

Peripheral Czech modal verb + infinitive constructions

Slavic Corpus Linguistics: The Historical Dimension

The Arctic University of Norway, Tromsø

František Martínek

Charles University in Prague, Faculty of Arts

April 21, 2015

Introduction

- New electronic sources filled with historical Czech texts make it possible to describe small shifts in the language development of Czech and in this way to understand language changes.
- new electronic sources: corpora, digitalized dictionaries and word indexes (“lexical archives”)
- peripheral phenomena: not enough evidence (proofs) available in the corpora of historical language

Motivation for this study (1/2)

1. Editing of the correspondence of Karel Havlíček (1821–1856)
 2. Characterization of the middle of the 19th century Czech
- Over 1100 letters; about a half of them included to a preliminary version of a corpus: over 292 000 tokens = 241 000 words
 - The corpus is running in a Sketch-Engine interface

Motivation for this study (2/2)

The Czech National Revival

Instability in many levels of language can be supposed:

spontaneous/intended neologization in word formation, literally translation, syntactic calquing...

- When did the stabilization come, after such an unrestrained development?
- R. Adam about correspondence of B. Němcová:
a „czechising of the syntax“ in the 1840es

Task of the talk

- semantical development in the field of the **volitive modality** (intention, necessity, possibility and ability – thus Miroslav Grepl):
- the development of one less frequent full verb to modal (*hodlat*) and one contrary development (*uspět*)
- Various Czech constructions of modal verbs with infinitive, expressing volitive modality, have been described already
 - For example, the infinitive in a syntactic structure with verbs *být* and *mít* expressing possibility like *Mám/Je kde spát.* ‚I have / There is a place for sleeping.‘ (Karlík - Štícha 2011); see also below.

Corpora and other sources (1/3)

- DIAKORP
- the entire duration of Czech, represented by more than 1 900 000 tokens – starting with the oldest texts and ending with almost today’s language
- here 26 hits of *hodlati* [query: “hodl.*”] from only 4 various sources, all of them with today’s meaning and with an infinitive valency.
- Sources - all after 1850:
 - Obrazy života* (journal; 1861, 12 hits)
 - Český študent* (journal; 1869, 8 hits)
 - Filosofská historie* (novel from Alois Jirásek; 1890, 1 hit)
 - Kouzelný dům* (novel from Karel Beneš; 1939, 5 hits)

Corpora and other sources (2/3)

- “The Old Czech Text Bank” (up to the end of 15th century), over 4,4 Mio. word forms
- “The Middle Czech Text Bank” (16th – 18th century), about 450 000 word forms
both: no evidence for *hodlati*

Other text collections of humanistic and baroque texts:

- texts between 1500 and 1620: about 600 000 word forms; supposed to be incorporated to DIAKORP
- texts between 1620 and 1780: transcribed by Pavel Kosek (Masaryk University, Brno)
both: no evidence for *hodlati*
- “Archive of the 19th century” developed in the Czech National Corpus Institute:
about 32 500 000 text words; a half of them: digitalized data from the *Příruční slovník jazyka českého*

Corpora and other sources (3/3)

- The Lexical Database of Humanistic and Baroque Czech
about 750 000 cards with excerpted words

Dictionaries:

- Václav Jan Rosa – 2nd half of the 17th century
- Josef Jungmann – 1834–1839
- František Štěpán Kott – end of the 19th century
- “academical” dictionaries from the 20th century, inclusive a database of their word index (cards with excerpted examples)

hodlati (1/4)

- An earlier form: *hodlovati*; rarely attested in the Old Czech:
- In Tomáš Štítný's writings (end of the 14th century) – the original meaning 'to adjust sth.'; joined by concrete, materially object
- In Řehoř Hrubý's z Jelení translations (about 1500) – semantic shift to: '[to tailor sth.] to make sth. appropriate'; joined by an abstract object
- designating of a specific "constructional activity" → a generalized action verb with the meaning 'to prepare'

hodlati (2/4)

- The form *hodlati* is first attested in Middle Czech (1514)
- ‘to prepare (sth., concrete → abstract)’
↓
‘to intent (to do sth., abstract)’, ‘to want (to do sth.)’
- In the second, grammaticalizational phase, taking place up to the 19th century, the bleached light verb becomes a modal, it stabilizes in this function and its collocability radically changes: now it collocates with verbal infinitives only.

hodlati (3/4)

Corpus of Havlíček's correspondence: 65 hits of *hodlati*

a) *hodlati* + infinitive: 64x

b) *hodlati* + directional adverbial (lexicalized ellipsis of the motion verb?): 1x

The 19th century lexical archive: between 600 hits 37* other than the group a):

- 16x with directional adverbial

- *A kam ty letos hodláš? Nenevštívíš Prahu?* (Němcová)
,Where do you want [to go] this year? Will you not visit Prague?'

