

Zusammengesetzte Indefinitpronomen in den slavischen Sprachen: Korpusbasierte Untersuchung der Grammatikalisierung und Distribution



Universität Regensburg

Björn Hansen / Roman Fisun
Institut für Slavistik
Fakultät Sprach-, Kultur- und Literaturwissenschaften

DFG Deutsche
Forschungsgemeinschaft

Struktur der Präsentation

0. Vorstellung des Projekts
1. Zusammengesetzte Indefinitpronomen
2. Korpusbasierte Inventarisierung
3. Grammatikalisierung
4. Distribution: Lizenzierung

Kontext

- DFG-Projekt (Leitung: B. Hansen, Universität Regensburg): „Zusammengesetzte Indefinitpronomen in den slawischen Sprachen. Semantische Karte der Indefinitheit der zweiten Generation“ (HA 2659/10-1)
- Dissertationen:
 - Martina Rybová „Die Beziehung zwischen Evaluativität und Indefinitheit am Beispiel der zusammengesetzten Indefinitpronomina im Tschechischen“ (2023)
 - Roman Fisun „Zusammengesetzte Indefinitpronomen in den slavischen Sprachen“ (2024)
 - MA: Irina Maykova „Zur Funktionalstilistik ausgewählter Indefinitpronomen im Russischen“(2023)
- Sprachen:
RU, UA, PL, CZ, HR

1. Zusammengesetzte Indefinitpronomen

Struktur der Indefinitpronomen in den slavischen Sprachen:

- nicht-reihenbildende IPs: HR *drugi, ostali, ini*; RU *tot ili inoj*; UA *odyn, pewnyj*...

- reihenbildende IPs:

präponiert

HR [i]_{INDEF} *tko*_{wer}
[i]_{INDEF} *šta*_{was}
[i]_{INDEF} *kakav*_{welcher}

kodifiziert

'irgendjemand'
'irgendetwas'
'irgendwelcher'

vs nicht-kodifiziert

HR [**malo**]_{INDEF} *tko*_{wer}
[**malo**]_{INDEF} *šta*_{was}
[**malo**]_{INDEF} *kakav*_{welcher}

'kaum jemand'
'kaum was'
'kaum welcher'

postponiert

UA *chto*_{wer} [s]_{INDEF}
*ščo*_{was} [s]_{INDEF}
*jakyj*_{welcher} [s]_{INDEF}

kodifiziert=„traditionell“

'jemand'
'etwas'
'irgendein'

vs nicht-kodifiziert

UA *chto*_{wer} [**zavhodno**]_{INDEF} 'jeder'
*ščo*_{was} [**zavhodno**]_{INDEF} 'irgendetwas'
*jakyj*_{welcher} [**zavhodno**]_{INDEF} 'jeder' (free-choice)

Definition

- Indefinite Funktion
- binäre Struktur: Modifikator + **k*-Pronominalwurzel
- ZIPs können nicht-pronominale Bedeutungen, einschließlich evaluativer, ausdrücken:
 - ❖ ‘negative Bewertung’
 - (1) RU *Tam **komu popalo** vydajut pasporta!* [Bulgakov: Master i Margarita]
 - DE Dieses Revier kenne ich, die geben ja jedem Hergelaufenen einen Ausweis. (Parasol-Corpus)
 - ❖ Quantitative Komponente
 - (2) HR *Nažalost, **rijetko tko** uspijeva putem hrane zadovoljiti dnevne potrebe ovih nutrijensa <...>. (hrWaC 2.2)*
 - ‘Leider schafft es kaum jemand, die erforderliche Tagesmenge dieser Nährstoffe über die Nahrung aufzunehmen.’

2. Inventarisierung

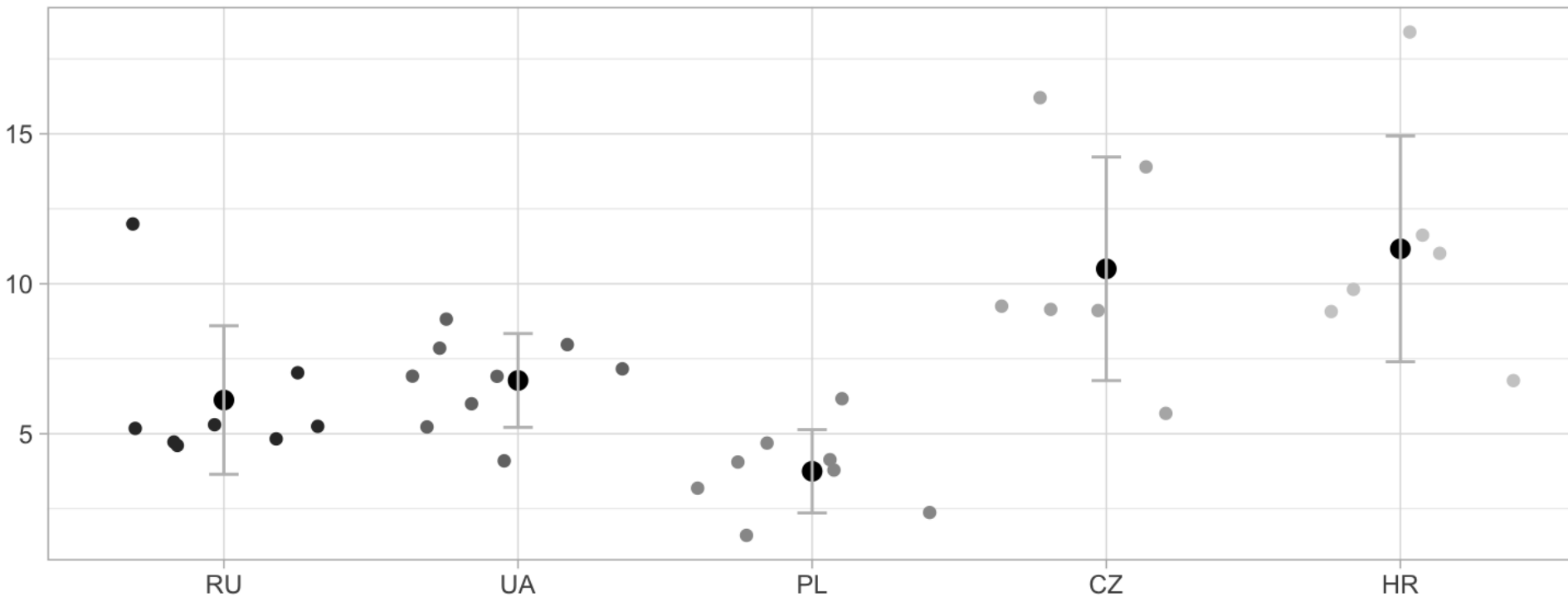
Wie viele Indefinitpronomen haben die Sprachen?

Probleme:

- Die Pronominalreihen können unvollständig sein:
RU **neka*, **nepočem*; UA **denaviščo*; PL **niekto*; CZ
**ledakolik*; HR **zaštogod*
- Verschiedene Grammatiken bieten unterschiedliche Listen von Indefinitmarkern an
- Es gibt keine systematisch erfassten Listen der nicht-kodifizierten Indefinitmarker
- Ist es möglich, eindeutig zu bestimmen, welche **Reihen** „traditionell“ (=kodifiziert) sind?
- Wie kann man die nicht-kodifizierten Reihen aus Korpora extrahieren?

