

OVERVIEW

- **How can we generate text in the right style?**
- **Our approach:**
 - A seq2seq model inspired from advances in Neural Machine Translation (NMT)
 - that jointly performs monolingual Formality Transfer [1] and cross-lingual Formality-Sensitive MT [2]
 - and leverages two types of supervision:
 - (1) sentence pairs with different styles in the same language, and
 - (2) translation pairs drawn from corpora of diverse style.
- **Findings:**
 - Formality Transfer: the joint model significantly improves formality transfer in both directions.
 - Formality-Sensitive Machine Translation: the joint model performs well without being trained on style-annotated translation examples.

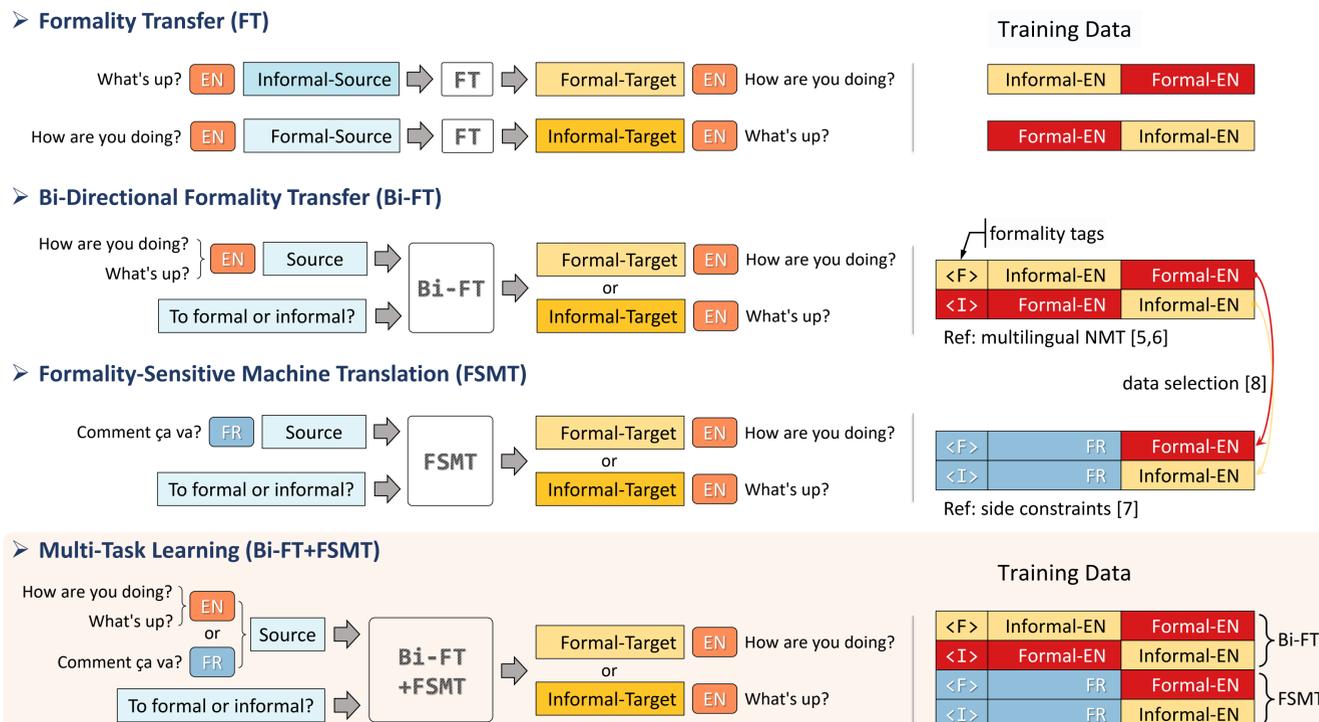
RESULTS – ① FORMALITY TRANSFER

Model	I→F		F→I	
	E&M	F&R	E&M	F&R
PBMT [1]	68.22	72.94	33.54	32.64
NMT Baseline [1]	58.80	68.28	30.57	36.71
NMT Combined [1]	68.41	74.22	33.56	35.03
FT	65.34	71.28	32.36	36.23
Bi-FT	66.30	71.97	34.00	36.33
Bi-FT (enhanced) + FSMT	71.36	74.49	36.18	38.34
	72.13	75.37	38.04	39.09

I: Informal / F: Formal
enhanced: training on E&M + F&R and ensemble decoding

- **Automatic Evaluation (BLEU):**
 - Bi-directional models perform better (FT vs. Bi-FT).
 - Multi-task learning yields further improvement (+FSMT).
- **Human Evaluation:**
 - Our model generates slightly more formal English and significantly more informal English than NMT Combined.

