

C2 “Case and referential context” of the SFB 732

Summary of the results in Phase 1

The term Differential Object Marking (DOM) is used to describe the phenomenon, found in many languages, that case marking of a direct object depends on certain semantic-pragmatic conditions. Whether this marking is obligatory, optional or ungrammatical is claimed to depend upon semantic factors such as animacy, definiteness or topicality which are ranked along universal implicational scales like the Animacy Scale in (1), the Referentiality Scale combining definiteness and specificity in (2), and the Topicality Scale in (3). A language cuts across each scale at one particular point – the language-specific transition point. In a number of languages DOM is obligatory for highly individuated objects, i.e. objects which are high on the animacy and/or referentiality and/or topicality scale. What makes this problem challenging is that there are multiple features triggering DOM. We combined these features under the term referential context and assumed that referential context is an architecture of semantic-pragmatic features which are responsible for the specification of underspecified lexical forms in context. Our goal was to find out how features of referential context interact and how they contribute to the specification of an underspecified lexical form of nouns in the context.

- (1) *Animacy Scale*: human > animate > inanimate
- (2) *Referentiality Scale*: personal pronoun > proper noun > definite NP > indefinite specific NP > indefinite non-specific NP > incorporated bare NPs
- (3) *Topicality Scale*: + topic > - topic

DOM has been investigated from functional, optimality theoretical and syntactic perspectives. Comrie (1975) and Bossong (1985) e.g. give a functional description of the phenomenon with reference to the scales. Aissen (2003) brings together scales and the preference type of a particular language in an optimality theoretical analysis in which two scales are combined via harmonic alignment. Analyses such as Brugè & Brugger (1996) focus on the syntactic side of DOM. However, all existing proposals concentrate on one or two features triggering DOM in particular languages. A comprehensive comparative analysis was not available when the project started.

We conducted detailed analyses of DOM in Mongolian, Turkish, Romanian, Spanish, Uzbek and Hindi. The analyses were based on grammars, corpus studies, questionnaires and interviews. The most important findings we gained are:

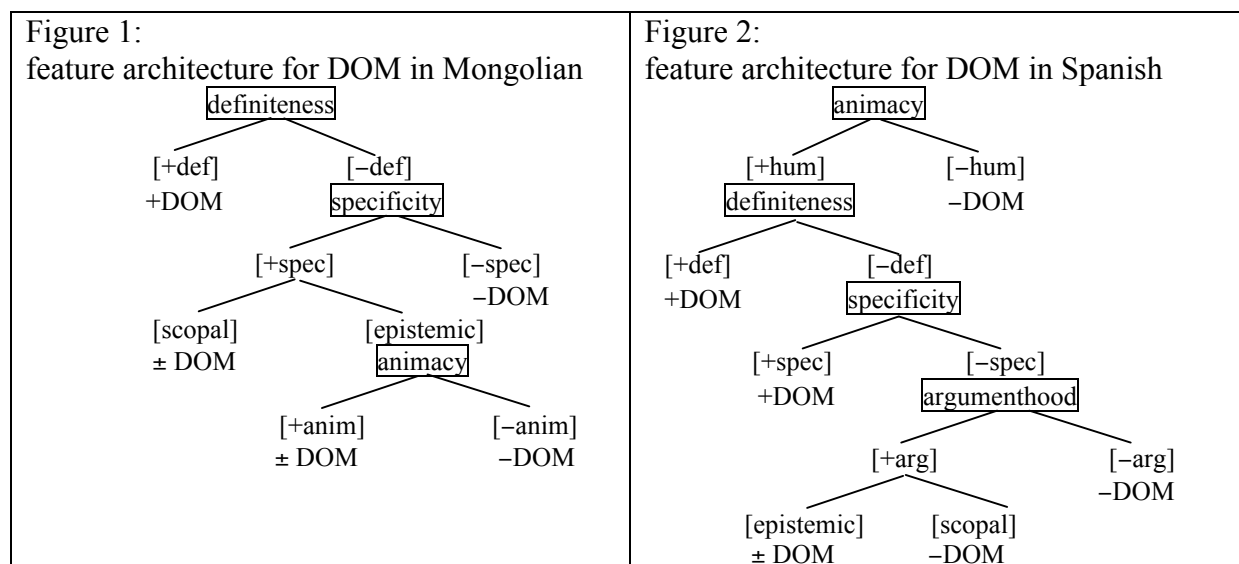
- (i) *Feature architectures*: While the same features of referential context (like animacy, definiteness, and specificity) tend to be relevant for DOM, most languages rank these features differently. We have developed detailed feature architectures which model the interaction of different features of referential context in particular languages.
- (ii) *Theoretical concept of specificity*: We argued that specificity as a feature involved in DOM is distinct from partitivity. We further developed the notion of specificity in terms of referential anchoring and argued that NPs can be anchored in different local domains.
- (iii) *New DOM triggers*: We found that differentially marked indefinite objects in Romanian tend to have a higher “referential persistence” in the following discourse. In the case of indefinites this feature is related to but distinct from the notion of specificity and may be understood as an additional feature on Topicality Scale.
- (iv) *Diachronic development of DOM and optionality*: We showed that optionality can be considered to be a reflex of diachronic change. Our diachronic studies revealed that

