SOME NOTES ON THE LEIPZIG-JAKARTA LIST IN THE LANGUAGE OF THE TOMSK TATARS

Tokmashev D.M.
National Research Tomsk Polytechnic University

In the 1950s, the linguist Morris Swadesh published a list of 200 words called the Swadesh list, allegedly the 200 lexical concepts found in all languages that were least likely to be borrowed from other languages. Swadesh later whittled his list down to 100 items. The Swadesh list, however, was based mainly on intuition, according to Martin Haspelmath and Uri Tadmor. Translations of the Swadesh list into a set of languages allow researchers to quantify the interrelatedness of those languages. The Swadesh list is named after the U.S. linguist Morris Swadesh. It is used in lexicostatistics (the quantitative assessment of the genealogical relatedness of languages) and glottochronology (the dating of language divergence).

The Leipzig-Jakarta list is a 100-word word list used by linguists to test the degree of chronological separation of languages by comparing words that are resistant to borrowing. The Leipzig-Jakarta list became available in 2009.

In 2016 we have undertaken a set of fieldtrips to the areas of traditional dwelling of the Tomsk Tatars funded by the SOAS ELDP project, aiming at gathering fieldtrip material including the Leipzig-Jakarta list. It's worth mentioning that the dialect of the Tomask Tatars has been undeservingly forgotten by linguists who take up Turkic varieties of Siberia. The last profound study of Eushta dialect, which comprises the Tomsk Tatar dialect cluster along with the idioms spoken Takhtamyshevo and Chyornaya Rechka was carried out in the 1960^s.

The results show that the dialect complex spoken by indigenous Turkic peoples who reside in the Tomsk area in the villages of Eushta, Chernaya Rechka, Takhtamyshevo, Boriki, Timiryazevo, Barabinka, Kaftanchikovo and some others, as well as the city of Tomsk, though retaining basic linguistic features of the North Kyptchak languages and very close to Kazan Tatar literary language, manifest a set of unique lexical data, which differ even within the Tomsk Tatar dialect cluster, e.g. the word "egg" has 3 different variants: *jymyrqa* (literary Kazan), *kukaj* and *tygam* which are regional and have no direct parallels in the adjacent Siberian Turkic cognates, which gives a broad field for subsequent comparative study.