Borrowing modals across Slavic

Igor Yanovich

DFG Center for Advanced Study “Words, Bones, Genes and Tools”, Universität Tübingen

FDSL 12 @ Humboldt-Universität zu Berlin

December 8, 2016
1. Borrowing modals: problem statement

2. Borrowing German *müssen* into Old Czech, Old Polish, Old Ukrainian

3. Polish *warto*, Ukrainian *varto*, Russian *stoit*: mutual influences

4. Conclusions and consequences
Borrowing modals: problem statement
A common misconception

“Grammatical elements cannot, or almost cannot, be borrowed”

- Wrong even for derivational and inflectional morphology
  see a recent overview in [Gardani et al., 2015]

- Borrowing of forms: conjunctions, *only*, complementizers, *modals*...

- Borrowing of patterns: perhaps even more widespread

Quick entry into the literature: [Aikhenvald, 2006], [Matras, 2009]
Borrowing modals: actually, quite a lot

**Borrowing of forms:**
- Russian *nado* ‘need’ ⇒ Evenki *na:da* [Grenoble, 2000]
- Persian *bā:yad* ‘have to’ ⇒ Iranian Azeri *bā:yad* [Kıral, 2005]
- Norse *munu* ⇒ Northern Middle English *mun* [Eitelmann, 2013] a.o.

**Borrowing of patterns:**
- GET-based modality in the Circumbaltic area, in South-East Asia
  [Enfield, 2003], [Kehayov and Torn, 2005], [van der Auwera et al., 2009], [Yanovich, 2016]
What we know and don’t know about borrowing modals

Modal borrowing hierarchy of [Matras, 2007]

deontic necessity > circ. necessity > circ. possibility > ability

- [Elšik and Matras, 2006]: over 70 Romani dialects in contact with two dozens different European languages
- [Matras and Sakel, 2007]: 30 languages under contact from around the world (Katanga Swahili, northern Kurdish, Indonesian...)

Problems and open questions for the hierarchy

- presupposes that modals have only one meaning
  - we don’t know how ambiguous modals get borrowed (cf. Evenki *na:da*)
- is based on synchronic descriptions
- doesn’t tell us anything about pattern borrowing
Further complications

Changes in modal meanings

- need  
- ability  
- circumstantial  
- epistemic  
- deontic

[Bybee et al., 1994], [van der Auwera and Plungian, 1998], and subsequent literature

- meanings of *can* today $\approx$ meanings of *may* in the 14th century

  corpus evidence about the 14th century: [Gotti et al., 2002]

$\Rightarrow$ work with primary historical texts is indispensable
Further complications

**Borrowing does not have to fill gaps**

Two broad motivations for borrowing:
- To fill a gap (e.g., European coffee from Arabic and Turkish)
- To align or even fuse elements from two systems

A striking example: from [Matras, 2009]

- Medieval Romani borrowed *tajśa* from Greek *taixiá* ‘tomorrow’.
- But Romani *tajśa* means both ‘tomorrow’ and ‘yesterday’.
- Indo-Aryan *kal* meant ‘one day away from today’; still exists in Hindi.

Thus Romani replaced *kal* with Greek-based *tajśa*, but preserved *kal*’s meaning.

⇒ we cannot assume that a borrowed modal filled some gap in the system
Further complications

Unclear specific motivation for borrowing

Some grammatical borrowings pertain to discoursive control
  e.g. conjunctions, modal particles (see [Matras, 2009])
    documented to occur as slips in the speech of individual bilinguals

But this does not help for many other frequent borrowings.
  Why *already* > *still*? Why *only* > *too*? Why superlatives > positives?
  Why modals in the first place?
  Why deontic necessity > circumstantial necessity?
  Which deontic — objective (common rules) or subjective (“my rules”)?

