

Formality Style Transfer (FoST)

Task definition: Generate a well-formed sentence that matches a desired formality attribute while preserving the meaning of the input

Example

INFORMAL Gotta see both sides of the story

FORMAL You have to consider both sides of the story

English (EN) resources:
Grammarly's Yahoo Formality Corpus (GYAFC) [1]

This work: How well can we perform FoST within different languages? -> **First work on multilingual ST!**

- Introducing XFORMAL: Benchmark for FoST in 3 languages
 - Brazilian Portuguese (BR-PT), Italian (IT), & French (FR)
 - 1000 informal segments paired with 4 formal rewrites
- Benchmark FoST models (8): Simple baselines (3), NMT-based approaches (3), & unsupervised models (2).
 - Automatic & Human evaluation on 3 dimensions: meaning preservation, formality transfer, & fluency

XFORMAL

Curation rationale

For each language:

- L6 Yahoo! Answers: Family & Relationship domain
- Pre-processing
- Detect informal answers
- Randomly sample 1000 answers for annotation

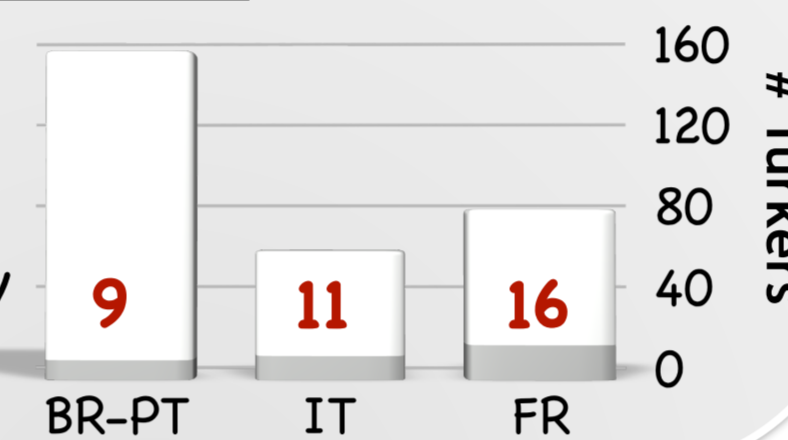
Procedures

Amazon Mechanical Turk workers (Turkers)

Given an informal excerpt generate its formal rewrite in the same language without changing its meaning

Quality Control

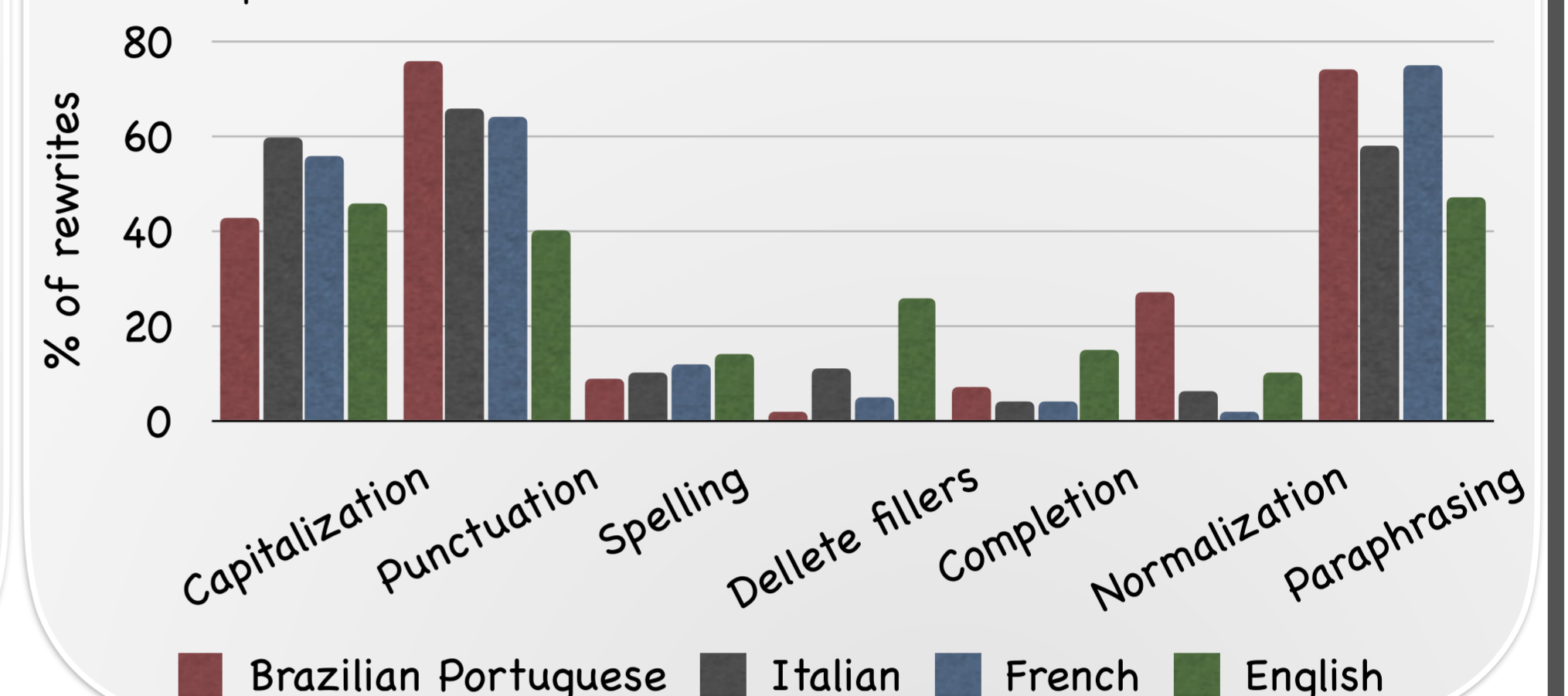
- Location Restrictions
- Qualification test
- Filtering by pilot study