particular steps in the development of DOM from one category to another on the Referentiality Scale are facilitated by intervening “transitional” categories, such as topicality, and specificity. At stages where such categories come into play DOM is optional.

- (v) *Interaction of DOM and verbal semantics:* We find that certain semantic properties of the verb interact with nominal semantic features in triggering DOM. This can be shown both synchronically and diachronically. This finding suggests that DOM not only expresses nominal properties of the object but also properties of the relation between verb and object. However, the interaction between verbal and nominal semantic features is it not fully understood.
- (vi) *Motivation for case marking:* The comparison of differential object marking with differential subject marking revealed that differential case marking applies not only to the subject and object of the same verbal predicate but can also apply to a subject of a subordinate clause to distinguish it from the subject of the main clause.

A more detailed presentation of our findings is given below.

(i) *Feature architectures:* We conducted very detailed investigations of DOM systems in Spanish (von Heusinger 2008, von Heusinger & Kaiser submitted), Mongolian (Guntsetseg in preparation, 2009, 2008) and Uzbek (von Heusinger & Klein 2009, submitted) and compared them with systems described in the literature for other languages like Hindi and Turkic languages. To account for language-specific patterns and cross-linguistic variation we discussed in de Swart & Klein (submitted) feature architectures as a unified representational format. Two representative examples of feature architectures capturing a complex interaction between different factors triggering DOM are given in Figure 1 from Mongolian and Figure 2 from Spanish.



As Figure 1 shows, DOM in Mongolian depends primarily on definiteness. If the direct object is realized as a personal pronoun, proper name or definite NP, DOM is obligatory. The marking of indefinite NPs is optional in the sense that the marked form indicates additional semantic features. The situation in Mongolian is representative of other DOM languages we studied: we always find cases of obligatory DOM, other cases where it is excluded and optional cases of DOM. Languages differ in the sets of feature values for which they require, admit or forbid DOM. In case of optionality additional features come into play and can trigger

DOM. Thus, the case morpheme in a given language does not mark a single feature, but rather it can signal different semantic features of referential context. For instance, the accusative marker *-(i)g* on direct objects in Mongolian (i) does not signal additional features if the NP is definite, (ii) signals the feature [scopally specific] for indefinite NPs irrespective of animacy or (iii) the feature [epistemically specific] for animate indefinite NPs. However, in addition to these features, the feature of referential persistence (cf. (iii) below) can trigger DOM independent of the NP type.

Feature hierarchies of the type we used above facilitate cross-linguistic comparison of DOM systems and typological generalizations. The generalization evolving from our comparison of feature architectures in different languages is that the ranking of features is language-specific, e.g. in Spanish, as is shown in Figure 2, animacy has priority over definiteness, whereas in Mongolian the priority is reversed.

