Transitivity and the diachronic development of differential object marking

Klaus von Heusinger, Udo Klein, Dildora Niyazmetowa
Workshop "Transitivity and case alternations"

Stuttgart, 14th and 15th of January, 2008

Abstract

The relation between transitivity and differential object marking (and case marking more generally) has been discussed mainly from a synchronic point of view. In this paper we discuss the relation between transitivity and DOM from a diachronic perspective.

The main claim is that the spread of DOM from one point on the definiteness/animacy scale to the next lower point is not homogenous accross verbs, but depends on the transitivity of the verb: the P arguments of higher-transitivity verbs are marked earlier than the P arguments of lower-transitivity verbs.

We provide diachronic evidence from the development of Spanish DOM and synchronic evidence from Uzbek DOM for this hypothesis.

1 Background: transitivity, prototypicality and markedness

Comrie (1979, 1989):

- based in part on text count in Givón (1979):

  - about 10% of subjects are indefinite
  - essentially equal number of definite and indefinite objects
  - therefore: if there is an indefinite NP in the clause, it is more likely to be the object than the subject

- "in actual discourse there is a strong tendency for the information flow from A to P to correlate with an information flow from more to less animate and from more to less definite"

- "the most natural kind of transitive construction is one where the A is high in animacy and deniteness, and the P is lower in animacy and deniteness; and any deviation from this pattern leads to a more marked construction." (p. 128)

This predicts/explains two types of phenomena:

- where subjects are differently marked when they are low in definiteness/animacy

- where objects are differently marked when they are high in definiteness/animacy – these are the DOM languages

Hopper and Thompson (1980):

- text count: "Our own statistics suggest that, in foregrounding, there is a marked tendency for O's to be individuated, i.e. to have properties associated with referentiality/definiteness and animacy."

- foregrounding ("material which supplies the main points of the discourse" p. 289) correlates with the ten transitivity features: participants, volitionality, agentivity, affectedness of O, individuation of O, kinesis, aspect, punctuality, affirmation, mode.

- "non-referential or indefinite O's show striking correlations with the verb morphology, case-markings, and word-order characteristic of 'intransitive' clauses in a number of languages." (p. 259)

- "These correlations suggest that the special markings on definite O's, found in many languages, are better interpreted functionally as signals of the high Transitivity of the clause as a whole – rather than as devices for distinguishing O's from A's" (p. 259)

Tsunoda (1985, p. 388) proposes the following scale of transitivity/affectedness:

- Type 1: direct effect on patient A) resultative: kill, break, bend,.... B) non-resultative: hit, shoot, eat

- Type 2: perception A) patient more attained: see, hear, find B) patient less attained: listen, look

- Type 3: persuit: search, wait, await

- Type 4: knowledge: know, understand, remember, forget

- Type 5: feeling: love, like, want

- Type 6: relationship: possess, have, lack

- Type 7: ability: capable, proficient, good

Næss (2007) proposes to unify the two functions of case marking as follows:

- "A prototypical transitive clause is one where the two participants are **maximally semantically distinct** in terms of their roles in the event described by the clause." (p. 30)

- "The **canonical** function of core case-marking is to discriminate between the participants in a fully transitive clause, that is, between Agents and Patients in maximal semantic distinction." (p. 167)

- This canonical function can extend along both the discriminatory and the indexing dimension.

  - distinguish, regardless of maximal semantic distinctness of arguments
  - index, regardless of maximal semantic distinctness of arguments

2 Transitivity and the transition points of DOM

The main claim is that the spread of DOM from one point on the definiteness/animacy scale to the next lower point is not homogenous accross verbs, but depends on the transitivity of the verb: the P arguments of higher-transitivity verbs are marked earlier than the P arguments of lower-transitivity verbs. In this paper we provide diachronic evidence from the development of Spanish DOM, and synchronic evidence from Uzbek DOM to support this claim.

2.1 Spanish

von Heusinger (to appear) presents and discusses the results of two corpus studies about the relation between verbs semantics and the development of differential object marking in Spanish. Here we summarise some of the main results.

