A subject relative clause preference in a split-ergative language: ERP evidence from Georgian

Ellen Lau\textsuperscript{a}, Michaela Socolof\textsuperscript{a,b}, Nancy Clarke\textsuperscript{a,c}, Rusudan Asatiani\textsuperscript{d}, Maria Polinsky\textsuperscript{a}

\textsuperscript{a} – University of Maryland, Department of Linguistics, 1401 Marie Mount Hall, College Park, MD 20742, USA

\textsuperscript{b} – McGill University, Department of Linguistics, 1085 Dr. Penfield, Room 111, Montreal, Quebec H3A 1A7, Canada

\textsuperscript{c} – Amazon, 27 Melcher St., Boston, MA 02210, USA

\textsuperscript{d} - Tbilisi State University, Department of General Linguistics, 13 Ilia Chavchavadze Avenue, T’bilisi 0179, Georgia

Email addresses:
ellenlau@umd.edu (corresponding author)
michaelasocolof@gmail.com
nancyj.clarke@gmail.com
rus_asatiani@hotmail.com
polinsky@umd.edu
Abstract
A fascinating descriptive property of human language processing whose explanation is still debated is that subject-gap relative clauses are easier to process than object-gap relative clauses, across a broad range of languages with different properties. However, recent work has argued that this descriptive generalization does not hold in Basque, an ergative language. These results have motivated an alternative generalization: that the preference is for gaps in morphologically unmarked positions, which would correspond to subjects in nominative-accusative languages, but to objects and intransitive subjects in ergative-absolutive languages. In the current study, we used ERP and self-paced reading measures in another ergative language, Georgian, to investigate whether this new proposal is supported by data from other languages with ergative-absolutive systems. The results instead show the classic pattern of subject relatives being easier to process than object relatives, as we observe a large anterior negativity and slower reading times when a relative clause is disambiguated to an object relative vs. a subject relative. These data thus appear to support the original descriptive generalization that object relative clauses are more costly to process than subject relatives, independent of the language’s case system.

Introduction

Theories of relative clause processing that have been put forth to explain these results have tended to focus either on the linear/temporal distance between the filler (in this case, the head noun) and the gap (inside the relative clause), or, alternatively, on the grammatical function of the relativized noun, whereby relativization on subjects is inherently easier to process. One obvious way to distinguish between these sets of theories is to look at languages that have prenominal relative clauses. In these languages, distance between the gap and filler will be shorter in object relative clauses, so if the subject processing advantage seen in languages with postnominal relative clauses, like English, is due to linear/temporal distance between the filler and gap, then we would expect object relative clauses to be easier. If, on the other hand, subjects are universally easier to process, then such languages should show the familiar preference for subject relative clauses.

The subject processing advantage has been consistently replicated in two languages with prenominal relative clauses: Japanese (self-paced reading time: Kanno & Nakamura 2001, Miyamoto & Nakamura 2003, Ishizuka et al. 2003; ERP: Ueno & Garnsey 2008) and Korean
(self-paced reading time: Kwon et al. 2006, Kwon 2008b; eye-tracking: Kwon et al. 2010; ERP: Kwon et al. 2013). For Chinese, however, mixed results have been reported: some studies of Mandarin have found a preference for subject relative clauses (self-paced reading time: Lin & Bever 2006, Chen et al. 2010), while others have found a preference for object relative clauses (self-paced reading time: Hsiao & Gibson 2003, Lin & Garnsey 2007, Chen et al. 2008, Lin 2010, Gibson & Wu 2013; maze-task: Qiao et al. 2012). For Cantonese, an object relative clause preference has been reported for child language (Yip & Matthews 2007). The reasons for this discrepancy are unclear, but it may be due to language-specific properties that impose further or competing constraints on processing.

Yet another dimension of cross-linguistic variation has to do with the encoding of subjects. The languages listed above all have nominative-accusative alignment, which means that their subjects, regardless of transitivity, are in the nominative case (additional subject encoding, for example, with dative experiencer subjects, is possible but such subjects are not included in the standard experimental stimuli). However, a large number of languages throughout the world have ergative-absolutive alignment, where the case of subject varies depending on transitivity (intransitive subjects appear in the absolutive case, transitive in the ergative case) and the same case—absolutive—can encode a subject and object.