Universität Regensburg

Reihenbildende IPs laut Grammatiken



\bar{x}	6,1	6,8	3,8	10,5	11,2
sd	2,5	1,6	1,4	3,7	3,8

„Traditionelle“ IPs

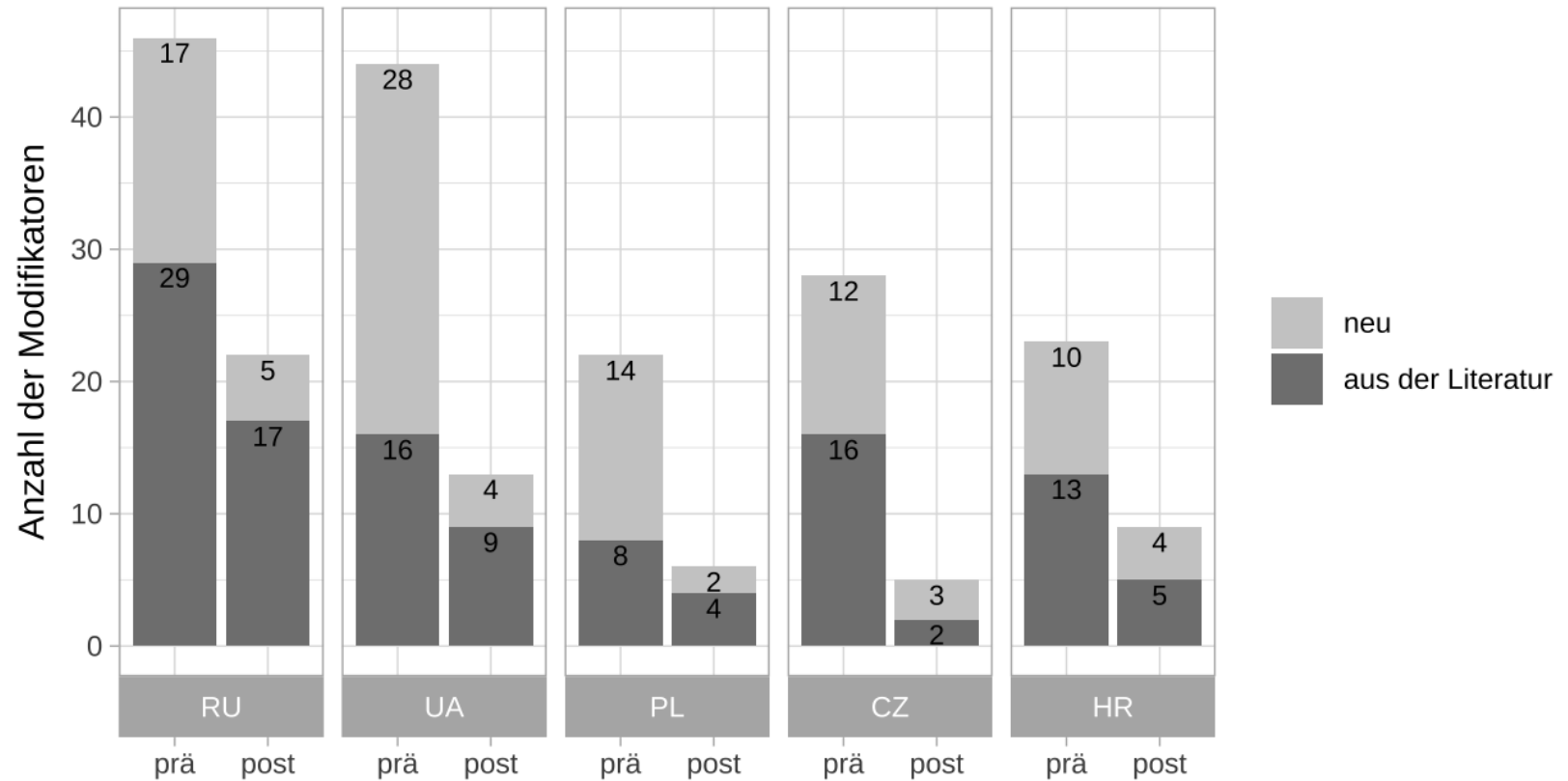
Sprache	RU	UA	PL	CZ	HR
PRÄ	ne koe (koj)	aby bud' de kazna chto zna	/	leda- ně- kde- málo- všeli- lec-	bito gdje- i- koje- ne- ma sva-/sav-/sve-
POST	-to -libo -nibud'	-s' -nebud'	-kolwiek -ś	-koli(v) -si	-god
Σ	5	7	2	8	8

Was ist in Korpora zu sehen?

- Verwendung von Webkorpora und (No)Sketch Engine: niedrige Frequenz, Umgangssprache, Analysetools
- Methodik: Berücksichtigung von Orthografie-Variationen und Kollokationsanalyse zur Extraktion neuer Reihen
- Varianz:
RU NIBUD': *нибудь, нить, нибуть, нидь, нть, ни-будь, небудь, небуть*
HR BILO: *bilotko* (361 Einträge), *bilo-tko* (4 Einträge) und *bilo tko* (36.585 Einträge)

Sprache	Korpus	Größe, Tokens	Software	Lemmatisierung
Russisch	ruTentTen11	18.280.486.876	SketchEngine	lemmatisiert, PoS-getaggt
Ukrainisch	ukTentTen14	2.734.851.744	SketchEngine	nicht lemmatisiert, keine PoS-Tags
Ukrainisch	ukTenTen20	3.282.586.754	SketchEngine	lemmatisiert, PoS-getaggt
Ukrainisch	GRAC	437.358.921	NoSketchEngine	lemmatisiert, PoS-getaggt
Kroatisch	hrWac 2.2	1.405.794.913	SketchEngine / NoSketchEngine	lemmatisiert, PoS-getaggt
Kroatisch	HNK	2.559.160	NoSketchEngine	lemmatisiert, PoS-getaggt

Unsere Ergebnisse: Quantitativ



HR Modifikatoren

präponierte Modifikatoren		postponierte Modifikatoren	
aus der kumulativen Liste	neue	aus der kumulativen Liste	neue
bilo/bilo;/bilo: 'sein _{ptcp} '	malo/malo: 'wenig'	:god;/god/:god 'passend'	:što 'was'
Bog zna/bog-zna;/ bogzna: 'Gott weiß'	mного/mного: 'viel'	bilo/-bilo/:bilo 'sein _{PTCP} '	treba 'notwendig.sein _{3SG} '
bud;/bud:/bud/budi;/ budi:/budi 'sein _{imp} '	ne-znam;/neznam/neznam: 'ich.weiß.nicht'	hoćeš/hoćeš/hoces 'du.willst'	valja 'notwendig.sein _{3SG} '
gdje: 'wo'	neznamo/ne znamo/ne znaš/neznaš/ne	mu drago/:mudrago 'angenehm.für.ihn'	:već 'schon'
i: 'und'	znaju/nezaju 'wir/du/sie.wissen.nicht'	te volja/ti volja 'du.magst'	
koje: 'welch'	neznano 'unbekannt'		
ma/ma: 'aber'	pogdje: 'auf.wo'		
makar 'obwohl'	rijetko 'selten'		
ne:/nie:/nje 'nicht'	tko zna/tko-zna; 'wer weiß'		
po: 'auf'	vrag zna/vrag će znati 'Teufel weiß /wird wissen'		
pone: 'auf.nicht'	znaš/znaš;/znamo/znamo; 'du/wir.wissen'		
što: 'was'			
sva:/sve:/sav: 'alles'			

Ergebnisse: qualitativ

- Modifikatoren, die auf quantitative Ausdrücke zurückgehen:
 - (3) HR *Nažalost, rijetko tko uspijeva putem hrane zadovoljiti dnevne potrebe ovih nutrijensa <...>. (hrWaC 2.2)*
 ‘Leider schafft es kaum jemand, die erforderliche Tagesmenge dieser Nährstoffe über die Nahrung aufzunehmen.’
- Modifikatoren, die in der Literatur nicht erwähnt wurden
 - (4) RU *Vychodit esli bud' kto pretenduet na to čto on bog - on dolžen razvivat'sja, no tut delema. (ruTenTen11)*
 ‘Es stellt sich heraus, dass wenn jemand behauptet, er sei Gott, er sich weiterentwickeln muss, aber hier liegt das Dilemm’
- Dialektale Varianten:
 - (5) UA – *Čy jila ty daščo [=deščo] dnes' ? (Vyparenjy pip, 1950–1990)*
 ‘Hast du heute etwas gegessen? – Nein, ich habe nichts gegessen.’
- Modifikatoren mit Vulgarismen und Euphemismen:
 - (6) UA *Spočatku doskažy istoriju, – prohnusavyla oficianka, kons'jeržka čy chrin joho zna čto. (O. Ul'janenko. Sofija, 2008)*
 ‘Erzähl zuerst die Geschichte zu Ende, – murmelte die Kellnerin, die Concierge oder wer auch immer sie war.’

