

# Parallel corpus as functional context of aspectual interpretation

– the case of Slavic biaspectual verbs  
in the comparative context of Finnish

Edyta Jurkiewicz-Rohrbacher *University of Regensburg, Germany*

Elżbieta Kaczmarska *University of Warsaw, Poland*

Alexandr Rosen *Charles University in Prague, Czech Republic*

Parallel Corpora as Digital Resources and Their Applications  
DHN2020 – Digital Humanities in the Nordic Countries  
Riga, October 20, 2020



# Outline

- 1 Biaspectual verbs in Slavic
- 2 About InterCorp
  - Basics
  - Content of *InterCorp*
  - Access to InterCorp
- 3 Examining comparative contexts
- 4 Pilot study
- 5 Discussion
- 6 References

# Outline

- 1 Biaspectual verbs in Slavic
- 2 About InterCorp
  - Basics
  - Content of *InterCorp*
  - Access to InterCorp
- 3 Examining comparative contexts
- 4 Pilot study
- 5 Discussion
- 6 References

# Biaspectual verbs in Slavic

- Aspect is an obligatory lexical category of Slavic verbs
- But some verbs, especially borrowings, appear in both perfective and imperfective contexts
- Prefixes turn some biaspectuals perfective, the original becomes imperfective [[Cockiewicz\(2007\)](#), [Piperski\(2018\)](#)]
- Experiments to test aspect: [[Perlin\(2010\)](#), [Chromý\(2014\)](#)]
- Our approach: compare parallel contexts using a parallel corpus
- Czech and Polish contrasted with Finnish, where aspect is interpretable from argument case marking and tense
- For an aspect-related contrastive study of Polish and Finnish see [[Jurkiewicz-Rohrbacher\(2019\)](#)]
- Other studies of Slavic aspect based on a parallel corpus: [[von Waldenfels\(2012\)](#)]

# Outline

- 1 Biaspectual verbs in Slavic
- 2 About InterCorp
  - Basics
  - Content of *InterCorp*
  - Access to InterCorp
- 3 Examining comparative contexts
- 4 Pilot study
- 5 Discussion
- 6 References

## Basics

- <http://www.korpus.cz/intercorp/>
- A part of the *Czech National Corpus* <https://www.korpus.cz>
- \* 2005
- At first as a service for linguistic and philological departments of Charles University
- On line since 2008
- New releases once per year

## The architecture of *InterCorp*

- Alignment: sentence-level
- Each text in Czech and at least one other language
- Tags and lemmas for most languages
- Rich metadata



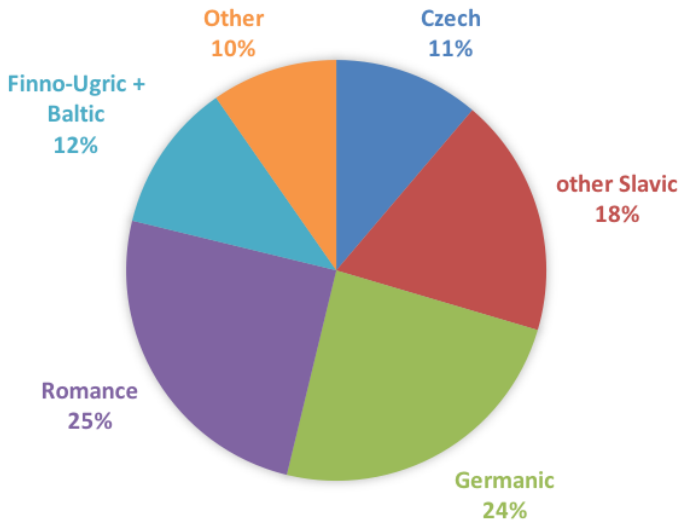
# Content (release 13)

## 40 languages + Czech

- 10 Slavic: **be**, **bg**, **hr**, **mk**, **pl**, **ru**, **sk**, **sl**, **sr**, **uk**
- 7 Germanic: **da**, **de**, **en**, **is**, **nl**, **no**, **sv**
- 6 Romance: **ca**, **es**, **fr**, **it**, **pt**, **ro**
- 5 Finno-Ugric + Baltic: **et**, **fi**, **hu**, **lt**, **lv**
- 12 other: **ar**, **el**, **he**, **hi**, **ja**, **ms**, **mt**, **rn**, **sq**, **tr**, **vi**, **zh**

- ➡ Only few texts are available in more than 20 languages
- ➡ Languages differ wildly in the volumes of text

