

Code-switching vs. borrowing: a morphosyntactic congruence perspective

Russian nouns in Southern Tungusic

Natalia Stoyanova, Russian Language Institute RAS, Moscow,
stoyanova@yandex.ru

TMP, Moscow (on-line), 7-9 October 2020

Introduction

- **Code-switching vs. borrowing:**

- code-switching ↔ occasional (nonce) borrowings ↔ established borrowings
- continuum? clear borderline(s)? if yes, where?

- **Data:**

- Nanai and Ulcha texts with Russian fragments
- Russian nouns: multi-word NPs vs. occasional vs. frequent

- **Starting points:**

- ?borrowing ↔ L-recipient morphosyntax?
- ?code-switching L-donor&L-recipient morphosyntax?
- morphosyntactic (in)congruence between L-recipient and L-donor:
 - influences borrowings rather than code-switches

1. Theoretical background

Code-switching vs. borrowing

→ How to differentiate between code-switching and borrowing?

→ Different approaches, different criteria

- frequency
- +/- phonetic adaptation
- +/- morphosyntactic integration
- semantics: +/- correlates in the recipient language (cf. cultural borrowings)
- included / not included in dictionaries
- speakers' reflection: perceived / not perceived as a “foreign word”
- used / not used by monolinguals
- ...

→ cf. an overview e.g. in Haspelmath (2009: 40–42)

Code-switching vs. borrowing

C. Myers-Scotton

- a very wide range of phenomena is attributed as code-switching, including “word-internal CS”
- **no clear-cut** between CS and borrowing
- today’s code-switches are tomorrow’s borrowings
→ Myers-Scotton 1992; 1993: 163–228;
2002: 153–160

Sh. Poplack: Nonce Borrowing Hypothesis

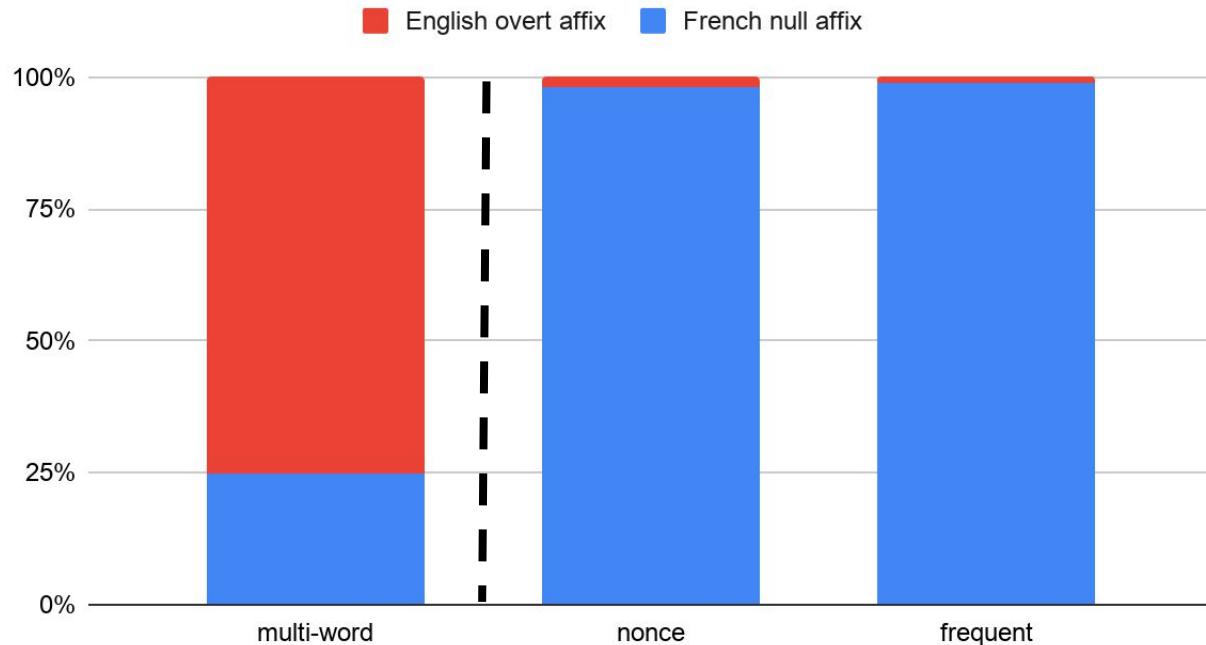
- a **clear-cut** between code-switching and borrowing
- CS and borrowing are essentially different phenomena
- morphosyntactic integration is the best indicator
- there are spontaneous one-word insertions (**nonce borrowings**) that behave in the same way as established ones and differently from code-switches
→ Poplack et al. 1988; Sankoff et al. 1990;
Poplack 2012; 2018

Nonce Borrowing Hypothesis (NBH)

“...Speakers not only code-switch spontaneously, but may also BORROW spontaneously, and these spontaneous borrowings assume *the morphological and syntactic identity of the recipient language* even PRIOR to achieving the social characteristics of established loanwords (recurrence in the speech of the individual, and dispersion across the community).” (Poplack 2012: 645)

NBH: Morphosyntactic integration

Plural marking in the Ottawa corpus (adapted from Poplak 2018: 147)



→ Variationist frequency-based approach

→ Nonce borrowings behave exactly as established ones

→ Both types of borrowings differ a lot from code-switches

Congruence in code-switching

- **Morphosyntactic (IN)CONGRUENCE as the main factor regulating structural types and restrictions on code-switching**

→ cf. Myers-Scotton & Jake 1995; Jake & Myers-Scotton 1997; Myers-Scotton 2002: 20, 306;
Sebba 2009 etc.



