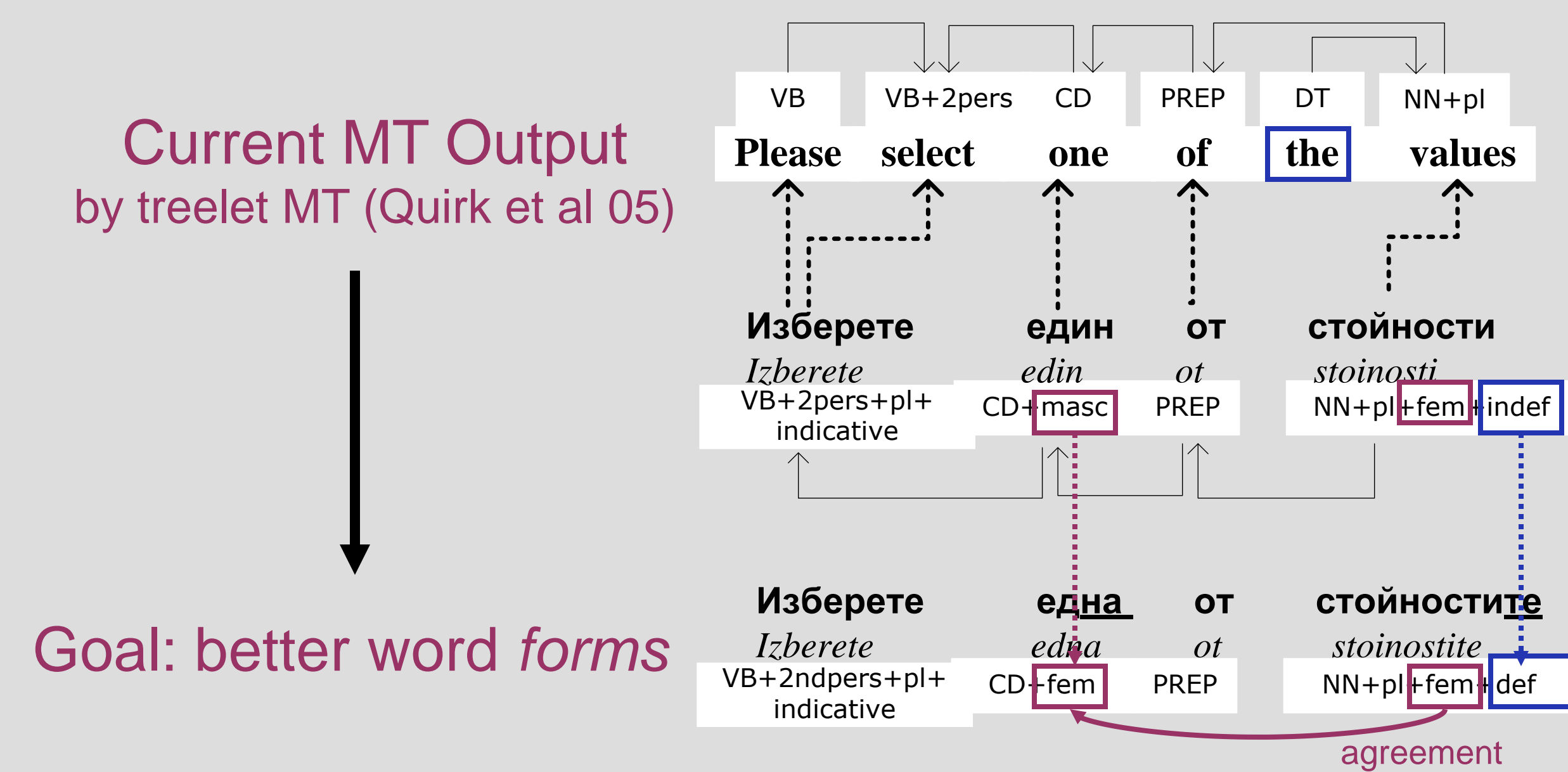


Applying Morphological Generation Models to Machine Translation

Kristina Toutanova, Hisami Suzuki, Achim Ruopp (Microsoft Research)

