

# More on the semantics of clitic doubling: principal filters, minimal witnesses, and other bits of truth

Mojmír Dočekal<sup>†</sup> and Dalina Kallulli<sup>‡\*</sup>

<sup>†</sup>Masaryk University in Brno, <sup>‡</sup>University of Vienna

## 1. Introduction

This paper deals with a phenomenon that in the generative paradigm has since Jaeggli (1982) come to be known as clitic doubling, and which is illustrated through the Albanian examples in (1). As its name suggests, clitic doubling involves the ‘doubling’ by a clitic pronoun of a DP that is a verbal argument inside one and the same propositional structure, which we take to be a clausal unit (Adger 2003).<sup>1</sup> The clitic and its associate (i.e. the DP it doubles) share the same case and phi-features (i.e. person, number and gender).<sup>2</sup> As the examples in (1) demonstrate, the associate can be instantiated by a full pronoun or a non-pronominal referring expression that can be a definite, indefinite, or proper noun.

- (1) a. Ana            **më**    pa    **mua**   në rrugë.  
Anna.the<sub>NOM</sub> me<sub>CL</sub> saw me<sub>FP</sub> in road  
‘Anna saw me in the street.’
- b. Ana            **e**                    lexoi **letrën**        deri në fund.  
Anna.the<sub>NOM</sub> CL.ACC.3S read letter.the<sub>ACC</sub> till in end  
‘Anna read the letter to the end.’
- c. Ana            **e**                    pa **Benin**        në rrugë.  
Ana.the<sub>NOM</sub> CL.ACC.3S saw Ben.the<sub>ACC</sub> in road  
‘Ana saw Ben on the road.’
- d. Do **ta**                            pija me kënaqësi **një uiski**.  
FUT SUBJ.CL.ACC.3S drink with pleasure a whisky  
‘I would gladly drink a whisky.’

Though clitic doubling constructions in many languages have since Kayne (1975) received a great deal of attention in the syntactic literature (for a recent review of the syntactic literature on clitic doubling, see Kallulli & Tasmowski 2008), modulo the detailed investigation in Gutiérrez-Rexach (1999) on clitic doubling in (varieties of) Spanish, research on the formal

---

\*For useful comments we wish to thank an EISS anonymous reviewer and the editor Chris Piñón.

<sup>1</sup>We assume that the associate of the doubling clitic is a DP (Abney 1987), not an NP, which is why bare nouns cannot be clitic doubled (see Kallulli 2000, and Kallulli & Tasmowski 2008).

<sup>2</sup>However, just like full pronouns, clitics are frequently underspecified for case and/or gender, as is the case with *më* in (1a), which could be either accusative or dative and with *e* in (1b), which is underspecified for gender. For details on the clitic paradigms in Albanian, see Kallulli (1995).

semantics of these constructions has been much less prolific. This is in part due to the perplexing fact that while dative clitic doubling (i.e. clitic doubling of dative objects) behaves analogous to object agreement marking (e.g. see Sportiche 1996 and references therein), clitic doubling of direct object DPs seems to be subject to various idiosyncratic language-specific semantic constraints, such as animacy and specificity in Romanian and Spanish (Farkas 1978, Suñer 1988, Dobrovie-Sorin 1990), definiteness in Modern Greek (Anagnostopoulou 1994a,b, Anagnostopoulou & Giannakidou 1995), topichood and/or givenness in Albanian (Kallulli 2000, 2008), which make it very hard if not altogether impossible to come up with a unitary semantic analysis.<sup>3</sup>

In this paper, we present novel data from Albanian showing that (i) the DP associated with the clitic (i.e. the ‘doubled’ DP) must be interpreted as generating admissible minimal witnesses, which in turn makes the DP topical; and that (ii) as a consequence of (i), clitic doubling systematically produces information structure effects in that the doubled DP is unequivocally interpreted as topical.

This paper is organized as follows. In §2, we present the core data, which among other things necessitate the search for an alternative analysis to the one proposed in Gutiérrez-Rexach 1999. In particular, we draw attention to the fact that, contrary to what Gutiérrez-Rexach claims, clitic doubling cannot be explained by appealing to the notion of principal filterhood. §3 then details our own analysis, the crux of which is that the clitic doubled DP must be interpreted as generating an admissible minimal witness, from which all other effects such as those involving information structure are derived.

## 2. Determiner types and clitic doubling: the view from Albanian

### 2.1. Strong and weak determiners

On the basis of their formal properties, Barwise and Cooper (1981:182) distinguish between two classes of quantifiers, namely, ‘strong’ and ‘weak’ ones. Strong quantifiers are those headed by the ‘strong’ determiners in (2a), and weak quantifiers are those headed by the ‘weak’ determiners in (2b).

- (2) a. strong determiners: the, both, all, every, each, most, neither  
 b. weak determiners: a, some, one, two, three, ..., many, a few, few, no, Ø

In Albanian, DPs headed by strong determiners can be clitic doubled; we already saw an instance of this in (1b). Two further examples of doubling with strong quantifiers are given in (3).<sup>4</sup>

- (3) a. (I) lexova të dy librat.  
 them<sub>Cl.ACC</sub> read.1s agr both books.the<sub>ACC</sub>  
 ‘I read both books.’  
 b. (I) lexova të gjithë librat.  
 them<sub>Cl.ACC</sub> read.1s agr all books.the<sub>ACC</sub>  
 ‘I read all (the) books.’

<sup>3</sup>However, some of these claims are controversial. For instance, the Greek example in (18b) in §2.3 shows that indefinites can clearly be clitic doubled in this language, a fact that has been acknowledged as a counterexample by Anagnostopoulou (1994b:4) herself.

<sup>4</sup>DPs headed by the strong determiners *every*, *each*, *most*, and *neither* may also be clitic doubled in Albanian. However, we postpone their discussion to §3.2.

DPs headed by weak determiners, except those headed by ‘no’ and bare nouns, may also be clitic doubled in Albanian, in which case the doubled DP invariably has wide scope.<sup>5</sup> An instance of clitic doubling with a weak quantifier was already provided in (1d); other examples are given in (4).

