**On the representation of morphosyntactic predictions: ERP evidence from Georgian**

Ellen Lau,[a] Maria Polinsky,[a] Nancy Clarke,[b] and Michaela Socolof[c]

University of Maryland College Park,[a] Amazon AWS AI,[b] McGill University[c]

*Abstract*

Much recent work in psycholinguistics argues that comprehenders rapidly generate expectations about the upcoming input. But so far, we know little about the representational format of these expectations. Do predictive inferences have the same status as more direct inferences from bottom-up input? We investigate this question using the case of morphosyntactic prediction in an ERP (event-related brain potential) study in Georgian. Our experimental design makes critical use of the grammatical properties of Georgian, a split-ergative language in which an ergative-marked nominal is most commonly found in the aorist (typically translated as past tense). Combining the ergative with temporal adverbials like 'tomorrow', which strongly predict a clause with non-past interpretation, allows us to create a situation in which the cues conflict not in themselves, but in the constraints that they place on the upcoming verb tense. If these cues drove predictive, incremental updates to the syntactic representation in the same way as bottom-up input, we might expect to see an immediate response to the conflict. However, the results of our ERP experiment provide no positive evidence of processing difficulty when contextual cues motivate conflicting predictive inferences about the tense of the upcoming verb. The late positivity ERP response associated with processing difficulty is only observed when direct evidence of unacceptability is encountered at the verb itself.

Keywords: context, ergativity, event-related brain potentials (ERP), morphosyntactic cues, prediction, split ergativity

In investigations of the human cognitive architecture that supports language comprehension, the majority of experimental studies have focused on just a few languages. However, more and more the field of psycholinguistics has come to recognize that less-commonly studied languages can afford unique opportunities to answer fundamental questions about these processing mechanisms. In this paper we describe such a case. We show how the grammatical properties of Georgian allowed us to investigate the representation of syntactic predictions in a way that would be possible in few other languages, in what to our knowledge is the first electroencephalographic (EEG) study of language processing in Georgian. To anticipate our conclusions, while our findings may be in need of further replication, it nonetheless seems simultaneously both astonishing and satisfying that brain responses could be consistent across typologically diverse languages.

*Background*

Psycholinguistic research in the last two decades has come to the conclusion that *prediction* is an important component of language comprehension. But, what is prediction? Several possibilities are on offer.

One weaker use of the concept of prediction simply equates it with context-sensitive comprehension. For example, imagine a toy comprehension system whose goal is to estimate the meaning intended by the speaker. Imagine that it is equipped with knowledge of the world, and knows, for instance, that events of the chasing of rabbits by dogs are more commonly viewed and discussed than events of the chasing of flies by dogs. Now imagine that the system is presented with the language input 'I saw a dog chasing a rabbit' or 'I saw a dog chasing a fly'. If, after receiving the last word of the sentence, the system uses world knowledge about what events are more likely to 'boost' the intended message representation in the rabbit-sentence, and if this results in the system 'selecting' this representation more quickly and effortlessly than in the fly-sentence, it would be an example of 'predictive language comprehension' under this relatively permissive view. Even simple lexical frequency effects might count as 'prediction' under this view, if the occurrence of a word or concept contributes to boosting a message because a word or concept that occurs frequently in the broadest context (all prior experience) is more likely to be intended than one that occurs less frequently.

A stronger use of the concept 'prediction'—stronger in the sense of characterizing a smaller set of processing models—uses it to refer to cases in which the context impacts estimation of an external cause before that external cause has actually generated any sensory input. For example, imagine that in the above example the sentence is presented incrementally, and at the point of receiving the input 'chasing', the system not only infers the lexical item (syntactic root) for 'chase' on the basis of this acoustic input, but also infers the lexical item for 'rabbit'. Even if these inferences both turn out to be correct, they can be distinguished by the fact that in one case the inferred lexical item 'chase' has already impacted the acoustic input, and in the other the inferred lexical item 'rabbit' has not yet had any impact on the acoustic input. And we can thus use the word 'prediction' to pick out this second class of inferences.

In this study, we adopt the stronger concept of prediction and investigate the hypothesis that comprehenders generate morphosyntactic predictions. We will explore questions that follow

from this hypothesis about the representation of the predictive inferences. It could be that, at the level of the syntactic representation being constructed during parsing, predictive (context) and non-predictive (direct evidence) inferences have qualitatively similar impacts—that is, predicting a morphosyntactic feature from context could have exactly the same kind of impact as hearing some direct acoustic or visual evidence for that feature. Alternatively, it could be that the impact of predictive inferences is qualitatively different because context is a less reliable cue.

