

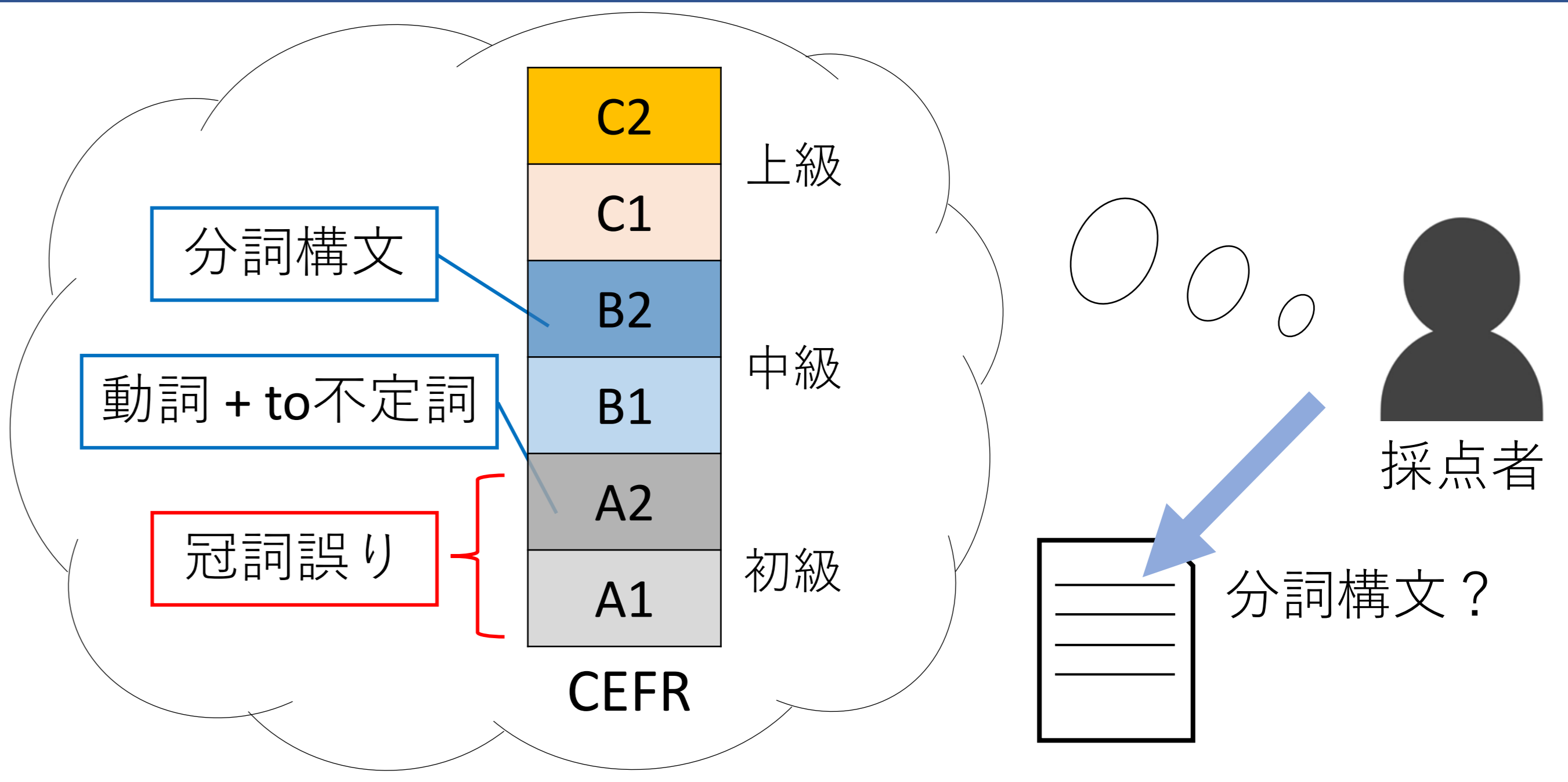
本研究の概要

エッセイの分散表現 + **文法特徴** で自動採点
 正しく使えている文法の情報 (**PF**): ×効果なし
 文法誤りの情報 (**NF**): ◎学習者のレベルが多岐
 (⇔ ×レベルが高い学習者)