- (4) Javën e shkuar (i) lexova dy / tre / disa / ca libra.  
 week.the past them<sub>CL.ACC</sub> read.I two / three / some / a few books  
 ‘Last week I read two books/some books.’
- (5) Nuk (\*/?e) lexova asnjë libër.  
 not CL.ACC.3S read-I not.one book  
 ‘I didn’t read any book’

As we explicate below, the fact that DPs headed by weak determiners can be clitic doubled is unexpected under the only available semantic analysis of clitic doubling in the literature, namely, the one in Gutiérrez-Rexach 1999 on clitic doubling of direct objects in Spanish, which is outlined in the following subsection.<sup>6</sup>

## 2.2. Gutiérrez-Rexach 1999

Gutiérrez-Rexach (1999) claims that accusative DPs in Spanish are clitic doubled if and only if they denote principal filters. That is, direct object clitic doubling is according to him subject to the constraint in (6), with ‘principal filter’ defined as in (7):

- (6) *The Principal Filter Constraint:*  
 The generalized quantifier associated with an accusative clitic has to be a principal filter. (Gutiérrez-Rexach 1999:326)
- (7) A generalized quantifier  $Q$  over  $E$  is a *principal filter* iff there is not necessarily empty set  $A \subseteq E$ , such that for all  $B \subseteq E$ ,  $Q(B) = 1$  iff  $A \subseteq B$ . The set  $A$  is called the *generator* of  $Q$  ( $A = \text{GEN}(Q)$ ). (Gutiérrez-Rexach 1999:326)

The problem with this analysis is that since weak quantifiers can never denote principal filters, as is obvious from the original definition of principal filter in Barwise & Cooper 1981:183 given below in (8), clitic doubling them should not be possible, contrary to fact.

- (8) *Definition.* A determiner  $D$  is *definite* if for every model  $M = \langle E, \llbracket \cdot \rrbracket \rangle$  and every  $A$  which  $\llbracket D \rrbracket(A)$  is defined, there is a non-empty set  $B$ , so that  $\llbracket D \rrbracket(A)$  is the sieve  $\{X \subseteq E \mid B \subseteq X\}$ . (Hence  $\llbracket D \rrbracket(A)$  is what is usually called the principal filter generated by  $B$ .)

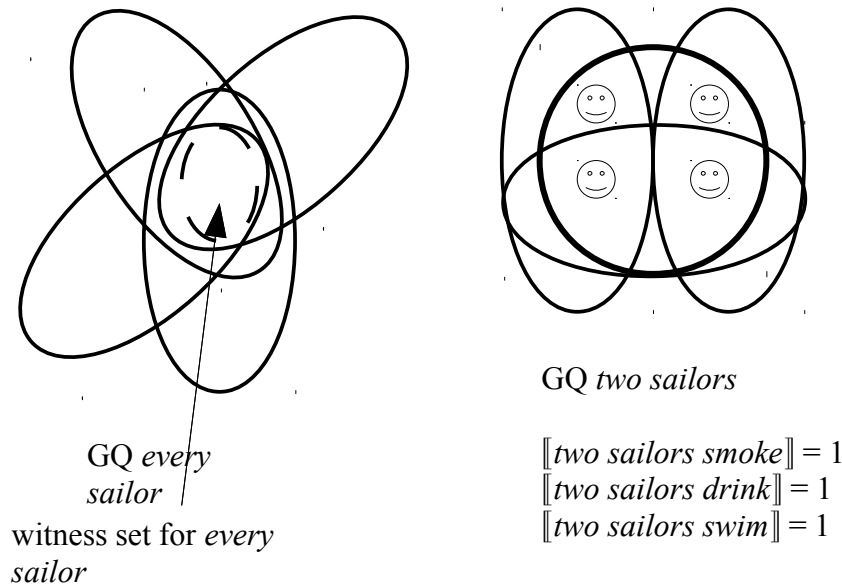
Specifically, the definition in (8) states that for a quantifier to be a principal filter, for *every* model there must be a non-empty set  $B$  (the set cannot be empty because it is the sieve) which belongs to the set of sets denoted by the quantifier  $X$ . To illustrate, consider the figures in (9), which show why *every* is a positive strong determiner (which generates the principal filter for the set in its restriction) but *two* is a weak determiner and can as such never denote a principal

<sup>5</sup>Following Kallulli 2000 and related literature, we assume that bare nouns cannot be doubled because while clitics are D-elements (alternatively: carry a D-feature), bare nouns lack a D-projection, which results in a feature mismatch causing the derivation to crash.

<sup>6</sup>As Gutiérrez-Rexach acknowledges, clitic doubling is possible with weak determiners in Spanish, too.

filter. A quantifier like *two sailors* denotes the family of sets which intersect with the set of sailors containing at least two sailors, but there is no non-empty set *B* which would be the subset of all such sets in every model. But exactly this situation is true for a quantifier like *every sailor*: there is a set *B* (the set of sailors) which is a subset of all sets in the family of sets denoted by the quantifier. This is why the principal filter hypothesis cannot help us in explaining the semantic conditions behind clitic doubling in Albanian: weak quantifiers in Albanian can be clitic doubled even though weak quantifiers can never generate principal filters.

(9)



Concluding this section, since weak quantifiers (which can never denote principal filters) can be clitic doubled in Albanian (as witnessed by our example (4)), an approach along the lines of Gutiérrez-Rexach 1999 is untenable.

### 2.3. Further specifications: determiner subtypes, clitic doubling, and information structure

In §2.1 we saw that both strong and weak quantifiers may be clitic doubled in Albanian. The data in (10) through (12), however, further complicate the picture; while none of the quantifiers in (10) through (12) are principal filters, some allow clitic doubling and some do not. More specifically, while weak monotone increasing quantifiers may be clitic doubled, as shown in (10), weak monotone decreasing quantifiers and non-monotone quantifiers cannot be clitic doubled, as illustrated in (11) and (12) respectively.

- (10) Javën e shkuar (?i) lexova tëpaktën dy libra / më shumë se dy libra.  
 week.the past them<sub>CL.ACC</sub> read.I at least two books/ more many than two books  
 ‘Last week I read at least two books/more than two books.’

- (11) Javën e shkuar (??/\*i) lexova të shumtën dy libra / më pak se dy libra.  
 week.the past them<sub>CL.ACC</sub> read.I at most two books/ more less than two books  
 ‘Last week I read at most two books/less than two books.’
- (12) Javën e shkuar (??/\*i) lexova tamam tre libra.  
 week.the past them<sub>CL.ACC</sub> read.I exactly three books  
 ‘Last week I read exactly three books.’