Scale of Affectedness and expected animacy of the direct object:

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

### 2.1.1 First corpus study: Bible (Samuel and Kings)

Percentage of DOM with definite human direct objects (number of all definite human objects in brackets; Bible translations of 1+2 Samuel and 1+2 Kings.).

class	verb	A: 14th cent.	B: 16th/17th cent.	C: 20th cent. (Euro)	D: 20th cent. (Am)
3	poner	25% (4)	50% (6)	83% (6)	100% (6)
	tomar	31% (19)	23% (17)	62% (24)	68% (25)
	sum	**30% (23)**	**30% (23)**	**67% (30)**	**74% (31)**
2	ver	35% (20)	41% (22)	83% (29)	75% (20)
	hallar	50% (4)	80% (5)	66% (3)	75% (4)
	sum	**38% (24)**	**48% (27)**	**81% (32)**	**75% (24)**
1	matar	59% (32)	85% (27)	92% (27)	100% (27)
	herir	62% (8)	48% (29)	83% (12)	81% (16)
	sum	**60% (40)**	**66% (56)**	**92% (39)**	**93% (43)**

Percentage of DOM with indefinite human direct objects (number of all indefinite human objects in brackets; Bible translations of 1+2 Samuel and 1+2 Kings):

class	verb	A: 14th cent.	B: 16th/17th cent.	C: 20th cent. (Euro)	D: 20th cent. (Am)
3	poner	0% (7)	0% (14)	14% (7)	0% (9)
	tomar	0% (8)	0% (14)	20% (5)	28% (7)
	sum	**0% (15)**	**0% (28)**	**17% (12)**	**13% (16)**
2	ver	0% (7)	02% (10)	50% (8)	56% (9)
	hallar	0% (4)	0% (3)	33% (3)	100% (3)
	sum	**0% (11)**	**15% (13)**	**45% (11)**	**67% (12)**
1	matar	7% (14)	14% (7)	87% (8)	100% (9)
	herir	–% (0)	0% (7)	100% (3)	100% (4)
	sum	**7% (14)**	**7% (14)**	**90% (11)**	**100% (13)**

The table below compares the percentage of DOM-marked human definite and indefinite direct objects of the three classes of verbs.
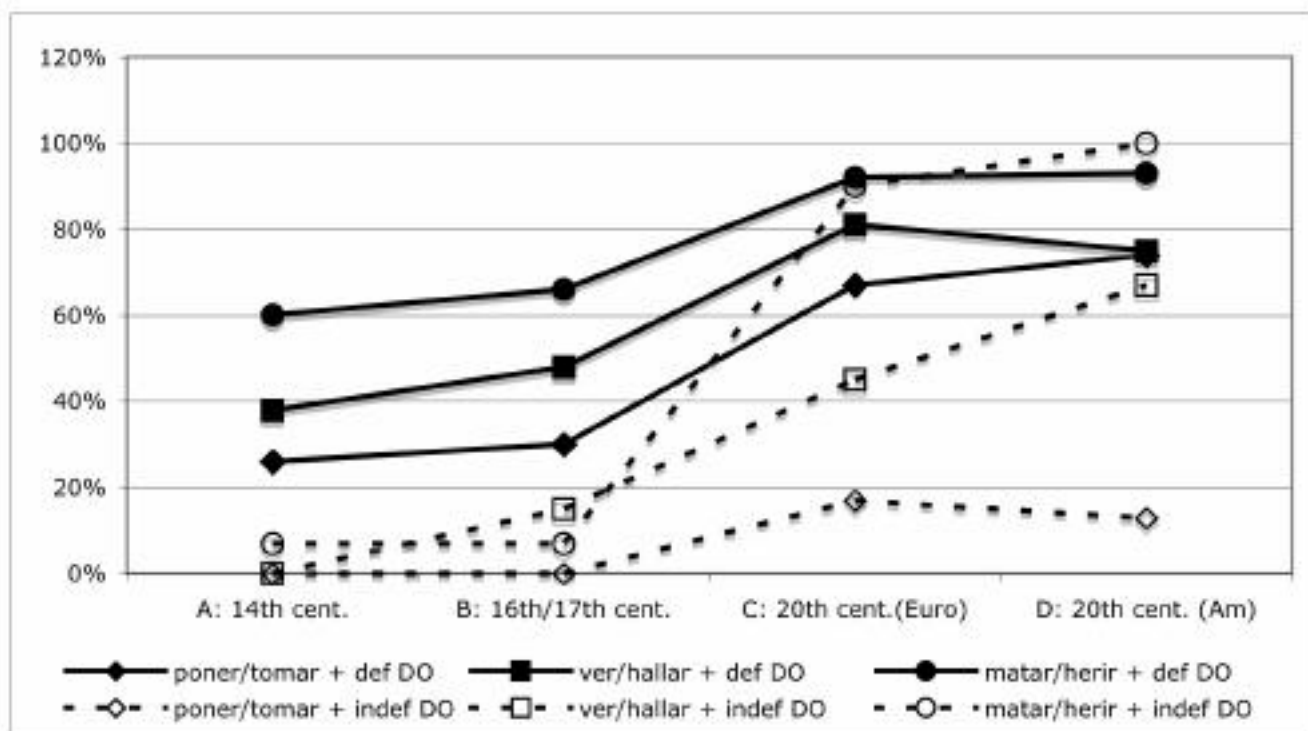
Question: Can these findings which were based on selected chapters of the Bible be confirmed by a larger corpus search?
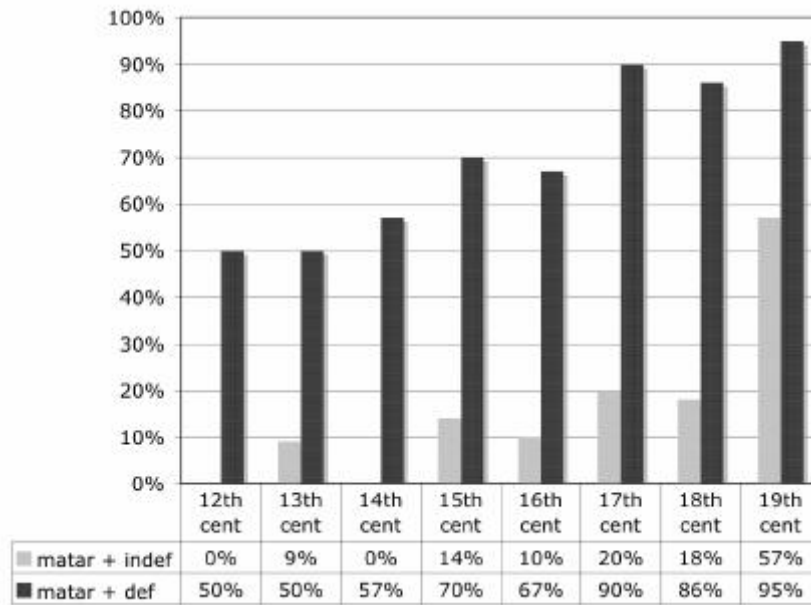
2.1.2 Second corpus study: Corpus del Español

