



From Solitary Corpus Analysis to Collective Insight: a Glimpse into Translation and Lexicography

Łukasz Grabowski (University of Opole)

Slovko 2023, Bratislava (18 October 2023)

Plan of my talk

Corpus linguists' work: towards expanding and scaling up corpus analyses

• Interdisciplinarity & collaboration

Case Study 1: Exploring repetition in translation

• Translation of reporting verbs from English into Polish and Italian

Case Study 2: Exploring English and Polish lexicographers' habits

• Which multi-word expressions get recorded in dictionaries?

Corpus linguistics and corpus linguists in the 2020s and beyond

• Reflecting upon and reimagining the corpus linguist's role in a collaborative landscape

Typical corpus linguists' work

- Corpus linguists are really just linguists who have decided to research language using corpora -> their research interests often align with those of other linguists (Ihrmark & Tyrkkö 2023)
 - Some of them use corpora to address theoretical questions,
 - Some turn to corpora to collect the data they need to produce an accurate description of a linguistic phenomenon
 - Others are preoccupied with how corpora should be compiled and structured to be representative of a population of language speakers and useful to a group of corpus users (Divjak et al. 2017)
- For most corpus linguists, corpus data analysis (text analysis) is part and parcel of their work
 - a corpus linguist seems to be a bit of a data scientist

Is corpus linguists a data scientist? (Ihrmark & Tyrkkö 2023)

Differences in how corpus linguists and data scientists view language

• Data scientists are often doing something via text analysis (e.g. identify trends, sentiment analysis, fraud detection), while **linguists are interested in learning more about language itself**

Differences in the type of training they typically receive

- Corpus linguists rarely receive explicit training in programming or statistics
 - although basic statistical concepts are often included to some degree
- Corpus linguists are often trained in the use of specific corpus tools to carry out their queries
 - o only some in the field engage with the things that happen "under the hood" (Mason 2018)
- Data scientists are often trained in programming and advanced statistics
 - rarely introduced to theoretical details about language

Towards expanding corpus analyses

- Simple descriptive statistics or concordance analysis is often not enough to explain more fine-grained patterns of language use in massive collections of linguistic data
 - if we put forward explanatory hypotheses rather than descriptive ones
 - if we deal with large swathes of linguistic data
- The range of statistical methods of corpus analysis as well as data visualization techniques have expanded considerably in recent years
- Corpus analysis can be also scaled up by incorporating computational methods from natural language processing (NLP) (Dunne 2022) and statistics

Two examples as a case in point

Research studies on:

- **repetition in translation of reporting verbs** (Mastropierro & Grabowski 2023, in progress)
- lexicographer's habits with respect to recording multi-word expressions in dictionaries (Maziarz, Grabowski, Piotrowski, Rudnicka & Piasecki 2023)

What skills should corpus linguists possess in the 2020s and beyond?

• How to succeed (or survive) in the coming era of Al-assisted (corpus) research, datasphere growth and attention economy?

Case study #1

Repetition in translation

• Mastropierro & Grabowski (2023-2024, in progress)

Repetition in translation

A corpus-based study - both descriptive and explanatory - dealing with literary translation, language use, and statistics

How are reporting verbs translated in literary novels and why they are translated the way they are

- RQ1: What are the **linguistic factors that can predict the replacement of reporting verbs** in the source-text with different target translations?
- RQ2: Which of the factors have a significant effect on either the avoidance or reproduction of reporting verbs' repetition in the English-to-Polish and English-to-Italian translation?

Repetition in translation

- "Avoiding repetition of words or phrases is part of a set of translation norms found to operate consciously or subconsciously on the translator" (Ben-Ari 1998: 2).
- Corpus-based studies of the translation of repeated features in literary texts (Čermáková & Fárová 2010; Čermáková 2015; Mastropierro & Mahlberg 2017, Čermáková & Mahlberg 2018, Mastropierro 2020) show that translators tend to avoid repetition in favour of variation.
 - Replacing repetition with lexical variety can have repercussion on the contribution repetition makes in the creation of stylistic effect

Repetition in translation

- Translators should restrain themselves from "synonimization" (i.e. using synonyms in an attempt to avoid repetition), they should preserve all repetitions found in the original, which is sometimes difficult given the fact that "the translators are nuts about repetitions" (Kundera 2015: 164)
- "The original's repetitions deserve respect" (Bray 2017)
- English speakers tend to choose lexemes with broad meaning rather than specific while speakers of Russian, Polish and other Slavic languages tend to choose specific lexemes, which is one of the reasons why they regard texts with multiple repetitions as ones with plain, simple, or even bad style (Piotrowski 1994: 95-96)
- Anecdotal evidence: avoiding repetition recommended when teaching writing at Polish primary/secondary schools
 - ► Use thesauri, don't repeat the same words

A study of repetition in translation

- Are there linguistic factors that can predict the replacement of a repeated lexical item with different target translations? (e.g. said -> powiedział, stwierdził, zaznaczył, odnotował, podkreślił ...)
 - ► Item frequency;
 - Number of different dictionary translations;
 - ► Number of different dictionary meanings;
 - ► Semantic category.
- What type of repeated items are more likely to be translated in different ways?

Earlier studies of reporting verbs in translation

► Why reporting verbs? Stylistic and functional significance!

- English-Italian (Mastropierro 2020; 2022), English-Spanish (Bourne 2002), English-Persian (Ardekani 2002), Czech-English and Czech-French (Nádvorníková 2020)
- Patterns of reporting verbs contribute importantly to characterisation (Ruano San Segundo 2016, Ruano San Segundo 2017, Eberhardt 2017, Mastropierro 2020).
- In literary texts, English reporting verbs are regularly substituted by words with emotive and evaluative meanings when translated from English into Russian (Urza (2019)
- In Russian-to-Polish translation of dialogues in literary texts (novels by V. Pelevin), there is a wide variety of Polish equivalents of the source-language reporting verb сказать 'say' Lubocha-Kruglik & Malysa (2019)

Translation criticism

- Defending her choices, APT explains that the Polish language, unlike English, is not able to bear the constant repetition of the form "he said" when recording the dialogue. (Umiński 2020, 2022: Three Translators/Trzy Tłumaczki. Warsaw: Marginesy)
 - It's about the translation of W. Faulkner's "The Sound and the Fury" into Polish by A. Przedpełska-Trzeciakowska

A study of repetition in translation

Translation from English into Polish and Italian;

► 11 + 11 target texts;

Reporting verbs in 11 English novels and their translations into Polish and Italian, retrieved via InterCorp (Čermák & Rosen 2012).

5 different authors (Rowling, Brown, Steinbeck, Adams, Tolkien);

Multiple translators

Data collection

- English novels with both Italian and Polish translations;
- ► \geq 500 reporting verbs (tokens) per book (both ST and TT);
- Parallel English-Italian and English-Polish queries in Intercorp

► CQL query for English:

[word="\""] [tag="NP"|word="he|she"] [tag="VBD"] | [word="'"] [tag="NP"|word="he|she"] [tag="VBD"] | [word="""]
[tag="NP"|word="he|she"] [tag="VBD"] | [word="'"] [tag="NP"|word="he|she"] [tag="VBD"] | [word="\""]
[tag="VBD"] [tag="NP"|word="he|she"] | [word="'"] [tag="VBD"] [tag="NP"|word="he|she"] | [word="""]
[tag="NP"|word="he|she"] | [word="'"] [tag="VBD"] [tag="NP"|word="he|she"] | [word="""]

Corresponding queries for Italian and Polish

Data collection

 English query paired with equivalent Italian/Polish query to obtain parallel concordance lines.

	InterCorp v15 - English 🔽		InterCorp v15 - Italian 🔽
Tolkien-Pan_prstenu_1	And I might say it to others , ' he added with a look at the stranger and the miller .	Tolkien-Pan_prstenu_1	E lo dico anche a voi ${\rm *}$, ${\rm aggiunse}$ lanciando uno sguardo al mugnaio ed al forestiero .
Tolkien-Pan_prstenu_1	Well-preserved indeed ! ' he snorted .	Tolkien-Pan_prstenu_1	E poi dicono che mi mantengo bene !? », sbuffò . « Io che mi sento tutto magro , come dire , teso ; rendo l'idea ? Come del burro spalmato su di una fetta di pane troppo grande .
Tolkien-Pan_prstenu_1	' No , it does not seem right , ' he said thoughtfully .	Tolkien-Pan_prstenu_1	« Hai ragione , non può essere normale » , disse pensoso .
Tolkien-Pan_prstenu_1	' Is n't that odd now ? ' he said softly to himself .	Tolkien-Pan_prstenu_1	», mormorò incantato .
Tolkien-Pan_prstenu_1	' I think , Bilbo , ' he said quietly , ' I should leave it behind .	Tolkien-Pan_prstenu_1	« Credo , Bilbo » , disse pacatamente , « che sarebbe meglio lasciarlo , quest' anello .
Tolkien-Pan_prstenu_1	' You will be a fool if you do , Bilbo , ' he said .	Tolkien-Pan_prstenu_1	« Sei un pazzo se lo fai , Bilbo » , disse ;
Tolkien-Pan_prstenu_1	' It will be my turn to get angry soon , ' he said .	Tolkien-Pan_prstenu_1	« Fra poco sarò io ad arrabbiarmi » , disse . « Guai a te se ripeti una sola volta quel che hai detto !
Tolkien-Pan_prstenu_1	' I do n't know what has come over you , Gandalf , ' he said .	Tolkien-Pan_prstenu_1	« Non capisco che cosa ti succeda , Gandalf » , disse ; « non ti ho mai visto così prima d' oggi .
Tolkien-Pan_prstenu_1	' I have never called you one , ' Gandalf answered .	Tolkien-Pan_prstenu_1	« Non ti ho mai accusato di esserlo » , rispose Gandalf , « e nemmeno io lo sono .
Tolkien-Pan_prstenu_1	' I am sorry , ' he said .	Tolkien-Pan_prstenu_1	« Mi dispiace » , <mark>disse</mark> , « ma mi sentivo così strano .

kon text Dotaz Korpusy Uložit Konkordance Filtr Frekvence Kolokace Zobrazení Nápověda Korpus: InterCorp v15 - English | Dotaz: ,] [tag=, |word=,] [tag=,] within <te... (3 129 výskytů) ~ Podrobnosti Výskytů: 3 129 | i.p.m.: Spočítat | ARF: 92,07 | Výsledek je setříděn 1 / 79 Výběr řádků: základní 🗸 InterCorp v15 - English Z InterCorp v15 - Polish adams-sbohem_a_dik " Thank you , " he called out , " thank you very ... " - Dziękuję ! - krzyknęła postać . - Bardzo wam ... adams-sbohem_a_dik adame-shohem a dik " It 's my big number , " she said adame-shohem a dik - To mój hit - powiedziała , zauważając zachwyt Forda adams-sbohem_a_dik " Then the planet got demolished , " he shouted adams-sbohem_a_dik - Potem planeta została zniszczona adams-sbohem_a_dik " What did she say ? " he asked anxiously adams-shohem a dik - Co ona powiedziała ? - z lekiem spytał Artur adams-sbohem_a_dik " Do n't worry about the umbrella , " she said to him as she climbed in . - Prosze się nie martwić parasolka - powiedziała wsiadając adams-sbohem_a_dik adams-sbohem_a_dik " No , London , " she said . adams-sbohem_a_dik - Nie , w Londynie - odparła " I think , " she said , with a momentary sigh of relief , " you should buy me a drink - Sądzę - powiedziała Fenny z westchnieniem ulgi - że powinien mi pan przed odjazdem pociągu adams-sbohem_a_dik adams-sbohem_a_dik before my train goes . " postawić drinka " I do hope you win, " she said with a smile that suddenly snapped together like a - Mam ogromna nadzieję, że pan wygra - powiedziała z uśmiechem, który rozkwitł jak origami adams-sbohem_a_dik adams-sbohem_a_dik piece of advanced origami, " the prizes are so nice . " dla zaawansowanych . - Nagrody są takie ładne ... " No , that 's all right , really , " he said with a wave of his hand , and she started with - Nie , proszę dać sobie spokój - powiedział machając ręką , kobieta zaczęła jednak z adams-sbohem a dik adams-sbohem a dik an awful deliberation to peel off five tickets , one by one przerażająca ślamazarnościa odrywać pieć losów . Jeden po drugim . adams-sbohem_a_dik " All right , " she said with a laugh , " you can ask me . adams-sbohem_a_dik - No dobrze - powiedziała ze śmiechem " I 'll let you see the prizes in a minute . " she said . and left . - Zaraz panu pokaże nagrody - dodała na odchodnym adams-sbohem a dik adams-sbohem a dik " Now we can relax , " she said with a slow smile which filled Arthur till he thought he - No to możemy odetchnąć - rzekła z uśmiechem , który tak wypełnił duszę Artura , iż zaczął się adams-sbohem_a_dik adams-sbohem_a_dik would burst bać, że zaraz pęknie jak balonik