Whereas nominative-accusative languages do not allow one to distinguish between the effects of a subject gap advantage versus an advantage of extracting form a position with unmarked (nominative) case, ergative-absolutive languages have the potential to distinguish between the two possibilities. Since ergative-absolutive languages separate case and grammatical function, these languages allow us to compare the processing of ergative and absolutive subjects. If it is the case that subjects are universally easier to process, then both ergative and absolutive subjects should show facilitation when compared with absolutive objects; on the other hand, if case is the relevant factor, then ergative subjects should be more difficult to process than absolutive subjects and should instead behave more like objects when it comes to subextraction.

Carreiras and colleagues (Carreiras, Duñabeitia, Vergara, de la Cruz-Pavia, & Laka, 2010) investigated subject vs. object relative clause preferences in Basque, which, like Japanese, Korean, and Chinese, has head-final relative clauses, but which also has an ergative-absolutive case system. In two separate experiments they measured self-paced reading reaction times and ERPs, in a disambiguation design where they looked for differential processing difficulty for disambiguation to an object relative clause vs. a subject relative clause.

Because Basque is heavily inflected, Carreiras et al. (2010) relied on an accident of form overlap in order to create a temporarily ambiguous structure. In particular, the first noun of the relative clause could be a plural absolutive, where the plural marker on the noun is –ak, or a singular ergative, where the singular determiner is –a- and the ergative marker is –k, yielding the same ending –ak. This form ambiguity in the context allowed the researchers to create sentences that were ambiguous between subject- and object-relative-clause analyses until the main clause auxiliary. The disambiguating cue is the singular vs. plural-marking on the auxiliary, viz., the singular ditu ‘has’ in the SRC condition the and plural-marked dira ‘are’ in the ORC, as illustrated in the example below.

(1) Subject relative clause (SRC)

[ ___i irakasle-ak aipatu ditu-en] ikasle-a-k,
  teacher-ABS.PL mentioned has-ADNOMINAL student-DET-ERG]
  lakun-ak ditu.
  friend-ABS.PL has

“The student that mentioned the teachers has friends.”
In their self-paced reading experiment, Carreiras et al. (2010) found a slowdown at the singular-marked verb that disambiguated to the SRC analysis relative to the plural-marked verb that disambiguated to the ORC analysis. In their ERP experiment, they found that ERPs to the singular-marked verb were more positive than ERPs to the plural-marked verb, over left-hemisphere sites, which they interpreted as a late positivity to SRCS of the type that is often reported for syntactic violations or garden paths. Together, they interpreted these results as evidence that ORCs are preferred to SRCS in Basque, in contrast to the typical SRC > ORC pattern observed in the previous literature. They suggested that the correct cross-linguistic generalization may not be that subject relative clauses are preferred, but that comprehenders prefer relative clauses headed by an unmarked case. As ergative is a marked case, this predicts that languages with ergative subjects will prefer object relative clauses that are headed by the unmarked absolutive case. Somewhat relatedly, subsequent ERP work by the same group in Basque argues that ergative and absolutive subjects are subject to qualitatively different agreement computations with the verb (Chow et al. 2018).

Although these results are intriguing, further work is needed to determine whether they generalize to other constructions and languages. Notably, the results of a recent self-paced reading study of relative clause processing in Georgian by Foley and Wagers (2017) appear to conflict with Carreiras et al.’s (2010) proposed generalization. Foley and Wagers used structures in which disambiguation to subject vs. object relative clauses occurred early in the relative clause, either by using clause-initial relative pronouns whose case marking reflected the case of the relativized argument or by positioning the co-argument early in the relative clause so that its case would disambiguate the structure (i.e. if the co-argument is an ergative-marked subject, the gap cannot also be a subject). They found faster reading times for cues to subject relatives than cues to object relatives even when cue informativity itself was factored out, like the traditional pattern and unlike what was reported by Carreiras et al. (2010) for Basque.

Here we aimed to extend the initial findings of Foley and Wagers (2017) with data from ERP, which provides greater temporal sensitivity and where different scalp distributions can allow some degree of qualitative discrimination between different response types.