⇒ we need to establish how modals get borrowed in the first place
Borrowing modals: problem statement

Main imperative

borrowing through the prism of a bilingual speaker (look at the micro-level!)
Borrowing German *müssen* into Old Czech, Old Polish, Old Ukrainian
### German müssens’s “success story”

#### Slavic descendants of müssens

<table>
<thead>
<tr>
<th>Language</th>
<th>Descendant</th>
<th>Date and Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Czech</td>
<td>muset</td>
<td>Already in the Dalimil chronicle (early 14th)</td>
</tr>
<tr>
<td>Polish</td>
<td>musieć</td>
<td>Already in Kazania Gnieźnieńskie (late 14th)</td>
</tr>
<tr>
<td>Ukrainian</td>
<td>musyty</td>
<td>Attested once in Moldovan documents in the early 15th; Attested safely in the 16th century</td>
</tr>
<tr>
<td>Slovak</td>
<td>musiet'</td>
<td>Historical Russian (the Dictionary of the Russian of the 11-17 cent. provides two citations, from 1488 and 1708)</td>
</tr>
<tr>
<td>Belarussian</td>
<td>music’</td>
<td></td>
</tr>
</tbody>
</table>

[Hansen, 2000]: müssens was borrowed because Slavic languages lacked a polyfunctional necessity modal without lexical-semantic overtones

- Based on the Old Church Slavonic situation, not the synchronic one
- Is not accompanied by a comprehensive analysis of the modal system

[Hansen, 2001] devotes 1.5 pages to the Old Polish musić, but does not comment on obvious semantic overlap with mieć
musiti in Old Czech

Recall our schema: 1) sociolinguistic situation; 2) meanings; 3) modal system; 4) contexts; 5) diachrony.

Short history of early Czech-German contacts

- Contacts among the nobility
  castle named Rýzmburk (=Riesenburg) started in the mid-13th
  Bruno von Schauenberg: bishop of Olomouc in the mid-13th

- Colonization of non-settled parts of the Czech lands
  quite extensive during the 13th

- Towns start to appear in the 13th, with German burghers participating
  Uničov founded in 1213 by a certain Theodoricus
  Litoměřice, 1253: Lutold the reeve, Johannes from Meissen, Henrich from Freiberg

⇒ plenty of routes for linguistic influence...
musiti in Old Czech: Dalimil

The Chronicle of the so-called Dalimil (henceforth Dalimil):

- an impressive rhymed history of the Czech lands
- one of the first long works in Old Czech: composed around 1310s
- expresses quite negative sentiment towards Germans; promotes ethnic-based solidarity for Czechs

...extensively features musiti

musiti ‘have to’ in Dalimil

musiti: a great variety of meanings

- circumstantial
- deontic objective
- deontic subjective
- future, conditional “would”
musiti in Old Czech: Dalimil

(1) **Objective deontic:**
“If you dared to say this about the emperor, ...”

\[
\text{musíš toho svým mečem dolíčiti musiti.2sg that.ACC one’s.own.INST sword.INST to.prove}
\]

‘you have to prove it with your sword’  
(Dal 7r)

(2) **Circumstantial:**
*Since nobody wrote a comprehensive chronicle of the Czechs before, ...*

Proto sě sám v to musím uvázati 
that’s.why REFL myself PART musiti.1sg take.upon

‘That’s why I have to take this upon myself.’  
(Dal 1r)

(3) **Future and “would”:**

Rod mój musil by toho škodu jmieti 
kin my musiti.PTPL SUBJ harm.ACC to.have

‘My kin would come to harm from that.’  
(Dal 3r)
What about the modal system in *Dalimil*?

- *jmieti* ‘have’: a common Slavic modal
  often deontic and futurate meanings

- In *Dalimil*, modal *jmieti* is extremely rare.
  However, it overlaps semantically with *musiti* there:

(4) “But ignore those words, because…”

jmám o Durynkovi mluviti
have.1sg about Durynk to.speak

‘(because now) I have to/will speak about Durynk’

(5) “Then the prince fenced the Prague town with walls, and his son Spytihněv…”

okolo svatého Jiřie hraditi jmieše
near saint George to.fence have.IMP.3sg

‘would/would have to fence [the place] near Saint George’
Old Czech *musiti* and *jmieti* in *Hradecký rukopis*

- **Hradecký rukopis** (further *Hrad*): a collection of texts written in mid-14th century ⇒ about 50 years later than *Dalimil*

- Both *musiti* and *jmieti* present

  *jmieti* in *Hrad* more prominent than in *Dalimil*

- Extremely clear semantic overlap: objective deontic

(6) V svatého Benedichta zákoně *musíš* jmieti utrpěnie in saint Benedict’s law *musiti.2sg* to.have hardship

  ‘By the Benedictine rule, you have to experience hardship’  
  (Hrad 4v:140)

(7) “If some woman gave birth to a son, the first firstborn, ...”