“Consider two team games, for example, football (soccer) and basketball <...> The achievement of codeswitchers is *to play both games at once* in a way that is satisfactory to the participants, while keeping (sufficiently if not absolutely) to the rules of both.” (Sebba 2009:40)

Congruence in code-switching

- Code-switching strategies (Sebba 2009):
 - morphosyntactic congruence: L1 morphosyntactic structure = L2 structure

⇒ L1/L2 morphosyntactic structure

HARMONIZATION

- morphosyntactic incongruence: L1 morphosyntactic structure \neq L2 structure

⇒ L1 morphosyntactic structure

NEUTRALIZATION (nativization)

⇒ a new morphosyntactic structure, e.g. bare forms **COMPROMISE STRATEGY**

Objectives

Testing the Nonce Borrowing Hypothesis

- 1) To apply the procedure proposed by Sh. Poplack to a new data sample
 - Neutralization strategy (“conflict sites”): L1-morphosyntax or L2-morphosyntax?
 - Expectation: CS \Rightarrow L1, (nonce) borrowings \Rightarrow L2
- 2) To expand this procedure to other types of contexts
 - Harmonization strategy: the rate of uses in congruent contexts
 - Expectation: CS \Rightarrow more uses in congruent contexts; (nonce) borrowings \Rightarrow no effect
 - Compromise strategy: the rate of bare forms in place of those with inflection
 - Expectation: CS \Rightarrow bare forms are actively used, nonce-borrowings \Rightarrow do not differ from established ones

2. Data

Textual collection

- **Nanai (Amur dialects) and Ulcha:**

- **Southern Tungusic** (closely-related to each other); Khabarovsk Krai (Russia)
- **Endangered:** a progressing shift to Russian, no transmission to children
→ see Gerasimova (2002); Kalinina & Oskolskaya (2016); Sumbatova & Gusev (2016)

- **Texts:**

- Spontaneous oral texts, recorded in the field under instruction “to tell something in Nanai/Ulcha”
- Collected from 47 speakers in 16 villages of Khabarovsk Krai, transcribed in ELAN, partly glossed
- Biographic stories, folklore, ethnographic descriptions etc.
- Contain a lot of Russian fragments
- **Ca. 27 h.:** Nanai – 43 708 tokens (158 texts), Ulch – 44 902 tokens (174 texts)
→ Sofia Oskolskaya, Anna Smetina, Natalia Stoyanova

Data annotation: Russian nouns

- Russian nouns in Southern Tungusic texts:
 - multi-word NPs and one-word noun insertions
 - with and without Tungusic inflection
 - with and without phonetic adaptation
- Sample size:
 - total number: 2888
 - final sample: **2186**

→ excluded: proper names, pronouns, NPs with false starts / self-corrections

Data annotation: sub-samples to compare

- **MULTI-WORD:** 227 **a clear case of code-switching**

→ Russian multi-word NPs with or without Tungusic inflection

английский язык ‘the English language’, порок сердца-ژی ‘with a heart disease’

- **NONCE:** 330 **???**

→ Russian nouns with or without Tungusic inflection: attested only once

китикэт ‘Kitekat (cat food)’, штрафбата-du ‘at the penal battalion’

- **FREQUENT:** 625 **a clear case of borrowing**

→ Russian nouns with and without Tungusic inflection: attested 10+ times

школа ‘school’, школа-du ‘at school’

- **TUNGUSIC:** 2492 **a monolingual benchmark**

→ Tungusic nouns extracted from a small glossed sub-sample of Nanai texts

Data annotation: morphosyntactic integration

→ RUSSIAN: with Russian inflection

(1) тракторист-ом ujləxə case-number
tractor_driver-ins work-pst
'(I) worked as a tractor driver' (ulc, agk)

(2) невод-а, jarako, хай-sal... čežə-sal [plural (nom/acc)]
net-nom.pl net what-pl net-pl
'different types of fishing nets' (gld, pnv)

→ [RUSSIAN: bare forms licensed by Russian syntactic rules]

(3) все время wāč-i-t adverbial uses with no correspondences in Tungusic
all.n time ???-prs-3pl
'(they) use (the stove) all the time' (ulc, spk)

Data annotation: morphosyntactic integration

→ TUNGUSIC: with Tungusic inflection

- (1) крышка-џі jaqč-i-n=gun **case**
lid-ins close-prs-3sg
'(she) covers (the jar) with a lid' (ulc, oax)
- (2) учитель-сәл=dә o-či **plural**
teacher-pl=emph become-pst
'(they) became teachers' (gld,lkb)
- (3) гостиница-пу **possessive**
hotel-1pl
'our hotel' (vmk)

Data annotation: morphosyntactic integration

→ [TUNGUSIC: bare forms licensed by Tungusic syntactic rules]

