Title: Evidence for a universal parsing principle in Santiago Laxopa Zapotec

Authors: K. Sasaki[1], S. Foley[1†], J. Pizarro-Guevara[1‡], F. Silva-Robles[2], M. Toosarvandani[1*], M. Wagers[1*].

Affiliations:

[1]Department of Linguistics, University of California; Santa Cruz, CA 95064, USA.

[2]Senderos; Santa Cruz, CA 95062, USA.

*Corresponding authors. Email: mtoosarv@ucsc.edu (MT), mwagers@ucsc.edu (MW).

†Current: Program in Linguistics, Princeton University; Princeton, NJ 08544, USA.

‡Current: Department of Linguistics, University of Massachusetts; Amherst, MA 01003, USA.

Abstract: Language comprehenders make predictions about what they will hear next, using parsing principles to resolve ambiguities. One principle, Parse Subject, has been attested in many experiments; however, these have predominantly tested a small handful of subject-initial languages, undersampling the world's linguistic diversity. We tested the universality of Parse Subject in Santiago Laxopa Zapotec, a verb-initial language, through a series of eye-tracking studies. Speakers inspected illustrations while they listened to recordings and we tracked their eye movements. The distribution of visual attention over time showed that Parse Subject interpretations were indeed the earliest mid-sentence interpretations entertained by participants. Such interpretations were routinely defeated by subsequent information, indicating that it is a defeasible principle. These findings support the universality of Parse Subject by incorporating data from an under-investigated language whose speakers have not historically contributed to our understanding of the language processing architecture.

One-Sentence Summary: A series of eye-tracking studies on Santiago Laxopa Zapotec provides evidence for the universality of a syntactic parsing principle guiding human language comprehension.

Main Text: Human language comprehension is characteristically fast and incremental. Interpretation can lag the perceptual signal by as little as 200 ms (*1, 2*). To achieve this feat, humans integrate information from the preceding context with knowledge about the structure of their language to estimate features of the upcoming input (*3*). They also take calculated risks and use parsing strategies for ambiguity reduction (*4*). Because local ambiguities proliferate in natural language, comprehenders can either wait for disambiguating input, potentially for a long time, or make a more rapid analysis guided by parsing principles that resolve ambiguous input with a particular bias. Such principles may derive from language-specific statistics, including detailed distributional knowledge of how the words of a language are arranged (*5*). But are there any principles shared by all language users, despite the superficial differences in their languages?

We tested for the universality of one parsing principle, Parse Subject, by probing for its presence in a language that should not favor it: Santiago Laxopa Zapotec, an Oto-Manguean language spoken by approximately 1,000 people in the Sierra Norte of Oaxaca, Mexico. Parse Subject directs comprehenders to interpret ambiguous noun phrase (NP) constituents first as the grammatical subject, all else being equal. While it has been validated in a wide variety of language processing studies (*6*), these have sampled almost entirely from a handful of national languages with large speaker populations (*7*). Crucially, in almost all these languages, the subject is typically placed first in the sentence. Thus, it is possible that Parse Subject only appears to be universal, because the languages in which it has been tested are mostly subject-initial.

Parse Subject should not be a universal principle if its effects were the by-product of distributional cues in the language. For example, comprehenders may analyze less frequent (less canonical) sentence types as maximally parallel to more frequent (more canonical) sentence types (cf. *8*). Under this highly language-particular view, the apparent universality of Parse Subject as a principle could arise from sample bias. The preference for a subject interpretation would be a product of comprehenders' experience with the more canonical subject-initial sentences of the languages where it has been tested. In contrast, Parse Subject is expected to be used by all language comprehenders in a theory that assigns the grammatical subject inherent priority, regardless of its linear position in the sentence (*9*). Subjects are the most prominent constituent in a sentence: they typically outrank other constituents in many grammatical and pragmatic processes, and the interpretation of other constituents often depends on them *(10, 11)*. Every comprehender should therefore, according to this universal view, attempt to link ambiguous constituents first with the subject role, in the absence of better information.

If Parse Subject behavior emerged in Santiago Laxopa Zapotec, this would count as strong evidence for the universality of the parsing principle. The language uses rigid verb-subject-object word order (Fig. 1A) and does not indicate grammatical roles in any other way, with case or verb agreement. Constructing a relative clause (RC) consequently creates an ambiguity that could be resolved by Parse Subject: the noun modified by an RC can be interpreted as either its subject or object (Fig. 1B). Parse Subject has been tested in the RCs of a few verb-initial languages, including the Mayan languages Ch'ol and Q'anjob'al (*12*) and the Austronesian languages Chamorro (*13*) and Niuean (*14*). However, these all have rich verb marking indicating the role of the modified noun. RCs in these languages are thus only sometimes ambiguous, unlike in Santiago Laxopa Zapotec.