Modifikatorenquellen in verschiedenen Sprachen

	RU	UA	PL	CZ	HR
‘wo’		de:/de/de;/da:		kde:/kde;	gdje:
‘auf.wo’					pogdje:
‘welch’	koe;/koe:/koj:/koe/koj	koje;/koje:/koj/koj;/koj:			koje:
‘?’				:koli(v)/:kol/:kolik/koliv	
‘?’			:kolwiek/kolwiek		
‘?’			:śkolwiek		
‘was’					što:
‘was’					:što

Modifikatoren, die auf ***k-Wurzel** zurückgehen, im Vergleich

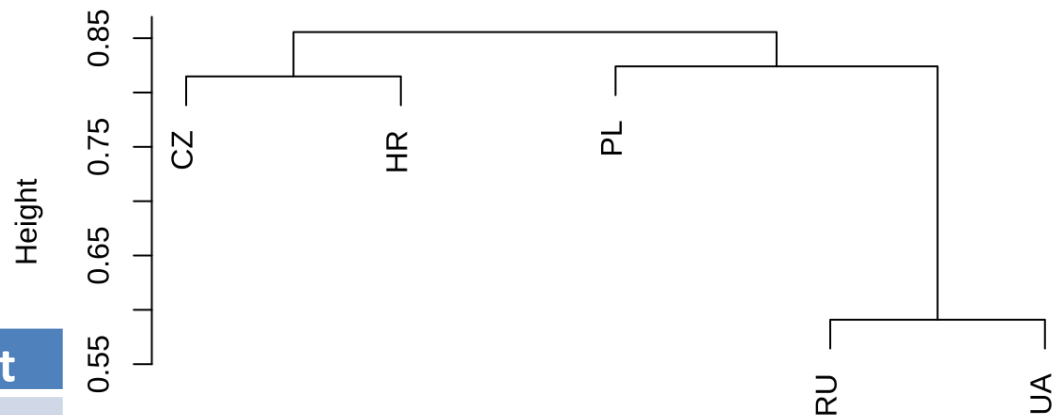
Lexikalisch-etymologische Ähnlichkeit der slavischen Sprachen nach Indefinitmarkern

$$\text{Jaccard}(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

$|A \cap B|$: Anzahl der gemeinsamen Elemente in Sprachen A und B

$|A \cup B|$: Anzahl der Invarianten (\approx mögliche Quellkonstruktionen)

Cluster Dendrogram



Sprachen	Jaccard-Koeffizient
RU und UA	0,4090909
UA und HR	0,1298701
RU und HR	0,125

distances
hclust (*, "average")

3. Grammatikalisierung

Grammatikalisierung

- Definition: „the increase of the range of a morpheme advancing from a lexical to a grammatical or from a less grammatical to a more grammatical status, e.g. from a derivative formant to an inflectional one” (Kuryłowicz 1965: 69)
- „primäre“ Grammatikalisierung
HR: Vollverb *htjeti wollen* > Hilfsverb des Futurs mit der Phonologischen Erosion *Pjevat ću*
- „sekundäre“ Grammatikalisierung
Klitisches Reflexiv *sę* im Altrussischen > RU Verbalpostfix *-sja*

Zielsetzung

- Hypothese prüfen: Überprüfung der Annahme, dass nicht-kodifizierte ZIPs einen (viel) niedrigeren Grammatikalisierungsgrad als kodifizierte aufweisen
- Fundierte Kodifizierungsvorschläge entwickeln
- Tatsächlichen Grad der nichtkodifizierten ZIPs messen

Grammatikalisierungsparameter von Lehmann (1982) und ihre Operationalisierungen für ZIPs

Lehmans Parameter (2015)	Praktische Operationalisierung
1. Integrität	<ul style="list-style-type: none"> a. Verlust an semantischer Integrität b. Verlust an morphosyntaktischer Integrität c. Phonologische Erosion
2. Paradigmatizität	<ul style="list-style-type: none"> a. Indefinitheit b. Präsenz der ontologischen Kategorien
3. Wählbarkeit	<ul style="list-style-type: none"> a. Unikalität des Modifikators b. Zunahme der Frequenz
4. Skopus	<ul style="list-style-type: none"> a. Syntaktische Kondensierung b. Koordinative Konstruktionen, c. Verletzung der Dependenzbeziehungen zwischen Modifikator und *k-Wurzel
5. Fügungseuge	<ul style="list-style-type: none"> a. Zusammenschreibung b. Abwesenheit der starken Autonomie c. Abwesenheit der Abtrennbarkeit d. Eingeschränkte Distribution der Modifikatoren
6. Stellungsfreiheit	Fixe Position des Modifikators in der Verbindung

Methode und Analyse

1. Kriterien 14 Grammatikalisierungskriterien (nach 6 Parametern)
2. Methode **Mixed-Methods:** Intuitiv & korpusbasiert, qualitativ & quantitativ
3. Annotation Binäre Skala zur Vereinfachung (+ / – / 0)
4. Ergebnis **Gesamtscore:** einfacher Durchschnitt der 6 Parameter
5. Grad 0 (keine Grammatikalisierung) bis 1 (volle Grammatikalisierung).

1. ‚Integrität‘ des Sprachzeichens

„sein konstitutioneller Umfang, die Komplexität seines Significans und seines Significatums“ (Lehmann o.D.)

nimmt bei der G. ab

Operationalisierungen:

a) Phonologische Erosion

UA *bozna* < *boh zna(je)*, *kazna* < *kat zna(je)*, *čortzna* < *čort zna(je)*, *čortbat'kazna* < *čorty bat'ka zna*

b) Verlust an semantischer Integrität:

- freie Verbindung:

(7) RU *On vzjal jabloko, kotoroe **popalo** emu v ruku.*
'Er nahm den Apfel, der ihm in die Hand geriet.'

- Wörterbuchdefinition:

(8) ,Случайно оказаться где-н., включиться куда-н., войти куда-н.' [Zufällig an eine bestimmte Stelle gelangen, sich anschließen, ohne Absicht irgendwo eintreten] (Ožegov o.D.)

- ZIP:

(9) RU *Sjuda možet zajti **kto popalo**.*
'Jedermann kann hierher kommen.'

2. Paradigmatizität

„Grad, zu dem [das Sprachzeichen] in ein Paradigma eingegliedert ist, und die formale und funktionelle Homogenität und Geschlossenheit dieses Paradigmas“ (Lehmann o. D.).

nimmt bei der G. zu

Koordinative Konstruktionen

- möglich

(10) RU *Odnako èto malo [kogo]*_k-Wurzel *i* und [kogda]*_k-Wurzel ostanavlivalo. (ruTenTen11)*

‘Das hat aber kaum jemanden bisher aufgehalten..’

