Our Methods

We thank Garrett Nicolai for the assistance with DTLM.

**Tools**
- DTLM (Nicolai et al, 2018): Combines discriminative transduction with character and word language models.

**Data**
- Grapheme-phoneme pairs in 15 languages. **LR setting**: 100 random pairs per language, **HR setting**: Full Dataset.
- Augmentation data: Unannotated Wikipedia corpus.

**Baselines**
- FST: Finite State Transducer, tuned on the size of n-grams.
- Transformer: Encoder-decoder architecture with self-attention, which requires extensive parameter tuning.

Introduction

**Grapheme to Phoneme (G2P)**

- *inconvénient*
- η̃σ̃ν̃j̃

**Phoneme to Grapheme (P2G)**

- inconvénient
- η̃σ̃ν̃j̃

→ 2020 (Task 1): Multilingual Grapheme-to-Phoneme (G2P) conversion.
→ Default setting: 3600 training instances.
→ We define a low-resource setting with 100 training instances.
→ We also perform Phoneme-to-Grapheme (P2G) conversion.

Key ideas: *Augment training data* using a combination of diverse models.

Discussion

**DTLM Ablation Results (Phoneme-to-Grapheme)**

- Demonstrates impact of Word Counts (WC) and character language models (LM) on P2G.
- WC helps more than target LMs.
- Without these two components, DTLM results are in the same range as baselines.

**Grapheme-to-Phoneme Test Results**

- Average WER (HR Setting)
  - DWLM: 18.8
  - Transformer: 17.6
  - Transformer+: 18.9

- Average WER (LR Setting)
  - DWLM: 93.8
  - Transformer: 89.1
  - Transformer+: 89.4

**Phoneme-to-Grapheme Test Results**

- Average WER (HR Setting)
  - DTLM: 18.5
  - Transformer: 17.7
  - FST: 18.4

- Average WER (LR Setting)
  - DTLM: 84.2
  - Transformer: 85.0
  - FST: 85.8

→ We introduce a method to synthesize additional training data from unannotated text.
1. Train FST and DTLM on for both G2P and P2G tasks.
2. Provide a word to the G2P models to produce phonemes.
3. Provide the phonemes to P2G models to produce graphemes.
4. Include the grapheme-phoneme pair in the new data if:
   a. The resulting graphemes for both FST and DTLM, match the initial word.
   b. The phonemes produced by FST and DTLM are the same.

**Results**

Tools & Data

**Data Augmentation**

- word list
- train data
- New data
- words
- unannotated
- new
- Chiel
- Discard
- Discard

→**DTLM**
  - japč
  - chiel

→**FST**
  - jal
  - FST
  - chal

Conclusion

- We proposed a novel data augmentation approach combining multiple string transduction methods.
- We explored both G2P and P2G tasks in both high-resource and low-resource settings.
- Our results demonstrate that the weakness of neural systems in low-resource settings can be mitigated through data augmentation.

Acknowledgments

- We thank Garrett Nicolai for the assistance with DTLM.
- We thank the shared task organizers for their effort.
- This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).