Raw data

Narzedzia główne Wstawianie Układ strony Formuły Dane Recenzia Widok

🔀 🔚 🤊 • (* - 🖃

HP and the Chamber of Secrets - Microsoft Excel użytek niekomercyjny

_				X	
	S	2	0	ı B	23

				nanj bane necenzja maon							
	D8	\bullet (° f_x ,	" all right "								• 47
1	A	В	С	D	E	F	G	Н	1	J	К
1	rowling	" Third time this week !	" he roared	across the table .	rowlingova -	- To już trzeci raz w tym tygodniu !	- ryknął	na niego p	oprzez sto	śł.	
2	rowling	" Do I look stupid ?	" snarled Uncle	Vernon , a bit of fried egg dang	rowlingova -	- Czy ja wyglądam na głupca ?	- warknął	wuj Verno	n . <mark>Z kr</mark> zac	zastego wa	isa zwisa
3	rowling	" There 's more in the frying	" said Aunt	Petunia , turning misty eyes on	rowlingova -	- Jest jeszcze trochę na patelni , syneczku	- odpowiedziała	ciotka Petu	inia , spog	lądając tkli	wie na sv
4	rowling	"Nonsense , Petunia , I neve	" said Uncle	Vernon heartily .	rowlingova -	- Ależ to nonsens , Petunio ! Kiedy ja był em w	- oświadczył	stanowczo	wuj Vern	on .	
5	rowling	" You 've forgotten the mag	" said Harry	irritably .	rowlingova-	hpot_tajemn	- Zapomniał	eś magiczr	ego słowa	a - odpowie	edział ze
6	rowling	" I meant ` please ' !	" said Harry	quickly .	rowlingova -	- Chodziło mi o " proszę " !	- powiedział	prędko Ha	rry.		
7	rowling	" HOW DARE YOU THREATE	" roared Uncle	Vernon , pounding the table w	rowlingova -	- Jak śmiesz grozić Dudleyowi !	- ryknął	wuj Verno	n , waląc p	pięścią w st	ół ostr
8	rowling	" All right ,	" said Harry	, " all right "	rowlingova -	- Dobrze , wuju , dobrze	- powiedział				
9	rowling	" This could well be the day	" said Uncle	Vernon .	rowlingova -	- To może być dzień , w którym dokonam najw	- rzekł	wuj Verno	n.		
10	rowling	" I think we should run thro	" said Uncle	Vernon .	rowlingova -	- Myślę , że dobrze by było jeszcze raz przejrze	- powiedział	wuj Verno	n.		
11	rowling	" In the lounge ,	" said Aunt	Petunia promptly , " waiting to	rowlingova -	- W salonie	- odpowiedziała	natychmia	st ciotka P	'etunia - go	otowa po
12	rowling	" They 'll love him !	" cried Aunt	Petunia rapturously .	rowlingova -	- Będą nim zachwyceni !	- zawołała	entuzjasty	znie ciot	a Petunia .	
13	rowling	" Excellent , Dudley ,	" said Uncle	Vernon .	rowlingova -	Znakomicie , Dudley	- pochwalił	go wuj Ver	non , po c	zym zwróc	ił się do I
14	rowling	" I 'll be in my bedroom , ma	" said Harry	tonelessly .	rowlingova -	- Ja będę siedział cicho w swojej sypialni , udaj	- odrzekł	Harry bezk	arwnym	tonem .	
15	rowling	" Exactly ,	" said Uncle	Vernon nastily .	rowlingova -	- Dokładnie	- powiedział	dobitnie w	uj Vernon	1.	
16	rowling	" I 'll announce dinner ,	" said Aunt	Petunia .	rowlingova -	- Oznajmię , że kolacja gotowa	- powiedziała	ciotka Petu	inia .		
17	rowling	" May I take you through to	" said Dudley	, offering his fat arm to an invi:	rowlingova -	 Czy mogę panią zaprowadzić do jadalni , pan 	- powiedział	Dudley, of	erując rar	nię niewid:	zialnej kc
18	rowling	" My perfect little gentlema	" sniffed Aunt	Petunia .	rowlingova -	- Mój doskonały mały dżentelmen !	- zagdakała	ciotka Petu	ınia .		
19	rowling	" And you ?	" said Uncle	Vernon viciously to Harry .	rowlingova -	- A ty ?	- warknął	wuj Verno	n , patrząc	: na Harry (ego.
20	rowling	" I 'll be in my room , makin	" said Harry	dully .	rowlingova -	- Ja będę siedział cicho w swoim pokoju , udaj	- powiedział	tępo Harry			
21	rowling	" I 'll be in my room , makin	" he said	•	rowlingova -	- Ja będę siedział cicho w swoim pokoju i udav	- wyrecytował				
22	rowling	" Too right , you will ,	" said Uncle	Vernon forcefully .	rowlingova -	- Tak jest i są ku temu powody	- rzekł	dobitnie w	uj Vernon	1.	
23	rowling	And you ,	" he snarled	at Harry .	rowlingova	A ty	- warknął	w kierunku	Harry'eg	o - nie pałę	taj się po
24	rowling	" I know what day it is ,	" sang Dudley	, waddling toward him .	rowlingova -	- A ja wiem , co dzisiaj jest , aha !	- zaśpiewał	Dudley , zr	nierzając	w jego stro	nę.
25	rowling	" What ?	" said Harry	, not taking his eyes off the spo	rowlingova -	- Co ?	- zapytał	Harry , nie	spuszczaj	ąc wzroku	z miejsca
26	rowling	" I know what day it is ,	" Dudley repeat	, coming right up to him .	rowlingova -	Wiem , co dzisiaj jest	- powtórzył	Dudley, po	odchodzą	c do niego	
27	rowling	" Well done ,	" said Harry	Carl all all all all all all all all all	rowlingova -	- Brawo !	- powiedział	Harry.			

Data collection

- Reporting verbs with minimum frequency ≥ 2 ;
- ► 23-79 ST reporting verbs, types (average 47.45 per text);
- ► 16,742 ST reporting verbs (tokens).
- Data preparation using custom-designed Python scripts

- What linguistic factors have an effect on the number of different translations a ST item is translated into?
- Outcome variable:
 - **Types** the number of different TT translations (types) for each ST item.
- ► Predictor variables:
 - ► **Freq** the frequency of occurrence of the ST item;
 - ► **Trans** the number of different possible TL translations of the ST item;
 - ► Senses the number of different possible meanings of the ST item;
 - **Verbtype** the category of reporting verb in Caldas-Coulthard's (1987) taxonomy.

- ▶ **Trans** the number of different possible TL translations of the ST item
- Treq (Translation Equivalents Database) (Skrabal & Vavrin 2017)
 - ► Lemmas;
 - Only verbs translations;
 - ▶ Proportion $\ge 4\%$



English

interject Search						
🔽 Lemma 🤊	Multiword ?	RegEx ?	✓ A = a ?			
▲ Frequency ▼	▲ Proportion ▼	▲ English ▼	🔺 İtalian 🔻			
10	35.7	interject	intervenire			
8	28.6	interject	interrompere			
2	7.1	interject	ribatté			
2	7.1	interject	interloquire			
1	3.6	interject	introdurre			
1	3.6	interject	intromettere			
1	3.6	interject	precisare			
1	3.6	interject	domandare			
1	3.6	interject	spiegare			
1	3.6	interject	intercedere			
28						

- ► Senses the number of different possible meanings of the ST item.
- ► Wordnet (Lexical Database of English):
 - ► Number of different senses reported.

 Verbtype - the category of reporting verb in Caldas-Coulthard's (1987) taxonomy.

- Caldas-Coulthard's (1987) taxonomy:
 - Seven main categories;
 - ► Both linguistic & paralinguistic reporting verbs;
 - ► Functional as well as semantic distinction.

Category	Sub-category	Examples
Neutral		say, tell
Structuring		ask, inquire, reply, answer
Metapropositional	Assertive	exclaim, proclaim, agree
	Directive	urge, instruct, order
	Expressive	accuse, lament, swear
Metalinguistic		narrate, quote, recount
Prosodic		cry, shout, scream
Paralinguistic	Voice qualifier	whisper, murmur, mutter
	Voice qualification	laugh, sigh, groan
Signalling discourse		repeat, add, go on, hesitate

Towards final data format (right bottom)

