

The Saliba/Logea language documentation project: data processing and its impact on linguistic analysis

Andrew Margetts & Anna Margetts

Saliba and Logea are two mutually intelligible dialects of an Oceanic language spoken in Milne Bay Province of Papua New Guinea. The project has been working towards a web-based database of texts with text-audio linkage and searchable annotations for certain types of information. In this presentation we discuss the impact that this treatment of texts can have on linguistic analysis of the data and demonstrate this on the basis of two research questions – changes in the use of the plural marker and the positioning of PPs.

In Saliba the plural marker *-ao* originally occurred only with NPs that have human referents but it is in the process of spreading to mark NPs more generally. Older speakers state that young people incorrectly use the plural marker when they apply it to nouns with non-human referents or to possessive constituents. The structure of the text database allows us to sort examples by speaker and investigate what age group uses the plural marker in this extended way. First results show that the extended use of the plural marker is found mainly but not exclusively with younger speakers. The search results also suggest that the extended use of the plural marker is restricted to the Saliba data and is not found in the Logea dialect.

Saliba/Logea postpositional phrases can occur at the beginning and/or the end of a clause. From basic transcripts it is often not clear whether a PP belongs to the preceding or the following clause. This is because (a) in many instances clauses follow each other without any overt indication of coordination or subordination, and (b) the postposition does not clearly specify the role of its object. So in (1) the PP *Magehao unai* ‘in/at/to/from Magehao’ could belong to the first or the second clause, and the sentence *He came up Magehao PP he asked* could mean ‘He came up **to Magehao** and asked’ or ‘He came up and **at Magehao** he asked’.

- (1) *Ye* *saema* *Magehao* *unai* *ye* *henamai*
3SG.SUBJ come.up Place.Name PP.SG 3SG.SUBJ ask
(a) [He came up to Magehao] and asked.
(b) He came up and [at Magehao he asked]. (WekuSinibu_01AC_123-24)

Which of the two readings is correct cannot be gleaned from the basic transcription in (1). A text notation including pauses may resolve the ambiguity of the transcription – but only if there are indeed pauses preceding or following the PP (the ambiguity remains if there are no pauses or if pauses are noted before and after the PP). A transcription with notation of intonation patterns may disambiguate but is extremely labour-intensive for an entire database. A transcription with text-audio linkage allows us to investigate pauses and intonation patterns in order to identify which reading is appropriate. It turns out, however, that some Saliba/Logea examples remain vague even when there is direct access to the audio recording and there may be an intentional vagueness where the PP is somehow shared between the preceding and the following clause. Such vague examples constitute a research topic worth investigating and need to be distinguished from examples which may only appear ambiguous in a text-only transcription.