Another important finding in our investigation of feature interactions was that case markers that qualify as markers of DOM in some contexts may not have this function in other contexts. In Turkish and Romanian we observed that the validity of case markers as markers of DOM can be overruled by formal requirements of the grammar. Thus in Turkish, as shown in Kornfilt & von Heusinger (2009), factors like distance to the verb and the presence of a nominal agreement suffix¹ override the otherwise robust correlation between overt case marking and specificity. An analogous example can be found in Romanian: as we showed in Chiriacescu (submitted b), the comparative construction appears to formally require PE-marking irrespective of animacy or other semantic features of the referential context.

(ii) *Theoretical concept of specificity*: In our investigation of DOM in Mongolian and Romanian we saw that the categorial distinction +/- specific integrated in the Referentiality Scale in (2) is insufficient to describe specificity-depending DOM on indefinite NPs. This motivated us to further develop the theory of specificity. On the basis of specificity marking in Romanian, Spanish and Turkish in Kornfilt & von Heusinger (2009) and von Heusinger (2007) we suggested the more flexible concept of referentially anchored specificity, according to which a specific indefinite introduces a new discourse item that has a (pragmatically salient) link to an already given discourse item. Specific indefinites thus receive a representation containing an underspecified link to an anchor in the context. The investigation of specificity marking in Russian in Geist (2008) and in Russian vs. Romanian in Geist & Onea (2007) brought additional evidence in favor of an understanding of specificity as relative referential anchoring. We showed that in Russian different indefinite pronouns used as determiners, which according to the literature mark specificity, encode different types of anchoring. In the diachronic study of DOM in Romanian in von Heusinger & Onea (2008) we found additional evidence for the assumption that specificity has not just a +/-value, as the Referentiality Scale in (2) suggests, but subdivides into categories that build an additional scale depending on the type of the anchor. Specificity in terms of anchoring to the speaker (i.e. epistemic specificity) is more referentially stable and must be placed higher on the scale than specificity in terms of anchoring to other items in the clause. This subdivision of specificity was supported by diachronic Romanian data which suggested that epistemic specificity was marked earlier than specificity in terms of anchoring to other items. Another important finding concerning the category of specificity documented in Kornfilt & von Heusinger (2009) was that partitivity has to be distinguished from specificity (contra Enç 1991) and that the accusative case marker in Turkish expresses specificity rather than partitivity. The fine-graining of semantic categories we pursued in the investigation of specificity was further explored in the domain of definiteness in Wespel's (2008) dissertation.

¹ If an NP lacks a lexical nominal head an agreement morpheme shows up instead and this morpheme, in turn, requires an overt case marker.

(iii) *New DOM triggers*: In our studies of DOM in Romanian (Chiriacescu & von Heusinger 2009, Chiriacescu submitted a, in preparation) we investigated the impact of the larger discourse on DOM marking. We identified the discourse trigger of DOM, which we call “referential persistence”, a term introduced by Givón (1983). We showed that besides animacy, definiteness and specificity, referential persistence can trigger case marking of post-verbal indefinite direct objects. In Chiriacescu (submitted b) and von Heusinger & Chiriacescu (2009) we showed that referential persistence designates a discourse pragmatic property that is weaker than topicality and reconstructs the informal description of “importance for the subsequent discourse”. Objects marked by case for referential persistence are taken up more frequently in the subsequent discourse than their unmarked counterparts.

(iv) *Diachronic development of DOM and optionality*: We investigated optionality in relation that optionality can be considered to be a reflex of diachronic change. In our study of the diachronic development of DOM in Spanish (von Heusinger 2008) and Romanian (von Heusinger & Onea 2008) we observed that DOM started from elements high on the Referentiality Scale such as personal pronouns and spread steadily to elements low on this scale such as indefinite NPs. Data from diachronic corpora show that this expansion of DOM toward the right of the Referentiality Scale (2) was driven by intervening “transitional” categories, such as topicality, and specificity. These transitional categories are only active for the category of NP to which DOM is developing: topicality for definite NPs, specificity for indefinite NPs. On such transitional stages we attested optionality in DOM.