Dataset analysis

Frequent edit operations across languages

- Consistent trends across language: edits cover the "noisy-text" sense of formality
- Delete fillers and completion-based edits are more frequent for EN
- Normalization-based edits are more frequent for BR-PT
- Paraphrase-based edits are more frequent for BR-PT, IT, FR compared to EN



Benchmarking Multilingual FoST

Simple baselines

Copy: informal input

Input: *n preciso pedir pois sei q ela vai vir atras!!*
Output: *n preciso pedir pois sei q ela vai vir atras!!*

Rule-based: hand-crafted rules

Input: *n preciso pedir pois sei q ela vai vir atras!!*
Output: *n preciso pedir pois sei q ela vai vir atras!!*
Fix casing, Normalize punctuation, Expand contractions etc.

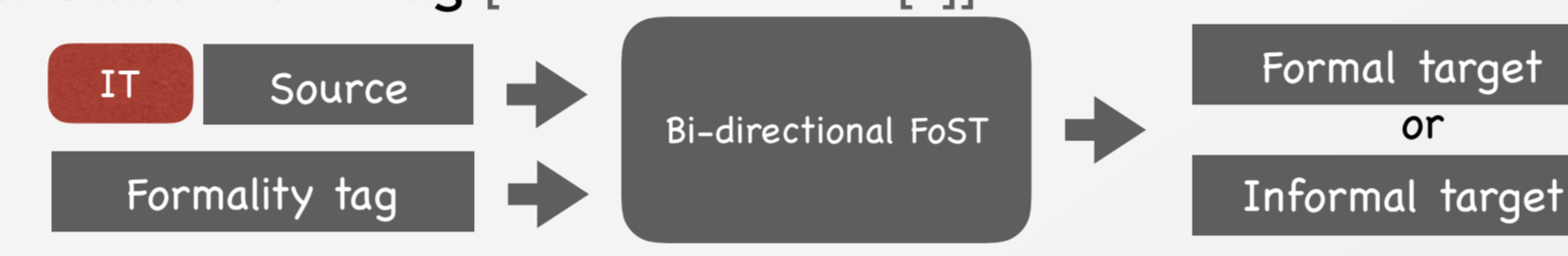
Round-trip MT: pivot to En & backtranslate

Input: *n preciso pedir pois sei q ela vai vir atras!!*
Pivot: *I don't need to ask because I know she's going to come back!!*
Output: *n preciso perguntar porque sei que ela virá atrás de mim!*