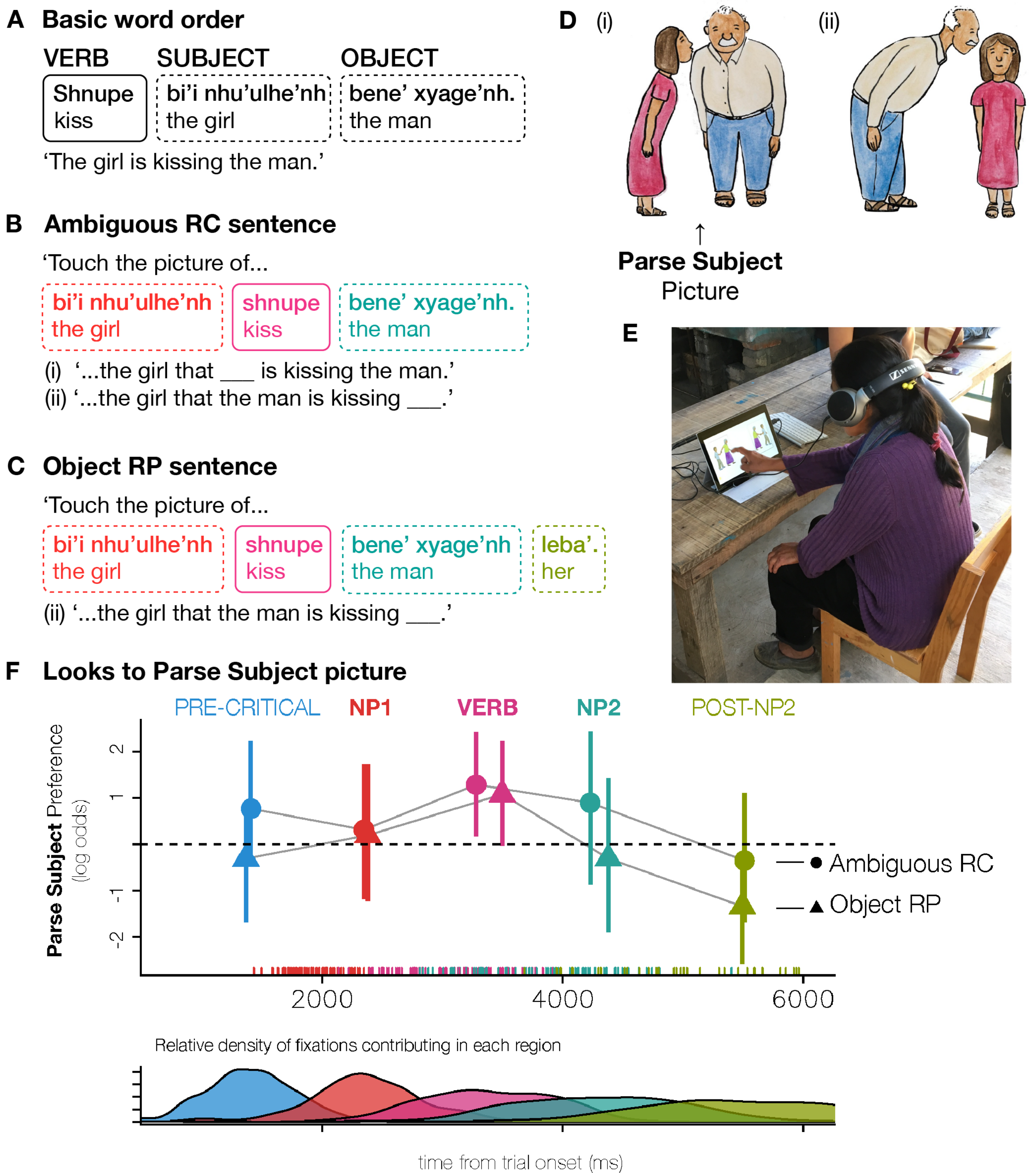

Fig. 1. Basic word order in Santiago Laxopa Zapotec with Exp. 1 design and eye-tracking results. Basic word order is fixed and unambiguous: the verb (solid box) is first, followed by the subject and then the object (dashed boxes) (**A**). In a relative clause (RC), the modified noun NP1 (orange) is first, followed by the verb (pink) and all other elements (**B**). NP1 could be either

subject or object, because these roles are not distinguished in any other way. RCs can be disambiguated using an object resumptive pronoun (ObjRP): NP1 must be interpreted as an object when an ObjRP (chartreuse) follows the post-verbal noun NP2 (turquoise) (**C**). In Exps. 1–4, participants heard sentences like those in (A)–(C), while viewing two illustrations: one depicted the Parse Subject interpretation for NP1, while the other depicted the reverse interpretation (**D**). Participants were seated with a tablet and eye-tracker situated upright on the table in front of them in a bright outdoor space; stimuli were played on over-ear headphones (**E**). In Exp. 1, participants showed an early preference to comprehend the ambiguous NP1 in RCs as a subject. In (**F**), the strength of the Parse Subject preference is represented on the *y*-axis as the difference, in log-odds, between looks to the Parse Subject picture and to the alternative depiction. Each plotting symbol corresponds to data acoustically bounded by critical linguistic events (NP1, VERB, NP2), with error bars marking 95% credible intervals. Inset ticks indicate acoustic onsets and the lower panel shows the distribution of fixation onsets contributing to each preference score.

We designed four experiments using picture matching with eye-tracking to test whether speakers of Santiago Laxopa Zapotec use Parse Subject. In Exp. 1, participants (n = 33) listened to recorded sentences containing an RC and were asked to choose the picture that was most compatible with what they heard (Fig. 1D–E). Participants ultimately selected the Parse Subject picture at chance (51%). There was considerable variation among participants and items, and thus we found no evidence for Parse Subject in picture selection [95% CrI: 21%, 81%]. By comparison, when participants heard unambiguous verb-subject-object sentences, they correctly selected the Parse Subject picture 91% of the time [87%, 99%].

Picture selection is only an endpoint measure of parsing, based on information integrated over a dynamically changing set of interpretations. The early output of Parse Subject could have been reanalyzed as listeners heard more of the sentence. To provide an incremental measure of parsing, we also recorded participants' eye movements to reveal early preferences as they listened to sentences. After hearing only NP1 and VERB, participants showed a robust preference for the Parse Subject picture (median log-odds difference in looks, 1.08 [0.21, 1.95]; Fig. 1F). We conclude that participants adopted Parse Subject interpretation before hearing further information.

In Exp. 1, we also included sentences whose structure mandate the reanalysis of initial Parse Subject decisions. RCs with object resumptive pronouns (ObjRPs) are unambiguous; the modified noun (NP1) is necessarily construed as the object (Fig. 1C). For these sentences, participants wrongly selected the Parse Subject picture 21% of the time [7%, 42%]. This error rate was twice the rate in unambiguous verb-subject-object trials, which suggests that participants parsed NP1 as a subject but were not always able to successfully reanalyze it at POST-NP2. We also found a characteristic cross-over pattern in fixation preferences: early looks to the (incorrect) Parse Subject picture (0.81 log-odds, [-0.06, 1.79]) were followed only later by looks to the correct picture (-1.27 log-odds, [-2.09, -0.44]; Fig. 1F).

We considered whether the results of Exp. 1 depended on RCs containing human subjects and objects, since animacy plays a potent role in parsing (*15*). The use of only high-animacy referents may have driven the early looking preferences. We conducted two follow-up experiments, identical to Exp. 1 except that sentences described either: (i) a human and an animal (Exp. 2, n = 34), or (ii) a human and an inanimate object (Exp. 3, n = 30). In both Exps. 2 and 3, we found nearly identical results to Exp. 1. Picture selection responses were equivocal between the two alternative interpretations for RCs (Exp. 2: 56% [40%, 71%]; Exp. 3: 49% [32%, 65%]). But, as in Exp. 1, the eye movement record revealed a clear preference for the Parse Subject picture at VERB (Fig. 2).