TASKS AND APPROACHES



RESULTS – ② FORMALITY SENSITIVE MACHINE TRANSLATION

Model	Training Data	BLEU		Formality Diff Range = [0,2]	Meaning Preservation Range = [0,3]	Comment
		FR-> Formal-EN	FR-> Informal-EN			
NMT 1	<F> Informal-EN Formal-EN	27.15	26.70	0.35	2.95	• English is selected by comparing to GYAFC data.
	<I> Formal-EN Informal-EN					
NMT 2	<F> Informal-EN Formal-EN	25.02	25.20	--	--	• The FR-EN data is the same as above. • GYAFC ≠ domain of translation data • ☺ for FT, ☹ for FSMT
	<I> Formal-EN Informal-EN					
	<F> FR Formal-EN					
NMT 3	<I> FR Informal-EN	25.24	25.14	0.32	2.90	• Tag embeddings and the attention are already learned from the informal-formal parallel data. • The FR-EN data is randomly selected.
	<I> Formal-EN Informal-EN					
PBMT	FR EN	29.12	29.02	0.05	2.97	• The FR-EN data is the same as above. • Re-ranking based FSMT [2]

- Neural models control formality significantly better than PBMT.
- The multi-task model perform well without data selection.

QUALITATIVE ANALYSIS – ① FORMALITY TRANSFER

informal (I) -> formal (F)			
Original (I)	chill out	sweetie everything will be just fine eventually	☺ Our model produces more formal/informal output by introducing phrasal level changes,
# 1 NMT Combined (F)	Can you chill out	sweetie everything will be just fine eventually.	➤ moving phrases around,
Bi-FT+FSMT (F)	Calm down,	sweetie, everything will be fine eventually.	➤ or making the output more complete (apply to formal).
Original (I)	Dakota Fanning.....	I know that she is only 12 but she is really famous.	
# 2 NMT Combined (F)	Dakota Fanning.i	know that she is only twelve, but she is famous.	
Bi-FT+FSMT (F)	I know that	Dakota Fanning is only twelve, but she is really famous.	
formal (F) -> informal (I)			
Original (F)	Therefore I would say that they do succeed	but not frequently.	I hope this is helpful.
# 3 NMT Combined (I)	So I would say that they do failing but not frequently,	I hope this is helps.	☹ Our model sometimes fails to preserve the meaning of the source.
Bi-FT+FSMT (I)	so i would say they do it	but not all the time,	hope this helps.
Original (F)	I am simply inquiring because people behave as though they	are no longer interested in them.	
# 4 NMT Combined (I)	I am just asking because people act as though they are no longer interested in them.		
Bi-FT+FSMT (I)	I'm just asking because people act like they	don't like them anymore.	

QUALITATIVE ANALYSIS – ② FORMALITY SENSITIVE MACHINE TRANSLATION

# 1	Reference	Refrain from the commentary and respond to the question, Chief Toohey.	
	NMT-3	You need to be quiet and answer the question, Chief Toohey.	☺ The multi-task model (NMT-3) in general makes very large changes while transferring the style, especially into informal.
Formal	NMT-1	Please refrain from any comment and answer the question, Chief Toohey.	
	PBMT	Please refrain from comment and just answer the question, the Tooheys's boss.	
	NMT-3	Shut up and answer the question, Chief Toohey.	
Informal	NMT-1	Please refrain from comment and answer the question, chief Toohey.	☹ In trying to make changes, NMT-3 occasionally change not only the style but also the meaning of the input.
	PBMT	Please refrain from comment and answer my question, Tooheys's boss.	
# 2	Reference	Try to file any additional motions as soon as you can.	
	NMT-3	You should try to introduce the sharks as soon as you can.	
Formal	NMT-1	Try to present additional requests as soon as you can.	
	PBMT	Try to introduce any additional requests as soon as you can.	➤ The re-ranking model (PBMT) usually generates very similar formal and informal translations.
	NMT-3	Try to introduce sharks as soon as you can.	
Informal	NMT-1	Try to introduce extra requests as soon as you can.	
	PBMT	Try to introduce any additional requests as soon as you can.	

DATA

- **FT data – GYAFC [1]:**
 - Grammarly's Yahoo Answers Formality Corpus (2 domains)
 - Entertainment & Music, Family & Relationships
 - 100K+ informal sentences from Yahoo Answers
 - Rewritten in a formal style via crowd-sourcing
 - 100K training / 10K dev / 5K test
- **FSMT data – OpenSubtitles2016 [3]:**
 - 20M+ least divergent French-English examples [4]
 - 20M training pool / 2.5K dev / 2.5K test

REFERENCES

- [1] Sudha Rao and Joel R. Tetreault. 2018. Dear sir or madam, may I introduce the GYAFC dataset: Corpus, benchmarks and metrics for formality style transfer. In NAACL-HLT.
- [2] Xing Niu, Marianna J. Martindale, and Marine Carpuat. 2017. A study of style in machine translation: Controlling the formality of machine translation output. In EMNLP
- [3] Pierre Lison and Jörg Tiedemann. 2016. Opensubtitles2016: Extracting large parallel corpora from movie and TV subtitles. In LREC.
- [4] Yogarshi Vyas, Xing Niu, and Marine Carpuat. 2018. Identifying semantic divergences in parallel text without annotations. In NAACL-HLT
- [5] Melvin Johnson et al. 2017. Google's multilingual neural machine translation system: Enabling zero-shot translation. TACL.
- [6] Xing Niu, Michael Denkowski, and Marine Carpuat. 2018. Bi-directional neural machine translation with synthetic parallel data. In NMT@ACL.
- [7] Rico Sennrich, Barry Haddow, and Alexandra Birch. 2016a. Controlling politeness in neural machine translation via side constraints. In HLT-NAACL
- [8] Robert C. Moore and William D. Lewis. 2010. Intelligent selection of language model training data. In ACL.