

AAACL - 2022

Phylogeny-Inspired Adaptation of Multilingual Models to New Languages

Fahim Faisal

Antonios Anastasopoulos

ffaisal@gmu.edu

<https://nlp.cs.gmu.edu/>

Going beyond the top-100 languages



Getting a LM for a new language

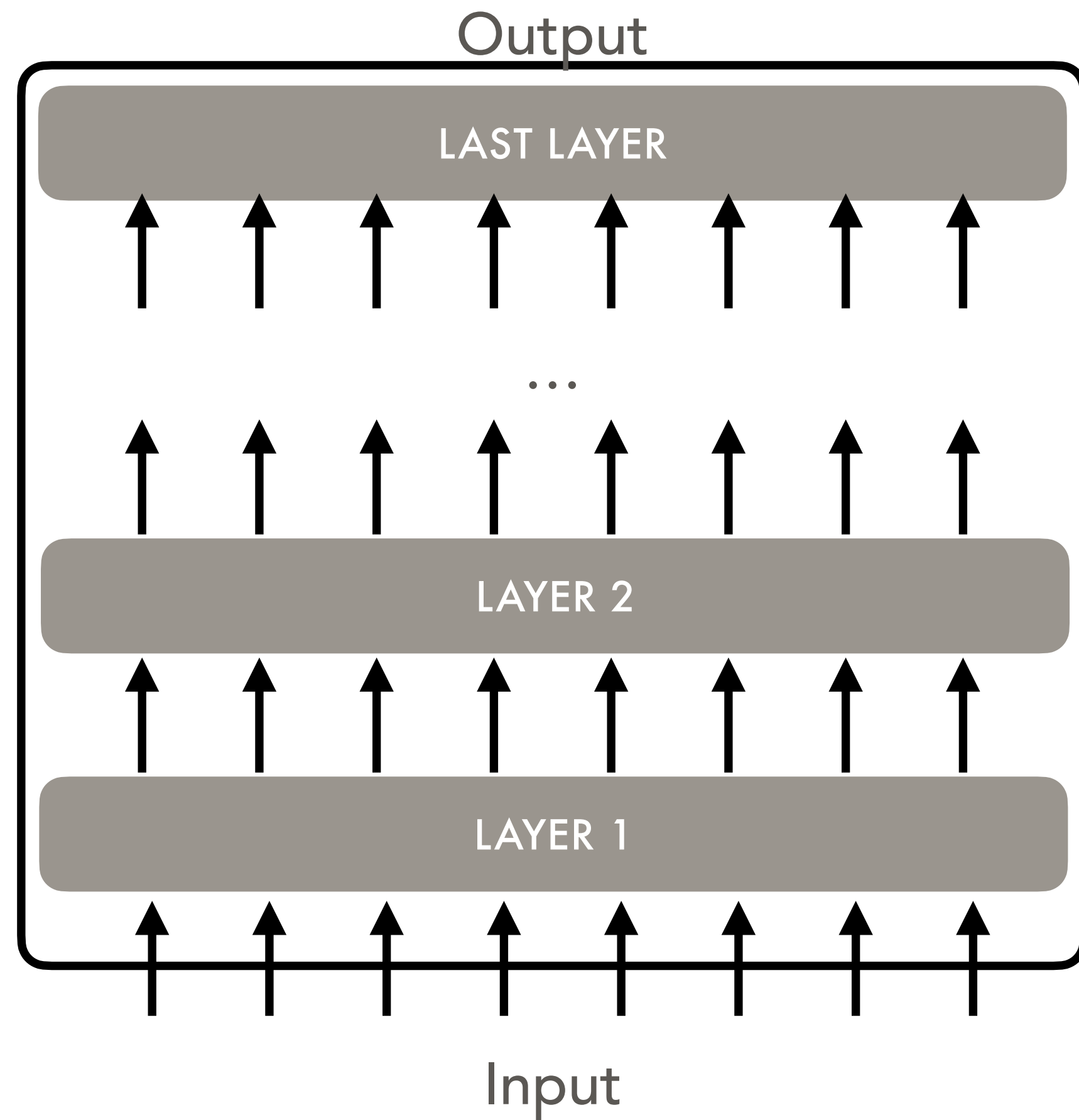
Option A: Train (a monolingual) one from scratch

Option B: Just use mBERT (zero-shot)

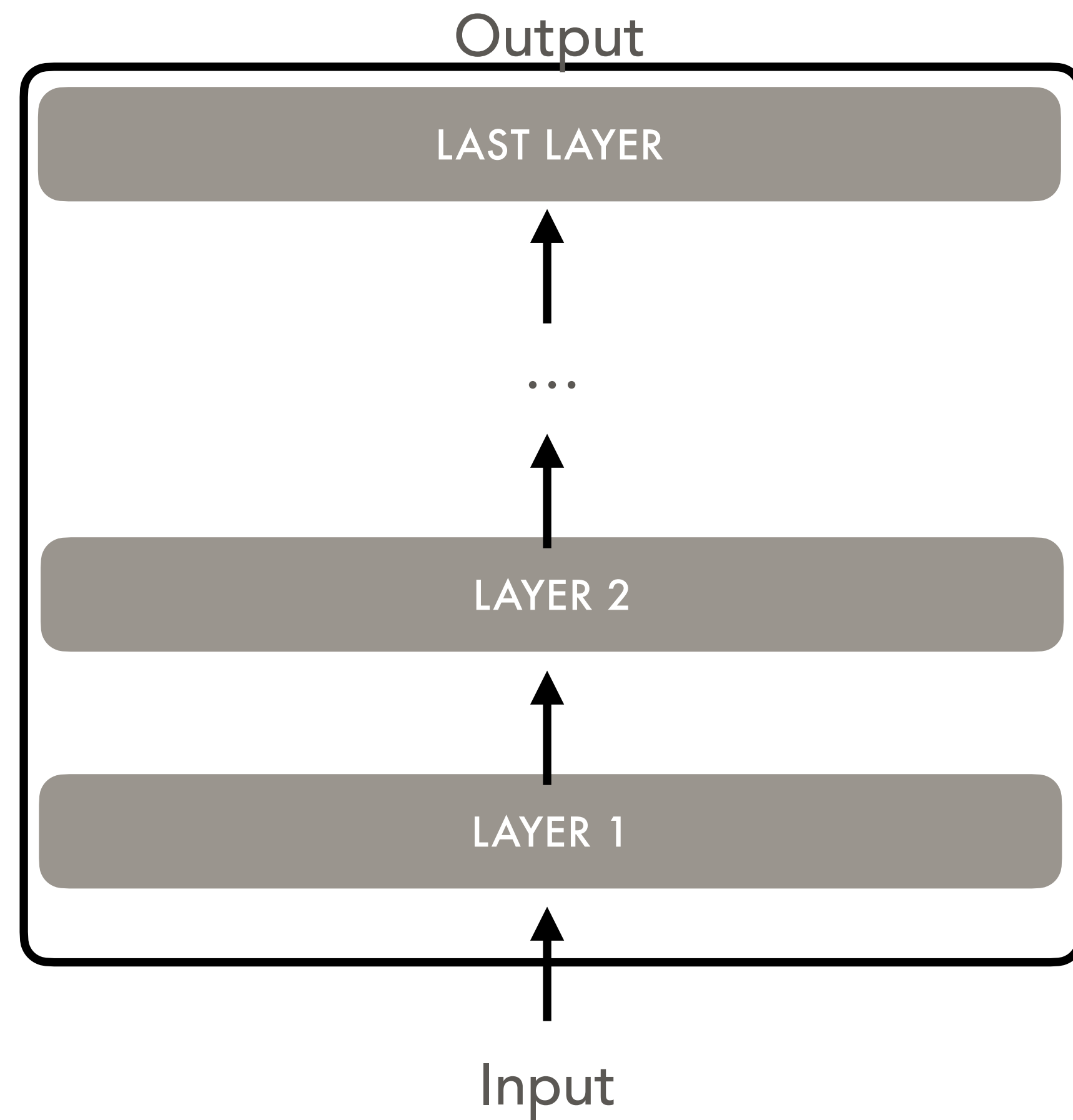
Option C: Continue training mBERT with same unsupervised objective