One simple thing we know about direct evidence for morphosyntactic features is that when direct acoustic/visual evidence conflicts with the syntactic representation inferred from the prior context, there is strong and immediate processing disruption in behavioral and neural responses. For example, in ergative languages like Hindi and Georgian, an ergative-marked subject cannot appear with a verb in the imperfective (Hindi) or present tense (Georgian). If an ergative-marked subject is followed by an imperfective verb in the same clause, we immediately see reaction time slowdowns and increased amplitude of an event-related brain potential (ERP) late posterior positivity response associated with processing difficulty incurred by grammatical computation (see Dillon et al. 2012 for Hindi, and see Swaab et al. 2012 for a general overview of ERP techniques in language).

Here, we asked whether analogous effects of processing difficulty are observed when the conflict is between two strong predictive inferences. If predictive inferences about morphosyntactic features have the same cognitive status as non-predictive inferences from direct evidence, then a cue that *generates a predictive inference* that conflicts with the prior context should cause immediate processing difficulty, just as direct acoustic/visual evidence for a morphosyntactic feature that conflicts with the prior context causes processing difficulty.

The experimental design we use to explore this question in Georgian is inspired by the observation by Dillon et al. (2012) that in ergative languages, both temporal adverbials and case-marking of the subject can serve to generate predictive inferences about the same morphosyntactic features of the upcoming verb. Dillon et al.'s ERP experiment was conducted in Hindi, where an ergative-marked subject can only co-occur with perfective verb forms. They showed rapid and robust ERP sensitivity to the ungrammaticality that arises when an ergative-marked subject was followed by a non-perfective future verb form (an early negativity and a late positivity). They also showed rapid ERP sensitivity when the contextual conflict was instead induced by a mismatching temporal adverbial (e.g. 'yesterday' followed by the same non-perfective future verb form), although with slight differences in the spatial profile of the response. Although Dillon et al. (2012) were interested in the nature of those subtle differences between the adverbial and case-marking cues, here we abstract across them to focus on a different question: what happens when their predictions conflict?

Georgian has the same form of contingency between the ergative subject marker (ə(ა)/-*m*(*a*)) and (aorist) tense morphology, as in (1-2) (although debate exists in the theoretical literature over whether the marker –*m*(*a*) in fact encodes ergative case vs. an Agent whose form is conditioned on the verb's tense).[1] Another grammatical form compatible with the ergative is the relatively

[1] See Klimov & Alekseev (1980), and see Harris (1981); Nash (2017); Polinsky (2020) for discussion in English and further references.

infrequent optative, an issue we will return to below. Before we proceed with the data, a note on terminology is in order. The ergative (form in მ(ა)/-*m*(*a*)) is sometimes referred to as the narrative, and the nominative (the form often but not always ending in -ი/-*i*) is also referred to as the absolutive. In what follows, we will be using the terms *ergative* and *nominative*, although nothing hinges on this terminology.

**Case Match:**

(1) თბილის-ში კაც-მა ბარ-ში დალი-ა ღვინო და...
tbilis-ši k'ac-ma bar-ši dali-a ɣvino da...
tbilisi-in man-ERG bar-in drink-AOR wine and
'In Tbilisi, the man in the bar drank wine and...'

**Case Mismatch:**

(2) *თბილის-ში კაც-მა ბარ-ში დალი-ევს ღვინო და...
tbilis-ši k'ac-ma bar-ši dal-evs ɣvino da...
tbilisi-in man-ERG bar-in drink-FUT wine and
'In Tbilisi, the man in the bar will drink wine and...'

In our study, instead of examining the violation responses at the verb, our question focused on what would happen *prior* to the verb if the predictions generated by the temporal adverbial and subject marker conflicted with each other—in other words, if a future temporal adverbial like 'tomorrow' predicting a non-aorist verb was followed by the ergative-marked noun predicting aorist morphology. As mentioned above, the only grammatical form compatible with the ergative and future-oriented adverbs such as 'tomorrow' is the relatively infrequent optative, and we assume that the expectation of the optative is lower than that of the past-oriented aorist.[2] Therefore, if comprehenders immediately computed the constraints of both the adverbial and the case and if predictive inferences about morphosyntactic features have the same cognitive status as non-predictive ones, we would expect to observe activity at the case-marked noun, associated with shifting the prediction for future morphology on the verb to a prediction for the rarer optative morphology.