  *jmieše* vyplatiti jeho beránkem

  *have.to.IMP.3sg* to.pay for.him with.a.lamb

  ‘then she had to pay for him with a lamb’  
  (Hrad 70r:355)
A difficulty: *musiti* already well attested in the earliest long texts (*Dalimil, Hrad*; also the 1st Czech redaction of the Bible, mid-13th) ⇒ we can’t see the pre-contact situation

*musiti* features a broad range of meanings already: circumstantial, obj. deontic, subj. deontic, futurate

*jmieti* is also polysemous. I’m not aware of circumstantial uses, though.

No semantic complementarity obvious in the data.
Synthesis on the Old Czech situation

- Really intensive Cz-Ge contact only starts in the early 13th century
- A highly polysemous *musiti* already in *Dalimil*, i.e. in the 1310s

\[\Rightarrow \text{it follows that } \textit{musiti} \text{ had to be borrowed with many meanings} \]
A century is just not enough for grammaticalization to develop such a wide array

My proposal: substitution plus variation

- *musiti* entered Czech as an alternative for *jmieti*
- *jmieti* and *musiti* coexisted, exhibiting sociolinguistic variation
- *musiti* faithfully copies *jmieti*’s meanings, but possibly adds circumstantial necessity, most likely present in its German source

Comparison with [Hansen, 2000]:

i) confirm polysemous borrowing; ii) update on the motivation: *not* gap filling
Quick tour through Old Polish and Old Ukrainian

**musieć in Old Polish**
- already in *Kazania Gnieźnieńskie* (late 14th)
- 5 instances of *musieć*: deontic, futurate
- in one case, *musieć* receives as a gloss... *mieć!*

⇒ looks similar to the Czech situation, but need more research

**musyty in Old Ukrainian/Belarusian**
- apparently only a single instance in the 15th (from Suceava, Romania)
- only one instance in the 1st Statute of the Grand Duchy of Lithuania, year 1529
  - deontic/circumstantial; impersonal construction
- significantly attested in the mid-16th:
  - (circumstantial; deontic)
    - Lutsjk Castle book: 3 instances
    - Peresopnytsjke Gospel: 4 instances just in Matthew
Further support for substitution plus variation: dialects

Dialectological maps from [Kuzmina, 1988]:

- Map 1: *maty* in the west, *musyty* in the center for ‘have to’
- Map 2: boxes show *musyty*-based constructions for ‘need’

A plausible scenario:

- After borrowing, variation between *musiti* and *jmieti* starts
- Over time, different dialects fix one of the variants
- Neighbor dialects tend to fix variation the same way
Polish *warto*, Ukrainian *varto*, Russian *stoit*: mutual influences
From worth to action

Symbouletic (=advice) modality  

(8) Anna to Mary: You really should go to that concert.
(9) 8 reported: Anna suggested/recommended that Mary go to the concert.

Grammaticalization of symbouletics from worth statements

Stage 1: WORTH(p) conveys p’s (literal or metaphorical) worth.

Stage 2: in situations where choice of action is at stake, statement “p is worthy” implicates “x taking action towards p would do the right thing” (≈ x should do p)

Stage 3: the implicature is reanalyzed as the primary meaning of WORTH

[Yanovich, 2013, Ch. 5.4] postulates this schema based on Russian stoit
From worth to action

- Not clear what triggers grammaticalization:

  (10) It’s worth going to that concert \(\neq\) You should go to that concert.

- In Russian, however, the evolution was extremely quick:

  1. **Zero-stage**: stoitj means ‘to cost’, ‘to be worth X’
     
     “You stoitj not to talk about”; “What did leading you stoitj to us?”

  2. **since 1830s**: action-guiding implicatures; the worth component present.
     
     “It stoitj not (the effort) even to mention them”

  3. **by the early 20th century**: stoit also without the worth component
     
     Example: “It would be good to ring the bells! Which day is it? Wednesday? If Wednesday, then stoit not.” \(\Rightarrow\) not about worth
The evolution of Russian \textit{stoitj/stoit}

During the implicature stage, many examples with “to \textit{stoitj} the effort”
⇒ a simple test for metaphorical worth: can we add “\textit{the effort}”?