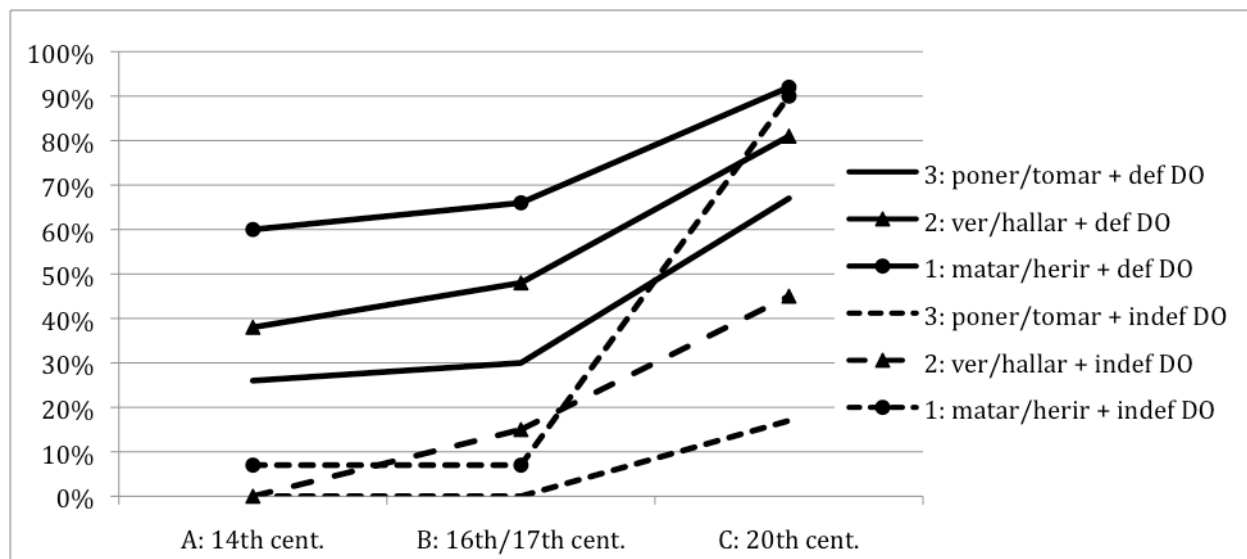
(v) *Interaction of DOM and verbal semantics*: The impact of verbal semantics on DOM, sometimes subsumed under the concept of Transitivity and Affectedness, was recognized in the literature but hasn’t been given a systematic analysis. We addressed this issue in our study of DOM in Spanish (von Heusinger 2008, von Heusinger & Kaiser submitted). First we reduced the complex notion of Affectedness to one distinct feature that was easy to control: We posited a scale of verb classes according to their preference to take animate direct objects with high transitivity verbs on the left side of the scale and low transitivity verbs on the right side: high transitivity verbs (class 1) have a high preference for human direct objects, class 2 has no preference, and low transitivity verbs (class 3) have a preference for inanimate direct objects.

(4) Scale of verb classes

Class 1	Class 2	Class 3
[+ human]	[± human]	[(±)/-animate]
>	>	
<i>matar</i> ‘kill’ <i>herir</i> ‘hurt’	<i>ver</i> ‘see’ <i>hallar</i> ‘find’	<i>poner</i> ‘put’ <i>tomar</i> ‘take’

We observed that in Spanish, the lexical class of the verb correlates with DOM in the following way (note that generally only human direct objects are case marked): class 1 verbs require obligatory DOM (with human direct objects) while class 2 verbs allow for variation and with class 3 verbs DOM is dispreferred. Further support for the scale in (4) came from our diachronic study. In von Heusinger (2008) we showed that the lexical semantics of the verb was an additional driving force in the diachronic evolution of DOM in Spanish. cf. Figure 3 from von Heusinger (2008).

Figure 3: Percentage of DOM comparing verb classes 1-3 with definite and indefinite human direct objects (Bible translations of 1+2 Samuel and 1+2 Kings)



This figure shows that the distribution of DOM depends on (I) the timeline, (II) the position on the Referentiality Scale, and (III) the verb class. We can see that DOM diachronically expanded rightwards on the scale of verb classes. Definite and indefinite objects of the higher-transitivity verbs (class 1 *matar/herir*) were marked earlier than the objects of lower-transitivity verbs (class 2 *ver/hallar* and 3 *poner/tomar*). The dependency of DOM on verb classes suggests that the case marker not only expresses nominal properties of the object, but also properties of the relation between verb and object. Similar effects of verb semantics can be observed in other DOM languages like Mongolian and Uzbek.

