But if FDef is satisfied by the presence of Σ or Gen, it does not govern article choice in plural generics, and other constraints must come into play to account for the cross-linguistic diversity in these contexts. Farkas and de Swart (2006) observe that in plural generics, the two components of determined reference, namely maximality and familiarity are in conflict: the discourse referent of the plural generic nominal in sentences like (6) and (7) always has the feature configuration [+Max] [-Fam]. It is [+Max], because Σ or Gen confers maximality upon the discourse referent. It is [-Fam] in generic generalizations, because variables (in)directly bound by an operator are never familiar (Kamp and Reyle 1993). It is [-Fam] in direct kind reference, because of the intensional nature of Σ . The fact that the discourse referent corresponding to the plural generic nominal is always maximal, and never familiar, suggests that the two

parameters of determined reference are independent in this context. We capture the two aspects of determined reference in the two constraints (10) and (11):

(10) *MaxMax*: Maximize maximality features of the discourse referent by reflecting them in the nominal projection.

(11) **Def/[-Fam]*: Avoid non-familiar definites.

MaxMax is a faithfulness constraint that requires a maximal discourse referent to be introduced by an NP with a definite article, even if determined reference is already reflected in Gen or Σ . It is more specific than *FDef*: satisfaction of *MaxMax* entails satisfaction of *FDef*, but not the other way around. **Def/[-Fam]* is a markedness constraint that penalizes the use of a definite article with non-familiar discourse referents. It is also more specific than *FDef*, because it only comes into play in languages that have a definite/indefinite contrast. These two constraints govern the choice of a definite or an indefinite form based on the semantic properties of the discourse referent. The constraints in (10) and (11) are assumed to be universal. The contrast of interest here is due to different orderings of the two constraints:

(12) *FDef* >> **Def/[-Fam]* >> *MaxMax* >> **Art* [English, Dutch]
FDef >> *MaxMax* >> **Def/[-Fam]* >> **Art* [Romance, Greek, Hungarian]

In plural generic contexts, *FDef* is satisfied by the sum operator or the generic operator, and therefore *FDef* is always satisfied in all the languages under consideration. We know that the plural discourse referent in direct kind reference as well as generic generalizations has the feature configuration [+Max][-Fam]. *MaxMax* favors the use of a definite article, whereas **Def/[-Fam]* penalizes it. Thus the two constraints conflict in environments involving plural generics. The ranking *MaxMax* >> **Def/[-Fam]* guarantees that languages like Romance, Greek and Hungarian use definite plural generics. The ranking **Def/[-Fam]* >> *MaxMax* implies that English and Dutch use non-definite (bare) plural generics. For generic generalizations, this is reflected in the following tableaux:

Tableau 1: indefinite plural generic generalizations (illustrated with English)

Meaning $\text{Gen}_x(\text{Dog}(x), \text{Pl}(x)), \text{Intel}(x)$ [+Max] [-Fam]	Form	FDef	*Def/-fam	MaxMax	*Art
	Dogs are intelligent			*	
	The dogs are intelligent		*		*

Tableau 2: definite plural generic generalizations (illustrated with French)

Meaning $\text{Gen}_x(\text{Dog}(x), \text{Pl}(x)), \text{Intel}(x)$ [+Max] [-Fam]	Form	FDef	MaxMax	*Def/-fam	*Art
	Des chiens sont intelligents		*		*
	Les chiens sont intelligents			*	*

We call languages like English ‘high Familiarity’ languages, because they block the definite article in plural generics. Romance, Hungarian and Greek are ‘high Maximality’ languages, because they reflect the maximal nature of the discourse referent in the use of the definite article, even though determined reference is already parsed by the generic construction. Our approach thus accounts for the intuition that the definite article of plural generics is semantically redundant. The OT analysis makes strong predictions as to article use in sentences expressing pseudo genericity and anaphoric genericity. The interested reader is referred to Farkas and de Swart (2006) for details. The account of plural generics raises two important questions that were not addressed in Farkas and de Swart (2006):

- Is article use in episodic contexts affected by the introduction of the new constraints MaxMax and *Def/[-fam]?
- Is article use in singular generics affected by the introduction of the new constraints MaxMax and *Def/[-fam]?

Given the similarities in article use in episodic contexts between English/Dutch on the one hand, and Romance/Hungarian on the other, the answer to both questions should be ‘no.’ Sections 3 and 4 show how we reach this answer and therefore how we account for cross-linguistic stability in article use in episodic and singular generic generalizations across the languages we study.