Option D: Adapters (*Pfeiffer et al. 2020*)

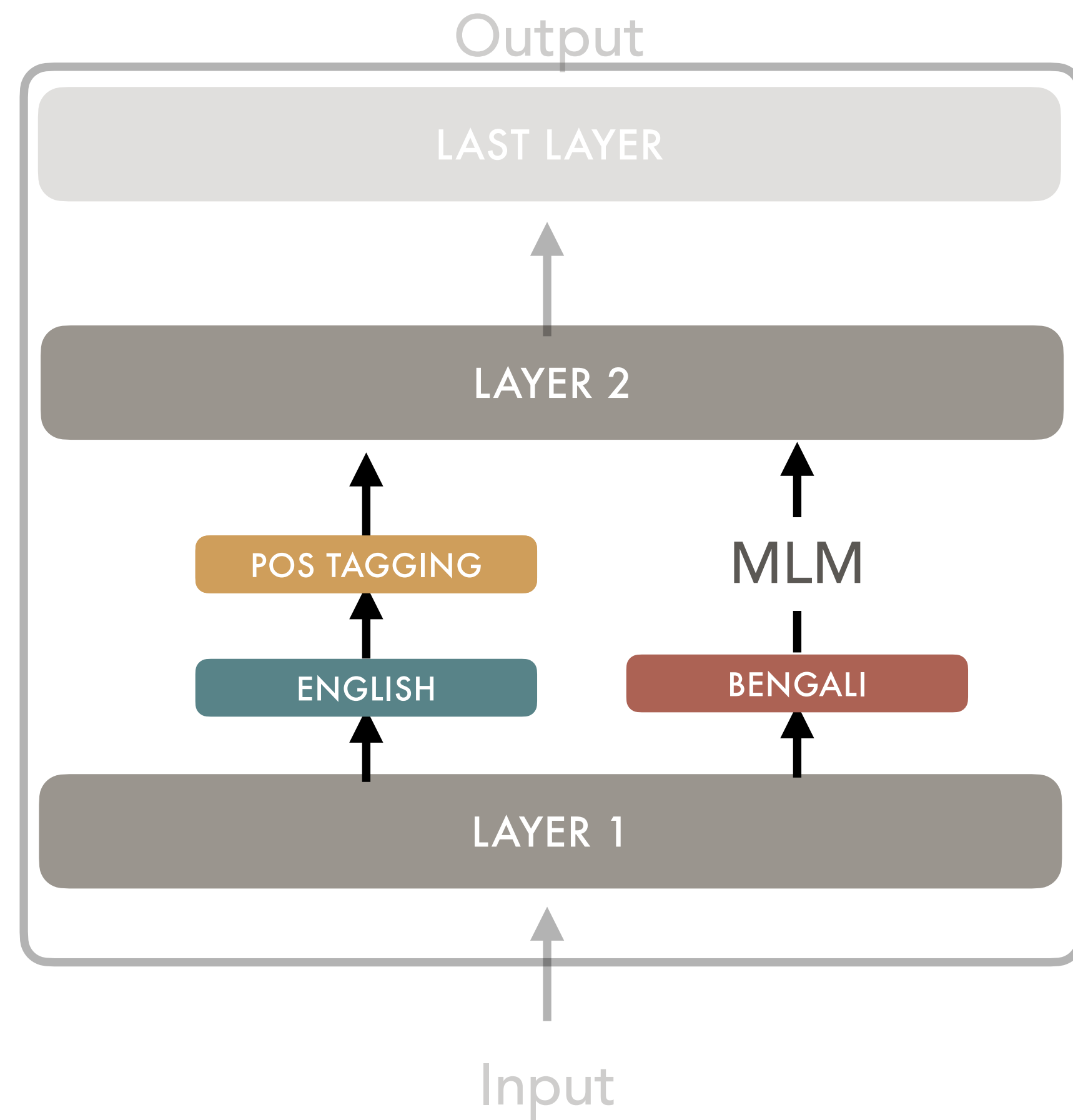
Revisiting Adapters



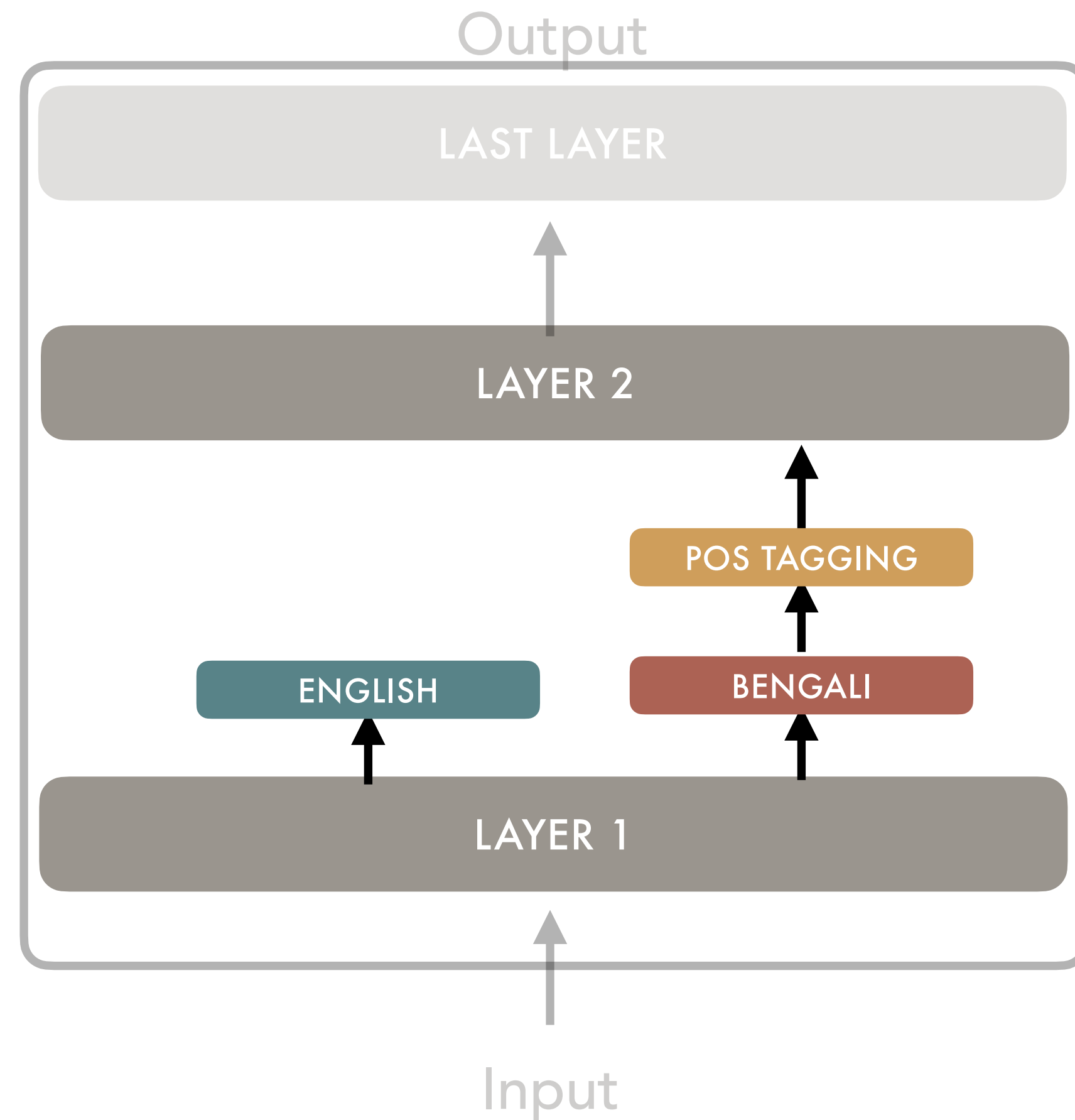
Revisiting Adapters



Revisiting Adapters



Revisiting Adapters



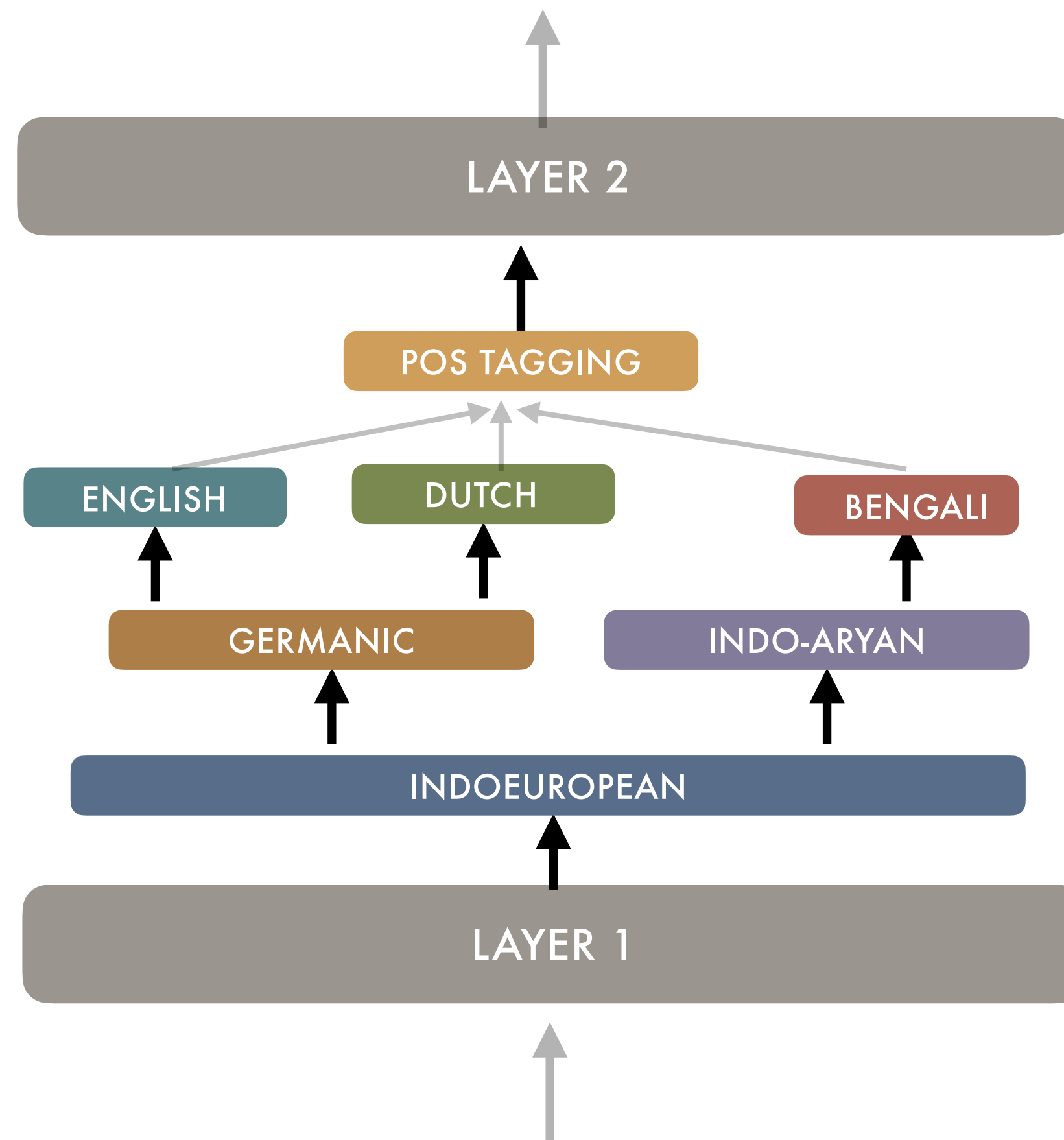