- (1) ti медведь piktə-n=də **N + POSSESSEE-poss**
that bear child-3sg=emph
'this bear cub' (ulc, aid)
- (2) койка oja-la-ni **N + postposition/spatial noun-poss**
bed on-loc-3sg
'on the bed' (gld, vsg)
- (3) доктур=də kəwə **N + NEG.EXIST**
doctor=emph neg.exist
'there is no doctor' (ulc, oab)
- (4) ńuŋgu парта **NUM + N.nom**
six desk
'six desks' (ulc, spd)

Data annotation: compromise strategy

→ BARE FORMS (NOM.SG) instead of Tungusic/Russian inflection

(1) магазин žobo-j-du-j=tani

shop work-prs-dat-refl.sg=coord

'when (I) worked at the shop' (gld, rchk)

expected:

(2) магазин-du / в магазин-е

shop-dat / in shop-loc

'at the shop' (Nanai) / 'at the shop' (Russian)

Data annotation: harmonization

→ Bare forms (NOM.SG) in congruent contexts

(1) печка=gdəl dāi=guni **Subject position**

stove=ptcl big=ptcl

‘I mean, the stove is really big!’ (ulc, lpd)

(2) плот anžu-xa-ni **[Direct object position (not for all nouns)]**

‘(he) built a raft’(ulc, aid)

- DOM in Russian: NOM = ACC vs. NOM≠ACC: depends on animacy, declension type
- DOM in Southern Tungusic: free variation between NOM and dedicated ACC forms (see Oskolskaya & Stoyanova 2017)

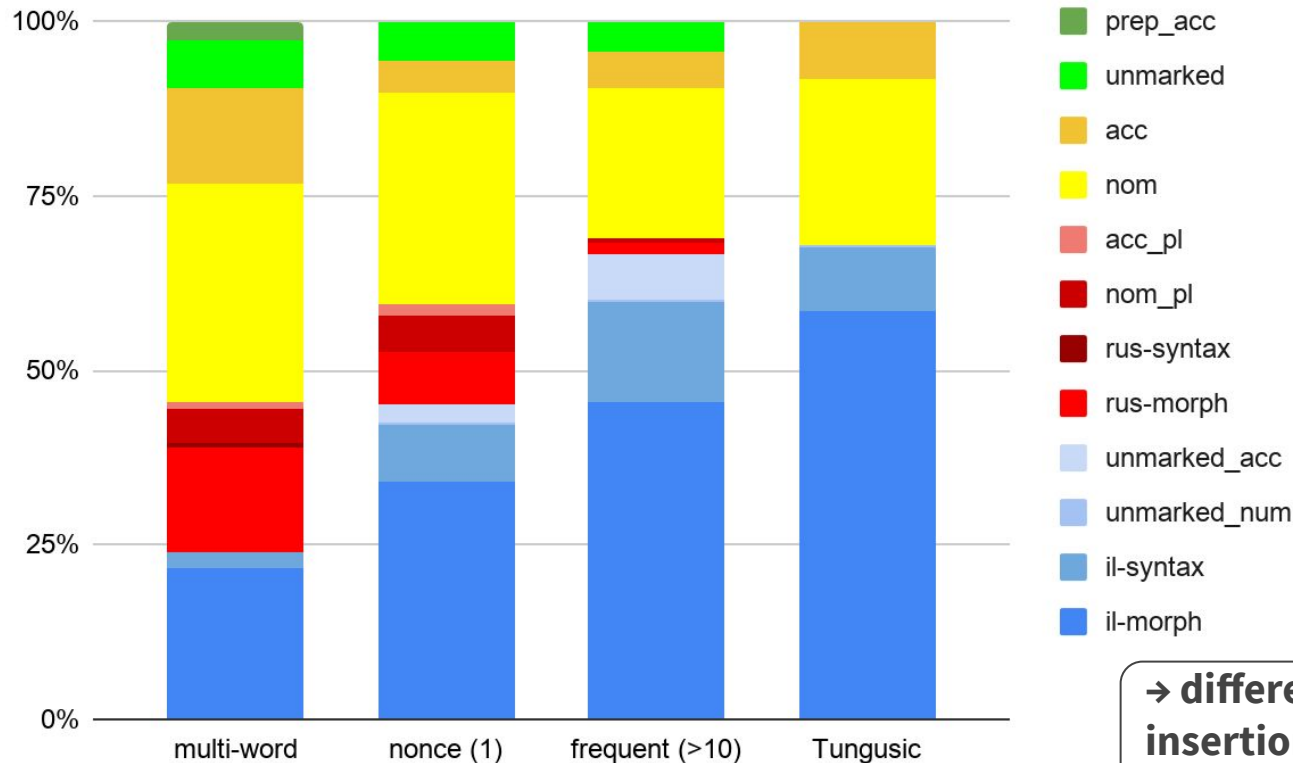
→ Only the contexts in which bare forms (NOM) are acceptable both in Russian and in Tungusic were included.

