1 Test Sentence Templates

1.1 Eliciting nouns given type of marking

1.1.1 CD + DOM-p marking

(1) Am văzut-o pe [MASK].
have seen-CD$_{3SG.FEM}$DOM-p [MASK].
'I/We have seen [MASK].'

(2) L-am văzut pe [MASK].
CD$_{3SG.MASC}$have seen DOM-p [MASK].
'I/We have seen [MASK].'

(3) Pe [MASK] l-am
DOM-p [MASK] CD$_{3SG.MASC}-have$
văzut.
seen
'[MASK], I/we have seen.'

(4) Am văzut-o pe [MASK]
have seen-CD$_{3SG.FEM}$DOM-p [MASK]
azi, dar pe Ileana, nu.
today but DOM-p Ileana not
'I/We have seen DOM-p [MASK] today, but Ileana I/we have not.'

(5) Pe [MASK] am văzut-o
DOM-p [MASK] have seen-CD$_{3SG.FEM}$
azi, dar pe Ileana, nu.
today but DOM-p Ileana not
'[NAME] I/We have seen today, but Ileana I/We have not.'

1.1.2 DOM-p marking

(6) Am văzut pe [MASK].
have seen DOM-p [MASK]
'I/We have seen [MASK].'

1.2 Eliciting types of marking given nouns

1.2.1 CD + DOM-p is obligatory

Proper nouns

(7) Am văzut-o [MASK] [NAME].
have seen-CD$_{3SG.FEM}$ [MASK] [NAME].
'I/We have seen [NAME].'

(8) Nu am văzut [MASK] azi [MASK]
not have seen [MASK] today [MASK]
[NAME], dar pe Ileana, da.
[NAME] but DOM-p Ileana yes
'I/We haven't seen [NAME] today, but I/We've seen Ileana.'

(9) [MASK] [NAME] am văzut [MASK].
[MASK] [NAME] have seen [MASK]
'[NAME], I/We have seen. (Someone else, I/We have not.)'

(10) [MASK] [NAME] am văzut [NAME]
[MASK] [NAME] have seen [MASK]
azi, dar pe Ileana, nu.
today but DOM-p Ileana not
'[NAME] I/We have seen today, but Ileana I/We have not.'

Personal pronouns

(11) Am văzut [MASK] [MASK] [PP].
have seen [MASK] [MASK] [PP]
'I/We have seen [PP].'

(12) Nu [MASK] am văzut azi [MASK] [PP]
not [MASK] have seen today [MASK] [PP]
azi, dar pe Ileana, da.
but DOM-p Ileana yes
'I/We haven't seen [PP] today, but I/we have seen Ileana.'

(13) [MASK] [PP] am văzut [MASK].
[MASK] [PP] have seen [MASK]
'[PP], I/we have seen.'

(14) [MASK] [PP] [MASK] am văzut azi, dar
[MASK] [PP] [MASK] have seen today but
pe Ileana, nu.
DOM-p Ileana not
'I/We haven't seen [PP] today, but I/We have seen Ileana.'

1.2.2 CD + DOM-p is optional

(15) [MASK] am văzut [MASK] [NOUN]
[MASK] have seen [MASK] [NOUN]
[MOD].
[MOD]
'I/We have seen [MASK] [NOUN] [MOD].'

1.2.3 CD + DOM-p is ungrammatical

Definite article *pe*-marking of objects is syntactically constrained by the presence of a definite article in the NP.

(16) [MASK] am văzut [MASK]
[MASK] have seen [MASK]
[NOUN$_{DEF}$].
[NOUN$_{DEF}$]

'I/We saw [NOUN$_{DEF}$].'

(17) [MASK] [NOUN + DEF.ART.] [MASK]
[MASK] [NOUN + DEF.ART.] [MASK]
am văzut.
have seen
'(It is) [NOUN + DEF.ART.] I/we have seen.'

Inanimates

(18) Am văzut [MASK] [MASK]
have seen [MASK] [MASK]
[NOUN$_{inanimate}$].
[NOUN$_{inanimate}$].
'I/We saw [MASK] [NOUN$_{inanimate}$].'

(19) Nu am vizitat orașul, dar am văzut
not have visited city-the but have seen
[MASK] [MASK] [NOUN$_{inanimate}$].
[MASK] [MASK] [NOUN$_{inanimate}$]
'I/We have not visited the city, but I/we have seen [NOUN$_{inanimate}$].'

1.2.4 DOM-p is obligatory

DOM-p object marking is restricted to nouns with a [+human] referent, thus being ungrammatical with non-human animates and with inanimates.

(20) Am văzut [MASK] [QUANT.].
have seen [MASK] [QUANT.]
'I/We saw [QUANT.].'