Table 5.1: Evolving distribution of INF-taking \textit{stoit}

<table>
<thead>
<tr>
<th></th>
<th>1841-1845</th>
<th></th>
<th>1914</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>POS</td>
<td>NEG</td>
<td>QUE</td>
<td>POS</td>
</tr>
<tr>
<td>\textit{čego ‘what.ACC’/ničego ‘nothing.ACC’}</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>other ACC noun</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>\textit{truda ‘effort.ACC’}</td>
<td>11</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>can add \textit{truda}</td>
<td>9</td>
<td>12</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td>cannot add \textit{truda}</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Total:</td>
<td>28</td>
<td>17</td>
<td>9</td>
<td>24</td>
</tr>
</tbody>
</table>

from [Yanovich, 2013]
Changes in *stoitj/stoit* across the lifespan

- For example, in the letters by Leo Tolstoy (1828–1910):

  (11) **can add the effort:**
  
  *Context: after enumerating several flaws in a text he was asked to provide advice on. “Perhaps sometimes also using incorrect language”.*

  (1873) No pro eto ne stoit govoritj.
  
  but about this.ACC NEG stoit to.speak
  
  ‘But it’s not worth mentioning. / But it shouldn’t be mentioned.’

  (12) **the first example where the effort cannot be added, with no obvious worth component:**
  
  *Context: discussing the new part of a building.*

  (1887) Po-moemu ne stoit lučšé i otdelyvatj.
  
  in.my.opinion NEG stoit better even to.decorate
  
  ‘In my opinion, we should better not even decorate it (any more)’

- In Dostoevsky (1821–1881), an example with new *stoit* in 1876.
Russian: a quick transformation of *stoit*

- Even in the 20th century, many *stoit* examples are ambiguous between a worth verb and a grammaticalized modal.

- But some clearly lack any worth component:

  (13) **Context:** But if objects and real, very specific circumstances serve as the cause, then one needs to use another preposition IZ-ZA. “The bus didn’t start the route IZ-ZA technical problems”

  (1980) *(ne stoit govoritj “PO texničeskim neispravnostjam”).*  
  *NEG stoit to.say due.to technical problems*

  ‘One **should not** say “PO technical problems”.’ *(from NatRusCorpus)*
Polish \textit{warto}: a long preservation of the implicature state

- Polish \textit{warto} ‘worth.ADJ’ $<$ German \textit{Wert}

- Already in Mickiewicz, we find potential for action-reanalysis of \textit{warto}:

  \begin{quote}
  (14) from \textit{Pan Tadeusz} (1834):
  \textit{Context}: While thinking so, she walked several times along the alcove, boldly and with joy. Then she lowered her head again.
  \end{quote}

  \textit{Warto by też pomyślić o Hrabiego losie... warto SUBJ also think about Hrabia.GEN fate.ACC}

  ‘Perhaps it makes sense to also think about the fate of Hrabia.’
  Why, maybe it’d work to get him Zosia? $< ... >$ for the young couple, she would be like a mother.

  $\Rightarrow$ arguably, the main point is action-related, not worth-related
Polish *warto*: a long preservation of the implicature state

- But even today we don’t find a Russian-like new modal in Polish:

  (15)  
  \[ \text{OK} \quad \text{warto}, \text{żebyś} \quad \text{poszła do lekarza}. \]
  \[ \text{warto} \quad \text{that.be.SUBJ} \quad \text{go} \quad \text{to doctor} \]
  ‘It’s worth it that you go to the doctor.’

  (16)  
  \[ \# \quad \text{warto}, \text{żebyś} \quad \text{poszła na ten \ ten koncert}. \]
  \[ \text{warto} \quad \text{that.be.SUBJ} \quad \text{go} \quad \text{to that concert} \]

- Russian: rapid grammaticalization from the implicature stage
- Polish: long preservation of the implicature stage
Ukrainian varto: grammaticalization following Russian

- Ukrainian varto < Polish warto
- Unlike in Polish, Ukr. varto undergoes grammaticalization

Lesja Ukrajinka’s letters from 1876 to 1897:
- 19 instances of varto, but all have a worth component

Valerjan Pidmohyljnyj (1901-1937):
- some examples apparently without the worth component

(17) Context: “< ... > the object of his desires was to get securely settled in life, and if that was possible without a degree,”

to vyšču školu varto vidtjaty
then higher school varto cut.away

‘then the university should be cut away’ (cf. #‘is worth cutting away’)

Continuation: “…as an unnecessary addition, similar to the appendix.”