3. Article choice in episodic contexts

In Farkas and de Swart (2006), we suggested that the ranking {FPl, Fdr, FDef} >> *Art accounts for the similarities in article use in episodic sentences in English/Dutch and Romance/Hungarian. (13) through (15) repeat the relevant examples from section 1 above:

- (13) a. The moon is round. The stars are shining. E
 b. La lune est ronde. Les étoiles brillent. F
 The moon is round. The stars shine.
 c. A hold kerek. A csillagok csillognak. H
 the moon round the star.Pl shine.Pl
- (14) a. A man and a child came in. The man was tall. E
 b. Un homme et un enfant entraient. L’homme était grand. F
 A man and a child came-in. The man was tall.
 c. Bejött egy férfi és egy gyerek. A férfi magas volt. H
 in.com.Past a man and a child the man tall was
- (15) a. Children and dogs were playing in the street. The children were noisy. E
 b. Des enfants et des chiens jouaient dans la rue. Les enfants faisaient du bruit. F
 Indef-pl children and indef-pl dogs played in the street. The children made
 indef-mass noise.
 c. Gyerekek és kutyák játszottak az utcán. A gyerekek hangosak voltak. H
 child.Pl and dog.Pl play.Past.Pl the street.on the child.Pl noisy.Pl were.Pl

In section 2.2 above, we argued that all the definite NPs in these examples have determined reference. Some achieve determined reference through uniqueness/maximality in the larger context, as in (13), while others achieve it through anaphoric reference to a discourse referent

introduced earlier in the discourse, as in (14), (15). This is the interpretive perspective. Under the generative perspective, we see that the constraint ranking $FDef \gg *Art$ requires parsing determined reference. If the construction is not marked as generic, the only way for a N headed nominal to satisfy the constraint $FDef$ is by the use of a definite article. In episodic contexts there is no conflict between the maximality and familiarity features of the relevant discourse referent. That is, the definite singulars and plurals in (13) - (15) are maximal (or unique) as well as familiar. The use of a definite form satisfies $FDef$ as well as $*Def/[-Fam]$ and $MaxMax$. We therefore predict the use of a definite form across the languages under consideration in these cases, given that the only relevant difference between the grammar of English/Dutch on the one hand and Romance/Hungarian on the other is the ranking of the constraints $*Def/[-Fam]$ and $MaxMax$.¹ This is illustrated for English in tableau 3:

Tableau 3: familiar episodic definites in ‘high Familiarity’ languages (illustrated with English for the second sentence of 14a)

Meaning man(x) [+Max] [+Fam]	Form	FDef	*Def/-fam	MaxMax	*Art
	a man	*		*	*
☞	the man				*

The result of the competition between definite and indefinite forms in episodic contexts is identical in Romance/Hungarian, because of the high ranking of $FDef$ in both language groups. Now, corpus research has shown that definites in English texts frequently introduce discourse new entities (Poesio and Vieira 1998), as in:

- (16) a. Fleet Street has been buzzing with the rumor that the prime minister is going to resign.
b. The first person to sail to America was an Icelander.

We find the same pattern in Romance/Hungarian:

- (17) a. A hír, hogy Mari beteg ijesztő. H
the rumor that Mari is sick is frightening
'The rumor that Mari is sick is frightening.'
b. A legjobb diákom Magyarországról jött.
the best student.my Hungary.from came
'My best student came from Hungary.'

Birner and Ward (1995) discuss the relevance of discourse new definites in existential contexts, cf. also McNally (this volume). The definites in (16)/(17) have determined reference because their discourse referent is unique relative to the model, so these non-familiar definites exemplify the feature combination [+Max][-Fam]. For discourse referents with this feature combination, the definite comes out as the optimal form in the grammar, as illustrated in tableaux 4 and 5, for English and Hungarian:

¹ Thanks to Larry Horn for pointing out problems with an earlier version of our analysis (see de Swart and Farkas 2005, and discussion in Horn 2005).

