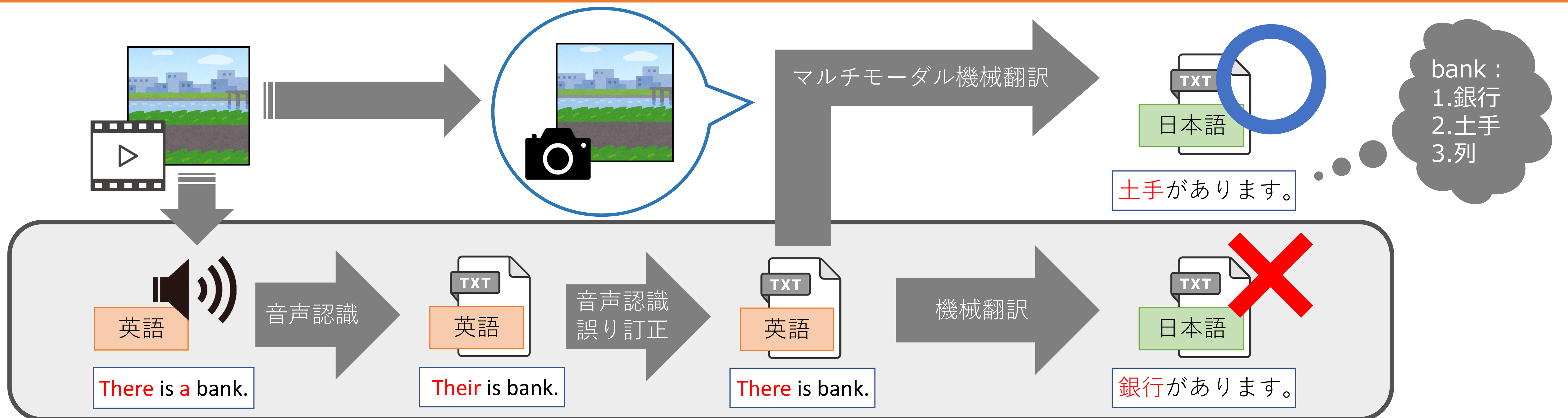


講演動画における英日字幕翻訳のための [S5-P07] マルチモーダル対訳コーパスの試作

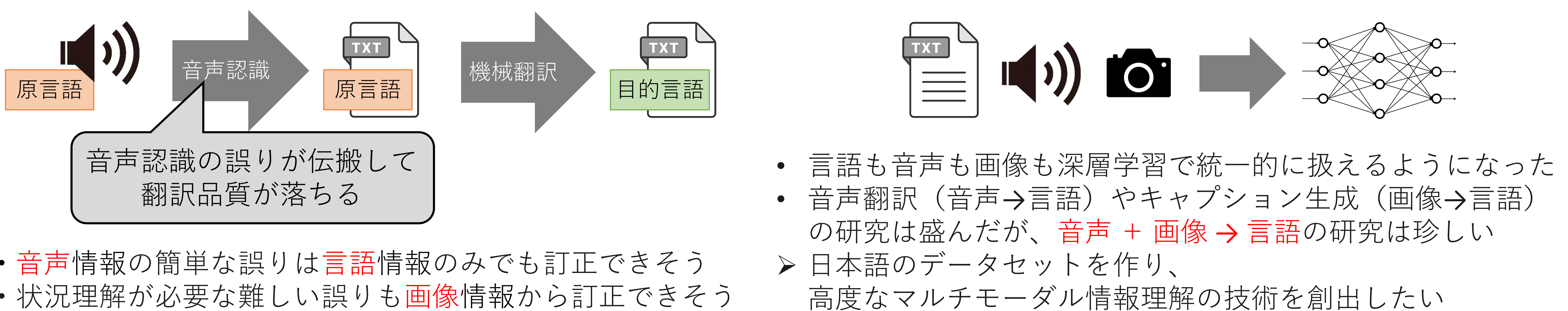
寺面杏優, 梶原智之, 二宮崇 (愛媛大学)

