

Differential Object Marking (DOM) in Spanish: Emergence and tendencies of the current system

J. Kabatek, Ph. Obrist, A. Wall (3-year term SNF project starting from 1st April 2016)

1. Scope and objectives

Spanish is a particularly important language for the study of DOM, both due to its many speakers and prominence and its importance for the scholarly description of the phenomenon (Bossong 1991). But despite the high degree of grammaticalization of what is traditionally referred to as "prepositional accusative" and it being subject to studies on many different aspects, there are still major research gaps: How did the current system emerge? In which direction is it currently evolving?

This project focusses on innovation in terms of empiricism (new data), methodology (new types of data, new visualizations) and theory (new findings on typological correlations between DOM and other elements). It will provide new insights about a major feature of Spanish grammar for Hispanic and Romance linguistics and beyond.

1.2 DOM in Spanish

Modern Spanish, "canonical" use

(1) Juan quiere a María. / *Juan quiere María.
Juan love3.SG DOM María.
Juan loves María.

(2) Juan lee un libro. / *Juan lee a un libro.
Juan read3.SG a Buch
Juan reads a book.

Complex variation in other contexts

(3) Cosecharon al maíz.
harvested3.PL DOM+DET corn.
They harvested the corn. (Kany 1969)

1.3 State of research & Goals

Emergence and spreading in the Middle Ages:

- marks objects accord. to their function in discourse: topicality
- the following grammaticalization process deviates to animacy, definiteness, specificity

Hypotheses on new developments:

- Change hypothesis** (e.g. Company Company 2002): DOM extends to inanimate objects.
- Stability hypothesis** (e.g. García García 2015): Current variation of DOM is essentially stable.

→ Research gap concerning the emergency of the current system (end of 17th to beginning of 19th century, cf. Kabatek in press)

→ Research gap concerning current tendencies: American vs. Peninsular Spanish, extension of DOM towards inanimate objects, processing of DOM (cf. Wall 2015)

Research question 1: Where and when did the current system of Spanish DOM emerge?

Research question 2: What are its current tendencies in the Spanish-speaking world (particularly in spoken language)?

2 Preliminary results

2.1 Research question 1

We consider Early Modern Spanish (approx. 1675-1825, cf. Octavio de Toledo 2013) to be a crucial period for the establishment of the new system → **new data**

The current rules do not apply to 18th century Spanish:

(4) ... para ir a recibir la Reyna [...]
to go to receive DET queen
... *to go and receive the Queen [of the Two Sicilies]*
(Mercurio1738, 13)

(5) ... dexàra a Madrid si no...
leave3SG.CON.IMP DOM M. if not
... *[he] would have left Madrid, if not for...*
(Aventuras, Madrid, 202)

→ „sorting out the variants“ (Stein 1994)

Differentiated description of the new system:

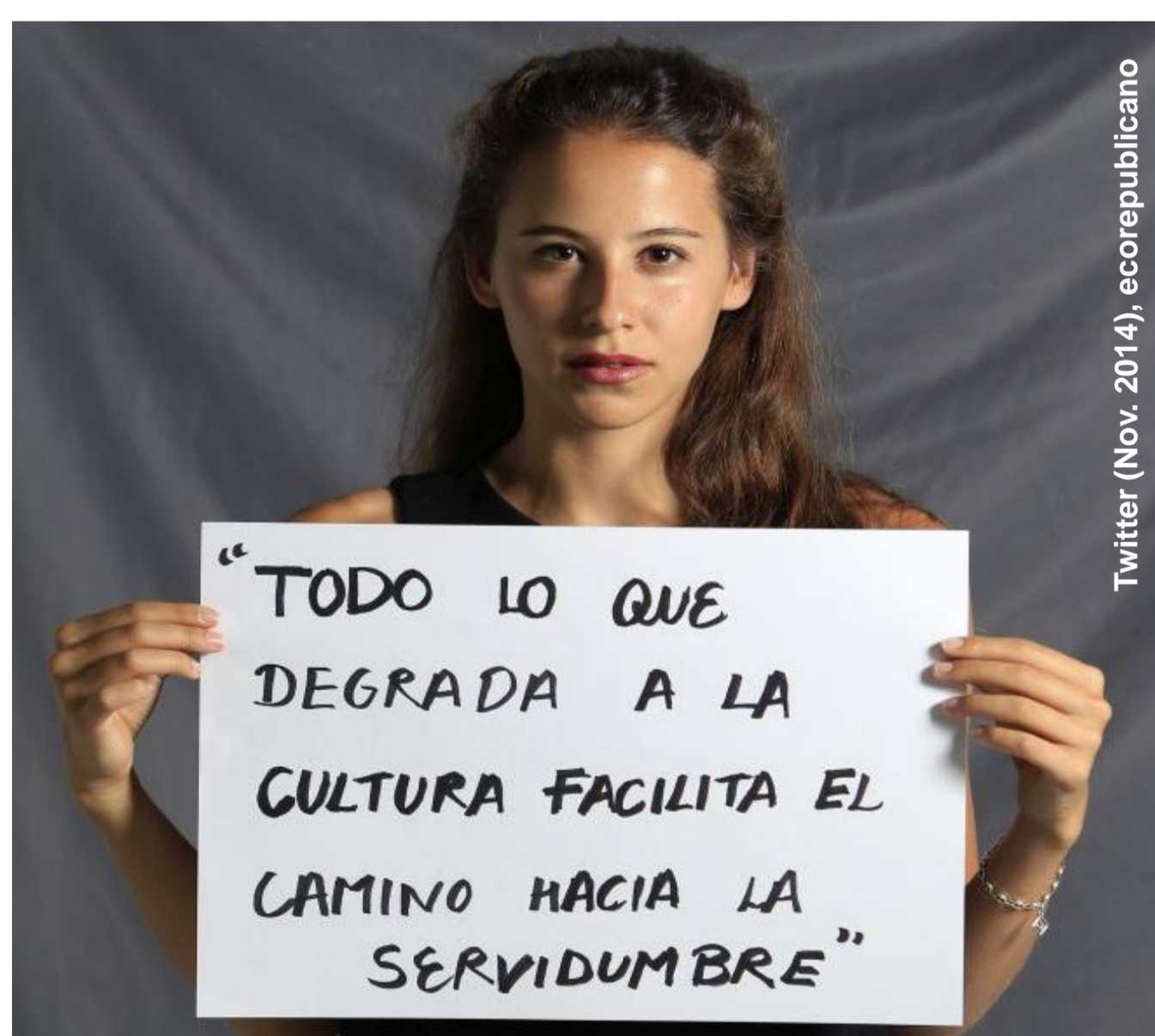
- Differential marking can be either relevant or irrelevant regarding changes in meaning:

A	B	C	D
mandatory	possible + relevant	possible - relevant	impossible

- “Symmetry approach” (argument structure)

primary order	derived (metonymic) order
S _{animate} – V – O _{inanimate}	S _{animate} – V – O _(animate)
S _{animate} – V _{DOM} – O _{animate}	S _{animate} – V _{DOM} – O _(inanimate)
S _{inanimate} – V _(DOM) – O _{inanimate}	

- Verb and noun classes may also play a role in triggering DOM with inanimate objects.



2.2 Research question 2

Existing work points out regional differences between American and Peninsular Spanish:

- Extension in Mexico and Argentina
- Replacement of the marker *a* by *onde* ('where') in Peru.

New tendencies: The scholarly dispute (change hypothesis vs. stability hypothesis) can only be resolved by incorporating new data and new types of data.

Preliminary results suggest a relatively stable system with divergent tendencies in regional standards (concerning areas B and C in the figure above)

Processing: Overgeneralized marking is more difficult to process than omission (Wall 2015)