Easy zero-shot adaptation to new languages at a low cost (additional parameters)

Avoids catastrophic forgetting

Performance comparable to full-model fine-tuning

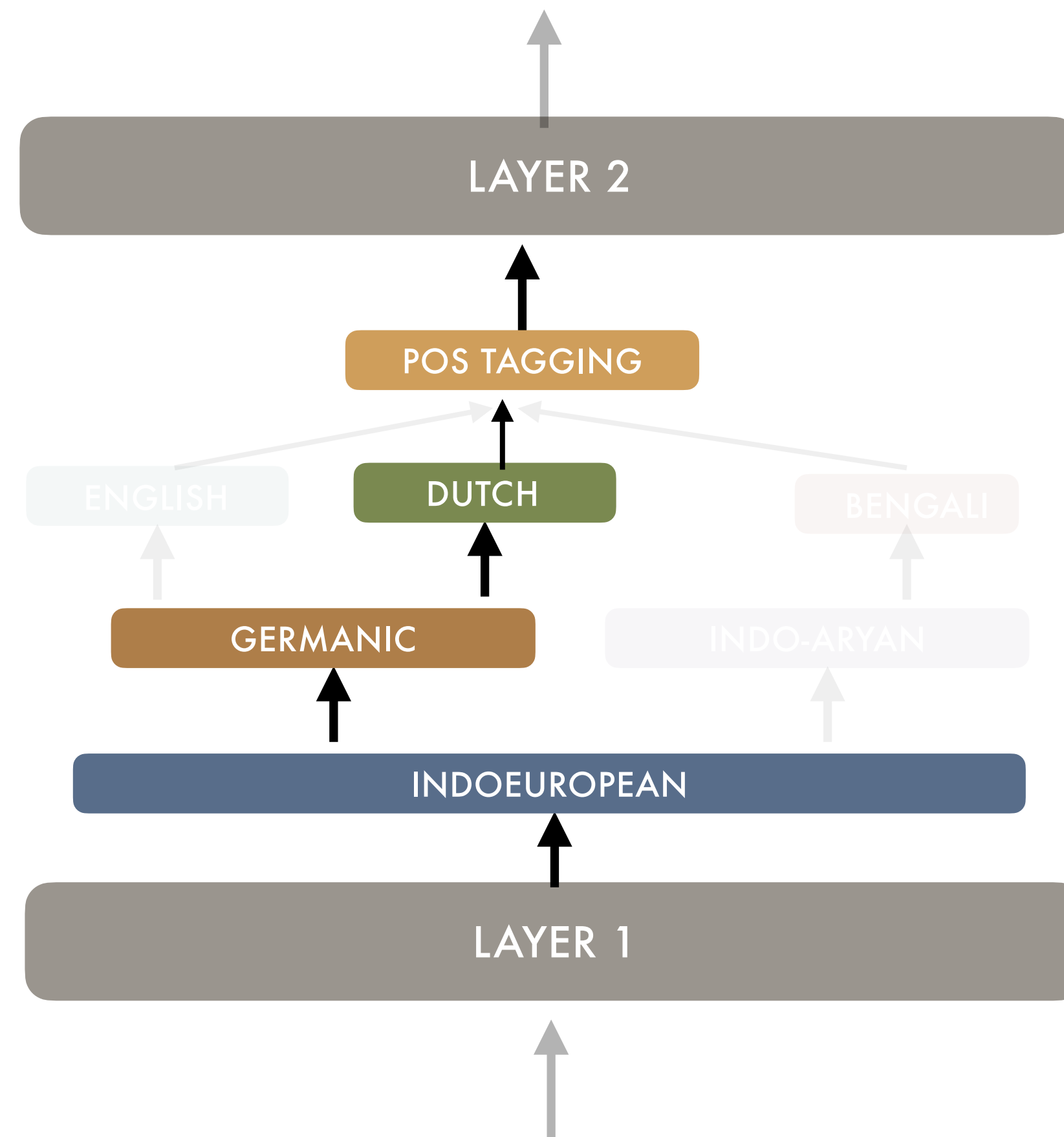
Can we do better?