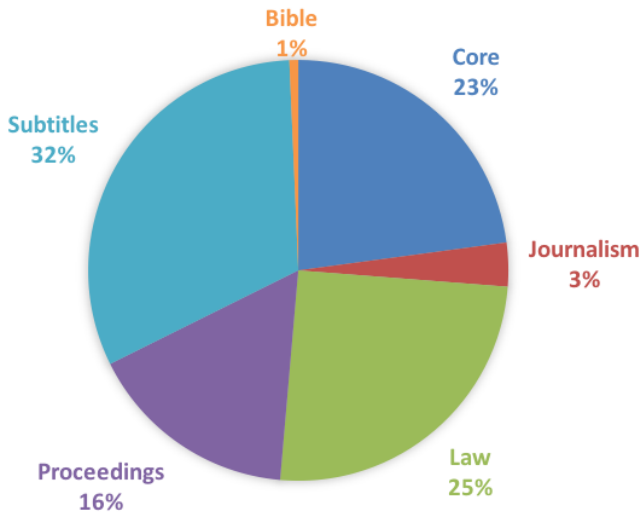
# Language groups



# Text types

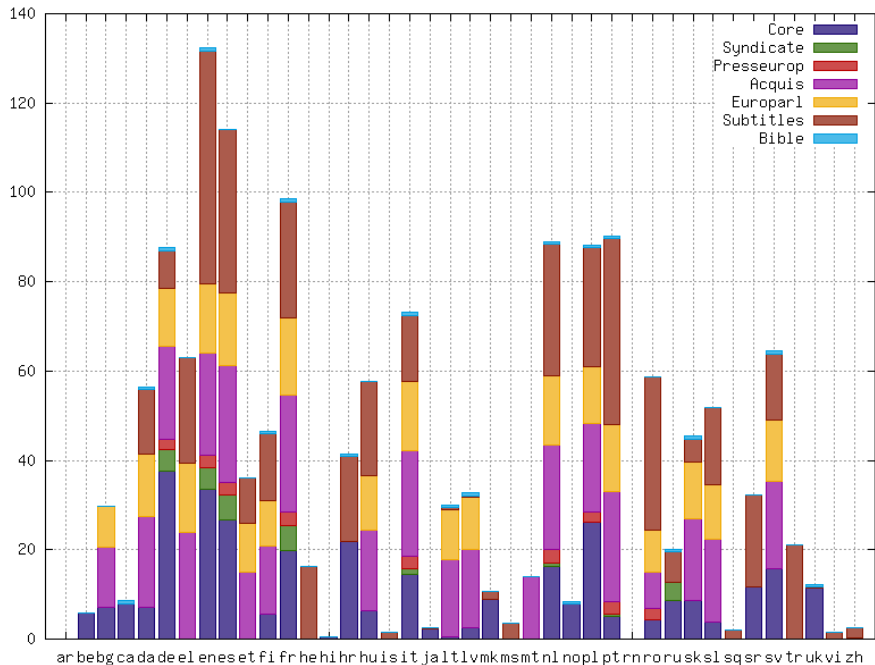
- **Total size** – 1.8 billion words
- **Core** – mostly fiction, proofread
- **Collections** – freely available texts
  - **Journalism**  
*Project Syndicate* <http://www.project-syndicate.org/>  
*VoxEurope* <http://www.voxeurop.eu/>
  - **Law**  
*Acquis Communautaire*  
<http://langtech.jrc.ec.europa.eu/JRC-Acquis.html>
  - **Parliament proceedings**  
*Europarl* <http://www.statmt.org/europarl/>
  - **Film subtitles**  
*Open Subtitles* <http://www.opensubtitles.org>
  - **The Bible**

# Text types

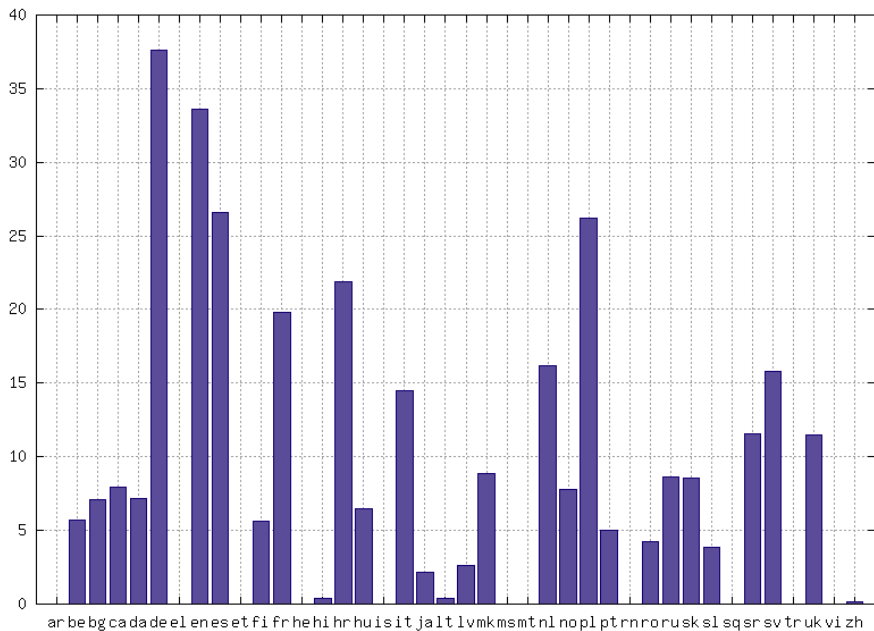


## Size in million words

	Czech	Other Slavic	Nordic +Baltic	Other	<b>Total</b>
Core	113,8	113,9	39,4	174,6	<b>441,7</b>
Journalism	6,7	6,4	0	45,9	<b>58,9</b>
Law	19,0	70,1	104,9	231,5	<b>425,5</b>
Proceedings	12,9	46,9	71,6	145,4	<b>276,8</b>
Subtitles	50,6	97,6	56,9	334,8	<b>539,8</b>
Bible	0.6	2,9	3,6	5.1	<b>12,1</b>
<b>Total</b>	<b>203,6</b>	<b>337,7</b>	<b>276,3</b>	<b>937,2</b>	<b>1754,8</b>



