

No leaked!

Machine translation gone wrong or normal bilingual speech?

Anna Verschik (Tallinn University)

EAAL conference, 18-19 April 2024, Tallinn

No lekke-d no leaks-NOM.PL

- A bilingual text advertising humidity absorber on car products trading site (acknowledgements to Geidi Kilp!)
- Looks funny to those who know both English and Estonian (or even only one of them)
- Research question: is it just an idiosyncratic combination of elements and patterns from both languages or the phenomena have been attested in normal, that is, naturalistic Estonian-English bilingual speech and writing?
- Comparison with earlier studies

https://www.autokaubad24.ee/et/tooted/talvekaubad/pingi-niiskuskoguja-300g-indikaator?fbclid=IwAR33O3VEeG2Z_dDaFEexFEb_x9MQ4f-M16kgjGcSlovteZzHvSTO5zdVh7g

 LOGI SISSE  0,00€

hooajaks

Uued tooted

27466

Autoakud

2335

Autokeemia, lisandid, remontained

10106

Autotarvikud, akulaadijad, tuuning

52549

Autopesu ja hoolduse tarvikud

3152

Auto varuosad universaalsed

11391

Auto klaasipühkijad ja nende tarvikud

1824

Autokaubad sügis/talve hooajaks

355

Jõulutuled, jõulutarbed

133

Akulaadijad

659

Esiklaasi kaitsekatted

48

Defa, Webasto, Calix soojendussüsteemid

1960

Koik tooted saadaval ainult e-poes



9,60€

10,20€ -6%

Tarneaeg väljastuspunkti 1tp / koju 2-3tp

LISA OSTUKORVI

Toote ID: 517380

Kood: PINGI

[Lisa võrdlusesse](#)

Innovaatiline niiskuseimaja eemaldab autosalongist kiirelt liigse niiskuse. Hoiab salongi kuivana, muudab udused aknad läbipaistvaks ja kõrvaldab ka halvad lõhnad. Toode on korduv kasutatav. Toode on vedelikku täis kui sinine indikaatortäpp on muutunud roosaks, siis kuumuta padjakest mikrolaine ahjus minimaalselt 600W juures umbes 4-minutit ja lase jahtuda. Peale seda on imaja valmis uueks kasutamiseks.

Over aeg a car interior accumulates quite a lot of humidity, fogging up windows ja making it peaaegu impossible to see road. Foggy windows are a not merely a nuisance – every car owner will agree that driving under such conditions on äärmiselt dangerous. innovative PINGI Car- Dehumidifier battles this problem valiantly, attracting excess moisture like a magnet, ja effectively absorbing it. toode eliminates excess moisture, ensuring comfort ja ohutus of starting your daily commute koos a clear windshield. toode will ka eliminate any bad odours that may have accumulated in car interior. Your sõiduk will have sama fresh smell as kui you just bought it. PINGI Car-Dehumidifier hits kaks birds koos üks kivi, making your driving experience safer, ja that much more enjoyable.

Departure points

- Contact-induced language change is not chaotic but is not governed by universal constraints
- It is useful if at least one of the languages in the pair has rich inflectional morphology (Estonian in this case)
- Bilingual speech and writing is not a combination of two monolingual lexicons and grammars
- Earlier research: Backus (2005) shows that CS items may affect morphosyntax

- *No lekked*
- Estonian *ei mingeid lekkeid, pole lekkeid* etc
lekke-id leak-PART.PL
- The English particle „drags along“ morphosyntactic properties
- English has no grammatical cases but has plural marking
- Hence, the Estonian noun receives only plural marker (technically, NOM.PL)

Contact phenomena in the text (1)

English preposition + Estonian noun

Estonian clause, English preposition

- EST *saadaval* **in** *kaks mõõdud* ‘available in two sizes’
- EST *sobib* **on** *esimene paneel* ‘fits on the first panel’

English clause and preposition, Estonian insertion

- ENG **an indicator on** *toode* ‘product’ (*toote peal*)
- ENG **to recharge by** *vahetus* ‘change’ (*vahetuse teel/kaudu*)
- ENG **will tell you it on** *aeg* (+ semantics?) ‘time’ (*õigel ajal*, literally ‘on the right time’)

Attested in naturalistic speech

*Shout out to **kõik kollase-d/ (...)** **värvilise-d riide-d** minu kapi-s*

Shout out to all:NOM yellow-NOM PL coloured-NOM PL clothes my closet-INES

‘Shout out to all yellow/(...)/coloured clothing items in my closet’
(Verschik and Kask 2019)

*kiri **from** Rimmel* ‘(a/the) letter from Rimmel’

Estonian *kiri Rimmeli-lt* (Rimmel-ABL)

(Verschik 2019)

Contact phenomena in the text (2)

- Estonian content word insertions in English matrix
- Estonian conjunctions (copying/borrowing of conjunctions is usual in language contacts)

Over **aeg (time)** a car interior accumulates quite a lot of humidity, fogging up windows **ja (and)** making it **peaaegu (almost)** impossible to see road. Foggy windows are a not merely a nuisance – every car owner will agree that driving under such conditions on **äärmiselt (extremely)** dangerous.