**Cue-conflict Control:**

(3) ხვალ კაც-ი ბარ-ში დალი-ევს ღვინო და...
xval k'ac-i bar-ši dal-evs ɣvino da...
tomorrow man-NOM bar-in drink-FUT wine and
'Tomorrow, the man in the bar will drink wine and...'

**Cue-conflict Mismatch:**

(4) *ხვალ კაც-მა ბარ-ში დალი-ევს ღვინო და...
xval k'ac-ma bar-ši dal-evs ɣvino da...
tomorrow man-ERG bar-in drink-FUT wine and
'Tomorrow, the man in the bar will drink wine and...'

---

[2] In the corpus data that we were able to check, the ratio of aorist and optative is about 8 to 1. (We are grateful to Irina Lobzhanidze for the numerical data.)

Having outlined the linguistic background of the experimental work presented here, we will now describe the methods and results of the study.

*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*

Our experiment was originally designed to include six conditions: (1-4) above, as well as two further conditions to confirm Dillon et al.'s (2012) observation that standard processing difficulty effects arise with tense-mismatching adverbials, here a past-tense adverbial (and nominative subject) with a present-tense verb. Therefore, we created 180 cue-type sets of six items each. All items began with a clausal adverbial, followed by the first noun which was rotated across a list of approximately 30 professions, followed by a locative, followed by the critical transitive verb, and a direct object; the first clause was always followed by a second clause connected to the first with a connective such as and/but/because. Unfortunately, due to experimenter error, the two extra adverbial conditions were accidentally created in such a way that items from both conditions appeared in an unacceptable form (specifically, in the 'acceptable' condition, 'yesterday' + nominative subject was combined with a past-tense perfective instead of an imperfective). Since this error made the contrast between those two extra conditions uninformative, we do not discuss them further in this paper.

In addition to the experimental items, there were 120 fillers and 60 items from another experiment (Lau et al. submitted). The 60 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 (similar to the relative clause stimuli used in another ERP experiment we conducted—see Lau et al., submitted). The remaining 60 fillers were sentences that began in other ways, usually with a sentence initial noun or a subordination marker.

The 180 item sets were distributed across six lists in a Latin Square design, such that each item could appear in each condition, but only one of these versions would occur on any given list. The 180 items from each list were combined and randomized with the 180 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*

The experiment was conducted at the Institute for Theoretical and Applied Linguistics at the Ivane Javakishvili Tbilisi State University. 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 with 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 (specifically, the *pop_artmwppth()* function), and eye-blink and eye-movement artifacts were identified using the step function artifact rejection routine provided by ERPLAB (specifically, the *pop_artstep()* function), followed by visual confirmation of the identified artifacts by the experimenters and exclusion. Participants for whom more than 50% of trials contained artifacts were excluded from further analysis. 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, using the *eeg_interp()* function provided by EEGLAB with the default method of spherical interpolation. A 100-ms pre-stimulus baseline was subtracted from all waveforms, and a 40-Hz low-pass filter was applied to the ERPs offline.

We conducted Type III SS repeated-measures ANOVAs on mean ERP amplitudes between 800-1000ms for the late positivity. For the latter, 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. submitted); however, we

note that the use of a less standard time-window means that the conclusions that can be drawn from the late positivity results are more tentative. 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*

As expected, we observed significant late positivity effects of unacceptability at the verb in both condition sets when the ergative case marker was followed by a future-tensed verb (Figure 1), both in the simple ergative control pair (significant interaction between condition and anteriority: $F(1,30) = 12.8$, $p < .05$) and in the future adverbial + ergative comparison of interest of interest (significant interaction between condition and anteriority: $F(1,30) = 8.6$, $p = .05$).