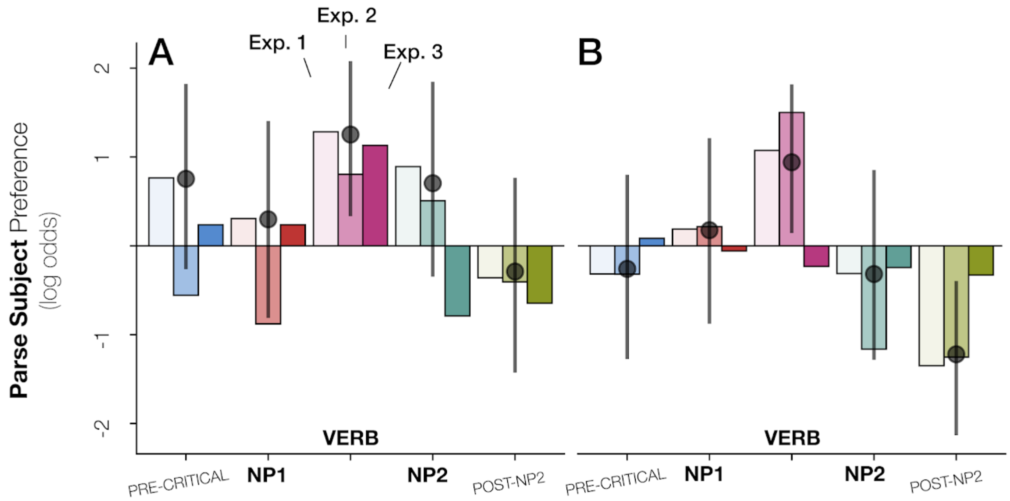

Fig. 2. Parse Subject looking behavior emerges in Exps. 1–3. Parse Subject looking behavior emerged in all experiments at VERB. When listening to VERB, after the ambiguous NP1, participants' attention was preferentially attracted to the Parse Subject picture. The height of the bars correspond to the Parse Subject preference. Filled symbols represent model medians across experiments for each region, with error bars for 95% Credible Intervals. There was less variability across experiments in the Ambiguous RC conditions (**A**) compared to the ObjRP conditions (**B**). In Exp. 3, for ObjRP conditions only, the evidence for early Parse Subject looks was not strong. (Uncontrolled) prosodic cues may have been present in the acoustic signal to guide the listener. In NP2 and POST-NP2 regions, a cross-over anti-Parse Subject preference emerges in ObjRP conditions.

The results of Exps. 1–3 imply that Parse Subject commitments are defeasible and that subsequent information can stochastically prompt their reanalysis. In our experiments, a central motivator for reanalysis may be the language-particular pressure to interpret the immediately post-verbal NP constituent as the subject. This pressure has been observed in at least one other verb-initial language, Chamorro (*13*), and conjectured to hold for another Zapotec language (*16*). Indeed, across Exps. 1–3, we observed a cross-over to an object reading in NP2 and POST-NP2 regions, not only for ObjRP sentences (Fig. 2B) whose syntax forces reanalysis, but also for Ambiguous RC sentences (Fig. 2A).

If reanalysis is responsible for the cumulative neutralization of early Parse Subject interpretations, then making reanalysis harder should increase Parse Subject interpretations. Reanalysis should be harder if NP1 and NP2 are semantically similar (cf. *17*). In Exps. 1–3, they always belonged to distinct animacy categories. In Exp. 4 (n = 103; picture selection only), we tested sentences in which NP1 and NP2 belonged to the *same* animacy category. We found a small, but systematic increase in Parse Subject interpretations. Ambiguous RCs elicited 63% Parse Subject interpretations [53%, 72%], compared to 51% across Exps. 1–3. ObjRP sentences in Exp. 4 erroneously elicited Parse Subject interpretations on 46% of trials [34%, 57%], compared to 32% across Exps. 1–3. These results thus show that an early Parse Subject interpretation can be entrenched by making reanalysis harder.

There is an alternative analysis of participants' early Parse Subject preference, compatible only with a language-particular view of Parse Subject. Comprehenders could be tracking word orders at a finer-grain size, using a noncanonical order to guide parsing. Non-verb-initial order is possible when a subject or object is fronted for focus (leaving a gap in argument position) or topic (resumed by a pronoun). If comprehenders tracked all such NP1–VERB–NP2 constructions, and if NP1 was more often a subject than an object, this information alone could prompt their early Parse Subject decisions. But non-verb-initial clauses, which are rare in production, do not show any bias toward mapping NP1 to the subject. In a spontaneous speech corpus of 1,942 sentences from the closely related San Bartolomé Zoogocho variety (*18*), 88% of sentences are verb-initial. In only a small percentage did NP1 precede the verb, as a subject in 5% and an object in 5%. In transitive sentences only, there was, in fact, an object bias: 14% were object-initial and 5% were subject-initial (81% were verb-initial). We conclude that, even if speakers were tracking word order statistics at a finer-grain, there is no evidence for a distributional bias toward NP1 subjects in production.

A second language-particular alternative involves participants accessing distributional information from their knowledge of Spanish, since they live in a bilingual Zapotec-Spanish context. Spanish comprehenders show Parse Subject incrementally (*19*), and in at least one study of a verb-initial language, this was shown to affect cumulative preferences. In Ch'ol, a Mayan language with some ambiguous RCs, (*12*) reported a stronger Parse Subject effect in picture selection than us (68% ± 2% SEM). However, when only monolingual Ch'ol speakers were considered, the result (59% ± 6%) was closer to our own. In our studies, 40% of participants self-reported using Spanish in several important facets of their life. When the analysis of eye movements was restricted to exclude those participants, we still found significant Parse Subject looking behavior at VERB in both Ambiguous RC (1.15 [0.37, 1.90]) and Object RP (0.97 [0.18, 1.72]) conditions (fig. S4). Therefore, our results are unlikely to be driven by transfer from participants' Spanish knowledge.

This study provides novel support for Parse Subject and shows that human language comprehension is strongly incremental, even for users of lesser studied, minority or indigenous languages. Including these language users incorporates not only cognitive diversity, but also grammatical diversity into psycholinguistic research. The unique phonological, morphological, syntactic, and semantic properties of their languages have the potential to clarify ongoing debates about how classes of information interact to support resilient language processing (*2–4, 15*) and how language learners develop that capacity (*20*).

References and Notes

1. W. D. Marslen-Wilson, Sentence perception as an interactive parallel process. *Science* **189**, 226–228 (1975).
2. E. Lau, C. Stroud, C. Plesch, C. Phillips, The role of structural prediction in rapid syntactic analysis. *Brain and Language* **98**, 74–88 (2006).
3. T. A. Farmer, M. Brown, M. K. Tanenhaus, Prediction, explanation, and the role of generative models in language processing. *Behavioral and Brain Sciences* **36**, 211–212 (2013).
4. L. Frazier, C. Clifton, *Construal* (MIT Press, Cambridge, MA, 1996).
5. J. Hale, A probabilistic Earley parser as a psycholinguistic model. *Second Meeting of the North American Chapter of the Association for Computational Linguistics* (2001).
6. N. Kwon, Y. Lee, P. C. Gordon, R. Kluender, M. Polinsky, Cognitive and linguistic factors affecting subject/object asymmetry: An eye–tracking study of prenominal relative clauses in Korean. *Language* **86**, 546–482 (2010).
7. P. Anand, S. Chung, M. Wagers, “Widening the net: Challenges for gathering linguistic data in the digital age” (Response to Dear Colleague Letter for SBE 2020: Future Research in the Social, Behavioral & Economic Sciences, 2011).
8. T. G. Bever, “Cognitive basis for linguistic structures” in *Cognition and the Development of Language* (Wiley, New York, 1970), pp. 279–362.
9. W. O’Grady, *Syntactic Development* (University of Chicago Press, Chicago, 1997).
10. E. L. Keenan, B. Comrie, Noun Phrase Accessibility and Universal Grammar. *Linguistic Inquiry* **8**, 63–99 (1977).