Furthermore, though quantifiers headed by strong determiners may be clitic doubled, there are contexts where they cannot be and contexts where they must be clitic doubled.<sup>7</sup> Consider the data in (13) through (16) (adapted from Kallulli 2000), noting in particular the complementarity of felicity conditions between the ‘minimal pairs’ in (13A)/(14B) versus (15B)/(16B), all of which mean ‘Anna read the book’:

- (13) A: What did Ana do? B: Ana (\*e) lexoi librin.  
 Anna CL.ACC.3S read book.the  
 ‘Anna read the book’
- (14) A: What did Ana read? B: Ana (\*e) lexoi librin.
- (15) A: Who read the book? B: Ana \*(e) lexoi librin.
- (16) A: What did Ana do with/to the book? B: Ana \*(e) lexoi librin.

What these data highlight is that when the VP or direct object DP is focus or part of the focus domain, clitic doubling is impossible but when the direct object DP is exempted from the focus domain (i.e. when the direct object DP is topical, or given) a doubling clitic is not only possible, but indeed obligatory. In other words, direct objects in Albanian need to be clitic doubled in order to be interpreted as topical (or given), a property which also accounts for two additional facts pointed out in Kallulli 2008, namely, that the object of the verb ‘to have’ may not be clitic doubled in Albanian existential constructions, as shown in (17), and that first and second person personal pronouns are invariably clitic doubled in this language.

- (17) (\*I) kishte minj në gjithë apartamentin.  
 them<sub>CL.ACC</sub> had mice in all apartment.the  
 ‘There were mice all over the apartment.’

Indeed we will argue that precisely this is what also accounts for doubling of indefinites (and other weak quantifiers), as in (18a) and (18b) for Albanian and Greek, respectively, as these are ‘non-novel’ indefinites in the sense of Krifka (2001).<sup>8</sup>

- (18) a. Do ta pija me kënaqësi një uiski.  
 FUT SUBJ.CL.ACC.3S drink with pleasure a whisky
- b. To pino exfaristos ena ouiskáki. (Kazasis and Pentheroudakis 1976:399)  
 it<sub>CL.ACC</sub> drink with pleasure a whisky  
 ‘I would gladly drink a whisky.’

Summarizing the discussion so far, we have shown that direct object clitic doubling in Albanian cannot be explained in terms of principal filterhood because weak quantifiers, which

<sup>7</sup>In fact as Kallulli (2000) shows, this also applies to doubling of indefinites (and/or other weak quantifiers).

<sup>8</sup>We discuss non-novel indefinites in some detail in §3.4.

are not principal filters, can be clitic doubled. Furthermore, weak quantifiers and monotone increasing quantifiers can be clitic doubled but monotone decreasing and non-monotone quantifiers cannot.<sup>9</sup> Finally, while both strong and weak quantifiers may be clitic doubled, there are contexts in which this is not possible, namely when they are part of the focus domain. What we are searching for, then, is some common property that would allow us to explain why strong determiners, weak determiners under wide scope interpretation and monotone increasing determiners can be clitic doubled and why monotone decreasing determiners and non-monotone determiners cannot be clitic doubled in Albanian, bearing in mind that clitic doubled expressions are invariably interpreted as topics in this language. Our hypothesis, then, naturally must mix purely semantic ingredients (monotone decreasing quantifiers can never be clitic doubled) with information structure approaches (even strong quantifiers cannot be clitic doubled when they are part of the focus domain/non-given).

### 3. Analysis

#### 3.1. The proposal

Since clitic doubled DPs in Albanian are invariably interpreted as topical, we believe the explanation for the data discussed in the previous sections should start with a discussion of topichood and the topical status of these DPs.

The distinction topic/non-topic, which we take to be more or less identical to the distinction topic/focus (cf. Hedberg 2006), closely corresponds to the subject/predicate distinction (see in particular Strawson 1974). In a crude approximation, then, a sentence such as (19a) says about John, its topic, that he has the property of loving Mary, whereas a sentence such as (19b) says about Mary, its topic, that she has the property of being loved by John. As is clear from the equivalence of (19a') and (19b'), both (19a) and (19b) have the same truth conditions, but they differ in how these truth conditions are communicated.

- (19) a. John loves Mary.  
 a'.  $\lambda x[\text{love}'(\text{John}, x)](\text{Mary})$   
 b. Mary, John loves.  
 b'.  $\lambda x[\text{love}'(x, \text{Mary})](\text{John})$

Crucially, topics refer to entities or to sets of entities (for plural topics) and focus refers to the properties of entities or of sets of entities denoted by topics. More formally stated, only entities or sets of entities, i.e. type  $\langle e \rangle$  or  $\langle e, t \rangle$  expressions, can be interpreted as topics. Focus, on the other hand, is always interpreted as a property, i.e. type  $\langle e, t \rangle$ , characterizing the topical entity, or also of type  $\langle \langle e, t \rangle, t \rangle$  – property of sets – for plural topics. Already from this assumption it follows that generalized quantifiers are not good candidates for topichood because their type  $\langle \langle e, t \rangle, t \rangle$  is not compatible with the entity (i.e. type  $\langle e \rangle$ ) status of topics (and in the case of the plural topics, which are of the type  $\langle e, t \rangle$ , functional application would reverse the subject/predicate asymmetry). It is then hardly surprising that some quantifiers cannot serve as sentence topics. We use the left dislocation construction (which is usually taken as a signal of topichood of the dislocated phrase) to show that unlike proper nouns, quantifiers like *every sailor* cannot be left-dislocated – cf. the contrast between (20a) and (20b). Weak

---

<sup>9</sup>By weak and strong quantifiers we also mean conjunctions and disjunctions of weak and strong quantifiers, because e.g. conjunction of strong quantifiers like *all the books and all the magazines* can be clitic doubled in Albanian. Also in this respect Albanian clitic doubling resembles partitivity contexts in English, where also conjunctions and disjunctions of strong quantifiers are allowed (see Keenan 1996). Thanks to the anonymous reviewer for pointing this out to us.

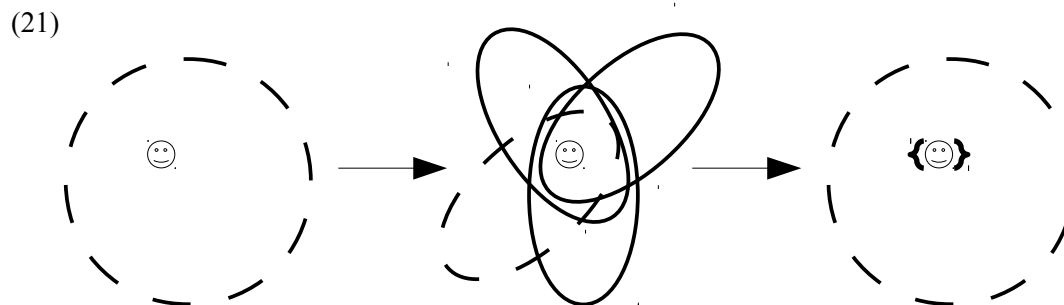
quantifiers like *three sailors* in (20c) on the other hand may be left-dislocated, which shows that at least some quantifiers allow for a topical interpretation.

