

# Differential argument marking in online communication games

NATALIA LEVSHINA

# Differential Object Marking (DOM)

Russian (personal knowledge)

- Inanimate masculine nouns: unmarked

<i>Ona</i>	<i>ubra-l-a</i>	<i>zont-Ø.</i>
she.NOM	remove.PFV-PST-F.SG	umbrella.M.SG.ACC
'She removed the umbrella.'		

- Animate masculine nouns: marked

<i>Ona</i>	<i>ubra-l-a</i>	<i>Putin-a.</i>
she.NOM	remove.PFV-PST-F.SG	NAME-M.SG.ACC
'She removed Putin.'		

# Differential Subject Marking (DAM)

Mangarrayi (northern Australia; Merlan 1982: 61)

a. (ergative prefix on inanimate A)

Na-ŋugu	ñim	ŋan-ga-ŋiñ.
N.ERG-water	submerge	3SG>1SG-AUX-PST.PUNCT
“Water covered/submerged me.”		

b. (no ergative marking on animate A)

Buy?	ñan-wu-na	ŋaɭa-bugbug?
show	3SG > 2SG-AUX-PST.PUNCT	F.NOM-old.woman
“Did the old woman show you (to him)?”		

# Popular claims about DM

- Human > Animal > Inanimate
- Definite > (Indefinite) Specific > Non-specific
- Given > New
- Pronoun > Noun
- 1 and 2 Person > 3 Person



NO (LESS) MARKING OF A  
(MORE) MARKING of P

(MORE) MARKING of A  
NO (LESS) MARKING of P

Silverstein (1976), Bossong (1991: 159), Comrie (1986: 94), Croft (2003: 132); Haspelmath (2021);  
but Bickel et al. (2015); Schmidtke-Bode & Levshina (2018); Levshina (2020)

# Actual scales in DOM

Observed 163 times  
in AUTOTYP 0.1

- **Human > Animal > Inanimate**
- **Definite > (Indefinite) Specific > Non-specific**
- **Given > New**
- **Pronoun > Noun**
- 1 and 2 Person > 3 Person



(MORE) MARKING OF P

NO (LESS) MARKING OF P

Silverstein (1976), Bossong (1991: 159), Comrie (1986: 94), Croft (2003: 132); Haspelmath (2021);  
but Bickel et al. (2015); Schmidtke-Bode & Levshina (2018); Levshina (2020)

# Actual scales in DAM

Observed 39 times  
in AUTOTYP 0.1

- Human > Animal > Inanimate
- Definite > (Indefinite) Specific > Non-specific
- Given > New
- **Pronoun > Noun**
- **1 and 2 Person > 3 Person**



NO (LESS) MARKING OF A

(MORE) MARKING OF A

Silverstein (1976), Bossong (1991: 159), Comrie (1986: 94), Croft (2003: 132); Haspelmath (2021);  
but Bickel et al. (2015); Schmidtke-Bode & Levshina (2018); Levshina (2020)

# The puzzle

How to explain the scale effects and the asymmetry of DOM and DAM?

Because of iconicity: formally marked  
objects are also conceptually marked  
(Aissen 2003)

WHY?

But what does 'markedness'  
really mean?  
And how to explain the  
asymmetry?



Because the speaker wants to help the addressee to distinguish between Subject and Object (Comrie 1989, de Hoop and Malchukov 2008)

Because of iconicity: formally marked objects are also conceptually marked (Aissen 2003)

WHY?

Because the speaker wants to help the addressee to distinguish between Subject and Object (Comrie 1989, de Hoop and Malchukov 2008)

But in most cases there is no actual ambiguity! Again, asymmetry?

Because of iconicity: formally marked objects are also conceptually marked (Aissen 2003)

WHY?

Because the speaker wants to help the addressee to distinguish between Subject and Object (Comrie 1989, de Hoop and Malchukov 2008)

Because the speaker wants to help the addressee to identify correctly the role of an NP (A or P) in the presence of weak cues (Levshina 2019, In press)

Because of iconicity: formally marked objects are also conceptually marked (Aissen 2003)

WHY?

Because the speaker wants to help the addressee to distinguish between Subject and Object (Comrie 1989, de Hoop and Malchukov 2008)

Because the speaker wants to help the addressee to identify correctly the role of an NP (A or P) in the presence of weak cues (Levshina 2019, In press)

Because of iconicity: formally marked objects are also conceptually marked (Aissen 2003)

Explains many cases of DOM,  
but not the asymmetry.

WHY?

Because the speaker wants to help the addressee to distinguish between Subject and Object (Comrie 1989, de Hoop and Malchukov 2008)

Because the speaker wants to help the addressee to identify correctly the role of an NP (A or P) in the presence of weak cues (Levshina 2019, In press)

Because of iconicity: formally marked objects are also conceptually marked (Aissen 2003)

WHY?

DOM is popular because a) human and specific/definite objects can be taken for subjects, and b) they make for good patients, they suffer and are salient. DAM is not so popular because disambiguation pressures and salience are in conflict. (de Hoop and Malchukov 2008)

Because the speaker wants to help the addressee to distinguish between Subject and Object (Comrie 1989, de Hoop and Malchukov 2008)

Because the speaker wants to help the addressee to identify correctly the role of an NP (A or P) in the presence of weak cues (Levshina 2019, In press)

Because of iconicity: formally marked objects are also conceptually marked (Aissen 2003)

WHY?