3. Results

DISCLAIMER:

→ Preliminary results reported in the abstract differ from the final ones

Results: general picture



blue - Tungusic morphosyntax

red - Russian morphosyntax

yellow - congruent contexts

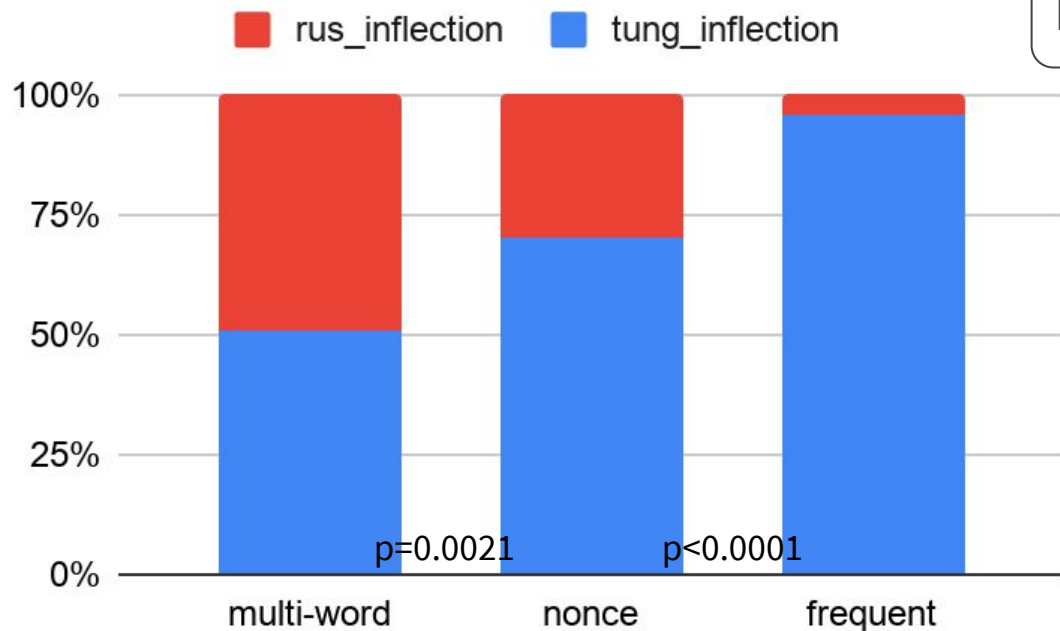
green - bare compromise forms

→ different types of Russian noun insertions behave in different ways

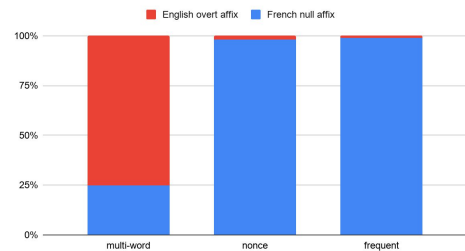
Results: neutralization strategy (morphosyntactic integration)

- (1) милиция-зи ηələču-mər TUNGUSIC structure
police.R-ins fear-cvb.sim.pl (ulc, oax)
- (2) милици-и ηələču-mər RUSSIAN structure
police.R-gen.Rfear-cvb.sim.pl (constructed)
'having a fear of policemen'

Results: Tungusic vs. Russian inflection



→ the picture differs from that reported by Poplack



% of uses with Russian inflection:

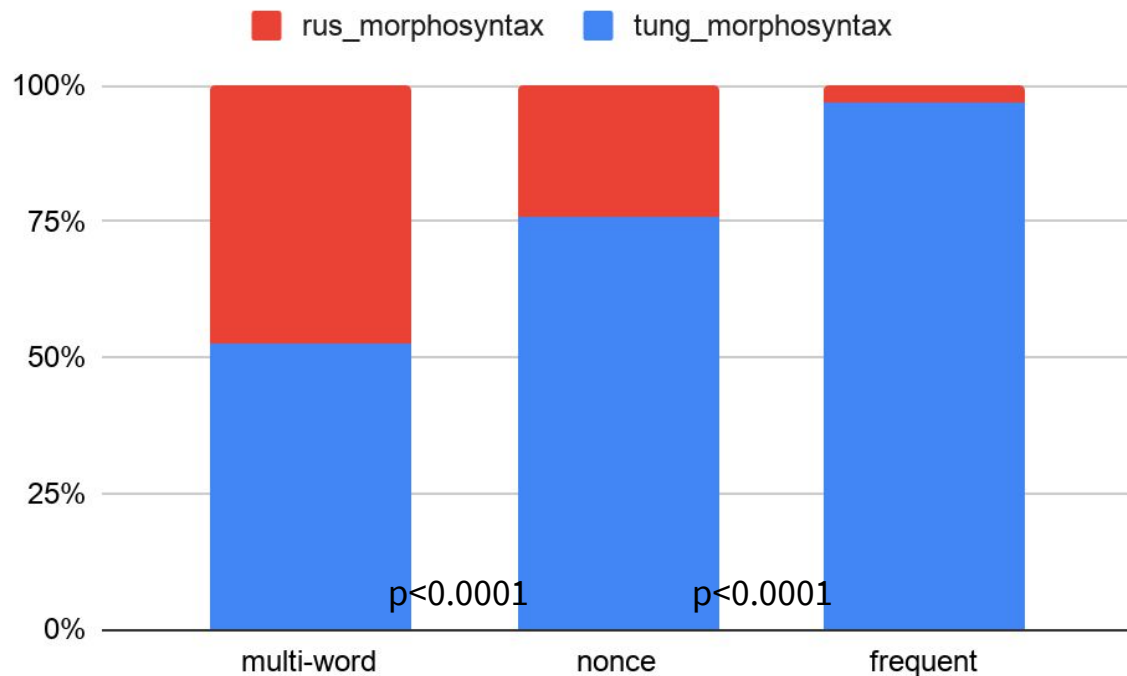
multi-word

> nonce

> established

Two-tailed Fisher's exact test

Results: Tungusic vs. Russian morphosyntax



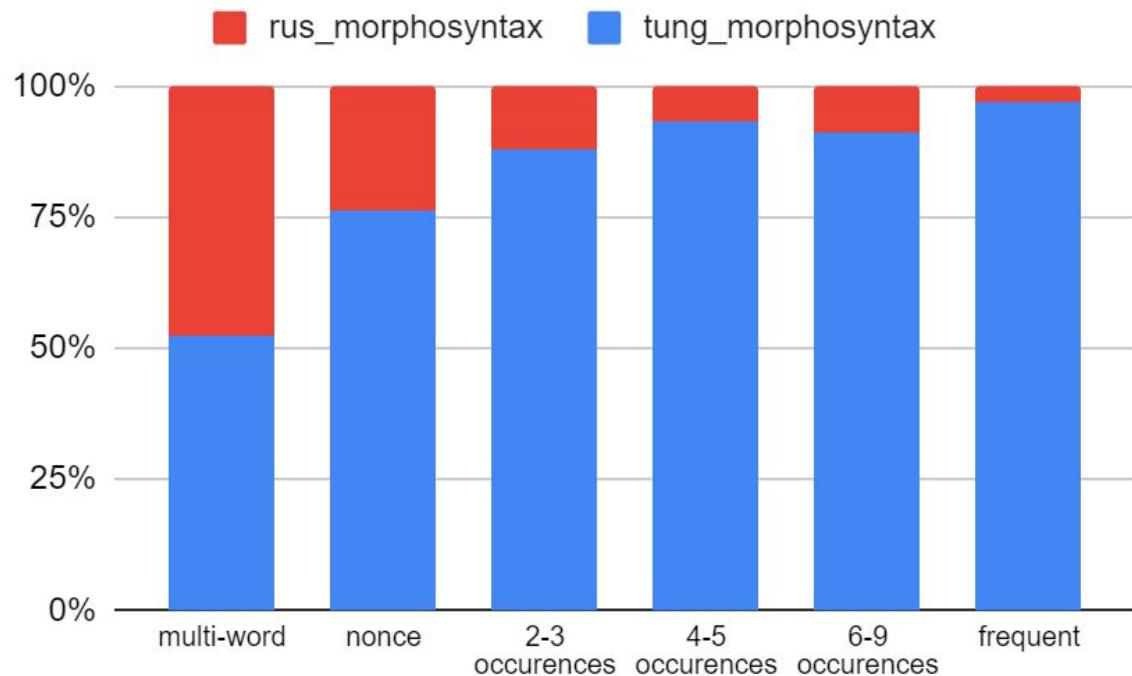
Russian inflection & bare forms licenced by Russian syntax

vs.

Tungusic inflection & bare forms licenced by Tungusic syntax

→ the effect is even more visible

Results: Tungusic vs. Russian morphosyntax (detailed)



+ The data on nouns of intermediate frequency

→ a gradual change

Results: compromise strategy (bare forms in incongruent contexts)

милиция

police.R

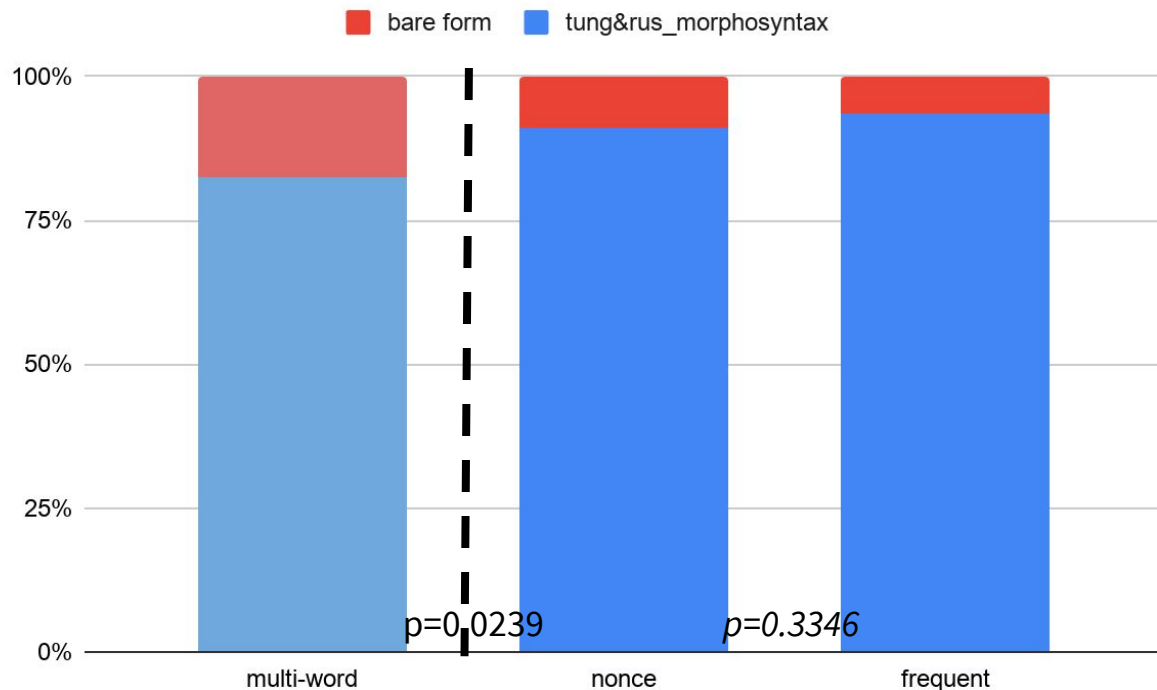
ɲələʃu-mər

fear-cvb.sim.pl (constructed)

