Sec 4.3 Mitigating error — Gender bias

Gender and data







Gender and data





Gender and data



- ~ 70% of the TED speakers is male
- Most of the ASR and MT data are generated by male speakers



Gender and translation

How do languages convey the gender of a referred entity?

I'm a good friend

English: Natural Gender Language

- Pronouns (he/she)
- Lexical gender (boy/girl)
- Gender-marked titles (actor/actress)

she is a good friend he is a good friend

Italian/French: **Grammatical Gender Languages**

• Overtly express feminine/masculine gender on numerous POS

è una buona amica (Fem.) è un_ buon_ amico » (Masc.)

Gender bias: a technical and ethical problem

"I'm a good friend"	Correct Italian translation	Most probable au
M: "Sono un_ buon_ amic <u>o</u> "		V
F: "Sono un <u>a</u> buon <u>a</u> amic <u>a</u> "		

Itomatic translation



Gender bias: a technical and ethical problem

"I'm a good friend"	Correct Italian translation	Most probable au
M: "Sono un_ buon_ amic <u>o</u> "		\checkmark
F: "Sono un <u>a</u> buon <u>a</u> amic <u>a</u> "		tho she
	Independ	lently from the spe

Itomatic translation



Gender bias: a technical and ethical problem

"I'm a good friend"	Correct Italian translation	Most probable au	
M: "Sono un_ buon_ amic <u>o</u> "		V	
F: "Sono un <u>a</u> buon <u>a</u> amic <u>a</u> "		tho she	
	Indonend	Independently from the sp	



Bias in the training data...

...pushes systems towards a "male default"... ...amplifying social asymmetries!

Itomatic translation





Gender bias and automatic translation

Machine Translation (text-to-text)
→ textual input does NOT always provide gender clues

Speech Translation (speech-to-text)
→ audio input can provide gender clues



Are ST systems able to exploit audio information to translate gender?

Gender bias and ST - exploiting audio features

- Bentivogli et al., "Gender in Danger? Evaluating Speech Translation Technology on the *MuST-SHE Corpus*", ACL 2020
 - MuST-SHE: a benchmark for the analysis of gender translation in MT and ST Ο

- **Derived from MuST-C** (2 language directions $En \rightarrow It$, $En \rightarrow Fr$)
- Gender-sensitive design: each segment contains 1+ English gender-neutral word translated into the corresponding masculine/feminine target word(s)
- **2 gender phenomena**: info-in-audio (*I'm a good friend*), info-in-content (*she is a good...*)

Gender bias and ST - exploiting audio features

- Bentivogli et al., "Gender in Danger? Evaluating Speech Translation Technology on the *MuST-SHE Corpus*", ACL 2020
 - MuST-SHE: a benchmark for the analysis of gender translation in MT and ST \bigcirc
 - Gender-sensitive evaluation methodology based on "gender swapping" Ο

- BLEU/Accuracy scores computed against **correct** and **wrong** references
 - Src: *I have been to London* (female speaker)
 - C-Ref: *Io sono stat<u>a</u> a Londra*,
 - W-Ref: *Io sono stato a Londra* \bigcirc
- Difference between correct and wrong reference as a measure of gender translation performance (the higher the better -- lower bias!)

Gender bias and ST - exploiting audio features

- Bentivogli et al., "Gender in Danger? Evaluating Speech Translation Technology on the *MuST-SHE Corpus*", ACL 2020
 - MuST-SHE: a benchmark for the analysis of gender translation in MT and ST \bigcirc
 - Gender-sensitive evaluation methodology based on "gender swapping" \bigcirc
 - **Comparison between end-to-end and cascade ST approaches** Ο
- Translation quality (BLEU): cascade better than e2e
- Gender translation (BLEU+gender swapping): the two perform on par
- Gender translation (Accuracy+gender swapping) on info-in-audio samples:
 - e2e much better than simple cascade \bigcirc
 - leveraging audio features sethical issues (vocally impaired, transgender)?

Gender bias and ST - exploiting speakers' info

- Gaido et al., "Breeding Gender-aware Direct Speech Translation Systems", Coling 2020
 - MuST-Speakers: annotation of MuST-SHE with speakers' gender information

tems", Coling 2020 ander information

Gender bias and ST - exploiting speakers' info

- Gaido et al., "Breeding Gender-aware Direct Speech Translation Systems", Coling 2020
 - MuST-Speakers: annotation of MuST-SHE with speakers' gender information \bigcirc
 - **Comparison of different e2e ST systems** Ο
- **Base**: Generic, "gender-unaware" ST model
- **Multi-gender**: single model informed of the speaker's gender via pre-pended gender tokens
- **Gender-specialized**: two models, fine-tuned on utterances spoken by men/women
- Overall translation quality (BLEU): small differences
- Gender translation (Accuracy+gender swapping) on info-in-audio samples (*I'm a good friend*):
 - **Specialized >> Multi-gender >> Base** Ο