11. We do not distinguish among alternative versions of Parse Subject that refer to syntactic relation, argument marking or semantic role, because our experiments only tested transitive verbs. What matters for our purposes is that there is a prominence-based parsing principle that operates to incrementally disambiguate. See *(12)* for further discussion.

12. L. E. Clemens, J. Coon, P. Mateo Pedro, A. M. Morgan, M. Polinsky, G. Tandet, M. Wagers, Ergativity and the complexity of extraction: A view from Mayan. *Natural Language and Linguistic Theory* **33**, 417–467 (2015).

13. M. W. Wagers, M. F. Borja, S. Chung, Grammatical licensing and relative clause parsing in a flexible word-order language. *Cognition* **178**, 207–221 (2018).

14. R. Tollan, D. Massam, D. Heller, Effects of case and transitivity on processing dependencies: Evidence from Niuean. *Cognitive Science* **43**, e12736 (2019).

15. L. Wang, E. Wlotko, E. Alexander, L. Schoot, M. Kim, L. Warnke, G. R. Kuperberg, Neural evidence for the prediction of animacy features during language comprehension: Evidence from MEG and EEG Representational Similarity Analysis. *Journal of Neuroscience* **40**, 3278–3291 (2020).

16. J. Foreman, P. Munro, Reconstructing a typologically marked strategy for Zapotec relative clauses. *Proceedings of the High Desert Linguistics Society* **6**, 141–154 (2004).

17. P. C. Gordon, R. Hendrick, M. Johnson, Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **27**, 1411–1423 (2001).

18. A. Sonnenschein, *A Descriptive Grammar of San Bartolomé Zoogocho Zapotec* (University of California, Los Angeles, 2004).

19. M. Betancort, M. Carreiras, P. Sturt, The processing of subject and object relative clauses in Spanish: An eye-tracking study. *Quarterly Journal of Experimental Psychology* **62**, 1915–1929 (2009).

20. J. D. Fodor, Learning to parse? *Journal of Psycholinguistic Research* **27**, 285–319 (1998).

21. S. Mathôt, D. Schreij, J. Theeuwes, OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods* **44**, 314–324 (2012).

22. E. S. Dalmaijer, S. Mathôt, S. Van der Stigchel, PyGaze: An open-source, cross-platform toolbox for minimal-effort programming of eye tracking experiments. *Behavior Research Methods* **46**, 913–921 (2014).

23. T. von der Malsburg, Saccades: An R package for detecting fixations in raw eye tracking data. https://github.com/tmalsburg/saccades (2019).

24. T. Benaglia, D. Chauveau, D. R. Hunter, D. Young, mixtools: An R package analysing finite mixture models. *Journal of Statistical Software* **32***,* 1–29 (2009).

25. Y. Huang, J. Snedeker, Evidence from the visual world paradigm raises questions about unaccusativity and growth curve analysis. *Cognition* **200**, 104251 (2020).

26. P. Bürkner, brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* **80**, 1–28 (2017).

27. D. Lewandowski, D. Kurowicka, H. Joe, Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis* **100**, 1989–2001 (2009).

28. M. Toosarvandani, "On reaching agreement early (and late)" in *Asking the Right Questions: Essays in Honor of Sandra Chung* (Linguistics Research Center, Santa Cruz, CA, 2017), pp. 127–141.

29. I. Sichel, M. Toosarvandani, Pronouns and attraction in Sierra Zapotec. *Syntax and Semantics at Santa Cruz* **4**, 97–116 (2020).

Acknowledgments: We are grateful to the many residents of Santiago Laxopa who participated in this project, and particularly to Alberto Díaz Robles, Raquel Díaz Robles, Raul Díaz Robles, Rosario Reyes Vázquez, Isidro Vázquez Jerónimo, and two other native speakers of Zapotec for teaching us about their language. We also thank our research assistants, Brianda Caldera and Azusena Orozco, and our illustrator, Roque Reyes Mendoza.

Funding:

National Science Foundation grant BCS-2019804 (MT, MW)

The Humanities Institute, University of California, Santa Cruz (MT, MW)

Committee on Research, University of California, Santa Cruz (MT, MW)

Vice Chancellor for Research, University of California, Santa Cruz (MT, MW)

Author contributions:

Conceptualization: KS, SF, JPG, MT, MW

Data Curation: KS, MT, MW

Formal Analysis: MW

Funding Acquisition: MT, MW

Investigation: KS, SF, JPG, FSR, MT

Project Administration: MT, MW

Resources: FSR

Software: JPG, MW

Supervision: FSR, MT, MW

Visualization: MW

Writing: MT, MW

Competing interests: Authors declare that they have no competing interests.

Data and materials availability: All data are available in the main text or the supplementary materials.

Supplementary Materials

Materials and Methods

Supplementary Text

Figs. S1 to S4

Materials and Methods

Participants

105 speakers of Zapotec were recruited across Exps. 1–3 (*ns* = 34; 36; 35, respectively) and 103 speakers in Exp. 4, some of whom overlapped with Exps. 1–3 participants. Data were collected in July 2018 and 2019 on two visits to the town of Santiago Laxopa (Oaxaca, Mexico), where nearly all speakers grew up and lived. Participants were recruited via word of mouth and public announcement, and they took the experiment in a quiet outdoor space at the local elementary school. They received small gifts in exchange for participation (an umbrella or raincoat and a small flashlight, totalling an approximate cash value of $10). Each participant responded to a brief language background questionnaire to probe what languages they used in different contexts (at home, at school, with family, with friends, and with co-workers). Participants in Exps. 1–3 ranged in age from 21 to 79 years (median: 41); in Exp. 4, from 18 to 78 years (median: 37).

Sentence materials and illustrations

In Exps. 1–3, sentences were created in factorial design that crossed three relative clause types (3) with the relative order of two NPs (NP1>NP2, NP2>NP1). The RC types were: V-S-O baseline (adjunct RC) *where-V-NP1-NP2*; Ambiguous RC, *NP1-(that)-V-NP2*; and Object RC, *NP1-(that)-V-NP2-RP*. 18 itemsets of 6 conditions (108 sentences tokens) were created in each experiment, and they were distributed via Latin Square in a within-subjects, within-items design. Based on target lexical items drafted in English, sentences were iteratively elicited in Zapotec in 1:1 fieldwork consultation between FSR and SF, JPG, KS, or MT. FSR produced the final versions of each sentence, and was recorded using a sound-attenuated booth (40-70 dB reduction) using a Zoom H4n Handy solid-state recorder with a AKG C520 head-worn condenser microphone.

Culturally-appropriate illustrations to depict each sentence set were created by Roque Reyes Mendoza (https://www.roquereyes.com/). Each itemset required three illustrations: two that depicted the two construals of the Ambiguous RC (NP1 acting on NP2; or NP1 acting on NP2) and one that neutrally established the two individuals mentioned in the trial. See Fig S1. All itemsets (text, audio) and illustrations are publicly available, see “Code availability” below.

Exp. 4 was similar to E1 except eye movements weren’t recorded. In lieu of a V-S-O baseline, we used a subject RP baseline (results not reported in manuscript).