- (20) a. John, we met him yesterday.  
 b. \*Every/\*no sailor, we met him yesterday.  
 c. Three sailors, we met them yesterday.

But how can a quantifier allow for a topic interpretation given the fact that the logical type of quantifiers is incompatible with the entity type of the topic? Following ideas in Endriss 2009 and Szabolcsi 2010, we assume that it is the witness sets of topical quantifiers that serve as their meaning.

The concept of witness sets goes back to Barwise & Cooper 1981. Intuitively, we think about sentences like *John is a sailor* in terms of set membership: the sentence is true if John is the member of the set of sailors (formally:  $j \in \{x \mid x \text{ is a sailor}\}$ ). But Barwise and Cooper (1981), following Montague (1973), argue exactly for the opposite perspective with respect to what is the function and what is its argument: a sentence like *John is a sailor* is true in their framework if and only if the property of being a sailor is one of the properties (family of sets) which John has (formally:  $\{x \mid x \text{ is a sailor}\} \in \{X \subseteq E \mid j \in X\}$ ).

The idea of witnesses (or witness sets) can be understood as restoring the former intuition: the witness set for *John* is the singleton set  $\{j\}$ , and the sentence *John is a sailor* is true if and only if the witness set for *John* is a subset of the set of sailors. Similarly for quantifiers: the witness set of the quantifier *every sailor* is the set of sailors, the witness set of the quantifier *two sailors* is the set of sets containing as members sets of two sailors and so on. Schematically, the meaning of the sentence ‘John is a sailor’, which the figure in (21) is supposed to depict, can be explained in the three stages given in (21).



The intuition behind topics is that they are referential, and witness sets can also be conceived of as the generalized quantifier’s referential contribution to the proposition (even if we accept the widely assumed view that quantifiers are non-referential expressions, as discussed in Heim & Kratzer 1998).

Let us compare the denotation of the generalized quantifiers with their witness sets. To visualize a generalized quantifier, we draw the denotations of the generalized quantifiers as in the table in (22). The order of rows starting from the bottom up represents the subset relation, hence  $\{a\}$  is below  $\{a, b\}$  and  $\{a, b\}$  below  $\{a, b, c\}$ , because  $\{a\} \subseteq \{a, b\} \subseteq \{a, b, c\}$ . The boldface represents the denotation of the quantifier: the denotation of the quantifier *at least one student* (if we assume that only atomic individuals  $\{a, b\}$  are students) is the family of sets containing at least one student:

$\{\{a\}, \{b\}, \{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \{a, b, c\}, \{a, b, d\}, \{a, c, d\}, \{b, c, d\}, \{a, b, c, d\}\}$

(22) monotone increasing NPs like *at least one student* (assume that  $\llbracket student \rrbracket$  is  $\{a, b\}$ )

		$\{a,b,c,d\}$			
	$\{a,b,c\}$	$\{a,b,d\}$	$\{a,c,d\}$	$\{b,c,d\}$	
$\{a,b\}$	$\{a,c\}$	$\{a,d\}$	$\{b,c\}$	$\{b,d\}$	$\{c,d\}$
	$\{a\}$	$\{b\}$	$\{c\}$	$\{d\}$	
		$\emptyset$			

As is clear from the table, an element of the GQ denoted by *at least one student* will contain other elements beside the denotation of the restrictor, i.e. for this quantifier the denotation can contain tigers, cars, cats and any other elements that include at least one student. Barwise and Cooper (1981) define *witness sets* as elements from which such alien bodies are removed. A principal filter is a superset of a witness set; it is the denotation of the quantifier if the quantifier lives on some set: the witness sets of the quantifier *at least one student* are the sets  $\{a\}$ ,  $\{b\}$ , and  $\{a, b\}$ .

Endriss and Hinterwimmer (2008) hypothesize that topical quantifiers are interpreted as *minimal witness sets* in such cases. While the witness sets for a monotone increasing quantifier such as *at least one student* in a model like (22) are  $\{a\}$ ,  $\{b\}$ , and  $\{a, b\}$ , the minimal witness sets are simply the two sets  $\{a\}$  and  $\{b\}$ .<sup>10</sup> For monotone decreasing quantifiers the minimal witness set is the empty set  $\emptyset$ . To illustrate, consider a variant of example (4), repeated below in (23). The minimal witness sets for the weak quantifier *disa libra* ‘some books’ are the singleton sets of all atomic entities which are books in the particular model. Assume these sets are  $\{a\}$ ,  $\{b\}$ , and  $\{c\}$ , which are interpreted as topical, with the rest of the sentence being interpreted as a property, as given in (24). Immediately, a type problem reveals itself: both the function and its argument are of the same type, namely,  $\langle e, t \rangle$ , so we cannot proceed with the beta reduction.

(23) Javën e shkuar (i) lexova disa libra.  
 week.the past them<sub>CL.ACC</sub> read.I some books  
 ‘Last week I read some books.’

(24)  $\lambda x[\text{read\_last\_week}(I, x)](\{a, b, c\})$

Endriss and Hinterwimmer (2008:90) solve this problem by hypothesizing that: ‘the elements of the minimal witness set corresponding to the topical quantifier are distributed over the elements of the set denoted by the comment’. Applied to our example in (23), for each of the individuals in the witness sets of books, the individual is fetched as an argument into the predicate consisting of the focus part of the sentence. Of course this yields an implausible reading for (23), namely, that I read all the books in the universe of discourse. Endriss and Hinterwimmer resolve this via an operation very similar to that of choice functions (see the

<sup>10</sup>We follow Barwise and Cooper (1981:103) in their definition of a witness set:

(i) A *witness set* for a quantifier  $D(A)$  living on  $A$  is any subset  $w$  of  $A$  such that  $w \in D(A)$ .

And we follow Szabolcsi (1997) in her definition of a minimal witness set  $M$ :

(ii) A *minimal witness set* is a set that is smallest among the witness sets of a generalized quantifier  $D(A)$ , i.e.,  $M$  is a minimal witness set of  $D(A)$  iff  $\neg \exists M' [M' \in D(A) \wedge M' \subset M]$ .



discussion in the next section). But what this properly describes is the expectation that monotone decreasing quantifiers cannot be topics because their minimal witness set is the empty set  $\emptyset$ , which as such cannot be fetched into the predicate at all.