DOM is popular because a) human and specific/definite objects can be taken for subjects, and b) they make for good patients, they suffer and are salient. DAM is not so popular because disambiguation pressures and salience are in conflict. (de Hoop and Malchukov 2008)

Nicely explains the quantitative asymmetry nicely, but not the qualitative differences.

Because the speaker wants to help the addressee to distinguish between Subject and Object (Comrie 1989, de Hoop and Malchukov 2008)

Because the speaker wants to help the addressee to identify correctly the role of an NP (A or P) in the presence of weak cues (Levshina 2019, In press)

Because of iconicity: formally marked objects are also conceptually marked (Aissen 2003)

WHY?

DOM constructions originate from other source constructions which involved marking on animate, definite, etc. NP, e.g. from topical constructions, especially in atypical sentence positions. Later, they spread to in-situ objects (Iemmolo 2010, Diessel 2019, Tal et al. In press, cf. Cristofaro 2019)

DOM is popular because a) human and specific/definite objects can be taken for subjects, and b) they make for good patients, they suffer and are salient. DAM is not so popular because disambiguation pressures and salience are in conflict. (de Hoop and Malchukov 2008)

Because the speaker wants to help the addressee to distinguish between Subject and Object (Comrie 1989, de Hoop and Malchukov 2008)

Because the speaker wants to help the addressee to identify correctly the role of an NP (A or P) in the presence of weak cues (Levshina 2019, In press)

Because of iconicity: formally marked objects are also conceptually marked (Aissen 2003)

WHY?

DOM constructions originate from marking of topicalized objects, especially in atypical sentence positions. Later, they spread to in-situ objects (Iemmolo 2010, Diessel 2019, Tal et al. In press, cf. Cristofaro 2019)

DOM is popular because a) human and specific/definite objects can be taken for subjects, and b) they make for good patients, they suffer and are salient. DAM is not so popular because disambiguation pressures and salience are in conflict. (de Hoop and Malchukov 2008)

What about DAM?  
This does not exclude the importance of predictability.



Because the speaker wants to help the addressee to distinguish between Subject and Object (Comrie 1989, de Hoop and Malchukov 2008)

Because the speaker wants to help the addressee to identify correctly the role of an NP (A or P) in the presence of weak cues (Levshina 2019, In press)

Because of iconicity: formally marked objects are also conceptually marked (Aissen 2003)

WHY?

DOM is popular because a) human and specific/definite objects can be taken for subjects, and b) they make for good patients, they suffer and are salient. DAM is not so popular because disambiguation pressures and salience are in conflict. (de Hoop and Malchukov 2008)

DOM constructions originate from marking of topicalized objects, especially in atypical sentence positions. Later, they spread to in-situ objects (Iemmolo 2010, Diessel 2019, Tal et al. In press, cf. Cristofaro 2019)

Pronouns lag behind nouns in language change, such as change from ergative to accusative alignment (Filimonova 2005)

Because the speaker wants to help the addressee to distinguish between Subject and Object (Comrie 1989, de Hoop and Malchukov 2008)

Because the speaker wants to help the addressee to identify correctly the role of an NP (A or P) in the presence of weak cues (Levshina 2019, In press)

Because of iconicity: formally marked objects are also conceptually marked (Aissen 2003)

WHY?

DOM is popular because a) human and specific/definite objects can be taken for subjects, and b) they make for good patients, they suffer and are salient. DAM is not so popular because disambiguation pressures and salience are in conflict. (de Hoop and Malchukov 2008)

DOM constructions originate from marking of topicalized objects, especially in atypical sentence positions. Later, they spread to in-situ objects (Iemmolo 2010, Diessel 2019, Tal et al. In press, cf. Cristofaro 2019)

Pronouns lag behind nouns in language change, such as change from ergative to accusative alignment (Filimonova 2005)

But what about other scales?

Because the speaker wants to help the addressee to distinguish between Subject and Object (Comrie 1989, de Hoop and Malchukov 2008)

Because the speaker wants to help the addressee to identify correctly the role of an NP (A or P) in the presence of weak cues (Levshina 2019, In press)

Because of iconicity: formally marked objects are also conceptually marked (Aissen 2003)