## Core - Fiction



## Number of texts (Core and Subtitles only)

	Czech	Other Slavic	Nordic +Baltic	Other	<b>Total</b>
Core	1 656	1 523	523	1 947	<b>5 649</b>
Subtitles	10 400	19 044	10 848	58 969	<b>99 261</b>



# Concordances

## *KonText*

<http://kontext.korpus.cz>

- Choice of *InterCorp* release
- Search filters:
  - languages, texts, publication year, text type
  - original/translation, language of the original
  - sex of the author and the translator
- Parallel queries, CQL
- Positive and negative filters on the concordances
- Export of concordances
- Sorting, frequency distribution, collocations
- Custom subcorpora

## Some studies based on *InterCorp*

- Romance: [Čermák et al.(2020)]
- German: [Káňa(2016)]
- Croatian: [Jirásek(2011)]
- English: [Malá & Brůhová(2018)]
- Finnish: [Čermáková & Fárová(2010)]
- Polish: [Kaczmarska(2019), Charciarek(2019)]

# Lexical lookup

*Treq – a database of translation equivalents*

<http://kontext.korpus.cz>

- Pairs of lexical equivalents extracted from word-aligned parallel texts
- cs/en ↔ any other language
- Filtering by text groups
- Queries by forms or lemmas
- Support of regular expressions
- Single forms or multi-word expressions

# Outline

- 1 Biaspectual verbs in Slavic
- 2 About InterCorp
  - Basics
  - Content of *InterCorp*
  - Access to InterCorp
- 3 Examining comparative contexts**
- 4 Pilot study
- 5 Discussion
- 6 References


# Parallel texts in Czech, Polish and Finnish

- **InterCorp v.12**: 60 mil. tokens
- **Core**: 34 texts, 3.7 mil. tokens
  - **Finnish**: Mika Waltari, Väinö Linna
  - **Czech**: Ladislav Fuks, Jaroslav Hašek, Bohumil Hrabal, Milan Kundera, Michal Viewegh
  - **Scandinavian**: Jostein Gaarder, Stieg Larsson, Astrid Lindgren
  - **Others**: Douglas Adams, Isabel Allende, Dan Brown, Lewis Carroll, Agatha Christie, Paulo Coelho, Umberto Eco, Francis Scott Fitzgerald, Anna Frank, William Golding, John Grisham, Peter Hoeg, Michel Houellebecq, Alan Alexander Milne, George Orwell, Orhan Pamuk, Matt Ridley, Joanne K. Rowling, Antoine de Saint-Exupéry, William Styron, J.R.R. Tolkien
- **Subtitles**: 18 mil. tokens
- **Europarl**: 14,2 mil. tokens
- 👉 **Searching in translations**: no Polish originals, few Finnish and Czech originals

- pl *Kiedy w genetyce **eksplodowały** nowe idee, Garrod nie żył już od*  
 when in genetics exploded<sub>PF/IPF</sub> new ideas Garrod NEG lived already since  
*dziesięciu lat.*  
 ten years (PF/IPF)
- cs *To už genetika **překypovala** novými myšlenkami a Garrod byl tehdy*  
 it already genetics brimmed<sub>IPF</sub> new ideas<sub>INS</sub> and Garrod was then  
*deset let mrtv.*  
 ten years dead (IPF)
- fi *Tuolloin perinnöllisyystiede **oli uusien haasteiden***  
 at.that.time genetics **be**<sub>PST</sub> new<sub>PL.GEN</sub> challenge<sub>PL.GEN</sub>  
*parissa, ja Garrod oli ollut kuolleena jo vuosikymmenen ajan.*  
 with and Garrod <sub>PST</sub> **be**<sub>PTCP</sub> dead<sub>ESS</sub> already decade<sub>GEN</sub> time<sub>GEN</sub> (IPF)
- en *By then, genetics **was exploding** with new ideas and Garrod had*  
*been dead for a decade.*<sup>1</sup> (fi: copula ⇒ STATIC)

<sup>1</sup>M. Ridley: *Genome: the autobiography of a species in 23 chapters*, translated into Polish by M. Koraszewska, into Czech by I. Müller, into Finnish by S. Sommersalo. ↻ 🔍 🔄