なぜ文法なのか？

- 評価の際の代表的な観点のひとつ
 - CEFR × **基準特性** [Hawkins+ 2012]
 - あるレベルの学習者に特徴的な文法項目が存在
 - 人間の採点者はこれらの項目を探しながら評価している
- [Bannò+ 2022]の手法をベースに検証する



実験

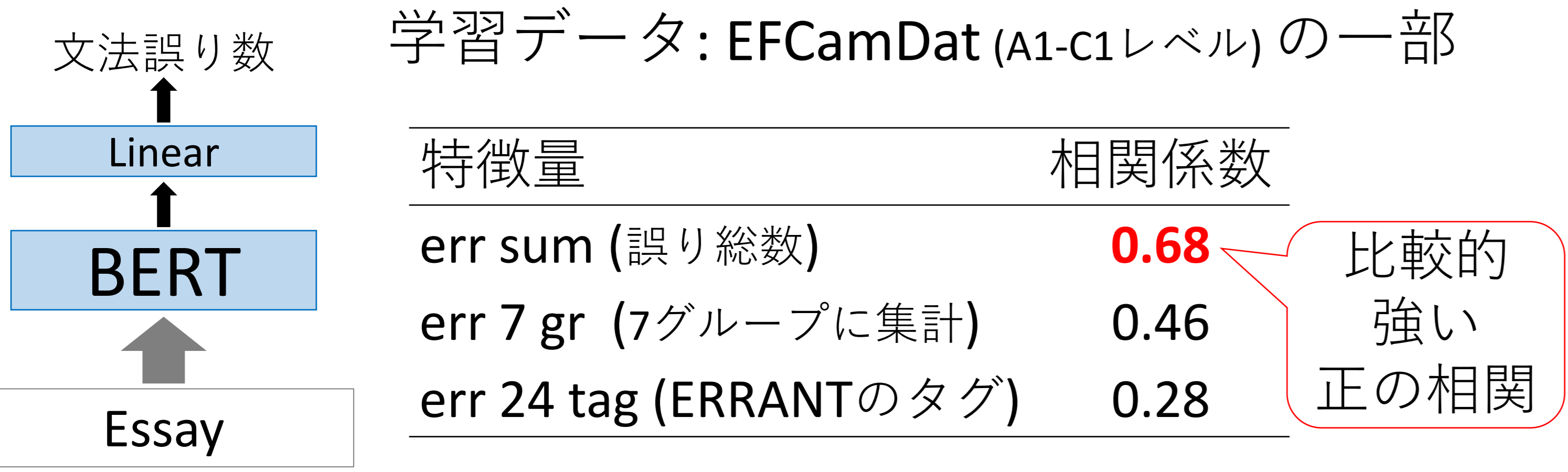
文法特徴量

PF item: 延べ数 (不定詞 = 2)
 type: 異なり数 (不定詞 = 1)
 NF ERRANT [Bryant+ 2017] の誤りタグ

... to play ...
 ... to study ...