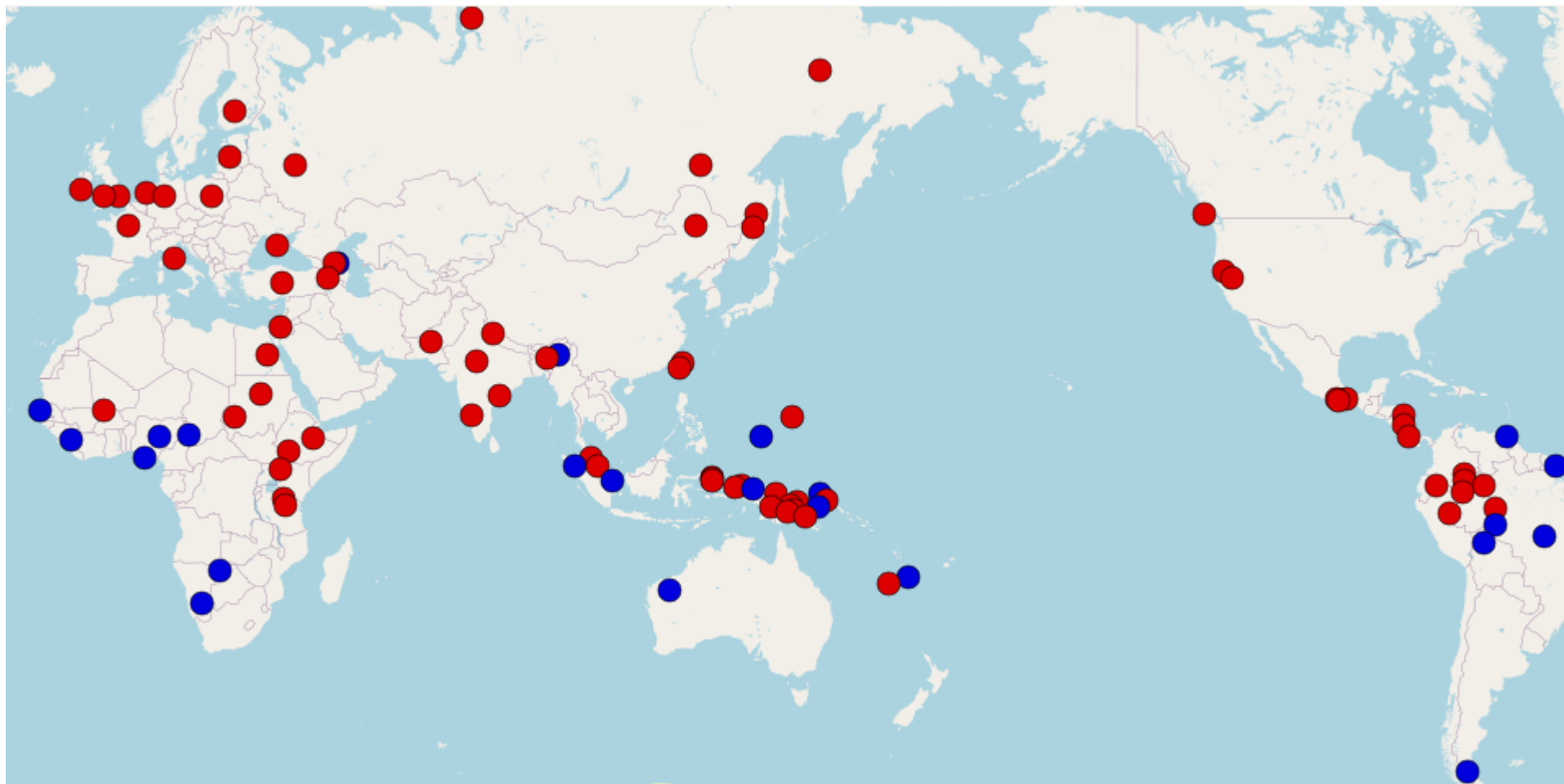
WHY?

DOM is popular because a) human and specific/definite objects can be taken for subjects, and b) they make for good patients, they suffer and are salient. DAM is not so popular because disambiguation pressures and salience are in conflict. (de Hoop and Malchukov 2008)

DOM constructions originate from marking of topicalized objects, especially in atypical sentence positions. Later, they spread to in-situ objects (Iemmolo 2010, Diessel 2019, Tal et al. In press, cf. Cristofaro 2019)

Pronouns lag behind nouns in language change, such as change from ergative to accusative alignment (Filimonova 2005)

Less need to mark subjects due to higher frequency of subject agreement than object agreement?



Subject agreement (red) vs. Object agreement (blue), based on  
Siewierska (2005) <https://wals.info/feature/102A#2/13.9/116.5>

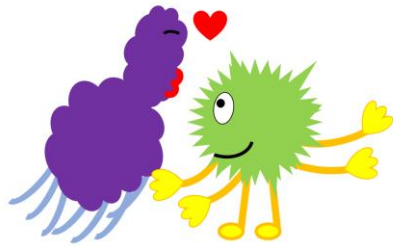
# Artificial Language Learning and DOM

- Fedzechkina et al. (2012):
  - a miniature language with optional object case marking
  - Subjects: always human, 50% of objects human and 50% inanimate.
  - Learners used significantly more case markers on atypical (animate) objects than on typical (inanimate) objects.
- Smith & Culbertson (2020) could not reproduce Fedzechkina et al's (2012) results.
  - However, when they ran an interactive experiment (a director-matching task), the sentences produced by the participants showed the expected DOM pattern.
  - Conclusion: "good evidence that participants' behaviour in actual communicative interaction [is] driven by efficient communication considerations": atypical meanings are expressed by atypical forms.
- This potentially supports many accounts: iconicity, disambiguation, predictability and salience...
- Tal et al. (2020) found no direct effect of topicality/givenness of animate objects on the chances of marking (but effect of atypical word order). But no communication was involved.
- No conclusive evidence so far...

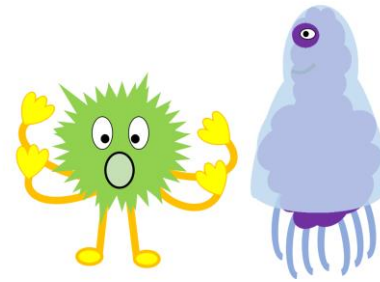
# Online communication game

- An artificial language with optional object marking (50% marked), some word order variation (SOV – OSV)
- Training phase: copying sentences in the input language
- Production phase: Director-matcher task with pictures. The goal is to earn as many points (correct matches) as possible.
- Two variants: with DOM and DAM.
- Many thanks to **Maarten van den Heuven** for programming the experiment in Frinex, and to **Iris Schmits and Marjolijn Dijkhuis** for technical support. The pictures were drawn by Iris.