- pl *Pytania **eksplodowały** w głowie Harry'ego jak fajerwerki i nie mógł*  
 questions exploded<sub>PF/IPF</sub> in head Harry<sub>GEN</sub> like fireworks and not could  
*się zdecydować, które zadać najpierw.*  
 REFL decide which ask first (PF/IPF)
- cs *V Harryho hlavě **vybuchlo** otázek jako při ohňostroji, a*  
 in Harry's head explode<sub>PF</sub> questions<sub>GEN</sub> like during firework and  
*nemohl se rozhodnout, kterou má vyslovit nejdřív.*  
 could<sub>NEG</sub> REFL decide which shall express first (PF)
- fi *Harryn päähän **räiskähti** kysymyksiä kuin ilotulitusta eikä*  
 Harry<sub>GEN</sub> head<sub>ILL</sub> explode<sub>MOM.PST</sub> question<sub>PL.PAR</sub> like firework<sub>PAR</sub> and.not  
*hän osannut päättää, mitä kysyisi ensin.*  
 he be.able<sub>PTCP</sub> decide<sub>INF</sub> what<sub>PAR</sub> ask<sub>COND</sub> first (PF)
- en *Questions **exploded** inside Harry's head like fireworks and he couldn't*  
*decide which to ask first.<sup>2</sup>* (CS: GEN, fi: PAR ⇒ IPF?)

<sup>2</sup>J. K. Rowling: *Harry Potter and the Philosopher's Stone*, translated into Polish by A. Polkowski, into Czech by V. Medek, into Finnish by J. Kapari. 

- pl *Pierwotne światło **eksplodowało**, obryzgując* *czasoprzestrzeń*  
 primal light exploded<sub>PF/IPF</sub> splattering<sub>ADV.PPLE</sub> space-time  
*czymś w rodzaju grudek budyniu.*  
 something<sub>INS</sub> in kind<sub>LOC</sub> cluds<sub>GEN</sub> custard<sub>GEN</sub> (PF/IPF)
- cs *Prvotní světlo **vybuchovalo** a rozstříkovalo se v časoprostoru jako*  
 primal light exploded<sub>IPF</sub> and splattered<sub>IPF</sub> REFL in space-time like  
*kapky jogurtu.*  
 drops yoghurt<sub>GEN</sub> (IPF)
- fi *Luomisen valo **räjähiti** singoten avaruusaikaa joka*  
 creation<sub>GEN</sub> light explode<sub>MOM.PST</sub> splatter<sub>INF2.INS</sub> space-time<sub>PAR</sub> every  
*suuntaan kuin herajuustomöykkyjä.*  
 direction<sub>ILL</sub> as whey.cheese.gob<sub>PL.PAR</sub> (PF)
- en *Primal light **exploded**, splattering space-time as with gobbets of*  
*junket.<sup>3</sup>* (simultaneous events ⇒ IPF?)

<sup>3</sup>D. Adams: *The Hitch Hiker's Guide to the Galaxy*, translated into Polish by A. Banaszczak, into Czech by J. Hollanová, into Finnish by P. Markkula



- pl *Eksploduje* ze stłumionym hukiem blisko mojej głowy.  
explodes<sub>PF/IPF</sub> with muffled sound near my head (PF/IPF)
- cs *Exploduje* mi s tlumeným výbuchem kousek nad hlavou.  
explodes<sub>PF/IPF</sub> me<sub>DAT</sub> with muffled sound bit above head (PF/IPF)
- fi *Se räjähtää* vähän minun pääni yläpuolella vaimeasti jysähtäen.  
it explode<sub>MOM.3SG</sub> a.little my head.my above silently thud<sub>INF3.INS</sub> (PF)
- da *Det **eksploderer** et stykke over mit hoved med et dæmpet knald.*<sup>4</sup>  
'It explodes with a muffled sound just above my head.'  
(historic present)

<sup>4</sup>P. Høeg: *Frøken Smillas fornemmelse for sne* (Danish original, translated into English as *Miss Smilla's Feeling for Snow*), translated into Polish by I. Zimnicka, into Czech by R. Novotný, into Finnish by P. Talvio-Jaatinen.

# Outline

- 1 Biaspectual verbs in Slavic
- 2 About InterCorp
  - Basics
  - Content of *InterCorp*
  - Access to InterCorp
- 3 Examining comparative contexts
- 4 Pilot study**
- 5 Discussion
- 6 References

# Pilot study

- 140 parallel contexts: Polish, Czech and Finnish
- 3 Polish biaspectual verbs:
  - *amputować* ‘amputate’ – 15 contexts
  - *autoryzować* ‘authorize’ – 21 contexts
  - *ratyfikować* ‘ratify’ – 51 contexts
- 1 Czech biaspectual verb (not balanced! work in progress. . . ):
  - *obětovat* ‘sacrifice’ – 53 contexts
- In Polish and Czech: annotated for tense, mood, and aspect
- In Finnish: annotated for tense, mood, finiteness, case marking of the object, case marking of another argument, temporal adverbial
- Polish vs. Czech: same or different aspect?
- Slavic vs. Finnish: what the morphosyntactic clues say about aspect

# Annotation issues

## All the three languages

- If possible, aspect determined by interpretation in (monolingual) context, including biaspectuals
- Auxiliaries and modals: aspect determined by the content verb
- Complex predicates: aspect determined by the matrix verb
- Deverbal nominals as equivalents – sometimes with aspect

## Finnish

- Object:
  - PARTitive  $\Rightarrow$  imperfective
  - TOTAL (NOM, ACC or GEN)  $\Rightarrow$  perfective
- Verb:
  - NEGated  $\Rightarrow$  PARTitive object  $\Rightarrow$  imperfective
  - MOMentaneous morpheme  $\Rightarrow$  perfective