(vi) *Motivation for case marking*: In the typological literature two main functions of case marking are distinguished (e.g. Comrie 1989): the Distinguishing Function and the Identifying Function of case. According to the Distinguishing Function, DOM contributes to distinguishing the object from the subject in a clause if the object is too similar to the subject in individuation. By contrast, the Identifying Function motivates case marking as an indication that the direct object is a highly individuated, strong object. We considered these Functions as motivation for differential case marking. The typical context of DOM (highly individuated subject and an individuated object) does not allow for comparing these two theoretical positions since both predict marking of individuated objects. We investigated the problem of the function of case marking in the domain of Differential Subject Marking (DSM). For DSM the Distinguishing and the Identifying strategies make different predictions and thus can be teased apart more easily, cf. de Hoop & Narasimhan (2005). According to the Identifying strategy, case marking applies if the subject is high on the individuation scales (1)-(3). Under the Distinguishing strategy, subjects with a low degree of individuation should receive case marking in order to be distinguished from objects. Our previous collaborative work on DSM in Turkish (von Heusinger & Kornfilt 2005) has shown that differential case marking of objects and subjects in this language has an Identifying Function; the genitive case as a DSM marker in embedded sentences is used to mark definite and indefinite subjects if the subject is specific, i.e. high on the Referentiality Scale.

In Guntsetseg & Klein (2009), Klein et al. (submitted) and von Heusinger et al. (submitted) we investigated the case marking strategy in DSM in Mongolian. DSM in Mongolian differs from standard cases like Turkish, because in Mongolian it applies across clause boundaries. DSM shows up on the subject of embedded clauses if it is adjacent to the subject of the matrix clause, cf. (5). While the matrix subject occurs in the nominative, the adjacent embedded subject can occur in the nominative or in the accusative case – the case which is typically used for direct object marking.

- (5) Bold *Tuya/(-g)* yav-sn-i daraa yav-na. <Mongolian>
 Bold.Nom *Tuya(-ACC)* go-Pst-Gen after go-Fut
 ‘Bold will go after Tuya went.’

We conducted a detailed empirical study of the conditions for case marking with accusative case and case omission (nominative) on embedded subjects. In 3 questionnaires, we tested 1300 Mongolian speakers. The results show that the omission of the accusative on embedded subjects depends on the animacy and on the definiteness of the embedded subject relative to the actual matrix subject: the accusative is more likely to be omitted if the embedded subject is lower in referentiality than the matrix subject, but is hard to omit if the embedded subject is higher than the matrix subject. We concluded that the motivation for such case marking is just to distinguish the embedded subject from the actual matrix subject in cases where the embedded subject is equal or higher in referentiality than the matrix subject. This is an unexpected result since normally case markers are assumed to either distinguish arguments of the same predicate from one another or to identify semantic or pragmatic properties of some argument independently of other arguments of the predicate. If our conclusion is correct, then the distinguishability of the arguments proposed in the literature for a transitive relation under the term “Distinguishing Function of Case” should be complemented by the distinguishability of two subjects (or agents) belonging to different clauses. The requirement to morphologically distinguish two adjacent subjects may be due to processing factors.

Methods

In our work we relied on grammars and comprehensive linguistic data elicited from native speakers. We tested our hypotheses using linguistic corpora and developed written questionnaires and web questionnaires. Questionnaires are a conventional method of testing syntactic hypotheses. However, applying them to testing semantic/pragmatic differences proved difficult since in judging semantic information informants need to provide very subtle intuitions. To learn more about the optimal design of questionnaires and statistical methods of evaluation we took advantage of expertise of Sam Featherston (SFB 441, Tübingen), who worked as a Gastwissenschaftler at the Institute for Linguistics for 3 months.

We profited from discussions with project **A1** about concepts of definiteness and topicality. These features of referential context play a crucial role in triggering DOM. We also discussed possibilities of testing referential persistence – the feature we discovered in Romanian as a DOM trigger. In order to discuss the impact of information structure on DOM in a broader context **A1** and **C2** organized a workshop on “Focus at the Syntax-Semantics Interface”.