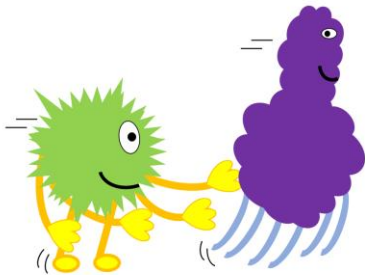
# Examples of stimuli



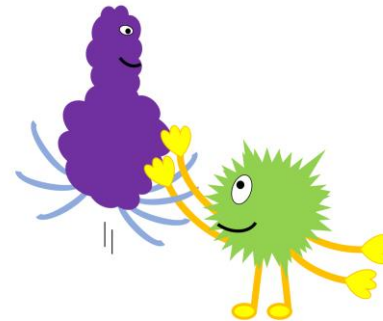
Tinga mgana mumi



Tinga mgana-ke uuhi



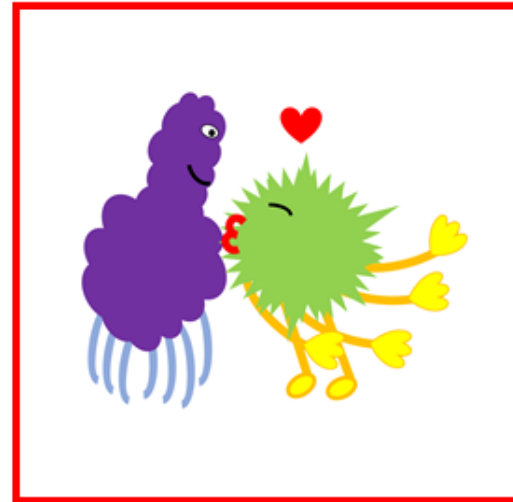
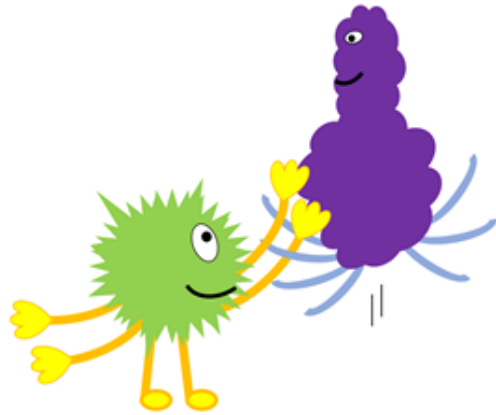
Mgana tinga-ke teki



Mgana tinga hopi

# Predictions

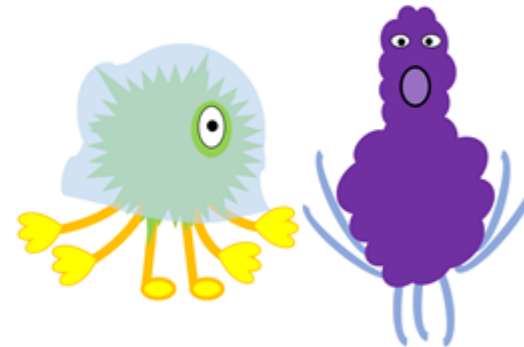
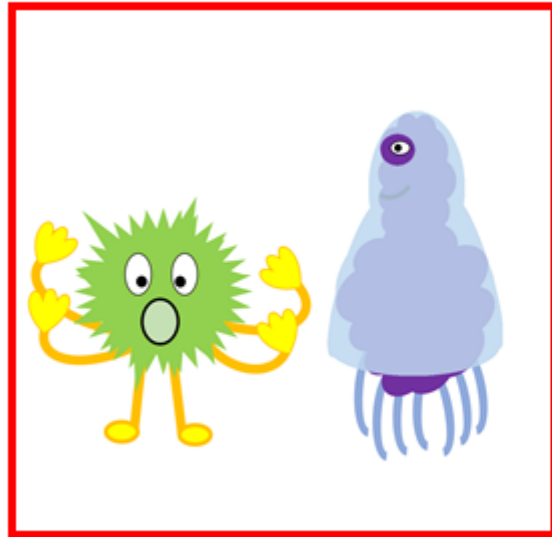
Same Agent, same Patient, different actions:  
The marker is used **the least** frequently





