

## Challenges in describing contact-induced language variation

In this paper, I would like to elaborate on the challenges of balancing the linguists' and the language community's wishes and expectations when describing contact-induced language variation, based on my experience in preparing a descriptive grammar of Lakurumau, an endangered Oceanic language spoken in Papua New Guinea.

The grammar is primarily based on the documentation corpus of Lakurumau, built between 2017 and 2019 and comprising 18 hours of recordings transcribed and translated into English. The corpus includes different text genres (Table 1.) and speakers from a wide age range (Table 2.). This variety has proved useful in documenting rare grammatical combinations, typical of specific text types, and contact-induced language change tendencies in the speech of younger speakers. Nowadays, the main contact language of Lakurumau is Tok Pisin, the lingua franca of Papua New Guinea. Knowledge of Tok Pisin is pervasive: everyone is bilingual and many households in Lakurumau have completely switched to it. The influence of Tok Pisin on Lakurumau is visible on three levels: code-mixing, lexical borrowings and grammatical calques. The first two phenomena are attested in all speakers, though to different extents. Grammatical calques from Tok Pisin are, for the most part, exclusive to younger speakers (under 22 years); cf. the use of alienable possessive forms instead of inalienable ones (ex. 1) and the use of the locative preposition with local nouns and place names, which do not need extra locative marking (ex. 2). Even heavier grammatical interference can be seen in speakers who have completely switched to Tok Pisin in their everyday life and only occasionally speak Lakurumau.

The corpus building did not raise problems as to the community's expectations and wishes; everyone was quite enthused to participate in the documentation project. I did receive occasional complaints as to some speakers I decided to include, who were seen as "bad speakers" by other community members, but in general my motivation of representing a wide range of Lakurumau speakers and varieties was well understood and accepted. The grammar writing, conversely, poses some problems with respect to the choice, whether to including data about code-mixing and lexico-grammatical influence from Tok Pisin, which are which are perceived as "bad language" by Lakurumau speakers. On the one hand, such information is useful for linguists and the grammar should reflect the real practices of the speakers. On the other hand, including "incorrect" language may actually be a problem for the community, which may feel that their language is misrepresented. Moreover, given that the grammar data are always cross-referenced and retrievable in the corpus, the speakers are identifiable and, especially the youngest ones, they may feel flagged as "bad speakers". It is true, that the descriptive grammar is primarily aimed at linguists. However, Lakurumau speakers have taken pride in having their language documented and described, and are expecting a book-length description of it, which the neighbouring languages Kara and Nalik already have. A solution would be to use a positive phrasing when describing contact-induced change, underlying the importance of dynamic processes of change, which occur in any living language.

### Tables and examples

Genres and sub-genres	Time (HH:MM:SS)
Narratives (including: personal and traditional narratives, fiction, history, exposition of cultural practices, stimuli)	09:54:13)
Dialogues	08:55:32
Procedural	00:48:50
Observational filming	00:50:00