The classification of topical quantifiers is further developed in Endriss (2009). Following Kadmon (1985), Endriss claims that all generalized quantifiers introduce (plural) discourse referents (i.e. minimal witness sets which are accessible in subsequent sentences). Even if all quantifiers create plural discourse referents, the latter are not of the same type for all quantifiers.<sup>11</sup> Quantifiers are distinguished according to their ability to create exhaustive and non-exhaustive plural discourse referents. Consider the meaning differences in the interpretation of the anaphor *they* in the following sentences:

- (25) a. Three fishermen sat by the river. *They* caught a lot of fish.  
 b. At least three fishermen sat by the river. *They* caught a lot of fish.

The first sentence in (25a) containing the quantifier *three fishermen* is compatible with a situation in which more than three fishermen sat by the river, but *they* in the second sentence in (25a) can only be understood as referring to the three fishermen in the preceding sentence. The sentence in (25b) is also compatible with a situation in which more than three fishermen sat by the river, but in this case, the pronoun *they* in the second sentence refers to the totality of fishermen that sat by the river. In other words, in a situation where six fishermen sat by the river and three of them caught a lot of fish, (25a) would be true but (25b) would be false. The anaphor *they* serves as a means of distinguishing between the discourse referent created by the quantifier *three fishermen* and the quantifier *at least three fishermen*. In (25a), the anaphor refers back to one of the witnesses created by the quantifier, whereas in (25b) the anaphor cannot refer to any witness of the quantifier *at least three fishermen*, but must instead refer to the maximal (exhaustive) intersection of the set of fishermen with the set of entities sitting by the river. This distinction between exhaustivity and non-exhaustivity is according to Endriss (2009) anchored in the lexical meaning of the determiners. The distinction between the non-exhaustive weak determiner *three* and the exhaustive monotone increasing determiner *at least three* lies not only in the relation = versus  $\geq$ , respectively, but crucially in the way the plural discourse referent  $X$  is created: for *three* it is any subset of the intersection of the noun  $P$  with the verb  $Q$  but for *at least three* it is the maximal intersection between  $P$  and  $Q$ , as shown in (26).

- (26) a.  $n = \lambda P \lambda Q. \exists X[|X| = n \wedge X \subseteq P \cap Q]$   
 b. at least  $n = \lambda P \lambda Q. \exists X[|X| \geq n \wedge X = P \cap Q]$

Endriss (2009) uses this variable ability of quantifiers to create discourse referents for a classification of quantifiers in terms of their ability to be interpreted as topics. Her main idea is that the aboutness function of a topic quantifier shouldn't change the semantics (i.e. the truth conditions) of the sentence. So she tests whether quantifiers interpreted as topics have the same meaning as their basic meaning. We refer the reader to the definition (6.1) on page 248 from Endriss (2009), where the exact mechanism is explained. For our purposes, it is sufficient to demonstrate the topicality condition on two types of quantifiers: monotone decreasing quantifiers (such as *at most three horses* in (27a)) and weak quantifiers (such as *three horses* in (27b)). What the topicality condition tests is whether the topic interpretation (via witness sets) equals the normal interpretation of the quantifier. The topic interpretation is on the left side of

<sup>11</sup>For a different hypothesis according to which only some quantifiers create discourse referents, see Kamp & Reyle 1993.

the equation of the formulas in (27), whereas the normal interpretation is on the right side. The minimal witness set for monotone decreasing quantifiers is the empty set and because the empty set is the subset of any set, the left part of (27a) is a tautology, hence it is not equal to the right side, which is simply the meaning of the quantifier *at most three horses* (the set of sets which include at most three horses). In (27b), on the other hand, the part on the left side is equal to the one on the right side because the witness set of the weak quantifier *three horses* (the set of sets containing three horses) is identical to the meaning of the quantifier.

- (27) a.  $\exists P[P = \emptyset \wedge P \subseteq Y] \neq \exists X[|X| \leq 3 \wedge X = \llbracket \text{horse} \rrbracket \cap Y]$   
 b.  $\exists P[P \subseteq \llbracket \text{horse} \rrbracket \wedge |P|=3 \wedge P \subseteq Y] = \exists X[|X| = 3 \wedge X \subseteq \llbracket \text{horse} \rrbracket \cap Y]$

The topicality condition draws a line between topicable and non-topicable quantifiers. With some simplification, weak quantifiers, indefinites, and the universal *all*-quantifier are topicable, whereas monotone decreasing quantifiers, non-monotone quantifiers, the universal quantifier *every*, and monotone increasing quantifiers are non-topicable. This is close (though not identical) to what we saw for the Albanian clitic doubling patterns. An exception are monotone increasing quantifiers which may be clitic doubled in Albanian, even though according to the topicality condition they are non-topicable. But as Endriss (2009) herself acknowledges, matters are not so simple and straightforward, since even a monotone increasing determiner such as the English *several* allows for a topical wide scope reading and non-exhaustive interpretation, as shown in (28) (from Endriss 2009, her example 6.44).

- (28) a. Several mathematicians were at the party yesterday. They danced all night.  
 b. The other mathematicians at the party only drank a lot.

The non-exhaustive interpretation of the quantifier *several mathematicians* shows that it can be interpreted as a vague bare numeral weak quantifier similar to *n*. Under such an interpretation, it can then meet the topicality condition.

To conclude this section, let us summarize our reasoning so far: if we put aside the information structure effects, the set of quantifiers which can be clitic doubled in Albanian consists of weak quantifiers (bare numeral and monotone increasing quantifiers) and strong quantifiers. Bare numerals and the *all* strong quantifier are uncontroversially argued to be good candidates to be topics by Endriss (2009). As for monotone increasing quantifiers, they allow for a topical interpretation under a non-exhaustive interpretation. Further scrutiny notwithstanding, we assume that the same process of reinterpretation is responsible for the acceptability of monotone increasing clitic doubled quantifiers in Albanian.

The next section is dedicated to those strong quantifiers which allow clitic doubling in Albanian but their status as topics (in a theory like Endriss') is at least controversial.

### 3.2. Presuppositional determiners

As was pointed out in §2.1, all DPs headed by strong determiners may be clitic doubled in Albanian. The fact that the DPs headed by the strong determiners *çdo* 'every', *secilin* 'each', *të shumtët* 'most', and *asnjërin* 'neither (one of)'/ 'none (of)' may be clitic doubled in Albanian – see the examples in (29) – is problematic for our attempt to explain the Albanian clitic doubling purely algebraically, because singular universal quantifiers are assumed to be non-topicable in Endriss' (2009) analysis, as was demonstrated by the ungrammaticality of (20b).