Comparison of matar (kill) and tomar (take).

Matar with human definite and indefinite direct objects:

matar	12th cent	13th cent	14th cent	15th cent	16th cent	17th cent	18thcent	19th cent
def	13	12	10	6	7	2	3	1
def +a	13	12	13	14	14	18	19	20
sum	26	24	23	20	21	20	22	21
indef	20	20	19	18	18	16	18	9
indef +a	0	2	0	3	2	4	4	12
sum	20	22	19	21	20	20	22	21

Percentages of DOM for matar (kill) with definite and indefinite human direct objects:

	12th cent	13th cent	14th cent	15th cent	16th cent	17th cent	18th cent	19th cent
matar + indef	0%	9%	0%	14%	10%	20%	18%	57%
matar + def	50%	50%	57%	70%	67%	90%	86%	95%

Tomar with human definite and indefinite direct objects:

tomar	12th cent	13th cent	14th cent	15th cent	16th cent	17th cent	18thcent	19th cent
def	32	25	16	9	1	2	2	2
def +a	14	24	16	14	6	1	5	15
sum	46	49	32	23	7	3	7	17
indef	28	5	8	37	3	5	15	9
indef +a	1	0	1	1	0	1	3	4
sum	29	5	9	38	3	6	18	13

Percentages of DOM for tomar (take) with definite and indefinite human direct objects:

	12th cent	13th cent	14th cent	15th cent	16th cent	17th cent	18th cent	19th cent
tomar + indef	4%	0%	11%	2%	0%	18%	17%	30%
tomar + def	30%	49%	50%	61%	80%	33%	67%	88%
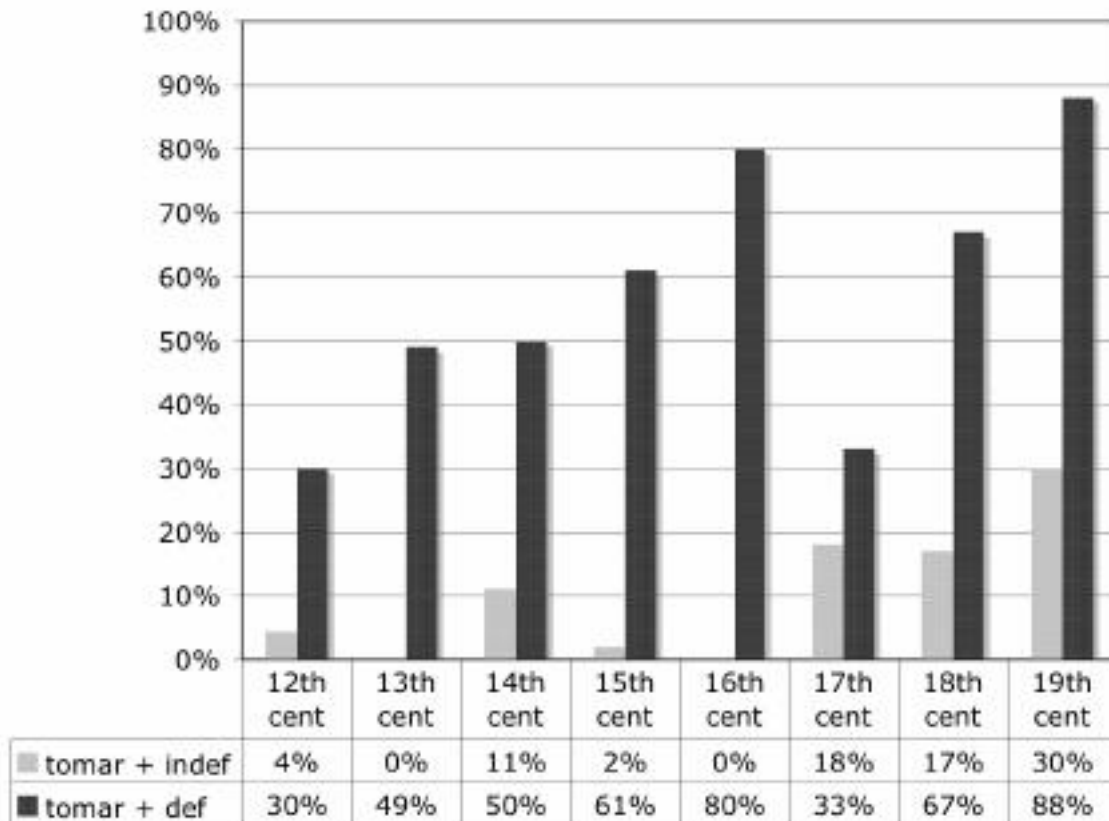
Summary:

- DOM develops first on arguments higher on the definiteness scale, and later on arguments which are lower.

- The development of DOM also depends on the verb class.

- Bible corpus: About 60% of definite human objects of `matar` are marked in the 16th century. Verbs of the second class reached a similar rate at least a century later. Verbs of the third class reached a similar rate at least two centuries later.

2.2 Uzbek

2.2.1 Preliminaries

If the main claim of this paper is correct (the P arguments of higher-transitivity verbs are marked earlier than the P arguments of lower-transitivity verbs), we predict that:

- if only arguments of some verbs are differentially object marked, then it is arguments of higher-transitivity verbs (Hindi?).

- if DOM is optional for a certain type of argument, then differentially marked objects of higher-transitivity verbs are judged better than formally marked objects of lower-transitivity verbs (Mongolian?)

- differential marking of a certain type of argument may become obligatory for (some) high-transitivity verbs while still being optional or ungrammatical for (some) low-transitivity verbs (Uzbek).

2.2.2 DOM in Uzbek

ACC marker is obligatory with pronouns, names and demonstrative/definite NPs as direct objects.

(1) a. U   me*(-ni)   tani-ma-di.
         3SG 1SG-ACC recognise-NEG-PRF

         S/he didn't recognise me.

    b. Biz Toschkent*(-ni) aylan-ma-dik.
         1PL Taschkent-ACC turn-NEG-1PL

         We did not walk through Tashkent.

    c. Biz bu   hikoya*(-ni) uqi-gan-miz.
         1PL DEM stories-ACC read-PST-1PL

         We read these stories.

With indefinite NPs referring to animate beings ACC is obligatory, too:

(2) Sen bitta muschuk*(-ni) urvor-ding-mi?
       2SG a   cat-ACC   run.over-PRF.2SG-Q
       Have you run over a cat?