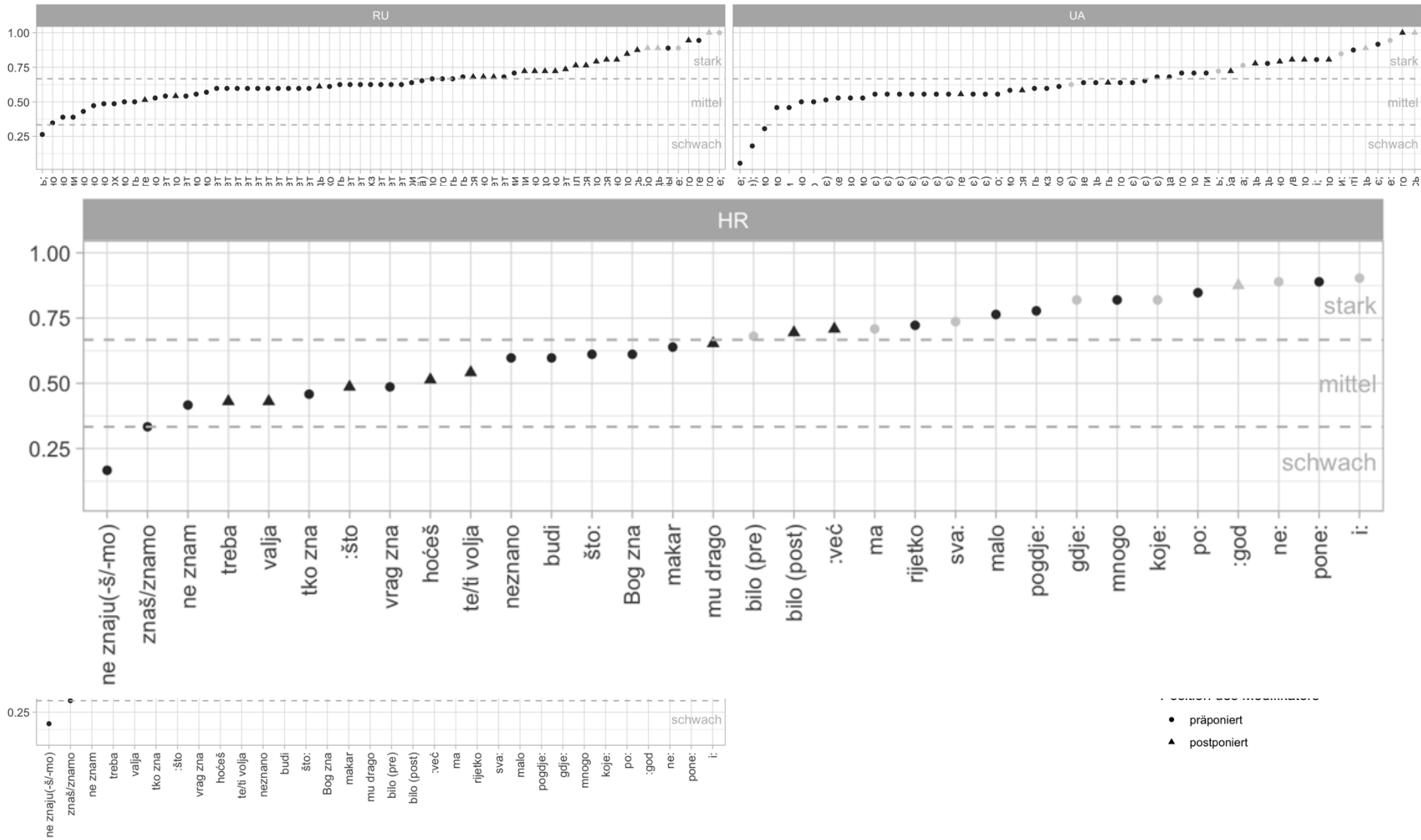
(11) HR *Bajdo nije zaslužio da se, zbog tko zna [kojih]*_k-Wurzel *i* und [čijih]*_k-Wurzel interesa, njegova nogometna genijalnost baca u drugi, treći plan. (hrWaC 2.2)*

‘Bajdo hat es nicht verdient, dass sein fußballerisches Genie wegen der Interessen anderer in den Schatten gestellt wird.’

- ausgeschlossen

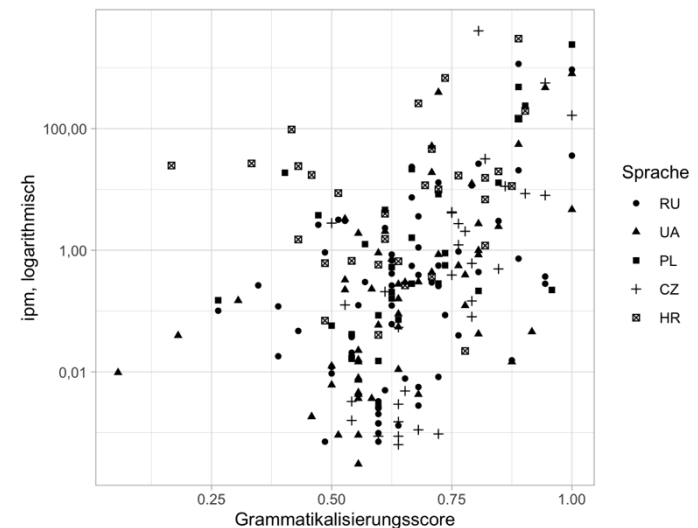
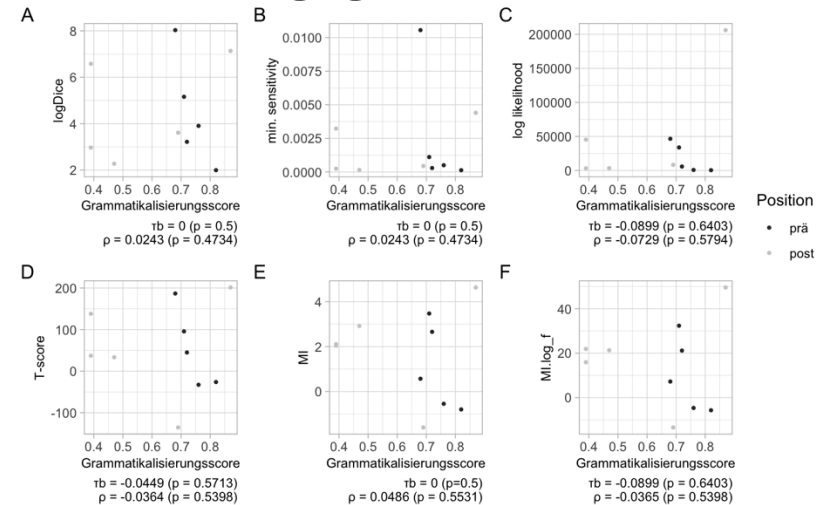
(12) HR *Pa cijelo društvo je ovisno o ne[kome]*_k-Wurzel ili oder o ne[čemu]*_k-Wurzel / *ne[kome]*_k-Wurzel ili oder [čemu]*_k-Wurzel . (hrWaC 2.2)*

‘Die gesamte Gesellschaft hängt von jemandem und etwas ab.’