- 21x ,intent' without infinitive verb

- *Co hodláš, královno?* (Wenzig, 1874) – ,What do you want to do, my queen?'
Zvěděla také, co hodlá. (Jirásek) – ,She came to know what he wanted/intended.'

*) Note that some texts do repeat, are quoted etc.

hodlati (4/4)

- In today's Czech, the verb *hodlat* belongs to the “modal verbs in wider sense”:
 - A potential additional meaning nuance in relation to the “basic” verb *chtít* ‘to want sth.’

uspěti (1/3)

- An opposing process – a lost of modal meaning
- Former meanings (up to 19th century):
 - ‘to cause that sth. [= an action] goes/runs faster’
 - ‘[to manage] to flee, to escape’
 - **‘to achieve, to manage sth.’ [obligatory with an infinitive]**
- During the 20th century, it has lost its ability to collocate with infinitive and also its scale of meanings has been reduced to a single meaning ‘to succeed’ (‘to fail’ in negation, respectively).

uspěti (2/3)

- This verb is first attested in Jungmann's dictionary (1834–39), although more forms derivated from this stem with similar meaning, like *prospěti* 'to benefit sb., to be good for sb.', already occurred in Old Czech.
- In an etymological dictionary by the Czech linguist Jiří Rejzek (2002), the verb *uspět(i)* is classified as a loan word from Eastern or Southern Slavonic languages.
- It is exactly by its unstable position in the system as well as due to the influence of the equally old noun *úspěch* 'success' (and also *spěch* 'hurry') that its semantic and functional shift can probably be explained.
 - one can find a very suspicious example of the derivate *úspěšnost* 'successfulness' in today's Czech in the "Lexical Database of Humanistic and Baroque Czech", dated to 1755: "[tkadlec] **úspěšnost** v díle zdržuje" '[a weaver] delays the success of the work'

uspěti (3/3)

The 19th century lexical archive: 2x modal verb, 18x not m. v.

Co jsem doposud vyzkoumati uspěl, ... (Český lid 1894) – ‚What I was able to search out‘

A přece neuspěl jsem probudit tu sílu Bohem danou ze spánku! (Zeyer) – ‚Thus, I did not succeed to wake up the God power from sleeping!‘

The lexical archive of the *Příruční slovník jazyka českého*:

A huge amount of modal verb usage (in fiction, till 1960s!): *Rozprodala [...] pozemky na Žižkově, které nebožtík [...] neuspěl zastavět.* – ‚She sold out of the estates in Žižkov which the deceased was not able to build up.‘ (1961)

Other – today not usual – usage, without agent as subject: *Plány neuspěly.* – ‚His plans did not come true.‘ (1956)

Conclusions

- small, subtle changes in collocability (semantical valency) regarding the grammatical system
- Possible reasons for use of *hodlat*:
 - “stylistic higher”, specific for more formal texts
 - distance from the intention of another person; irony
 - source of expresivity (peripheral lexical unit)
- The change of *uspět*:
 - reduction of not related meanings
 - as a modal verb till 1960s!

Perspectives

- More sources of linguistic data should be elaborated for historical Czech; thus, it will be possible to describe the shifts in older language more effectively.
- This is the case for the Present-Day Czech, where there have been described some very interesting changes in meaning by means of utilizing extensive synchronic language corpora (ČNK SYN):
- Ivana Kolářová – the usage of several verbal forms of *(po)dívat se*, *hledět* and *koukat (se)*, all meaning ‘to look (at sth.)’, a modal meaning ‘let’s do ...’ arises and some forms, e.g. the imperatives *hled’*, *dívej* etc. have changed to modal particles expressing an appeal
- Mirjam Fried – the shift of the conditional conjunction *jestli* to a modal particle expressing doubts: [Direct speech:] “*Jen jestli tu práci udělal!*” / “*Jestli tu práci vůbec udělal!*” ‘I doubt he did/finished the work/the task.’
- Björn Hansen (and others) – a new meaning of the verb *muset*, ‘not to like’, developed in a special syntactic construction with negation *Karla Gotta nemusím*. ‘I do not like Karel Gott (very much).’

Thank you for your attention.
frantisek.martinek@ff.cuni.cz

Sources

- *Český národní korpus - Diakorp* [on-line]. Available under <https://kontext.korpus.cz>.
- *Lexikální databáze humanistické a barokní češtiny* [on-line]. Available under <http://madla.ujc.cas.cz>.
- *Příruční slovník jazyka českého* [on-line]. Available under <http://psjc.ujc.cas.cz>. [The excerption cards used for newer Czech 20th century dictionaries are also available here.]
- *Vokabulář webový* [on-line]. Available under <http://vokabular.ujc.cas.cz>: contains a digitalized version of Václav Jan Rosa's dictionary, Josef Jungmann's *Slovník česko-německý*; the "Old Czech Text Bank" (over 4,4 Mio. tokens) and the "Middle Czech Text Bank" (over 450 000 tokens), among others.