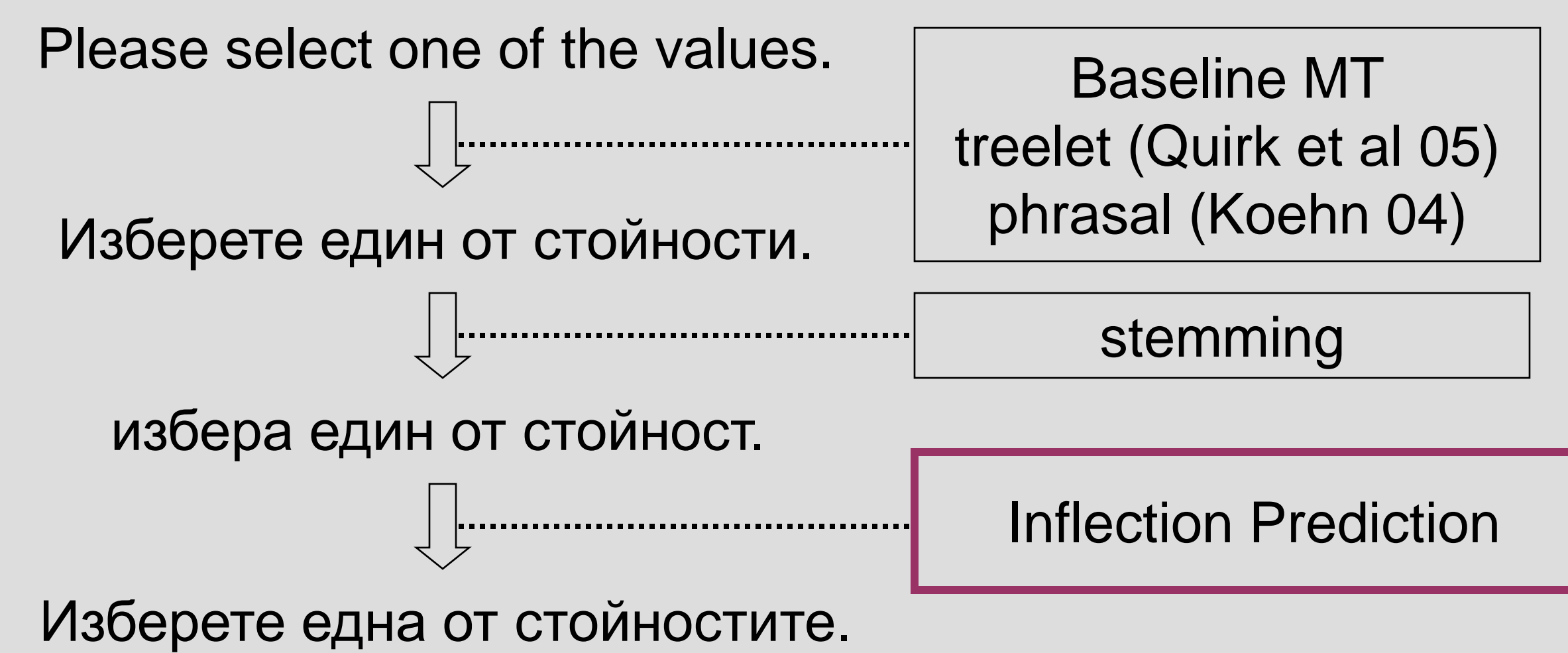
Motivation

Better MT output for morphologically rich languages