Method

Exps. 1–3 were presented on Microsoft Surface Pro tablets running a custom Python script developed and deployed in OpenSesame (*21*). Eye movements in Exps. 1–3 were collected with a Tobii Pro Nano (Tobii AB, Sweden), sampling at 60Hz and controlled via the PyGaze library (*22*). After participants gave verbal consent and answered a few questions about their language background, the experiment began. First the eye tracker was calibrated, and then participants received instructions and completed five practice trials. The 18 target trials were interspersed randomly with 15 filler trials and organized into two blocks separated by a break screen. See Figure S1 for details of each trial structure. The average total duration of the two blocks was 13.5 minutes. At the conclusion of the experiment, participants were debriefed about their experience.

Exp. 4, also deployed in OpenSesame, was presented on Nexus 10 tablet using the touch tracking method and software developed by (*13*). No eye movements were recorded but the trial structure was the same as in Exps. 1–3.

Analysis

Eye movement recordings (Exps. 1–3) were pre-processed first by excluding all eye tracker samples without a TRUE-valued validity code in either the left or the right eye and then by parsing into fixations using the *saccades* library (*23*) in R (R Foundation for Statistical Computing, Vienna, Austria). We did not apply a trackloss criterion to trials. Fixations were then categorized, for each participant, into three areas of interest (AOI: left picture, right picture, center) using the EM algorithm for mixture of univariate normals implemented in the *mixtools* library (*24*). See fig. S2. Across Exps. 1–3, 8 participants could not be analyzed because of a problem recording their eye movements or clearly identifying AOI in the data, leaving 97 participants in the final sample (*n*s = 33; 34; 30; respectively in Exps. 1–3). In Exp. 2, two itemsets using the verb *bsegu* ('trip') were excluded from analysis because subsequent fieldwork revealed that it was likely intransitive. In Exp. 3, three itemsets (5, 6, 17) were excluded from analysis because of low accuracy on V-S-O controls in the choice task.

A set of five region-of-interest (ROI) intervals defined the critical regions in Ambiguous RC and Object RP sentences: *Pre-critical* (NP1 onset, less 1000 ms), *NP1* (NP1 onset to NP1 offset), *VERB* (V onset to V offset), *NP2* (NP2 onset to NP2 offset) and *Post-NP2* (NP2 offset, plus 1000 ms. See fig. S2.

The dependent variable in our eye-tracking analysis was whether or not participants were looking at the Parse Subject picture during each of the five intervals. Left and right pictures were mapped onto Parse Subject picture or Non-Parse Subject picture (which varied by trial). Within each trial, fixation onsets and offsets were multiply coded with respect to these intervals. If more than 30% of a fixation's duration fell within an interval, it was counted as a look in that interval. This follows the analytic strategy of (*25*), who observe that this 'binarization' analysis is appropriate for small time windows because it better matches the dynamics of how a picture is inspected: saccades only typically occur once or twice a second and our acoustic ROI, across experiments, range between 650-708 ms for the *VERB* ROI; and 748 - 1165 ms for either *NP1* or *NP2* ROI. See fig. S2.

The binary dependent variable- i.e., whether or not a participant was looking at the Parse Subject picture during each interval- was modeled in a single mixed-effects Bayesian logistic regression with three population-level predictors and all their higher-order interactions: GapType (whether the RC was a Ambiguous RC or an ObjRP; sum coded), FirstArg (which of two referents NP1 mapped onto; sum coded), and ROI (Pre-critical, NP1, VERB, NP2, and Post-NP2; ordered and Helmert-coded). Full random effects structure grouped by Subject and Item was included. Models were estimated using the *brms* library (*26*). We used the default priors in *brms* (v. 2.13.0), with the exception of the intercept, whose prior was set to Normal(0,1); and the correlations between random effects, set to the Lewandowski-Kurowicka-Joe prior *lkj*(eta = 2) (*27*). Reported point estimates are medians over the posterior distributions, and credible intervals are highest posterior density intervals (95%). The same strategy was used for analyzing the picture selection data in Exps. 1-4, although the models did not include ROI [since there is only one picture selection judgment per trial].

Figure 1 represents the resulting model for Exp. 1 by computing the Parse Subject Preference directly from the posterior samples. The Parse Subject Preference is the difference, expressed in log-odds, between (a) looks to the Parse Subject picture for NP1>NP2 trials (=Parse Subject "hits"); and (b) looks to the same picture for NP2>NP1 trials (=Parse Subject "false alarms"). The following describes this calculation algorithmically, using English translations of Zapotec sentences:

<u>Calculating the **Parse Subject Preference**</u>

Estimate a multilevel model with the following formula and priors described above:

```
critical_look ~ 1 + RC_Type*FirstArg*ROI +
                              (RC_Type*FirstArg*ROI|Subject) +
                              (RC_Type*FirstArg*ROI|Item)
```

`critical_look` is defined as whether or not there is a look, for that condition in that ROI, to the picture where the Non-elder HUman participant is acting as the Subject. `RC_Type` refers to whether the sentence is an Ambiguous RC or contains an Object RP. ROI refers to the acoustic ROI defined above and illustrated in Figure S3. `FirstArg` refers to NP1, the HU participant, precedes NP2, or vice versa. In Exp. 1, NP2 = ELder human; Exp. 2, NP2 = ANimal; Exp. 3, NP2 = INanimate. `*` here is conventional shorthand for the full interaction (A*B => A + B + A:B).

This model is fit with *brms*/Stan, with the priors described in text above.

Take the posterior samples (n = 4000) for each coefficient and use the regression equation to calculate, for each ROI and RC type, the preference in logits to look at the picture where Subject = HU for a sentence where NP1 > NP2 and NP1 = HU.

(i) the-girl.**NP1** [call.**VERB** the-man.**NP2** …]

Subtract the preference to look at the *identical* picture, where Subject = HU, for a matched sentence where NP2 > NP1 (and NP1 = HU).

(ii) the-man.**NP2** [call.**VERB** the-girl.**NP1** …]

Figure 2 plots this score, with the median Parse Subject Preference as bar heights. The dark points and 95% CrI represent a fourth model that aggregates data across all experiments, and includes Experiment as a fourth fixed effect.

Code availability

All materials and code, including audio recordings (.wav), illustrations (.png), presentation scripts (OpenSesame), plaintext data (.csv), R Markdown analysis scripts, and de-identified participant characteristics (.xlsx) are archived at https://osf.io/2wgd8/ (DOI: 10.17605/OSF.IO/2WGD8).

Figures

Figures S1-S4

- S1 demonstrates the timing details of the trial.
- S2 exhibits the analytic strategy for coding fixations into visual areas of interest (AOIs).
- S3 exhibits how acoustic regions were parsed into regions of interest (ROIs).
- S4 shows the effect of Zapotec language balance on the Parse Subject Preference.

