The time course of processing cataphora in a *pro*-drop language: The case of Slovenian

Arthur Stepanov (University of Nova Gorica) & Matic Pavlič (University of Ljubljana)

Previous research on processing cataphora (backwards anaphora) fruitfully explored the idea that the pronoun triggers a forward-looking active search for an "antecedent" in order to establish a coreference relation, similarly to a fronted wh-phrase and its base-generated syntactic position in filler-gap dependencies ([1-2]). In processing terms, coreference is usually understood as a relation between elements in a mental model instantiated by correspondent lexical cues and based, among other, on congruence of morphological features such as gender and number. [3] showed that, because of the active search, a cataphoric coreference relation with a (forward) antecedent is established prior to making use of the gender and number information on the pronoun. As a result, sentences with feature-incongruent incoming NPs as "antecedents" (as in *When **she**$_f$ was fed up, **the boy**$_m$ visited **the girl**$_f$ very often*) were more difficult to read than their congruent counterparts (cf. *When **he**$_m$ was fed up, **the boy**$_m$ visited **the girl**$_f$ very often*), as revealed in reading time slow-downs and eye-movement regressions. Pro-drop languages in which pronoun features are computed based only on the verbal inflection raise a host of questions for the active search hypothesis that have not been addressed so far in the literature: i) how is a cataphoric coreference established online in the absence of an overt pronoun cue? ii) does *pro* initiate an active search similarly to an overt pronoun? iii) does *pro*, similarly to a pronoun, trigger a reference set specifically targeting an incoming NP as a target antecedent? A key issue in addressing these matters is the timing of computing *pro* when processing a cataphora sentence. If, due to absence of an overt pronoun, computation of *pro* is delayed beyond the first potential 'antecedent', we might expect the absence of an incongruence effect. On the other hand, such effect is expected if the relevant features of *pro* is computed early enough similarly to the overt pronoun. Slovenian, a pro-drop language, presents an excellent opportunity to tease apart these possibilities.

Experiment We conducted a self-paced reading experiment with the aim to determine whether a grammatical number-incongruence effect obtains with subject *pro* in Slovenian. We also asked whether this effect is sensitive to the actual number feature of *pro*. Materials (see (1)) included 24 sentence sets of 4 items each organized in a 2 x 2 design crossing factors ±Congruency and Number (sg/pl) on the first auxiliary, as well as 48 filler sentences, each followed by a comprehension question. Participants were 33 native speakers of Slovenian (21 female; *M* = 36.3, *SD*=14) who took the experiment on the Ibex Farm platform.

Results The time course of reading sentences in all four conditions is depicted in Figure 1. Overall reading times were higher in the non-congruent conditions (a singular auxiliary followed by a plural NP (Npl) and a plural auxiliary followed by a singular NP (Nsg)) than in the congruent conditions (Table 1). Furthermore, congruency interacted with number: there was no difference in reading times between sentences with singular and plural auxiliaries in the congruent conditions, but sentences with singular auxiliaries were read slower (about 40 ms per word) than those with plural auxiliaries in the non-congruent conditions. The slowdown occurred primarily in the post-antecedent regions, namely, the verbal phrase following the main subject, in line with previous results on English (see above refs). The observed congruency effect indicates that *pro* is computed early on the basis of featural information in a preverbal auxiliary and enters the computation of a cataphoric dependency in a fashion similar to that found in overt pronoun languages. The differential effect of the Number limiting non-congruent behavior only to sg-Npl contexts, but not to pl-Nsg contexts, points to a more intricate aspect of grammar-parser interaction. Adopting the growing consensus view in the (experimental) semantic literature that, in the grammatical component, plural number is

semantically unmarked, hence, includes the meaning of singular [4], we tentatively suggest that the parser makes use of this information online by accommodating the plural auxiliary with either singular or plural number on the antecedent as compatible alternatives. This opens an interesting direction of further research regarding interaction between the grammar and the parser, building an online representation of *pro* and the role of morphological markedness constraints in sentence processing.

