

Can Bible translations be used as a source of language data in grammar writing?

When writing a grammar, it is desirable to base one's analytic conclusions and illustrative examples on natural texts, on a par with elicited examples. However, field linguists are not always in a position to collect large corpora of spontaneous texts that would be representative of various linguistic phenomena and suitable for a convincing quantitative analysis. A possible source of natural texts is the published literature in a target language, e.g. fiction. Still, it is often the case that minor languages mainly function orally, and the only published source available for a given language is a translation of the Bible (Heider et al. 2011: 54).

While Bible texts are now often used to create and to study parallel corpora (Wälchli 2007; Christodouloupoulos & Steedman 2015), their place is still somewhat controversial in field linguistics. Some linguists explicitly note that Bible texts do not exhibit any significant grammatical differences from the natural spontaneous texts and can be used for analysis (Heider et al. 2011; Khachaturyan 2015: 7), while others see Bible translations as unreliable and not fully representative since they may use archaic language, they are stylistically marked and they may contain calques or errors, cf. Wälchli (2007: 131). Still, it is an empirical question whether and how Bible translations differ from the natural texts in a given language. This issue is related to a more general question whether field linguists can use the Bible translations as a source of data for language documentation and ultimately for reference grammars.

In this talk, I address both questions by analyzing parallel Bible translations in Guinean and Liberian varieties of the Kpelle macrolanguage (Mande, West Africa). Crucially, I compare Bible translations with natural texts in these dialects collected by myself in the Republic of Guinea in 2008-2018, as well as taken from the published sources (for Liberian Kpelle). I provide a case study of quotative constructions, i.e. linguistic expressions introducing reported discourse (Güldemann 2008). These syntactic expressions are good candidates for exploring the reliability of Bible texts since (a) they show strong language-internal and crosslinguistic variation in African languages; (b) they are quite frequent in both colloquial and religious texts and (c) one might expect that if Bible translations differ from other texts, the divergence is likely to be found in the domain of syntax.

I demonstrate that quotative markers pattern differently in the two translations. First, quotatives tend to be used as independent predicates in Guinean Bible and as quotative clause linking markers in Liberian translation. Another crucial difference is that quotative markers tend to co-occur with semantically empty pronominal addressees in Guinean translation, which is much less frequent in Liberian translation. Crucially, this difference cannot be attributed to the source languages of translation, i.e. French and English. In contrast, I show that the tendencies similar to those found in Bible translations are attested in other Guinean and Liberian texts.

I conclude that Bible translations in the two Kpelle dialects pattern similarly to natural texts with respect to the syntax of quotatives. This implies that Bible translations may indeed be used for corpus-based descriptions of minor languages, although it is always desirable to complement Bible translations with natural spontaneous texts in one's analytic work, as well as to explicitly test the distribution of particular phenomena in religious vs. other texts.

References

Christodouloupoulos, Christos and Mark Steedman. 2015. A massively parallel corpus: the Bible in 100 languages. *Language Resources and Evaluation* 49(2): 375–395.

Güldemann, Tom. 2008. *Quotative indexes in African languages. A synchronic and diachronic survey*. Berlin: Mouton de Gruyter.

Heider, Paul M., Hatfield, Adam, and Jennifer Wilson. 2011. Repurposing Bible translations for grammar sketches. *Studies in the Linguistic Sciences: Illinois Working Papers*: 51–65.

Khachaturyan, Maria. 2015. Grammaire du mano. *Mandenkan* 54: 1–252.

Wälchli, Bernhard. 2007. Advantages and disadvantages of using parallel texts in typological investigations. *Sprachtypologie und Universalienforschung*. 60(2). 118–134.