Supplementary Text

Relative clauses (RCs) are clauses that modify a noun. Santiago Laxopa Zapotec exclusively has head-external RCs: the modified noun (or "head" nouns), only appears preceding the RC, which is enclosed in square brackets in (1).

(1)	**bi'i**	**nhu'ulh=e'nh**	[shlill	bene'	xyag=e'nh]
	child	**female=DEF**	call.CONT	person	male=DEF

'the girl that called the man'
'the man that called the girl'

There is no overt relative pronoun (e.g., *who* or *which)* or complementizer (e.g., *that*) that intervenes. The RC can follow the head noun directly.

The role of the head noun inside the RC, as subject or object, can be indicated by a gap. Since Santiago Laxopa Zapotec has basic verb-subject-object order, a RC like (1) is ambiguous. The head noun could be the subject (and *bene' xyage'nh* 'the man' the object) or it could be the object (and *bene' xyage'nh* the subject). Disambiguation is, in principle, possible by adding a resumptive pronoun (RP). These are pronouns in an argument position, as a subject (2a) or an object (2b), interpreted as referring to the head noun.

(2)	a.	**bi'i**	**nhu'ulh=e'nh**	[shlill=**ba'**	bene'	xyag=e'nh]	
		child	**female=DEF**	call.CONT=**3.HU**	person	male=DEF	
		'the girl that called the man'					
	b.	**bi'i**	**nhu'ulh=e'nh**	[shlill	bene'	xyag=e'nh	**leba'**]
		child	**female=DEF**	call.CONT	person	male=DEF	**3.HU**
		'the girl that the man called'					

For a head noun linked to a different argument position within the RC either a gap or a RP is possible. In (3a), the head noun is the object of a preposition (*tse* 'of'), and in (3b), it is the possessor of *xhikw* 'dog'.

(3)	a.	**bi'i**	**nhu'ulhe**='nh	zike	wiya'anh	tse=**ba'**	
		child	**female**=DEF	love.CONT	dancer	of=**3.HU**	
		'the girl that the dancer loves'					
	b.	**bi'i**	**xyag=e'nh**	udi'in	xhikw=**ba'**	bi'i	nhu'ulhe='nh
		child	**male=DEF**	bite.COMP	dog=**3.HU**	child	girl=DEF
		'the boy whose dog bit the girl'					

Since RP in Santiago Laxopa Zapotec are ordinary pronouns, those in (2–3) are, in principle, interpretable as referring to some other individual, as long as the head noun can still be linked to another argument position inside the relative clause

The resumptive pronouns are drawn from the inventory of third person pronouns, which in Santiago Laxopa Zapotec exhibit a four-way animacy contrast. Elder humans are distinguished from non-elder humans, which are, in turn, distinguished from both animals and inanimates (*28*).

(4)

	Clitic	Strong
3.EL(DER)	*=e'*	*le'*
3.HU(MAN)	*=ba'*	*leba'*
3.AN(IMAL)	*=(e)b*	*leb*
3.IN(ANIMATE)	*=(e)nh*	*lenh*

These pronouns also come in two series: "clitic" and "strong" pronouns. Clitic pronouns are weak elements, which attach to the verb; strong pronouns can stand on their own. As shown in (2a), a subject pronoun must be a clitic. An object pronoun can only be a clitic, however, if the subject also is (*29*). Thus, in (2b), the object pronoun is a strong pronoun.

While there are no relative pronouns in Santiago Laxopa Zapotec, it is possible for a "classifier" to appear between the head noun and the relative clause, as in (5a–b).

(5) a. bi'i wegu'=nh **bi='nh** dzesbanh beku'=nh
child young=DEF **CL:HU=DEF** wake.CONT dog=DEF
'the young person that is waking up the dog'
'the young person that the dog is waking up'

b. beku'=nh **bi'a=nh** dzesbanh bi'i wegu'=nh
dog=DEF **CL:AN=DEF** wake.CONT child young=DEF
'the dog that is waking up the young person'
'the dog that the young person is waking up'

There is a unique classifier for each animacy category: *be* (elder), *bi* (human), *bi'a* or *be* (animal), and *de* (inanimate). The presence of this classifier does not affect the interpretation of the RC.