## Polish: *amputować* ‘amputate’

- 15 hits, 11 annotated as perfective, 4 as imperfective
- The most common Czech equivalent also biaspectual
- Czech and Polish agree in 11 contexts: 10 perfective and 1 imperfective
- Where Czech and Polish agree on perfective, Finnish differs only in 1 context ⇒ **90% agreement**
- For imperfective, too few examples

Typically, all agree (10 hits):

pl **Amputowano** prawą dłoń. (PF/IPF)

cs *Pravá ruka je **amputovaná** nad zápěstím.* (PF/IPF)

fi *Oikea käsi **amputoitiin** ranteen alapuolelta.*  
(PAST, TOT, PASS ⇒ PF)

en *The right hand has been **amputated** above the wrist.<sup>a</sup>*

---

<sup>a</sup>Subtitles: *Valkyrie*

Disagreements are rare (1 hit):

- pl      *Lekarz mówił, że może będzie trzeba **amputować**.*      (PF/IPF)
- cs      *Doktor říkal, že bude možná muset **amputovat**.*      (PF/IPF)
- fi      *Lääkäri sanoi, että heidän saattaa olla pakko **amputoida** se.*  
(PRES, TOT, INF1 ⇒ PF)
- en      *The doctor said he might have to amputate.<sup>a</sup>*

---

<sup>a</sup>Subtitles: *Son of Rambow*

## Czech: *obětovat* ‘sacrifice’

- 54 hits, 35 annotated as perfective, 17 as imperfective, 2 as ambiguous
- Most frequent Polish equivalents:
  - perfective *poświęcić*
  - imperfective *poświęcać*
- Czech and Polish agree in 38 contexts: 25 perfective and 13 imperfective
- Like in *amputować*, Finnish often agrees with Czech and Polish – in 24 contexts: 16 perfective ⇒ **75% agreement** and 8 imperfective



- cs *Otec a já jsme si na začátku našeho “odchodu do podzemí” pořídili improvizovaný nočník – pro nedostatek nočních váz jsme k tomu účelu **obětovali** zavařovací láhve.* (PF/IPF)
- pl *Tata i ja sprawiliśmy sobie od początku naszego ukrywania się zaimprovizowany nocnik, w tym sensie, że z braku nocnika **poświęciliśmy** na ten cel szklany słoik weka.* (PF)
- fi *Silloin kun tulimme tänne, isä ja minä teimme itse itsellemme yöastian; ja kun meillä ei ollut sopivampaa kapinetta, me **uhrasimme** tähän tarkoitukseen lasitölkin.*  
(PAST, TOT, 1 PL, GOAL ⇒ PF)
- en *On the day of our arrival, Father and I improvised a chamber pot, **sacrificing** a canning jar for this purpose.<sup>5</sup>*

---

<sup>5</sup>Anna Frank

- cs **Obětovala** mu nejlepší léta, vařila chop suey, pomáhala s firmou a určitě se nemohl dočkat, až z ní serve kimono. (PF/IPF)
- pl Ona **ofiarowała** mu swoje najlepsze lata, gotowała sushi, pomagała w interesach. I pewnie kiedyś nie mógł doczekać się, aż zerwie z niej kimono. (PF)
- fi **Antanut** parhaat vuotensa, kokannut, auttanut firmassa – ja palvellut kai vaimonakin. (PERF, TOT, 3SG ⇒ PF)
- en She's **given** him her best years, cooked his chop suey, helped with the business, and I bet at one time he couldn't wait to rip off her kimono.<sup>6</sup>

<sup>6</sup>Subtitles: *Alfie*

- cs *Žádná země **nebude chtít** své muže **obětovat** nadarmo, ani Anglie.* (PF/IPF)
- pl *Żaden kraj nie będzie na darmo **poświęcać** swoich ludzi, a zwłaszcza nie dla interesów innego, również Anglia nie będzie tego robić.* (IPF)
- fi *Mikään maa **ei aio uhrata** miehiään turhan takia eikä ainakaan toisten hyödyksi. Englantikaan ei aio sitä tehdä.*  
(PRES, NEG, 'intend', PAR, INF1, REASON ⇒ IPF)
- en *No country **sacrifices** its men without reason, and certainly not in the interests of another, and England is no exception.<sup>7</sup>*

---

<sup>7</sup>Anna Frank

- cs *Jako by se **obětovaly** pro větší dobro celého těla.* (PF/IPF)
- pl *Jak gdyby **poświęcały** się dla dobra całego organizmu.* (IPF)
- fi *Aivan kuin solut **uhraisivat** itsensä paremman kehon puolesta.*  
(COND, TOT, 3PL, REASON ⇒ PF)
- en *It is as if they **sacrifice** themselves for the greater good of the body.<sup>8</sup>*

---

<sup>8</sup>Matt Ridley *Genome: the autobiography of a species in 23 chapters*, Harper Collins, New York, 1999