- (29) a. (E) lexova çdo libër.  
 CL.ACC.3S read.1s every book  
 ‘I read each book.’
- b. (E) lexova secilin libër.  
 CL.ACC.3S read.1s each.the<sub>ACC</sub> book  
 ‘I read each book.’
- c. (I) lexova më të shumtët (e librave).  
 them<sub>CLACC</sub> read.1s sup most.pl.the agr books.the<sub>ACC</sub>  
 ‘I read most (books).’
- d. Nuk (e) lexova asnjërin (libër) / (prej librave).  
 not CL.ACC.3S read.1s neither.the<sub>ACC</sub> book / from books.the<sub>DAT</sub>  
 ‘I read neither (book / of the books).’

However, the fact that the DPs headed by these strong determiners may be clitic doubled is not necessarily an argument against Endriss’ 2009 framework because the minimal witness set for the universal quantifiers is the same for singular and plural; the minimal witness set for the quantifiers *every book*, *all books* and *each book* in any model would be the same, namely the set of all the books in the universe of discourse. Other things being equal, we would then expect universal quantifiers to be good topics independently of their grammatical number, with the consequence that the English facts in (20b) are surprising. Indeed Endriss 2009 attributes the non-topicality of singular universal quantifiers to a clash between the denotation of the minimal witness set and the morphological number of the quantifier (and the morphological number of the resumptive pronoun in the case of left-dislocated noun phrases like in (20b)). According to her, because the minimal witness set denotes plurality and the grammatical number on the quantifier (and the resumptive pronoun) is singular, there arises a mismatch, which is the reason for the ungrammaticality of singular universal quantifiers in syntactic positions associated with topics. It could be hypothesized that, for some reason, this clash in number doesn’t arise in Albanian, hence the singular universal quantifiers are topical in this language, perhaps due to the fact that clitics are very light elements (although they are not underspecified for number in Albanian). We would need to examine this hypothesis more thoroughly, but for now suffice it to mention that, as is well-known, a simple mapping between singular grammatical number and non-plurality of its denotation isn’t always viable. For instance, collective nouns like *team* or *government* in English denote pluralities but are grammatically singular (even though they may determine plural agreement). Similarly, from a cross-linguistic perspective, Slavic languages exhibit singular verbal agreement with subject DPs headed by numerals higher than four.

Another problematic case is the determiner *asnjërin* ‘neither (one of the)’/‘none (of the)’. As the determiner is negative strong, its minimal witness set is simply the empty set. Being so, we would expect that clitic doubling of a DP headed by this determiner shouldn’t be possible, contrary to fact – see (29d). We assume that the reason for this is the presuppositional behavior of determiners like *neither*. In this respect, there is a common core for all four quantifiers discussed in the present section: all these determiners are presuppositional, i.e., they presuppose the non-emptiness of the set denoted by their noun argument. We cannot go into the details of the presuppositional treatment of quantifiers here but would nonetheless like to mention that there is an ongoing discussion between the presuppositional and non-presuppositional treatment of quantifiers, which as far as we know has not been resolved yet. But starting at least with Barwise & Cooper (1981), it is usual to treat the determiners *the*, *both*, and *neither* as presuppositional. Nevertheless other determiners such as *every*, *all*, and *most* are also sometimes considered presuppositional, so basically all strong determiners are argued to

presuppose the non-emptiness of their restrictor (see Diesing 1992 and Heim & Kratzer 1998). For instance, Heim and Kratzer (1998:172) cite the following paradigm, originally due to Lumsden (1988), to test the presuppositionality of determiners. Filling the gap in (30) with the strong determiners *every*, *each*, *most*, or *neither* (as opposed to weak determiners like *two*, *no*, ...) leads to a presupposition that the speaker assumes that there are mistakes. We can then safely conclude that *neither* is different from *no* (and recall from §2.1 that *no* cannot be clitic doubled in Albanian, e.g. (5)), in that the former is presuppositional, whereas the latter is not.

(30) If you find \_\_\_ mistake(s), I'll give you a fine reward.

We assume that our approach, which relies on topicality being explained via minimal witnesses, should be enriched with some presuppositional theory of topichood like the one in Cresti (1995). According to Cresti (1995), topical constituents bear some kind of existence presupposition. Although we would have been happier to treat Albanian clitic doubling in a purely algebraic fashion, data like (29) convincingly show that a mixed, semantico-pragmatic theory is needed. Moreover, there is a common denominator between topicality defined in terms of minimal witnesses and topicality defined in terms of presupposition theories, namely the constraint on the non-emptiness of the restrictor (be it non-emptiness of minimal witnesses or non-emptiness as a presupposition). We leave the proper investigation of this common link for future work.

### 3.3. The wide scope of clitic doubled indefinites

As is well-known, indefinites may receive either wide or narrow scope with respect to other scope taking elements in the same sentence. For instance, an indefinite expression such as *një libër* 'a book' in (31a) can have either a wide scope or a narrow scope reading with respect to the implication. Under a wide scope reading there must be (at least) one book in the particular bookshop such that if Ben buys it, the book will ruin him financially. Under the narrow scope reading, if Ben buys (at least) one book in the bookshop, he will be broke (i.e. any book in the bookshop is so expensive that buying it will spell financial disaster for him). The predicate logic formulas corresponding to the wide scope and the narrow scope readings of the indefinite are given in (31b) and (31c), respectively.

- (31) a. Në qoftë se Beni do të blejë një libër në këtëlibrary,  
 in be that Ben FUT SUBJ buy a book in this bookshop  
 atëhere s'do të ketë më asnjë grosh.  
 then not-FUT SUBJ have.he more not.one cent  
 'If Ben buys one/any book in this bookshop, then he will be broke.'  
 b.  $\exists x[\text{book}'(x) \wedge [\text{buy}'(\text{Ben}, x) \rightarrow \text{broke}'(\text{Ben})]]$   
 c.  $[\exists x[\text{book}'(x) \wedge \text{buy}'(\text{Ben}, x)] \rightarrow \text{broke}'(\text{Ben})]$

In contrast, the sentence in (32a), which differs from the one in (31a) only in that the indefinite is clitic doubled, lacks the narrow scope reading for the indefinite. That is, unlike in (31a), in (32a) the indefinite must scope over the implication, as shown in (32b) (versus the unavailable narrow scope reading in (32c)), which says that the witness set of books (in that particular bookshop) contains such a member that if Ben buys that member, he will be broke. This is equivalent to the predicate logic formula in (31b) but we use the minimal witness set way to express the meaning because it explains the obligatory wide scope interpretation straightforwardly.