Tableau 4: non-familiar episodic definites in ‘high Familiarity’ languages (illustrated with English 16b)

Meaning first person to sail to America(x) [+Max] [-Fam]	Form	FDef	*Def/-fam	MaxMax	*Art
	a first person to sail to America	*		*	*
☞	the first person to sail to America		*		*

Tableau 5: non-familiar maximal episodic definites in ‘high Maximality’ languages (illustrated with Hungarian 17b)

Meaning best student(x) [+Max] [-Fam]	Form	FDef	MaxMax	*Def/-fam	*Art
	egy legjobb diákom	*	*		*
☞	a legjobb diákom			*	*

As tableaux 4 and 5 illustrate, English and Hungarian differ in their relative ranking of the constraints *Def/[-Fam] and MaxMax, but are alike in ranking FDef above both. The feature [+Max] in the input requires determined reference to be parsed in order to satisfy the high ranked constraint FDef. In episodic contexts, a definite form is the winner in both languages because unlike in generic environments, there is no alternative way of parsing determined reference.² In sum, the role of FDef in the grammar is crucial in accounting for stability of article choice in episodic contexts across the languages under consideration, and the cross-linguistic variation found in plural generics.

As the unmarked member of the opposition, indefinite NPs do not impose determined reference, and just introduce a discourse referent. Even though the indefinite has such a weak, unmarked meaning, it can express a contrast with a definite NP. Suppose the speaker wants to introduce just a discourse referent, and wishes to remain uncommitted with respect to whether the referent has determined reference. In this case, the constraint FDef is vacuously satisfied in the OT syntax, but *Def/-fam blocks the use of the definite form. This is illustrated in tableau 6.

Tableau 6: non-familiar, non maximal episodic indefinites (illustrated for English with the first sentence of 14a)

² For the sake of completeness, we should mention the case of [+Fam][-Max] discourse referents. Imagine a situation in which we have set up three discourse referents a, b and c, and we know that a, b, c are children. If we now want to refer to a, neither the indefinite *a child*, nor the definite *the child* would be appropriate, rather the partitive *one of the children* would be used. This suggests that a full grammar of English involves more candidates than just the definite and indefinite articles, and more constraints than what we have here. An account of the three-way contrast between definites, indefinites and partitives is beyond the scope of this paper.

Meaning child(x) [-Max] [-Fam]	Form	FDef	*Def/-fam	MaxMax	*Art
☞	a child				*
	the child		*		*

There is no distinction between high familiarity and high maximality languages, for the difference in ranking between *Def/-fam and MaxMax is irrelevant in this case. Tableau 6 shows that the definite form is marked, and cannot be used to express an unmarked meaning.

As a result of their unmarked status, indefinite NPs are not incompatible with familiarity or maximality of the discourse referent *per se*. As far as the interpretation of an indefinite form is concerned, the discourse referent introduced by the indefinite is taken to be non-familiar and non-maximal by conversational implicature. For Horn (2005), Farkas (2006) and de Swart (2006) the definite and indefinite articles form the scale in <the, a>. Within the approach to scalar implicatures developed by Horn (1984) and Levinson (2001), lack of maximality and familiarity specifications characterizing indefinites is strengthened to non-maximality and non-familiarity as the result of an I-implicature (Horn 1984, Levinson 2001). There are, however, contexts where these implicatures do not arise and where indefinites are compatible with unique or maximal interpretations. Farkas (2006) discusses examples with ‘singleton indefinites’ like (18a), taken from Schwarzschild (2002) to support her claim that non-uniqueness can be cancelled. Similarly, Bosveld-de Smet (1998) provides examples of French indefinite plurals that do not necessarily block maximality (18b):

- (18)a. We will watch *a movie that Phil said he liked the best*.
 b. Des étudiants ont manifesté.
 Indef_pl students have demonstrated.

Example (18a) invites an interpretation where there is only one movie Phil said he liked the best. In (18b), we focus on the collective participation in a single event, and we do not get any indication as to the question of whether all the students were involved, or only a subset of them was. The fact that ‘the’ and ‘les’ are not used in (18a) and (18b) respectively indicates lack of determined reference marking and therefore speaker indifference to it. Since the indefinite is the unmarked member of the definite/indefinite pair <the, a> the OT syntax governing article choice does not involve a constraint selecting for indefiniteness, but only constraints governing the use of definite forms.