Follow Phylogeny for Parameter Sharing



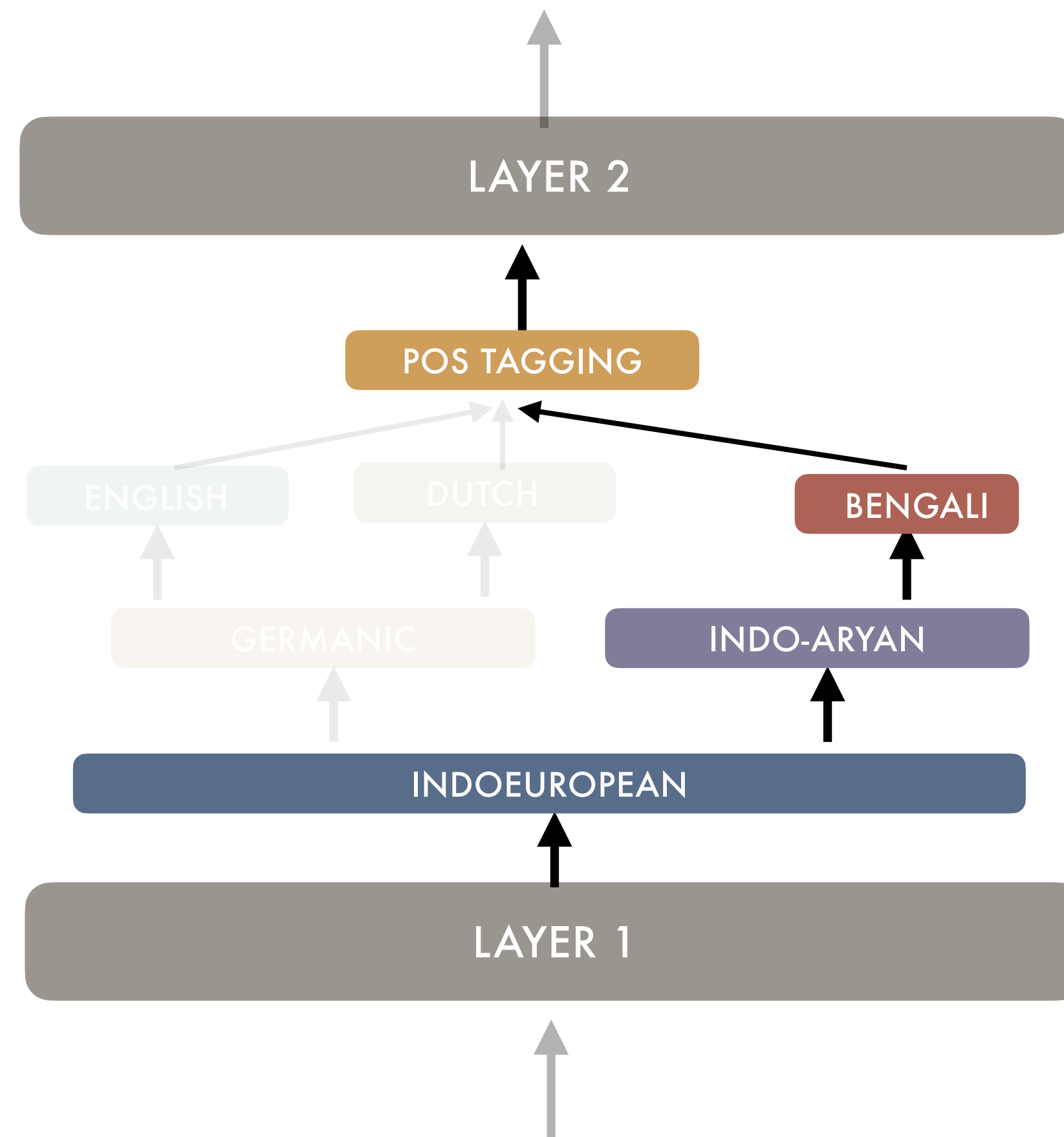
Follow Phylogeny for Parameter Sharing