BARE FORM

‘having a fear of policemen’

Results: bare forms instead of Tungusic/Russian inflection



multi-word NPs:

→ *a bit* more bare forms

nonce & frequent

borrowings:

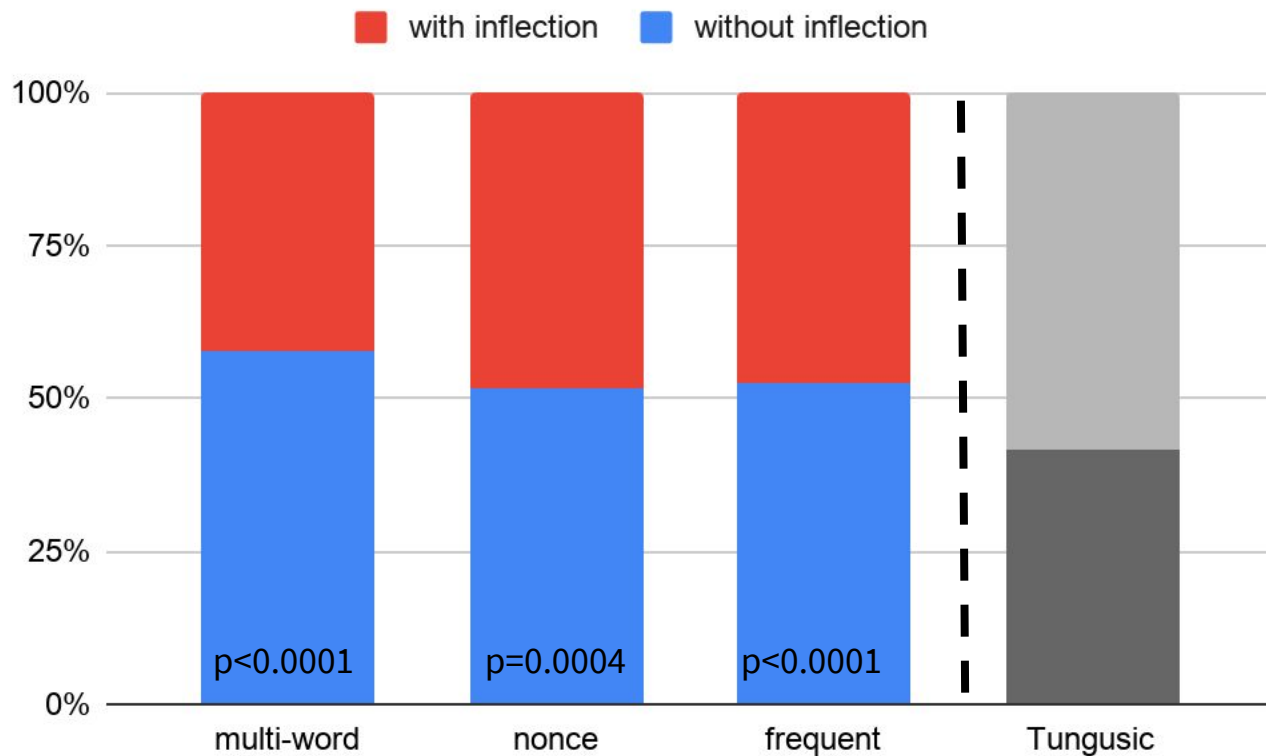
→ fewer bare forms

→ behave in the same way

Results: harmonization strategy (bare forms in congruent contexts)