With **B1** our collaboration was both theoretical and methodological. On the theoretical level we discussed the issue of incremental specification of nominal features in context, and on the methodical level we shared experience in the development of questionnaires for testing semantic hypotheses, which was a new methodology for both projects.

C1 and **C2** are both interested in referential categories like specificity and definiteness. We discussed with **C1** interactions between specificity and definiteness and the impact of these features on morphological marking in languages that exhibit the so-called double definiteness.

The results of our analysis of DOM marking in Romanian as a marker of referential persistence (Chiriacescu submitted, Chiriacescu & von Heusinger 2009, von Heusinger & Chiriacescu 2009) were used as one topic for the new project **C4**.

We collaborated with experts in particular languages and language families: for South Asian languages with Miriam Butt in Konstanz, for Spanish with Manuel Leonetti in Madrid, for Turkic languages with Jaklin Kornfilt in Syracuse, USA and Marcel Erdal in Frankfurt. We discussed DOM and DSM in Altaic languages with the audience of the Workshop on “Altaic Formal Linguistics” in London (2008) and Nagoya (2009).

To broaden our knowledge about DOM in other languages we organized two workshops on “Case Marking in Bantu and Romance” and “Differential Case Marking”, where we discussed DOM in Bantu, Romance languages, Finnish, Manipuri and Russian with our collaborators. To further deepen our understanding of the DOM system in Romance and Altaic languages we organized separate workshops on “Differential Marking in Romance” and on “The DP structure of Altaic languages”.

We discussed two linguistic issues with international experts: theories of grammar to explain the multiple factors involved in DOM and the treatment of optionality. Concerning theories of grammar we collaborated with the research group on Optimal Communication at the faculty of linguistics of the Radboud University Nijmegen, chair Prof. Helen de Hoop. In our joined workshops we discussed advantages and disadvantages of different theories of grammar, in particular rule-based approaches and Optimality Theory. Peter de Swart, a collaborator of the Nijmegen group, visited our institute for 3 months in 2008 to prepare of a joined publication (de Swart & Klein submitted).

We collaborated with Miriam Butt and the members of her SFB project “The role of semantic fields in the development of postpositions and Case Markers” in Konstanz on the topic of semantics of DOM in South Asian languages. To get new insights into the treatment of optionality in case marking we organized workshops on “Transitivity and Case Alternations” and “Case Variation”, to which we invited the German experts Beatrice Primus and Elisabeth Stark as well as the international experts Judith Aissen, Arto Anttila, Jóhanna Barðdal, Andrej Malchukov and Áshild Næss.

Publications:

- Chiriacescu, S. & von Heusinger, K. 2009. *Pe*-marking and referential persistence in Romanian. In: E. Onea & A. Riester (eds.), *SinSpeC. Working Papers of the SFB 732*. Vol. 3. (Proceedings of the Workshop Focus at the Syntax-Semantics Interface). Online Publikationsverbund der Universität Stuttgart (OPUS). 1-19.
- Geist, L. & Onea, E. 2007. Specificity and implicatures. In: Maria Aloni, Paul Dekker and Floris Roelofsen (eds.) *Proceedings of the Sixteenth Amsterdam Colloquium* December 17-19, 2007. 109-114.
- Geist, L. 2008. Specificity as referential anchoring: evidence from Russian. In: Grønn, A. (ed.): *Proceedings of SuB12*. Oslo: ILOS 2008. 151-164
- Guntsetseg, D. 2008. Differential Object Marking in Mongolian. In: F. Schäfer (ed.), *SinSpeC. Working Papers of the SFB 732*. Vol. 1. Online Publikationsverbund der Universität Stuttgart (OPUS). 53-69.
- Guntsetseg, D. 2009. Differential Object Marking in (Khalkha-)Mongolian. In: R. Shibagaki & R. Vermeulen (eds.). *Proceedings of the 5th Workshop on Altaic Formal Linguistics (WAFL 5)*. *MIT Working Papers in Linguistics (MWPL)* 58. 115-129.
- Guntsetseg, D. and Klein, U. 2009. Two types of case alternation in Mongolian. In: *Current Issues in Unity and Diversity of Languages. Collection of the papers selected from the CIL 18*. Linguistic Society of Korea. 3114-3128.
- Kornfilt, J. & von Heusinger, K. 2009. Specificity and partitivity. In: R. Shibagaki & R. Vermeulen (eds.). *Proceedings of the 5th Workshop on Formal Altaic Linguistics (WAFL 5)*. *MIT Working Papers in Linguistics (MWPL)* 58. 19-40.
- von Heusinger, K. & Chiriacescu, S. 2009. Definite “Bare” nouns and *pe*-marking in Romanian. In: M.T. Espinal, M. Leonetti & L. McNally (eds.), *Proceedings of the IV Nereus International Workshop “Definiteness and DP Structure in Romance Languages”*. Arbeitspapier Nr. 124. Fachbereich Sprachwissenschaft, Universität Konstanz 2009. 63-82.