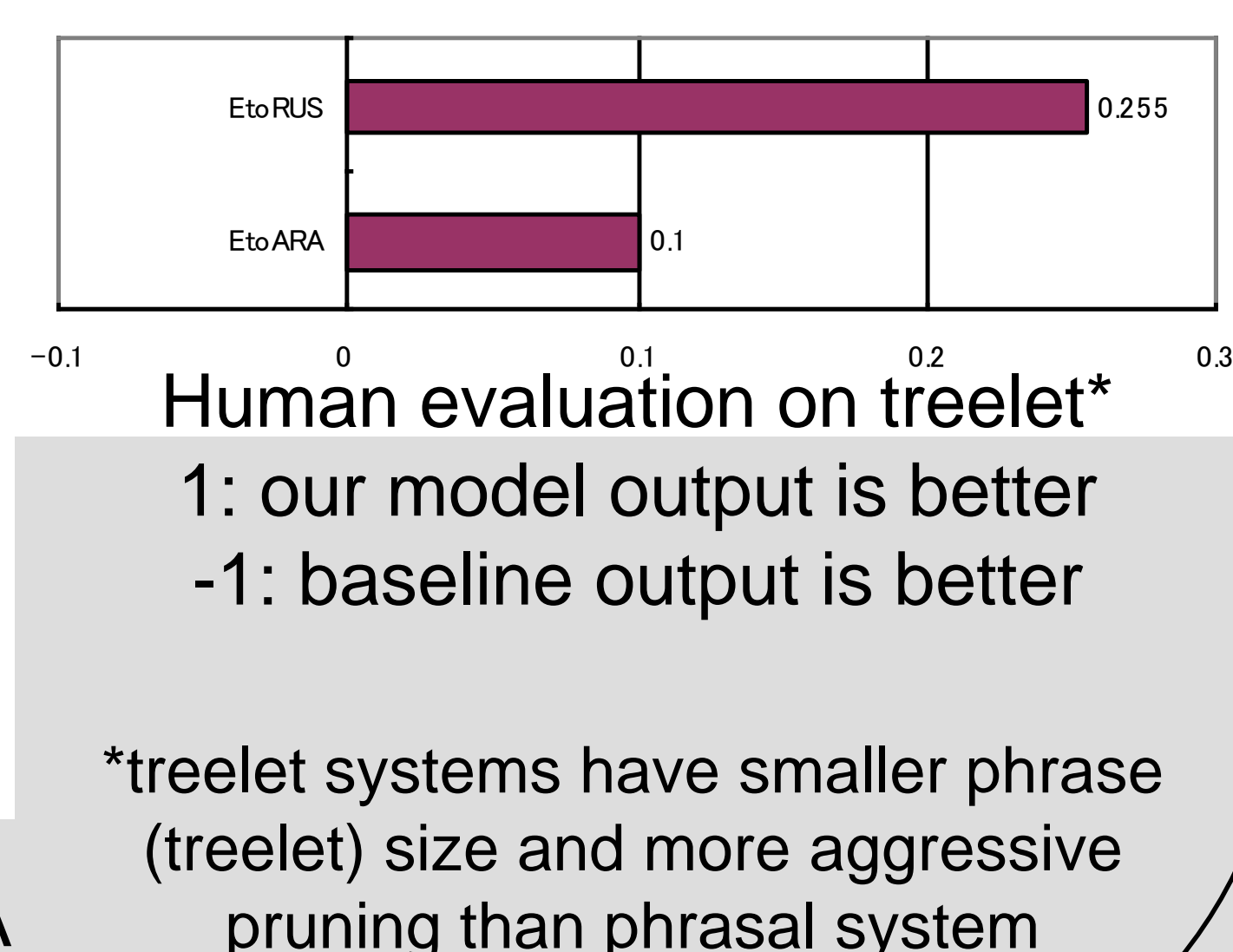
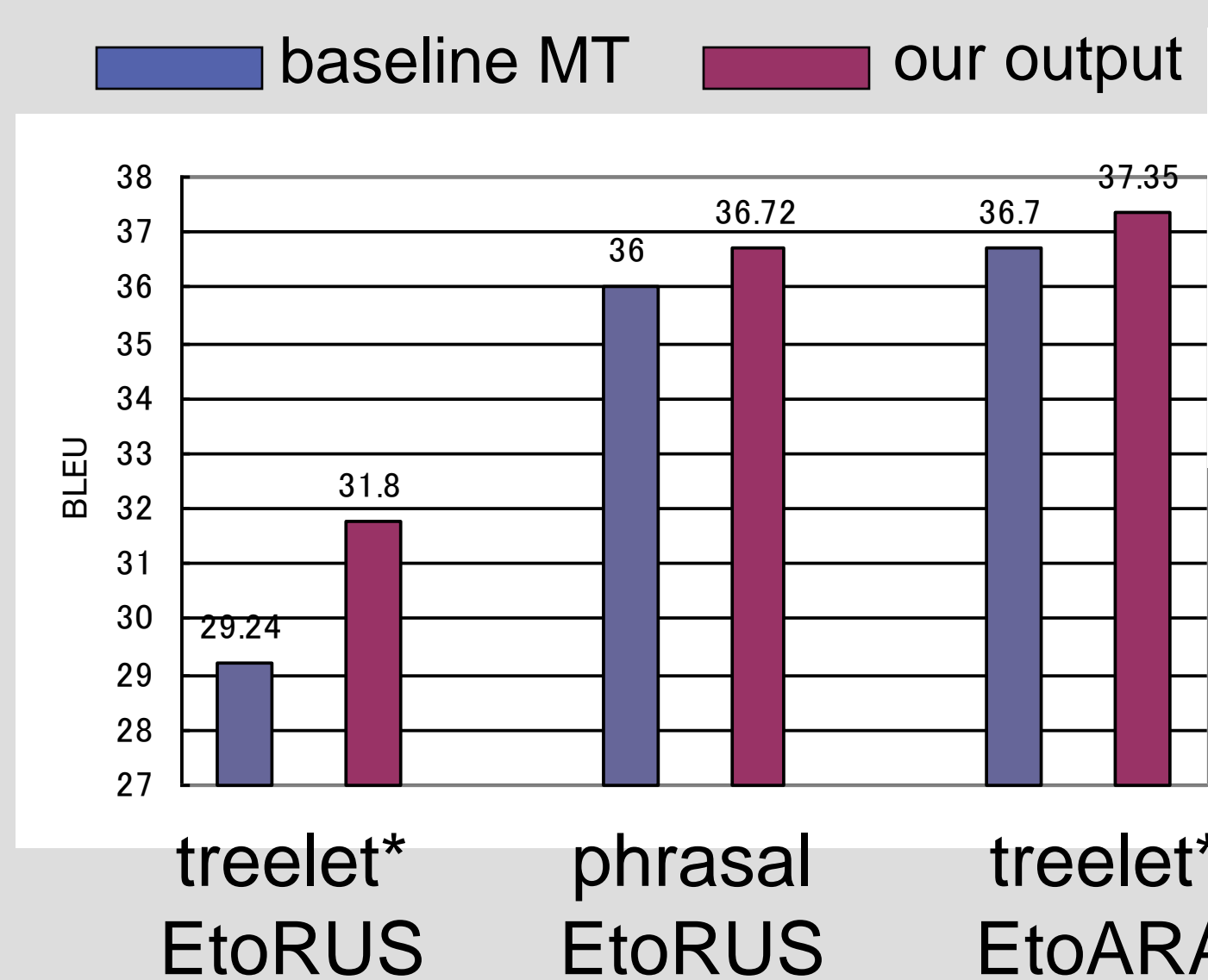


