Russian numerals in Moksha and Hill Mari speech[1]

Irina Khomchenkova[1,2,3] & Polina Pleshak[1,2]
1 – Lomonosov Moscow State University; 2 – Institute of Linguistics, RAS;
3 – Vinogradov Russian Language Institute, RAS

This paper deals with the occurrences of Russian numerals in the spontaneous speech of Moksha-Russian and Hill Mari-Russian bilinguals. We argue that these are not borrowings but rather instances of code-switching (CS). We propose an analysis in terms of the Matrix Language Frame (MLF) Model (Myers-Scotton 1993), basing on our comparative corpus study of the numeral phrases with CS in Moksha and Hill Mari.

Both languages in question are Finno-Ugric languages spoken in the Volga region of Russia and both have a long history of contacts with Russian (Bereczki 1968; Johanson 2000). However, as far as we know, there are no works on CS in these languages. The data come from spoken corpora collected in the villages of Lesnoje Tsibajevo and Lesnoje Ardashevo in the Republic of Mordovia (Moksha corpus: 20103 tokens) and in the village of Kuznetsovo in the Mari El Republic (Hill Mari corpus: 44297 tokens).

Our first aim is to describe the internal morphosyntax of constructions with Russian numerals. The constructions with cardinal numerals in Moksha and Hill Mari differ from Russian ones. In Russian, numerals require genitive case on the noun, whereas in Moksha and Hill Mari numerals do not influence noun case. The rules of noun number marking are different in all the three languages (Sidorova 2018). The fact that the structures are incongruent leads to the formation of Russian embedded language (EL) islands (in MLF model) in both languages (1), (2). EL islands obey the rules of the EL in their internal syntax. In the external syntax they follow the rules of the matrix language (ML) – Hill Mari and Moksha, respectively.

HILL MARI

(1) *st'ip'end'ij-žǝ̂*   ***dv'enacat' rubl'-ej***   *ǝ̂l'-ǝ̂*   *v*   *m'es'ac*
scholarship-POSS.3SG   twelve   ruble-GEN.PL   be-AOR.3SG   in   month
'The scholarship was twelve rubles per month'.

MOKSHA

(2) *t'ejǝ-st*   *er'av-ǝl'-ø*   *sa-m-s*   ***p'at' časov-sǝ* /   ***p'at' čas-cǝ*
PRON.DAT-3PL.POSS need-PQP-SG   come-INF-ILL   five   hour-GEN.PL-IN   five hour-IN
'They had to come at 5'.

As we see in (2), in Moksha, in contrast to Hill Mari, it is possible to add ML morphological markers reflecting the external syntax. It is an instance of The Double Morphology Principle (Hok-Shing Chan 2009). Moreover, as (2) shows, it is obligatory to insert a constituent following a Russian syntactic pattern. No such constraints are attested for Hill Mari.

The constructions of another type are also attested, namely, ML+EL constituents. These are mainly constructions with ordinal numerals. In these constituents, Russian words are inserted into the ML frame:

HILL MARI

(3) ***p'ervǝ̂j sm'enǝ̂*-*m=ät*   *nänge-ä*
first   shift-ACC=ADD   lead-NPST.3SG
'And he is leading the first shift'.

MOKSHA

(4) *kul'-at*   *što li*   *al'i*   **tr'icat'**   **p'ataj**   **s'ekcijɛ-s'**
hear-NPST.2SG or.not   or   thirty   five   section-DEF.SG
'Do you hear or not, the thirty fifth section'.

Auer and Muhamedova (2006) analyze Kazakh (< Turkic) sentences with Russian insertions analogous to (3) and (4) as instances of EL islands and not as two single word insertions, since "there is a relationship of dependency between the two words" (Auer, Muhamedova 2006: 44). They argue that the matrix language (in this case having no gender system) can have an impact on the embedded language (in this case having a gender system), because neither *pervâj sm'enə* 'first shift' in (3) nor *tr'icat' p'ataj s'ekcijɛ* 'thirty fifth section' in (4) are grammatical in Russian due to gender mismatches. We disagree with this point of view. We argue that these are ML+EL constituents, since all members of the constituent are taken from the EL in their bare form (or in the "default" masculine gender), but no island is formed: the internal syntax is that of the ML.

The second aim of this paper is to compare the preferences in the CS of numerals in Hill Mari and Moksha. We will discuss the following factors:

a) Syntactic type of numerals: ordinal (Russian numerals constitute 30,4% of all numerals in Hill Mari data, 51% in Moksha) > cardinal (HM: 4,2%, M: 28%) > collective (HM: 2,1%, M: 0%).

b) Semantic type of numerals: 1000, 100 > 20+ > 10+ > 5+ > 5- (Matras 2007: 51). In both languages the following hierarchy is applicable: more than 10 > less than 10.

c) Type of the context.

To sum up, although in both languages numeral systems are conserved, there is a strong tendency to use Russian numerals in some contexts to express larger quantities. We consider that ordinal numbers are switched more frequently because of structural reasons: they occupy different position in the NP. We argue that the formation of EL islands is influenced by the types of dependencies established in a construction (quantifier-like cardinal numbers vs. adjective-like ordinal numbers). The obligatoriness of islands can vary among different structures within one language, as well as across languages. These constraints may shed light on the "tightness" of syntactic relations in a construction and on the congruence of the structures in the languages. In the talk, we will elaborate on these points and provide more language examples.

Abbreviations

2, 3 – 2 and 3 person, ACC – accusative, ADD – additive, ATTR – attributivizer, DAT – dative, DEF – definite, GEN – genitive, ILL – illative, IN – inessive, INDEF – indefinite, INF – infinitive, NPST – non-past tense, PL – plural, POSS – possession, PQP – plusquamperfect, PRON – pronominal stem, REFL – reflexive, SG – singular

References

Auer, Peter and Muhamedova, Raihan (2006). 'Embedded language' and 'matrix language' in insertional language mixing: Some problematic cases. Italian Journal of Linguistics / Rivista di linguistica 17(1). [Special Issue: "Commutazione di codice e teoria linguistica"]. P. 35–54.
Bereczki, Gábor (1968). Wichtigere lautgeschichtliche Lehren der russischen Lehnwoerter im Tscheremissischen // Congressus Secundus Internationalis Fenno-Ugristarum. Pars I. Acta Linguistica. Helsinki. P. 70-76.
Hok-Shing Chan, Brian (2009). Code-switching between typologically distinct languages. In Bullock B.E., Toribio A. (Eds.) The Cambridge Handbook of Linguistic Code-switching. Cambridge: Cambridge University Press.
Johanson, Lars (2000). Linguistic convergence in the Volga area. Studies in Slavic and General Linguistics, 28, P. 165-178.

Matras, Yaron (2007). The borrowability of structural categories // Empirical Approaches to Language Typology, 38, 31.

Myers-Scotton, Carol (1993). Duelling languages: Grammatical structure in code-switching. Oxford: Clarendon Press. 265 p.

Sidorova, Maria A. (2018). Chislovoe markirovanie suschestvitel'nogo v gornomarijskih kolichestvennyh konstrukcijah [Number marking of nouns in Hill Mari numeral constructions]. Acta Linguistica Petropolitana, №2. P. 400–427.