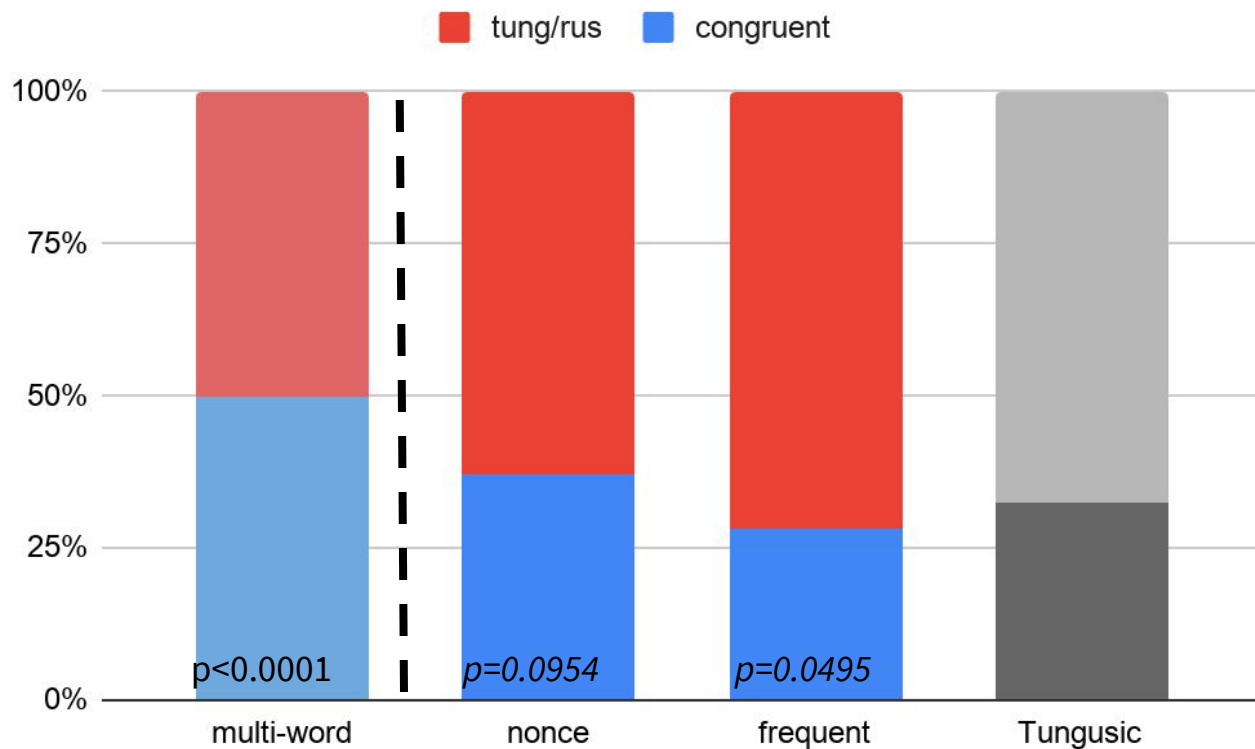
<u>конференция</u>	xodu-wa-mi	BARE FORM
conference.R	finish-ds-cvb.sim.sg	
'when <u>the conference</u> finished' (ulc, lpd)		

Results: with vs. without inflection



multi-word CS & nonce
& frequent borrowings
→ behave in the same
way
→ differ from Tungusic
nouns
→ all avoid inflection

Conflict sites vs. congruent contexts



multi-word NPs:

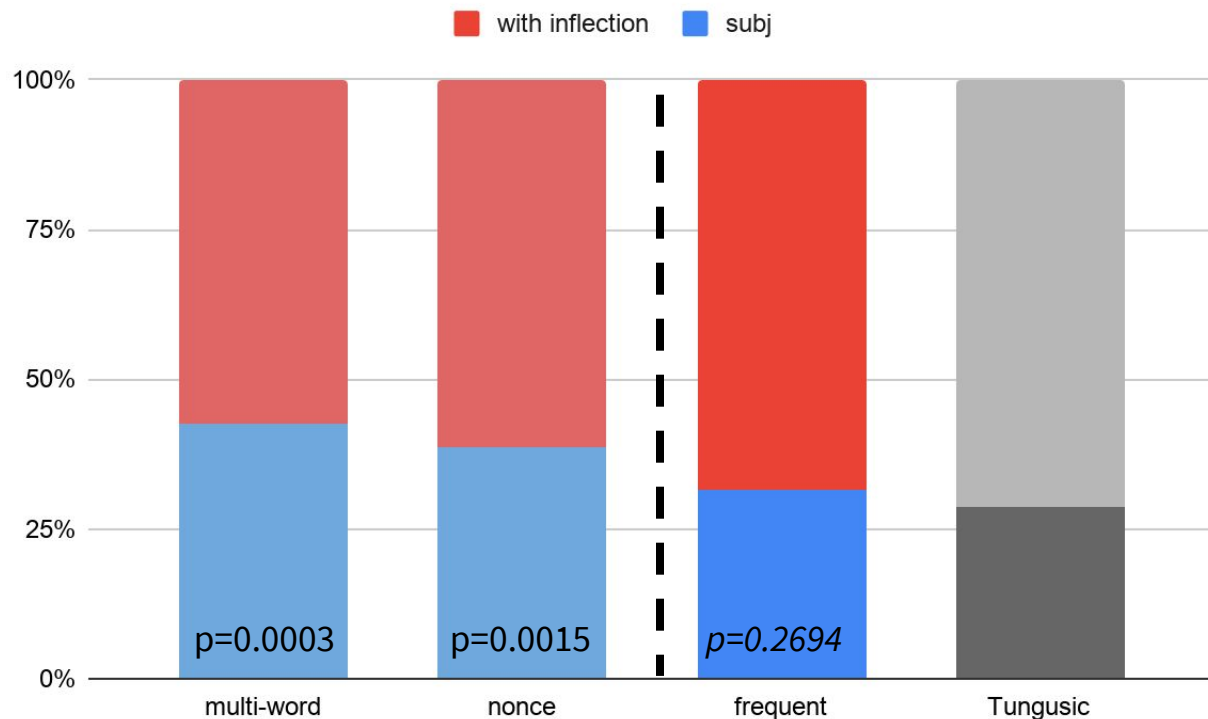
→ more uses in
congruent context

nonce & frequent

borrowings:

→ do not show any
effect

Inflection vs. nom.sg (subject)



multi-word CS and
nonce borrowings:

→ a bit more uses in
congruent context

frequent borrowings:

→ do not show any
effect

Results: Summary

- **NEUTRALIZATION STRATEGY:** Tungusic vs. Russian morphosyntax
 - multi-word NPs - higher rate of uses with Russian morphosyntax
 - frequent borrowings - higher rate of uses with Tungusic morphosyntax
 - nonce borrowings - intermediate result
 - nonce borrowings do not cluster with frequent ones
- **COMPROMISE STRATEGY:** bare forms
 - the use of bare forms instead of inflected ones is more typical of multi-word NPs
 - nonce and frequent borrowings behave in the same way
- **HARMONIZATION STRATEGY:** the rate of uses in congruent contexts
 - inconsistent results
 - no effect for any type of Russian insertions?
 - higher rate of congruent contexts for multi-word NPs / for multi-word NPs and nonce borrowings?