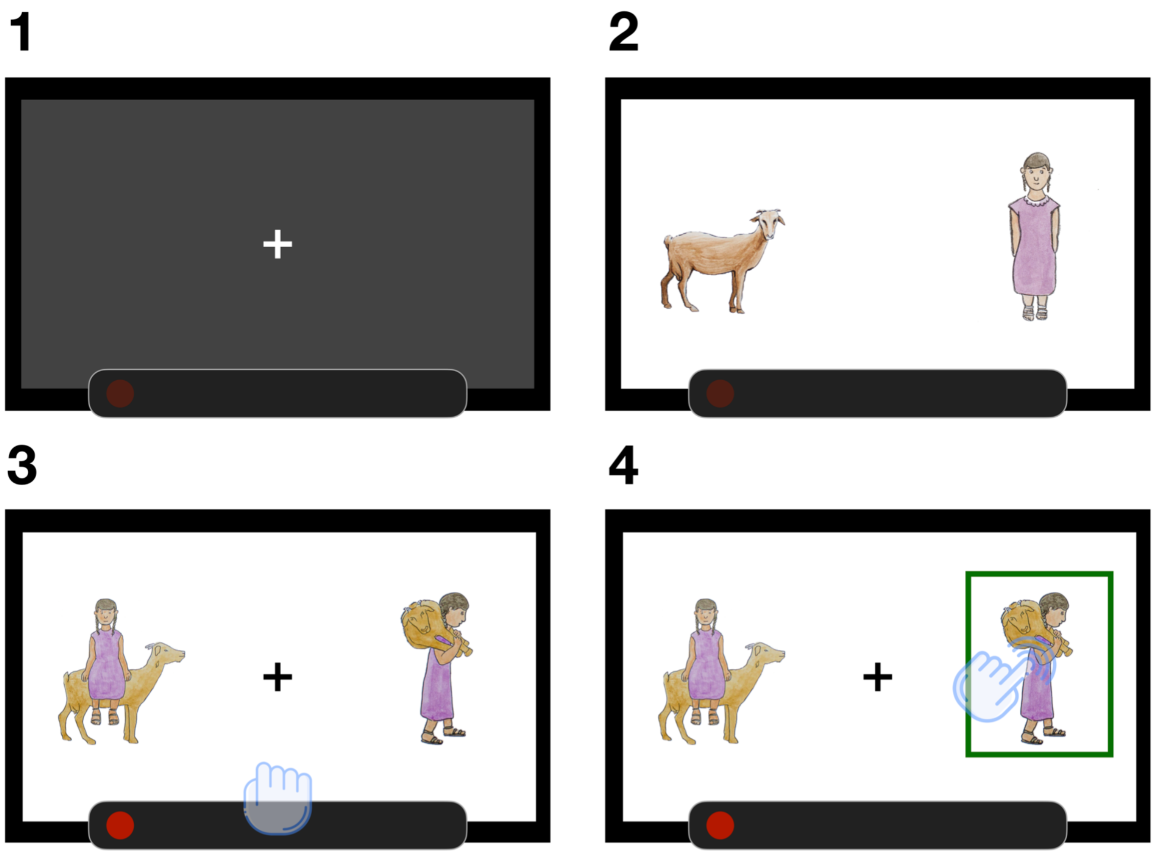

Fig. S1.
Each trial began with drift correction followed by (1) a blank screen for 1000 ms. (2) A context screen was displayed simultaneously with the audio onset of a context setting sentence. For example, the context sentence for this illustration was: *Nhi ze tu chiv nha' tu bi'i do'lhiz. Bi'a kafe chiv'enh. Bi'i do'lhize'nh yazba' vestidw moradw.* ("Here's a goat and a girl. The goat is brown. The girl is wearing a purple dress.") (3) Following the offset of the context sentence, the eye tracker began recording and the choice screen was displayed simultaneously with the audio onset of the target sentence. (4) Participants could select either picture at any point. Once they touched a picture, it was highlighted in a green outline. The trial advanced after 2000 ms had elapsed from picture selection, during which time participants could change their selection (and reset the timer).

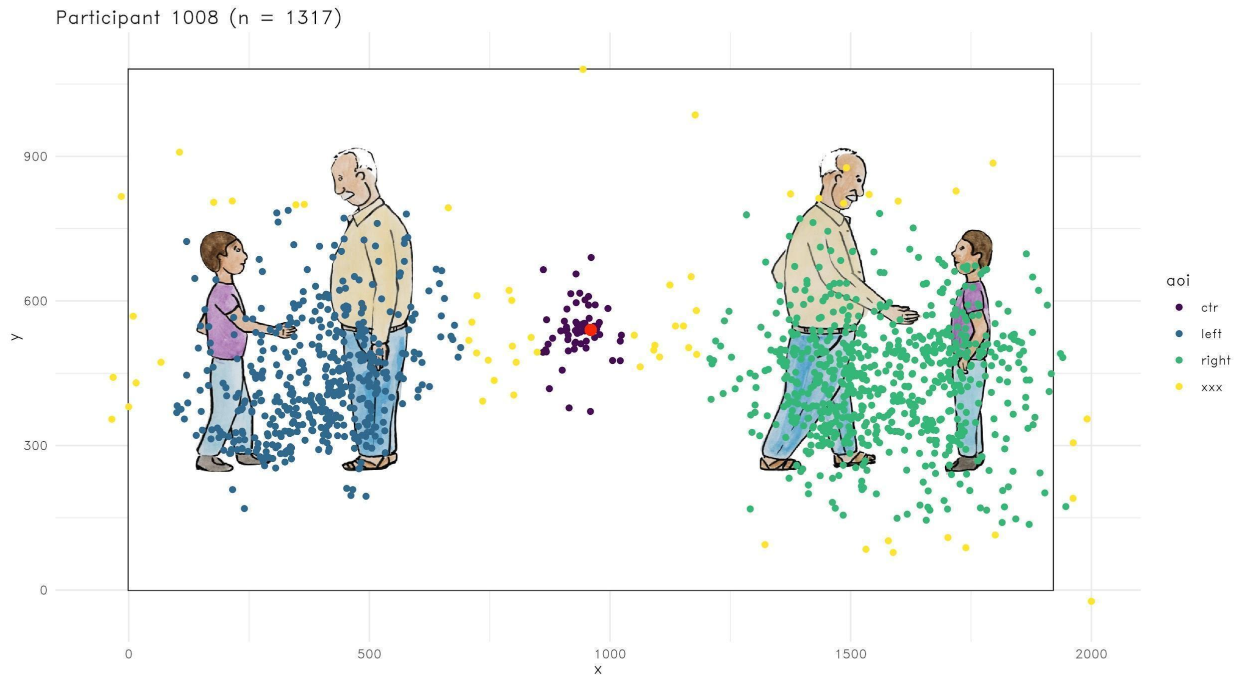

Fig. S2.

Participant 1008's fixations (n = 1317) recorded across an entire session and superimposed on a sample image. EM mixture modeling with three Normal distributions (univariate along the x-axis) categorized fixations into 3 AOI, indicated here by coloring ('left', 'right', or 'ctr'). AOI 'xxx' indicates points that were extreme in their mode (> 2 absolute standard deviates), and which we did not classify as belonging to either left, right or center AOIs. For five participants, a 3-mixture categorization did not produce a good fit; they were re-fit with 2-Normal mixture that only identifies left and right AOIs.