4. Article choice in singular generics

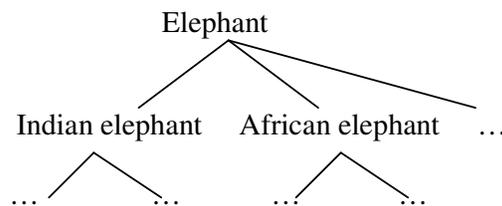
In this section we return to generic environments, concentrating this time on singular generics. Recall from section 2 that cross-linguistic variation is found with plural generics: the English/Dutch group contrasts with Romance/Hungarian/Greek in that the former uses indefinite forms in generic plurals while the latter languages use a definite form. In this section we address the question of why this contrast is neutralized in singular generics, where both groups of languages use a definite form when direct reference to a well established kind

is involved (section 4.1), and an indefinite or a definite form in generic generalizations (section 4.2).

4.1 Direct reference to kinds

Following Carlson's 1977 seminal work, we have assumed in section 2.3 above that the ontology comprises, besides ordinary individuals, special kind-level individuals as well. Singular generics denote atomic kinds, special in that their realizations are individuals, whereas the realizations of ordinary individuals are stages, that is, spatio-temporal 'slices' of individuals. Atomic kinds are special also in that many kinds are structured into subkinds that form taxonomical structures as schematized in Figure 1:

Figure 1: The elephant and its subkinds



Every node in this tree is an atomic kind, with the highest one being the most inclusive.

Since we have kind-level atoms in the ontology we expect NPs to be able to refer to them. And indeed, we have such direct kind-level reference to well-established kinds in (19):

- | | |
|--|---|
| (19) a. <i>The elephant</i> is widespread. | E |
| b. <i>De olifant</i> is wijdverbreid. | D |
| c. <i>L'éléphant</i> est commun. | F |
| d. <i>Elefantul</i> e răspîndit. | R |
| e. <i>Az elefant</i> gyakori. | H |

The representation of the meaning of the sentences in (19) is given in (20):³

$$(20) \exists!k(\text{Indian elephant}(k) \wedge \text{widespread}(k))$$

According to the semantics in (20), the definite (italicized) NPs in (19) denote one particular atomic kind level entity, namely the highest node in the tree in Figure 1.⁴ We see in (19) that

³ We adhere strictly to Carlson's original view here, and do not follow the neo-Carlsonian treatment of singular generics developed by Chierchia (1998) and Dayal (2004). Their proposals are driven by the parametrized reference to individuals in Chierchia (1998). Given that we use the typology of Hendriks et al. (2007: chapter 7), we can maintain Carlson's original insights without losing the contrast between reference to atomic kinds by definite singulars (in English, Romance, etc.) and bare singulars (in Hindi, Chinese, etc.). Note that not all languages allow kind-level predicates in combination with singulars. For instance, our Greek informant resisted the use of the definite singular in contexts like (21). We don't have anything to say about the requirement for a plural form in such cases.

⁴ One can, of course, refer to a subspecies of elephant by a singular definite, such as *the Indian elephant*. In that case too the NP refers to the unique atomic kind identified by its descriptive content.

the languages in both our groups uniformly use singular definite NPs to refer to such atomic kinds. This is as expected, for the discourse referent introduced by the NP is a well-established kind, which qualifies as a unique, familiar entity. The relevant discourse referent is [+Max] and [+Fam]. The mechanism imposing the choice of a definite form over an indefinite one in (19) is thus the same as the one used in episodic contexts (cf. section 3 above). Given the high ranking of FDef in all the languages under consideration here, the system we set up so far predicts uniform definite singular forms.

Article use with generics referring to atomic kinds thus parallels article selection in episodic contexts, and contrasts with what we find in generic plurals referring to plural kinds. Plural kinds involve the sum operator Σ , and we have assumed that all sentences involving this operator satisfy FDef. With singulars, no special generic operator comes into play in the expression of direct kind reference. As a result, the only way to parse determined reference and satisfy FDef is by the use of the definite article.

Given the possibility of taxonomic structures such as the one above, we predict the possibility of indefinite singular subkind (or taxonomic) reference. This happens when the value of the kind-level discourse referent is not narrowed down (yet) to a particular node in the relevant taxonomic tree. In such cases, the NP does not have determined reference since its denotation is one of the nodes of the taxonomical structure identified by its descriptive content, and therefore is not fixed. In the absence of determined reference, FDef is automatically satisfied. Our account predicts the use of an indefinite singular form in such cases in all the languages under consideration, by the same mechanism that leads to an indefinite singular form in case of ordinary indefinite individual reference. This is indeed what we find, as exemplified in (21), whose representation is given in (22):