Korrelation zwischen dem Grammatikalisierungsgrad und anderen Parametern

- keine Korrelation mit der Kollokationsstärke der Verbindung ‚Modifikator+*k-Wurzeln‘
- schwache, statistisch signifikante positive Korrelation mit der Frequenz



Fazit zur Grammatikalisierung

Beitrag zur Theorie:

- Neues Verfahren zur Grammatikalisierungsmessung entwickelt
- Ergebnis: Grammatikalisierungsgrad korreliert mit der **Frequenz** und der **Quelle**, nicht mit der **Kollokationsstärke** „Modifikator+*k-Wurzel“

Beitrag zur Slavistik:

- Grammatikalisierungsgrade aller ZIPs bestimmt
- Bestätigt: Kodifizierte Formen sind stärker grammatikalisiert
- Neu: Stark grammatikalisierte „neue“ Formen als Kandidaten für Kodifizierung identifiziert

4. Distribution: Lizenzierung

Semantische Klassifikation der IPs

1. Referierende **unbestimmte** IPs (Specific Unknown)

- Existenz wird angenommen, die Identität ist für Kommunikanten (Sprecher und Hörer) unbekannt

Angenommene Distribution:

- Frei verwendbar, typisch für episodische Kontexte

(14) UA *Ščos' (*ščo-nebud') rizko zašurchotilo pravoruč[...]* (ukTenTen20)

'Etwas raschelte plötzlich rechts von mir....'

Semantische Klassifikation der IPs

2. Nicht-spezifische IPs:

- Referent ist nicht **individualisiert** und wird aus seiner Klasse nicht gewählt
- Benötigen einen speziellen Kontext = **Lizenzierung**
 - (15) UA *Jakščo vy po **jakych-nejbud'** pryčynach vybraly druhyj variant, to vam varto čytaty dali.* (ukTenTen20)
'Wenn Sie sich aus irgendeinem Grund für die zweite Option entschieden haben, sollten Sie weiterlesen.'
- Lizenzierender Kontext in (15): Konditionalsatz (NSI). Auch Negation (\sim NPIs), modale Kontexte (\sim FCIs) usw.
- FCIs (free-choice indefinites): eine beliebige Auswahl aus einer offenen Menge aus ('jeder Mögliche')
 - (16) UA *Viz'my **bud'-jake** jabluko.*
'Nimm irgendein Apfel.'
- RCIs (random choice indefinites): Indifferente Wahl aus einer Menge ohne Präferenz, oft abwertend ('der Erstbeste', 'egal wer')
 - (17) RU *Sjuda možet zajti **kto** popalo.*
'Hier kann jeder hereinkommen.'

Lizenzierung

- Traditionelle Auffassung: qualitative Lizenzierung
„Elemente der Klasse X sind in den Kontexten A, B, C zulässig, aber in anderen ausgeschlossen“
- Padučeva (2015, 2018): Kontexte der aufgehobenen Assertion
- Definition: Kontexte, in denen der Sprecher keine Verantwortung für die Wahrheit der Äußerung übernimmt
- 25 Typen von KaA: Negationen, Fragen, modale Konstruktionen, Vergleichsstandards u. a.

Kontexte der aufgehobenen Assertion (Padučeva 2015, 2018)

Universität Regensburg

Lizenzierung verschiedener Arten unbestimmter Pronomen nach Padučeva

	referentiell	NSI	NPI	FCI
0.1. assertorischer Kontext	+	-	-	
1.1. direkte Negation		-		+
1.2.1. Negation des übergeordneten Infinitivs		-		
1.2.2. Negation des übergeordneten Prädikats mit Konjunktion		-		
1.5. Präposition mit impliziter Negation		-	+	
4.2. Konstruktion mit Adverbialpartizip				+
9.1. Wunsch		+	-	+
9.2. Bitte, Vorschlag, Imperativ		+	-	+
10. Unsicherheit: Annahme, Irrealität, Meinung		+	-	+
11. Vergleichsstandard				

Auf russischem Material unterscheidet Padučeva (2015, 2018) vier Klassen von IPs:

- referentielle: *koe-kto, kto-to*
- nicht-referentielle (NSIs): *kto-ninud'*
- negative Polaritätspronomen (NPIs): *kto-libo*
- freier Wahl (FCIs): *kto ugodno, ljuboj*

Problem:

- Forschungslücke: keine realen Daten, quantitative Korpusanalysen
- Bisher: Korpusbeispiele dienen nur der **Illustration** von Theorien

Forschungsfragen:

- Ist Lizenzierung im Korpus nachweisbar?
- Welche Muster zeigen sich dabei?
- Lassen sich daraus Rückschlüsse auf Pronomentypen ziehen?

Reihen der IPs

- 51 Reihe kodifizierter (fett gedruckt) und nicht kodifizierter (= „neuer zusammengesetzten“) IPs

	RU	UA	PL	CZ	HR
präponiert	aby Bog vest´ čert-te malo koe- ne pojmi	bahato kazna- nevidomo malo Boh zna(je)	bądź (nie)byle cholera wie lada mało	bůchví: jedno kdoví: lhostejno málo: nevím: sotva:	bilo Bog zna gdje- i- koje- malo mного
postponiert	by to ni bylo -to -libo -nibyd´ ugodno popalo	zavhodno popalo b (to) ne bulo -s´ -nebud´	bądź popadło -ś -kolwiek śkolwiek	chceš/chcete -koli(v) je libo	-god bilo

Datengewinnung und Analyse

Formalisierung der 26 Kontexte (CQL-
Abfragen)

Manuelle Datenbereinigung

Extrahierung von Frequenzen der 51
pronominalen Reihen

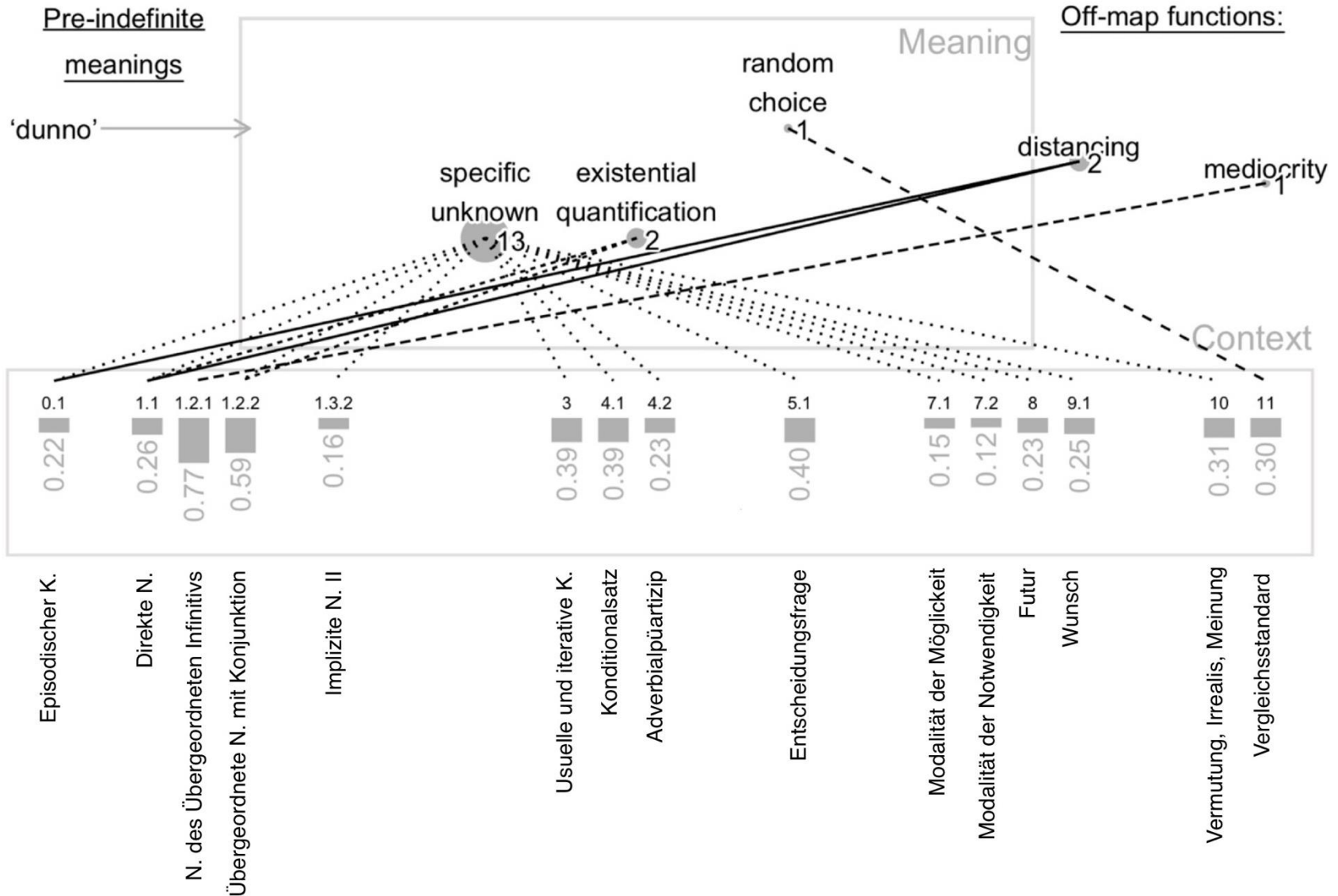
Normalisierung (mod. PMI, Skala 0–10)