	rzędzia główne		, ,			MOOK.											∞ 🕜 =	o Ma
C2	*	f_r ryknął																
Α	В	с	D	E	F	G	н	1	J	К	L	M	N	0	Р	Q	R	1
word	count	variant	variant cou	int														
roared	10	ryknął	6															
roared		ryczał	2															
roared		krzyknął	1															
roared		wrzasnął	1															
snarled	7	warknął	6															
snarled		burknął	1															
said	947	odpowiedziała	18															
said		oświadczył	14															
said		zapomniał	1															
said		powiedział	259															
said		rzekł	51															
said		pochwalił	2															
said		odrzekł	31															
said		powiedziała	56															
said		warknąf	7															
said		wyrecytowaf	1															
said		zapytaf	108															
said.		oupowiedzian kondenał	52															
said		krzyknąt	19															
said		wyjąkai	12															
said		zachrypiaf	1															
said		przerwai	11															
said		nruknat	14															
said		odezwał	10															
sald		powtorzyi	0															
	rzędzia główne	Wstawianie Układ str	ony Formuły	Dane	PL, Recenzja V	_1_HP_CS_Po /idok	olishtypes - I	Microsoft Ex	cel użytek niek	omercyjny						-	0 0 - 0	× 7 83
C2	rzędzia główne • (Wstawianie Układ str <i>f</i> w {'wskaza	ony Formuły ał', 'wzγwał', 'e	Dane odparł', 'r	PL, Recenzja W ucieszył', 'p	1_HP_CS_Po /idok olecił', 'zag	olishtypes - I gadnąť, 'wy	Microsoft Ex sapal', 'zap	cel użytek niek perzył', 'uciął'	omercyjny , 'zaklikať	, 'ciągnął',	, 'wypalił', 'ı	wykrzykną	, 'powital	, 'uciszył',	'skarcil', 'zr	o 🕜 🗖	× ਭ ಜ
C2	rzędzia główne • (Wstawianie Układ str fr {'wskaza 'zdał', 'p	ony Formuły ał', 'wzywał', 'o przedstawił', 'v	Dane odparť, 'r vpadť, 'z	PL Recenzja W ucieszył', 'p denerwow	_1_HP_CS_Po //dok olecił', 'zag ał', 'chciał',	olishtypes - I gadnąť, 'wy , 'przyznať,	Microsoft Ex sapał', 'zaj 'zwariowa	cel użytek niek perzył', 'uciął' ł', 'rzekł', 'zap	omercyjny , 'zaklikał' zewnił', 'z	, 'ciągnął', awotał', 'p	, 'wypalił', 'n comrukiwał	vykrzykną ', 'odrzekł'	', 'powital 'wyznai',	', 'uciszył', 'stracił', 'o	'skarcil', 'zr burzyl', 'do	alazť, dať,	× 37 23
C2	rzędzia główne • (B	Wstawianie Układ str fr {'wskaza 'zdał', 'p	ony Formuły ał', 'wzywał', 'o przedstawił', 'v	Dane odparť, 'r vpadť, 'z	PL Recenzja W ucieszył', 'p denerwow	_1_HP_CS_Po Adok olecił', 'zag ał', 'chciał', C	olishtypes - I gadnąť, 'wy , 'przyznať,	Microsoft Ex sapal', 'zaj 'zwariowa	cel użytek niek perzył', 'uciął' ł', 'rzekł', 'zap	omercyjny , 'zaklikał' bewnił', 'z	, 'ciągnął', awotał', 'p D	, 'wypalii', 'n comrukiwał E	vykrzykną ', 'odrzekł' F	', 'powital 'wyznař, G	', 'uciszył', 'stracił', 'o H	'skarcił', 'zr burzył', 'do I	o alazi', dai', J	× 77 × 12
C2 A source_E	• (B N_source_	Wstawianie Układ str	ony Formuły ał", 'wzywał", 'o przedstawił", 'v	Dane odparť, 'r vpadť, 'z	PL, Recenzja W ucieszył', 'p denerwow	1_HP_CS_Po /idok olecił', 'zag ał', 'chciał', C	olishtypes - I gadnąť, 'wy , 'przyznať,	Microsoft Ex sapal', 'zap 'zwariowa	cel użytek niek perzył', 'uciął' ł', 'rzekł', 'zap	omercyjny , 'zaklikał' pewnił', 'z TT	, 'ciągnął', awołał', 'p D _types_Va	, 'wypalił', 'womrukiwał E Ir_1_freq	wykrzykną ', 'odrzekł' F	', 'powital 'wyznař, G	', 'uciszył', 'stracił', 'o H	'skarcil', 'zr burzył', 'do I	ialazi', dai',	× 77 4 *
C2 A source_E said	rzedzia główne	Wstawianie Układ str fr ('wskazz 'zdał', 'p EN_crTT_types 947 ('wskazał', 'wzyw	ony Formuły ał', 'wzywał', 'o orzedstawił', 'v wał', 'odparf', '	Dane odparť, 'r vpadť, 'z ucieszyť,	PL, Recenzja W ucieszył', 'p denerwow , 'polecił', 'z	1_HP_CS_Pc /idok olecił', 'zag ał', 'chciał', C agadnął', '	olishtypes - I gadnął', 'wy , 'przyznał', wysapał', 'z	Microsoft Ex sapal', 'zap 'zwariowa :aperzył', 'u	cel użytek niek perzył', 'uciął' ł', 'rzekł', 'zap uciął', 'zaklika	omercyjny , 'zaklikał' bewnił', 'z TT, ť, 'ciąg	, 'ciągnął', awołał', 'p D _types_Va _114	, 'wypalił', 'v omrukiwał E ir_1_freg 947	wykrzykną ', 'odrzekł' F	', 'powital 'wyznał', G	', 'uciszył', 'stracił', 'o H	'skarcil', 'zr burzył', 'do I	alazi', dai', J	× 7 83 7 4
C2 A source_E said whispere	rzędzia główne	Vstawianie Układ str // ('wskaza' 'zdał', 'p EN_crTT_types 947 ('wskazał', 'wzyu 23 ('szepnął', 'wysz	ony Formuły ař, 'wzywał', 'e orzedstawił', 'e wał', 'odparf', ' eptař, 'oznajn	Dane odparť, 'r vpadť, 'z ucieszyť, niť, 'zapy	PL, Recenzja W ucieszył', 'p denerwow , 'polecił', 'z rtał'}	1_HP_CS_Po //dok olecił', 'zag ał', 'chciał', C agadnął', '	olishtypes - 1 gadnął', 'wy , 'przyznał', wysapał', 'z	Microsoft Ex sapał', 'zaj 'zwariowa saperzył', 'u	cel użytek nieko perzył', 'uciął' ł', 'rzekł', 'zap uciął', 'zaklika	omercyjny , 'zaklikał' bewnił', 'z TT, ľ', 'ciąg	, 'ciągnął', awołał', 'p D _types_Va _114 _4	,'wypalił', 'v omrukiwał E ir_1_freq 947 23	wykrzykną ', 'odrzekł' F	', 'powital 'wyznał', G	', 'uciszył', 'stracił', 'o H	'skarcil', 'zr burzył', 'do I	alazi', J	× 37 83 4 •
C2 A source_E said whispere shouted	rzędzia główne ▼ (B N_source_ ed	Wstawianie Układ stru 'zdał', 'g N_crT_types 947 {'wskazał', 'wzyw 23 {'szepnął', 'wysz 23 {'krzyknął', 'rykn	ony Formuły ał", 'wzywał', 'w rrzedstawił", 'w wał', 'odpart", ' eptał', 'oznajn rął', 'zawolał', '	Dane odparť, 'u vpadť, 'z ucieszyť, niť, 'zapy szydziť, '	PL, Recenzja W ucieszył', 'p denerwow . 'polecił', 'z tał'} 'powiedział	1_HP_CS_Pc fidok olecil', 'zag al', 'chcial', C agadnął', ' ', 'wrzasną	olishtypes - 1 gadnął', 'wy 'przyznał', wysapał', 'z ł ¹ }	Microsoft Ex sapal', 'zag 'zwariowa :aperzył', 'u	cel użytek nieko perzył', 'uciął' ł', 'rzekł', 'zap uciął', 'zaklika	omercyjny , 'zaklikał' pewnił', 'z TT, f', 'ciąg	, 'ciągnął', awołał', 'p D _types_Va 114 4 6	, 'wypalił', 'n omrukiwał E or_1_freq 947 23 23	wykrzykną ', 'odrzekł' F	', 'powital 'wyznał', G	', 'uciszył', 'stracił', 'o H	'skarcil', 'zr burzył', 'do I	ialazi', dai', J	× 7 * * *
C2 A source_E said whispere shouted muttered	rzędzia główne ▼ (B IN_source_ ed	Wistawianie Układ str. //w ('wskaza' 'zdai', 'p N. ot TT_types 947 ('wskaza', 'wzyw 23 ('szepnaj', 'wzys 23 ('szepnaj', 'wysz 23 ('krzyknaj', 'rykn 20 ('warknaj', 'rykn	ony Formuły ał', 'wzywał', 'w przedstawił', 'w vał', 'odparf', ' eptał', 'oznajn iął', 'zawołał', ' namrotał', 'beł	Dane odparť, 'u vpadť, 'z ucieszyť, niť, 'zapy szydziť, ' kotať, 'ję	PL, Recenzja W ucieszył', 'p denerwow , 'polecił', 'z tał') 'powiedział ,knął', 'napi	1_HP_CS_Po fidok olecił', 'zag ał', 'chciał', C agadnął', ' ", 'wrzasną sał', 'powt	olishtypes - 1 gadnął', 'wy , 'przyznał', wysapał', 'z ął'} arzał', 'mru	Microsoft Ex sapal', 'zaj 'zwariowa :aperzył', 'u knął', 'wyji	oel użytek niek perzył', 'uciął' ł', 'rzekł', 'zap uciął', 'zaklika ąkał', 'mrucza	omercyjny , 'zaklikal' pewnił', 'z TT, f', 'ciąg ił'}	, 'ciągnął', awołał', 'p D _types_Va 114 4 6 9	, 'wypalił', 'n comrukiwał E or_1_freq 947 23 23 23 20	wykrzykną ', 'odrzekł' F	', 'powital 'wyznař, G	', 'uciszył', 'stracił', 'o H	'skarcil', 'zr burzył', 'do I	∂ ⊂ alazť, dať, J	
C2 A source_E said whispere shouted muttered asked	rzędzia główne ▼ (B IN_source_ ed	Vistawianie Układ str. // ('wskazz 'zdał', 'p 947 ('wskazał', 'wzyu 23 ('szepnał', 'wzyu 23 ('krzyknął', 'rykn 20 ('warknął', 'wym 19 ('zwrócił', 'zapyt	ony Formuły ał', 'wzywał', 'w przedstawił', 'w wał', 'odpart', ' eptał', 'oznajn iął', 'zawolał', 'beł ał'}	Dane odparť, h vpadť, 'z ucieszyť, niť, 'zapy szydziť, ' kotať, 'ję	PL, Recenzja W ucieszył', 'p denerwow ,'polecił', 'z tał') 'powiedział ,knął', 'napi	1_HP_CS_Pc //dok olecił', 'zag ał', 'chciał', C agadnął', ' ', 'wrzasną sał', 'powt	olishtypes - 1 gadnął', 'wy 'przyznał', wysapał', 'z ł'} arzał', 'mru	Microsoft Ex sapať, 'zaj 'zwariowa aperzyť, 'u knąť, 'wyji	cel użytek niek perzyłt, 'uciął' ł', 'rzekł', 'zap uciął', 'zaklika ąkał', 'mrucza	omercyjny , 'zaklikal' pewnił', 'z TT, F, 'ciąg ił'}	, 'ciągnął', awołał', 'p D types Va 114 4 6 9 2	, 'wypalii', 'n comrukiwał E 947 23 23 20 19	vykrzykną ', 'odrzekł' F	', 'powital 'wyznał', G	', 'uciszył', 'stracił', 'o H	'skarcił', 'zr burzył', 'do I	alazi', J	
C2 A source_E said whispere shouted muttered asked yelled	rzędzia główne (B SN_source_ ed	Wstawianie Układ str. // ('wskaz' 'zdał', 'p EN_cr TT_types 947 ('wskazał', 'wzyw 23 ('szepnął', 'wzys 23 ('krzyknął', 'rykn 20 ('warknął', 'rykn 19 ('zwróckł', 'zapyt 16 ('zawył', 'krzykn.	ony Formuly al', 'wzywał', 'v przedstawił', 'w wał', 'odpart', ' eptał', 'oznajn rął', 'zawołał', ' amrotał', 'beł ał'} ął', 'zapiał', 'ry	Dane odparť, h vpadť, 'z ucieszyť, niť, 'zapy szydziť, ' kotať, 'ję	PL, Recenzja W ucieszył', 'p denerwow , 'polecił', 'z rtał') 'powiedział knął', 'napi wołał', 'ode	1_HP_CS_Pc Rdok olecił', 'zag ał', 'chciał', C agadnął', 'i gadnął', 'y r, 'wrzasną sał', 'powt	olishtypes - 1 gadnął', 'wy 'przyznał', wysapał', 'z arzał', 'mru zał', 'wrzasi	Microsoft Ex sapal', 'zaj 'zwariowa aperzył', 'u knął', 'wyji nął')	cel użytek niek perzył', 'uciął' ł', 'rzekł', 'zap uciął', 'zaklika ąkał', 'mrucza	omercyjny , 'zaklikał' pewnił', 'z TT, ł', 'ciąg ł''}	, 'ciągnął', awota'', 'p D types Va 114 4 6 9 2 8	, 'wypalil', 'n comrukiwał er 1 freg 947 23 23 20 19 16	wykrzykną ', 'odrzekł' F	', 'powita' 'wyznał', G	', 'uciszyl', 'stracif', 'o H	'skarcit', 'zr burzył', 'do I	alazi', J	
C2 A source_E said whispere shouted muttered asked yelled added	rzędzia główne	Wistawianie Układ str. // ('wskaza', 'wskaza', 'wzyw 947 ('wskaza', 'wzyw 'zdał', 'g 23 ('krzykara,', 'rykn 'zdał', 'rykn 20 ('warknąt', 'wym 'gwróch', 'zapyt', 'krzykna' 16 ('zawyt', 'krzykna' 'strzykna'	ony Formuly af', 'wzywał', 'd przedstawił', 'v wał', 'odparł', ' eptał', 'oznajn ramrotał', 'beł ał'} ał', 'zapiał', 'ry	Dane odparť, 'v vpadť, 'z ucleszyť, niť, 'zapy szydziť, ' kotať, 'ję knąť, 'za	PL, Recenzja W ucieszył', 'p denerwow 'polecił', 'z tał'} 'powiedział knął', 'napi wołał', 'ode	1_HP_CS_Pc Ridok olecił', 'zag ał', 'chciał', C agadnął', ' ', 'wrzasną sał', 'powt szwał', 'rycz	olishtypes - 1 gadnął', 'wy ,'przyznał', wysapał', 'z ł'} arzał', 'mru zał', 'wrzasi	Microsoft Ex sapal', 'zaj 'zwariowa aperzył', 'u knął', 'wyji nął')	cel użytek niek perzył', 'uciął' ł', 'rzekł', 'zap uciął', 'zaklika ąkał', 'mrucza	omercyjny , 'zaklikal' pewnił', 'z TT, f', 'cląg	, 'ciągnął', awotał', 'p D types Va 114 4 6 9 2 8 1	, 'wypalil', 'v comrukiwał er_1 freg 947 23 20 19 16 15	wykrzykną ', 'odrzekł' F	', 'powital 'wyznał', G	', 'uciszyl', 'stracif', 'o H	'skarcił', 'zr burzył', 'do I	alazi', J	
C2 A source_E said whispere shouted muttered asked yelled added told	B N_source_ ed	Votawianie Układ stru '(wskaza', 'ky stadat', 'p 947 ('wskaza', 'wzyw 23 ('szenar)', 'wzyw 23 ('krzyknaf', 'wzyw 23 ('krzyknaf', 'rykn 20 ('warknaf', 'yznyt) 16 ('zawyf', 'krzykna 15 ('dodaf') 15 ('donarjmif', 'szep	ony Formuly al', 'wzywał', 'd przedstawił', 'w wał', 'odparł', ' eptał', 'oznajn iął', 'zawolał', ' namrotał', 'beł ał'} ał', 'zapiał', 'ry pnął', 'wyjaśnił	Dane odparť, 'v vpadť, 'z ucieszyť, hiť, 'zapy szydziť, ' kotať, 'ję knąť, 'za	PL Recenzja W ucieszył', 'p denerwow 'polecił', 'z tał') 'powiedział knął', 'napi wołał', 'ode adał', 'stwi	1_HP_CS_Pc Fidok olecił', 'zag ał', 'chciał', C agadnął', 'v ', 'wrzasną sał', 'powt zwał', 'ryci ardził', 'poi	olishtypes - 1 gadnął', 'wy 'przyznał', wysapał', 'z ął'} arzał', 'mru zał', 'wrzasz wiedział'}	Microsoft Ex Isapal', 'zaj 'zwariowa isaperzył', 'u knął', 'wyji nął')	cel uzytek nieko perzył', 'uciął' ł', 'rzekł', 'zap uciął', 'zaklika gkał', 'mrucza	omercyjny , 'zaklikal' pewnił', 'z TT, F, 'ciąg	, 'ciągnąt', awołał', 'p D types Va 114 4 6 9 2 8 1 6	, 'wypalił', 'n comrukiwał er_1_freg 947 23 23 20 19 16 15 15	vykrzykną ', 'odrzekt' F	', 'powital 'wyznał', G	', 'uciszył', 'stracił', 'o H	'skarcił', 'zr burzył', 'do I	alazi', J	
C2 A source_E said whispere shouted muttered asked yelled added told hissed	Reference in the second s	Wstawianie Układ str zdałł, je PAT (Wskazz 2017, je PAT (Wskazzł, Wszy 23 (krzykagł, wszy 23 (krzykagł, wsz 23 (krzykagł, wsz 20 (krzykagł, wsz 20 (krzykagł, wsz 19 (zwródł, zapył) 16 (zawył, krzyka 15 (dodałł) 15 (dodałł) (krzyka	ony Formuly af', 'wzywał', 'o rrzedstawił', 'w wał', 'odpart', ' eptał', 'ozwajn iamrotał', 'ozwajn amrotał', 'beł ał'} ał', 'zapiał', 'ry pnął', 'wyjaśnił iął')	Dane odparť, 'u vpadť, 'z ucieszyť, hiť, 'zapy szydziť, ' kotať, 'ję knąť, 'za	PL Recenzja W ucieszył', 'p denerwow , 'polecił', 'z tał') 'powiedział , 'napi wołał', 'napi adał', 'stwi	1_HP_CS_Pc Ridok olecił', 'zag ał', 'chciał', C agadnął', 'r ', 'wrzasną sał', 'powt izwał', 'ryc: erdził', 'por	olishtypes - 1 gadnąť, 'wy , 'przyznať, wysapať, 'z łť'} arzať, 'mru zať, 'wrzase wiedziať}	Microsoft Ex sapał', 'zaj 'zwariowa 'aperzył', 'u knął', 'wyji nął')	cel utytek nieko perzył', 'uciął' ł', 'rzekł', 'rzekł', 'zap uciął', 'raklika ąkał', 'mrucza	omeroyjny , 'zaklikał' pewnił', 'z TT, ł', 'ciąg	, 'ciągnąt', awołat', 'p D types_Va 114 4 6 9 2 8 1 2 8 1 6 2	, 'wypalii', 'i eomruiwał er_1_freq 947 23 23 20 19 16 15 15 11	vykrzykną ', 'odrzeki' F	', 'powital 'wyznał', G	', 'uciszył', 'stracił', 'o H	'skarcit', 'zr burzył', 'do I	o 🕜 📼 alazť, dať, J	