- von Heusinger, K. & Klein, U. 2009. Two indefinite Articles in Uzbek. In: R. Shibagaki & R. Vermeulen (eds.). *Proceedings of the 5th Workshop on Altaic Formal Linguistics (WAFL 5). MIT Working Papers in Linguistics (MWPL) 58*. 153-164.
- von Heusinger, K. 2008. Verbal Semantics and the diachronic development of Differential Object Marking in Spanish. *Probus* 20(1). 1-33.
- von Heusinger, K. 2007. Referentially anchored indefinite NPs. In: I. Comorovski & K. von Heusinger (eds.). *Existence: Syntax and Semantics*. Dordrecht: Springer. (Studies in Linguistics and Philosophy 84) 273-292.
- von Heusinger, K. & Onea, E. 2008. Triggering and blocking effects in the diachronic development of DOM in Romanian. *Probus* 20(1). 71-116.

Submitted:

- Chiriacescu, S. (submitted a). DOM in Romanian and the referential form - mental accessibility interplay. *Bulletin of the "Transilvania" University of Brasov*. Vol. 2 (51). Series IV: Philology and Cultural Studies. (submitted 02.07.09)
- Chiriacescu, S. (submitted b). Indefinite NPs and pe-marking in Romanian. In: *Lucrarile celui de-al doilea Simpozion International de Lingvistica*, ed. N. Saramandu & M. Nevaci & C. I. Radu. Bucharest: Bucharest UP, 2009. (submitted 20.02.09)
- de Swart, P. & Klein, U. (submitted). Case and referential properties. *Lingua*. (submitted 24.07.09).
- Klein, U. & Guntsetseg, D. & von Heusinger, K. (submitted). Case in conflict: embedded subjects in Mongolian. To appear in Lamers, M. and de Swart, P. *Case, Word Order, and Prominence: Psycholinguistic and Theoretical Approaches to Argument Structure*. Studies in Theoretical Psycholinguistics. Springer. (submitted 25.05.09)
- von Heusinger, K. & Kaiser, G. A. (submitted). Affectedness and Differential Object Marking in Spanish. *Morphology*. (submitted revised version 1.12.09)
- von Heusinger, K. & Klein, U. (submitted). The distribution of two indefinite articles in Uzbek. In: C. Ebert & S. Hinterwimmer (eds.). *Funny Indefinites*. Dordrecht: Springer. (Studies in Linguistics and Philosophy) (submitted 25.05.08)
- von Heusinger, K. & Klein, U. & Guntsetseg, D. (submitted). The case of accusative embedded subjects in Mongolian. *Lingua*. (submitted 27.05.09)

Dissertation projects:

- Chiriacescu, S. in preparation. Indefinites and their discourse functions. Ph.D. dissertation, University Stuttgart. The dissertation will be completed by the end of phase 1 of the SFB.
- Guntsetseg, D. in preparation. Differential Case Marking in Mongolian. Ph.D. dissertation, University Stuttgart. The dissertation will be completed by the end of phase 1 of the SFB.
- Wespel, J. 2008. Descriptions and their domains. Ph.D. dissertation, University Stuttgart, published in the SinSpeC. *Working Papers of the SFB 732*. Vol. 2. Online Publikationsverbund der Universität Stuttgart (OPUS). 1-221.