Hierarchische Clusteranalyse

([lс="ніж"]|[tag="А.с.*](Rc)"][tag="S.*"]?
 ([lс="хто|к(ого|им|ому|ім)|що|ч(ого|ому|им|ім)|як(ий|ого|ому|ім|е|а|ої|ій|у|ою|і|их|им|ими)|котр(ий|ого|ому|им|ім|е|а|ої|ій|у|ою|і|их|ими)|скільки|ох|ом|ома)|як|де|навіщо|коли)-
 завгодно"]|[lс="хто|к(ого|им|ому|ім)|що|ч(ого|ому|им|ім)|як(ий|ого|ому|ім|е|а|ої|ій|у|ою|і|их|им|ими)|котр(ий|ого|ому|им|ім|е|а|ої|ій|у|ою|і|их|ими)|скільки|ох|ом|ома)|як|де|навіщо|коли"]|[lс="завгодно"]) within (</>
 !containing([lс="не"]|[lс="немає?"])|[lemma="(у|в|с)тримувати(ся)?|(у|в|с)тримати(ся)?|виключати(ся)?|виключити(ся)?|заперечувати(ся)?|заперечити(ся)?|запобігти(ся)?|запобігати(ся)?|бракувати(ся)?|унеможливити(ся)?|унеможливлювати(ся)?"]|[lemma="позбавляти(ся)?|позбавити(ся)?|віднімати(ся)?|відіймати(ся)?|відняти(ся)?|відійняти(ся)?|відшукати(ся)?|відібрати(ся)?|забирати(ся)?|забрати(ся)?|припиняти(ся)?|(при|у|с)пинити(ся)?|(у|в)?трачати(ся)?|позбуватися|(у|в)?тратити(ся)?|позбутися(ся)?|перешкодити(ся)?|перешкоджати(ся)?|забувати(ся)?|забути(ся)?|забороняти(ся)?|заборонити(ся)?|заборонувати(ся)?|(у|в)никати(ся)?|(у|в)никнути(ся)?|псувати(ся)?|з(і|о)псувати(ся)?|(і)?зникати|зникнути|обмежувати(ся)?|обмежити(ся)?|обмежувувати(ся)?|звільняти(ся)?|звільнити(ся)?|(при|зу|с)?пняти(ся)?|(при|зу|с)?пинити(ся)?|відмовляти(ся)?|зрікатися|відмовити(ся)?|зректися|відпадати(ся)?|відпасти|переставати|перестати|придушувати(ся)?|придушити(ся)?|губитися|загублювати(ся)?|за?губити(ся)?|пропадати|пропасти|(у|в|від)?хилитися|(у|в)хилитися|(у|в)хильнутися|(у|в)сувати(ся)?|відстороняти(ся)?|відсторонювати(ся)?|усунути(ся)?|відсторонити(ся)?|відхилити(ся)?|відхилювати(ся)?|відхилити(ся)?|(за)?лишати(ся)?|(за)?лишити(ся)?|заперечувати(ся)?|заперечити(ся)?|відкидати(ся)?|відкинути(ся)?|відучити(ся)?|відзвичуювати(ся)?|відчувати(ся)?|відзвичаїти(ся)?|рятувати(ся)?|позбутися|ліквідувати(ся)?|знищувати(ся)?|протестувати|опротестувати(ся)?|блокувати(ся)?|заблокувати(ся)?|делегітимізувати(ся)?|заглушати(ся)?|заглушити(ся)?|закінчуватися|закінчити(ся)?|закривати(ся)?|закрити(ся)?|нівелювати(ся)?|знівелювати(ся)?|знімати(ся)?|зняти(ся)?|нейтралізувати(ся)?|перекрити(ся)?|скасовувати(ся)?|скасувати(ся)?|стирати(ся)?|стерти(ся)?|руйнувати(ся)?|зруйнувати(ся)?|нищити(ся)?|знищувати(ся)?|знищити(ся)?"]|[lс="навіть"]|[lс="чи"]|[lс="ледве"]|[lс="що|чи"]?|[lс="сумнівно|марно|даремно|спірно|малоймовірно|[за]?ледве"]|[lemma="(за)?сумніватися|сумнівно"]|[lс="перше?|раніше"]|[lс=","]?|[lс="ніж"]|[lс="перед"]|[lс="тим"]|[tag="S.*"]])|((lс="максимум|щонайбільше|принаймні|щонайменше|максимально|мінімально|мінімум|максимум")|[lс="не"]|[lс="більше|менше"]|[lс="як|ніж"]?|[lс="хоча?"]|[lс="би?"]|[lс="як"]|[lс="мінімум"]])[tag="M.*"])|((lс="часто|іноді|зазвичай|щохвилино|щохвилини|щосекунди|щосекундно|щомісяця|щомісячно|щотижня|щорік|щороку|щорічно|щогодини|щогодинно|щомоментно|завжди|щоднини")|[lс="кожного"]|[lс="разу|дня|року|моменту|місяця"]|[lс="кожен|кожний"]|[lс="раз|день|рік|момент|місяць"]|[lс="кожну"]|[lс="мить|хвилину|секунду"]|[lс="кожної"]|[lс="мити"]|[lс="час"]|[lс="від"]|[lс="часу"]|[lс="з"]|[lс="року"]|[lс="в"]|[lс="рік"]|[lс="рік"]|[lс="у"]|[lс="рік"]|[lс="з"]|[lс="дня"]|[lс="на"]|[lс="день"]|[lс="день"]|[lс="у"]|[lс="день"]])|[lс="якщо|якби"]|[tag="V..g.*"&lemma=".*ти(ся)?"&lс!=".*ть|. *ться|. *ти"]|((lс="би?")|[lс="щоби?|аби"]|[lemma="кожен|скрізь|(в|у)сюди|завжди|(у|по)всюди"]|[lс="\"?"]|((lemma="з?могти|можна|можливо|можливий|з?(у|в)міти|мабуть|ймовірно|здатний")|[lemma="бути|мати"]])|[lс="змож(а|у)|можливість"]|[lemma="бути"]|[lс="в"]|[lс="стані|змозі"]|[lemma="мати"]|[lс="право"]|[lemma="бути"]|[lemma="спроможний"])|((lemma="треба|потрібно|необхідно|мусити|слід|доводиться|довестися|(з|ви)мушений|потрібний|повинний|належати")|[lemma="обов'язок|борг|зобов'язання|необхідний|неминучий|зобов'язаний|зобов'язаність"])|([tag="V..n.*"&lс=".*ти(ся)?"]|[lemma="мати"]])|[lemma="мати"]|[tag="V..n.*"&lс=".*ти(ся)?"]])|[tag="V...f.*"&lemma=".*ти(ся)?"]|((lemma="(за)?хотіти(ся)?|(по)?бажати(ся)?|прагнути(ся)?|(по)?мріяти(ся)?|з?воліти(ся)?|потребувати(ся)?|(по)?ждати(ся)?|збиратися|зібратися|наміряться|наміритися|(за)?планувати(ся)?"]|[lс="добре|здорово|непогано|чудово|корисно|лєво"]|[lс="би?")|[lemma="бути|мати"]|[lс="наміри?|плани?|мрі(я|ї|ю)|бажання|прагнення"])|((lemma="(по)?просити(ся)?|благати(ся)?|заблагати(ся)?|пропонувати(ся)?|запропонувати(ся)?|молити|(в|у)прошувати(ся)?|(в|у)прохувати(ся)?|(за)?кликати(ся)?|заохочувати(ся)?|запрошувати(ся)?|прикликати(ся)?|зазивати(ся)?|просьба|прохання|благання|пропозиція|мольба|моління|(за)?кликання|запрошення|заохочування|прикликання"]|[tag="V..m.*"&lemma=".*ти(ся)?"])|((lemma="дозволяти(ся)?|дозволити(ся)?|поступатися|поступитися|допускати(ся)?|допустити(ся)?|зволювати(ся)?|зволити(ся)?|дозволено|вільно"]|[lс="готов(ий|а|е|і)?|згоден(а|і|е)?"]|[lemma="дати|давати"]|[lс="зможу|право"]|[lemma="давати"]|[tag="V..n.*"&lс=".*ти(ся)?"])|((lemma="(по)?думати(ся)?|(в|у)важати(ся)?|сподіватися|сумніватися|припускати(ся)?|припустити(ся)?|передбачити(ся)?|передбачити(ся)?|боятися|здаватися|здати(ся)"]|[lс=","]?|[lс="дивно|можливо|мабуть|хіба|безсумнівно|очевидно|звичайно|либонь|(на)?певн(о|е)|допевне|мабути|надісь|надійсь|правдоподібно"]|[lс=","]?|[lс="найімовірніше|швидше|скоріш"]|[lс="за"]|[lс="все"]|[lс=","]?|[lс="певен|пєвн(а|о|і)|впєвнєн(ий|а|о|і)"]|[lс="можє"]|[lс=","]?|[lс="без"]|[lс="сумніву"]|[lс=","]?|[lemma="надія|припущєннє|безперєчнє"]|[lemma="з"]|[lemma="у?пєвнїстю"])))