C2 A source_E said whispere shouted muttered asked yelled added told hissed roared	B N_source_ ed	Wittenkinie Układ stru Jr. (wskazał vadał, r y N. of TL, types 947 (wskazał, 'wzyw 947 (wskazał, 'wzyw 23 (%zrapał, 'wzyw 23 (%zrapał, 'wzyw 23 (%zrapał, 'wzyw 19 (żwródił, 'zapyt 16 (żawył, 'krzyknał, 'krzyknał, 'krzyknał, 'krzyknał, 'krzyknał, 'krzyknał, 'krzyknał,' krzyknał, 'krzyknał,' krzyknał,' krzyknał, 'krzyknał,' krzyknał,'	ony Formuly af', 'wzywał', 'o rzedstawił', 'w vał', 'odpart', ' eptał', 'oznajn iął', 'zawolał', 'namrotał', 'beł ał'} ał', 'zapiał', 'ry onął', 'wyjaśnił iął') cał, 'krzyknął',	Dane odparť, 'u vpadť, 'z ucieszyť, hiť, 'zapy szydziť, ' kotať, 'ję knąť, 'za ', 'opowi 'ryknąť}	PL Recenzja W ucieszył', 'p denerwow , 'polecił', 'z tał') 'powiedział knął', 'napi wołał', 'ode adał', 'stwie	1_HP_CS_Pc Rdok olecił', 'zaga ał', 'chciał', C agadnął', 'i ', 'wrzasną sał', 'powt izwał', 'rycz erdził', 'pou	olishtypes - 1 gadnął", 'wy ('przyznał', wysapał', 'z ł(') arzał', 'mru zał', 'wrzase wiedział'}	Microsoft Ex sapał', 'zaj 'zwariowa :aperzył', 'u knął', 'wyji nął')	cel uzytek nieko serzył', 'uciął' ł', 'rzekł', 'zap uciął', 'zaklika ąkał', 'mrucza	omeroyjny , 'zaklikał' pewnił', 'z TT, f', 'ciąg	, 'ciągnął', awotał', 'p D types, Va 114 4 6 9 2 8 1 4 6 2 4	, 'wypalif', 'n komrukiwał er_1_freq 947 23 20 19 16 15 15 11 10	vykrzykną ', 'odrzekł' F	', 'powital 'wyznał', G	', 'uciszył', 'stracił', 'o H	'skarcil', 'zr burzył', 'do I	o oo o	
C2 A source_E said whispere shouted muttered asked yelled added told hissed roared called	B B Source_ ed	Witsonkanie Uikład strużych Jak (Wokka, ż żdał ?, N., cr TL, types 947 (Wokkazał, Wrzyw 23 (Stezepanł, Wrzyw 23 (Stezepanł, Wrzyka 10 (Warzkanł, Ywym 19 (zwych, Krzykan 15 (dodał) 16 (zwych, Krzykan 15 (dodał) 11 (syknął / krzyka 10 (Wrzykanł, Ykazyka) 10 (Wrzykanł, Ykazyka) 12 (Stezasnął / wym)	ony Formuly af, 'wzywat', 'k rzzedstawit', 'w eptat', 'odpart', ' eptat', 'oznajn af', 'zawolat', ' hamrotat', 'bet af' af', 'zapiat', 'ry ungt', 'wyjaśnił af' tat', 'krzyknąt', 'krzyknąt', ' kawolat'	Dane odparť, 'n vpadť, 'z ucieszyť, iř, 'zapy szydziť, ' kotať, 'ję knąť, 'zar ', 'opowi 'ryknąť}	PL Recenzja W ucieszył', 'p denerwow , 'polecił', 'z tał'} 'powiedział knął', 'napi wołał', 'ode adał', 'stwie	1_HP_CS_Pc Rdok olecił', 'zag ał', 'chciał', C agadnął', 'i ', 'wrzasną sał', 'powt szwał', 'ryc: erdził', 'por	olishtypes - 1 gadnął', 'wy , 'przyznał', wysapał', 'z it'} arzał', 'mru zał', 'wrzasł wiedział'}	Microsoft Ex sapal', 'zaj 'zwariowa :aperzył', 'u knął', 'wyj: nął')	cel uzytek nieko serzył", 'uciął" ł', 'rzekł", 'zap uciął", 'zaklika ąkał", 'mrucza	omercyjny , 'zaklikał' pewnił', 'z TT, r', 'ciąg	, 'ciągnąt', awotał', 'p D types, Va 4 6 9 2 8 1 6 2 4 3	, 'wypalil', 'n somrukiwał er_1_freq 947 23 20 19 16 15 15 11 10 0	vykrzykną ', 'odrzekł' F	', 'powital 'wyznał', G	', 'uciszył', 'stracił', 'o H	'skarcil', 'zr burzył', 'do I	₽ Palazť, dať, J	
C2 source_E said whispere shouted muttered asked yelled added told hissed roared called snapped	Kanala and a second secon	Wittenkinie Układ strużenie J.K. (Wickaz: 'zdał', rp. N. of TL, types 947 (Wickazał', Wzyu 23 (Iszepraf, Wzyz 23 (Iszepraf, Wzyz 20 (Warknaf, 'zapyt 16 (izawyt', krzykni 15 (idoałr) 16 (izawyt', krzykni 15 (idoałr) 15 (idoarr) 11 (isyknaf, 'krzykni 10 (Wirzsanf, 'krzykni 10 (Wirzsanf, 'pryc	ony Formuly aff, 'vzywaf', 'v vzrzedstawif', 'v eptaf', 'ocnajn aff, 'canajn aff, 'zavolaf', 's aff, 'zapiaf', 'ry aff, 'zapiaf', 'ry jaf, 'zapiaf', 'ry jaf, 'zapiaf', 'ry jaf, 'zapiaf', 'ry jaf, 'zapiaf', 's avolaf', hn aff)	Dane odparť, 'r vpadť, 'z ucieszyť, iť, 'zapy szydziť, ' kotať, 'ję kotať, 'ję knąť, 'za ', 'opowi 'ryknąť'}	PL, Recenzja W ucieszył', 'p denerwow 'polecił', 'z tał'} powiedział knął', 'napi wołał', 'ode adał', 'stwie	1_HP_CS_Pc Adok olecil', 'zag al', 'chcial', C agadnął', 'r r, 'wrzasną sał', 'powt szwał', 'ryc: erdził', 'pow	olishtypes - 1 gadnął', 'wy 'przyznał', wysapał', 'z ił'} arzał', 'mru zał', 'wrzase wiedział'}	Microsoft Ex sapal', 'zaj 'zwariowa aperzył', 'u knął', 'wyj: nął')	cel uzytek nieko perzyłł, 'uciął' ł', 'rzekł', 'zap uciął', 'zaklika ąkał', 'mrucza	omercyjny , 'zaklikal' pewnił', 'z TT, F, 'ciąg	, 'ciągnąt', 'p D types, Va 114 4 6 9 2 8 1 6 2 8 1 6 2 4 3 2 2	,'wypalif','i komrukiwał E 947 23 20 19 16 15 15 11 10 9 9	wykrzykną ', 'odrzekł' F	', 'powitał 'wyznał', G	', 'uciszyl', 'stracił', 'o H	'skarcił', 'zr burzył', 'do I	alazť, dať, J	
C2 source_E said whispere shouted muttered asked yelled added told hissed roared called snapped gasped	K B K Source	Wissendarie Uiklad 3th John (Verkaz, z zdať, ?) Varkaz, ? (verkaz, z zdať, ?) Stependt, Vorkaz, ?) verkaz, ?) 23 (krzyknaž, *) verkaz, ?) 20 (warknaž, *) verkaz, ?) 20 (krzyknaž, *) verkaz, ?)	ony Formuly af, 'vzywaf', 'v vař, 'odparf, ' eptař, 'odparf, ' eptař, 'canajn ař} ' zawolař, 'bel ař ař zař, 'zawolař tař, 'zawolař hnař ', 'zaszlochař, ' zaszlochař,	Dane odparť, 'u vpadť, 'z ucieszyť, niť, 'zay kotať, 'ję knąť, 'zav ', 'opowi 'ryknąť}	PL, Recenzja W ucieszył', 'p denerwow , 'polecił', 'z tał') 'powiedział knął', 'napi wołał', 'ode adał', 'stwie ił', 'jęknął',	1_HP_CS_Pc Rdok olecił', 'zag ał', 'chciał', C agadnął', '' r, 'wrzasną sał', 'powt izwał', 'ryc: erdził', 'pow	olishtypes - 1 gadnął', 'wy 'przyznał', wysapał', 'z ł(') arzał', 'mru zał', 'wrzasz wiedział'}	Microsoft Ex sapal', 'zaj 'zwariowa :aperzył', 'u knął', 'wyji nął')	cel uzytek nieko perzyłł, 'uciął' ł', 'rzekł', 'zap rciął', 'zaklika ąkał', 'mrucza	omercyjny , 'zaklikal' pewnił', 'z TT, ř, 'ciąg	, 'ciągnąt', 'p D types, Va 114 4 6 9 2 8 1 1 6 2 8 1 6 2 4 3 2 6	, 'wypalił', 'n komrukiwał rfreq 947 23 23 20 19 16 15 15 11 10 10 9 9 9	vykrzykną , 'odrzekt' F	', 'powital 'wyznał', G	', 'uciszył', 'stracił', 'o H	'skarcil', 'zr burzył', 'do I	alazť, J	
C2 A source_E said whispere shouted muttered asked yelled added told hissed roared called snapped gasped gasped	Kanala and a second and a	Witsenderic Uilded 199 / (Viskazař (Viskazař / todař / (Viskazař (Viskazař 947 (Viskazař, Vizty) 23 (Stapenař, Visza 947 (Viskazař, Vizty) 23 (Stapenař, Visza 91 (Viskazař, Vizty) 19 (Stavráční, Vizty) 19 (Stavráční, Vizty) 16 (Stavrý, Krzhvař, Vizty) 15 (dodař) 15 (dodař) 10 (Viztsaněř, Vizty) 9 (Viztnař, Vizty) 9 (Viztnař, Vizty) 7 (Viztnař, Vizty) 9 (Viztnař, Vizty) 7 (Viztnař, Vizty)	ony Formuly aff, 'vzywaf', 'v vzrzedstawif', 'v eptaf', 'oznajn emrotaf', 'zanajn aff', 'zapiaf', 'ty aff', 'zapiaf', 'ty aff', 'zapiaf', 'ty aff', 'zapiaf', 'ty anaff', 'wyjaśnii naff') ', 'zasiochaf', ', 'zasiochaf', ', 'zasiochaf',	Dane odparť, 'u upadť, 'z ucieszyť, hiť, 'zapy szydziť, ' kotať, 'ję kotať, 'ję kotať, 'zar ', 'opowi 'ryknąť' ' 'wydysza	PL Recenzja W ucieszył', 'p denerwow 'polecił', 'z tał'} 'powiedział knął', 'napi wołał', 'ode adał', 'stwie tł', 'jęknął', '	1_HP_CS_Pc Fick	olishtypes - 1 gadnął', 'wy 'przyznał', wysapał', 'z łt'} arzał', 'mru zał', 'wrzase wiedział'}	Microsoft Ex 'sapal', 'zaj 'zwariowa :aperzył', 'u knął', 'wyji nął'}	cel uzytek nieko perzył", 'uciął" I', 'rzeki", 'zap uciął", 'zaklika ąkał", 'mrucza	omercyjny , 'zaklikat' rewnit', 'z TT, r, 'ciąg	, 'ciągnąt', 'p D types, Va 114 4 6 9 2 8 1 6 2 8 1 6 2 4 3 2 2 4 3 2 2 6 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	, 'wypalii', 'i eomrukiwał E yr_1_freq 947 23 23 20 19 16 15 15 15 15 11 10 10 9 9 7	wykrzykną , 'odrzekł' F	^r , 'powital 'wyznał', G	', 'uciszył', 'stracił', 'o H	'skarcił', 'zr burzył', 'do I	alazť, dať, J	
C2 A source_E said whispere shouted mutterec asked yelled added told hissed roared called snapped gasped snarled growled	Kenter Contraction of the second	Witzension Uikid str Image: Comparison of the strength of the stre	ory Formuly aff, 'vzywaf', 'c rzzedstawif', 'v waf', 'odparf', ' potaf', 'oznajn aff', 'zawolaf', hamrotaf', 'beł aff aff', 'zawolaf', hnął? ', 'zaszlochaf', nazł', warczał	Dane odparť, 'u vpadť, 'z ucieszyť, iť, 'zapy szydziť, ' kotať, 'ję knąť, 'zar ', 'opowi 'ryknąť} 'wydysza ', 'ryknąť	PL, Recenzja W ucieszył', 'p denerwow , 'polecił', 'z tał') 'powiedział knął', 'napi wołał', 'ode adał', 'stwie ił', 'jęknął', '	1_HP_CS_Pc Rock olecit', 'zaga at', 'chciat', C agadnął', 'i (, 'wrzasną sał', 'powt szwał', 'rycz erdził', 'pow 'wyjąkał'} odpowiedz	olishtypes - 1 gadnął', 'wy 'przyznał', wysapał', 'z tł'} arzał', 'mru zał', 'wrzase wiedział'}	Microsoft Ex rsapal', 'zaj 'zwariowa aperzył', 'u knął', 'wyj: nął'}	cel uzytek nieko serzył', 'uciął' t', 'rzekł', 'zap uciął', 'zaklika ąkał', 'mrucza	omercyjny , 'zaklikal' sewnił', 'z TT, ř', 'ciąg	, 'ciągnąt', awotať', 'p Lypes, Va 114 4 6 9 2 2 8 1 6 2 4 3 2 2 6 2 4 3 2 6 2 6 2 6	, 'wypalit', 'i comrukiwat emrukiwat 947 23 20 19 16 15 11 10 10 9 9 7 7	vykrzykną ', 'odrzekt' F	^r , 'powital 'wyznał', G	', 'uciszył', 'stracił', 'o H	'skarcit', 'zr burzyt', 'do I	alazť, dať, J	
C2 A source_E said whispere shouted added told hissed called snapped gasped snarled growled went on	Compared a givene Compared a givene Compared a givene Compared a givene	Witzendarie Uilda din Jr (wskaz tadar); Versitz (wskaz tadar); 947 ('wskazaf, 'wzyu 23 ('szepanj,' wzyu 23 ('szepanj,' wzyu 16 ('zawy); 'krzyt 16 ('zawy); 'krzyt 15 ('doadar); 15 ('doadar); ('arananj,''yryt 10 ('krzytanj,' zapyt 10 ('krzytanj,' zapyt 9 ('warkanj,' put); 9 ('warkanj,''put); ('warkanj,''put); 7 ('warkanj,''put); ('krzytanj,''put);	ony Formuly alf, 'wzywal', 'a vzrzedstawił', 'w vał', 'odparf', ' petaf', 'oznaja', 'zawolaf', 'sawolaf', 'zawolaf', 'sawolaf', ' am'otaf', 'sawolaf', 'b af', 'zapiaf', 'r ang', 'wysśnił 'af', 'szystochaf', ', 'zaszlochaf', 'n '	Dane odparť, 'u vpadť, 'z ucieszyť, iiť, 'zapy szydziť, ' kotať, 'ję knąť, 'za ', 'opowi 'ryknąť} 'wydysza f, 'ryknąť	PL, Recenzja W uccieszył', 'p denerwow 'polecił', 'z tał'} 'powiedział knąť', 'napi wołał', 'ode adał', 'stwie iť, 'jęknął', '	1_HP_CS_Pc Fook fook olecil', 'zaga at', 'chcial', C agadnął', ' r, 'wrzasną sał', 'powt sał', 'powt erdzil', 'pon 'wyjąkał'} odpowiedz	olishtypes - 1 gadnął', 'wyy , 'przyznał', wysapał', 'z tł'} arzał', 'wrzase wiedział'} ział'}	Microsoft Ex sapał', 'zaj 'zwariowa aperzył', 'u knął', 'wyj: nął')	cel uzytek nieko perzyłł, 'uciął' ł', 'rzekł', 'zap uciął', 'zaklika ąkał', 'mrucza	omercyjny , 'zaklikal' pewnił', 'z TT, f', 'ciąg	, 'ciągnął', 'p D U types Va 114 6 9 2 8 1 6 2 4 3 2 6 2 6 2 6 2 6 2 6 2 2 6 2 2	,'wypalif', 'y komrukiwat 947 23 20 19 166 15 15 11 10 10 9 9 7 7 6	vykrzykną ', 'odrzeki' F	^r , 'powital 'wyznał', G	', 'uciszył', 'stracił', 'o H	'skarcit', 'zr burzył', 'do I	₽ alazť, dať, J	
C2 A source_E said whispere shouted muttered asked yelled added told hissed roared called snapped snapped snapred growled growled went on repeated	Revenue of the second sec	Witzendorf Uilded str Image: Comparison of the strength of the str	ony Formuly aff, 'wzywaf', 'c rzedstawif', 'w waff, 'conajn qf', 'zawolaf', ' namrotaf', 'bel aff) qf', 'zawolaf', ' naŋf', 'wyjaśnii qf' naŋf', 'wyjaśnii qf' ', 'zaszlochaf', '', 'zaszlochaf', naŋf')	Dane Jodparl', 'i upadi', 'z szydził', 'z szydził', 'z knął', 'zał ', 'opowi 'ryknął', 'z wydysza '', 'ryknął'	PL, Recenzja W uucieszył', 'p denerwow 'polecił', 'z tał'} 'powiedział knął', 'napi wołał', 'ode adał', 'stwie ił', 'jęknął', '	1_HP_CS_Pc Rook olecit', 'zag at', 'chciat', C agadnął', 'r ', 'wrzasną sał', 'powt zzwał', 'rycz erdzit', 'pow 'wyjąkał'} odpowiedz	olishtypes - 1 gadnął', 'wyy 'przyznał', 'z wysapał', 'z łt'} arzał', 'wrzasz wiedział'} ział'}	Microsoft Ex sapał', 'zaj 'zwariowa aperzył', 'c knął', 'wyji nął')	cel uzytek nieko perzył", 'uciął" t', 'rzekł", 'zap uciął", 'zaklika ąkał", 'mrucza	omercyjny , 'zaklikal' sewnit', 'z TT, ?, 'ciąg	, 'ciągnąt', awota', 'p D types, Va 114 4 6 9 2 8 1 6 2 8 1 6 2 4 3 2 2 6 2 2 6 2 2 1	,'wypalif', 'somrukiwat E 947 23 20 19 16 15 15 15 11 10 10 9 9 7 7 7 6 5	vykrzykną ', 'odrzeki' F	^r , 'powitał 'wyznał', G	', 'uciszył', 'stracił', 'o H	'skarcit', 'zr burzyt', 'do I	alazi', dai', J	
C2 A source_E said whispere shouted muttered asked yelled added told hissed roared called snapped gasped gasped growled went on repeated croake4	Control of the second sec	Witsendorf Uilde str Image: Comparison of the strength of the stre	ony Formuly alf, 'wzywał', 'w rzzedstawił', 'w eptaf, 'ochparf,' eptaf, 'ochparf,' eptaf, 'zawolaf, tar, 'zawolaf,' tał', 'zawolaf,' tał', 'zawolaf', tał', 'zawolaf', hnąf' , 'zaszlochaf, naf' ; szaszlochaf, naf' ; szaszlochaf, iczał', 'warczał ;)	Dane Dodparl', 'i vypadł', 'z szydził', 'rapy szydził', 'rapy szydził', 'rapy 'opowi 'ryknął' ' 'wydysza ', 'ryknął'	PL, Recenzja W uucieszył', 'p denerwow , 'polecił', 'z tał') 'powiedział knął', 'rapi wołał', 'ode adał', 'stwie ił', 'jęknął', '	1_HP_CS_Pc fick olecit', 'zaga 4', 'chciat', C agadnął', 'n c agadnął', 'n agadn agadnął', 'n agadnął', 'n agadnał', 'n agadnął', 'n agadnął', 'n agadnął', 'n agadn	olishtypes - 1 gadnął', 'wy 'przyznał', wysapał', 'z ił'} arzał', 'mru zał', 'wrzase wiedział'}	Microsoft Ex sapał', 'zaj 'zwariowa aperzył', 'u knął', 'wyja nął'}	cel uzytek nieko serzyłł, 'uciął' ł', 'rzekł', 'zap uciął', 'zaklika ąkał', 'mrucza	omercyjny , 'zaklika' TT, TT, 'cląg	, 'ciągnąt', awołat', 'p D Lypes Va 114 4 6 9 2 8 1 4 6 2 4 3 2 6 2 4 3 2 6 2 6 2 1 1	,'wypalif', 'v somrukiwat E rr_1_freq 947 23 20 19 16 15 15 11 10 10 9 9 9 7 7 6 5 5	vykrzykną ', 'odrzeki' F	', 'powital 'wyznař, G	', 'uciszył', 'stracił', 'o H	'skarcit', 'zr burzyt', 'do I	alazi', Ja'',	