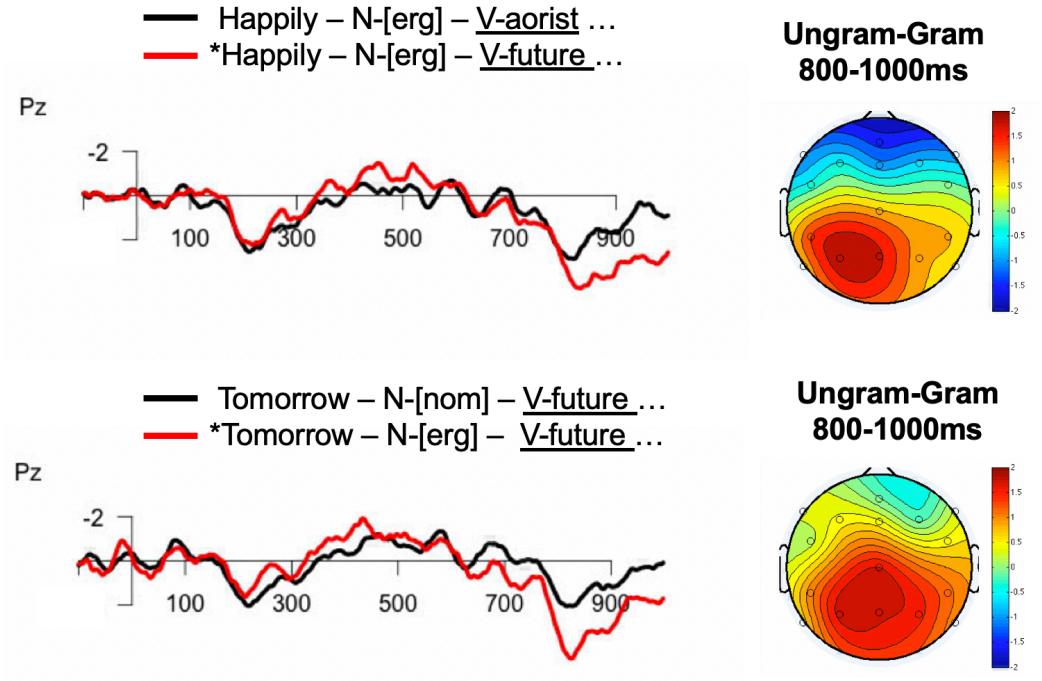


Figure 1. Scalp topoplots and ERP waveforms at electrode Pz time-locked to the verb position in the four conditions.

Our question of primary interest was whether neural effects of processing difficulty would be observed earlier in the sentence when an aorist-predicting ergative subject followed a non-aorist a conflicting adverbial, one that predicted a future tense. We saw no evidence of such processing difficulty; there were no reliable differences during the late positivity time-window (no main effect of condition, nor an interaction between anteriority and condition; *p*s > .1), and little numerical difference in the waveforms anywhere in the response to the subject noun.

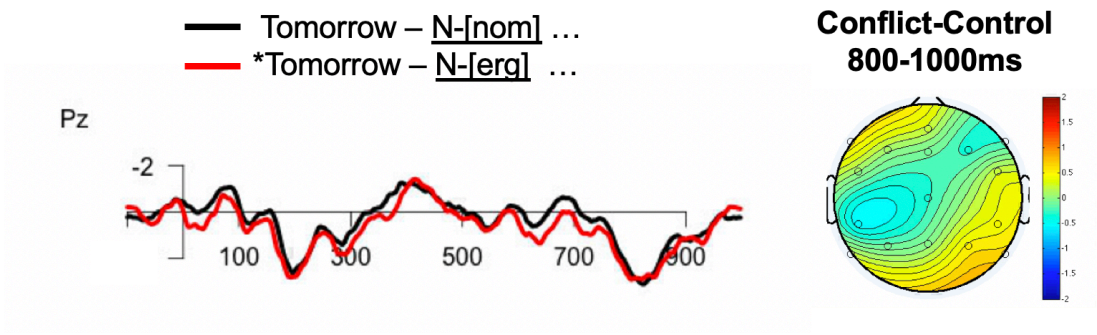


Figure 2. Scalp topoplots and ERP waveforms at electrode Pz time-locked to the noun position in the cue-conflict condition and its control.

We also examined the responses to the end-of-sentence acceptability judgment. Responses to the simple ergative cue control pair showed a robust contrast in acceptability judgments: 69% of the ergative + aorist judgments were acceptable, while only 17% of the ergative + future judgments were acceptable. However, in the cue-conflict pair, while the unacceptable ergative + future sentences yielded low ratings (10%), so did the acceptable nominative + future sentences (11%). In revisiting the stimuli, we realized that due to experimenter error, a tense mismatch systematically occurred in this condition well after the critical region, in the second clause of the sentence. Thus, even though the critical first clause was acceptable, the sentence was globally unacceptable. Below we discuss whether this post-critical region error could have had an indirect impact on responses to earlier parts of the sentence.

