

Three Approaches to Prefix and Suffix Statistics in the Languages of the World

Harald Hammarström

MPI / Radboud U Nijmegen

It has long been estimated that suffixing is more common than prefixing in the languages of the world. It is of interest to get detailed statistics on this tendency in order to understand whether the preference is localized to a specific region of the world or whether it calls for a general explanation in terms of the human mind or human communication, (see Himmelmann 2005 *inter alia*).

The classic approach to gathering data on the prefix/suffix preference is for a human to read grammatical descriptions. This approach is represented by Dryer (2005), who classifies 894 languages into one of five categories *Little or No Inflectional Morphology*, *Predominantly Suffixing*, *Moderate Preference for Suffixing*, *Approximately Equal Amount*, *Moderate Preference for Prefixing*, *Predominantly Prefixing*.

A second possibility is to let a computer read a grammatical description (originally intended for humans to read) and crudely count the number of times the words *prefix* and *suffix*¹ occur. We have access to a large collection of OCR:ed grammatical descriptions spanning 3014 languages.

A third possibility is to go directly to raw text. A simple observation is that the words in a suffixing language have less variability in the number of different segments that appear at the end of words, and *mutatis mutandis* in a prefixing language. There are various ways of turning this observation into a measure of suffixation/prefixation/neither. We have access to a collection of (partial) bible translations for over 800 languages².

We will present a comparison the three approaches in terms of theoretical advantages and perform a full empirical investigation of how well the three correlate.

References

Dryer, Matthew S. 2005. Prefixing Versus Suffixing in Inflectional Morphology. In Bernard Comrie, Matthew S. Dryer, David Gil & Martin Haspelmath (eds.), *World Atlas of Language Structures*, 110-113. Oxford University Press.

¹Or the equivalent in whatever language the grammatical description is written in.

²We wish to thank the individual translation organizations and Thomas Mayer and associates at LMU-Marburg for access to these.

Himmelmann, Nikolaus. 2005. Asymmetries in Word Architecture: Another look at the Suffixing Preference. Paper Presented at the 6th Biennial Meeting of the Association for Linguistic Typology, Padang, West Sumatra, Indonesia.