

Gaze behaviour reveals how cross-linguistic differences in case alignment shape the dynamics of sentence planning

Sebastian Sauppe (sebastian.sauppe@uzh.ch)¹, Elisabeth Norcliffe², Agnieszka E. Konopka³,

Stephen C. Levinson⁴

¹University of Zurich, ²University of Surrey, ³University of Aberdeen, ⁴Max Planck Institute for Psycholinguistics

The time course of sentence planning can vary in response to, for example, extralinguistic factors such as the visual salience of referents [1]. In recent years, evidence has accumulated that languages' grammatical properties can exert a strong influence on sentence production processes [2–4]. It remains largely unknown, however, whether adaptations to grammatical affordances are the same across languages, mainly due to a lack of comparative cross-linguistic production research [4,5]. Here, we extend the field's empirical basis by investigating how differences in case marking alignment affect the time course of planning in Yélî Dnye (a language isolate of Papua New Guinea) and Japanese. While nominative-accusative languages use the same case (the nominative) to mark the agents of both transitive and intransitive verbs, ergative languages use a dedicated case (the ergative) to exclusively mark the agents of transitive verbs. Previous studies on Hindi and Basque found that planning sentences with ergative case alignment induces earlier and more extensive relational encoding at the outset of formulation because speakers must determine the “who-does-what-to-whom” structure of the event early to select the appropriate case marker on the initial argument [6,7]. However, in those languages, ergative case is conditioned by language-specific grammatical factors beyond syntactic transitivity (grammatical aspect in Hindi and lexical specification in Basque), leaving it unresolved whether it is ergativity itself which drives earlier relational encoding or another co-dependent grammatical property of the languages tested so far. We conducted an eye-tracked picture description task with speakers of Yélî Dnye, a language with strictly ergative case marking ($n = 45$), and speakers of Japanese, a language with nominative-accusative case marking ($n = 56$). The default word order in both languages is Agent-Patient-Verb (APV). Speakers of the two languages described 52 coloured cartoon scenes depicting transitive (two-participant) events while their eye movements were recorded (60 Hz sampling rate). Binomial mixed-effects growth curve regressions modelled the time course of visual attention to the agent and patient characters in the scenes, controlling for a range of nuisance predictors. In an early time window (200–800 ms), associated primarily with conceptual encoding, Yélî Dnye speakers prioritised relational encoding to a greater degree than Japanese speakers. This is evidenced by a weaker preference for fixations to agents by Yélî Dnye speakers, who distributed their visual attention more between the agents and other aspects of the depicted events (Figure 1). Differences between the languages are reflected in model comparisons favouring models that included a language predictor and its interactions with the polynomial time terms for agent fixations. Our results support the conclusion that it is ergative case marking per se that consistently induces extensive early relational coding (in Yélî Dnye, but also in Hindi [7] and Basque [6,8]). Thus, across diverse languages, the uptake of visual information relevant for utterance preparation is shaped by the affordances imposed by a language's grammar from the outset. Although the Papua and New Guinea area is home to a fifth of the world's languages [9], currently only a single published study on sentence processing in a Papuan language exists [3]. Our study thus also provides a window into the psycholinguistics of languages of the most linguistically diverse region of the world, broadening our understanding of how typological variation shapes the time course of language processing and thereby contributing to making psycholinguistic theory more robust.

References

- [1] Myachykov A, Thompson D, Garrod S, Scheepers C. Referential and visual cues to structural choice in visually situated sentence production. *Front Psychol.* 2011;2: 396.
- [2] Nordlinger R, Garrido Rodriguez G, Kidd E. Sentence planning and production in Murrinhpatha, an Australian “free word order” language. *Language.* 2022;98: 187–220.
- [3] Sarvasy HS, Morgan AM, Yu J, Ferreira VS, Momma S. Cross-clause planning in Nungon (Papua New Guinea): Eye-tracking evidence. *Mem Cognit.* 2023;51: 666–680.
- [4] Jaeger TF, Norcliffe EJ. The cross-linguistic study of sentence production. *Lang Linguist Compass.* 2009;3: 866–887.
- [5] Norcliffe E, Harris AC, Jaeger TF. Cross-linguistic psycholinguistics and its critical role in theory development: early beginnings and recent advances. *Language, Cognition and Neuroscience.* 2015;30: 1009–1032.
- [6] Egurtzegi A, Blasi DE, Bornkessel-Schlesewsky I, Laka I, Meyer M, Bickel B, et al. Cross-linguistic differences in case marking shape neural power dynamics and gaze behavior during sentence planning. *Brain Lang.* 2022;230: 105127.
- [7] Sauppe S, Choudhary KK, Giroud N, Blasi DE, Norcliffe E, Bhattamishra S, et al. Neural signatures of syntactic variation in speech planning. *PLoS Biol.* 2021;19: e3001038.
- [8] Isasi-Isasmendi A, Andrews C, Flecken M, Laka I, Daum MM, Meyer M, et al. The Agent Preference in Visual Event Apprehension. *Open Mind.* 2023;7: 240–282.
- [9] Palmer B, editor. *The languages and linguistics of the New Guinea area.* Berlin, Germany: De Gruyter Mouton; 2018.

Figures and Tables

Figure 1: Grand mean proportions of Yélî Dnye and Japanese speakers’ agent and patient fixations over time. Ribbons indicate the 95% confidence interval. The dotted vertical lines indicate the analysis time windows (200-800 and 800-2300 ms after stimulus onset); the solid vertical lines show the grand mean speech onsets for each language.