- | | | |
|---------|---|---|
| (21) a. | <i>A (kind of) elephant</i> is widespread. | E |
| b. | <i>Een (soort van) olifant</i> is wijdverbreid. | D |
| c. | <i>Un(e) (espèce d')éléphant</i> est commun. | F |
| d. | <i>Un (soi de) elefant</i> e rãspîndit. | R |
| e. | <i>Egy (fajta) elefant</i> gyakori. | H |

$$(22) \exists k(\text{elephant}(k) \wedge \text{widespread}(k))$$

The analysis we have developed predicts, correctly, that a kind-denoting indefinite NP as in (21) can only get a taxonomical reading. This is so because the use of an indefinite form implicates that the definite could not have been chosen, and therefore that the conditions for definiteness are not met. This means that the relevant discourse referent must be assumed not to have determined reference and therefore not to be unique. Direct kind reference involves uniqueness, so a taxonomic reading is the only way non-unique kind reference can be achieved.

To sum up, the use of definite/indefinite forms in sentences involving direct singular kind reference is regulated, in our view, by exactly the same mechanism that decides between these forms in ordinary episodic contexts. We therefore correctly predict that our two groups of languages will not contrast in the choices they make. A definite singular form will be interpreted as involving unique atomic kind reference, denoting the node in the taxonomic kind identified by the descriptive content of the NP. An indefinite singular form will be interpreted taxonomically. In that case the kind-level discourse referent does not have determined reference: any subnode in the taxonomic tree identified by the descriptive content is a possible value for the relevant discourse referent. Thus, *the Indian elephant*, we predict, when used generically, can only refer to the species **Indian elephant**. Kind reference with *an Indian elephant* on the other hand can only involve reference to some subkind of Indian

An interpretation in terms of atomic kind reference is the only way a definite singular in the languages under consideration can acquire generic reference. The predicate here, however, is individual-level rather than kind-level and therefore we are dealing with a mismatch between the type of the variable (kind-level) and the type the predicate requires as its argument (individual-level). Such mismatches can be fixed by allowing either for the possibility of lowering the nominal (via a realization relation) or for the possibility of raising the predicate. We adopt the first option here although nothing in what follows rests on this choice:

- (27) a. *The cocker-spaniel* is intelligent.
 b. $\exists!k(\text{cocker-spaniel}(k) \wedge \text{Gen}_s [\exists x(\text{R}(x,k),s)] [\text{intelligent}(x,s)])$

The definite nominal introduces a kind-level variable whose individual realizations are brought into play by a Carlsonian realization relation *R*, and indirectly bound by the generic operator. Using the *R* relation here amounts to having a type-shift that takes kinds to their individual realizations. We differ from Chierchia and Dayal in that its use is not free. The type shift is triggered, strictly, by the type mismatch between the predicate and its argument. The semantic representation in (27b) triggers the use of the definite article in the (c) examples of (23)-(26). According to the OT constraint system, the selection of a definite article is in order because we have here reference to a unique and familiar entity, namely the atomic, well-established kind *cocker-spaniel*. The combination of the features [+Max] [+Fam] requires marking by a definite article in languages with a high ranking of *FDef*. Given that *FDef* is ranked high in all the languages under consideration, we predict no cross-linguistic variation in this case, and indeed we find none.

Indefinite singulars are the outcome of two possible expressive optimization processes. Under one reading, the (b) sentences of (23)-(26) have a taxonomic interpretation, where the indefinite refers to a subkind of dog. This interpretation is spelled out in (28):

- (28) a. *A cocker-spaniel* is intelligent. (taxonomic reading)
 b. $\exists k(\text{cocker-spaniel}(k) \wedge \text{Gen}_s [\exists x(\text{R}(x,k),s)] [\text{intelligent}(x,s)])$

The indefinite nominal here introduces a kind-level variable, just like in (27). The main difference is that the discourse referent is not unique and familiar, because now it refers to one of the subbranches in the subkind structure of cocker-spaniels, but we do not know which one. The feature combination [-Fam][-Max] triggers the use of an indefinite article, just like in the case of (21) in section 4.1 above.