Semantische Karte der RU Reihe um *ne pojmi*

RU *ne pojmi*

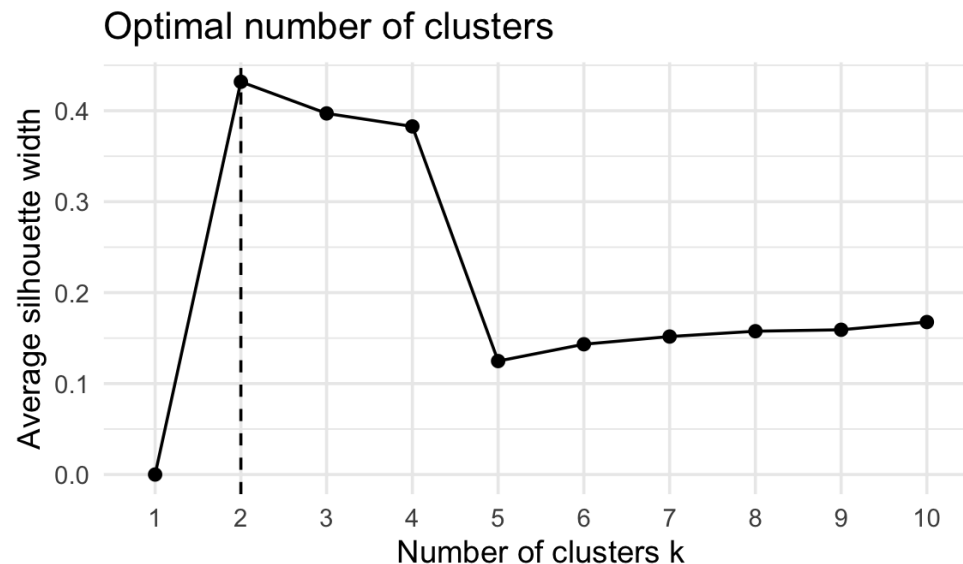


Hierarchische Clusteranalyse

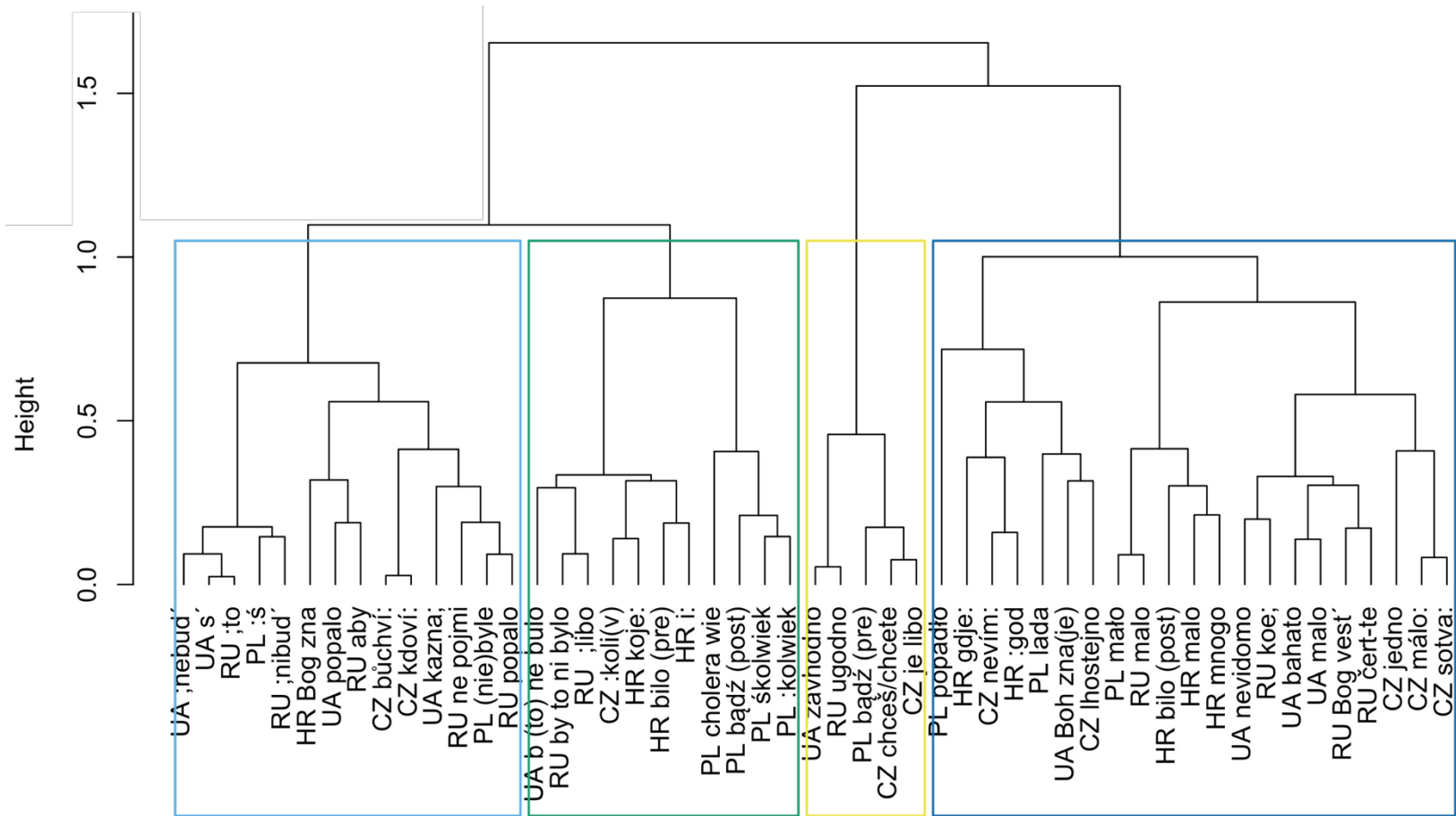
- **Daten:** relative Häufigkeiten der Pronomen in den Kontexten (mod. PMI)
Metrik: Pearson-Korrelation (Vergleich der Profile, nicht der Werte)
Methode: Clusteranalyse nach Ward (kompakte Cluster, geringere Streuung innerhalb der Cluster)
Ergebnis: Gruppierung der Serien nach Ähnlichkeit der Verteilungen

