Given Best EN→CS systems on par with humans at seg. level.

New This includes discourse connectives and their AltLexes.

New Information structure (subjects in focus) also preserved.

New "Doc-level" systems not any better than seg-level ones.

(IAA: 66-93%)

A Test Suite and Manual Evaluation of Document-Level NMT at WMT19

Kateřina Rysová, Magdaléna Rysová, Tomáš Musil, Lucie Poláková and Ondřej Bojar

Charles University, Faculty of Mathematics and Physics Institute of Formal and Applied Linguistics Prague, Czech Republic



Introduction

- We provided a test suite for WMT19 aimed at assessing discourse phenomena of English→Czech MT systems participating in the News Translation Task.
- We carried out an extensive manual annotation of the MT outputs and identified types of translation errors relevant to document-level translation.
- We also compared the systems' performance with respect to the observed phenomena.

Manual Evaluation

President Bush and Soviet leader Mikhail Gorbachev will hold an informal meeting in early December, a move that should give both leaders a political boost at home.	Prezident Bush a sovětský vůd setkání, které by mělo dát obě		lají na začátku prosince neformální ru doma.	
1) The White House is purposely not calling the meeting a summit so that there won't be any expectation of detailed negotiations or agreements \triangleleft 1.	Bilý dům úmyslně neříká schůzku summitem, takže nebude žádné očekávání podrobných jednání či dohod.			
Rather, senior administration officials said ₁▶ that the unexpected meeting was scheduled at Mr. Bush's request because of his preference for conducting diplomacy through highly personal and informal meetings with other leaders ◄ 1.	Úředníci vyšších správních úřadů uvedli, že nečekané setkání bylo naplánováno na žádost pana Bushe kvůli jeho preferenci vést diplomacii prostřednictvím vysoce osobních a neformálních setkání s ostatními vůdci.			
	1 Exp Expansion.Substitutio	on.Arg2-as-subst		
	Původní vztah: O An	10 O Ne	poznámky	
	Zachován v překladu: O And	10 O Ne		
2 > The two leaders will meet on Dec. 2 and 3, alternating the two days of meetings between a U.S. and a Soviet naval vessel in the Mediterranean Sea \triangleleft 2.	Oba vůdci se sejdou 2. a 3. pr námořním plavidlem ve Středo		ůzek mezi USA a sovětským	
2▶ The unusual seaborne meeting won't disrupt plans for a formal summit meeting next spring or summer, at which an arms-control treaty is likely to be completed ◄ .				
	2 Ent			
	Původní vztah: O An	no ○ Ne	poznámky	
	Zachován v překladu: O An	no ○Ne		
In announcing the meeting yesterday, Mr. Bush told reporters at the White House that neither he nor Mr. Gorbachev expects any "substantial decisions or agreements." Instead, he said that the purpose is simply for the two to get "better acquainted" and discuss a wide range of issues without a formal agenda.	"podstatná rozhodnutí nebo d	omě novinářům, že ani on, an lohody". Místo toho, on říkal, i o široké škále otázek, aniž by f	i pan Gorbačov neočekávají žádná že účelem je jednoduše, aby se dva formální agendu.	
3Despite the informal nature of the session and the calculated effort to hold down expectations, the meeting could pay significant political dividends for both leaders <3.	Navzdory neformální povaze z setkání mohli zúčastnit význar		žet se nad očekáváním by se na ba vůdce.	
$_3$ > $_4$ > Mr. Gorbachev badly needs a diversion from the serious economic problems and ethnic unrest he faces at home $_4$ 0 $_4$ 1.	Pan Gorbačov špatně potřebuj nepokojů, kterým čelí doma.	ije odklon od vážných ekonom	ických problémů a etnických	
	3 Imp Contingency.Cause.R	leason		
	Původní vztah: O An	no ○ Ne	poznámky	
	Zachován v překladu: O An	no ○Ne		
American officials have said ₄▶ that a meeting with the leader of the U.S. could help bolster his stature among Soviet politicians and academics, whose support he needs ❖ .	Američtí představitelé uvedli, mezi sovětskými politiky a aka	, že setkání s vůdcem USA by ademiky, jejichž podporu potře	mohlo pomoci posílit jeho postavení ebuje.	
	4 Imp Contingency.Cause.R	Result		
	Původní vztah:	no O Ne	poznámky	
	Zachován v překladu: O An	no ○Ne		
s⊳ For his part, Mr. Bush has been criticized regularly at home for moving too slowly and cautiously in reacting to Mr. Gorbachev's reforms and the historic moves away from communism in Eastern Europe <₅ .			pomalu a opatrně pohyboval v od komunismu ve východní Evropě.	
5► A face-to-face meeting with Mr. Gorbachev should damp such criticism, though it will hardly eliminate it ◄s.	Tváří v tvář osobnímu setkání sotva odstraní.	í s panem Gorbačovem by tato	kritika měla být vlhká, i když ji	

- most of the errors in the translation cannot be found by a simple comparison with the reference translation, a bilingual evaluation is needed
- AltLexes are "multi-word discourse connectors" such as for this reason or attributed that to
- vast majority of AltLexes translated correctly, across all systems
- most discourse connextors also translated correctly
- the information structure is also mostly correct

Conclusion

- most errors (equally across systems) were detected in the area of word order and contextual (non-)boundness of the subjects
- for the observed phenomena, the systems performed with only a minor differences among each other and they reached the quality of the reference
- the two context-aware systems did not outperform the others in translating the followed document-level phenomena
- strong sentence-level systems seem to be performing slightly better than context-aware ones

Annotation Results

AltLexes

	adequate	missing	wrong
CUNI-Transf-2018	***	公公公公公	
CUNI-DocTransf-Marian	*****		
online-B	****		★☆☆☆☆
CUNI-DocTransf-T2T	****	公公公公公	★☆☆☆☆
CUNI-Transf-2019	****	公公公公公	★☆☆☆☆
reference	*****		★☆☆☆☆

Discourse Connectors

	adequate	wrong place	inadequate
CUNI-Transf-2018	***		
CUNI-DocTransf-Marian	****		
online-B	****		公公公公公
CUNI-DocTransf-T2T	***		公公公公公
CUNI-Transf-2019	***		公公公公公
reference	****		

Information Structure

	context		rhema		placement		
	yes	no	yes	no	yes	no	
CUNI-Transf-2018	11	1	2	0	14	0	•
CUNI-DocTransf-Marian	17	3	5	3	14	5	
online-B	6	1	0	1	3	1	
CUNI-DocTransf-T2T	17	1	1	2	13	3	
CUNI-Transf-2019	6	1	0	1	6	O	
reference	19	4	1	6	19	3	

context – contextually non-bound subject in both English and Czech
rhema – subject is the focus proper in both English and Czech
placement – correct placement according to Czech word order rules



We acknowledge support from the Czech Science Foundation project no. GA17-06123S (Anaphoricity in Connectives: Lexical Description and Bilingual Corpus Analysis) and the EU project H2020-ICT-2018-2-825460 (ELITR). This study has utilized language resources distributed by the LINDAT/CLARIN project of the Ministry of Education, Youth and Sports of the Czech Republic (project LM2015071). This research was partially supported by SVV project number 260 453.