The other possible interpretation of the (b) sentences of (23)-(26) involves the generic operator *Gen*. These sentences express a generalization over individual dogs, and don't involve reference to kinds. We refer to such interpretations as strict generic generalizations in contrast to the definite generic in (27):

- (29) a. *A cocker-spaniel* is intelligent. (strict generic generalization)
 b. $\text{Gen}_s[\exists x(\text{cocker-spaniel}(x,s))][\text{intelligent}(x,s)]$

The indefinite singular in (31a) introduces a discourse referent *x* that is novel because it is indirectly bound by the generic operator *Gen*. *Gen* ranges over a set of dog-size situations (situations with a dog in them). Due to a general constraint against vacuous quantification (de Swart 1993, 1996, section 2.3 above), this set of situations cannot be a singleton set. If individuating the relevant situations amounts to individuating dogs, then the non-singleton requirement on the set of situations amounts to a non-singleton requirement on the set of dogs involved. Given that the singular definite imposes determined reference upon a discourse

referent with an atomic value, a definite singular cannot be used in a strict generic generalization. Therefore, a semantic representation of the (c) sentences of (23)-(26) in terms of a Gen operator indirectly binding the discourse referent introduced by the definite singular is not available. However, generic generalizations over individual-level predicates can be conveyed with a definite singular generic, because there is an interpretation in terms of kind reference that can be assigned to the (c) sentences of (23)-(26), along the lines of (27).

4.2.3 Stage-level predicates

We now come to the problem of article selection in generic generalizations that involve stage-level predicates. The singular generic in the (a) sentences of (23)-(26), just like its plural counterpart, introduces a discourse referent indirectly bound by Gen. Gen here quantifies over typical dog-sized situations that have a hungry dog in them. In such situations, the sentence claims, the dog-stage is generally dangerous. The representation is given in (30), where x^s stands for a stage-level instantiation of the dog variable x .

- (30) a. A *dog* is dangerous when it is hungry. (strict generic generalization)
 b. $\text{Gen}_s [(\exists x \text{ dog}(x,s)) [\text{Hab} (\exists e \text{ hungry}(x^s,e)) (\text{dangerous}(x^s,e))]]$

Note that this example involves not only the generic operator Gen, but an additional habitual operator Hab. Hab ranges over stage-level predicates and creates an individual-level property, along the lines of Carlson (1977). Languages in both groups choose *indefinite* singular forms here, even though they contrast when it comes to the choice of the article in the corresponding plural generic case. The explanation follows along the same lines as that of the (b) examples of (23)-(26) above. The indefinite article is called for in strict generic generalizations because the definite singular creates a singleton set, which is incompatible with the plurality requirement on the Gen operator. Cross-linguistic stability is predicted here because the languages under consideration all have the same semantics of the definite article.

A question that remains to be answered is why languages in both groups choose *indefinite* singular forms here, i.e. why don't we have (31) as the counterpart to (23a), just like we have (23c) as the counterpart to (23b):

- (31) *The dog* is dangerous when it is hungry. (*Gen/ $\sqrt{\text{specific dog}}$)

(31) is a well-formed sentence, but it refers to a specific dog that has the habit of getting dangerous when it is hungry. It cannot be used to predicate this habit of the atomic kind *dog*. We conjecture that this has to do with the fact that the predicate involved is stage-level. Remember that the semantic representation of (23c/27a) in (27b) involves the lowering of an atomic kind variable to its individual realizations. Somehow, this process of lowering is blocked by the presence of a stage-level predicate. Note also that (30a) only has the semantic representation in (30b). The sentence does not have a taxonomic reading similar to (28a), along the lines of (28b). We conclude that reference to atomic kinds is generally unavailable with a stage-level predicate.⁶

⁶ It is possible that this process is not blocked for all languages. A Bulgarian example doing just what (31) cannot do is given in (i):

- (i) Kucheto e opasno kogato e gladno.
 dog.the is dangerous when it is hungry.

According to what we have said so far, we express generic generalizations over individual level predicates with definite as well as indefinite forms (23b, c). There is a difference between strict and non-strict generic generalizations, as we can see in (27) and (29), but the semantic representations are truth-conditionally equivalent. What frequently goes unnoticed is that the discourse effects of definite and indefinite singular generics are distinct. The strict generic generalization in (29b) does not involve direct reference to a kind, while the more complex interpretation in (27b) does involve such reference. So whether the speaker chooses the definite or the indefinite form depends on whether she intends to convey direct reference to a well established kind as part of the message or not. If yes, a definite singular will be used, because the referent has the features [+Max] [+Fam]. Otherwise the form of the singular NP will be indefinite. Some relevant examples that illustrate the distinction between well-established (i.e. familiar) kinds are in (32) (inspired by Krifka et al. 1995):

- (32) a. A green bottle is attractive to customers.
 b. The green bottle is attractive to customers. (*Gen/√specific bottle)
 c. The Coca Cola bottle is attractive to customers.