C2 A source_E said whispere shouted muttered asked yelled added told hissed roared called snapped gasped snapped growled went on repeated croaked	Compared as glownee Compared as glownee Compared as glownee Compared as glownee	Wissendard Utility S Crosslaw Crosslaw Crosslaw Crosslaw Sec TT_types Crosslaw 947 Crosslaw 23 Starepart, Vrsyz 23 Starepart, Vrsyz 23 Starepart, Vrsyz 23 Starepart, Vrsyz 15 Crosslaw, Yangy, Karykon 15 Crosslaw, Yangy, Karykon 15 Crosslaw, Yangy, Karykon 10 Krazykand, Yangy 9 Wartoard, Yangy 9 Wartoard, Yangy 9 Wartoard, Yangy 9 Wartoard, Yangy 9 Vartoard, Yangy 9 Stardoard, Yangy 9 Stardoard, Yangy 9 Stardoard, Yangy	ony Formuly af', 'wzywał', 'a vzrzedstawił', 'w eptaf', 'odparf', eptaf', 'odparf', eptaf', 'odparf', 'at', 'awolaf', at', 'awolaf', 'at', 'awolaf', 'at', 'awolaf', 'at', 'kzzyknąf', 'saziochaf',	Dane Dodpari', 'i ucieszyf', 'zapy kotał', 'zapy kotał', 'ję kotał', 'ję wodysza ', 'ryknął' '	PL Recenzja W uccieszył', 'p denerwow , 'polecił', 'z tał') powiedział knął', 'napi wołał', 'ode adał', 'stwi if, 'jęknął', 'i	1_HP_CS_Pc Rook olecii', 'zaga 4', 'chciai', C agadnął', 'n c agadnął', 'n c zwał', 'powt zwał', 'rycz ardził', 'pow wyjąkał') odpowiedz	olishtypes - I gadnął', 'wy 'przyznał', wysapał', 'z łt'} arzał', 'mru zał', 'wrzase wiedział'}	Microsoft Ex sapať, 'zaj 'zwariowa :aperzyť, 'c knąť, 'wyj: nąť'}	cel uzytek nieko serzyłł, 'uciął' ł', 'rzekł', 'zap uciął', 'zaklika uciął', 'mrucza	, 'zaklika' sewnii', 'z F, 'ciąg	, 'ciagnal', 'p D D types_Va 114 4 6 9 9 2 8 8 1 1 6 6 2 2 4 3 2 6 6 2 2 1 1 1 2 2 2 3 2 2 6 2 2 1 1 2 2 2 3 2 2 3 2 2 3 2 2 3 2 2 3 2 3	, 'wypalit', 'i e r_1 freq 947 23 23 20 19 16 15 11 10 10 9 9 9 7 7 6 5 5 6	vykrzykną , 'odrzekt' F	',' powital 'wyznał', G	', 'uciszyl', 'stracił', 'o H	'skarcit', 'zr burzyt', 'do I	2 2 alazt', Jai alazt', J	
C2 A source_E said whispere shouted muttered asked yelled added told hissed roared called snapped gasped gasped growled went on repeated croaked breathed coalled	Constant of the second se	Witsendorf Uikled str Image: Comparison of the strength of the str	ony Formuly af', 'wzywał', 'w vrzedstawił', 'w eptaf', 'odparf', ' eptaf', 'cawolaf', tamrotaf', 'beł af} af', 'zawolaf', ' tar, 'krzyknąf', ' 'zawzolaf', hnąf') ' zawzolaf', 'r zawolaf', tar, 'krzyknąf', 'zawzolaf', hnąf') pnąf', 'odpowi uścarfi.	Dane odpari', 'i vpadi', 'z ucieszył', 'z ucieszył', 'z szydził', 'g szydził', 'g wydysza 'ryknął' 'wydysza 'ryknął' wydysza catalogo (stato) 'ryknął'	PL, Recenzja W ucceszył', 'p denerwow , 'polecił', 'z tał'} powiedział "knął', 'napi wołał', 'ode adał', 'stwie adał', 'stwie t', 'jęknął', '	1_HP_CS_PC Rook olecit', 'zaga 4', 'chciat', C agadnął', 'r c agadnął', 'r wrzasną sał', 'powt i', 'wrzasną sał', 'powt wyjąkał') odpowiedz	olishtypes - 1 ;adnął', 'wy ;'przyznał', wysapał', 'z wysapał', 'mru zał', 'wrzasz wiedział'}	Microsoft Ex sapal', 'zaj 'zwariowa :aperzył', 'u knął', 'wyji nął')	cel użytek nieko serzyłł, 'uciął' ł', 'rzekł', 'zap uciął', 'zaklika uciął', 'mrucza	vomercyjny , 'zaklikal' TT, F, 'ciąg	, 'ciągnął', ' D D Lypes Va 114 4 6 7 2 8 8 8 114 4 6 6 2 2 4 3 2 6 6 2 2 6 6 2 2 1 1 1 3 3 2 2 6 6 2 2 1 1 1 2 2 3 2 2 6 6 2 2 3 2 2 3 2 2 3 2 2 3 3 2 2 3 3 3 2 2 3	,'wypalit', '\ bomrukiwat F r_1_freq 947 233 200 199 166 155 151 111 100 100 9 9 9 7 7 6 5 5 5 5	vykrzykną , 'odrzekł	', 'powital 'wyznał', G	', 'uciszyf', stracif, 'o H	'skarcit', 'zr burzyt', 'do I) alazi', j	
C2 A A said whispere shouted whispere shouted added told hissed gasped growled growled growled growled breathed breathed breathed breathed breathed	receive glowne v (B N_source_ ed d	Wissendard Utidat sn C Cysakaz R. of TL_types Cysakaz 947 Cysakaz 923 Sizepanji, 'viyaz 23 Sizepanji, 'viyaz 10 Varafranji, 'knym 15 Groadnanji, 'save 16 (zavarji, 'krazkna', 'krazkna', 'krazkna', 'krazkna', 'krazkna', 'krazkna', 'krazkna', 'Save', 'krazkna', 'bark', 'bark', 'krazkna', 'bark', 'bark', 'bark', 'krazkna', 'bark'	ony Formuly af', 'wzywał', 'a vzrzedstawił', 'w vał', 'odpart', ' eptaf', 'oznajn amrotał', 'beł af' af', 'zawolat', 'beł af' af', 'zawolat', 'beł af' af', 'zawolat', 'beł af', 'zawolat', 'beł af', 'zawolat', 'beł af', 'zawolat', 'beł awolat', 'beł warczał', 'warczał) pagł', 'odpowi jaśnił')	Dane Dodpari', 'n vpadł', 'z szydził', 'z kotał', 'g kotał', 'g kotał', 'g kotał', 'g wydysza ', 'nyknął' 'wydysza edział')	PL Recenzja W ucieszył', 'p denerwow 'polecił', 'z tał') 'powiedział knąť', 'napi wołał', 'ode adał', 'stwie ił', 'jęknął', '	1, HP_CS_PC Robe Robe robecit', 'zag al', 'chciat', C agadnąt', '' ", 'wrzasną sat', 'powt vzwat', 'rycz erdzit', 'pow wyjąkat'} odpowiedz	olishtypes - I şadnqi', 'wy 'przyznai', wysapai', 'z wysapai', 'y arzai', 'mru zai', 'wrzasi wiedziai'}	Microsoft Ex sapat', 'zayariowa zwariowa aperayf', 'u knqf', 'wyji	cel uzytek nieko perzyłł, 'uciął' ł', 'rzekł', 'zap uciął', 'zaklika gkał', 'mrucza	comercyiny , 'zaklika' pewnił', 'z rrr, 'ciąg ł', 'ciąg kił')	,'ciagnal','p D D 114 4 6 9 9 2 8 11 6 6 2 8 1 1 6 6 2 2 6 6 2 1 1 1 3 3 2 2	vwypalił', 'v eomrukiwał r_1_freq 947 23 23 20 19 16 15 15 15 15 15 15 15 15 15 5 5 5 5 5	vykrzykną F	', 'powital 'wyznał', G	', 'uciszyf', 'u stracif', 'o H	'skarcit', 'zr burzył', 'do I	2 2 alazi', Jailazi', J	
C2 A A said whispere shouted mutterec asked yelled added told hissed called snapped gasped gasped snapped gasped called called snapped gasped conception c	v (B Vone ed d d d d d d d d d d d d d d d d d	Witsendorf Uikled str Image: Constraint of the strength of the str	ony Formuly aff, 'wzywat', 'k zrzedstawit', 'v waf', 'odpart', ' eptaf', 'oznajn grt, 'zawolaf', amotaf', 'be amotaf', 'be amotaf', 'tawolaf', 'amotaf', 'zawolaf', 'yaszlochaf', 'yyaszlochaf', 'yyaszlochaf', 'yyaszlochaf', 'yyaszlochaf', 'y	Dane jodpari', 'n' podpari', 'n' szydzi', 'zapy szydzi', 'zapyt 'nyknąt'} 'wydysza edziat') if', 'zapyt	PL Recenzja W ucieszył', 'p denerwow 'polecił', 'z tał') wołał', 'ode adał', 'stwie ił', 'jęknął', ' ', 'szukał', 'i sał')	1_1HP_CS_PC Road olecif,'taga C C agadnąf', 'thciaf' f, 'wrzasną saf', 'powt zwaf', 'nycc zwaf', 'powt wyjąkaf')	olishtypes - I gadnął', 'wy 'przyznał', wysapał', 'z tł'} arzał', 'mru zał', 'wrzasz wiedział'}	Microsoft Ex sapal', 'zayariowa 'zwariowa aperzył', 'k aperzył', 'wyj	cel uzytek nieko serzyłł, 'uciął' ł', 'rzekł', 'zap uciął', 'zaklika ąkał', 'mrucza	, 'zaklika' , 'zaklika' TT, F, 'ciąę	, 'ciągnąt', 'p D D types, Va 114 4 6 9 9 2 8 8 11 6 2 2 8 3 2 2 6 2 2 6 2 2 6 2 2 6 2 2 1 1 1 3 3 2 2 4 4 3 2 2 6 2 2 6 2 2 6 2 2 6 2 2 6 2 2 3 3 2 2 6 2 2 3 3 2 2 2 3 3 3 2 2 3 3 3 3	v wypalił', vi teomruki wa 947 23 20 19 16 15 11 10 10 9 9 9 9 7 7 6 5 5 5 5 5 5	vykrzykną F	', 'powital 'wyznaf', G	', 'uciszyf', 'o Stracif', 'o H	'skarcil', 'zr burzył', 'do I	♥ 🖗 📼 alazi', j	
C2 A A said whisperers asked yelled added total total hissed roared gasped snapped gasped snapped gasped croak	v (v	Wissendari Utidat sin C (Vaskaz) Variat (Vaskaz) 947 (Vaskaz) 923 ('szepari, 'uspaz) 23 ('szepari, 'uspaz) 10 ('variat, 'uspaz) 15 ('dodar) 15 ('aspaz), 'kspaz, 'kspaz) 10 ('ksrayna], 'tspaz) 15 ('ksrayna], 'tspaz) 16 ('ksrayna], 'tspaz) 17 ('ksrayna], 'tspaz)	ony Formuly aff, 'vzzywaf', ' orzedstawii', 'u eptaf', 'oznajn eptaf', 'oznajn eptaf', 'oznajn eptaf', 'oznajn eptaf', 'oznajn eptaf', 'aznajn ', 'zaziotaf', 'beł af') ', 'zaziotaf', 'w ezavolaf') hnął ', 'zaziotaf', 'arzenaj pnął', 'odpowi jaśnif', 'zakp olaf')	Dane 1 podpari', 'i, ucieszyi', 'zay szydzil', 'zapyt i', 'apyt '', 'opowi ''ywydysza edział'} i', 'zapyt	Recenja W denerwow 'polecii', 'polecii', 'tat' powiedział howiedział', 'apie powiedział , 'apie powiedział , 'stwin i', 'szukaf', 'stwin af')	1, LHP CS, Pc (SAR) Notek af, 'chciaf', 'chciaf', C agadnąf', 'r, 'wrzasną saf', 'powt zwaf', 'powt wyjąkaf') wdpowiedz	olishtypes - I gadnął', 'wy 'przyznał', wysapał', 'z tł'} arzał', 'mru zał', 'wrzasł wiedział'}	Microsoft Ex sapal', 'zayariowa zwariowa aperayf', 'uyia aperayf', 'uyia	cel utytek nieko serzył", 'uciął" r, 'rzekł", 'zapr uciął", 'zaklika ąkał", 'mrucza	, 'zaklika'i evevni', 'z TT, F, 'ciąg	, 'ciagnal', D D D 1 114 4 6 9 9 2 8 8 1 6 6 2 8 8 1 6 6 2 2 6 6 2 2 1 1 3 3 2 2 4 4 2 2	Vyopalit ⁷ , ¹ emrukiwaba 23 23 20 20 29 20 20 20 20 20 20 20 20 20 20 20 20 20	wykrzykną F	','powital 'wyznar', G	, 'uciszyf', 'uciszyf', 'uciszyf', 'uciszyf', 'uciszyf', 'uciszyf', 'uciszyf', 'uciszyf', 'uciszyf', 'uciszyf'	skarcil, 'zr burzyl', 'do I	♥ 🖓 📼 alazť, j j	
C2 A source_E said whispere shouted mutteree asked velled told told told told told told told snapped gasped snarled graveled went on repeated breathere sneered croaked breathere sneered toreaked breathere sneered toreaked to told to told to told to told told told told told told told told	v (B Source_ edd	Witsendorf Uilded th Image: Control of the state of the sta	ony Formuly aff, 'wzywaf', ' zrzedstawit', 'u vaf', 'oznajni iqf', 'zawolaf', iamortaf', 'bei aff', 'zawolaf', iamortaf', 'tawolaf', iamortaf', 'zawolaf', iamortaf', 'zawolaf', iamortaf', 'zawolaf', izaziotaf', 'zawolaf', 'zaszlochaf', 'zaszlochaf', 'zaszlochaf', 'zaszlochaf', 'zaszlochaf', 'zaszlochaf', 'zaszlochaf', 'zaszlochaf', 'zaszlochaf', 'zaszlochaf', 'zakp otaf'))	Dane Dane vrpadł', 'z vrypadł', 'z vrypadł', 'z vryknął' ' 'ryknął' ' ' 'ryknął' ' ' ' ' ' ' ' ' ' ' ' ' '	PL Recencija W denerwow , 'polecil', 'z tat') powiedział , 'napie , 'powiedział , 'napie , 'powiedział , 'napie , 'sowież, 'stwi- , 'szukat', 'stwi- ał')	1_1HP.CS.Pc/ski Noke C agadnąf, 't ', 'wrzasną saf', 'powt zzwał', 'tyc: zzwał', 'tyc: zzwał', 'tyc: wyjąkał'}	sadnąt', wysa ('przyznat'', ('przyznat'', 'przyznat'', 'przyznat', 'przyznat', 'przyznat', 'mrzasz ('przyznat', 'mrzasz ('przyznat'', 'mrzasz))	Microsoft Ex sapal', 'zay aperzyf', 'uyja aperzyf', 'wyja	cel utytek nieko perzyłł, 'uciął' ł, 'rzekł', 'zapi uciął, 'zaklika ąkał', 'mrucza	, 'zakika'ika'ika'ika 'zakika'ika'ika 'zakika'ika'ika 'zakika'ika'ika 'zakika'ika'ika 'zakika'ika'ika'ika 'zaki	, 'ciągną", 'p swolat", 'p Lypes Va 114 4 6 9 2 8 8 1 4 4 6 6 2 2 4 3 2 2 6 6 6 2 2 1 1 3 2 2 4 2 2 4 4 2 2 2 2 2 2 2 2 2 2 2 2	vvvpalif, vr emruktivati 947 - 23 23 23 23 23 23 23 23 23 23 23 23 23 2	wykrzykną F	', 'powital' 'wyznaf', G	', 'uciszyf', 'stracif', 'o	iskarcii, iz	♥ 🖗 == alazť, Jałł, J	
C2 A asource_E said mutterec asked added dddd dddd dddd hissed called snapped snapped snapped snapped breathed breathed breathed breathed snapped snap	v (B N source ed d	Wissendari Utidat str. Image: Constraint of the strength of the st	ony Formuly aff, "vzywaff, c przedstawii", 'v waf', 'conajn (gf, 'zawolaf', ' namrotaf', 'beł af') aff, 'zawiaf', ' nagf, ' , 'zasziochaf', nagf) , 'zasziochaf', nagf) , 'zasziochaf', magf) , 'zasziochaf', magf) , 'zasziochaf', magf) , 'zasziochaf', magf) , 'zasziochaf', jaśnif')	Dane Ddpart ¹ , 1 ¹ pypadt ¹ , 2 ² szydzil ¹ , 'zapy 'ryknąt ² 'wydysza dział ¹ if', 'zapyt	Recenja V V Recenja V V denerwow 'polecii', 'polecii', ' powiedział powiedział powiedział powiedział powiedział powiedział ('stwi ', 'szukaf', 'stwi af')	1_1HP CS.Pc/solo	;adnąt', 'wysaka , 'przycnał', (f) (f) (f) (f) (f) (f) (f) (f) (f) (f)	Microsoft Ex sapat', 'zay 'zwariowa aperzył', 'u knąť, 'wyja	cel utytek nieko perzył', 'uciął' , 'rzekł', 'zap rciął', 'zaklika gkał', 'mrucza	, 'zaklika'i , 'zaklika'i TT, T', 'ciąę	, 'ciągnąt', Dy Dy Dy Dy 114 4 6 9 2 8 8 1 1 6 2 8 8 1 6 2 2 6 6 2 2 1 1 3 3 2 2 4 4 2 2 1 1 2 2 1 1 3 3 2 2 2 1 1 1 4 1 4 1 4 1 4 1 4 1 4 1 1 4 1 4	Vwypalif, Y, Yapalif, Yapal	vykrzykną F	','powital 'wyznaf', G	', 'uciszyf', 'o	i skarci, iz burzył, do	© @ □ alazť, j	
C2 A source_E_said whispere asked added told hissed called sanared gasped gasped gasped went on repeated breathed explanered croaked breathed stammerr	reed	Witsendorf Uikid str Image: Constraint of the strength of the stre	ony Formuly af, 'vzywaf', 'v vzrydstawi', 'v vaf', 'ocnajng', 'zawala', ' namrotaf', 'bel af', 'zawala', 'krykngf', taf', 'zawolaf', 'ngf', 'zawolaf', 'ngf', 'zawolaf', 'ngf', 'zawolaf', 'ngf', 'zawolaf', 'ngf', 'zawolaf', 'zaszochaf', 'ark jachif', 'yychngf', 'zakp olaf')	Dane Dane ypadł', 'z ypadł', 'z yzydźł', 'z szydźł', 'z wodział', 'g wydysza edział', 'ryknął ł', 'zapyt	Recently U work of the second seco	1_1HP.CS.Pc/skie Skok C G agadnąf, 't 'f, 'wrzasną saf,' powt zzwał, 'rycy character wychyckaf') y odpowiedz	;gadnaf', 'wysana'', 'z 'przyznaf', ' I'Przyznaf', ' 'przyznaf', ' ' I'P' I'P' I'Zar', 'mru Szaf', 'mru Szaf', 'mru Szaf', 'mru Szaf', ' ' Szaf' ' ' Szaf' ' ' Szaf' ' ' Szaf' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' '	Microsoft Ex sapat', 'zay aperzyf', 'uyj knąf', 'wyj	cel uzytek nieko serzył", 'uciął" r', 'rzekł", 'zag uciął", 'rzekłika ąkał", 'mrucza	, taklika ika ika ika ika ika ika ika ika ika	, 'ciagnal', 'p types_Va types_Va 4 4 4 6 9 2 8 8 1 1 4 4 6 2 2 6 6 2 6 2 1 1 1 3 2 4 4 2 2 4 1 1 4 4 4 5 8 8 1 1 4 4 4 4 4 6 9 2 2 4 4 4 4 4 4 6 9 2 2 4 4 4 4 4 6 9 2 2 4 4 4 4 4 6 9 2 2 4 4 4 4 6 9 2 2 4 4 4 6 9 2 6 6 2 6 2 6 2 6 2 6 2 6 2 6 2 6 2 6 2 6 2 6 2 6 2 6 2 6 2 6 2 2 4 4 4 4 4 6 9 2 6 6 2 6 2 6 2 2 4 4 4 4 4 4 4 4 4 4 4 4 4	'wypaif', 'wy	ivykrzykną F	''yowitaf	', 'uciszyf', 'stracif', 'o	i skaril, z	♥ ♥ ♥ ■ alazt', Jalazt', J	