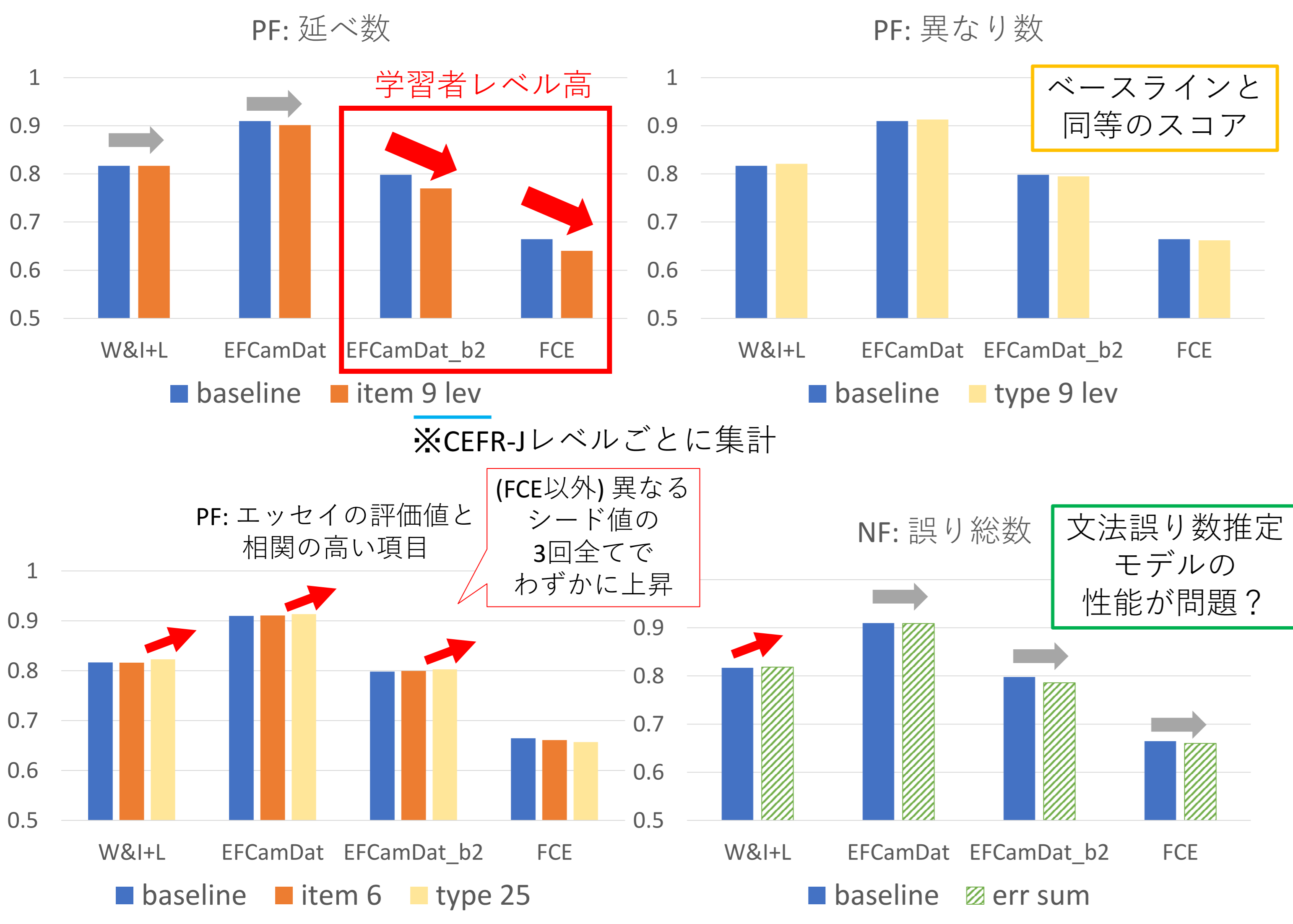
★ NFだけでなく, **PFを考慮したい**

文法誤り数推定モデル



