

Heuristic based approach





**Comparison baseline: Pre-training with equivalent
amount of English-Mandarin monolingual data**

Why?

Loss of structure



“I love playing games”

“Mein mohabbat playing games”

Perplexity	Data	RNNLM	DLM
Heuristic 1	Dev	76.63	78.54
	Test	63.54	65.73
Heuristic 2	Dev	76.13	78.27
	Test	63.3	65.4
Fair Baseline	Dev	75.11	74.5
	Test	62.53	61.95

Heuristic based approach

Perplexity	Data	RNNLM	DLM
Heuristic 1	Dev	76.63	78.54
	Test	63.54	65.73
Heuristic 2	Dev	76.13	78.27
	Test	63.3	65.4
Fair Baseline	Dev	75.11	74.5
	Test	62.53	61.95

Comparison baseline: Pre-training with equivalent amount of English-Mandarin monolingual data

“I love playing games”

“*Mein mohabbat* playing games”

Why?
Loss of structure

Why not directly use RNNLM?