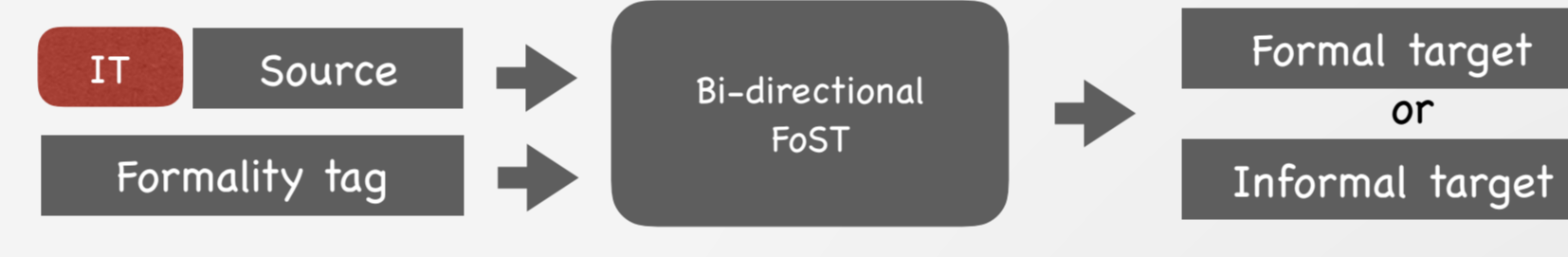
Formalization effect of MT: AWS translation service

NMT-based approaches

Translate Train Tag [TRANSLATE-TRAIN & [2]]



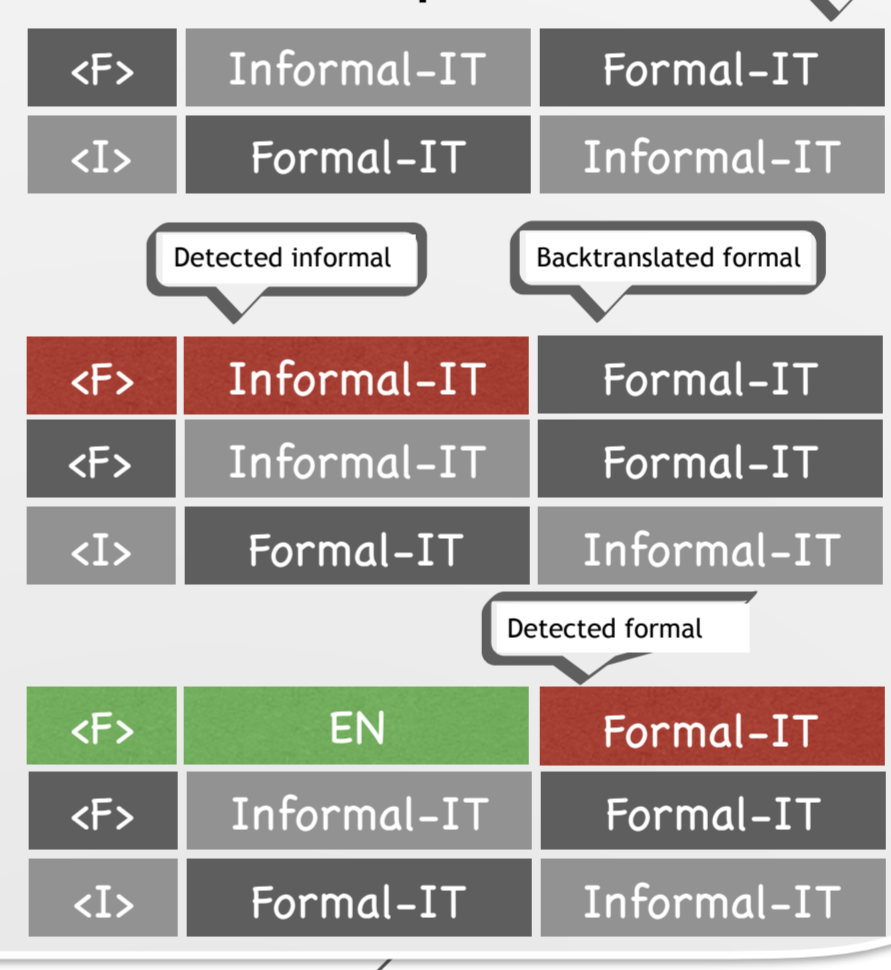
Backtranslate [TRANSLATE-TRAIN [3]]



Multi-task tag [TRANSLATE-TRAIN & [2]]