4. Discussion and conclusions

Discussion and conclusions

NONCE BORROWING HYPOTHESIS

code-switching \neq [nonce-borrowing = borrowing]

neutralization strategy (morphosyntactic integration) — **IS NOT confirmed**

code-switching > nonce-borrowing > borrowing

compromise strategy (bare forms) — **confirmed**

code-switching \neq [nonce-borrowing = borrowing]

harmonization strategy (congruent contexts) — **partly confirmed**

code-switching = nonce-borrowing = borrowing

code-switching \neq [nonce-borrowing = borrowing]

[code-switching = nonce-borrowing] \neq borrowing

Discussion and conclusions

- **The nonce borrowing hypothesis** was tested on the data of **Russian nouns in Southern Tungusic texts**
 - The same type of contexts as in the papers by Poplack et al. (those with morphosyntactic integration) were checked
 - unexpectedly, the hypothesis **is not confirmed** by our Tungusic data
- Lack of data? No control on dispersion across texts and speakers!
- A specific feature of **languages with rich agglutinative morphology**: real code-switches more easily undergo morphological integration?
- It explains why the rate of multi-word NPs marked with Tungusic affixes is higher than expected
 - It does not explain why the rate of nonce-borrowings marked with Tungusic affixes is lower than expected
- cf. similar results in Stammers & Deuchar (2012): integration of English verbs in Welsh

Discussion and conclusions

- **The nonce borrowing hypothesis** was tested on the data of **Russian nouns in Southern Tungusic texts**
 - It was **expanded** to other types of contexts:
 - congruent contexts — inconsistent results
- how to measure “congruence” in the most appropriate way?
- contexts with compromise strategy (bare forms) — confirmed
- **A general picture is more complicated and less uniform than one might expect**

References

Герасимова А. Н. 2002. Нанайский и ульчский языки в России. Сравнительная характеристика социолингвистической ситуации // Языки коренных народов Сибири, 12. С. 246–257.

Калинина Е. Ю., Оскольская С. О. 2016. Ульчский язык // Михальченко В. Ю. (Гл. ред.) Язык и общество. Энциклопедия. М.: Азбуковник. С. 293–296.

Оскольская С.А., Стойнова Н.М. 2017. Дифференцированное маркирование объекта в нанайском языке // Acta Linguistica Petropolitana, XIII(3). С. 336–370.

Сумбатова Н. Р., Гусев В. Ю. 2016. Ульчский язык // Михальченко В. Ю. (Гл. ред.) Язык и общество. Энциклопедия. М.: Азбуковник. С. 513–515.

Haspelmath M. 2009. Lexical borrowing: Concepts and issues // Haspelmath M., Tadmor U. (Eds.) Loanwords in the World's Languages: A Comparative Handbook. Berlin: Mouton de Gruyter. P. 35–54.

Jake J. L., and C. Myers-Scotton. 1997. Codeswitching and compromise strategies: Implications for lexical structure // International Journal of Bilingualism, 1(1), 25–39.

Myers-Scotton C. 1992. Comparing codeswitching and borrowing // Journal of Multilingual & Multicultural Development, 13(1–2). P. 19–39.

Myers-Scotton C. 1993. Duelling languages: Grammatical structure in codeswitching. Oxford: Clarendon.

Myers-Scotton C. 2002. Contact linguistics: Bilingual encounters and grammatical outcomes. Oxford: Oxford University Press.

References

- Myers-Scotton C., and J. L. Jake. 1995. Matching lemmas in a bilingual language competence and production model: Evidence from intrasentential code switching. *Linguistics*, 33 (5), 981–1024.
- Poplack Sh. 2012. What does the Nonce Borrowing Hypothesis hypothesize? // *Bilingualism: Language and Cognition*, 15(3). P. 644–648.
- Poplack Sh. 2018. *Borrowing: Loanwords in the speech community and in the grammar*. Oxford: Oxford University Press.
- Poplack Sh., Sankoff D., and C. Miller. 1988. The social correlates and linguistic processes of lexical borrowing and assimilation // *Linguistics*, 26(1), 1988. P. 47–104.
- Sankoff D., Poplack Sh., and S. Vanniarajan. 1990. The case of the nonce loan in Tamil // *Language Variation and Change*, 2. P. 71-101.
- Sebba M. 2009. On the notions of congruence and convergence in code-switching // *The Cambridge Handbook of Linguistic Code-switching*. Bullock B. E., Toribio A. (Eds.) Cambridge: Cambridge University Press.
- Stammers J.R. and M. Deuchar. 2012. Testing the nonce borrowing hypothesis: Counterevidence from English-origin verbs in Welsh. *Bilingualism: Language and Cognition*, 15. P. 630-643.