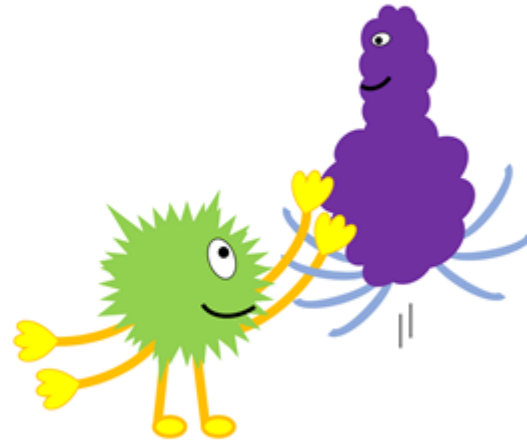
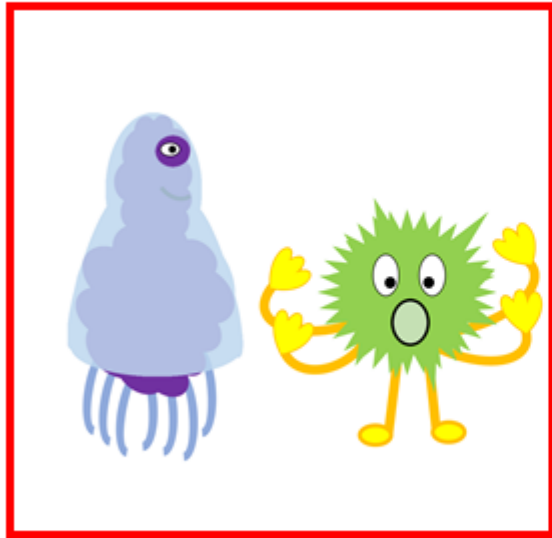
# Predictions

Different Agent and Patient, same actions:  
The marker is used **the most** frequently