In Polish, all feminine forms of verbs were reduced to the masculine form (*powiedziała*

=>powiedział)

	A1 -	fr	source EN	word			
Plik	Narzędzia główne	Wstawianie	Układ strony	Formuły	Dane	Recenzja	Widok
	- (*					P	L_1_HP_CS_Polishtypes_tw - Microsoft Ex

	A	В	С	D	E	F	G
1	source_EN	source_EN_count	TT_types	Var_1_freq	Var_2_trans	Var_3_senses	Var_4_verb_type
2	said	947	114	947	3	11	N
3	whispered	23	4	23	3	1	Vier
4	shouted	23	6	23	4	4	Pros
5	muttered	20	9	20	4	2	Vier
6	asked	19	2	19	6	7	Str
7	yelled	16	8	16	5	2	Pros
8	added	15	1	15	2	6	Sdis
9	told	15	6	15	3	8	N
10	hissed	11	2	11	3	4	Vion
11	roared	10	4	10	2	6	Vion
12	called	10	3	10	6	28	Mprop
13	snapped	9	2	9	1	13	Mprop
14	gasped	9	6	9	4	1	Vion
15	snarled	7	2	7	2	4	Vion
16	growled	7	6	7	5	1	Vion
17	went on	6	2	6	0	5	Sdis
18	repeated	5	1	5	2	6	Sdis
19	croaked	5	1	5	6	3	Vion
20	breathed	5	3	5	1	9	Vier
21	explained	5	2	5	4	3	Mprop
22	sneered	5	4	5	5	2	Mprop
23	cried	4	2	4	3	7	Pros
24	howled	4	2	4	2	4	Vion
25	stammered	4	1	4	6	1	Vion
26	sighed	4	1	4	3	2	Vion
27	began	4	1	4	4	10	Sdis

X-	-	0	*	÷

Ph	Narz	ędzia główne	Wstawiani	e Układ st	rony Form
	A1	• (- 1	types	
	A	В	С	D	E
1 t	ypes	freq	trans	senses	verbtype
2	114	947	3	11	N
3	4	23	3	1	Vier
4	6	23	4	4	Pros
5	9	20	4	2	Vier
6	2	19	6	7	Str
7	8	16	5	2	Pros
8	1	15	2	6	Sdis
9	6	15	3	8	N
10	2	11	3	4	Vion
11	4	10	2	6	Vion
12	3	10	6	28	Mprop
13	2	9	1	13	Mprop
14	6	9	4	1	Vion
15	2	7	2	4	Vion
16	6	7	5	1	Vion
17	2	6	0	5	Sdis
18	1	5	2	6	Sdis
19	1	5	6	3	Vion
20	3	5	1	9	Vier
21	2	5	4	3	Mprop
22	4	5	5	2	Mprop
23	2	4	3	7	Pros
24	2	4	2	4	Vion
25	1	4	6	1	Vion
26	1	4	3	2	Vion
27	1	4	4	10	Sdis

- Generalised linear model with negative binomial regression: overdispersion in our data (variance > mean); outcome variable is a count variable
- Two analyses: combined Italian data and combined Polish data (some novels with less then 15/20 observations per predictor, cf. Hair et al. 2009: 176);
- Preliminary analyses to find the best-fitting model (lowest AIC).
 - ► Tailor-made R scripts

Results: Italian

AIC: 1469.2

Overall (left): Freq, Verbtype, and Trans (borderline: p=0.072) are significant predictors of Type; Senses doesn't have a significant effect.

Coefficients:

	Estimate Std. Error z value Pr(> z)
	(Intercept) 0.02452 0.10289 0.238 0.8117
	verbtypeN 0.30647 0.16639 1.842 0.0655.
	verbtypePros -0.17199 0.11635 -1.478 0.1393
Df Doviance Pesid Df Pesid Dov Dr(>Chi)	verbtypeSdis -0.22725 0.11531 -1.971 0.0488 *
$\frac{1}{1}$	verbtypeStr -0.17462 0.15643 -1.116 0.2643
NULL 402 1037.25 Norhtuno 7 1007 77 LEE EKK 08 < 70-16 ***	verbtypeVier -0.15484 0.11431 -1.355 0.1755
laafraa 1 220 12 LEL 226 86 < 20-16 ***	verbtypeVion -0.06859 0.08700 -0.788 0.4305
trane 1 2.22 LE2 2L2 &L 0.07262	verbtypeMlin -0.77609 1.03921 -0.747 0.4552
114113 1 3.22 433 243.04 0.01202 .	logfreq 0.48366 0.02844 17.005 <2e-16 ***
	trans 0.05505 0.03059 1.800 0.0719.

- How often a reporting verb occurs, what type of verb it is, and the number of available translation equivalents influence the chances of seeing that verb translated into multiple TL items (positive association).
 - ▶ e.g. (right) for a one-unit change (increase) in **logfreq**, *types* increase e^{0.48}-fold
 - a verb (e.g. screamed) has its verb frequency logfreq value of 2. If we increase the frequency by one unit (to a logfreq value of 3), we would expect that the number of translation equivalents (types) increases by approximately 48% to its previous value

Results: Polish

AIC: 1575.5

Overall (left): Freq, Verbtype, Senses, and interaction Trans:Senses are significant predictors of Type; Trans doesn't have a significant effect.

 Df Deviance Resid. Df Resid. Dev Pr(>Chi)

 NULL
 451
 4697.9

 verbtype
 7
 3651.0
 444
 1046.9
 < 2e-16 ***</td>

 logfreq
 1
 832.5
 443
 214.5
 < 2e-16 ***</td>

 trans
 1
 2.4
 442
 212.1
 0.12230

 senses
 1
 5.0
 441
 207.1
 0.02538 *

 trans:senses
 1
 6.5
 440
 200.5
 0.01054 *

Coefficients:

Estimate Std. Error z value Pr(>[z]) (Intercept) -0.255314 0.149713 -1.705 0.088129 . verbtypeN 0.451657 0.134111 3.368 0.000758 *** verbtypePros -0.158915 0.100394 -1.583 0.113440 verbtypeSdis -0.250851 0.104430 -2.402 0.016301 *

verbtypeStr -0.412646 0.129258 -3.192 0.001411 ** verbtypeVier -0.195429 0.105511 -1.852 0.063996. verbtypeVion -0.119487 0.084124 -1.420 0.155502 verbtypeMlin 0.181544 0.597759 0.304 0.761350 logfreq 0.592083 0.023207 25.513 < 2e-16 *** trans 0.098669 0.031383 3.144 0.001666 ** senses 0.029276 0.017031 1.719 0.085623. trans:senses -0.008774 0.003374 -2.600 0.009310 **

- How often a reporting verb occurs, what type of a verb it is, and the interaction between the number of senses and translation equivalents influence the chances of seeing that verb translated into multiple TL items
 - e.g. (right) for a one-unit change (increase) in **logfreq**, *types* increase $e^{0.59}$ -fold
 - a verb (e.g. screamed) has its verb frequency logfreq value of 2. If we increase the frequency by one unit (to a logfreq value of 3), we would expect that the number of translation equivalents (types) increases e^{0.59}-fold

Results: Italian (without said)

AIC: 1629.2

 Overall (left): Freq, Verbtype, and Trans are significant predictors of Type; Senses doesn't have a significant effect.

Df Deviance Resid. Df Resid. Dev Pr(>Chi) NULL 451 480.44 verbtype 7 48.069 444 432.37 3.452e-08 *** logfreq 1 193.561 443 238.81 < 2.2e-16 *** trans 1 4.544 442 234.26 0.03303 *

Coefficients:

Estimate Std. Error z value Pr(>|z|)(Intercept) 0.08956 0.09806 0.913 0.3610 verbtypeN 0.28627 0.18131 1.579 0.1144 verbtypePros -0.14079 0.10604 -1.328 0.1843 verbtypeSdis -0.20807 0.10576 -1.967 0.0491 * verbtypeStr -0.12397 0.14089 -0.880 0.3789 verbtypeVier -0.12695 0.10379 -1.223 0.2213 verbtypeVier -0.12695 0.10379 -1.223 0.2213 verbtypeVion -0.07096 0.07966 -0.891 0.3731 verbtypeMlin -0.80951 1.00278 -0.807 0.4195 logfreq 0.43751 0.03137 13.949 <2e-16 *** trans 0.05982 0.02782 2.150 0.0315 *

► The presence or absence of said doesn't change the results.