## Polish: *ratyfikować* ‘ratify’

- 51 hits, 48 annotated as perfective, 3 as imperfective
- In 44 contexts biaspectuals in both Polish and Czech, including deverbal nouns
- Czech and Polish agree in 48 contexts: 45 perfective and 3 imperfective
- Monoaspectual Czech equivalents rare: in 7 contexts, then Finnish agrees on perfective in 3 contexts
- Finnish often agrees with Czech and Polish – in 44 contexts: 43 perfective and 1 imperfective

- pl *Choć należy przyznać, że **ratyfikowała** tę konwencję z jednym zastrzeżeniem: w przypadku konfliktu pomiędzy postanowieniami tej konwencji a normami prawa islamskiego, ...*  
(PF/IPF)
- cs *Je pravda, že **podepsala** Úmluvu s jednou výhradou: v případě konfliktu mezi články Úmluvy a normami islámského práva ...*  
(PF)
- fi *Se on eittämättä **ratifioinut** yleissopimuksen yhdellä varauksella: jos yleissopimuksen lausekkeet ja islamilaisen lain vaatimukset ovat keskenään ristiriidassa, ...*  
(PERF, TOT, 3PL ⇒ PF)
- en *Admittedly it has ratified the Convention with one reservation: in the event of a conflict between the clauses of the Convention and the standards of Islamic law, ...<sup>9</sup>*

<sup>9</sup>Europarl: Günter Verheugen

- pl Senat go **ratyfikuje**. (PF/IPF)
- cs Ujistili mě, že ji senát určitě **potvrdí**. (PF)
- fi Se **ratifoidaan** varmasti senaatissa.  
 (PRES, TOT, PASS, LOCATION ⇒ PF)
- en I have assurances it will be **ratified** by the senate.<sup>10</sup>

---

<sup>10</sup>Subtitles: *Star Wars: Episode I – The Phantom Menace* ▶ ◀ ☰ ▶ ◀ ☰ ▶ ◀ ☰ ▶ ☰ ◀ 🔍 ◀

- pl *Po co **ratyfikować** Traktat, który jeszcze bardziej chroni interesy rynku kosztem interesów pracowników?* (PF/IPF)
- cs *Proč **schvalujeme** smlouvu, která ještě více posiluje zájmy trhu na úkor zájmů zaměstnanců?* (IPF)
- fi *Miksi **hyväksytään** sopimus, joka entisestään vahvistaa markkinoiden etuja työntekijöiden kustannuksella?*  
(PRES, TOT, PASS ⇒ PF)
- en *Why **approve** a Treaty which further strengthens the interests of the market against those of working people?<sup>11</sup>*

---

<sup>11</sup>Europarl



- pl *Mam jedną prośbę: niezależnie od tego, jak będzie pani negocjować i co się wydarzy, proszę na czas informować Parlament i zapewnić, byśmy, niezależnie od tego, czy **ratyfikujemy** traktat lizboński, czy też nie, rozważyli ponownie ewentualną możliwość stworzenia czegoś na kształt “handlowego traktatu lizbońskiego”.* (PF/IPF)
- cs *Paní komisařko, mám jednu prosbu: ať budete jednat jakkoli a ať se stane cokoli, informujte prosím Parlament včas a buďte ujištěna, že bez ohledu na to, zda **budeme mít** Lisabonskou smlouvu či ne, budeme se znovu zabývat tím, zda by bylo možné předem vytvořit určitou “kvazilisabonskou smlouvu” pro oblast obchodu.* (IPF)
- fi *Arvoisa komission jäsen, minulla on yksi pyyntö: riippumatta siitä, miten neuvottelette ja mitä tapahtuu, olkaa niin ystävällinen ja tiedottakaa parlamentille ajoissa asioista ja varmistakaa se, että huolimatta siitä, **tuleeko** Lissabonin sopimus **voimaan** vai ei, tarkastelemme mahdollisuutta luoda kaupankäyntiä varten jonkinlainen Lissabonin sopimusta läheisesti muistuttava väline ennakoivasti.* (PRES, 0, 3SG, light verb ⇒ PF)
- en *Commissioner, I do have one request: however you negotiate and whatever happens, please inform Parliament in good time and make sure, too, that whether we **get** the Lisbon Treaty or not, we take another look at whether we can perhaps create some kind of ‘quasi Lisbon Treaty’ in advance, relating to the area of trade.<sup>12</sup>*

<sup>12</sup>Europarl: Erika Mann

## Polish: *autoryzować* 'authorize'

- 21 hits, 10 annotated as perfective, 11 as imperfective
- Czech biaspectual equivalent just once
- Czech and Polish agree in 12 contexts: 6 perfective and 6 imperfective
- 2 contexts without Czech equivalent
- 19 contexts with unambiguous Czech equivalents: 9 perfective, 10 imperfective
- Finnish agrees with Czech and Polish in 10 contexts: 7 perfective and 3 imperfective

For 9 Czech PFs, Czech and Finnish agree in 7 contexts (1/2):