Given that the use of a definite article is driven by determined reference, we expect to find definite generics with kinds that are well-established on the basis of world knowledge (as in 32c). However, another way to establish determined reference is via a discourse anaphoric relation, as we see in the examples (2) and (3) in section 1. Example (33) shows that discourse anaphoric reference can trigger the use of a definite generic singular:

- (33) “**We need a new bottle** – a distinctive package that will help us fight substitutions...we need a bottle which a person will recognize as *a Coca-Cola bottle* even when he feels it in the dark. *The Coca-Cola bottle* should be so shaped that, even if broken, a person could tell what it was...” wrote the company’s legal counsel in 1915, urging management to develop packaging that could be protected by trademark and patent laws.
 [http://www.cdf.org/cdf/atissue/vol2_2/cocacola/cocacola.html]

A new subkind of *bottle* is introduced in the first sentence of (35), and defined as *a Coca-Cola bottle*. This new subkind is then picked up with the definite generic *the Coca-Cola bottle* in the subsequent discourse. The definite singular has determined reference, because discourse familiarity a single value for the atomic discourse referent is available in the context. The role of discourse anaphoricity in (32) is reminiscent of the discourse contrast between definite and bare plural generics we find in languages like English. Farkas and de Swart (2006) discuss examples such as the Dutch (34):

- (34) Dus de dieren die wij *dinosaurussen* noemen zijn echt uitgestorven. (...)
 Het is niet precies bekend hoe *de dinosaurussen* zijn uitgestorven.
 So the animals we call dinosaurs are really extinct. (...)
 It is not exactly known how the dinosaurs died out.
 [http://www.dinosaurus.net/AlgemeneInfo/Uitsterven/uitsterven.htm]

More systematic cross-linguistic research will have to determine the variation ranging from Greek (cf. footnote 5 above) via English/Romance/Hungarian to Bulgarian.

4.4 Discourse effects

The similarities between (33) and (34) make it clear that article choice is sensitive to discourse anaphoricity, not only in episodic contexts, as is well-known, but also in generic contexts. Singular and plural generics behave alike in their preference for a definite article in contexts where kind reference has been established in the preceding discourse.

5. Conclusion

In this paper we provided an account of cross-linguistic stability and variation in article choice across two groups of languages, English/Dutch, on the one hand, and Romance/Hungarian/Greek on the other. We started from reviewing our earlier account of the contrast between these two groups with respect to the form chosen in plural generics. The analysis relies crucially on the system of constraints independently proposed in Hendriks et al. (2007, chapter 7), supplemented by two constraints that take into account the two components of determined reference, maximality and familiarity. Plural generic NPs introduce maximal and non-familiar variables. The contrast between the two language groups reduces to the fact that in English/Dutch the non-familiar feature blocks the use of a definite form while in Romance/Hungarian/Greek it does not.

The rest of the paper is devoted to an account of cross-linguistic stability in article choice across the languages under consideration when it comes to episodic contexts as well as generic singulars. The heart of the matter in episodic contexts concerns the fact that in such contexts the article is the only vehicle that can encode determined reference while in generic environments the nature of the interpretation of the variable is encoded in the generic operator. In the case of singular generics, the use of the indefinite form in all the languages under consideration is connected to the clash between uniqueness (associated with definite singular forms) and the property of being bound by a generic operator. In the account that emerges, article choice is a matter that crucially binds together the number of the NP and the semantic environment in which it occurs.

There are many questions that arise at this point. The most pressing one is extending the typology to a larger group of languages, and a larger group of nominal types (including, e.g. partitive constructions). The tools and analyses in the present paper, we hope, will provide a useful stepping stone in understanding these larger issues.