Results: Polish (without said)

AIC: 1451.2

Overall (left): Freq, Verbtype, Senses, and interaction Trans:Senses are significant predictors of Type; Trans doesn't have a significant effect.

Df Deviance Resid. Df Resid. Dev Pr(>Chi) NULL 440 694.45 verbtype 7 163.29 433 531.16 < 2e-16 *** logfreq 1 337.79 432 193.37 < 2e-16 *** trans 1 2.41 431 190.96 0.12064 senses 1 4.89 430 186.07 0.02706 * trans:senses 1 5.57 429 180.51 0.01830 * **Coefficients**:

Estimate Std. Error z value Pr(>|z|)(Intercept) -0.167841 0.154303 -1.088 0.27671 verbtypeN 0.349472 0.159390 2.193 0.02834 * verbtypePros -0.121462 0.099935 -1.215 0.22421 verbtypeSdis -0.234794 0.101897 -2.304 0.02121 * verbtypeStr -0.310430 0.133016 -2.334 0.01961 * verbtypeVier -0.161899 0.103769 -1.560 0.11872 verbtypeVier -0.161899 0.103769 -1.560 0.11872 verbtypeVion -0.120793 0.081902 -1.475 0.14026 verbtypeMlin 0.169165 0.582811 0.290 0.77162 logfreq 0.556015 0.029994 18.537 < 2e-16 *** trans 0.092077 0.030819 2.988 0.00281 ** senses 0.025857 0.016774 1.541 0.12321 trans:senses -0.007977 0.003329 -2.397 0.01655 *

► The presence or absence of *said* doesn't change the results

Results: verb types and number of translations

Italian

Verbtype	Estimate
Neutral	0.30647
Metapropositional	0.02452
Voice qualification	-0.06859
Voice qualifier	-0.15484
Prosodic	-0.17199
Structuring	-0.17462
Signaling discourse	-0.22725
Metalinguistic	-0.77609

Polish

Verbtype	Estimate
Neutral	0.451657
Metalinguistic	0.181544
Voice qualification	-0.119487
Prosodic	-0.158915
Voice qualifier	-0.195429
Signaling	
discourse	-0.250851
Metapropositional	-0.255314
Structuring	-0.412646

- Neutral verbs (e.g. said, told) more likely to be translated into different TT verbs
- Structuring (e.g. asked, replied) and signalling discourse (e.g. added, repeated) verbs less likely to be translated into different TT verbs

Conclusions - case study # 1

- For both Italian and Polish data, the positive correlation between Freq and Types has been confirmed
- For Polish data, the interaction between Senses and Trans has a significant effect on Types, but not Trans on its own
 - ► those verbs which are polysemous have many translation equivalents
 - polysemy (along with homonymy) may be realized differently across languages (Srinivasan & Rabagliati 2014; Zuercher 2019)
- ► By learning what can prompt translation of repeated reporting verbs (e.g. *said*) into a wider lexical variety, we can improve translation strategies to deal with repetition, making them more sensitive to the **stylistic effects of the original**

Case study #2

(English and Polish) lexicographers' habits with respect to recording multi-word expressions in dictionaries

• Maziarz, M., Grabowski, Ł., Piotrowski, T., Rudnicka, E., Piasecki, M. (2023). "Lexicalisation of Polish and English word combinations: an empirical study". *Poznan Studies in Contemporary Linguistics*. https://doi.org/10.1515/psicl-2023-2002

MWEs (multi-word expressions) vs MWLUs (multi-word lexical units)

In theory, and in practice (lexicography), it is difficult to determine when MWEs are treated as word combinations and when as vocabulary units

- blood bank, blood pressure, black lung, black coal, early night, cold night, early grave, mass grave
- Iew morski 'sea lion', ssak morski 'sea mammal', dziecko szczęścia 'whiz-kid/child of fortune', przyczyna szczęścia 'the cause of happiness/luck'

Various linguistic criteria are taken into consideration

- formal-linguistic, frequency-driven/distributional (statistical), and psycholinguistic ones
 - frequency of distribution in text and speech, (non-)standard morphosyntactic patterns, constituent (non-)substitutability, semantic (non-)compositionality, among others (Zgusta 1971, Bright 1992, Biber et al. 1999, Hunston and Francis 2000, Svensén 2009)

MWEs vs MWLUs

Any idiosyncrasy in those criteria may be a signal that a process of **lexicalisation** has started (Baldwin and Bond 2002, Baldwin and Kim 2010, Ježek 2016: 25-28, Constant et al. 2017: 838).

 the more idiosyncrasies, the closer to a lexical unit (LU) status a given word combination (MWE) gets

Lexicalisation is a syntax-to-lexicon process of transgression of a purely compositional word combination (a syntactic unit) into a single semantic or pragmatic unit, exhibiting word-like behaviour (Ježek 2016: 25-28, Constant et al. 2017: 838)

- ► a gradable process best described on a scale
 - purely compositional word combinations created ad hoc on one end, through collocations, to fixed phrases and idioms on the other end

Problems with lexicalisation

Lexicalisation of MWEs is a scalar and gradable process, yet in lexicography and lexical databases construction (e.g. wordnets) it is a dichotomous (Yes or No) process

- clear, operational procedures for including MWEs in dictionaries or lexical databases are scarce (there are exceptions though)
- Iexicographers rely on their linguistic intuition, linguistic experience and competence, contemporary and previous sources of information (dictionaries, books etc.) to decide which MWE should be granted an MWLU status

More problems

Lexicographers are a very special type of native speakers!

- They work with tens of thousands of lexical items, and their experience is far wider than that of a typical native speaker
- A structure that may appear to be semantically opaque to a "lay" native speaker (e.g. rare or non-modern) may be transparent to a lexicographer,
 - semantically transparent word combinations are often excluded from the dictionary
- A lexicographer may include transparent word-combinations because of publishing demands or pressure of competition
 - unlike a linguist, a lexicographer works under enormous pressure of time, and may include some items because there is simply no time
- Dictionaries are usually produced by teams, which often results in uneven treatment of individual items.
 - one has to treat dictionary data with caution!

Our study (Maziarz et al. 2023)

We explore **specific factors contributing to the classification/treatment of MWEs as MWUs**

We **review a number of lexicality criteria** and select the ones with the highest potential to be operationalised

semantic compositionality (substitutability/paraphrasing, linguistic intuition), measures of collocational strength, word length (attribute and head)

We explore the criteria in a series of experiments conducted using a multivariate statistical method called Latent Class Analysis (LCA)

Our study (2)

We study two samples of Polish and English MWEs

- ▶ **bigrams**: adjectives and nouns (A+N) or nouns and nouns (N+N),
- we contrast the MWEs extracted from selected monolingual dictionaries with those formed by ourselves, and verify which of the MWEs should be accorded MWU status
- ► DATASET: <u>https://clarin-pl.eu/dspace/handle/11321/853</u>

We deconstruct the English and Polish lexicographers' habits in order to better understand how they deal (what criteria guide their decisions?) with registering MWEs in dictionaries

Research material

Two samples of MWEs - lexicalised and non-lexicalised ones (350 MWEs)

- The first sample extracted based on experts' opinion / purposive sampling - from monolingual Polish (e.g. bułka paryska, lew morski, dziecko szczęścia) and English dictionaries (e.g. floor plan, flight bag, belly laugh)
 - (Dunaj 1996: Słownik współczesnego języka polskiego, SWJP; Pearsall 2001: The New Oxford Dictionary of English)
 - ► 102 English and 91 Polish MWEs recorded in dictionaries (lexicalised, MWLUs)
- The second (control) sample created by us and modelled on those MWEs in the first sample (np. twarda bułka, ssak morski, przyczyna szczęścia, clear plan, garbage bag, loud laugh) with the intention to obtain non-lexicalised items
 - ► 76 English andi 91 Polish MWEs not recorded in dictionaries (not lexicalised ones)
 - We verified if they were recorded there (11 English MWEs and 1 Polish MWE were found to be listed in the dictionaries)

Methodology. Procedures and study stages

We operationalised **semantic compositionality** by testing two custom-designed criteria, i.e., **intuition and paraphrase**, and by using statistical methods (selected **measures of collocational strength: LL, PMI and Jaccard**) for checking word order fixedness and word combination specificity

We also checked **how long (in letters) the syntactic head / its complement is** (the measure highly correlated with word frequency (Piantadosi et al. 2011, Strauss et al. 2007), word familiarity and with the degree of polysemy (Grzybek 2014)

combination of semantic compositionality with syntactic fixedness

Lexicality was identified with the presence of an MWE in a dictionary

Methodology. Procedures and study stages

Two teams of annotators (one team for Polish MWEs, and one for English ones), with three annotators in each

Intuition was verified by a test question: "Can you derive the meaning of an MWE from the meanings of its component parts and their syntactic relations?"

- ► If 'yes', the MWE is compositional;
- ► if 'no', the MWE is non-compositional

Paraphrase was verified by a test question: "Can you paraphrase an MWE using a relative clause built of MWE components and of function words (prepositions and conjunctions) combined using regular syntactic rules?"

Lexicality: for manually created MWEs, we checked whether they are recorded in the English or Polish dictionaries

- ► English: Lexico.com (OED), Merriam-Webster
- Polish: WSJP, SJP (PWN), SJP Dor, Dunaj 1996 (SWJP)



We want to identify hidden dependencies between the studied criteria and indicators

▶ we run Latent Class Analysis

Why?

We want to predict the presence of MWEs in two classes: recorded and not recorded in dictionaries

Results (1). Inter-annotator agreement

From all pairs of annotations, we randomly selected the annotations of particular pairs and recorded them in two separate subsets

- Subsets with paired annotations and used Cohen's kappa κ (Cohen 1960)
 - ► Range is the confidence interval at 95% probability

Criterion	Agreement	CI	Annotation quality
Intuition	<i>κ</i> = 0.35	[0.351, 0.354]	"Fair"
Paraphrase	<i>κ</i> = 0.64	[0.638, 0.639]	"Substantial"

Table 1. Inter-annotator agreement (Cohen's kappa) for Polish and English MWEs

Given the minimal (intuition) and moderate (paraphrase) inter-annotator agreement, we averaged their decisions: if they replied "YES" and "NO", then we proposed a new category "I DON'T KNOW".

Results (2). Collocation measures and word length

For each MWE, we calculated its collocation strength measures: Log Likelihood (LL), Pointwise Mutual Information (PMI), and Jaccard.

- ► High values of PMI and Jaccard -> a strong collocation;
- ► High values of LL -> a weak collocation with free word order
- We used different corpora for Polish and English and analyzed both data sets separately
- For English, we merged some corpora from the NLTK package : The Reuters-21578 benchmark corpus, Project Gutenberg Selections, Brown Corpus, Australian Broadcasting Commission 2006, Twitter Samples (50 million tokens)
- For Polish, we used the KGR10 Corpus with its 4 billion words (Kocoń and Gawor 2019)

All variables

<u>Semantic and syntactic features included three nominal variables</u> (intuition, paraphrase, and MWE status), **two ordinal variables** (HWL and AWL) and **three continuous variables** (collocation measures):

Semantic' features:

- Intuition ("YES" "DON'T KNOW" "NO"),
- Paraphrase ("YES" "DON'T KNOW" "NO"),
- MWE Status ("Dict" / "non-Dict"),

Syntactic' features:

- HWL and AWL (values: "short" vs. "long", i.e. "1-4 letters" and "5+ letters", respectively, for English and "1-5", "6+" for Polish),
- Collocation measures: LL, PMI and Jaccard.

Criteria		Polish	English	
Lexicality	MWE status	" <mark>Dict</mark> "	91 [50%]	102 [58%]
		"non-Dict"	91 [50%]	76 [42%]
Semantic	Intuition	"YES"	45 [25%]	115 [65%]
		"DON'T KNOW"	86 [47%]	27 [15%]
		"NO"	51 [28%]	36 [20%]
	Paraphrase	"YES"	101 [56%]	111 [63%]
		"DON'T KNOW"	22 [12%]	36 [20%]
		"NO"	59 [32%]	31 [17%]
Syntactic		mean Jacc	0.0240	0.0072
		mean LL	54.7	62.0
		mean PMI	5.22	7.44
	AWL	"short"	73 [40%]	46 [26%]
		"long"	109 [60%]	132 [74%]
	HWL	"short"	63 [35%]	62 [35%]
		"long"	119 [65%]	116 [65%]

Table 2. Statistics for variables related to syntactic and semantic features of MWEs

Statistics for all variables

Word Length (HWL and AWL): For Polish, the median of the phrase head and the attribute is 6; for English it is 5.

Word length was transposed into the ordinal variable:

- "short" -> MWEs whose head word length (HWL) / attribute word length (AWL) was shorter than the median
- "long" -> MWEs whose HWL and AWL scores were equal or longer than the median

Latent class analysis (LCA)

LCA is a statistical technique (Lazarsfeld 1950) used to identify hidden (unseen) groups or classes in a population based on observed nominal variables

nominal variables: lexicality (Status = Dict or Non-Dict), intuition (Yes, No, I don't know and paraphrase (Yes, No, I don't know)

The probability of belonging to a particular class is assigned to each word combination

In practice, we want to **select the model** (from among the ones of various number of hidden classes) **with the lowest AIC and BIC scores** (quality=data matching and model simplicity respectively)

- ▶ it means selecting a model with an optimum number of classes
- ▶ we used R library PoLCA (Linzer and Lewis 2011)

Results (3a). LCA for Polish

3 class model as an optimum one (lowest AIC and BIC values), high p-value of bootstrapped model fits statistics

Class #1 - "compositional MWEs not found in dictionaries"

- word combinations out of a dictionary (Status = "non-Dict"),
- paraphrasable ("YES"),
- vague intuition ("DON'T KNOW", in rare cases "YES")
- e.g. czerwony dywan, kozie mleko, twarda bułka

Class #2 - "(partly) compositional MWEs found in dictionaries"

- word combinations from dictionaries
- either vague intuition of compositionality ("DON'T KNOW") or strong intuition ("YES"),
- partly/fully paraphrasable ("DON'T KNOW", "YES")
- e.g. gorąca krew, gruba forsa, krótka pamięć

Class #3 - "non-compositional MWEs found in dictionaries"

- word combinations come from dictionaries (Status = "Dict"),
- non-paraphrasable ("NO"),
- clearly non-compositional (Intuition = "NO")
- e.g. gabinet cieni, klucz francuski, beczka śmiechu













Results (3b). Polish - word length

- Long attributive words were characteristic for Class #1 (*t*-test *p*-values < 0.005), a likely indicator of compositional MWEs
- Shorter attributive words were characteristic for Classes #2 and #3: partly and non-compositional MWEs
 - Such a relationship did not occur with syntactic heads (not statistically significant -> most probably due to our procedure of constructing non-lexicalised MWEs on the basis of syntactic heads of MWEs found in dictionaries)



Results (3c). Polish - collocational measures LCA with co-variant variables (measures of collocational strength)

Two classes were differentiated by PMI and Jaccard measures Class #1 (compositional)

• lower PMI scores = syntactically free word combinations (p<0.001).