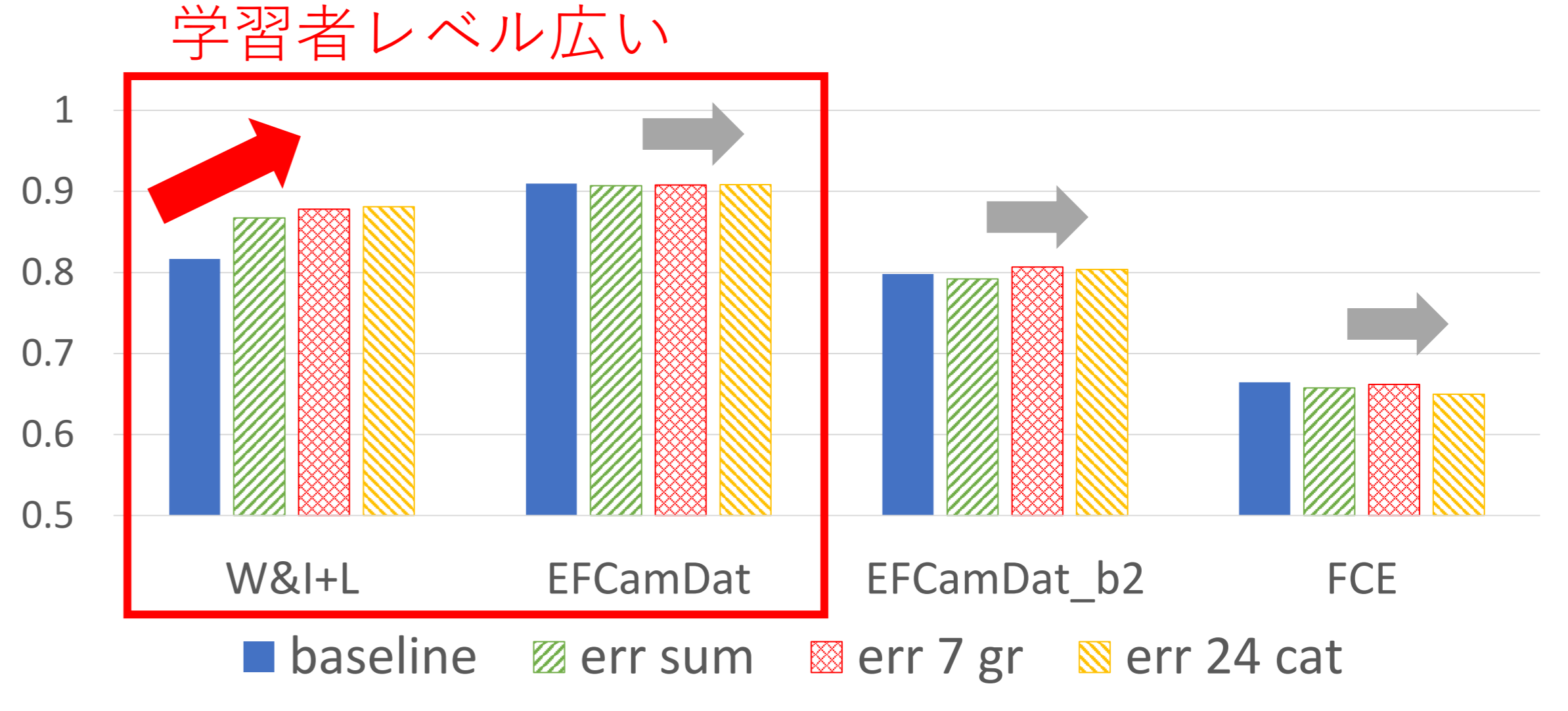
- 推定する誤りのカテゴリ数が増えると, 相関係数が低下
- 各文法誤りの頻度が低くなるものが影響