(1)   a.   Ko   je   osamljen,   stric   kliče   sorodnike   po   večkrat   na   dan.
           When   is$_{aux\text{-}sg}$ lonely$_{sg}$ uncle$_{sg}$ calls$_{sg}$ relatives$_{pl}$ several   times   a   day
           "When he is lonely, the uncle calls relatives several times a day"   **[+congr, +sg]**

      b.   Ko   je   osamljen,   sorodniki   kličejo   strica   po   večkrat   na   dan.
           When   is$_{aux\text{-}sg}$ lonely$_{sg}$ relatives$_{pl}$ call$_{pl}$ uncle$_{sg}$ several   times   a   day
           "When he is lonely, the relatives call the uncle several times a day"   **[-congr, +sg]**

      c.   Ko   so   osamljeni, stric   kliče   sorodnike   po   večkrat   na   dan.
           When   are$_{aux\text{-}pl}$ lonely$_{pl}$ uncle$_{sg}$ calls$_{sg}$ relatives$_{pl}$ several   times   a   day
           "When they are lonely, the uncle calls the relatives several times a day"   **[-congr, -sg]**

      d.   Ko   so   osamljeni, sorodniki   kličejo   strica   po   večkrat   na   dan.
           When   are$_{aux\text{-}pl}$ lonely$_{pl}$ relatives$_{pl}$ call$_{pl}$ uncle$_{sg}$ Several   times   a   day
           "When they are lonely, the relatives call the uncle several times a day"   **[+congr, -sg]**
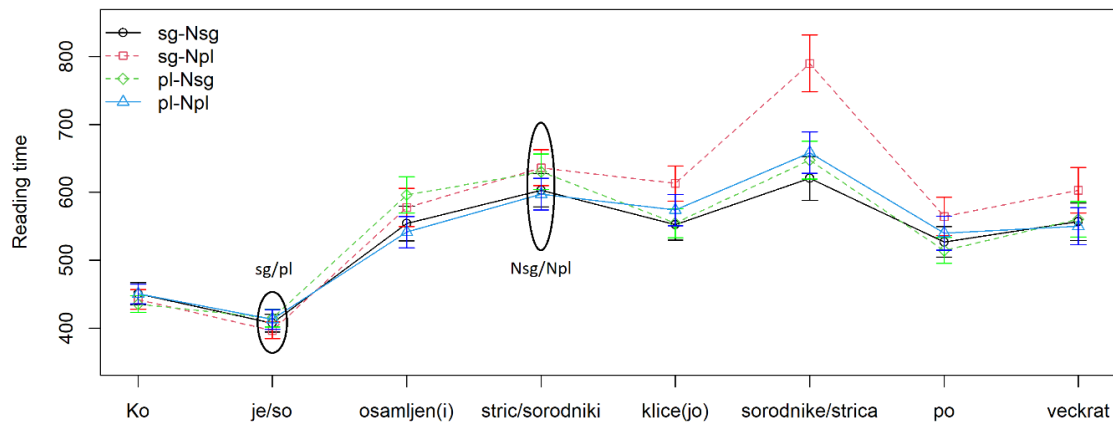

Figure 1. Time course of self-paced reading (the last two regions are not shown)

Main effect	CONGR.	Conditions	RT (ms)	SE	NUM	Main effect
χ^2=11.102 p=0.0008***	+congr	pl. – Npl.	4751	220	PL	χ^2=6.461 p=0.01*
		sg. – Nsg.	4849	286	SG	
	-congr	sg. – Npl.	5254	299	SG	
		pl. – Nsg.	4976	242	PL	
Interaction CONGR*NUM: χ^2=5.0685, p=0.024*						

Table 1. Contrasts across the Congruence and Number factors, total reading times

Selected references
[1] Cowart W, Cairns HS. Evidence for an anaphoric mechanism within syntactic processing. *Memory & Cognition.* 1987;15: 318–331. **[2]** Kazanina N, et al. The effect of syntactic constraints on the processing of backwards anaphora. *Journal of Memory and Language.* 2007;56: 384–409. **[3]** van Gompel RPG, Liversedge SP. The influence of morphological information on cataphoric pronoun assignment. *Journal of Experimental Psychology* 2003;29: 128–139. **[4]** Sauerland U, et al. 2005. The Plural is Semantically Unmarked. In Stefan Kepser and Marga Reis (eds.), Linguistic Evidence, 409–30. de Gruyter.