For Dutch input

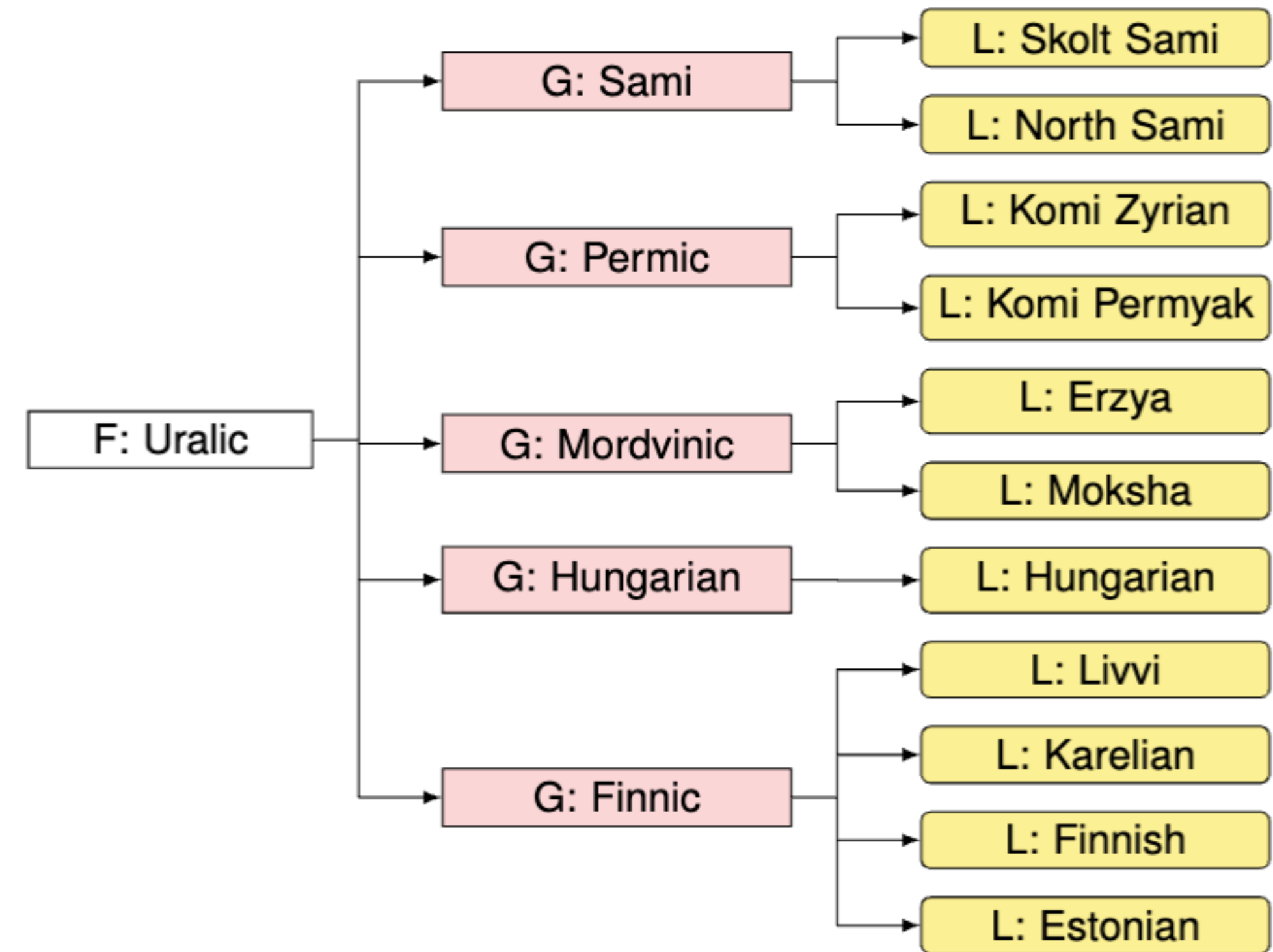
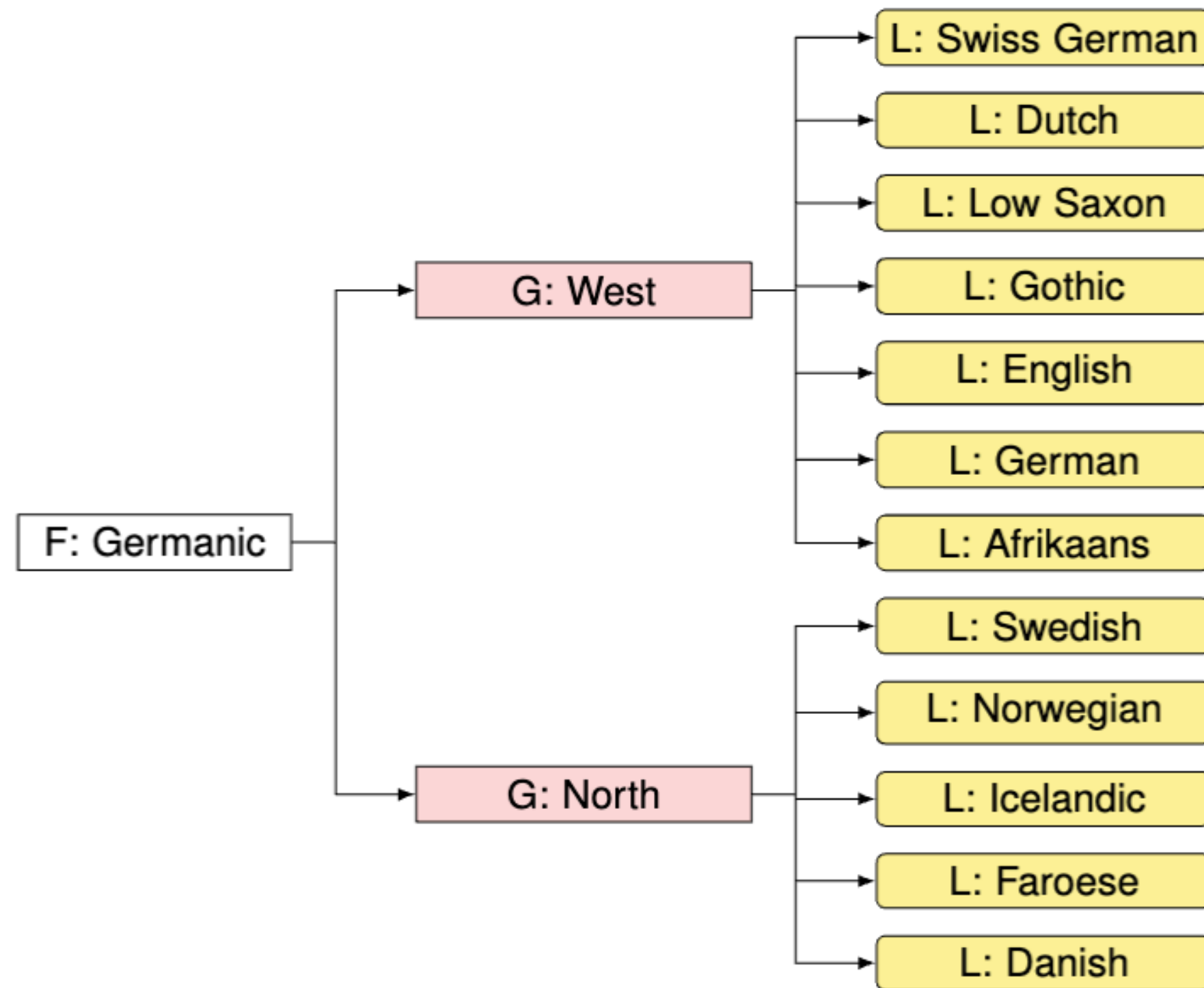