採点モデル



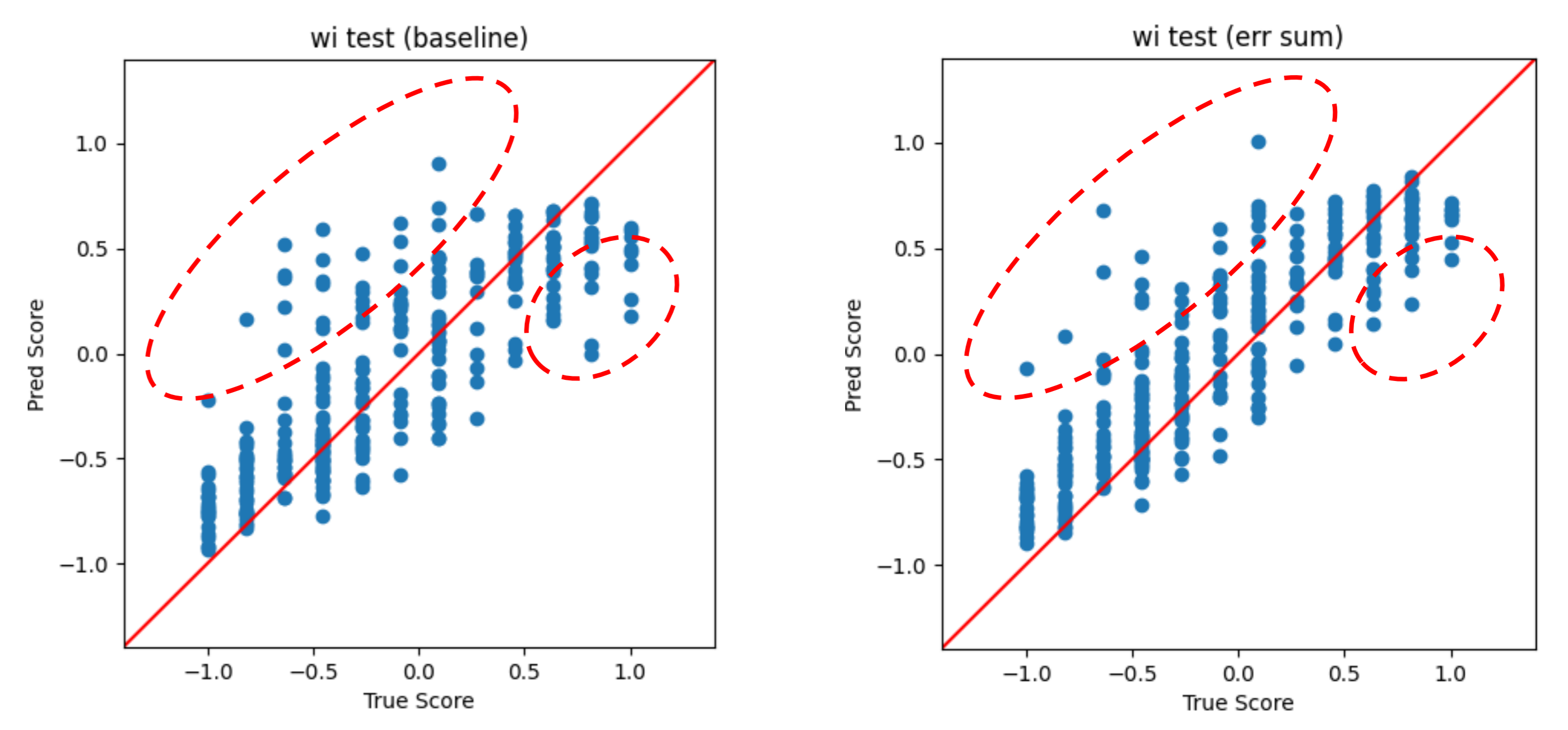
オラクルデータによるNFの検証

データセットに付与されている文法誤り情報を使用



- 学習者レベルが**広い**
- 学習者レベルが**高い**
- W&I+L
- EFCamDat
- 性能の向上なし
- アノテーションの一貫性・網羅性
- 文法誤りの数がW&I+Lより少ない

- 誤り情報は, 集計するよりもタグのままがよい
- 正解スコアと予測スコアの差異が大きい点で効果的



NFは, 採点性能の向上に寄与する
 ⇔しかし, 誤りの推定に課題

PFは, 本研究の活用手法では効果なし
 ⇒今後の課題