Our Approach

- Build an independent model to inflect words
- Inflect the n-best output of the baseline system
- Works both with treelet and phrasal MT systems

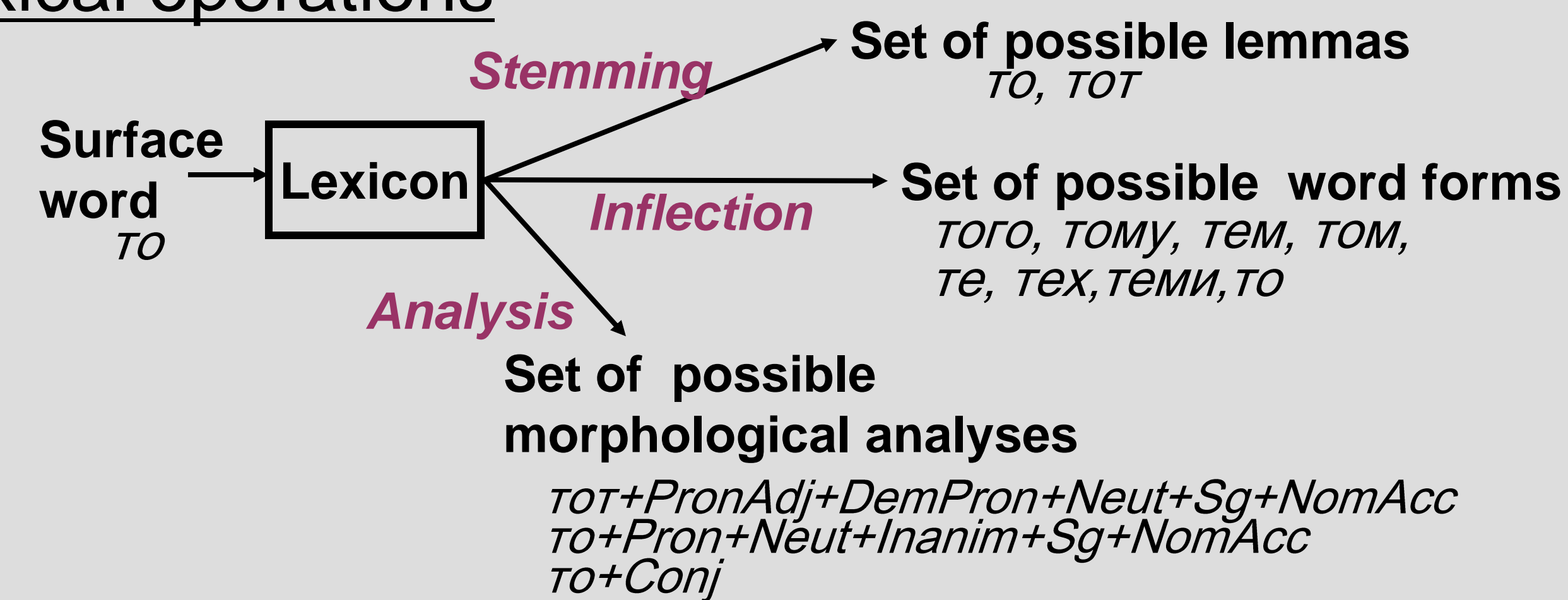


- Result: better in both BLEU and human evaluation



Inflection Prediction Model

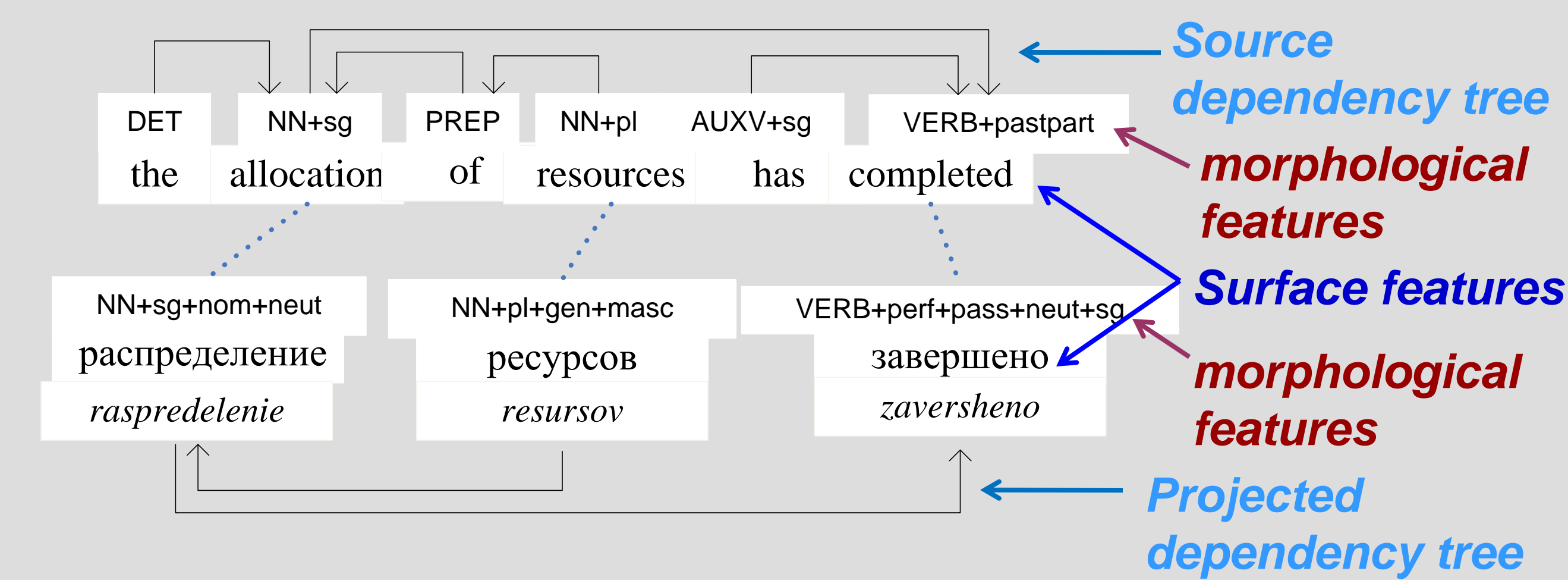
Lexical operations