Follow Phylogeny for Parameter Sharing

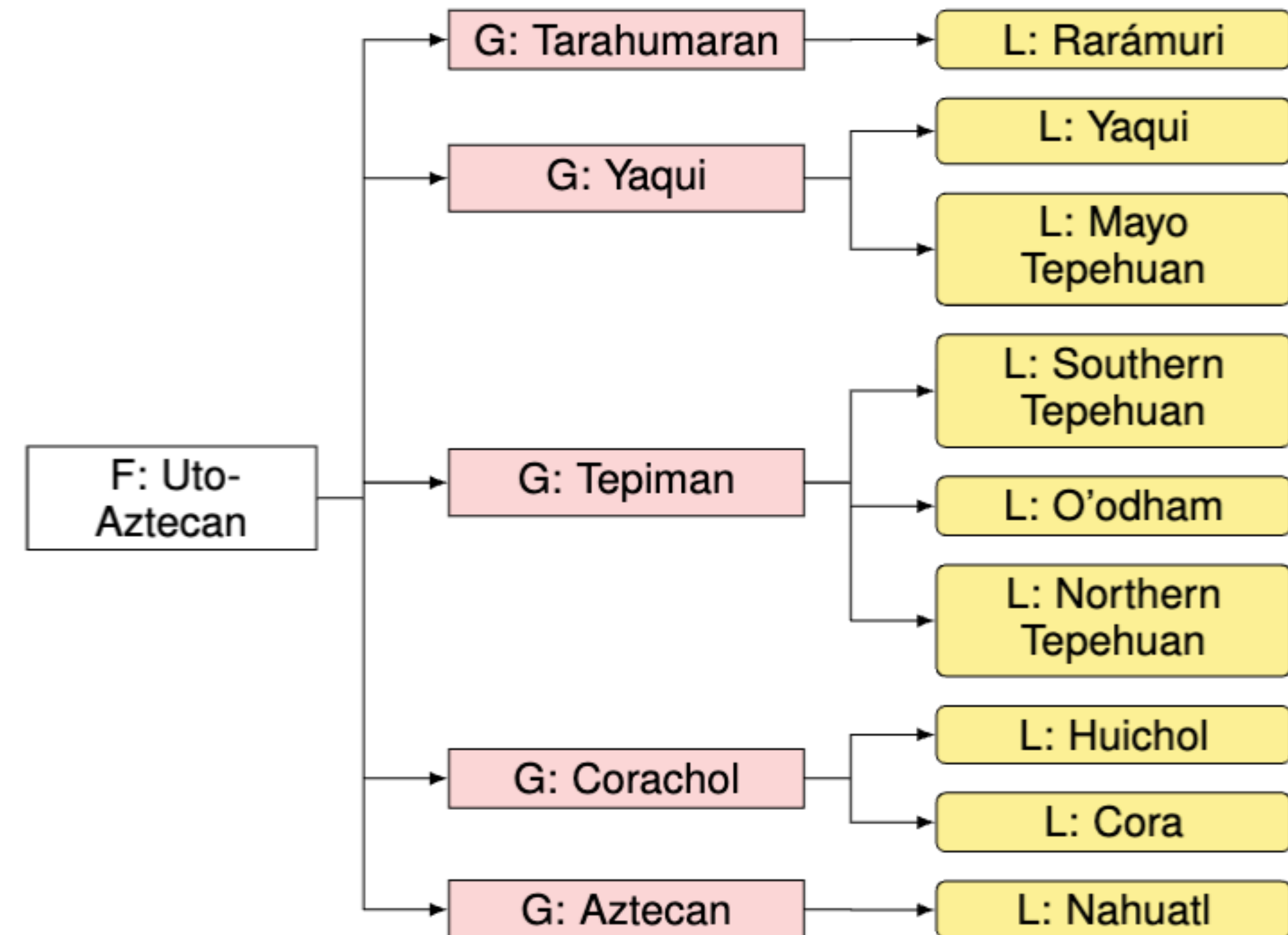
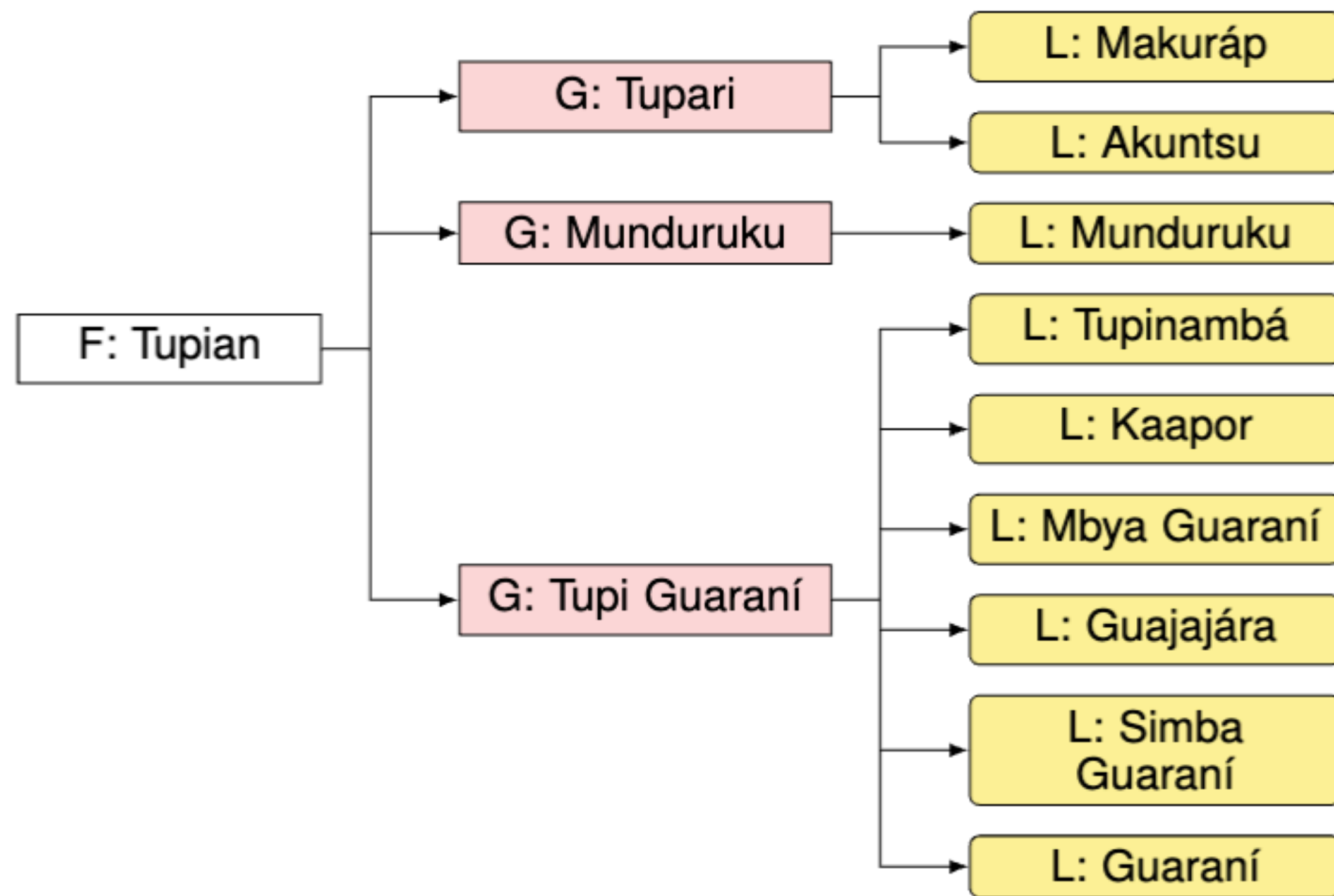
For Bengali input