- (32) a. Në qoftë se Beni do ta blejë një libër në këtë library,  
 in be that Ben FUT SUBJ.CL.ACC.3S buy a book in this bookshop,  
 atëhere s'do të ketë më asnjë grosh.  
 then not-FUT SUBJ have.he more not.one cent  
 'If Ben buys a certain book in this bookshop, he will be broke.'  
 b.  $\exists P[\text{book}'(P) \wedge \min(P, \text{book}') \wedge \forall x[[P(x) \rightarrow \text{buy}(\text{Ben}, x)] \rightarrow \text{broke}'(\text{Ben})]]$   
 c.  $*[\exists x[\text{book}'(x) \wedge \text{buy}'(\text{Ben}, x)] \rightarrow \text{broke}'(\text{Ben})]$

Following Endriss and Hinterwimmer (2008), we take this semantic property of clitic doubled indefinites to follow from their topical interpretation. Endriss and Hinterwimmer postulate the rule in (33), where  $\alpha_T$  is the topical quantifier,  $Q$  is the comment and  $\min(P, \alpha_T)$  is to be read as 'P is a minimal witness set of  $\alpha_T$ '. Accordingly, our example in (23), repeated here again for ease of reference, would have the interpretation in (34): there is a minimal witness set of the quantifier *some books* and for some of the atoms in this set it holds that I read every atom last week.

- (23) Javën e shkuar (i) lexova disa libra.  
 week.the past them<sub>CL.ACC</sub> read.I some books  
 'Last week I read some books.'

- (33)  $\exists P[\alpha_T(P) \wedge \min(P, \alpha_T) \wedge \forall x[P(x) \rightarrow Q(x)]]$   
 (34)  $\exists P[\text{some\_books}'(P) \wedge \min(P, \text{some\_books}') \wedge \forall x[P(x) \rightarrow \text{read\_last\_week}'(I, x)]]$

This mechanism is in fact almost identical to the choice function treatment of indefinites (see e.g. Reinhart 1997, Winter 2000): it selects one of the elements from the minimal witness set of the quantifier and this element is interpreted as having wide scope over other operators in the sentence. What is new about it is that it explains the link between topicality and wide scope phenomena: if the quantifier is clitic doubled, then it is topical and receives wide scope. The reason is that the topical quantifier is interpreted as its witness set, from which one of its members is picked up via existential closure. This member is then distributed over the predicate (i.e. the focus part of the sentence).

### 3.4. Non-novel indefinites

In §2.3, we noted that clitic doubled indefinites, as in (18), repeated below for ease of reference, are so-called 'non-novel' indefinites (Krifka 2001).

- (18) a. Do ta pija me kënaqësi një uiski.  
 FUT SUBJ.CL.ACC.3S drink with pleasure a whisky  
 b. To pino exfaristos ena ouiskáki. (Kazasis & Pentheroudakis 1976:399)  
 it<sub>CL.ACC</sub> drink with pleasure a whisky  
 'I would gladly drink a whisky.'

Contra Heim's (1982) view, Krifka (2001) argues that indefinites may pick up discourse referents that exist in the input context. For a discourse referent to exist in the input context, it must either have been mentioned before in the immediate context, or its existence must in some way be presupposed (e.g. through sensory salience, via world knowledge, or typically through accommodation). Crucially, such non-novel indefinites must be deaccented, an idea that is in

tune with the well-known observation that across languages, ‘given’ (and therefore topical) information systematically correlates with lack of phonetic prominence (Ladd 1980, Schwarzschild 1999 and references therein). For Krifka, primary evidence for non-novel indefinites stems from adverbial quantification in connection with the so-called ‘requantification problem’ (Rooth 1985, 1995, von Stechow 1994), whereby the domain of quantification is given by the deaccented indefinite, which forces the assumption that indefinites may pick up existing discourse referents and ‘requantify’ over them, as illustrated in (35) (examples from Krifka 2001).

- (35) a. A freshman usually wears a báseball cap.  
       ‘Most freshmen usually wear a baseball cap.’  
       b. A fréshman usually wears a baseball cap.  
       ‘Most wearers of baseball caps are freshmen.’

That the clitic doubled indefinites in the sentences in (18) are non-novel is supported by several diagnostics. First, they are deaccented (i.e. the nuclear pitch accent cannot be borne by the clitic doubled expressions). Secondly, the clitic doubled indefinite in either sentence picks up a discourse referent whose existence in the input context is obviously presupposed, as can be seen by the fact that the sentences in (18) can be uttered felicitously in either of the contexts in (36). Finally, while the clitic doubled indefinite in (18a,b) functions as a kind of quotation in the context of (36a), it stands in a part-whole relationship with the indefinite ‘a drink’ in (36b), and is presupposed through accomodation in the context of (36c).

- (36) a. What about a whisky? / Would you like a whisky?  
       b. What about a drink? / Would you like a drink?  
       c. I have just stepped out of work.

Looking back at the ‘requantification’ sentences in (35) from the perspective of topics interpreted as minimal witnesses, we immediately see a problem. Deaccented indefinites in the examples are not interpreted as we would expect: (35a) according to our assumptions would claim that there is a witness of the set of freshmen who usually wears the baseball hat, which of course contradicts the meaning the sentence has. This problem can be resolved by the assumption that quantificational adverbs like *usually* quantify over pairs of situations and specify to which degree the situations denoted by the topical phrase are contained in the situation denoted by the rest of the sentence. In other words, (35a) means that most situations containing a freshman are situations with the freshmen wearing a baseball cap. The set of situations containing the denotation of the topical noun phrase is called *indirect aboutness topic* by Endriss and Hinterwimmer (2008) because the topical expressions identify the set of the situations and is the real topic of the sentence over which the quantificational adverb ranges.

#### 4. Summary and conclusion

In this paper, we have provided a formal semantic analysis of direct object clitic doubling in Albanian, which confirms and renders precise previous intuitions about this phenomenon (Kallulli 2000, 2008). Specifically, we have shown that clitic doubled direct object DPs must be interpreted as generating admissible minimal witnesses, which in turn makes these DPs topical. We consider clitic doubling to be a syncategorematic strategy for marking the clitic doubled DPs as topical, which renders weak quantifiers (at least in their narrow scope interpretation) and monotone decreasing quantifiers ungrammatical with clitic doubling, as these quantifiers

cannot be interpreted as topics (in the sense of topics being interpreted as minimal witness sets). In some cases, though, namely those involving the DPs headed by the strong determiners *çdo* ‘every’, *secilin* ‘each’, *të shumtët* ‘most’, and *asnjërin* ‘neither (one of)’/‘none (of)’, this purely algebraic approach must be supported by a presuppositional analysis of quantifiers. Finally, future work will have to deal with whether and to what extent this analysis can also account for clitic doubling of dative DPs, which is obligatory in all possible contexts in this language.