1.2.5 DOM-p is ungrammatical

Definite and specific inanimates

(21) Am văzut [MASK] [NOUN + DEF.ART.].
have seen [MASK] [NOUN + DEF.ART.]
'I/We saw [NOUN + DEF.ART.].'

Indefinite and specific inanimates

(22) Am văzut [MASK] [NOUN +
have seen [MASK] [NOUN +
INDEF.ART.].
INDEF.ART.]
'I/We saw [NOUN + INDEF.ART.].'

1.3 Follow-up task: Eliciting types of marking given nouns

The following templates generated five test sentences each for our follow-up task, where we introduced an additional [MASK] token to account for the '-' symbol. This is a different test set than for the other marker prediction task, since a larger context is needed to balance out the masked-to-given text proportion. This way, we avoid having a too high mask rate – this is why the 'accuracy before' scores without the extra mask token are higher than for the other task in Appendix 1.2.

(23) În drum spre teatru, [MASK] [MASK] am
in way to theatre [MASK] [MASK] have
văzut [MASK] [NAME] pe stradă și
seen [MASK] [NAME] on street and
părea foarte grăbit.
seemed very hurried
'On my way to the theatre I saw [NAME] and he/she seemed to be in a hurry.'

(24) Nu am văzut [MASK] [MASK] azi
not have seen [MASK] [MASK] today
[MASK] [NAME] pe stradă, dar am
[MASK] [NAME] on street, but have
văzut-o pe Ileana si părea foarte
seen Ileana and seemed very hurried
grăbită.

'I did not see [NAME] on the street today, but I have seen Ileana and she seemed to be in a hurry.'

(25) [MASK] [NAME] nu am văzut [MASK]
[MASK] [NAME] not have seen [MASK]
[MAKS] azi pe stradă, dar am văzut-o
[MASK] today on street but have seen
pe Ileana și părea foarte grăbită.
Ileana and seemed very hurried
'[NAME] I/we did not see today on the street, but I/we saw Ileana and she seemed to be in a hurry.'

(26) [MASK] [MASK] am văzut [MASK] [PP]
[MASK] [MASK] have seen [MASK] [PP]
azi, dar nu și pe fratele lui/ei.
today, but not and *pe* brother his/hers
'[PP] I/we saw today, but not his/her brother.'

(27) Când [MASK] [MASK] am văzut
when [MASK] [MASK] have seen
[MASK] [NOUN + MODIFIER] în
[MASK] [NOUN + MODIFIER] in
revistă, am fost tare surprinsă.
magazine I was very surprised
'When I saw [NOUN + MODIFIER] in the magazine today, I was really surprised.'

Model	Version	Accuracy
mBERT	uncased	0%
	cased	5%
RoBERT	small	23.75%
	base	13.75%
	large	**25%**
Romanian BERT	uncased	21.25%
	cased	26.25%

Table 1: Accuracy scores per model on predicting the case marker and the clitic, when the noun is given.

2 Effect of elicitation strategy

To test the limits of MLM tasks for grammatical inquiry, we also explore the alternate approach of masking the grammatical marking for Romanian DOM. While this is a more direct test for this grammatical phenomenon, it does come at the cost of higher constraints, a reduced pool of suitable candidates (the accusative marker *pe*, on one hand, or the clitic paradigm on the other hand) to predict from, as well as increased task difficulty from having to mask a high proportion of the tokens in the sentence, due to the multiple morpheme realisation.

For this, we needed 16 templates, shown in Appendix 1.2. We manually generate five test sentences for each, setting the sample size for this task at n = 80. It should be noted that sentences were labeled as a whole, as opposed to the predictions for each [MASK] token being judged and labeled independently. This is especially relevant for these sentences as they contain multiple [MASK] tokens, where all masked positions had to be filled with tokens that yield a grammatical and semantically valid sentence in order for the sample (i.e. the sentence) to be labeled as 'correct'.

The results presented in Table 1 show that this approach was much less successful, with accuracy scores going as low as 0% for mBERT-uncased and no higher than 26.3% overall across models, indicating a failure to elicit grammatical sentences altogether. The process of manual annotation revealed that this might be due to tokenization issues, in particular related to the hyphen symbol ('-') involved in the orthography of clitic doubling. Many outputs contained this symbol alone in positions where a clitic was expected. We performed a follow-up task and found that accuracy scores drastically increased with an additional [MASK] token to account for both the clitic and the hyphen. This result is shown in Table 2. The templates for this task are shown in Appendix 1.3.

Model	Version	Before	After
mBERT	uncased	0%	22%
	cased	0%	12%
RoBERT	small	0%	40%
	base	0%	80%
	large	**12%**	92%
Romanian BERT	uncased	0%	88%
	cased	4%	88%

Table 2: Accuracy scores of each model version in the follow-up task, before and after introducing an additional [MASK] token in the '-' position.