

Variations in the choice of subject references in Lelepa, an Oceanic language

Yuchen Li

(Australian National University)

Keywords: Oceanic languages, Quantitative linguistics, Syntax, Corpus linguistics, Subject reference

In Lelepa, an underdescribed Oceanic language spoken in central Vanuatu, syntactic subjects are described as being expressed obligatorily through a subject proclitic at the beginning of the verb complex and optionally with a preverbal subject noun phrase (NP), headed by either a lexical noun or a free pronoun (see 1 and 2 below) (Lacrampe, 2014). However, evidence has suggested that subject proclitics are sometimes deleted. Hence, Lelepa features four different subject expressions: a sole proclitic, a lexical NP, a free pronoun, or zero subject, whose contexts are nonetheless not well understood. This usage-based study aims to find out what factors relate to the presence of lexical or pronominal NPs, what factors relate to the absence of subject proclitics, and how the findings accord with or differ from referential choice in other languages.

This study draws on a 57-minute sample of Lelepa natural speech, collected from 2006 to 2010 and archived in ELAR as part of a larger open-access corpus, for both inferential and descriptive statistical analyses (Lacrampe, 2017). Following additional coding of the transcribed data using the GRAID annotation schema (Haig & Schnell, 2014), two logistic regression models and a recursive partitioning analysis were implemented. Categorical fixed effects were selected based on the reference grammar (Lacrampe, 2014), cross-linguistic studies on pronominal variation in Oceanic languages (e.g. Meakins, 2015; Schnell & Barth, 2020), and preliminary observation of the data. Speaker was coded as a random effect where applicable.

The results support existing description that subject proclitics are the most prevalent subject expression, although they are also found deleted in the presence of a coreferential lexical NP. Subject NPs are commonly present when there is a referent switch, especially when the proclitic has an ambiguous form due to syncretism. Furthermore, referents with higher topicality, such as human or direct-speech referents, show a preference particularly for free pronoun references. Overall, Lelepa speakers' realisation of subject references mostly reflects the optimal choice that prevents both ambiguity and redundancy. This study complements existing description of Lelepa morphosyntax, and extends our typological understandings of referential choice in Oceanic languages. The data preprocessing and statistical modelling also demonstrate practical approaches to performing quantitative analyses of small and/or unbalanced datasets.

Examples:

1. *a=ga* *to* *wara-e* *to*.
1SG.NOM=IRR stay place-ADD STAT
'I will stay there.' (Lacrampe, 2014, p. 377)
2. *konou a=ga* *to* *wara* *to*.
1SG **1SG.NOM=IRR** stay here STAT
'I will stay here.' (Lacrampe, 2014, p. 160)

References

- Haig, G., & Schnell, S. (2014). *Annotations using GRAID (Grammatical Relations and Animacy in Discourse). Manual Version 7.0*.
https://www.academia.edu/10328418/Haig_Geoff_and_Stefan_Schnell_2014_Annotations_using_GRAID_Grammatical_Relations_and_Animacy_in_Discourse_Manual_Version_7_0
- Lacrampe, S. (2014). *Lelepa: Topics in the grammar of a Vanuatu language* [Ph.D. thesis, The Australian National University]. <https://openresearch-repository.anu.edu.au/handle/1885/12893>
- Lacrampe, S. (2017). *Description and documentation of Lelepa, an endangered language of Central Vanuatu*. Endangered Languages Archive. <http://hdl.handle.net/2196/00-0000-0000-000F-BAF7-9>
- Meakins, F. (2015). Not obligatory: Bound pronoun variation in Gurindji and Bilinarra. *Asia-Pacific Language Variation*, 1(2), 128–161. <https://doi.org/10.1075/aplv.1.2.02mea>
- Schnell, S., & Barth, D. (2020). Expression of anaphoric subjects in Vera'a: Functional and structural factors in the choice between pronoun and zero. *Language Variation and Change*, 32(3), 267–291.