# Predictions

Different Agent and Patient, different actions:  
in-between

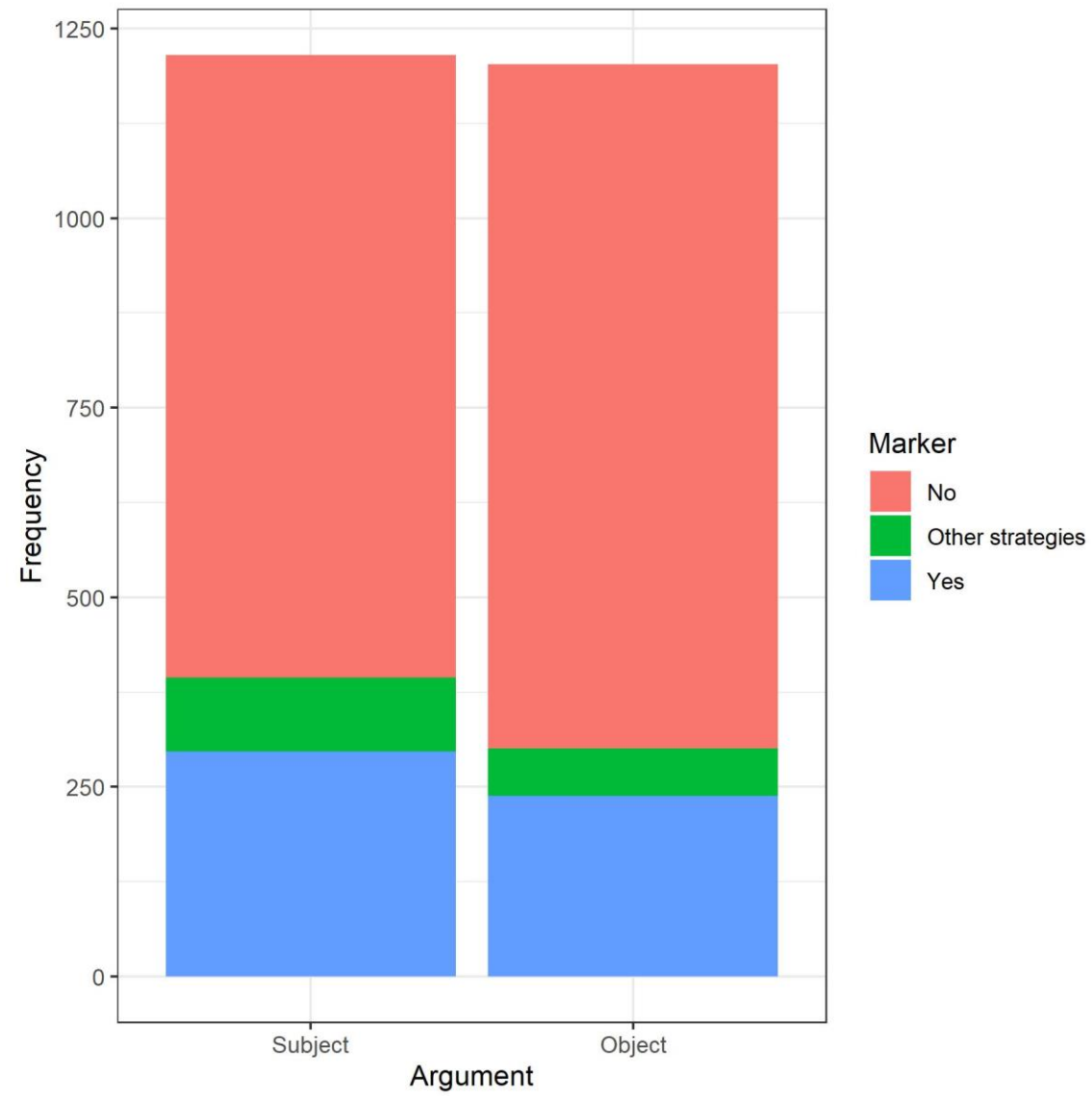


# Data

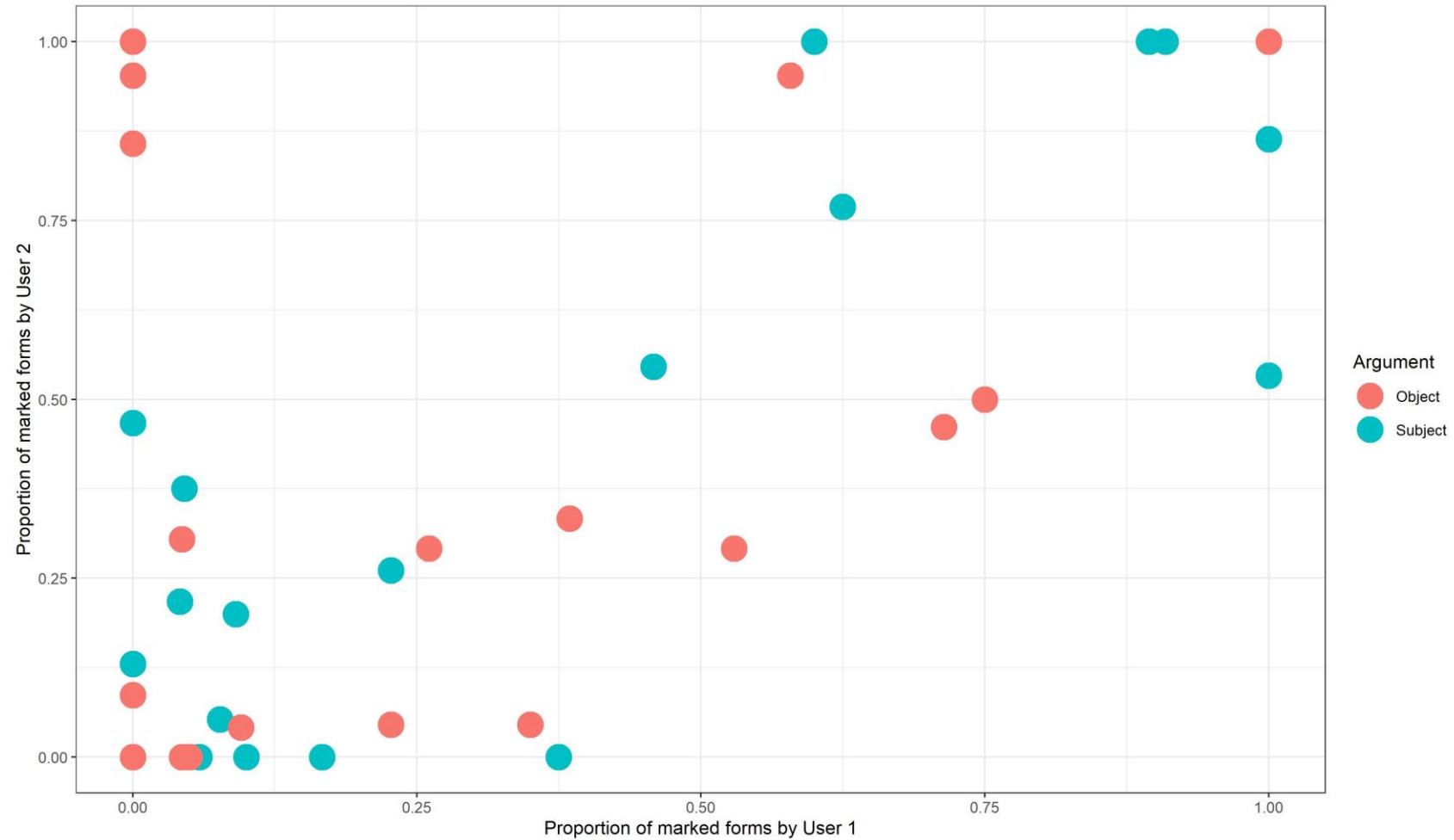
- 28 pairs for DAM, 27 pairs for DOM
- Approximately 2,500 sentences
- Less than 5% errors (incorrect guesses)
- Most participants understood that the experiment has something to do with subject/object marking.
- Excluded: sentences without two arguments, with English or Dutch verbs, with a marker on the wrong argument.

