Quantifying contact-induced change: A multivariate typological approach

Kaius Sinnemäki & Noora Ahola
(University of Helsinki & University of Helsinki)

Keywords: language contact, language typology, possessive noun phrases, sampling

In this presentation we illustrate a multivariate typological approach for researching and quantifying language contact effects and discuss preliminary results based on it. In contact research a central aim is to analyze whether contact between different languages has resulted in linguistic changes, and if yes, to what kind. A useful way to classify such changes is into matter vs. pattern borrowing, that is, borrowing of actual forms vs. patterns, such as prosody, case alignment, or word order (cf. Matras & Sakel 2007, Sakel 2007). However, typological studies on contact effects are rare, usually demonstrating contact effects or areal phenomena rather than quantifying contact effects.

To detect contact effects, we employ the sampling method developed by Di Garbo et al. (2021) and analyze 15 language sets representing nine of the ten continents in the *AUTOTYP* database (Bickel et al. 2022). Each set contains three languages: (1) the Focus Language, (2) the Neighbor Language (unrelated to the Focus but in contact with it), and (3) the Benchmark Language (related to the Focus but not in contact with it or the Neighbor). If the Focus and Neighbor share features that the Focus and Benchmark do not, this is taken as a sign of possible contact effect in the Focus Language.

We apply this method to evaluating pattern borrowing (or pattern convergence) in adnominal possessive constructions. Four features were analyzed: locus of marking, conditioning features, boundness, and word order. Locus of marking is about whether the syntactic dependency between the possessor and the possessee is morphologically marked on the possessor, on the possessee, on both or on neither (Nichols 1992). In English, the possessor is morphologically marked as in *Mary's mother*; in Abun (1a) it is the possessee (Berry & Berry 1999: 77-78):

(1) a. *ji   bi   nggwe*   b. *ji   syim*
          1SG   POSS   garden   1SG arm
          'my garden'   'my arm'

The conditioning features refer to factors that may condition language-internal variation in locus of marking, such as head marking of alienable possession in Abun (1a) and zero marking for inalienable possession (1b). Boundness concerns whether morphological marking is done by bound morphology, a clitic or an independent word. Word order refers to the order between the possessor and the possessee.

Our preliminary analyses suggest that there were contact effects in 13 languages (87%). Changes occurred in all features, most often in locus of marking (9/15) and the least often in boundness (2/15). In six languages, changes were observed in two or more features, in five traces of contact effects occurred in one feature only.

The results suggest that the sampling method together with the multivariate typological analysis can detect language contact effects in all studied features. The contact effects are not prominent across all features, nor do they occur in each of the language sets. However, the approach enables quantifying contact effects in well-delimited typological domains at least in preliminary ways. We also discuss briefly how the magnitude of the contact effects could relate to the contact ecology.

Contact scenarios under study

The language sets are listed below by *AUTOTYP* continents, the first language being the Focus, followed by the Neighbor and the Benchmark (language family is given in parentheses).

Africa: Mursi (Surmic), Hamer-Banna (South Omotic), Tennet (Surmic); Kambaata (Afro-Asiatic), Wolaytta (Ta-Ne-Omotic), Xamtanga (Afro-Asiatic)
Australia: Mawng (Iwaidjan), Kunbarlang (Gunwinyguan), Iwaidja (Iwaidjan); Yawuru (Nyulnyulan), Karadjeri (Pama-Nyungan), Bardi (Nyulnyulan)
Western and Southwestern Eurasia: Zazaki (Indo-Eur.), Turkish (Turkic), Western Balochi (Indo-Eur.)
New Guinea and Oceania: Papapana (Austrones.), Rotokas (North-Bougainville), Marshallese (Austrones.); Alorese (Austrones.), Adang (Timor-Alor-Pantar), Lewoingu Lamaholot (Austrones.)
South and Southeast Asia: Burmese (Sino-Tibetan), Mon (Austroasiatic), Kurtöp (Sino-Tibetan); Kupwar Marathi (Indo-European), Kupwar Kannada (Dravidian), Chakma (Indo-European)
Northern and Central Asia: Wutun (Sino-Tibetan), Bonan (Mongolic-Khitan), Cantonese (Sino-Tibetan); Sibe (Tungusic), Uighur (Turkic), Even (Tungusic)
South America: Western Toba (Guaicuruan), Wichí Noctén (Matacoan), Kadiwéu (Guaicuruan); Yuhup (Naduhup), Macuna (Tucanoan), Nadëb (Naduhup)
Central America: Garifuna (Arawakan), Galibi (Cariban), Ashéninka Perené (Arawakan)
Western North America: Nuxalk (Salishan), Kwak'wala (Wakashan), Okanagan (Salishan)

References
Berry, Keith & Christine Berry (1999), *A Description of Abun: A West Papuan Language of Irian Jaya*. Canberra: Australian National University.
Bickel, Balthasar, Johanna Nichols, Taras Zakharko, Alena Witzlack-Makarevich, Kristine Hildebrandt, Michael Rießler, Lennart Bierkandt, Fernando Zúñiga & John B. Lowe (2022), The AUTOTYP database (v1.0.1), https://doi.org/10.5281/zenodo.6255206.
Di Garbo, Francesca, Eri Kashima, Ricardo Napoleão de Souza & Kaius Sinnemäki (2021), Concepts and methods for integrating language typology and sociolinguistics, in S. Ballarè, and G. Inglese (eds), (2021), *Tipologia e Sociolinguistica: verso un approccio integrato allo* studio *della variazione: Atti del Workshop della Società Linguistica Italiana 20 settembre 2020*, Milano: Officinaventuno, 143-176.
Matras, Yaron & Jeanette Sakel (2007), Introduction, in Y. Matras, and J. Sakel (eds), (2009), *Grammatical Borrowing in Cross-linguistic Perspective*, Berlin: Mouton de Gruyter, 1-13.
Nichols, Johanna (1992), *Linguistic Diversity in Space and Time*. Chicago, IL: The University of Chicago Press.
Sakel, Jeanette (2007), Types of loan: matter and pattern, in Y. Matras, and J. Sakel (eds), (2009), *Grammatical Borrowing in Cross-linguistic Perspective*, Berlin: Mouton de Gruyter, 15-29.