概要：8万件の音声&画像つき英日対訳コーパスを作成 → 音声認識の誤り訂正で翻訳品質を改善

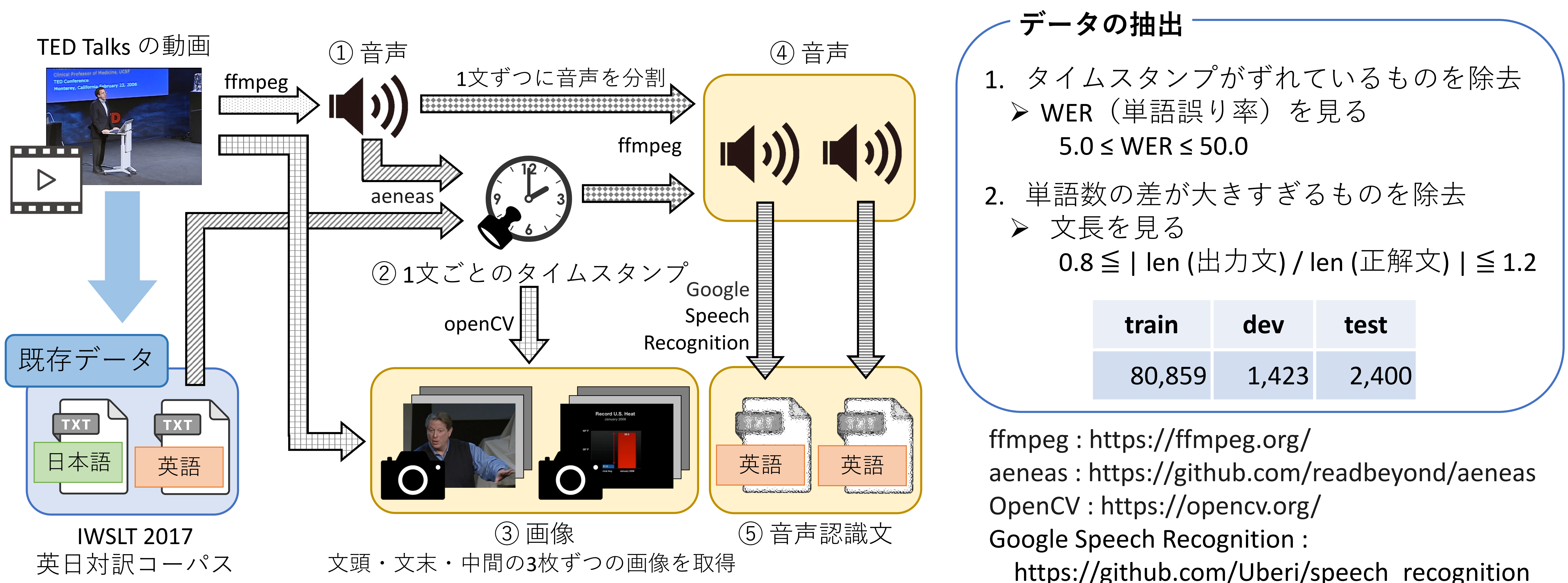
本研究の全体像：音声認識の誤り訂正と画像を用いて音声翻訳の性能を上げる



背景：複数のモダリティを組み合わせた機械翻訳の高度化に期待が高まっている

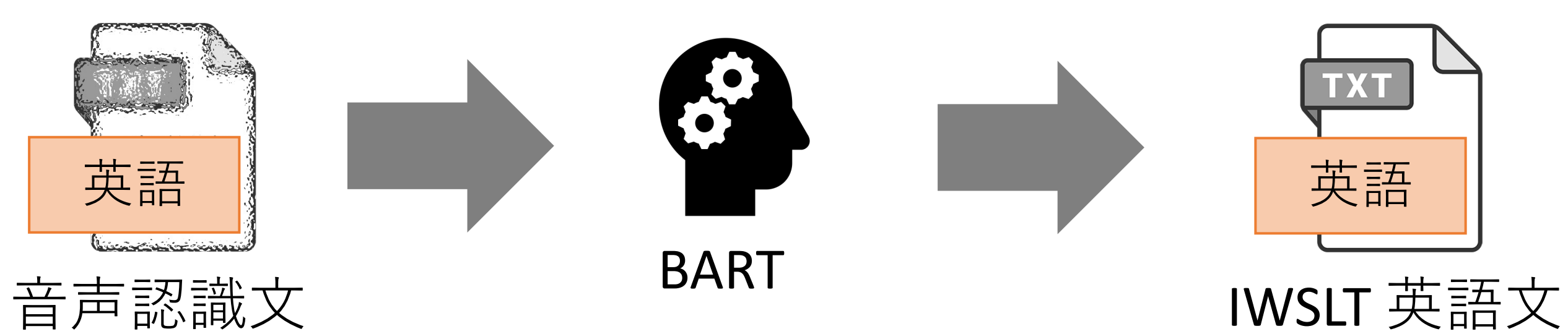


5つ組データセットを作成 (音声・画像・音声認識文・英語文・日本語文)



実験：音声認識の誤りを訂正すると翻訳品質が向上

