

## The Subject Gap Preference in a split-ergative language: Reading time evidence from Georgian

Steven Foley & Matt Wagers, University of California Santa Cruz

**Overview** • Relative clauses with subject gaps (SRCs) are routinely found to be easier to process than relative clauses with object gaps (ORCs). However, the underlying sources of this **Subject Gap Preference (SGP)** are difficult to disentangle (Kwon et al. 2010). According to the *Subjecthood Hypothesis*, SRCs are easier because gaps in subject position are universally more accessible (Keenan & Comrie 1977). But typological undersampling has led to a confound: RC-processing data have mostly come from NOMinative–ACCusative languages, where morphological case and syntactic structure covary. According to the *Unmarked Case Hypothesis*, morphological informativity directly correlates with processing difficulty (Polinsky et al. 2012). So, NOM–ACC languages show an SGP because subject gaps are associated with uninformative, default case (NOM), and object gaps with informative, dependent case (ACC). ORCs will be more taxing because it’s costly to project the specific transitive structure that licenses the ACC gap.

To test these theories, we conducted two self-paced reading experiments on Georgian. Though nearly absent from sentence-processing research (cf. Skopeteas et al. 2012), it is ideal for this issue: as a split-ergative language, Georgian totally disassociates case and grammatical position. Nonetheless, we found a clear SGP. RC complexity primarily tracked structural subjecthood, not unmarked case.

**Processing ERG relatives** • In an ERGative–ABSolutive language, unmarked ABS case appears on intransitive subjects and transitive objects, while dependent ERG case appears on transitive subjects. This morphological makes no difference for the Subjecthood Hypothesis; all languages should have an SGP. The Dependent Case Hypothesis, though, predicts an advantage for absolutive gaps — i.e., a preference for *object gaps* in transitive RCs. Recent research on ergative languages has yielded mixed results: an SGP in Ch’ol and Q’anjob’al (Clemens et al. 2015); an apparent ABS preference in Basque (Carreiras et al. 2010); a combination of both in Avar (Polinsky et al. 2012).

(1)

	<i>S<sub>Trans.</sub></i>	<i>S<sub>Unacc.</sub></i>	DO	Alignment
FUT...	NOM		DAT	NOM–ACC
AOR...	ERG	NOM		ERG–ABS
PERF...	DAT	NOM		‘DAT–ABS’

Enter Georgian. Case alignment in this language depends on the clause’s tense–aspect–mood features (TAM). Arguments show either a NOM–ACC pattern, or one of two ERG–ABS patterns (1). Consequently, cases vary greatly in their informativity for parsing.

**Design** • We tested whether the difficulty of RC processing in Georgian depends on syntactic structure or case informativity with two self-paced reading experiments ( $N_{\text{SUBJ}} = 56$ ). In Experiment 1 (described here), 36 item sets were constructed in a  $2 \times 3$  design, crossing extraction site (SRC, ORC) and TAM / case alignment (FUTURE: NOM–ACC; AORist: ERG–ABS, PERFect: DAT–ABS). Table (2) gives the relevant regions from an example item set. For padding/spillover, experimental items also had a sentence-initial adverb before this DP, and a four-word matrix clause continuation after it.

(2)

