Given Best EN $\rightarrow$ CS systems on par with humans at seg. level. New This includes discourse connectives and their AltLexes. Information structure (subjects in focus) also preserved. New New "Doc-level" systems not any better than seg-level ones.

(IAA: 66–93%)

## **A Test Suite and Manual Evaluation**

**Annotation Results** 

# 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 $\rightarrow$ 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**



#### AltLexes

CUNI-Transf-2018 CUNI-DocTransf-Marian online-B CUNI-DocTransf-T2T CUNI-Transf-2019 reference

adequate missing wrong \*\*\*\*\* किंक्रिक्रेक्रिक 

### **Discourse Connectors**

CUNI-Transf-2018 CUNI-DocTransf-Marian online-B CUNI-DocTransf-T2T CUNI-Transf-2019

inadequate wrong place adequate क्रिक्रेक्रेक्रे 

	Prezident Bush a sovětský vůdce Michail Gorbačov uspořádají na začátku prosince neformáln setkání, které by mělo dát oběma vůdcům politickou podporu doma.				
	Bílý dům úmyslně neříká schůzku summitem, takže nebude žádné očekávání podrobných jednání či dohod.				
Bush's request because of his preference for conducting diplomacy through highly		etkání bylo naplánováno na žádost ednictvím vysoce osobních a			
1	1 Exp Expansion.Subs	stitution.Arg2-as-subst			
	Původní vztah:	🔿 Ano 🔿 Ne	poznámky		
	Zachován v překladu:	○ Ano ○ Ne			
The two leaders will meet on Dec. 2 and 3, alternating the two days of meetings between a ( and a Soviet naval vessel in the Mediterranean Sea $<\!\!2^2$ .	Oba vůdci se sejdou 2. námořním plavidlem ve	a 3. prosince a střídají dva dny s Středozemním moři.	chůzek mezi USA a sovětským		
	Neobvyklé námořní setkání nenaruší plány na formální summit na jaře nebo v létě, na kt bude pravděpodobně uzavřena smlouva o kontrole zbraní.				
	2 Ent				
	Původní vztah:	○ Ano ○ Ne	poznámky		
	Zachován v překladu:	○ Ano ○ Ne			
nnouncing the meeting yesterday, Mr. Bush told reporters at the White House that neither or Mr. Gorbachev expects any "substantial decisions or agreements." Instead, he said that purpose is simply for the two to get "better acquainted" and discuss a wide range of issues out a formal agenda.	Bush oznámil včera v B "podstatná rozhodnutí i "lépe seznámit" a disku	ílém domě novinářům, že ani on, nebo dohody". Místo toho, on řík tovat o široké škále otázek, aniž b	ani pan Gorbačov neočekávají žádn al, že účelem je jednoduše, aby se d y formální agendu.		
Despite the informal nature of the session and the calculated effort to hold down to hold down the calculated effort to hold down to h	Navzdory neformální po setkání mohli zúčastnit	ovaze zasedání a vypočtené snaze významné politické dividendy pr	držet se nad očekáváním by se na o oba vůdce.		
	Pan Gorbačov špatně po nepokojů, kterým čelí d	otřebuje odklon od vážných ekono oma.	omických problémů a etnických		
	3 Imp Contingency.Ca	use.Reason			
		0 0	poznámky		
	Původní vztah:	○ Ano ○ Ne			
	Původní vztah: Zachován v překladu:				
rican officials have said 4> that a meeting with the leader of the U.S. could help bolster	Zachován v překladu: Američtí představitelé u	○ Ano ○ Ne	by mohlo pomoci posílit jeho postav třebuje.		
rican officials have said 4> that a meeting with the leader of the U.S. could help bolster 4 stature among Soviet politicians and academics, whose support he needs 4 .	Zachován v překladu: Američtí představitelé u	○ Ano ○ Ne uvedli, že setkání s vůdcem USA i y a akademiky, jejichž podporu po			
rican officials have said 4> that a meeting with the leader of the U.S. could help bolster 4 stature among Soviet politicians and academics, whose support he needs 4 .	Zachován v překladu: Američtí představitelé v mezi sovětskými politiky	○ Ano ○ Ne uvedli, že setkání s vůdcem USA i y a akademiky, jejichž podporu po			
erican officials have said 4> that a meeting with the leader of the U.S. could help bolster A stature among Soviet politicians and academics, whose support he needs 4.	Zachován v překladu: Američtí představitelé t mezi sovětskými politiky 4 Imp Contingency.Ca	Ano Ne uvedli, že setkání s vůdcem USA y a akademiky, jejichž podporu po use.Result Ano Ne	střebuje.		
rican officials have said 4> that a meeting with the leader of the U.S. could help bolster 4 stature among Soviet politicians and academics, whose support he needs 4.	Zachován v překladu: Američtí představitelé u mezi sovětskými politiky 4 Imp Contingency.Ca Původní vztah: Zachován v překladu: Pan Bush byl pravidelní	<ul> <li>Ano Ne</li> <li>vvedli, že setkání s vůdcem USA i y a akademiky, jejichž podporu po nuse.Result</li> <li>Ano Ne</li> <li>Ano Ne</li> <li>Ano Ne</li> <li>ě krjitizován doma za to, že se příl</li> </ul>	střebuje.		
rican officials have said ₄► that a meeting with the leader of the U.S. could help bolster stature among Soviet politicians and academics, whose support he needs ◀ .	Zachován v překladu: Američtí představitelé u mezi sovětskými politiky 4 Imp Contingency.Ca Původní vztah: Zachován v překladu: Pan Bush byl pravidelní	<ul> <li>Ano Ne</li> <li>vvedli, že setkání s vůdcem USA i y a akademiky, jejichž podporu po nuse.Result</li> <li>Ano Ne</li> <li>Ano Ne</li> <li>Ano Ne</li> <li>ě krjitizován doma za to, že se příl</li> </ul>	vřebuje.		

- 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

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	0
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

• strong sentence-level systems seem to be performing slightly better than context-aware ones



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.

Presented on August 1, 2019 at WMT19, collocated with ACL 2019 in Florence, Italy.