Before we discuss our study, some comments on the grammatical structure of Georgian are in order. Georgian is a split-ergative language—that is, subjects only appear with ergative case in certain constructions. In Georgian, the split is based on verbal aspect: ergative case shows up when the verb is in the aorist series, which has strong perfective connotations. In Georgian, the ergative case is always overtly marked with the suffix -m(a) (-m after a vowel, -ma after a

1 Some researchers characterize Georgian alignment as active-inactive (split-intransitive) rather than ergative. On this approach, the main distinction is between agentive and non-agentive subjects (e.g. Harris 1981). The former are always in the agentive form (the form in -m(a) described below), the latter are themes, and grouped together with nominative-marked objects which are also themes. Georgian does have a number of verbs which are semantically intransitive (‘yawn’, ‘dance’) but take the ergative subject in the aorist series, which is the main reason for classifying it as active. However, it remains to be seen if such verbs are genuinely intransitive or have a covert object. (They were not included in our study.) In this work, we proceed with characterizing the language as split-ergative.
consonant), which makes it quite visible and distinct. The accompanying object appears in the nominative case. Another important property that played a critical role in our design has to do with a relatively free word order found in Georgian; in declarative sentences, all orders are possible as long as the verb is not in the first position. In order to provide a closer point of comparison with the prior Basque study, we took advantage of this relatively free word order, to create structures in which the disambiguation point between subject vs. object relative clause (the case-marked co-argument) occurred closer to the end of the clause. Finally, the reader should keep in mind that the head noun precedes a relative clause modifying it. Relative clauses can include a relative pronoun romeli ‘which’ that matches the case of the extracted constituent, or the invariable complementizer rom ‘that’. We used the latter, to make the materials more ambiguous and challenging.

To illustrate the key properties of Georgian that govern our design, we provide an example item set below. In both conditions, sentences began with an adverbial clause and a leading ergative-marked noun phrase. At this point in the sentence, a reader would likely presume that a simple matrix clause lies ahead. However, at the complementizer rom, readers can determine that they have likely hit the edge of an embedded clause, more specifically, a relative clause from which the subject noun has been extracted. Through the complementizer and the verb, it is ambiguous whether the extraction is from the subject or object position of the embedded clause. This ambiguity is resolved by the case-marking on the subsequent noun inside the relative clause. If that noun is marked absolutive, then the gap must correspond to an (ergative) subject of a transitive clause; if that noun is marked ergative then the gap must correspond to an (absolutive) object of a transitive clause. In the examples below, the matrix clause has a transitive verb (‘buy’) in the aorist, so the subject is ergative and the object is nominative (‘the baker bought a new house’). The ergative matrix subject is modified by a relative clause in which this argument is either in the subject, or object, position.

(2) Subject relative clause (SRC) condition

ცინა ქვირას ხაბაზმა ფ’ოსთ’ის გარეთ რომ ცააქცია ფ’ოსთ’ის ალიონ-ი მშვენიერი სახლი
last week-DAT baker-ERG post.office-GEN outside COMP trip.3SG.AOR mailman-ABS

‘Last week the baker [that ___ tripped the mailman outside the post office] bought a new house.’

(3) Object relative clause (ORC) condition

ცინა ქვირას ხაბაზ-მა [ფ’ოსთ’-ის გარეთ-რომ ც’aakcia პ’ოს’ალიონ-ი]
last week-DAT baker-ERG post.office-GEN outside COMP trip.3SG.AOR mailman-ABS

‘Last week the baker [that the mailman tripped ___ outside the post office] bought a new house.’
In sum, our goal was to investigate the hypothesis that the traditional SRC > ORC preference is flipped in ergative-absolutive languages, as claimed for Basque by Carreiras et al. (2010). As in the prior work, we conducted both an ERP experiment and a self-paced reading experiment, in order to take advantage of the strengths of both. ERP provides an extremely time-sensitive index of processing differences. However, because there is no straightforward correspondence between the direction of ERP voltage changes and the overall amount of neural activity, in some cases in which differences are observed it can be difficult to disambiguate which condition is inducing more processing cost, as discussed above. Self-paced reading is a less time-sensitive measure, but as a reaction time measure, the directionality of differences can be more straightforwardly interpreted with the assumption that larger RTs indicate greater processing difficulty.