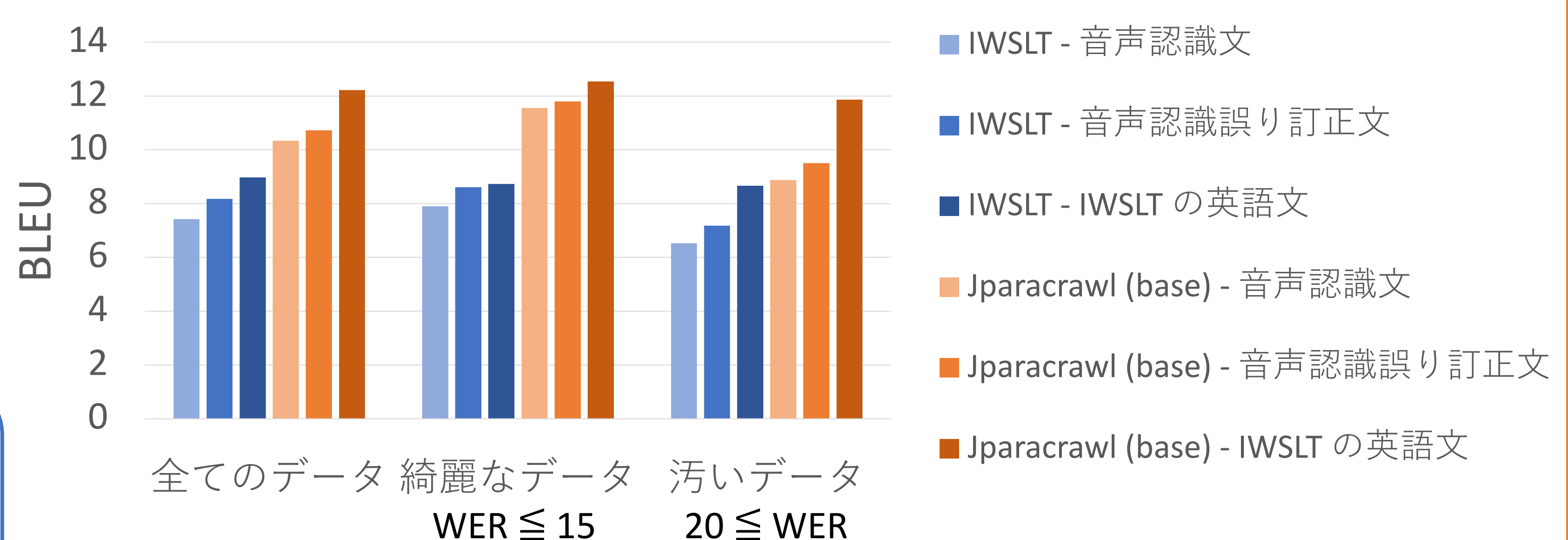
BARTを使って音声認識の誤り訂正を行う



音声認識誤り訂正の例

日本語文
音声認識文
音声認識誤り訂正文
IWSLT の英語文

それだけ聞けば私は十分でした
I was all I needed to hear
And that was all I needed to hear.
That was all I needed to hear.



～今後の展望～

画像も併用して翻訳品質のさらなる向上を目指す！！