Task: Given a sequence of stems, predict word inflection
 Model: Maximum Entropy Markov Model

$$p(\bar{y} | \bar{x}) = \prod_{t=1}^n p(y_t | y_{t-1}, y_{t-2}, x_t), y_t \in \text{Inflection}(t)$$

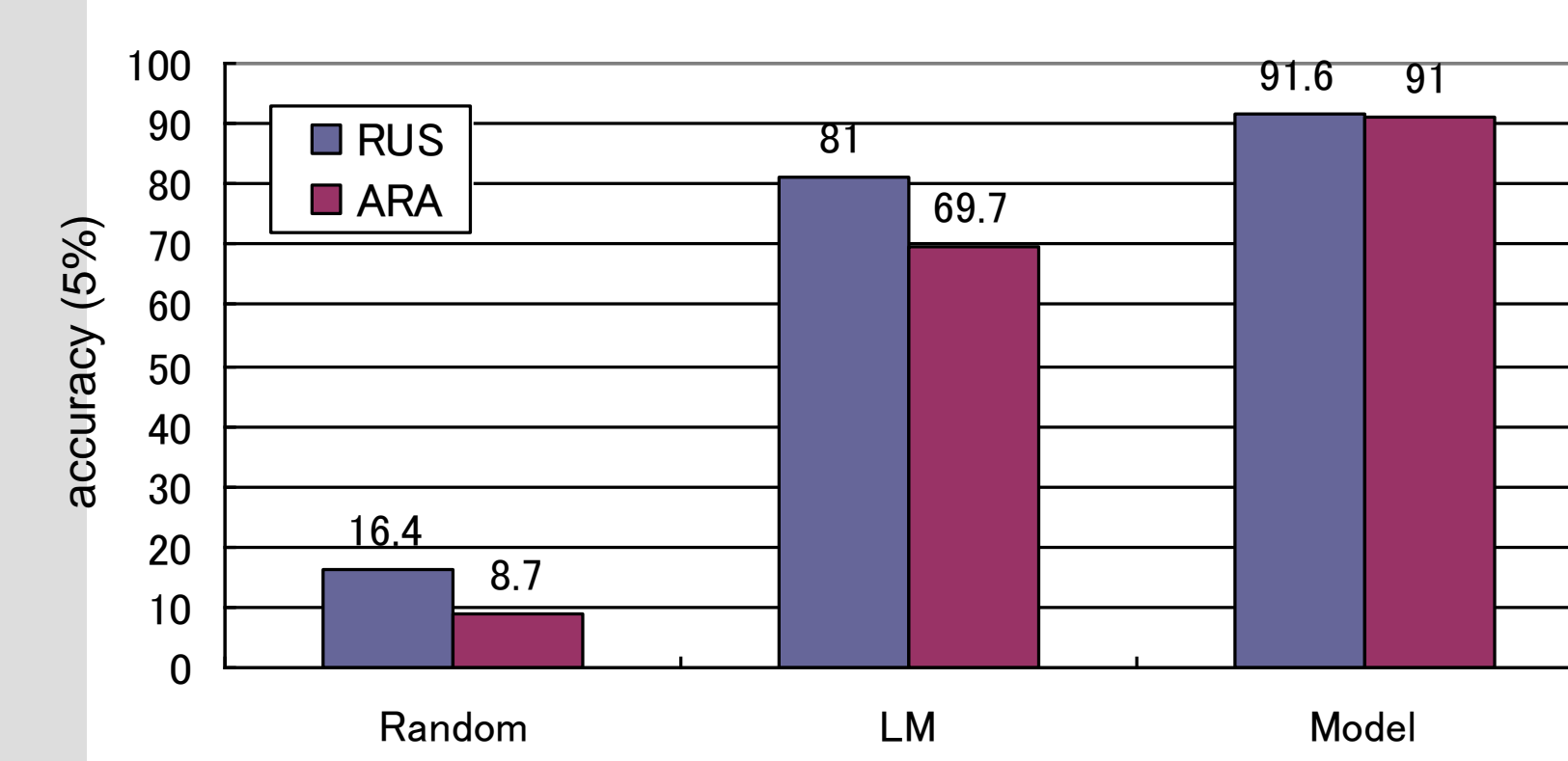
x_t = context at position t , obtained from the linguistic annotation (provided by the treelet MT system)



