

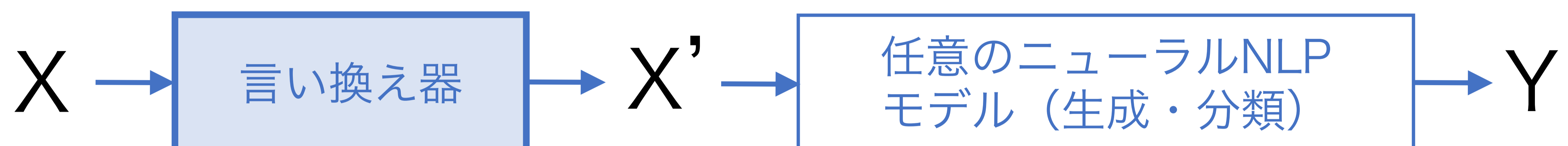
背景 : Input Sensitivity

ニューラルNLPモデルは入力表現に敏感 (得意な表現と不得意な表現がある)

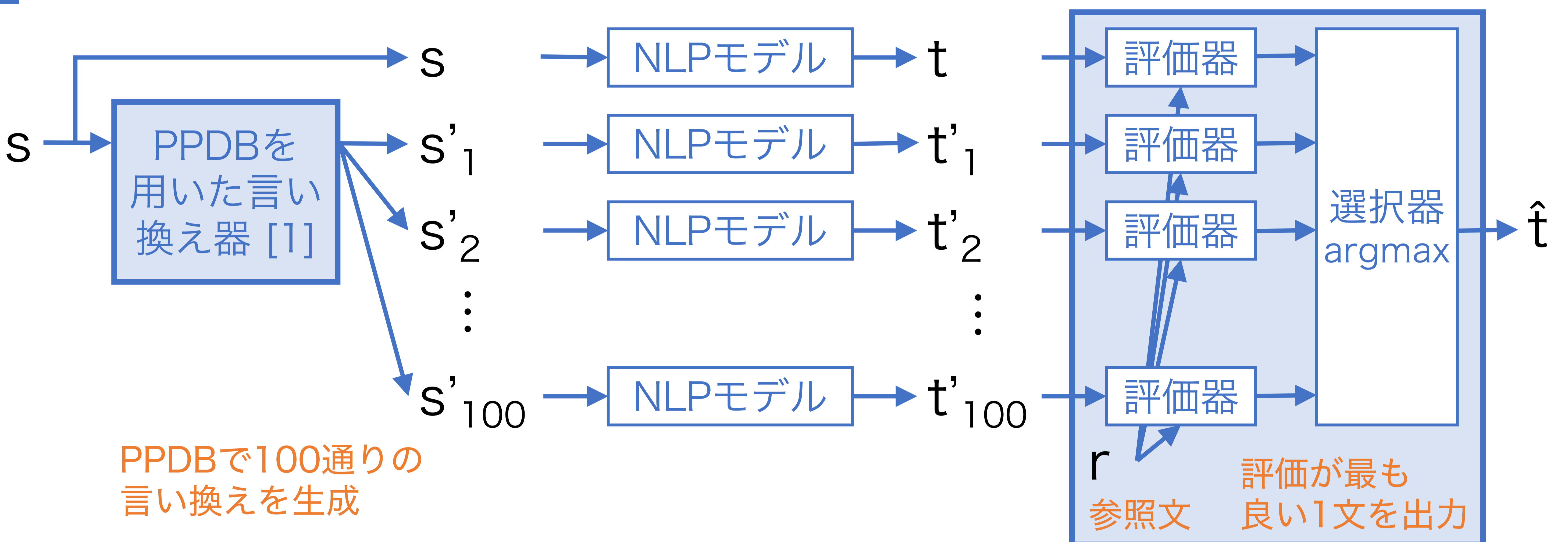
入力文	Google 翻訳
私は阪大を受けます。	I will receive Osaka University.
私は阪大を受験します。	I take Osaka University.
私は, 阪大を受験します。	I will take the Osaka University examination.

提案 : 言い換えによる前編集

- ◆ 入力文の意味を保持しつつ得意な表現へ前編集できれば、より良い出力を得られる
- ◆ 自然言語を入力とするタスク全般に適用可能な方法



実験 : PPDBで100通りの言い換えを生成 (オラクル評価)



翻訳実験(1) : WMT2017 優勝モデル [2] (対訳 30~50M)

BLEU	英→中	英→露	英→独
WMT2017 優勝モデル	32.48	30.16	27.05
+前編集	40.12	38.75	35.08

最大8ポイント改善!

翻訳実験(2) : Marian デフォルト設定 (対訳 3 M)

BLEU	英→捷	英→露	英→独
Marian デフォルト設定	8.76	13.59	18.60
+前編集	16.00	22.82	28.74

WMT2017 優勝モデルより強い

分類実験 : BERT による感情分析 (SST-5)

	Acc.	入力例文	Pred.	Gold
元の入力文	54.8	It's a movie forged in the fires of chick flick hell.	pos.	neg.
+前編集	84.9	It's a movie forged in the firing of chick flick hell.	neg.	neg.

今後 : 言い換えして1回翻訳 or 複数の言い換え+翻訳からQEで選択