Estonian content word insertions in naturalistic speech

- In this combination, not impossible but probably infrequent because bilingual speakers investigated so far are Estonians with English L2 and the tendency is either for English insertions in Estonian matrix or code-alternation
- Anecdotal evidence: an e-mail (2013) sent by my student to me (our language was English but she is fluent in Estonian)

*„Hi Anna, when would be a good day to collect my **juhendaja arvamus** (supervisor’s opinion) and get your signature to the **aruanne** (report)? Also what room should I go to get the **kinnitatud väljavõte läbitud ainetest** (confirmed certificate of passed subjects)?“*

Contact phenomena in the text (3)

Mixed copies

- Multi-word items (compounds, fixed expressions) in general may yield insertion/alternation, so-called loan-translation (selective copies, word for word renditions) and mixed copies (Johanson 2002)
- MC: the combination and overall meaning is from language A while some elements are from language A and others from language B

as if as > *as kui as*, cf. Estonian *kui* ‘if’

over time > *over aeg*, cf. Estonian *aeg* ‘time’

Idioms (subject to mixed copying)

two birds with one stone > *kaks birds koos üks kivi* cf. Estonian *koos* ‘with/together’, *üks* ‘one’, *kivi* ‘stone’

- Numerals + partitive sg *aks lindu* ‘two birds’ (a compromise construction – English insertion retains plural but no Estonian case marking added; *kaks birds* ‘i?’) –recall also *in kaks mõõdud* ‘in two sizes’ (circumstantial evidence from Yiddish-Estonian CS: štrajtn tsvišn kaks erakonnad ‘dispute between two parties’)
- Unusual choice of preposition, probably because *koos* means also ‘together’ (*lähme koos* ‘let’s go together’)
- Monolingual Estonian: instrumental (cf. the equivalent idiom *kaks kärbest ühe hoobiga* ‘two flies with one blow’)

Kaks birds with üks kivi? Kaks birds ‘i with üks kivi? Kaks bird ‘i with üks kivi?

Kaks birds ühe kiviga?

Mixed copies in naturalistic speech

*See oli nagu **jõulu-look***

‘It was like a Christmas look’ (Estonian *jõulu* ‘Christmas, of Christmas’)

*Siis tõlkis mulle 8 aastane **fluent-rääkija** ära selle lause*

‘Then a 8 years old fluent speaker translated the sentence to me’
(Estonian *rääkija* ‘speaker’)

(Verschik and Kask 2019)

***hard-töötav** ‘hard-working’* (Estonian *töötav* ‘working’)

English models (combination of items and overall meaning)

- Mixed copied are present in Estonian-English bilingual speech and writing
- But they are of a different nature
- No multi-word conjunctions (*as if as*)
- Mixed copies of idioms have not been attested

„Wrong“

- One instance of Estonian N GEN + N NOM

toode kirjeldus ‘product description’, cf. *toote kirjeldus*

- Unproductive declension type (NOM *toode*: GEN *toote*)?
- Non-existent verbal stem

*Will never **leke*** ‘will never leak’, cf. Estonian *Ei leki kunagi*

- Estonian verb ***lekkima, lekkida, lekib*** ‘to leak’

lekib ‘leaks’, *ei leki* ‘does not leak, will not leak’

Will never lekib? Will never leki? (sounds less probable because English *never* yields affirmative: real example *ma never ostan seda* ‘I never buy this’)

- Is the logic here that in English *leak* is both verb and noun?

Discussion and preliminary conclusions

- Primary candidates for insertions are (1) semantically specific items (mostly nouns, term-like items); (2) expressive words and figurative expressions; (3) items crucial for shaping and directing the discourse (discourse markers, incl. conjunctions)
- I.e., based on a type of meaning!
- Do insertions in this case fit into these categories (yes for conjunctions)? Does machine translation differentiate between various types of meaning?

- We need acceptability tests in order to establish whether bilingual constructions feel right (cf. method elaborated in Verschik 2006, Bahtina et al. 2021)
- More general: the text is not idiosyncratic or chaotic, it has many characteristics of naturalistic bilingual Estonian-English speech and writing
- The question is about frequency and distribution

References

- Backus, Ad 2001. The role of semantic specificity in insertional codeswitching: evidence from Dutch-Turkish. – Codeswitching Worldwide II. Ed. Rodolfo Jacobson. Berlin, New York: Mouton de Gruyter, 125–154.
- Backus, Ad 2005. Codeswitching and language change: One thing leads to another? – International Journal of Bilingualism 9, 3/4, 307–340.
- Bahtina, Daria, Kask, Helin, Verschik, Anna. 2021. English adjectives and Estonian nouns: Looking for agreement? Frontiers in Psychology,
<https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2021.735232/full>
- Johanson, Lars 2002. Contact-induced linguistic change in a code-copying framework. – Language Change: The Interplay of Internal, External and Extra-Linguistic Factors (Contributions to the Sociology of Language 86). Eds. Mari C. Jones, Edith Esch. Berlin: Mouton de Gruyter, 285–313.
- Kask, Helin. 2021. English-Estonian Code-Copying in Estonian Blogs and Vlogs. Tallinn: Tallinn University Press.
- Verschik, Anna. 2006. Convergence in Estonia's Russian: directional vs. static vs. separative verb. International Journal of Bilingualism. 10, 383–404. doi: 10.1177/13670069060100040101
- Verschik, Anna. 2019. English-Estonian Code-Copying in Blogs: Combining a Contact Linguistic and Cognitive Approach. — Cognitive Contact Linguistics, Berlin, 51—80.
- Verschik, Anna, and Kask, Helin. 2019. Estonian–English code alternation in fashion blogs: structure, norms and meaning, in Multilingualism in the Baltic States. Societal Discourses, Language Policies and Contact Phenomena, ed. S. Lazdina and H. F. Marten (London: Palgrave Macmillan UK)

- Thank you for tähelepanu!
- Aitäh for your attention!