Sample features

RUS: [PrevStem=x, Case_Inflection=y]
 [AlignedWords="will", Tense_Inflection=future]
 ARA: [Conj(AlignedWords)=true, Conj_inflection="و"]
 [POS=NOUN & POS(NextStem)=NOUN, Det_inflection=true]

Accuracy on Reference translation



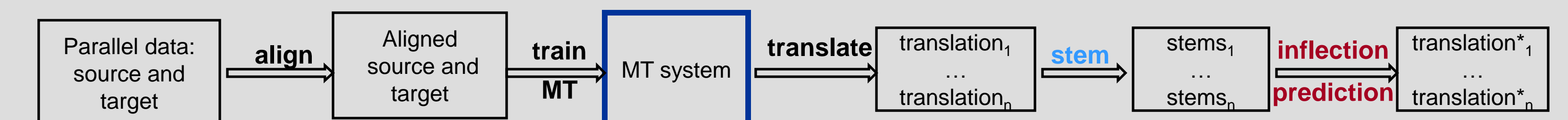
Given perfect word choice and word order, how accurately can we predict the inflected form?

Integrating Inflection Prediction models with MT systems

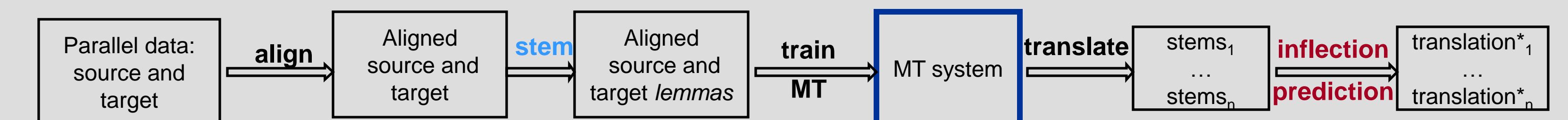
Studied 3 methods of integration

To what extent should the translation problem be factored into the sub-problems of predicting stems and predicting inflections?

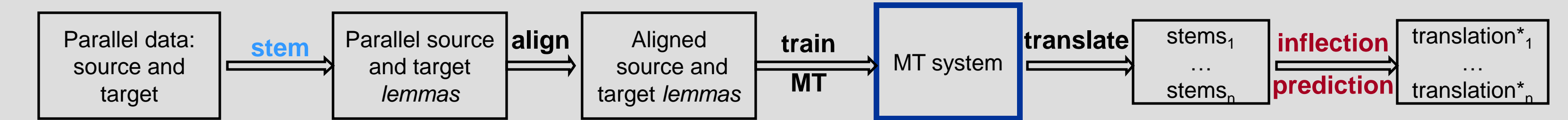
(1) No factoring: the base MT system predicts fully inflected forms



(2) Factor in translation but not in alignment: align using inflected target, then stem and train the base MT system to predict stems



(3) Factor in translation and alignment: align and train base MT system using stemmed target sentences

