Modifier attachment problem: a processing experiment on Lithuanian

The problem of modifier attachment ambiguity has been extensively studied — primarily due to noteworthy cross-linguistic variation: some languages prefer high attachment (HA, attachment to the head noun, N_H), the others prefer low attachment (LA, attachment to the dependent noun, N_D), while the third group does not show a clear preference (e.g. Frazier, Fodor, 1978; Cuetos, Mitchell, 1988; Grillo, Costa, 2014). In this study, we look at the processing of complex NPs with clausal modifiers in Lithuanian, analyzing sentences with HA and LA disambiguated by gender or number agreement, and ambiguous constructions (AMB). Lithuanian is an interesting case because, being a head-initial language, it puts genitive dependent nouns before the head nouns. We examined relative clauses (RCs) following complex NPs and participial constructions (PrtCs) preceding them.

Method. We constructed 24 target sentence sets, as in (1a-c) and (2a-c). In every set, there were sentences in three conditions (AMB, HA and LA). Half of the sentences included an RC (N_{SUBJ} V (N_D N_H)$_{OBJ}$ $RC_{D/H}$), the other half — a PrtC (N_{SUBJ} V $PrtC_{D/H}$ (N_D N_H)$_{OBJ}$). N_D and N_H were matched for animacy, RCs and PrtCs were matched for length (in the number of syllables). 9 out of 24 NPs had the part-whole relationship between the nouns, and 6 out of 24 had kinship terms. We ran a pretest with a separate group of Lithuanian speakers checking that target sentences in the AMB condition were rated as plausible in both interpretations. Lithuanian nouns are inflected for case and number and belong to the M or F gender. Participles and heads of RCs (*kuris* 'who, that') agree with the noun they modify in number and gender. This was used to disambiguate HA and LA sentences. Participles also agree in case, so in the PrtC sentences, the verbs required a genitive object, and N_D, N_H and the participle were in genitive case. There were also 48 filler sentences.

We collected online data (word-by-word reading times using the self-paced reading methodology) and offline data (answers to comprehension questions). In the target sentences, the answer indicated where the modifier was attached (i.e. whether unambiguous sentences were understood correctly as HA or LA and how ambiguous sentences were interpreted). The experiment was conducted on spellout.net/ibexfarm/. 49 speakers of Lithuanian (aged 19-63, SD=14,1) participated in it.

Results and discussion. Mixed-effect models were used for the statistical analysis (with a condition — AMB, HA, LA — as a predictor). Offline results are presented in Table 1. In AMB sentences, HA was chosen significantly more often than LA both in PrtCs and in RCs. In LA and HA sentences, there were significantly more errors in the former than in the latter both in PrtCs and in RCs. As for online results, there were no significant differences in PrtC conditions, while in the RC conditions, LA sentences were processed significantly more slowly than HA and AMB ones (see Fig. 1).

Thus, independently from the modifier position, a strong HA preference was found in offline measures. In particular, 48% of PrtCs and 35% of RCs with LA were interpreted incorrectly (similar results were obtained for unambiguous LA sentences in Russian, see Chernova et al. (2016a,b)). When the PrtC modifier precedes the complex NP, LA and HA are equally difficult to process online — presumably because in the LA sentences (which could be more difficult to read because this is not the preferred interpretation), participial constructions are immediately before the N_D. When the RC modifier follows the complex NP, LA sentences are more difficult to process — presumably because the modifier is separated from the N_D. These results might be interesting in a cross-linguistic perspective. Other languages where dependent nouns precede head nouns, which were analyzed in the modifier attachment ambiguity debate — for example, Chinese, Basque and Swedish (Shen, 2006; Leeser, Prieta, 2015; Ehrlich et al., 1999) — showed an LA preference.

(1) RCs: a. Ambiguous sentence (**AMB condition**):

*Vyras taip ir nepataisė komod-**os** dėž-**ės**, kur-**i***
husband still not.fixed dresser$_{GEN.SG.F}$ drawer$_{GEN.SG.F}$ that$_{NOM.SG.F}$
taip erzino jo žmoną.
so irritated his wife.
'The husband still has not fixed the drawer$_i$ of the dresser$_j$ that$_{i/j}$ irritated his wife so much.'

b. Unambiguous sentence with HA (**HA condition**):

*Vyras taip ir nepataisė stal-o dėž-**ės**, kur-**i**…*
husband still not.fixed desk$_{GEN.SG.M}$ drawer$_{GEN.SG.F}$ that$_{NOM.SG.F}$
'The husband still has not fixed the drawer$_i$ of the desk$_j$ that$_i$...'

c. Unambiguous sentence with LA (**LA condition**):

*Vyras taip ir nepataisė stal-**o** dėž-ės, kur-**is**…*
husband still not.fixed desk$_{GEN.SG.M}$ drawer$_{GEN.SG.F}$ that$_{NOM.SG.F}$
'The husband still has not fixed the drawer$_i$ of the desk$_j$ that$_j$...'

(2) PrtCs: a. AMB:

*Petras neatpažino [Šiaurę seniai persikėl-usi-**ų**] kaimyn-ų vaik-ų.*
Petras not.recognized to North long.ago move$_{PST.PA-GEN.PL}$ neighbor$_{GEN.PL}$ child$_{GEN.PL}$
'Petras did not recognize the children$_i$ of the neighbors$_j$ who$_{i/j}$ moved up North long ago.'

b. HA:

*Petras neatpažino [Šiaurę seniai persikėl-usi-**o**] kaimyn-ų vaik-**o**.*
Petras not.recognized to North long.ago move$_{PST.PA-GEN.SG.M}$ neighbor$_{GEN.PL}$ child$_{GEN.SG}$
'Petras did not recognize the child$_i$ of the neighbors$_j$ who$_i$ moved…'

c. LA:

*Petras neatpažino [Šiaurę seniai persikėl-usi-**ų**] kaimyn-ų vaik-**o**.*
Petras not.recognized to North long.ago move$_{PST.PA-GEN.PL}$ neighbor$_{GEN.PL}$ child$_{GEN.SG}$
'Petras did not recognize the child$_i$ of the neighbors$_j$ who$_j$ moved…'

Table 1. Interpreting target sentences.

	Relative clauses				Participial constructions			
	HA answers		LA answers		HA answers		LA answers	
AMB sentences	175	89%	21	11%	141	72%	55	28%
HA sentences	183	93%	13	7%	170	87%	26	13%
LA sentences	69	35%	127	65%	94	48%	102	52%

Figure 1. Average word-by-word RTs (in ms) for RC target sentences.