⇒ no special effort required to leave school; no value in cutting it
Halyna Pahutjak (1958–): often no worth component

(18) from *The Gothic of Urizh* (2009)

Tam, na hori, vin dumav, ščo dejaki tajemnycji ne varto
there on mountain he thought that some mysteries NEG varto
rozkryvaty:
to.solve

‘While there on the mountain, he thought that some mysteries should not
be solved:’

*Continuation*: “one should (Ukr. *slid*) make peace with the reality that
does not want to submit itself to the human senses.”

⇒ *varto* is used synonymously with *slid* ‘should’
⇒ this is not about a lack of worth in solving mysteries
**warto, stoit, varto: the complex interplay**

1. Ukrainian borrowed *varto* from Polish
2. Polish may have provided an “implicature target” for Russian
3. Then Russian pulled Ukrainian towards the “action target”
**warto, stoit, varto: the complex interplay**

- **Polish → Ukrainian:** borrowing of form
- **Polish → Russian:** borrowing of pattern (?)
- **Russian → Ukrainian:** borrowing of pattern

**Open question:** why did Russian change, but Polish didn’t?

**Heavy Russian → Ukrainian influence in another case:** [Yanovich, 2015]

- late 19th century Ukr. *maty* ‘have’: futurate, necessity, possibility
- after mid-20th century: *maty* loses possibility
- Russian lacks a three-way, fut-nec-pos ambiguity
- but Russian lacks nec-fut ambiguity, too!

⇒ erosion in Ukrainian, but not all the way to system identity
Conclusions and consequences
Conclusions and consequences

Conclusions

Polysemy is important in modal borrowing

- Czech *musiti*: borrowed as polysemous
- Ukrainian *varto*: adding a new meaning under contact
- Ukrainian *maty*: losing an old meaning under contact

Role of gap filling unclear

- Czech *musiti*: clearly not gap filling
- Russian *stoit*: “grammatical accommodation” to Polish? gap filling? perhaps both?
Conclusions

Pressure on bilinguals to increase alignment is clear in some cases
- Czech *musiti*: form alignment with German
- Ukrainian *varto, maty*: pattern alignment with Russian
- Polish and Russian: questionable (no widespread bilingualism)

BIG remaining open question
Why modality???

Important issue for future contact-linguistic research
Tension between:
- The reality of Matras’s borrowing hierarchy
- The fact that modals can get borrowed with multiple meanings
Conclusions and consequences

Consequences

Our contribution today

We’ve shown with fine-grained historical data that:

- modals can be borrowed with several meanings
- modals under contact can align their sets of meanings

⇒ an important link for solving a long-standing puzzle

Overlap between epistemic and root modality

[van der Auwera and Ammann, 2013]
Thank you!
This work would not have been feasible if not for the enormous efforts of people creating historical corpora and digitizing historical texts. In particular, I would like to sincerely thank: the members of the language history department of the Institute of the Czech language, who maintain an excellent set of digitized dictionaries and an extensive Old Czech corpus; the creators of the PolDi and KTS corpora of Old Polish, respectively at Regensburg and the Old Polish department of the Institute of the Polish language; the team of the Russian National Corpus; and, last but not least, the often anonymous enthusiasts who scanned and digitized an enormous number of public-domain books in Ukrainian and Polish that I benefitted from.

Many thanks to Natasha Korotkova, Anna Szabolcsi, and Basia Tomaszewicz for comments, judgements, and discussions which turned out to be very relevant to this project even though not all of them were conceived so at the time; and to my colleagues at the “Words, Bones, Genes and Tools” research center and at the UniTübingen Institute of the Prehistory, who help me to contextualize archaeological research that can bear on linguistics.

I gratefully acknowledge the support by DFG under FOR 2237 (DFG Center for Advanced Study “Words, Bones, Genes and Tools”).
Grammars in contact: A cross-linguistic perspective.

The evolution of grammar: Tense, aspect and modality in the languages of the world.
University of Chicago Press, Chicago.

Remembering (ge)mun.

Markedness and language change: the Romani Sample.
Mouton de Gruyter, Berlin.

Linguistic epidemiology: semantics and grammar of language contact in mainland Southeast Asia.

Borrowed Morphology.
Walter de Gruyter, Berlin.

Variation in central modals.
Peter Lang, Bern.

Morphosyntactic change: the impact of Russian on Evenki.