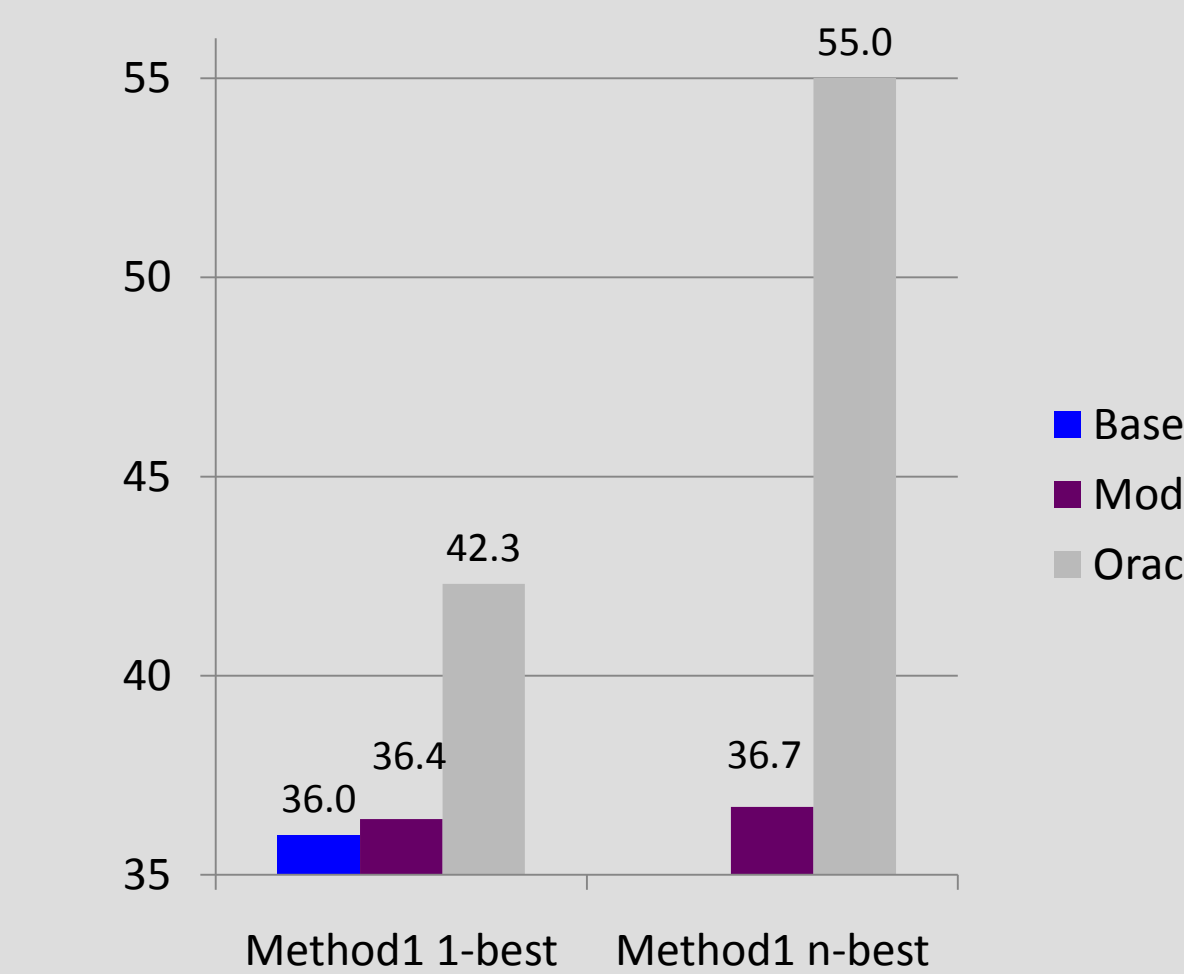


$$trans_i = \operatorname{argmax}_{trans_i \in \text{Infl}(stems_i)} \lambda_{IM}(trans_i | stems_i) + \lambda_{LM} \log pr_{LM}(trans_i)$$

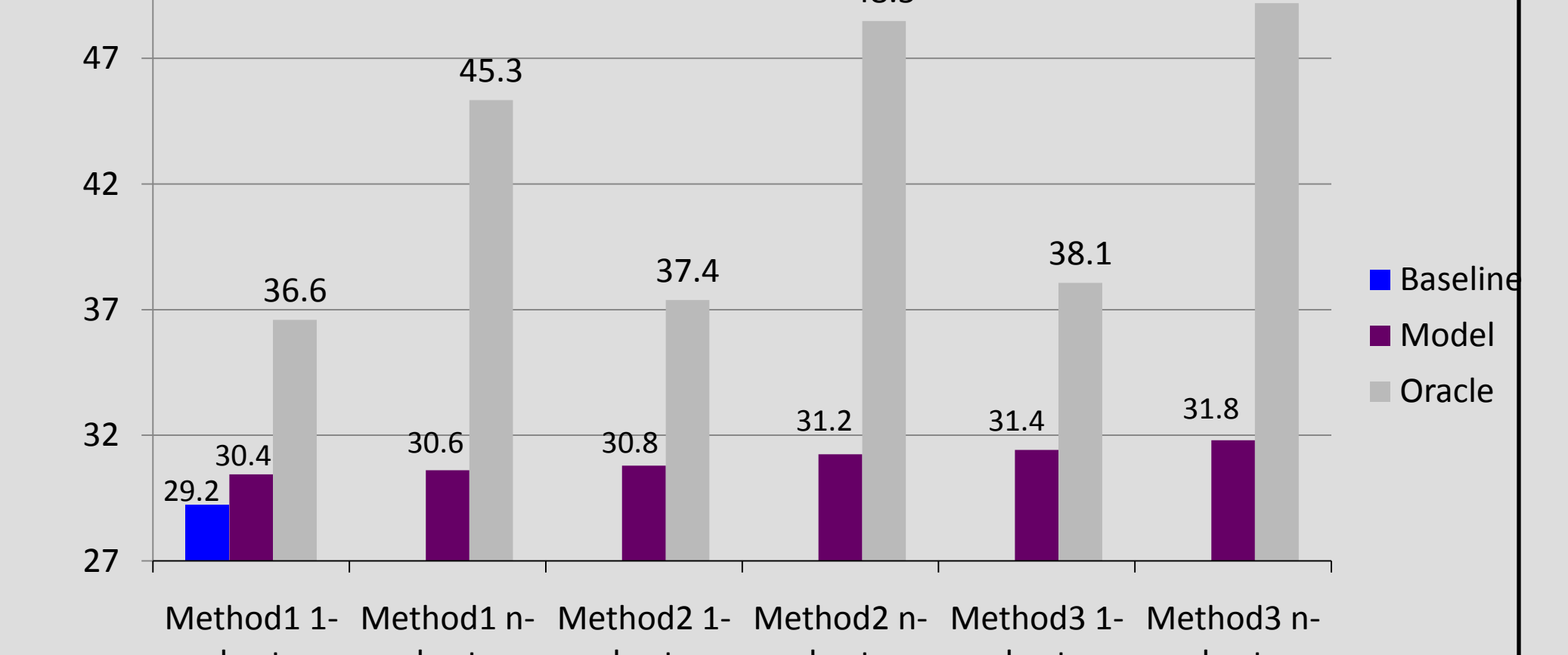
$$trans^* = \operatorname{argmax}_{i=1..n} \lambda_{MT} score(stems_i) + \lambda_{IM} \log pr_{IM}(trans_i | stems_i) + \lambda_{LM} \log pr_{LM}(trans_i)$$