Silhouettenanalyse der Cluster:

- Optimale Aufteilung: 2 Cluster \Rightarrow schwache Strukturierung der Daten
- Für die weitere Analyse wurden 4 Cluster ausgewählt (Schwellenwert = 0,25)



Dendrogramm der Pronomencluster nach Verwendungskontexten



Universität Regensburg

- Cluster A: heterogen, unterteilt sich in NSIs und RCIs; die Unterteilung ist statistisch instabil, spiegelt aber die Theorie wider
- FCIs (Cluster C) bilden einen klaren Cluster (Ausnahme: CZ *koli(v)*)
- Andere Cluster: enthalten gemischte Typen

Fazit:

- ✓ Nur FCIs zeigen eine statistisch zuverlässige Verteilung
- ✗ Für andere Klassen ist das Lizenzieren eher eine Tendenz als eine Regel

A	B	C	D
UA nebud´	UA b (to) ne	UA zavhodno	PL popadło
UA -s´	bulo	RU ugodno	HR gdje
RU to	RU by to ni bylo	PL bądź (pre)	CZ nevím
PL -ś	RU libo	CZ	HR god
RU nibud	CZ koli(v)	chceš/chcete	PL lada
	HR koje	CZ je libo	UA Boh zna(je)
	HR bilo (pre)		CZ lhostejno
	HR i		PL mało
	PL cholera wie		RU malo
	PL bądź (post)		HR bilo (post)
HR Bog zna	PL školwiek		HR malo
UA popalo	PL kolwiek		HR mnogo
RU aby			UA nevidomo
CZ bůchví			RU koe
CZ kdoví			UA bahato
UA kazna			UA malo
RU ne			RU Bog vest´
pojmi			RU čert-te
PL (nie)byle			CZ jedno
RU popalo			CZ málo
			CZ sotva
14	11	5	21

Fazit

1. Die Bestimmung des semantischen Typs slavischer IPs allein anhand ihrer kontextuellen Distribution ist nicht möglich (Ausnahme: FCIs)
2. Ausnahme: FCIs. Sie bilden ein statistisch zuverlässiges Cluster
3. Lizenzierung ist quantitativ, nicht qualitativ: Unterschiede zwischen Kontexten zeigen sich in den Häufigkeiten, nicht in absoluter Zulässigkeit oder Unzulässigkeit
4. Lizenzierung ist eine Tendenz, kein Regelwerk: Daten großer Korpora zeigen eher eine Neigung zu bestimmten Kontexten als eine strikte Regel
5. Notwendigkeit einer Revision der Theorie: Bestehende Modelle der Lizenzierung müssen unter Berücksichtigung quantitativer Korpusdaten präzisiert werden

Literatur

- Barić, Eugenija et al. (1997) *Hrvatska gramatika*. Zagreb
- Bondareva, G. A. (2010) Sostavnye mestoimenija v ruskom jazyke. (Diss. kand. Voronež: Voronežskij gosudarstvennyj universitet)
- Dudko, I. V. (2010) Neoznacheni zaimennyky v ukrainskii movi: istoriia ta suchasnyi stan. In: *Naukovyi chasops NPU imeni M. P. Drahomanova* 6, 104–111
- Fisun, R. & Hansen, B. (2025) A corpus based study on the inventory and the grammaticalisation of compound indefinite pronouns in Croatian. *Die Welt der Slaven* 69(2), 293-330. URL: https://www.researchgate.net/publication/390201462_A_corpus-based_study_on_the_inventory_and_grammaticalisation_of_compound_indefinite_pronouns_in_Croatian [25.05.2025]
- Fisun, R. S. (2024) Zusammengesetzte Indefinitpronomen in den slavischen Sprachen (Doktorarbeit, Universität Regensburg)
- Giannakidou, A. (1998) *Polarity sensitivity as (non)veridical dependency*. Amsterdam
- Kuryłowicz, J. (1965). The Evolution of Grammatical Categories. *Diogenes*, 13(51), 55–71
- Lehmann, Ch. (2015[1982]). *Thoughts on Grammaticalization* (3rd edition). (Classics in Linguistics 1). Berlin
- Lehmann, Ch. (o. D) Grammatikalisierung URL: https://www.christianlehmann.eu/ling/ling_theo/grammatikalisierung.php [6.08.2023]
- Maykova, Irina (2023) Zur Funktionalstilistik ausgewählter Indefinitpronomen im Russischen. Publikationssever UR. URL: <https://epub.uni-regensburg.de/54410/2/Masterarbeit%20Maykova%20digitale%20Version.pdf> [25.05.2025]
- Ožegov, S. I. (o. D.). Tolkovyj slovar' Ožegova onlajn. URL: <https://slovarozhegova.ru/> [12.03.2023]
- Padučeva E., (2018) Russkie mestoimenija svobodnogo vybora. *Russian Linguistics* 3(42), 291–319
- Padučeva, E. V. (1985) *Vyskazyvanie i ego sootnesennost' s dejstvitel'nost'ju*. Moskva
- Padučeva, E. V. (2015) Snjataja utverditel'nost' i neveridikal'nost' (na primere russkich mestoimenij otricateľ'noj poljarnosti). *Russian Linguistics* 2(39), 129–162
- Rybová, M. (2023) Die Beziehung zwischen Evaluativität und Indefinitheit am Beispiel der zusammengesetzten Indefinitpronomen im Tschechischen (Doktorarbeit, co-tutelle UK Prag – Universität Regensburg)
- Silić, J. & I. Pranjković (2007) *Gramatika hrvatskoga jezika za gimnazije i visoka učilišta*. Zagreb
- Sketch Engine (o. D.). URL: <https://www.sketchengine.eu> [12.03.2023]
- Sokolova, S. V. (2007) Dinamičeskie processy v sisteme mestoimennykh slov sovremennogo russkogo jazyka. (Diss. kand. Moskva: MGU)
- Testelec, Ja. G. & E. G. Bylinina (2005) O nekotorych konstrukcijach so značenijem neopredelennykh mestoimenij v ruskom jazyke: amal'gamy i kvazireljativy. Doklad k seminaru „Teoretičeskaja semantika“ IPPI RAN



Universität Regensburg

DFG Deutsche
Forschungsgemeinschaft

Danke für die Aufmerksamkeit!