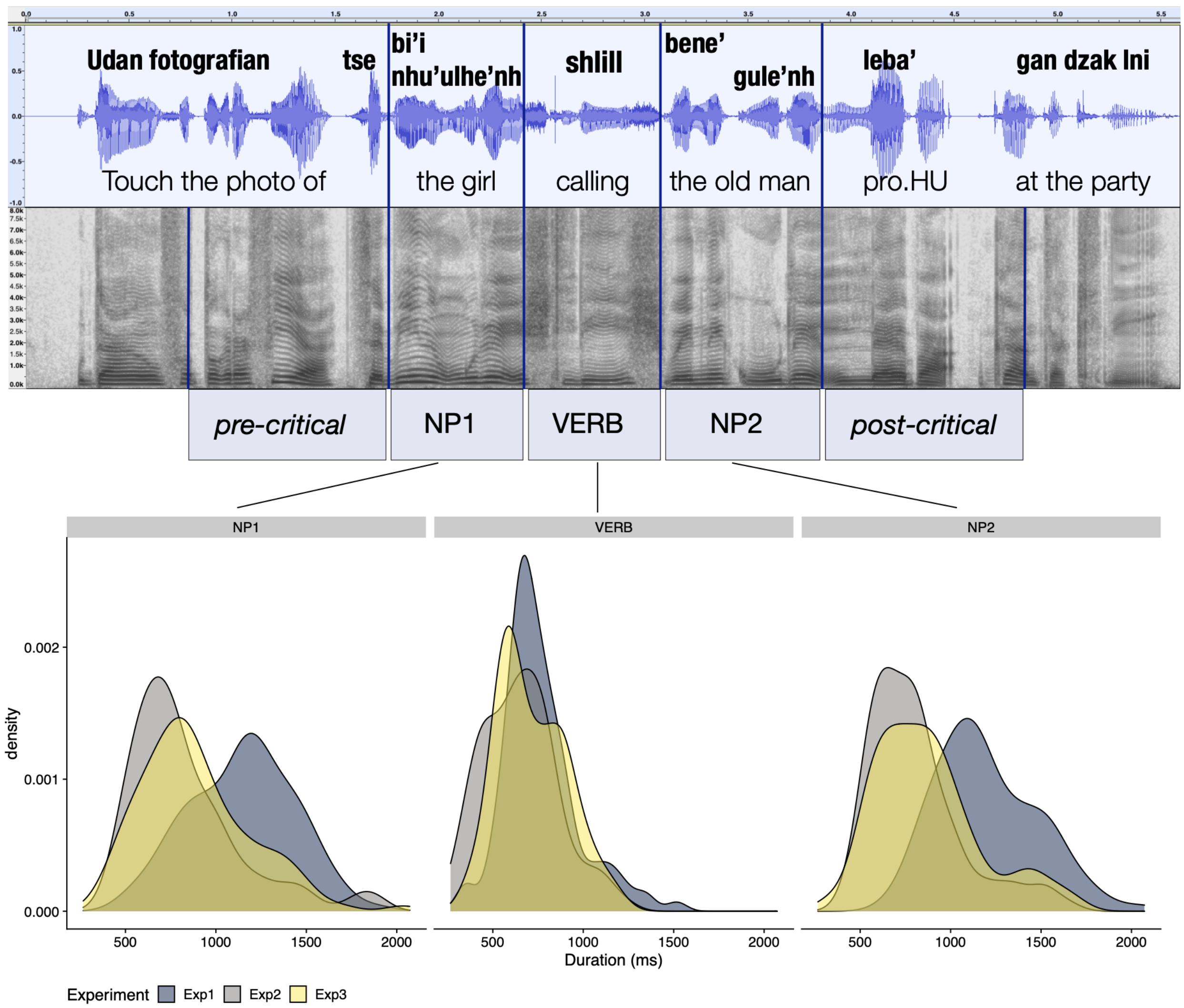

Fig. S3.

Core regions of interest were defined by the acoustic onsets and offsets of three linguistic events: *NP1*, *VERB*, and *NP2*. *Pre-critical* region extends 1000 ms prior to the onset of NP1 (Ambiguous RC and Object RP conditions) or Verb (V-S-O Baseline), and *Post-critical* region to 1000 ms post NP2 offset. Top panel illustrates segmentation for an example Object RP sentence. Bottom panel shows the distribution of ROI durations (by experiment). Note that NP1 and NP2 durations are characteristically longer in Exp. 1, because all Elder and Non-human Elder NPs begin with an obligatory classifier.

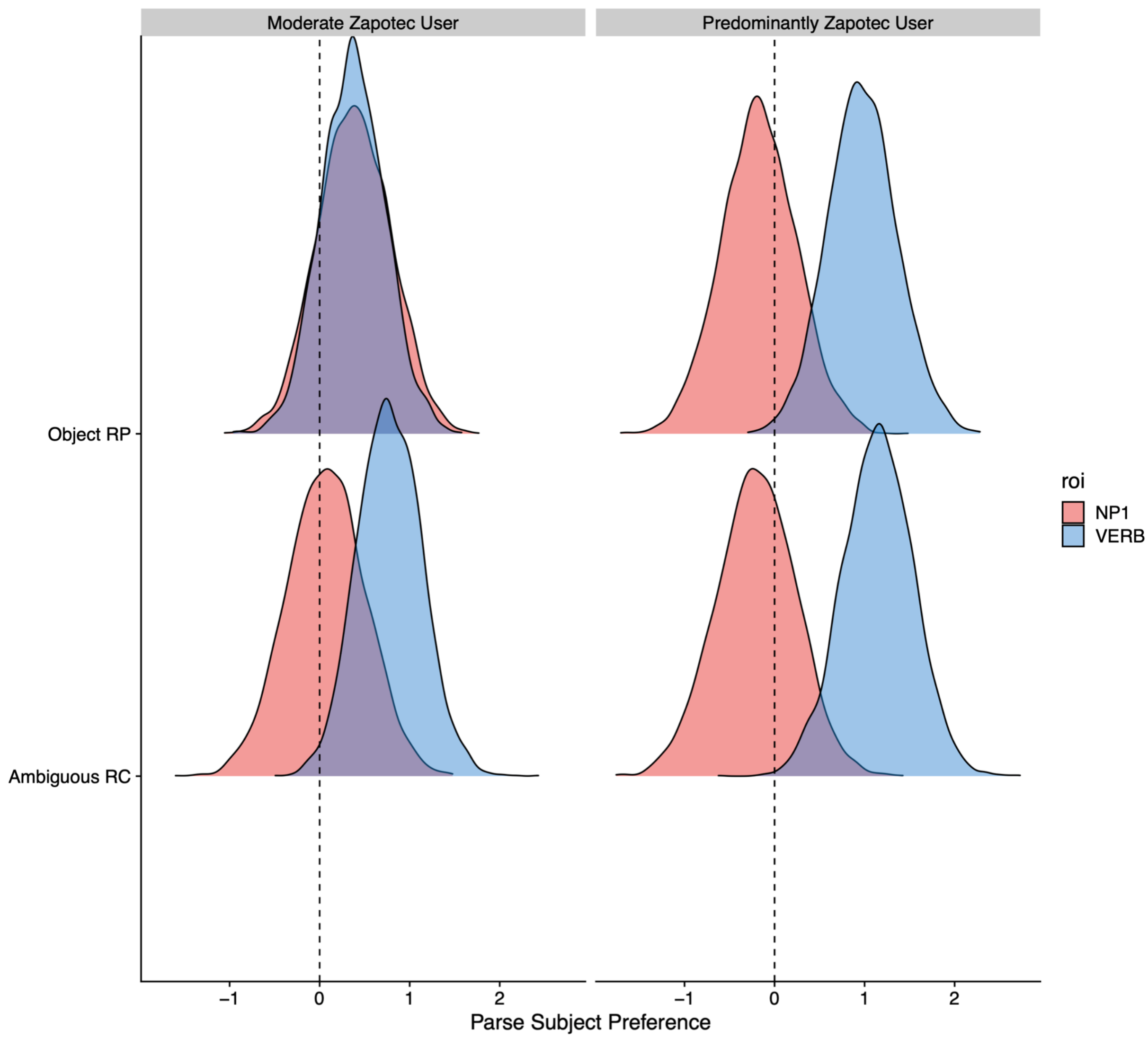

Fig. S4.

The Parse Subject Preference emerges most strongly at the Verb for participants predominantly use Zapotec (who self-report using Zapotec all, or nearly all, the time; a slight majority, 60%). The posterior sample densities above represent the same regression across Exps. 1–3 described in the text, but add as a fixed effect an indicator variable based on our self-reported Zapotec strength variable: *Moderate Zapotec Users*, left column, report either speaking or hearing Spanish in more than one aspect of their daily life (at home, with friends, at work) whereas Predominantly Zapotec Users, right column, only report at most one context in which they either speak or hear Spanish. There is a clear positive shift from no preference at NP1 (red; centered near 0, dashed line) to a strong Parse Subject Preference at the Verb (blue) for the Predominantly Zapotec Users in both Object RP and Ambiguous RC sentences. In contrast, Moderate Zapotec Users show a weaker shift, and only in Ambiguous RC sentences. This supports the claim that access to Spanish distributional information is not driving the Parse Subject effect, on the assumption that the Predominantly Zapotec Users have a weaker representation of that information.