**Experiment 1 - EEG**

*Participants*

EEG data were collected from a total of 46 participants, for which they received monetary compensation. All the participants were right-handed native speakers of Georgian. Written informed consent was obtained from all participants. Datasets from 15 participants had to be excluded from further ERP analysis due to excessive artifact during the epoch (artifact identification procedures detailed below). This number is slightly larger than typical for EEG studies, primarily because imperfect climate control during the hot Georgian summer resulted in large sweat artifacts in a significant number of participants. Data from the remaining 31 participants (9 male, mean age 23.6y) were carried forward for subsequent analysis.

*Materials*

We constructed 60 item sets in which sentences were initially ambiguous as to whether the relative clause contained a subject or object gap, and where relative clause structure was disambiguated by the case-marking on the embedded noun, as illustrated above in (2)-(3). In each sentence both the subject and the object were animate (either both animals or both humans). The stimuli capitalized on the aforementioned free word order of Georgian; the most common orders are verb-final and verb-medial (Skopeteas et al. 2009), with both subject and object equally possible in the postverbal domain.

Because in this experiment the participant’s task was to judge the sentences for acceptability, half of the experimental items included grammatical errors at some later point in the sentence after the critical relative clause region. These subsequent errors were always fully independent from the relative clause structure itself and were distributed equally across ORC and SRC conditions so that the clause structure did not predict the likelihood that the sentence would be judged acceptable.

In addition to the experimental items, there were 120 fillers and 180 items from another experiment. The 180 items from the other experiment, as well as half of the fillers, began with either a locative adverbial or a temporal adverbial, followed by a noun, just like the relative clause stimuli. The remaining 60 fillers were sentences that began in other ways, usually with a sentence initial noun or a subordination marker.

The 60 item sets were distributed across two lists in a Latin Square design, such that each item could appear in the ORC and SRC condition, but only one of these versions would occur on any given list. The 60 RC items from each list were combined and randomized with the 300 additional items, such that each participant saw a total of 360 items in the experimental session,
where half of the items were designed to be judged acceptable and half were designed to be judged unacceptable.

**Procedure**

During the experiment, participants were seated in a chair in a quiet room. Stimuli were visually presented on a computer monitor in white 18-point text on a black background. Each trial began with a 1000ms fixation cross. After a 200ms blank screen, the words of the sentence were presented with a constant 600ms stimulus onset asynchrony (SOA), where each word appeared for 500ms separated by a 100ms blank screen. The final word stayed on the screen for a duration 600ms, followed by a blank screen of 200ms. Then the probe screen appeared, asking whether the sentence was acceptable or not. Participants responded using the ‘F’ and ‘J’ keys on the keyboard, where ‘F’ indicated acceptable and ‘J’ indicated unacceptable. The experiment was preceded by a brief practice session with filler sentences to ensure that participants understood the task and were comfortable with the presentation format. Five breaks were evenly spaced across the experiment to allow participants to rest.

**Electrophysiological Recording**

Sixteen Ag/AgCl electrodes were held in place on the scalp by an elastic cap (BrainVision): AFz, F7, F3, Fz, F4, F8, FC5, FC6, Cz, CP5, CP6, P7, P3, Pz, P4, P8. Bipolar electrodes were placed above and below the left eye and at the outer canthus of the right and left eyes to monitor vertical and horizontal eye movements. Responses were referenced to the left mastoid. The ground electrode was positioned on the scalp between Fz and Cz. Impedances were maintained at less than 10 kΩ for all scalp and ocular electrode sites and less than 2 kΩ for the mastoid site. The EEG signal was amplified by a portable BrainVision V-Amp system and continuously sampled at 512 Hz by an analog-to-digital converter.