The ACC marking of indefinite inanimate objects is more complex since it depends on a number of parameters.

The first parameter is partitivity: if an object is to be interpreted partitively, then the ACC marker is obligatory:

(3)　　Rasta-da　　　besch-ta kitob bor. Bitta kitob*(-ni) kecha　　　uq-di-m.
　　　　bookshelf-LOC five-CL　book exist a　　book-ACC yesterday read-PRF-1SG
　　　　There are five books on the shelf. One of the books I read yesterday.

(4)　　Men bitta moschina(*-ni) sot-ib　　　ol-di-m.
　　　　1SG a　　car-ACC　　　sell-PTCP get-PRF-1Sg
　　　　I bought one of the cars.

The reverse does not hold, i.e. not every ACC marked indefinite inanimate object must be interpreted partitively.

(5)　　Men kecha　　bitta rus-cha　　kitob-ni　u'q'i-di-m.
　　　　1SG 3SG-DAT a　　Russian-in book-ACC read-PRF-1SG
　　　　Yesterday I read a Russian novel [not necessarily partitive].

Secondly, if a dirct object is modified by a (restricted) relative clause then the ACC marker seems obligatory.

(6)　　Men hozir Ispaniya-da sot-ib　　ol-gan　bitta kitob-im-ni　　u'q'i-yap-man.
　　　　I　　now　Spain-LOC sell-GER get-PST a　　book-1SG-ACC read-PRES-1SG
　　　　I'm reading a book I bought in Spain.

(7)　　Men Farhod tavsiya　　　q'il-gan　bitta DVD-ni　sot-ib
　　　　I　Farhod recommandation make-PST a　　DVD-ACC sell-GER
　　　　ol-di-m.
　　　　get-PRF-1SG
　　　　I bought a DVD-ACC which Farhod recommended.

Thirdly, the direct object of a perfective construction must be marked with ACC.

(8)　　a.　　U　kecha　　bitta kitob uqi-di
　　　　　　　3SG yesterday a　　book read-PRF:3SG
　　　　　　　He has read a book yesterday.

　　　　b.　　U　kecha　　bitta kitob*(-ni) uqi-ib　　tugat-di
　　　　　　　3SG yesterday a　　book-ACC read-GER finish-3SG
　　　　　　　He finished reading a book yesterday.

Fourthly, when none of the above apply, then ACC marking depends on the type of verb and the individuation (type of modification) of the referent.

With a first class of verbs (repair, erase, break, etc.) the ACC marking is obligatory, even if the object is **not** partitive or modified by a relative clause.

(9)　　a.　　Men bitta stol*(-ni)　tuzat-di-m.
　　　　　　　1SG a　　table-ACC repair-PRF-1SG
　　　　　　　I have repaired a table. (not necessarily partitive).

　　　　b.　　U　bitta suz*(-ni)　uchir-di.
　　　　　　　3SG a　　word-ACC delete-PRF
　　　　　　　S/he deleted a word (not necessarily partitive).

The ACC is obligatory even if the speaker does not have a specific entity in mind.

(10)　　Farhod bitta moshina*(-ni) tuzat-ib-di.
　　　　Farhod a　　car-ACC　　　repair-EVID-PRF
　　　　(I have heard that) Farhod has repaired a car.

With a second class of verbs, the ACC is grammatical if the indefinite inanimate object is intended to be interpreted partitively or is modified somehow, and ungrammatical if it is not modified.

(11) U   men Ispaniya-da sot-ib   ol-gan   bitta kitob*(-ni)   u'q'i-di.
      S/he I   Spain-LOC sell-GER get-PST a   book-1SG-ACC read-PRF

      S/he has read a book I bought in Spain.

(12) Men bitta machsus kitob(-ni) u'q'i-di-m
      I   a   special   book-ACC read-PRF-1SG

      I've read a special book. [not necessarily partitive]

(13) Men bitta rus-ch'a   kitob(-ni) u'q'i-di-m
      I   a   Russian-in book-ACC read-PRF-1SG

      I've read a Russian book.[not necessarily partitive]

(14) Men bitta kitob(*-ni) u'q'i-di-m.
      1SG a   novel-ACC book-PRF-1SG

      I read a book.

With a third class of verbs the ACC marker is grammatical only if the object is to be interpreted partitively or if it is modified by a relative clause. Otherwise the ACC is ungrammatical.