Experimental setup

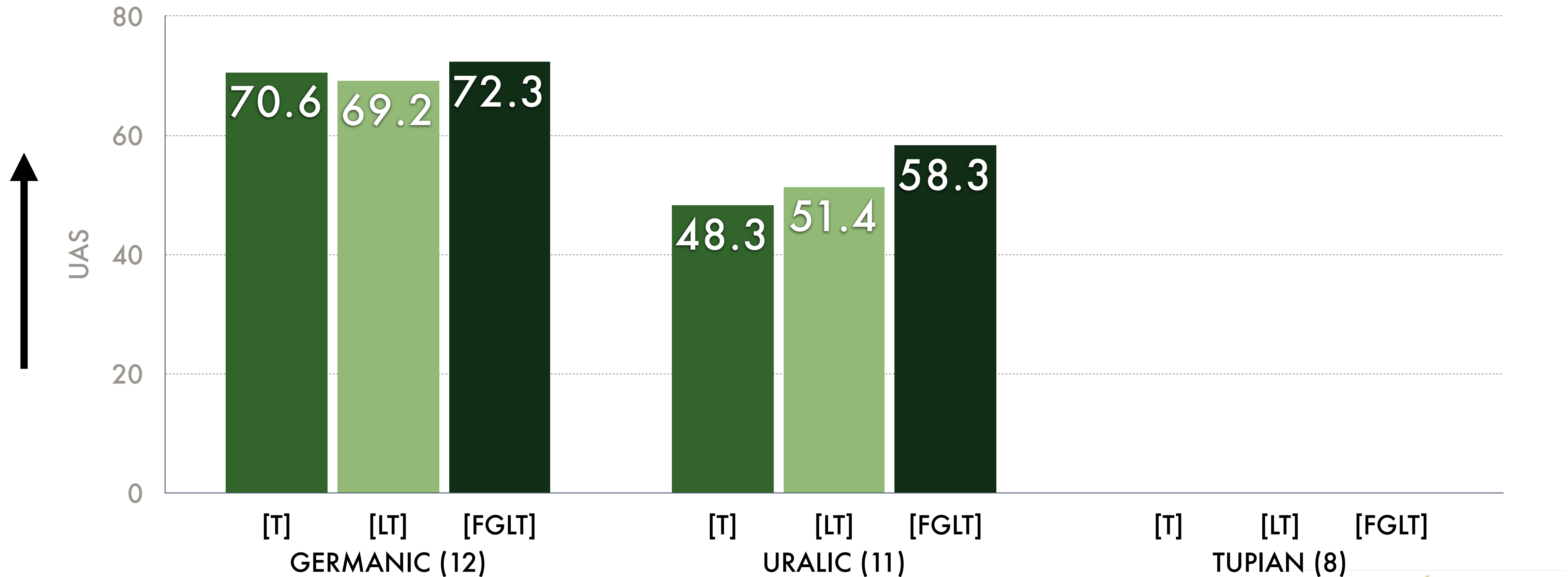


Experimental setup



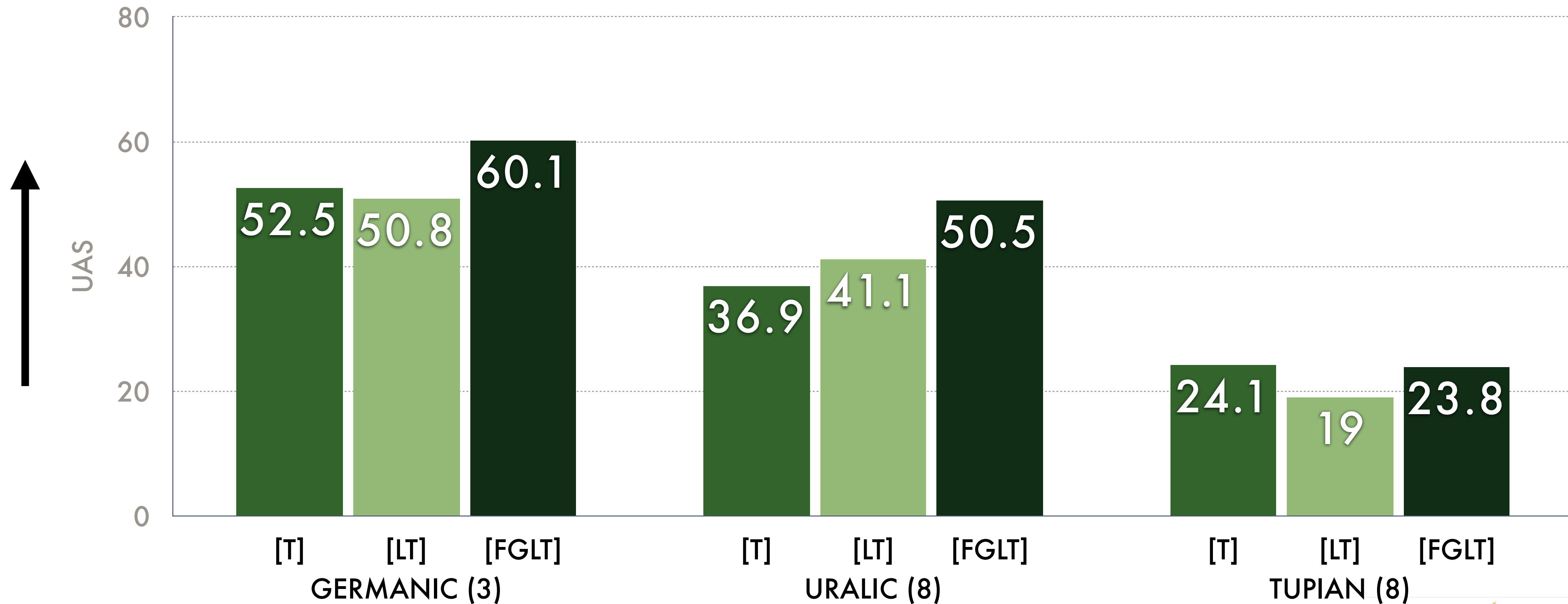
Results

DEPENDENCY PARSING



Results on unseen languages

DEPENDENCY PARSING

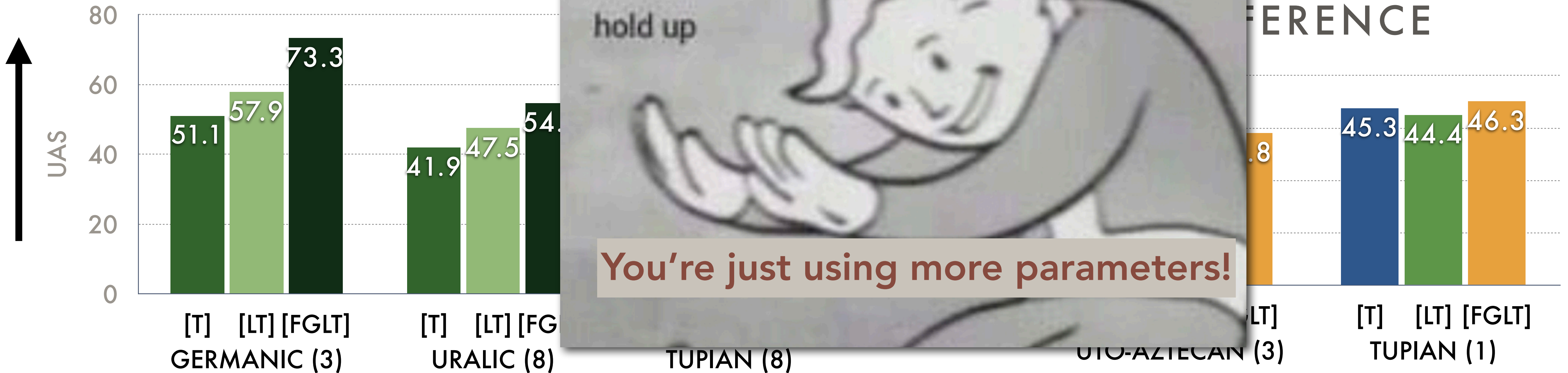


Much larger improvements for *new, unseen* languages

Results on unseen languages

PARTS-OF-SPEECH TAGGING

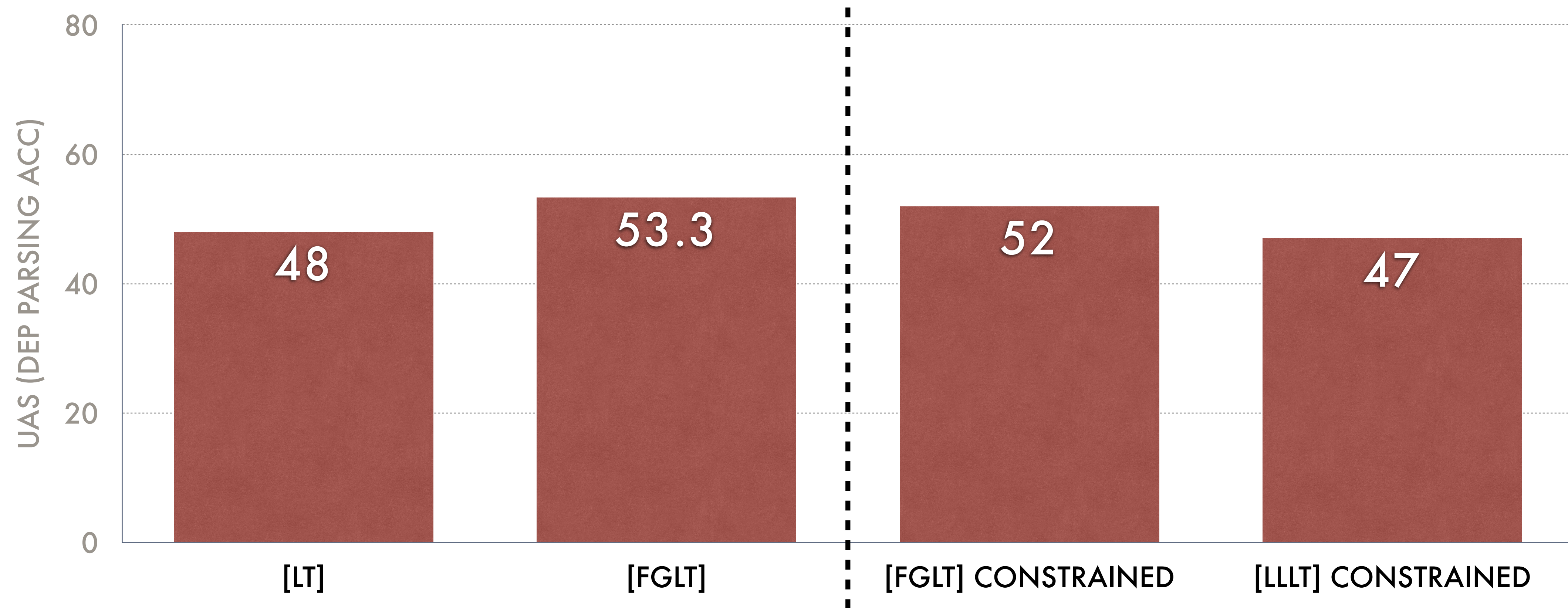
NATURAL LANGUAGE REFERENCE



Much larger improvements for *new, unseen* languages

Ablations: Parameters

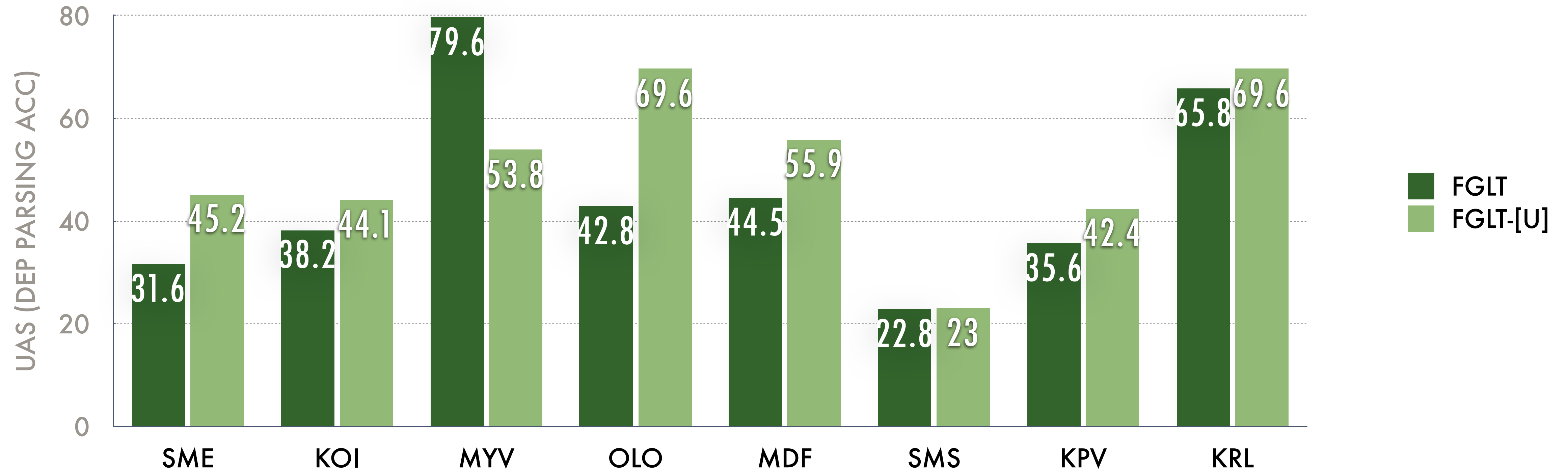
DEPENDENCY PARSING ON URALIC LANGS



Even constraining to the same number of parameters, still improvements!