**Analysis**

As our recordings were conducted in an environment without electrical shielding, two notch filters were applied offline to the continuous data (50Hz and 100Hz) to minimize line noise. We also applied offline a more standard bandpass filter (Butterworth, order 2) of .1-20Hz. We then extracted epochs time-locked to the onset of the critical word from -100:1000ms. Averaged ERPs were formed from these epochs, after rejecting trials free of ocular and muscular artifact, using preprocessing routines from the EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014) toolboxes. Muscle potential, sweat, and alpha wave artifacts were identified using the peak-to-peak artifact rejection routine provided by ERPLAB, and eye-blink and eye-movement artifacts were identified using the step function artifact rejection routine provided by ERPLAB, followed by visual confirmation of the identified artifacts by the experimenters. Epochs with artifacts were excluded. In three datasets, one electrode (different for each dataset) contained a disproportionate number of epochs containing peak-to-peak fluctuations of 100 μV or more and was therefore replaced with an interpolated value from surrounding electrodes. A 100-ms pre-stimulus baseline was subtracted from all waveforms, and a 40-Hz low-pass filter was applied to the ERPs offline. ERP data are made publicly available for further analyses on the first author’s website (http://ling.umd.edu/~ellenlau/public_data_archive/Georgia).

We conducted Type III SS repeated-measures ANOVAs on mean ERP amplitudes in two time-windows: 300-500ms for the LAN and 800-1000ms for the late positivity. We focused on the later end of the traditional time-window in which late positivities are observed (~600-1000ms) because the complexity of Georgian morphology would be likely to increase the
processing time associated with basic morphological decomposition, and because a syntactic violation manipulation in the other sub-experiment elicited a late positivity in this later time-window (Lau et al. 2018; in prep.). In order to quantify the topography of the effects, we included the factor of anteriority in all analyses (anterior electrodes: AFz, F7, F3, Fz, F4, F8, FC5, FC6; posterior electrodes: Cz, CP5, CP6, P7, P3, Pz, P4, P8).

Results

Behavioral accuracy on acceptability judgements across the relative clause conditions was 63% and was similar for subject relative clauses (61%) and object relative clauses (64%). These accuracies were somewhat low, which can be attributed to several factors. First, our stimuli were quite complex, and the overall length and syntactic complexity may have led to lower ratings (see Alexopoulou & Keller 2007, Sprouse & Almeida 2017 on the effect of the length of dependencies on ratings). Second, and more importantly in our view, Georgian participants are not used to sentence rating, and that lack of familiarity with test-taking is reflected in relatively low ratings (see Clemens et al. 2015 for similar observations). In their study, Foley and Wagers (2017) observed a similar trend even though their stimuli were simpler than the ones used in our study.

In ERPs to the critical noun, an increased negativity over anterior electrodes was observed for the object relative clause compared to the subject relative clause (Figures 1 and 2). This resulted in a significant interaction between condition and anteriority in the 300-500ms time-window (F(1,30) = 6.0, p < .05). In the 800-998ms time-window we observed a marginal interaction between condition and anteriority (F(1,30)=4.05, p = .053). This interaction appeared to be driven by the combination of a residual anterior negativity and a small late posterior positivity in the object relative clause condition.

Figure 1. ERP responses at the point of disambiguation to object relative clause vs. subject relative clause analysis at representative frontal and posterior electrode sites. Scalp maps illustrate the distribution of the difference between the two conditions (ORC – SRC) in the LAN time-window and the late positivity time-window.
Discussion

The primary finding from the ERP experiment was that ambiguous relative clauses elicit an increased anterior negativity when disambiguated to the ORC analysis compared to the SRC analysis. As prior ERP work has associated anterior negativities with increased processing cost, these results appear to indicate that Georgian does not fit the generalization proposed by Carreiras et al. (2010) in which SRCs should be dispreferred in an ergative language.

Although logically possible, we believe it is unlikely that these differences are due to inherent differences in processing the ergative and absolutive case. A separate manipulation in the same session (reported in Lau et al., 2018 and in prep.) required us to vary the case of the matrix subject between ergative and absolutive (Tomorrow doctor-erg… or Tomorrow doctor-abs…). In this manipulation, we observed no hint of differences between the two conditions at the matrix subject position.