## References

- Adger, David. 2003. *Core syntax: A minimalist approach*. Oxford: Oxford University Press.
- Anagnostopoulou, Elena. 1994a. *Clitic dependencies in Modern Greek*. Doctoral dissertation, Salzburg University.
- Anagnostopoulou, Elena. 1994b. On the representation of clitic doubling in Modern Greek. *EUROTYP Working Papers*, Theme Group 8, 5:1–66.
- Anagnostopoulou, Elena, and Anastasia Giannakidou. 1995. Clitics and prominence, or why specificity is not enough. *Chicago Linguistics Society* 31:1–14.
- Barwise, Jon, and Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4:159–219.
- Cresti, Diana. 1995. *Indefinite topics*. Cambridge, MA: MIT dissertation.
- Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.
- Dobrovie-Sorin, Carmen. 1990. *The syntax of Romanian*. Berlin: Mouton de Gruyter.
- Endriss, Cornelia. 2009. *Quantificational topics*. Dordrecht: Springer.
- Endriss, Cornelia, and Stefan Hinterwimmer. 2008. Direct and indirect aboutness topics. *Acta Linguistica Hungarica* 55(3–4):297–307.
- Farkas, Donka. 1978. Direct and indirect object reduplication in Rumanian. *Papers from the Seventeenth Regional Meeting of the Chicago Linguistics Society* 14:88–97.
- Gutiérrez-Rexach, Javier. 1999. The formal semantics of clitic doubling. *Journal of Semantics* 16(4):315–380.
- Hedberg, Nancy. 2006. Topic-focus controversies. *The architecture of focus*, ed. by Valeria Molnar and Susanne Winkler. Berlin: Mouton de Gruyter.
- Heim, Irene. 1982. *The semantics of definite and indefinite noun phrases*. Amherst, MA: UMass dissertation.
- Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.
- Jaeggli, Oswald. 1982. *Topics in Romance syntax*. Dordrecht: Foris.
- Kadmon, Nirit. 1985. The discourse representation of noun phrases with numeral determiners. *Proceedings of NELS 15*, ed. by Stephen Berman, Jae-Woong Choe and Joyce McDonough. Amherst, MA: GLSA.
- Kallulli, Dalina. 1995. *Clitics in Albanian*. [University of Trondheim Working Papers in Linguistics 24.] Dragvoll: University of Trondheim.
- Kallulli, Dalina. 2000. Direct object clitic doubling in Albanian and Greek. *Clitic phenomena in European languages*, ed. by Frits Beukema and Marcel den Dikken, 209–248. Amsterdam: John Benjamins.
- Kallulli, Dalina. 2008. Clitic doubling, agreement, and information structure. *Clitic doubling in the Balkan languages*, ed. by Dalina Kallulli and Liliane Tasmowski, 227–255. Amsterdam: John Benjamins.
- Kallulli, Dalina, and Liliane Tasmowski (eds.) 2008. *Clitic doubling in the Balkan languages*, 1–32. Amsterdam: John Benjamins.
- Kamp, Hans, and Uwe Reyle. 1993. *From discourse to logic*. Dordrecht: Kluwer Academic Publishers.
- Kayne, Richard. 1975. *French syntax*. Cambridge, MA: MIT Press.
- Kazazis, Kostas, and Joseph Pentheroudakis. 1976. Reduplication of indefinite direct objects in Albanian and Modern Greek. *Language* 52:398–403.

- Keenan, Edward. 1996. The semantics of determiners. *The handbook of contemporary semantic theory*, ed. by Shalom Lappin, 41–63. Oxford: Blackwell.
- Krifka, Manfred. 2001. Non-novel indefinites in adverbial quantification. *Logical perspectives on language and information*, ed. by Cleo A. Condoravdi and Gerard Renardel de Lavalette, 1–40. Stanford, CA: CSLI Publications.
- Ladd, Robert. 1980. *The structure of intonational meaning: Evidence from English*. Bloomington: Indiana University Press.
- Ladusaw, William. 1982. Semantic constraints on the English partitive construction. *Proceedings of the First West Coast Conference on Formal Linguistics*.
- Lumsden, Michael. 1988. *Existential sentences*. London: Croom Helm.
- Montague, Richard. 1973. The proper treatment of quantification in ordinary English. *Approaches to natural language*, ed. by Jaakko Hintikka, Julius Moravcsik and Patrick Suppes, 221–242. Dordrecht: Foris.
- Reinhart, Tanya. 1981. Pragmatics and linguistics. An analysis of sentence topics. *Philosophica* 27:53–94.
- Reinhart, Tanya. 1997. Quantifier scope: how labour is divided between QR and Choice Functions. *Linguistics and Philosophy* 20:335–397.
- Rooth, M. 1985. *Association with focus*. Amherst, MA: UMass dissertation.
- Rooth, Mats. 1995. Indefinites, adverbs of quantification and focus semantics. *The generic book*, ed. by Gregory N. Carlson and Francis Jeffrey Pelletier, 265–291. Chicago: Chicago University Press.
- Schwarzschild, Roger. 1999. GIVENness, AvoidF and other constraints on the placement of accent. *Natural Language Semantics* 7:141–177.
- Sportiche, Dominique. 1996. Clitic constructions. *Phrase structure and the lexicon*, ed. by Johan Rooryck and Laurie Zaring, 213–87. Dordrecht: Kluwer.
- Strawson, Peter. 1974. *Subject and predicate in logic and grammar*. London: Methuen.
- Suñer, Margarita. 1988. The role of agreement in clitic doubled constructions. *Natural Language and Linguistic Theory* 6:391–434.
- Szabolcsi, Anna. 1997. Strategies for scope taking. *Ways of scope taking*, ed. by Anna Szabolcsi, 109–154. Dordrecht: Kluwer.
- Szabolcsi, Anna. 2010. *Quantification*. Cambridge: Cambridge University Press.
- Von Stechow, Kai. 1994. *Restrictions on quantifier domains*. Amherst, MA: UMass dissertation.
- Winter, Yoad. 2000. What makes choice natural? *Reference and anaphoric relations*, ed. by Urs Egli and Klaus von Stechow. Dordrecht: Kluwer.