# Variables

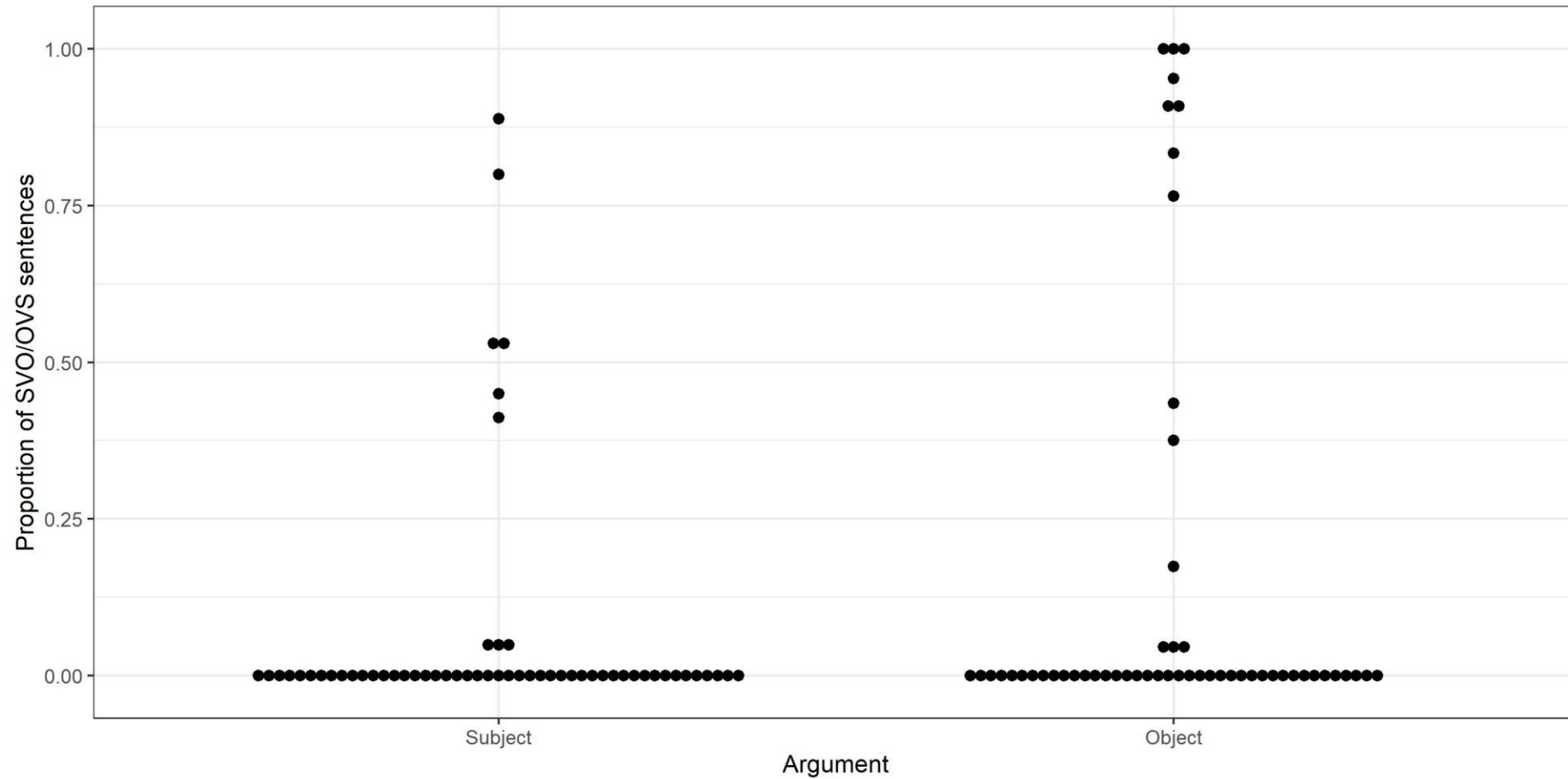
- Response: marked or unmarked relevant argument (subject or object)
- Predictors:
  - Argument (Subject or Object)
  - Stimulus Type (Different A and P, Different Actions, all different)
  - Trial
  - Verb position (verb-final vs. verb-medial) in the output.
  - Who is the agent?
  - Does the previous response (by the other participant) contain the marker?