At the same time, it remains a logical possibility that our interpretation of the ERP difference as an increased anterior negativity for the ORC could be incorrect, and that instead the ERPs reflect an increased anterior positivity for a more costly SRC analysis. This possibility is particularly pertinent because Carreiras et al. (2010) observed a left-lateralized difference spanning anterior and posterior electrodes in which SRCs were more positive than ORCs and interpreted this as a P600 associated with processing cost in SRCs. We believe that such an interpretation is less possible in the current case, because the strongly frontal distribution of the effect is so different from the posteriorly distributed P600 standardly observed for syntactic manipulations in the ERP literature. Furthermore, we observed a trend towards a more standardly distributed P600 effect for ORCs in the later time-window. However, in Experiment 2, we directly evaluated this possibility by investigating reaction time measures to the same SRC and
ORC materials, in which increased processing cost can be more unambiguously related to increased RTs.

**Experiment 2 – Self-paced reading**

**Participants**
Self-paced reading data was collected online from a total of 33 participants (mean age = 22.1 years), for which they received monetary compensation. All the participants were monolingual native speakers of Georgian. Written informed consent was obtained from all participants. The majority of participants were current or former undergraduate students living in Tbilisi.

**Materials**
Materials were a subset of 30 of the relative clause items used in the EEG experiment, along with 136 fillers and 20 items from another experiment. The other sentences were mostly simple clauses (i.e. did not include relative clauses), though they began similarly to the relative clause items, with an adverbial phrase followed by the subject noun. As in the EEG experiment, experimental items were distributed across lists in a Latin Square design such that each participant saw 15 items from each of the two conditions and no participant saw more than one version of each item.

**Procedure**
Items were presented word-by-word in a self-paced moving window paradigm (Just, Carpenter & Woolley, 1982) using the IBEX software (Drummond, 2016). Each trial began with a screen in which a row of dashes masked the words in the sentence. Participants revealed the first word and each subsequent word by pressing the space bar. When a new word was revealed, the previous word was re-masked, so that only one word was ever visible at a time. Participants were instructed to read as naturally as possible at their normal reading speed and to make sure that they understood the sentences they were reading. 25% of items were followed by a two-choice comprehension question to encourage participants to attend to the stimuli.

**Analysis**
Reaction times exceeding a threshold of 2.5 seconds were excluded as outliers from further analysis; this impacted 2.1% of all reading times. The regions of analysis were the disambiguating case-marked noun and the subsequent verb (the spillover region). As we had only two conditions, we conducted pairwise comparisons between SRC and ORC conditions in those two regions.

**Results**
Reading times for the disambiguating noun and the three words before and after are presented in Figure 3. At the disambiguating noun, we observed significantly slower reading times for the ORC condition relative to the SRC condition ($t(31) = 3.07, p < .05$, mean difference $= 85$ms). This was also the case in the following spillover region ($t(31) = 2.78, p < .05$, mean difference $= 55$ms).
Discussion

The goal of Experiment 2 was to provide converging reaction time evidence on Georgian speakers’ preference for disambiguation to subject vs. object relative clauses. The results of the self-paced reading experiment showed a clear reaction time cost for disambiguation to ORCs. This supports our interpretation of the ERP results of Experiment 1 in which the cost of ORC disambiguation was reflected by an increased anterior negativity (AN).

General Discussion

The current study evaluated processing cost differences for subject vs. object relative clauses in Georgian, in the context of recent proposals that ergative-absolutive languages do not show the ORC processing cost observed in most other languages (Carreiras et al., 2010; Chow et al., 2018). In two experiments using ERPs and self-paced reading respectively, we presented native speakers of Georgian with sentences which were initially ambiguous between subject and object relative clause analyses and looked for signs of differential processing cost at the point of disambiguation. Both experiments indicated that Georgian speakers preferred disambiguation to the subject relative as opposed to object relatives. In particular, we observed an increased anterior negativity for disambiguation to the ORC structure in ERPs, and increased reaction times for disambiguation to the ORC structure in self-paced reading. Together, these results suggest that the generalization about SRC vs. ORC preference in ergative-absolutive languages cannot be stated as simply as an overall preference for ORCs due to the morphologically unmarked form of their objects.