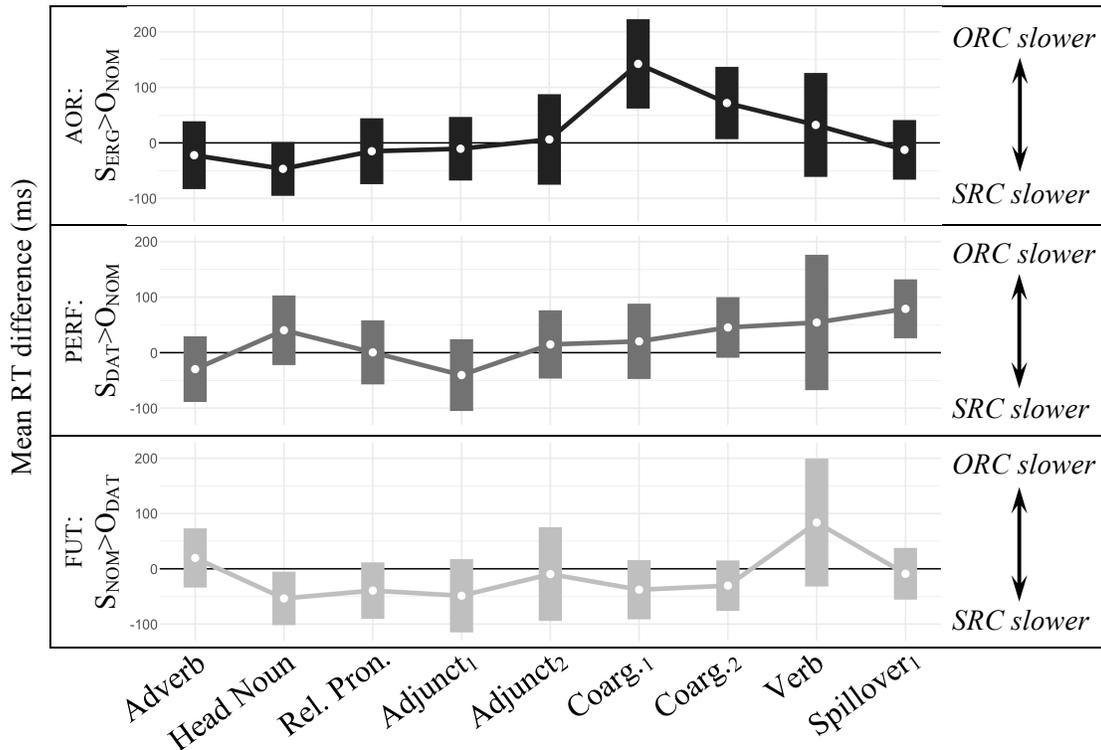
TAM	RC Type	Head	Rel. Pronoun	Adjunct	Coargument	Verb
FUT: $S_{\text{NOM}} > O_{\text{DAT}}$	SRC ORC	<i>cxeni</i>	<i>romel-ic</i> <sub>NOM/S</sub> <i>romel-sac</i> <sub>DAT/O</sub> <i>romel-mac</i> <sub>ERG/S</sub> <i>romel-ic</i> <sub>NOM/O</sub> <i>romel-sac</i> <sub>DAT/S</sub> <i>romel-ic</i> <sub>NOM/O</sub>	<i>bnel tje-ši</i>	<i>čkar-∅ irem-s</i> <sub>DAT/O</sub>	<i>naxavs</i> <sub>FUT</sub>
					<i>čkar-i irem-i</i> <sub>NOM/S</sub>	
AOR: $S_{\text{ERG}} > O_{\text{NOM}}$	SRC ORC				<i>čkar-i irem-i</i> <sub>NOM/O</sub>	<i>naxa</i> <sub>AOR</sub>
					<i>čkar-ma irem-ma</i> <sub>ERG/S</sub>	
PERF: $S_{\text{DAT}} > O_{\text{NOM}}$	SRC ORC				<i>čkar-i irem-i</i> <sub>NOM/O</sub>	<i>unaxavs</i> <sub>PERF</sub>
		<i>čkar-∅ irem-s</i> <sub>DAT/S</sub>				
		horse	which <sub>CASE/ROLE</sub>	dark forest-in	quick deer <sub>CASE/ROLE</sub>	see <sub>TAM</sub>

“...the horse [RC which \_\_\_ {will see, saw, has seen} the quick deer in the dark forest ]...” or  
 “...the horse [RC which the quick deer {will see, saw, has seen} \_\_\_ in the dark forest ]...”

**Results** • Figure (3) reports differences in reading times between the gap site conditions for each region ( $RT_{\text{ORC}} - RT_{\text{SRC}}$ ), calculated by participant. White points indicate the mean differences; bars

represent 95% confidence intervals. If a bar falls completely above 0ms, that region was read significantly slower in the ORC condition (i.e., there's an SGP). If it's below 0ms, the SRC condition was read slower (i.e., there's an object gap preference). Only RTs for words up to the first spillover region are shown here.

(3)



A clear SGP manifests in several ways. In the AOR condition, the ORC's ERG coargument is read significantly slower than the SRC's NOM coargument ( $p < 0.001$ ). This ERG coargument cost was replicated in Experiment 2, which examined another type of relative clause. Since the ERG coargument eliminates the possibility of a subject gap, this finding is in line with the Subjecthood Theory: ORC-disambiguation taxes the parser. Crucially, there is no *general* ERG cost (cf. Polinsky et al. 2012) — ERG and NOM relative pronouns are not read any differently ( $p = 0.613$ ), and a post-hoc analysis reveals ERG and NOM head nouns aren't either ( $p = 0.154$ ). Therefore, this effect cannot be due to the informativity of dependent ERG case.

An ORC-disambiguation cost is also seen in the PERF and FUT conditions. Here, the verb's TAM morphology disambiguates the syntactic role of NOM and DAT DPs, and thus also the gap site. In the PERF relatives, we see an SGP in the first spillover region ( $p < 0.005$ ). In the FUT, there appears to be a similar effect at the verb region, but because of a spurious baseline difference starting before the relative clause, it doesn't reach significance.

**Conclusion** • We found a strong SGP in Georgian, despite its split-ergative case system. In both experiments, RTs slowed at ORC disambiguation points: ERG coarguments, and certain RC-final verbs. Contributing to the growing crosslinguistic perspective on relative clause processing, this study provides strong evidence in favor theory which privileges a gap's structural position, rather than the case morphology associated with it. Subjecthood, then, must be a driving factor for comprehenders of Georgian, even though the language's case system is radically dissociated from syntactic role.

**References** • Carreiras, M., et al. (2010). Subject relative clauses are not universally easier to process. *Cognition* 115. • Clemens, L. E., et al. (2015). Ergativity and the complexity of extraction. *NLLT* 33. • Keenan, E. & B. Comrie (1977). Noun phrase accessibility and universal grammar. *LI* 8. • Kwon, N., et al. (2010). Cognitive and linguistic factors affecting subject/object asymmetry. *Language* 86. • Polinsky, M., et al. (2012). Subject preference and ergativity. *Lingua* 122. • Skopeteas, S., et al. (2012). Case inversion in Georgian: Syntactic properties and sentence processing. In *Case, Word Order and Prominence*, Springer.