References

- Abbott, B. (1999). Support for a unique theory of definite descriptions. *Proceedings of SALT IX*, 1-15.
- Abbott, B. (2003). A reply to Szabó's "Descriptions and Uniqueness." *Philosophical Studies* 113: 223-31.
- Beaver, D. and H. Lee (2003). Input-output mismatches in OT, in: R. Blutner and H. Zeevat (eds.) *Optimality theory and pragmatics*, Palgrave/MacMillan, p. 112-153.
- Behrens, L. (2005). Genericity from a cross-linguistic perspective, *Linguistics* 43, 275-344.
- Blutner, R. (2001). Some aspects of optimality in natural language interpretation, *Journal of Semantics* 17, 189-216.
- Bosveld-de Smet, L. (1998). *On mass and plural quantification: the case of French 'des'/'du' NPs*. PhD Dissertation, University of Groningen.
- Carlson, G. (1977). *Reference to kinds in English*, PhD University of Massachusetts at Amherst; published 1980 New York: Garland.
- Carlson, G. and J. Pelletier (1995) (eds.). *The generic book*, University of Chicago Press.
- Chierchia, G. (1998). Reference to kinds across languages, *Natural Language Semantics* 6, 339-405.

- Cohen, A. (2001). On the generic use of indefinite singulars, *Journal of Semantics* 18, 183-209.
- Dayal, V. (2004). Number marking and (in)definiteness in kind terms, *Linguistics and Philosophy* 27, 393-450.
- Farkas, D. and Sugioka (1983). Restrictive if/when clauses, *Linguistics and Philosophy* 6 :225-258.
- Farkas, D. (2002). Specificity distinctions, *Journal of Semantics* 19: 1-31.
- Farkas, D. (2006) The unmarked determiner, in L. Tasmowski and S. Voegeler (eds.) *Indefiniteness and Plurality*, Amsterdam: Benjamins, 81-105.
- Farkas, D. and H. de Swart (2003). *The semantics of incorporation : from argument structure to discourse transparency*. Stanford: CSLI Publications.
- Farkas, D. and H. de Swart (2006). Article choice in plural generics, *Lingua* (in press).
- Ferreira, M. (2005). Bare habituals and plural definite descriptions, *Proceedings of Sinn und Bedeutung* 9, Nijmegen, 102-115. See: <http://www.ru.nl/ncs/sub9/>
- Hawkins, J. (1991). On (in)definite articles: implicatures and (un)grammaticality predictions, *Journal of Linguistics* 27, 405-442.
- Hendriks, P. and H. de Hoop (2001). Optimality theoretic semantics, *Linguistics and Philosophy* 24, 1-32.
- Hendriks, P., H. de Hoop, I. Krämer, H. de Swart and J. Zwarts (2007). *Conflicts in Interpretation*, ms. Groningen/Nijmegen/Utrecht.
- Hoop, H. de (2007). On (in)animate noun phrases (this volume).
- Horn, L. (1984). Toward a new taxonomy for pragmatic inference: Q- and R-based implicature. In: D. Shiffrin (ed.). *Meaning, form and use in context*, Washington: Georgetown University Press, p. 11-42.
- Horn, L. (2005). <the, a>: (In)definiteness, implicature and scalar pragmatics, ms. Yale University.
- Kamp, H. et U. Reyle (1993). *From discourse to logic*. Dordrecht : Kluwer Academic Press.
- Krifka, M. et al. (1995) Genericity: an introduction, in: G. Carlson and F. Pelletier (eds.) *The Generic Book*, Chicago: University of Chicago Press, 1-124.
- Levinson, S. (2001). *Presumptive Meanings*, Cambridge: MIT Press.
- McNally, L. (2007). Properties, entity correlates of properties and existentials (this volume).
- Ojeda, A. (1993). *Linguistic individuals*, Stanford : CSLI Publications.
- Poesio, M. and R. Vierira (1998). A corpus-based investigation of definite description use. *Computational Linguistics* 24 : 183-216.
- Schwarzschild, R. (2002). Singleton indefinites, *Journal of Semantics* 19, 289 – 314.
- de Swart, H. (1996). (In)definites and genericity, in : M. Kanazawa et al. (eds.). *Quantifiers, deduction and context*, CSLI Publications, p. 171-194.
- de Swart, H. (2006). Aspectual implications of the semantics of indefinite plurals, in L. Tasmowski and S. Voegeler (eds.) *Indefiniteness and Plurality*, Amsterdam: Benjamins, 161-189.
- de Swart, H. and D. Farkas (2005). Généricité et indéfinitude. Une analyse dans la théorie de l'optimalité, in: C. Dobrovie-Sorin (ed.). *Noms nus et généricité*, Paris: Presses Universitaires de Vincennes, 98-126.
- de Swart, H. and J. Zwarts (2006). Less form, more meaning: why bare nominals are special *Lingua* (in press).
- Warner, G. and B. Birner (1995). Definiteness and the English existential, *Language* 71, 722-742.