Pseudo-parallel data

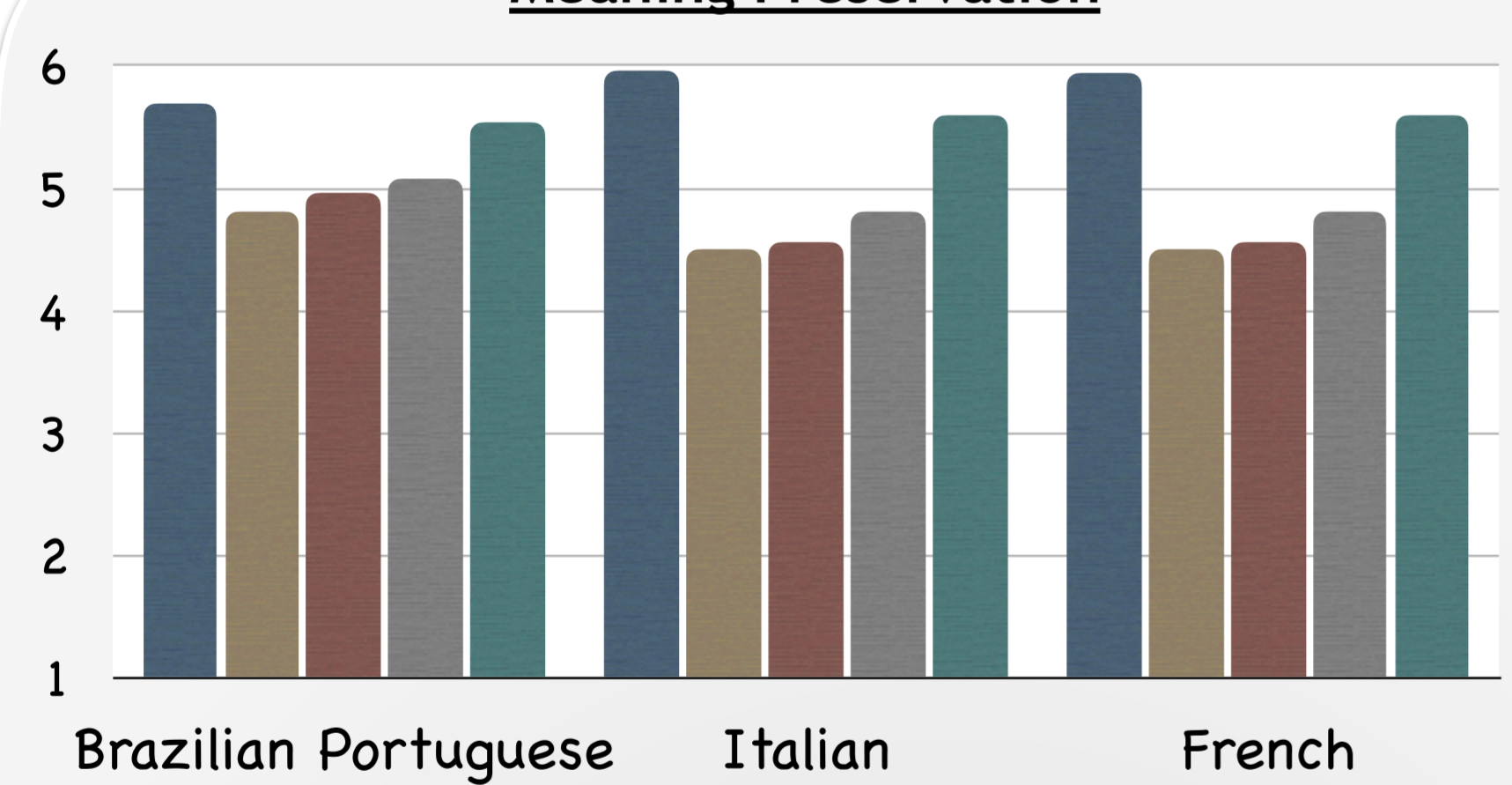


Training Data

- Machine-translated GYAFC
 - GYAFC translated through AWS
 - Family & Relationship
 - 100K+ informal-formal pairs
- Data augmentation
 - OpenSubtitles [4] (En-x)
 - Filter out noise w/ Bicleaner [5]
 - Formal vs. Informal (target side): mBERT regression [6] & formality annotations [7]
- Backtranslate
 - GYAFC [1] translated through AWS + 2M backtranslated informal
- Multi-task tag
 - GYAFC [1] translated through AWS + 2M En-Formal-It