# Alignment of participants re. marking

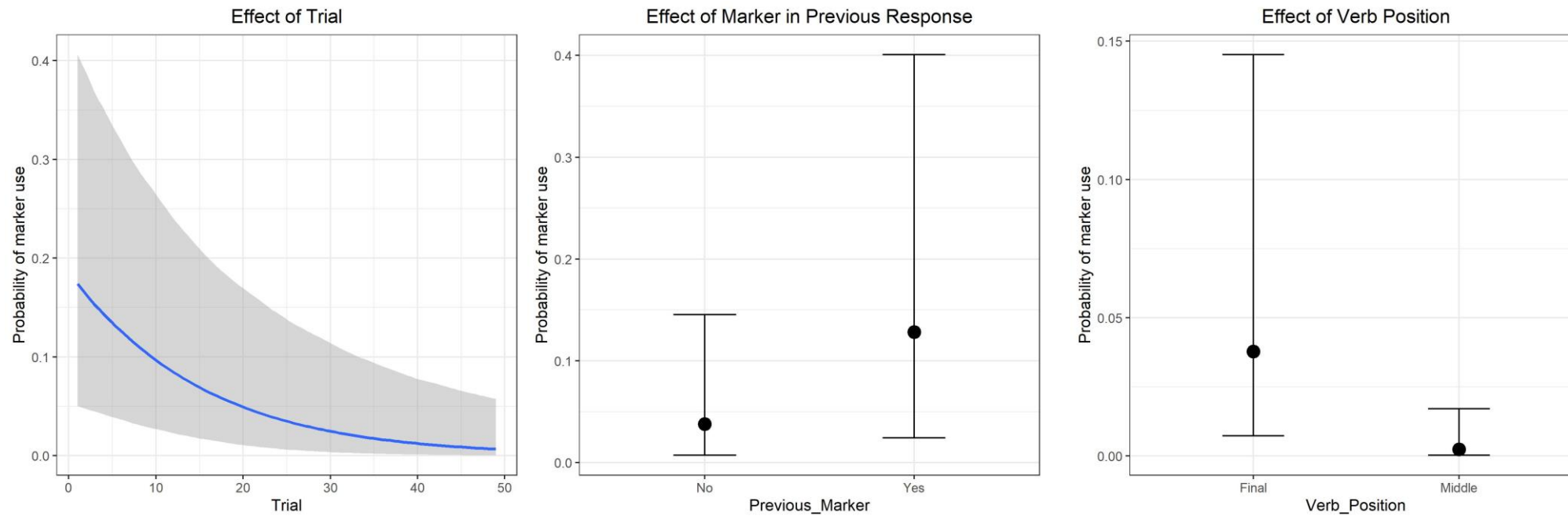


# Proportion of verb-medial sentences



# Bayesian regression results

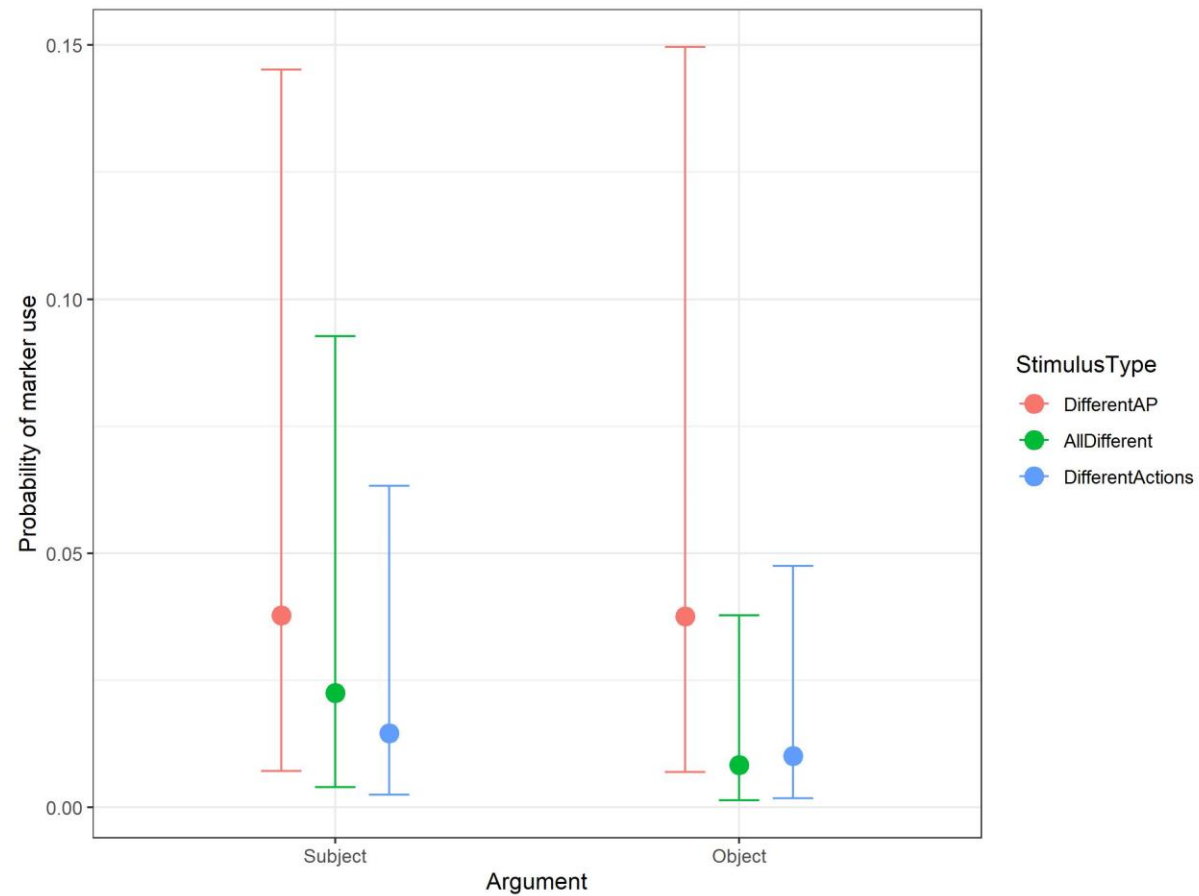
```
Marker ~ (1 + Trial | UserID) + (1 | StimulusID) + Argument + Trial +  
Previous_Marker + StimulusType + Verb_Position + Agent +  
Argument:StimulusType
```



Bayes pseudo- $R^2 = 0.70$  (0.68, 0.72)



# Interaction between Stimulus Type and Argument



# Summary

- The proportions of marked arguments in the users' output were much lower than in the input language. This is different from the previous DOM experiments. Probably, due to the shorter learning phase.
- In many pairs, the individual marking rates were similar – a sign of mutual alignment and emerging conventions.
- As the game went on, they used marking less often, other factors controlled for. Possibly, the effect of learning gradually vanished.
- There was a priming effect of the marker in the previous sentence.
- Participants sometimes used SVO sentences without any case marking – probably, influence of the default order in Dutch or English.
- As predicted, some of the participants used case marking more frequently in ambiguous situations where only A and P were different than in unambiguous ones.
- The marking rates were similar for Subject and Object, although the role of the number of visual cues in unambiguous contexts (actions only or actions and A and P) was slightly different.
- The marking was differential for 39% of users for Subject, and 56% of users for Object. But this difference is not significant. More data are needed to say anything about the DAM/DOM asymmetry.

# Discussion

- The aliens did not differ in animacy, predictability of A/P roles or information status. There were no agreement markers. The disambiguating function is the only factor playing a role here.
- Bickel et al. (2015) argued for a processing bias to process the first base-form noun phrase as the agent (which according to them explains why ergative languages are rare). We could expect then that subject marking would be used less often than object marking. However, there were no indications of such an asymmetry.
- We should focus on other factors (information structure, word order, frequency, agreement, ...) to explain the asymmetry.