- pl *Zdajecie sobie sprawę chłopcy, że depesze **autoryzuje** kmdr Owynn.* (PF/IPF)
- cs *Jistě víte, pánové, že nesmíte odesílat žádné zprávy, dokud je **neschválí** velitel Owynn.* (PF)
- fi *Ymmärrätte varmasti, ettette saa lähettää juttuja - ennen kuin komentaja Owynn  **antaa luvan.*** (PRES, TOT, 3SG, TIME ⇒ PF)
- en *I'm sure you boys realise that you're not to file any stories until Cdr. Owynn **gives you the word.**<sup>a</sup>*

<sup>a</sup>Subtitles: *In Harm's Way*

For 9 Czech PFs, Czech and Finnish agree in 7 contexts (2/2):

pl *Kto to **autoryzował**?* (PF/IPF)

cs *Kdo to **schválil**?* (PF)

fi *Kuka **antoi luvan** tähän?* (PAST, TOT, 3SG ⇒ PF)

en *Who **authorised** this?<sup>a</sup>*

<sup>a</sup>Subtitles: *Clear and Present Danger*

For 11 Czech IPFs, Czech and Finnish agree only in 3 contexts:

pl	Nie <b>autoryzowaliśmy</b> żadnych wy...	(PF/IPF)
cs	Nebudeme <b>schvalovat</b> žádn...	(IPF)
fi	Emme ole <b>valtuuttaneet</b> ...	(PAST, TOT, 3SG ⇒ IPF)
en	We did not <b>authorize</b> ... <sup>a</sup>	

<sup>a</sup>Subtitles: *The Librarian: The Curse of the Judas Chalice*

For 11 Czech IPFs, Czech and Finnish disagree in **8** contexts:

pl ***Autoryzuje*** *akcję awaryjną.* (PF/IPF)

cs ***Povolují*** *akci Blackbird 1195.* (IPF)

fi ***Vahvistetaan*** *hätätoimi Blackbird 1195.*  
(PRES, TOT, PASS ⇒ PF)

en ***Authenticate*** *emergency action. Blackbird 1195 ...<sup>a</sup>*

---

<sup>a</sup>Subtitles: *Transformers*

For Czech IPFs, Finnish is ambiguous in 2 contexts:

- pl *To całkowite darowanie winy, **autoryzuje** natychmiastowe zwolnienie.* (PF/IPF)
- cs *Toto je plná milost. **Opravňuje** vás k okamžitému propuštění.* (IPF)
- fi *Tämä on täydellinen armahdus. Vapaudutte välittömästi.* (0 ⇒ PF/IPF)
- en *Mr. Barnes . . . this is a full pardon. It **authorises** your immediate release.<sup>a</sup>*

---

<sup>a</sup>Subtitles: *Virtuosity*

# Summary of the pilot study

	<i>obětovat</i>			<i>autoryzować</i>			<i>ratyfikować</i>			<i>amputować</i>			TOTAL		
	pl	cs	fi	pl	cs	fi	pl	cs	fi	pl	cs	fi	pl	cs	fi
PF	27	36	37	10	9	10	46	45	42	11	13	13	94	103	102
ImPF	21	17	9	11	11	11	2	6	9	4	1	2	38	35	31
PF/ImPF	5	0	7	0	1	0	3	0	0	0	1	1	8	2	8
<b>TOTAL</b>	<b>53</b>			<b>21</b>			<b>51</b>			<b>15</b>			<b>140</b>		
agree	24			10			44			10			88		
	38			12			48			11			109		
PF	51%	68%	70%	48%	43%	48%	90%	88%	82%	73%	87%	87%	67%	74%	73%
ImPF	40%	32%	17%	52%	52%	52%	4%	12%	18%	27%	7%	13%	27%	25%	22%
PF/ImPF	9%	0%	13%	0%	5%	0%	6%	0%	0%	0%	7%	7%	6%	1%	6%
<b>TOTAL</b>	<b>100%</b>			<b>100%</b>			<b>100%</b>			<b>100%</b>			<b>100%</b>		
agree	45%			48%			86%			67%			63%		
	72%			57%			94%			73%			78%		



# Outline

- 1 Biaspectual verbs in Slavic
- 2 About InterCorp
  - Basics
  - Content of *InterCorp*
  - Access to InterCorp
- 3 Examining comparative contexts
- 4 Pilot study
- 5 Discussion**
- 6 References

# Discussion

- The ideal situation: biaspectual in L1, monoaspectual in L2
- But sometimes the aspectual interpretation of the L1 biaspectual differs from its monoaspectual equivalent, then L3 may help
- Comparison with a language without the grammatical category of aspect is important for both sides
- Perspectives: add Czech biaspectuals, extend to more verbs