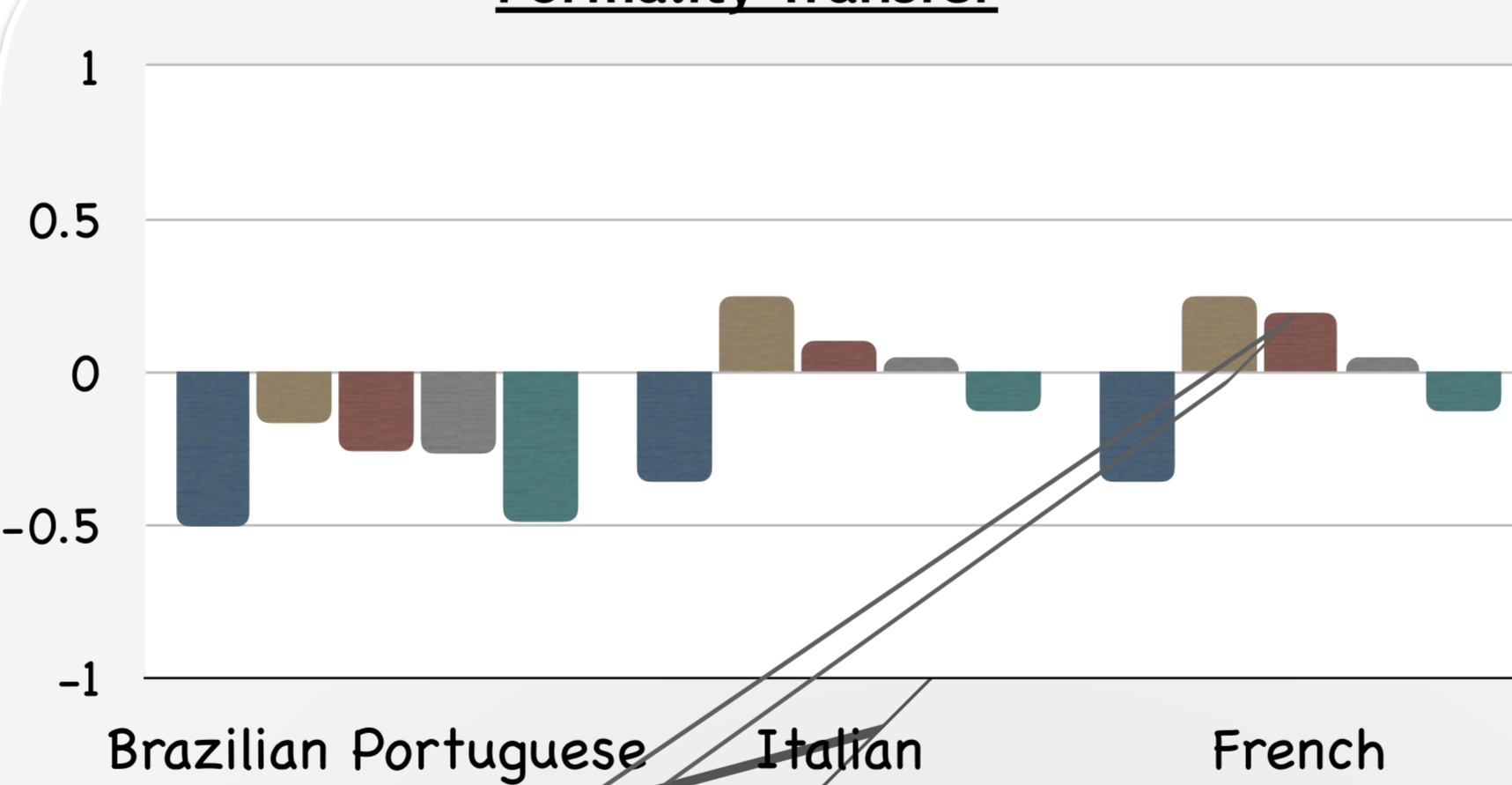
Human Evaluation Results

Meaning Preservation



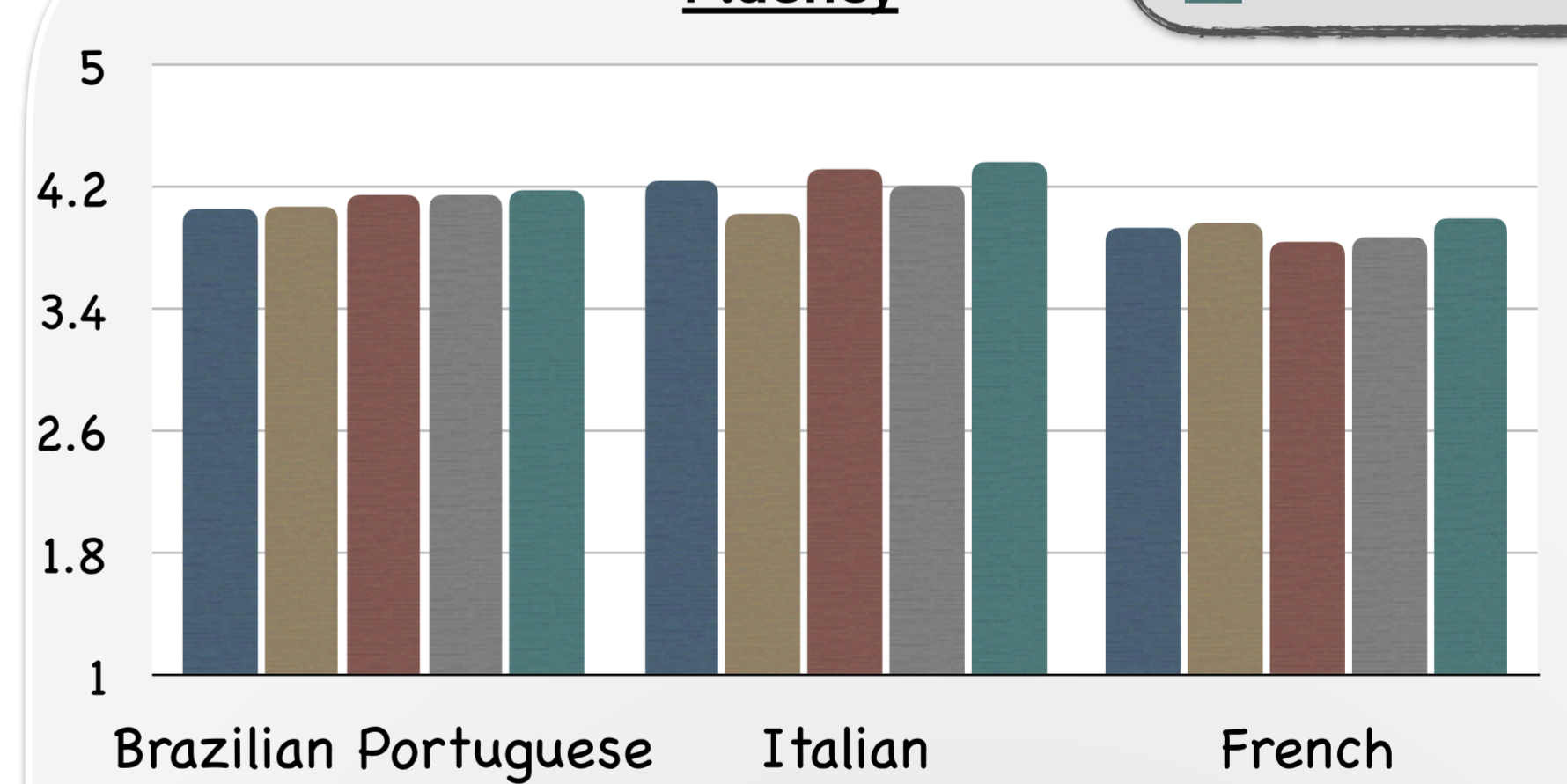
Outputs are meaning preserving on average

Formality Transfer



Outputs are concentrated around neutral formality levels

Fluency



Outputs are comprehensible on average

REFERENCES

- 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.
- Xing Niu, Sudha Rao, Marine Carpuat. 2018. Multi-Task Neural Models for Translating Between Styles Within and Across Languages. In COLING.
- Rico Sennrich, Barry Haddow, Alexandra Birch. 2016. Improving Neural Machine Translation Models with Monolingual Data. In ACL.
- Pierre Lison and Jörg Tiedemann. 2016. Opensubtitles2016: Extracting large parallel corpora from movie and TV subtitles. In LREC.
- Gema Ramírez-Sánchez, Jaume Zaragoza-Bernabeu, Marta Bañón, Sergio Ortiz Rojas. 2020. Bifixer and Bicleaner: two open-source tools to clean your parallel data. In EAMT.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In NAACL.
- Elie Pavlick, Joel Tetreault. 2016. An Empirical Analysis of Formality in Online Communication. In TACL.

SUMMARY

- Introduce XFORMAL
 - Evaluation dataset of informal-formal pairs in FR, IT, BR-PT
 - MTurk & multiple levels of quality control
- Benchmarking of FoST
 - conservative edits on the informal input
 - simple baselines perform comparable to advanced models
- Future work should look at...
 - models that do not heavily rely on supervised data
 - automatic evaluation methods that generalize beyond English

Request XFORMAL at: <https://github.com/Elbria/xformal-FoST>