3. Methods & Research plan

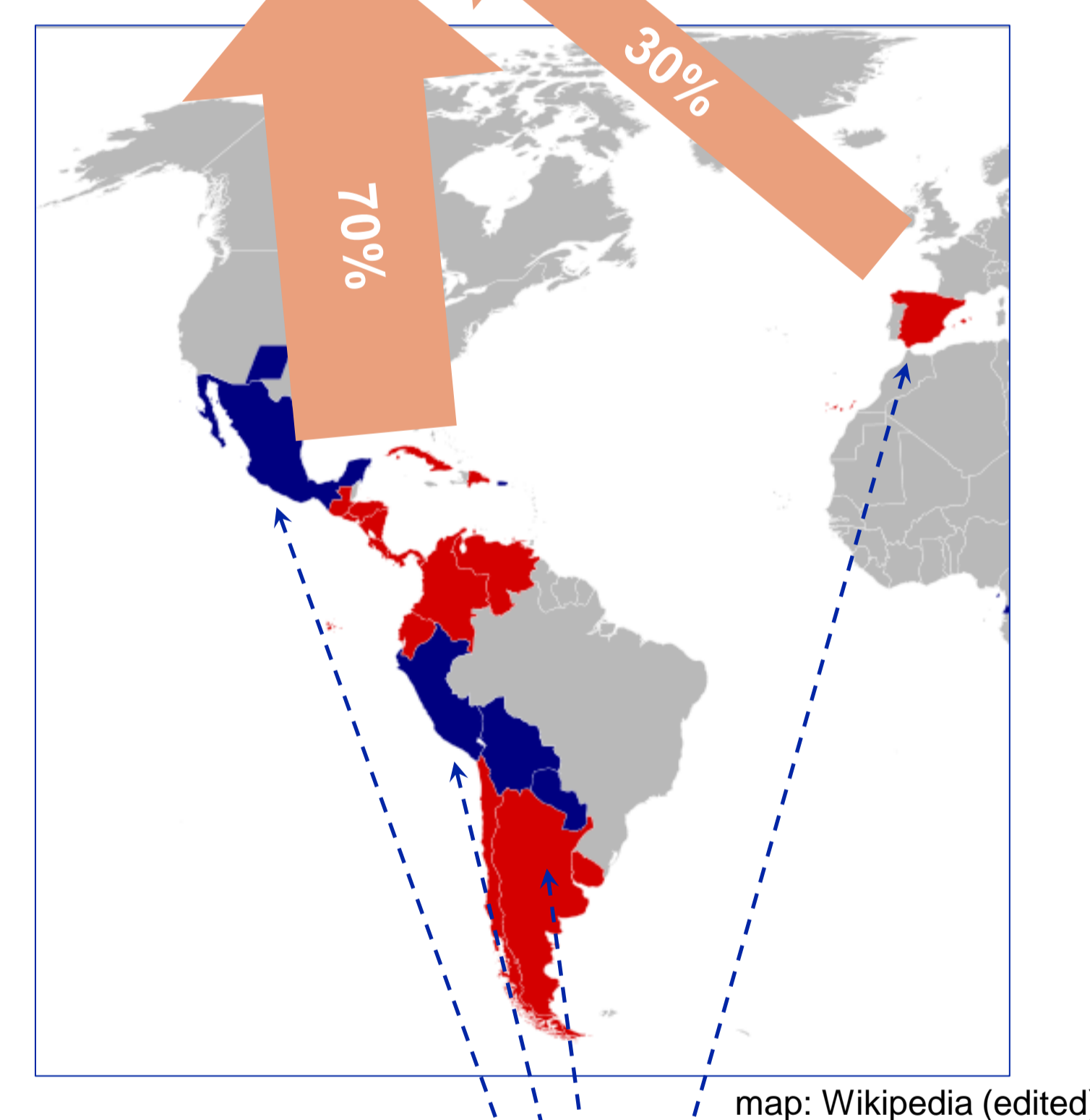
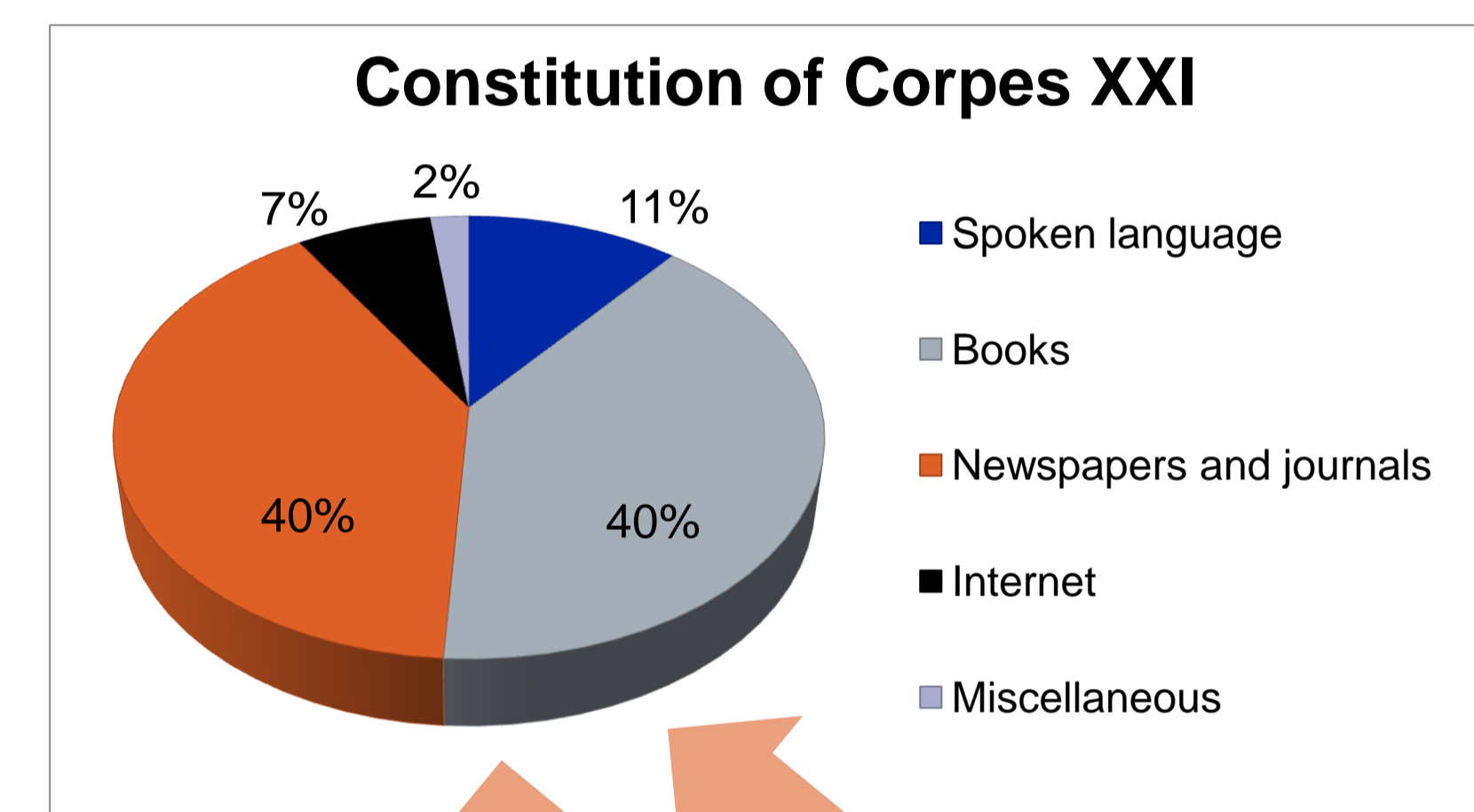
3.1 Diachronic corpus research

- CORDE
- CHARTA
- CORDIAM

- Queries based on verb types, object types and argument structure
- Motion charts

3.2 Synchronic corpus research

- Gigacorpora (EsTenTen, 9,5 billion tokens), other web-based corpora
- CORPES XXI: Comprehensive data, differentiated by country of origin, starting 2001 (25 million tokens per year)



3.3 Experiments

- Combined methods
- Parallel series of equivalent experiments in four relevant areas (acceptability judgements, production, reading time, picture verification)
- Testing of hypotheses emerging from corpus work, processing

Collaborations

- Jun.-Prof. Dr. Álvaro Octavio de Toledo y Huerta (Munich)
- SFB Bedeutungskonstitution (Tübingen)
- Prof. Dr. Klaus von Heusinger, Jun.-Prof. Dr. Marco García García (Cologne)
- Prof. Dr. Concepción Company Company, UNAM, Mexico

Contact data

Prof. Dr. Johannes Kabatek; Romanisches Seminar
kabatek@rom.uzh.ch / <http://www.rose.uzh.ch/de.html>