As reviewed in the introduction, some theories propose that object relative clauses bear an extra processing cost that holds across languages; for example, that the extra cost for ORCs lies in the complexity of the message or scenario that they convey, relative to SRCs (e.g. Keenan & Comrie, 1977; MacWhinney, 1977). These classic theories naturally account for the current results, as they demonstrate that yet another language, with its own idiosyncratic morphosyntactic properties, fits the predominant ORC-more-costly-than-SRC pattern observed for most other languages.
These theories still need to provide an explanation for why Carreiras et al. (2010) observed the reverse pattern in self-paced reading times in Basque. However, as discussed in the introduction, several additional factors in that study could plausibly have masked the independent cost of ORCs. First is the fact that the temporary ORC/SRC ambiguity in that study was introduced by using a morphological ambiguity between the ergative singular ending and the absolutive plural ending, such that baseline differences in frequency between these morphemes could have driven parsing commitments that would have downstream consequences before comprehenders are even aware that they are in a relative clause. Second is the fact that the critical disambiguation region was a main clause verb whose transitivity differed across the two conditions, potentially introducing variability that would be independent of relative clause processing itself (see Longenbaugh & Polinsky, 2017, for further discussion).

It is less clear to us whether or not the Basque ERP results from Carreiras et al. (2010) in fact conflict with the current ERP results. One possibility is that the same factors that drove the increased reading times for disambiguation to SRC in Basque also drove an increased positivity in the ERP. For example, if the differences in morphological frequency discussed above led to an ORC analysis, disambiguation to the SRC analysis at the main verb could induce reanalysis processes that have been associated with late positivities. However, as we noted in the discussion above, the left-lateralized, slightly anterior topographical distribution of the SRC – ORC difference in Carreiras et al. (2010) is not typical of the late positivities associated with reanalysis, which are reliably central-posterior. Therefore, an alternative interpretation of the Basque ERP results is that they reflect an increased anterior negativity for disambiguation to the ORC analysis, as in the current results. We note that this second interpretation would require additional assumptions about how different sources of processing cost are mapped to self-paced reading times and ERP, respectively, and therefore we remain agnostic here about which is the better account of the Basque data.

One question is the extent to which these differences in ORC/SRC processing cost are proximally driven by differences in the frequency of the constructions. However, this assumption is not supported by corpus data. In the Corpus of New and Modern Georgian (http://corpora.iliauni.edu.ge), we examined 785 relative clauses with rom.2 Out of those, 32 percent were relative clauses with the intransitive subject gap, 26 percent had transitive subject gap, and 28 percent had the object gap (the remaining 14% were other gaps). These results are comparable to the relative frequencies of relative clauses in other ergative languages, where the most common relative clause structures are intransitive subject RCs, while transitive subject and object RCs are fairly comparable in frequency (Carreiras et al. 2010; Longenbaugh & Polinsky, 2017). They are also comparable to the distribution of relative-clause types in English (Gordon & Hendrick 2005). Crucially, the frequency of transitive-subject relative clauses and that of object relative clauses were comparable. Thus, the conditional likelihood of a SRC vs. an ORC, based on past experience, was well-matched prior to the moment at which participants encountered the disambiguating noun.

Conclusion

Here we have reported some of the first ERP work on online comprehension of Georgian. Our results make an important contribution to ongoing debates about language processing differences across languages, suggesting that in contrast to a recent proposal, the classic

---

2 This corpus is the largest currently available for Georgian and it is more annotated than some other corpora, for example, the Georgian National Corpus (available at http://gnc.gov.ge/gnc/page), which is still under development and is smaller than the corpus we used.
preference for subject relative clauses over object relative clauses is mirrored in at least some ergative-absolutive languages. More broadly, we note that with continued improvements in the portability and usability of the EEG technology, it is now much more feasible for psycholinguists to bring their equipment to new populations of speakers and languages who can most effectively resolve core questions about language processing.

Acknowledgments. This work was supported in part by NSF grant BCS-1619857 to Maria Polinsky, NSF grant BCS-1749407 to Ellen Lau, and by the College of Arts and Humanities at University of Maryland. We would like to thank Anna Namyst and Bill Idsardi for technical assistance. We are grateful to Zurab Baratashvili, Irina Lobzhanidze, Tamar Makharoblidze, Irakli Salia for help with the Georgian data. All errors are our responsibility.

References