(15) Men Farhod tavsiya   q'il-gan   bitta DVD-ni   sot-ib
      I   Farhod recommandation make-PST a   DVD-ACC sell-GER
      ol-di-m.
      get-PRF-1SG

      I bought a DVD-ACC which Farhod recommended.

(16) Men bitta machsus kitob(*-ni) sot-ib   ol-dim
      I   a   special   book-ACC sell-GER get-PRF-1SG

      I've bought a special book.

(17) Men bitta rus-ch'a   kitob(*-ni) sot-ib   ol-di-m
      I   a   Russian-in book-ACC sell-GER get-PRF-1SG

      I've bought a Russian book.

(18) Men bitta book(*-ni) sot-ib   ol-dim.
      1SG a   book-ACC sell-GER get-PRF-1SG

      I bought a book.

The differential object marking of indefinite inanimate direct objects is summed up in the following table:

	a N	a ADJ N	a special N	partitive	RC
repair, delete	+	+	+	+	+
read, show	–	±	±	+	+
buy, eat	–	–	–	+	+

In summary: A number of transitivity factors are involved in the ACC marking of indefinite inanimate DOs.

- Perfectivity: Indefinite objects in explicitly perfective constructions are obligatorily marked with ACC.

- Affectedness: Verbs which imply a change of an independently existing object seem to require ACC marking. Verbs which do not imply such a change of an independently existing object do not allow ACC of an unmodified indefinite inanimate object.

- Individuation: (i) Partitivity of indefinite objects is expressed by means of ACC. (ii) ACC marking of indefinite inanimate objects is also sensitive to the presence (and type) of modification.

The fact that the indefinite inanimate objects of higher-transitivity verbs that are obligatorily marked, while the indefinite objects of some lower-transitivity verbs cannot be marked provides evidence for the claim that the arguments of higher-transitivity verbs are marked earlier than arguments of lower-transitivity verbs.

3 Modelling diachronic change and further questions

A possible analysis of synchronic and diachronic DOM:

- DOM languages have two language-specific rules for combining transitive verbs and direct objects – call them the high- and low-transitivity rules respectively.

- The rules apply if verb and argument(s) satisfy certain language-specific semantic conditions.

- The rules can be viewed either as adding or as requiring a certain encoding of the argument(s). E.g. the high-transitivity rule may require the P argument to be ACC-marked, while the low-transitivity rule requires the P argument to be realised as unmarked for case.

- If an argument can be combined by means of either rule, then DOM is optional.

- Development of DOM along the definiteness/animacy scales involves three stages:

  – the conditions of the high-transitivity rule are reanalysed, allowing for some arguments which could only be combined by the low-transitivity rule to be combined by the high-transitivity rule

  – for a certain type of arguments there is a the competition between the two rules, which results in preferring the high-transitivity rule, and

  – the preference for using the high-transitivity rule is grammaticalised.

Further questions:

- What is the unifying feature behind the components of transitivity? Hopper and Thompson (1980) suggest foregrounding, Næss (2007) suggests maximal semantic distinctness.

- Although transitivity plays an important role in the development of DOM, it seems that the overt marker always ends up correlating with the definiteness/animacy of the argument, and not with the properties of the verb. If this is correct, then why should it be so?

- Are we being viciously circular if we determine the language-specific notion of affectedness on the basis of formal structure, and then use the notion of language-specific affectedness in order to explain formal structure?

- How to embedd language-specific notions of affectedness into a general (formal?) theory of affectedness?

- How to best account for the multi-dimensional nature of DOM?

References

Bernard Comrie. Definite and animate direct objects: a natural class. *Linguistica Silesiana*, 3:13–21, 1979.

Bernard Comrie. *Language universals and linguistic typology*. University of Chicago Press, Chicago, 2 edition, 1989.

Talmy Givón. *On understanding grammar*. Academic Press, New York, 1979.

Paul J. Hopper and Sandra A. Thompson. Transitivity in Grammar and Discourse. *Language*, 56:251–299, 1980.

Åshild Næss. *Prototypical transitivity*. John Benjamins, Amsterdam/Philadelphia, 2007.

Tasaku Tsunoda. Remarks on transitivity. *Journal of linguistics*, 21:385–396, 1985.

Klaus von Heusinger. Verbal Semantics and the Diachronic Development of Differential Object Marking in Spanish. *Probus*, 20, to appear.