Is it language sharing or network depth?

Ablations: Upsampling

DEPENDENCY PARSING ON URALIC LANGS



Upsampling by simple repeating sentences does better

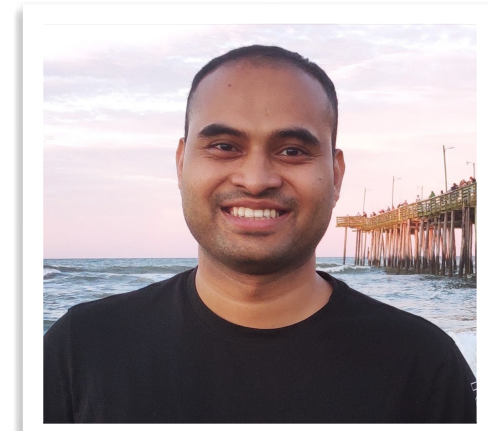
- Adapter-based approach to leverage language phylogenetic information for better cross-lingual adaptation.
- Exact same parameter count but smaller adapters with parameter sharing across related language improves performance in true-zero-resource scenarios.

Thank you!

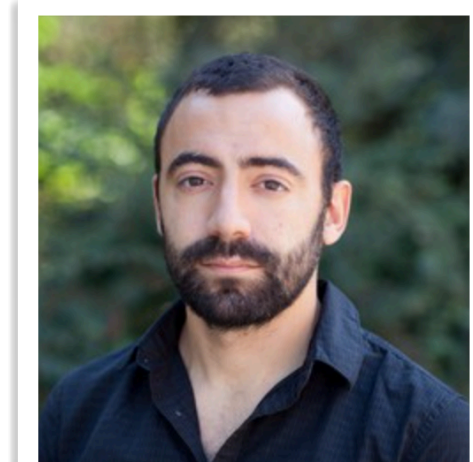
**Code
&
Dataset**



https://github.com/ffaisal93/adapt_lang_phylogeny



FAHIM FAISAL (ffaisal@gmu.edu)



Antonios Anastasopoulos (antonis@gmu.edu)