Class #3 (non-compositional MWEs)

higher PMI scores = syntactically fixed 0.8 Probability 0.6 0.4 of obability Jacc 0.2 PM non-compositional MWEs are syntactically fixed 0.0 class 1 ("compositional") class 2 ("partly compositional") class 3 ("non-compositional") 10 to a higher degree than compositional ones PMI

nd MWE class for medium Jacc and LL values

Results (4). LCA for English

3 class model as an optimum one (the lowest AIC and BIC values), high p-value of bootstrapped model fits statistics

Class #1 - "compositional MWEs"

- mostly not recorded in dictionaries ("non-Dict"),
- paraphrasable ("YES"),
- intuitively compositional ("YES")
- e.g. electric fire, huge glasses, noisy party, short break

Class #2 - "partly compositional MWEs found in dictionaries"

- mostly found in dictionaries ("Dict")
- ambiguous status in the paraphrase test ("DON'T KNOW" and rarely "YES")
- uncertain intuition regarding compositionality ("DON'T KNOW" and rarely "YES")
- e.g. *dirty work, dry rot, first finger, floor exercise, flower bed* **Class #3** "non-compositional MWEs found in dictionaries"
- found in dictionaries ("Dict"),
- non-paraphrasable ("NO"),
- intuitively non-compositional ("NO")
- e.g. blood money, dragon tree, dry run, garden cress, rotten apple





'NO" 📒 'DON'T KNOW' 📕 'YES'



class2

MWE class

class3

0,2

0.0

classf



INTUITION

0.75





LCA: comparison of Polish and English MWEs



Discussion

Polish lexicographers don't tend to include compositional MWEs as lexical entries in their dictionaries; the criteria of paraphrase and intuition are important for them

- if MWEs are not clearly and unambiguously paraphrasable and compositional, then they are recorded in dictionaries
 - approximately ¹/₃ of compositional MWEs in Class #1 are recorded in the dictionaries of English (Merriam-Webster and Pearsall 2001); in Class #1 of Polish compositional MWEs not even a single one was recorded in the consulted Polish dictionaries (SJP, SJPDor, WSJP and Wilga)

English lexicographers tend to record also compositional and partly compositional MWEs

► 94% were recorded in the dictionaries of English

English MWLUs are more heterogeneous than the Polish ones

 English compositional MWEs covers word combinations not recorded in dictionaries but, unlike the class of Polish compositional MWEs, it also includes (1/3) compositional MWEs recorded in dictionaries

Conclusions - case study # 2

We obtained a **collective insight into the degree of semantic compositionality of MWEs and into certain criteria guiding lexicographers' decisions** as to which MWEs should be recorded, as MWUs, in dictionaries of Polish and English

We obtained **further evidence that lexicalization of MWEs is gradable in nature**

The findings are potentially useful for extending existing dictionaries or lexical databases, notably wordnets, with new MWLUs

Corpus linguistics in the 2020s (and beyond)

Corpus linguistics has become increasingly embedded in statistics, data science and NLP

research has become increasingly collaborative and interdisciplinary

Corpus linguists benefit a lot from research and development projects

e.g. CLARIN-ERIC or national CLARIN infrastructures offering access to computational tools for linguistic data analysis and large textual datasets Corpus linguistics in the 2020s (and beyond)

Advances in research on application of large language models (LLMs) may bring further changes to the corpus linguists' work

potential (not fully known yet) of using LLMs for linguistic data analysis

- integration of LLMs into corpus analysis tools (Anthony 2023, in progress) as a way to overcome some limitations of corpus analysis tools
- Lew (2023): the quality of COBUILD-style AI-generated (by ChatGPT Plus) lexicographic definitions is comparable to the ones written by human lexicographers
- practical difficulties in using LLMs for corpus-based research (Crosthwaite & Baisa 2023)
 - knowing the data, authenticity, replicability, multimodality, safety (data privacy), hallucinations

AntConc 4.2.4 (Anthony 2023)

	AntConc
Target Corpus	KWIC Plot File View Cluster N-Gram Collocate Word Keyword Wordcloud ChatAl
Name: AmE06	Tokens prompt: 655 completion: 38 total: 693 session total: 2158
Files: 500 Tokens: 1017879	How can I help you?
AmE06_A01.txt AmE06_A02.txt AmE06_A03.txt AmE06_A04.txt AmE06_A05.txt AmE06_A06.txt AmE06_A07.txt AmE06_A08.txt AmE06_A09.txt	You How is the word "gauge" used in the KWIC results? Image: State of the formation of the forma
AmE06_A10.txt AmE06_A11.txt AmE06_A12.txt AmE06_A13.txt AmE06_A14.txt AmE06_A15_txt	Inc. Legacy Max Tokens 50 Temp 1.0 Model gpt-3.5-turbo Source KWIC
AmE06_A16.txt AmE06_A17.txt AmE06_A17.txt AmE06_A19.txt AmE06_A20.txt AmE06_A21.txt AmE06_A22.txt AmE06_A22.txt AmE06_A23.txt AmE06_A24.txt AmE06_A25.txt	Based on the results, is the word "gauge" generally used with positive or negative sentiment?
Progress	Clear Submit
	Time taken (creating chatai results): 1.6353 sec

Corpus linguists in the 2020s (and beyond?)

A linguist in the first place!

The goal of any corpus analysis is to obtain new knowledge of language, be it a language-system or language-in-use

(Some) background in statistics

- to understand, at least conceptually: descriptive and inferential statistics and methods popular in computational linguistics/NLP; specificity of linguistic data (types of variables etc.), estimation of the quality of a statistical model of linguistic data to understand how to interpret different visualizations of linguistic data

(Some) programming skills sufficient to

- **understand programming scripts** (in R, Python or what comes next)
- use collaborative programming platforms (Jupyter Notebook, Google Collab etc.)
 - use visual or low-code programming platforms for linguists, e.g. KNIME (Ihrmark & Tyrkkö 2023)
 - https://pladan.clarin-pl.eu/ (under construction: Pęzik, Kaczyński, Grabowski, Jałowiecki, 2023)
- know how to extract linguistic data or use language tools via API

Collaboration skills

- working with specialists from other disciplines and/or stakeholders in interdisciplinary research projects (e.g. Baker et al. 2019)
 - it also facilitates social impact and public engagement

https://pladan.clarin-pl.eu/



Platforma Danologiczna

Use Cases Resources Q

Contact Us

Login

Platforma Danologiczna CLARIN-PL

Biblioteka zastosowań narzędzi, korpusów i modeli CLARIN-PL w warsztacie badawczym

Usecase query...





Wyznaczanie słów kluczowych w tekście

2023-10-08 Łukasz Grabowski

Scenariusz: Wyznaczanie słów kluczowych w tekście

Wersja: 16-08-2023

Zasoby Clarin: SpokesBiz, Morphodita lub WCRFT, Respa

Poziom trudności: średni







Modelowanie tematyczne krótkich tekstów

2023-10-04 Łukasz Grabowski

Scenariusz: Modelowanie tematyczne krótkich tekstów

Wersja: 05-08-2023

Zasoby Clarin: sentencembeddings, Shortextopic

Zasoby spoza Clarin: możliwe do wykorzystania - Spacy, BERTopic

Poziom trudności: średni

Liczba odsłon: 3



Wyszukiwanie nazw własnych w tekstach - Liner

2023-10-03 Konrad Kaczyński, Łukasz Grabowski

Scenariusz: Wyszukiwanie nazw własnych w tekstach - Liner

Wersja: 01-10-2023

Zasoby Clarin: Korpus Dyskursu Parlamentarnego, NER/Liner

Zasoby spoza Clarin:

Poziom trudności: podstawowy

Liczba odsłon: 3





Opis:

≔

Q

▲ Korpus Paralela - odnajdywanie ekwiwalnatów za pomocą wektorów słów.ipynb ☆ Plik Edytuj Widok Wstaw Środowisko wykonawcze Narzędzia Pomoc <u>Zmiany nie zostana zapisane</u>



Połącz T4 🔻	1
-------------	---

- {x} 1. Z korpusu Paralela pobieramy wybrany podkorpus
- 2. Z pobranych danych generujemy próbkę 1000 par segmentów PL i ENG
 - 3. Liczymy podobieństwo cosinusowe dla par zdań z podkorpusu
 - 4. Generujemy alignment słów/fraz PL i ENG za pomocą wektorów słów/zdań
 - 5. Generujemy słownik zawierający najczęściej występujące tłumaczenia
 - 6. Zapisujemy dane do pliku xlsx

+ Kod + Tekst 🔥 Skopiuj na Dvsk

- Kroki scenariusza:
- Pobierz z korpusu Paralela wybrany zbiór tekstów
 - [] # Instalujemy biblioteki, których google collab nie ma preinstalowanych

```
!pip install jsonlines
            !pip install progress
            Collecting isonlines
             Downloading jsonlines-4.0.0-py3-none-any.whl (8.7 kB)
            Requirement already satisfied: attrs>=19.2.0 in /usr/local/lib/python3.10/dist-packages (from jsonlines) (23.1.0)
            Installing collected packages: jsonlines
            Successfully installed jsonlines-4.0.0
            Collecting progress
             Downloading progress-1.6.tar.gz (7.8 kB)
             Preparing metadata (setup.py) ... done
<>
            Building wheels for collected packages: progress
             Building wheel for progress (setup.py) ... done
Created wheel for progress: filename=progress-1.6-py3-none-any.whl size=9610 sha256=f8b19cd1470656ebcaaa59fdc3b974046b749a0aafca9d7432fc5da4ddb6b179
             Stored in directory: /root/.cache/pip/wheels/a2/68/5f/c339b20a41659d856c93ccdce6a33095493eb82c3964aac5a1
2
            Successfully built progress
```

Conclusions

On the individual level, corpus analysis is often a solitary work

- working in a multi-disciplinary research team helps obtain the collective insight into complex linguistic or social problems involving language use
 - cyberhate (Kopytowska et al. 2017), perception of science by Poles (Dzimińska et al. 2023)
- Corpus linguists' toolkit and skills have become a mix of those of data scientists, computer programmers, statisticians and skilled collaborators
- Technological developments bring about social changes that impact the ways we do corpus linguistic research
 - attention economy (Hyland 2023): better and better research methods may tip the scales in the race for getting published and becoming visible
 - in 2022, Elsevier alone received 2.7 million paper submissions (2.5 million in 2021); over 600,000 papers published in 2021
 - datasphere growth (statista.com, 2023): data volume will grow from circa 125 ZB in 2023 to more than 180 ZB in 2025 (where 1 ZB equals a billion TB)

Final note

Corpus linguistics is not the sole approach to the study of language -> rationalistic "armchair" linguistics continues and is also likely to evolve in the future

- Its cumulative knowledge has led to contemporary developments
- It may provide a critical outlook on current developments in corpus linguistics

The period of Al-assisted (corpus) research has already started

Corpus linguists should navigate through it with an open-mind and caution!

Thank you!

Q & A

e-mail: lukasz(at)uni.opole.pl

References (selected)

Crosthwaite, P. & Baisa, V. (2023). Generative AI and the end of corpus-assisted data-driven learning? Not so fast! Applied Corpus Linguistics, 3(3). https://doi.org/10.1016/j.acorp.2023.100066

Daniel, I. & Tyrkkö, J. (2023). Learning text analytics without coding? An introduction to KNIME. Education for Information 39 (2023) 121–137

Dunne, J. (2022). Natural Language Processing for Corpus Linguistics (Elements in Corpus Linguistics). Cambridge: Cambridge University Press, 96 p.

Hirschberg, J. & Manning, Ch. (2015). Advances in natural language processing. Science, 349(6245), 261-266.

Hyland, K. (2023). Academic publishing and the attention economy. Journal of English for Academic Purposes, 64. https://doi.org/10.1016/j.jeap.2023.101253

Jurafsky, D. & Martin, J. (2023). Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. (Third edition e-book: draft of January 7, 2023). Accessible at https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf (accessed on 19 July 2023).

Lew, R. (2023, June 12). ChatGPT as a COBUILD lexicographer. Accessible at https://doi.org/10.31219/osf.io/t9mbu

NCES. (n.d.). Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2020, with forecasts from 2021 to 2025 (in zettabytes). In Statista - The Statistics Portal. Accessible at https://www.statista.com/statistics/871513/worldwide-data-created/

Winter B. (2019). Statistics for Linguists: An Introduction Using R. London: Routledge, 310 p.

Selected references - Case study # 1

Ben-Ari, N. (1998). The ambivalent case of repetitions in literary translation. Avoiding repetitions: A 'universal' of translation. Meta, 43(1), 68–78.

Caldas-Coulthard, C. R. (1987). Reported speech in Written narrative texts. In Discussing Discourse, edited by M. Coulthard, 149–67. Birmingham: University of Birmingham.

Čermák, F., & Rosen, A. (2012). The case of InterCorp, a multilingual parallel corpus. International Journal of Corpus Linguistics, 17(3), 411–27.

Čermáková, A. (2015). Repetition in John Irving's novel A Widow for One Year. A Corpus stylistic approach to literary translation. International Journal of Corpus Linguistics, 20(3), 355–77.

De Sutter, G., & Lefer, M.-A. (2019). On the need for a new research agenda for corpus-based translation studies: A multi-methodological, multifactorial and interdisciplinary approach. Perspectives, 28(1), 1–23.

Mastropierro, L. (2020). The translation of reporting verbs in Italian: The case of the Harry Potter series. International Journal of Corpus Linguistics, 25(3): 241–69.

Nádvorníková, O. (2020). Differences in the lexical variation of reporting verbs in French, English and Czech fiction and their impact on translation. Languages in Contrast, 20(2), 209–34.

Piotrowski, T. (1994). Problems in bilingual lexicography. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.

Selected references - Case study # 2

Baldwin, T. and Kim S.N. 2010. "Multiword Expressions". In: Indurkhya, N., Damerau, F. J. (Ed.). Handbook of Natural Language Processing (Vol. 2). CRC Press.

Bauer, L. 2019. "Compounds and multi-word expressions in English". In: B. Schlücker (Ed.), Complex Lexical Units: Compounds and Multi-Word Expressions. Berlin: de Gruyter, 45-68.

Corpas Pastor, G. and Colson, J-P. (eds.). 2020. Computational Phraseology. Amsterdam: John Benjamins.

Hanks, P. 2013. Lexical Analysis: Norms and Exploitations. Cambridge, MA: MIT Press.

Hunston, S. and Francis, G. 2000. Pattern Grammar: a corpus-driven approach to the lexical grammar of English. Amsterdam: John Benjamins.

Moon, R. 1998. Fixed Expressions and Idioms in English. A Corpus-Based Approach. Oxford: Clarendon.

Pęzik, P. 2018. Facets of prefabrication. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

Sinclair, J. 1991. Corpus, concordance, collocation. Oxford: Oxford University Press.

Wray, A. 2002. Formulaic language and the lexicon. Cambridge: CUP.

Zgusta, L. 1971. Manual of Lexicography. Prague: Czechoslovak Academy of Sciences.