*Discussion*

The results of our ERP study failed to show evidence that a morphosyntactic prediction that conflicts with the context generates the same processing difficulty as direct evidence for a conflicting morphosyntactic feature. We observed standard late positivity ERP effects of conflict when an ergative-marked subject was followed by direct evidence of unacceptability, in the form of a future-tensed verb. However, we saw no difference in the ERP response when a future-tense predicting temporal adverbial (e.g., 'tomorrow') was followed by an aorist-predicting ergative subject.

The most interesting explanation for these results is that predictive morphosyntactic inferences from context and direct morphosyntactic inferences from bottom-up phonological input do not have the same cognitive status. In other words, morphosyntactic prediction is not just an 'early version' of the same operations that would be triggered by direct evidence. How might the consequences of predictive cues be different? One example would be a parser that has something akin to 'parallel activation' and 'selection' stages, where predictive cues could act to weight the support of some parses higher or lower during a stage in which multiple parses remain under consideration, while direct evidence not only updates their weights but also triggers selection of a particular parse (compare Toscano et al., 2010, for a similar approach to the parsing of sound signals). Then, the late positivity might be taken to be diagnostic of hitting a selection problem when the direct evidence from the non-aorist verb has ruled out other grammatically acceptable analyses. In contrast, the predictive cue from the ergative subject acts to markedly shift weight

away from the future-tensed analysis, but does not trigger selection. This kind of explanation plays out the common intuition that predictive cues provide less solid evidence than direct bottom-up input, and thus that their role in parsing is accordingly weaker: they serve to bias rather than to choose.

Another possibility is that the particular kind of predictive inference we tested, morphosyntactic prediction, simply isn't routinely computed at all. This explanation would be interesting because according to some popular 'Bayesian' or 'predictive coding' approaches, prediction is an inherent architectural property of cognition, and should be observed whenever reliable cues are available. Clear evidence for morphosyntactic prediction is difficult to obtain because processing difficulty when a prediction is violated can usually be equally well explained by a bottom-up view where the direct evidence forces selection of an infrequent/low-probability analysis. In older work, one of us has argued that morphosyntactic prediction for verbal agreement explains certain error patterns in agreement comprehension known as agreement attraction (e.g. *the key to the cabinets are on the table*; see Wagers, Lau and Phillips, 2009), yet later work has suggested that these error patterns are variable (Hammerly, Staub & Dillon, 2019). A case that has received a lot of attention recently in the ERP literature is the dependence of the form *a/an* in English on the predicted phonological form of the subsequent word (e.g. *On a windy day the boy flew a/*an kite*). Delong et al. (2005) and Urbach et al. (2020) have observed ERP sensitivity at the determiner as a function of the predicted form of the noun (cf Nieuwland et al. 2018). However, note that this case (and parallel ones manipulating grammatical gender on a prenominal marker) involves prediction of *lexical* information (the form or gender of a predicted noun) rather than the prediction of grammatical properties like the tense or aspect of the clause.

A third possibility is that predictive morphosyntactic inferences are generated, but that our materials were presented too rapidly to detect them, or that the conflicts they generated were resolved too easily. For example, because the temporal adverbial was immediately followed by the ergative subject, it could be that there was not enough time to generate the prediction for a future-tense verb at the adverbial. In this case, the parser would be effectively receiving the adverbial cue and the ergative marker cue simultaneously and trying to use them jointly to predict the continuation. Or, it could be that the one low-probability form consistent with both the future adverbial and the ergative marker, namely, the optative, was still accessible enough that it was not especially costly to shift from a future-tense prediction to an optative prediction when the ergative-marked subject was encountered. To address this possibility in future work, it would be helpful to start with simple sentence-continuation experiments which could show how often Georgian speakers combine the ergative and the optative (see also fn. 2).

Finally, it is important to acknowledge several shortcomings and errors in our experimental design that could suggest less interesting explanations for our failure to observe a cost for prediction conflict. First, our original design included an additional pair of conditions designed to confirm standard processing difficulty effects when combining a future adverbial alone with an aorist verb, but errors in stimuli creation compromised this pair. Second, due to an unfortunate error, the grammatical control conditions with the nominative-marked subject and a future-tensed verb often continued with a conjoined second clause in the past tense, resulting in global unacceptability later in the sentence. This error occurred well after the analysis regions, and the late positivity observed at the critical verb shows that these post-critical errors did not lead to

participants adopting a strategy of 'giving up' on analyzing the sentences in general. However, it remains possible that these errors or other general properties of our stimuli set altered the predictive inferences that participants generated such that the results don't generalize to the case of natural language comprehension.