Data: computer manual domain (# sentence pairs)

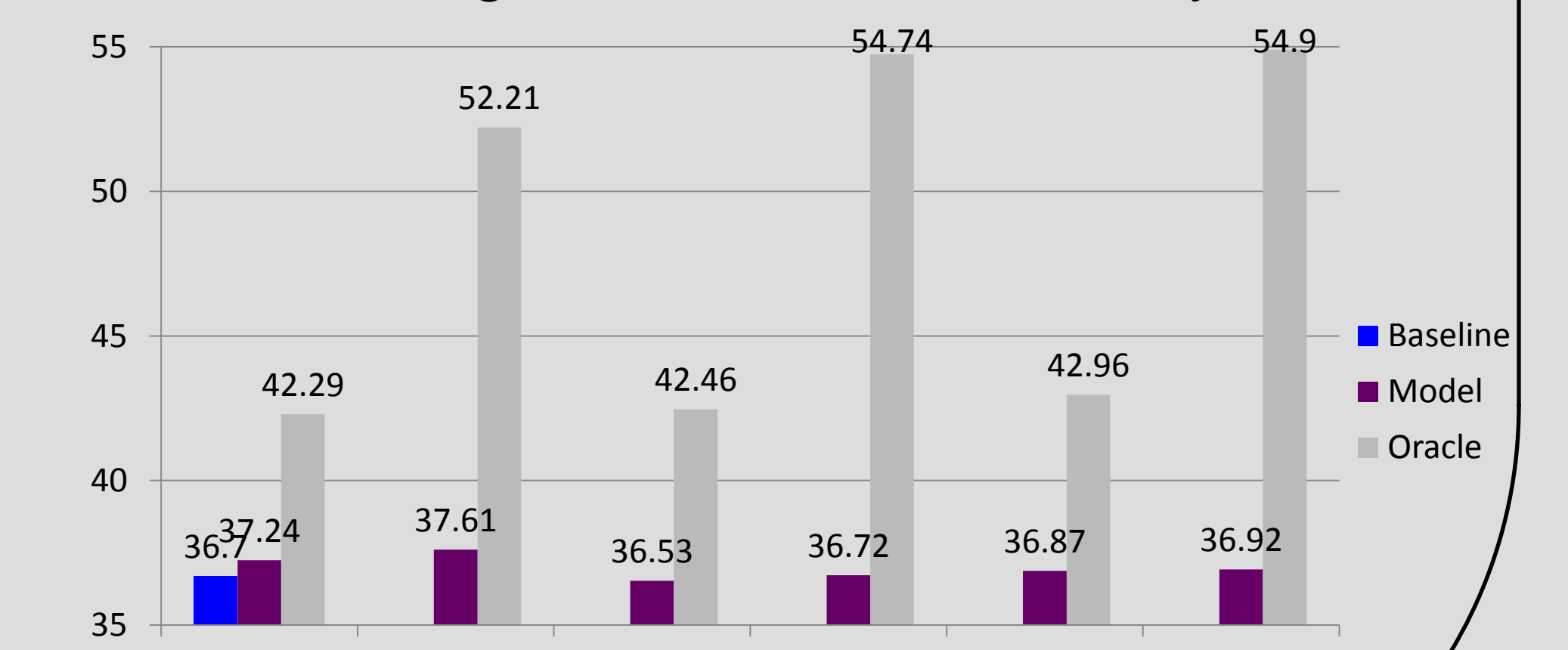
	Eng-Rus	Eng-Ara
train	1.6M	463K
lambda	2K	2K
dev	1K	1K
test	4K	4K



Results: English-Russian w/ phrasal system



Results: English-Russian w/ treelet system



Results: English-Arabic w/ treelet system