# Paldies! 'Thank. 1 PL you!'

InterCorp v12 - Czech ✓	InterCorp v12 - Polish ✓	InterCorp v12 - Finnish ✓
Děkujeme vám .	Dziękujemy wam .	Kiitos .
" Děkujeme vám ! " kvitoval to Thorin .	- Dzieki ci , Gandalfie - rzekł Thorin .	» Kiitos vaan » , sanoi Thorin .
„ Děkujeme .	– Dziękuję .	- Kiitos .
" Děkujeme vám , Roaku , synu Krákův .	- Dzieki ci , Roaku , synu Karka !	» Kiitos vain , Roac Carcin poika .
Děkujeme vám . "	Dziękujemy za uwagę .	Kiitos . "
" Děkujeme , biskupe , " ozval se sekretář .	- Dziękujemy , księżę biskupie - powiedział sekretarz watykański .	» Kiitos , piispa » , sanoi secretarius .
" Děkujeme , ano , " odpověděl Langdon .	- Tak , bardzo by śmy prosili - odparł Langdon .	» Kiitos kyllä » , vastasi Langdon .
" Děkujeme , Pottere , tentokrát jsi nám posloužil ! "	" Brawo , Potter , jesteśmy ci wdzięczni ! "	Luihuit puolestaan taputtivat käsiään , kun hän kulki heidän ohitsean , vihelsivät ja hurrasivat : " Kiitos Potter , me ollaan sinulle velkaa ! "
" Děkujeme mockrát ! "	- Dziękujemy ślicznie !	» Kiitos , kiitos ! »
" Děkujeme ti , Slartibartfaste , " řekl Béda pevně .	- Dziękujemy ci , Slartibartfost - rzekł stanowczo Benjy .	" Kiitos , Romppanen " , Benjy vingahti tiukkaan sävyyn .
Později , když všichni řekli Kryšťůfkovi Robinovi " Sbohem " a " Děkujeme " , Pú s Prasátkem se ubírali vlahým večerem společně domů .	Jakis czas później , gdy wszyscy już powiedzieli " Do widzenia " i " Dziękuję " Krzysztofowi Robinowi , Phi z Prosiaczkciem udali się razem do domu i szli czas jakis zamyśleni , milczący , w złotym blasku wieczoru .	Myöhemmin kun kaikki olivat sanoneet Risto Reippaalle » Näkemiin » ja » Kiitos » , Puh ja Nasu astelivat ajatuksissaan illan kultaisessa kajossa kotiinpäin .
„ Děkujeme .	" Zaczynam mieć ich już dość " .	" Kiitos , mutta me alamme saada tarpeeksemme . "

# Outline

- 1 Biaspectual verbs in Slavic
- 2 About InterCorp
  - Basics
  - Content of *InterCorp*
  - Access to InterCorp
- 3 Examining comparative contexts
- 4 Pilot study
- 5 Discussion
- 6 **References**



Charciarek, A. (2019).

Využití paralelního korpusu v translatoologii (na základě česko-polského intercorpu).

*Bohemistyka*, 19(2), 194–216.



Chromý, J. (2014).

Impact of tense on the interpretation of bi-aspectual verbs in Czech.

*Studie z aplikované lingvistiky / Studies in Applied Linguistics*, 5(2), 87–97.



Cockiewicz, W. (2007).

Na peryferiach aspektu.

*LingVaria*, 2(4), 9–25.



Jirásek, K. (2011).

Využití paralelního korpusu InterCorp k získávání ekvivalentů pro chorvatsko-český slovník.

In F. Čermák, editor, *Korpusová lingvistika Praha 2011: 1 – InterCorp*, pages 45–55. Nakladatelství Lidové noviny, Praha.



Jurkiewicz-Rohrbacher, E. (2019).

*Polish verbal aspect and its Finnish statistical correlates in the light of a parallel corpus.*




Ph.D. thesis, University of Helsinki, Finland.



Kaczmarska, E. (2019).

*Metody ustalania ekwiwalentów czasowników wyrażających stany emocjonalne w przekładzie czesko-polskim na materiale z korpusu równoległego InterCorp.*

Wydział Polonistyki Uniwersytetu Warszawskiego.

-  Káňa, T. (2016).  
*Česká a německá deminutiva na pozadí korpusových dat.*  
Masarykov univerzita, Brno.
-  Malá, M. & Brůhová, G. (2018).  
English presentative semantic patterns as seen through a parallel translation corpus.  
*Languages in Contrast*, **19**(2), 232–255.
-  Perlin, J. (2010).  
Ile jest we współczesnej polszczyźnie czasowników dwuaspektowych?  
*Linguistica Copernicana*, **1**(3), 165–171.



Piperski, A. (2018).

The grammatical profiles of Russian biaspectual verbs.

In M. Kopotev, O. Lyashevskaya, and A. Mustajoki, editors, *Quantitative Approaches to the Russian Language*, chapter 6, page 115–136. Routledge, London and New York.



von Waldenfels, R. (2012).

Aspect in the imperative across Slavic – a corpus driven study.

In A. Grønn and A. Pazelskaya, editors, *The Russian Verb*, volume 4(1) of *Oslo Studies in Language*, page 141–154. University of Oslo.



Čermák, P., Kratochvílová, D., Nádvorníková, O., & Štichauer, P., editors (2020).

*Complex Words, Causatives, Verbal Periphrases and the Gerund: Romance Languages versus Czech (A Parallel Corpus-Based Study)*.

Karolinum, Charles University Press.





Čermáková, A. & Fárová, L. (2010).

Keywords in Harry Potter and their Czech and Finnish translation equivalents.

In F. Čermák, A. Klégr, and P. Corness, editors, *InterCorp: Exploring a Multilingual Corpus*, volume 13 of *Studie z korpusové lingvistiky*, pages 177–188. Nakladatelství Lidové noviny, Praha.