Converging evidence with analogous configurations is certainly needed, and evidence from behavioral designs may be especially helpful as there it is possible to present many fewer items per condition than in ERP. In a pilot online self-paced reading experiment (n=33) using similar items we have similarly observed no clear evidence of reading time slowdown associated with the cue conflict created by the future adverbial and the ergative-marked subject, even though those trials were now much rarer (5 out of 186 total sentences)[3].

*Conclusions*

We hope that this paper illustrates the untapped potential of languages with rich morphological systems like Georgian for investigating central questions about the language comprehension architecture. Here the properties of Georgian allowed us to ask an important and relatively unexplored question in psycholinguistics: do morphosyntactic predictive inferences from the context have the same cognitive status as direct inferences from the input? Although our conclusions are tentative, we hope our investigation points the way to further work on an important question for researchers that are committed to a predictive processing architecture.

*Acknowledgments*

This work was supported in part by NSF grant BCS-1619857 to Maria Polinsky, NSF grant BCS-17949407 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 and Cass Lowry for detailed comments on an earlier draft of this paper. We are grateful to Rusudan Asatiani, Zurab Baratashvili, Irina Lobzhanidze, Tamar Makharoblidze, and Irakli Salia for help with the Georgian data. We are indebted to the faculty of the Institute for Theoretical and Applied Linguistics at the Ivane Javakishvili Tbilisi State University for help with participant recruitment and the administration of the experiment. All errors are our responsibility.

*References*

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature neuroscience*, *8*(8), 1117-1121.

[3] It is worth noting that in the self-paced reading experiment, all items were designed to be acceptable, and therefore the five future-adverbial + ergative sentences continued with the acceptable optative verb. However, the rarity of these trials together with the presence of 20 fillers where the future adverbial was followed by the more frequent future tense, makes us dubious that participants learned an 'optative strategy' within the experiment.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, *134*(1), 9-21.

Dillon, B., Nevins, A., Austin, A. C., & Phillips, C. (2012). Syntactic and semantic predictors of tense in Hindi: An ERP investigation. *Language and cognitive processes*, *27*(3), 313-344.

Hammerly, C., Staub, A., & Dillon, B. (2019). The grammaticality asymmetry in agreement attraction reflects response bias: Experimental and modeling evidence. *Cognitive psychology*, *110*, 70-104.

Harris, A. C. (1981). *Georgian syntax: A study in relational grammar.* Cambridge: Cambridge University Press.

Klimov, G. A. & Alekseev, M.E. (1980). *Tipologija kavkazskix jazykov.* Moscow: Nauka.

Lau, E., Socolof, M., Clarke, N., Asatiani, R., Polinsky, M. (submitted). A subject relative clause preference in a split-ergative language: ERP evidence from Georgian.

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in human neuroscience*, *8*, 213.

Nash, L. (2017). The structural source of split ergativity and ergative case in Georgian. In J. Coon, D. Massam, & L. Travis (eds.) *The Oxford Handbook of Ergativity,* 175-203. Oxford: Oxford University Press.

Polinsky, M. (2020). Introduction. In M. Polinsky (ed.) *The Oxford Handbook of Languages of the Caucasus*, 1-25. Oxford: Oxford University Press.

Swaab, T. Y., Ledoux, K., Camblin, C. C., & Boudewyn, M. A. (2012). Language-related ERP components. In S. J. Luck & E. S. Kappenman (Eds.), *Oxford library of psychology. The Oxford handbook of event-related potential components* (p. 397–439). Oxford University Press.

Toscano, J. C., McMurray, B., Dennhardt, J., & Luck, S. J. (2010). Continuous perception and graded categorization: electrophysiological evidence for a linear relationship between the acoustic signal and perceptual encoding of speech. *Psychological Science* 21(10),1532-1540.

Urbach, T. P., DeLong, K. A., Chan, W. H., & Kutas, M. (2020). An exploratory data analysis of word form prediction during word-by-word reading. *Proceedings of the National Academy of Sciences*, *117*(34), 20483-20494.

Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, *61*(2), 206-237.