

モダリティ表現辞書の構築とその 日本語作文支援への応用

Bor Hodošček[†], Bekeš Andrej^{††}, 仁科喜久子[†]

[†]東京工業大学, ^{††}リュブリャーナ大学

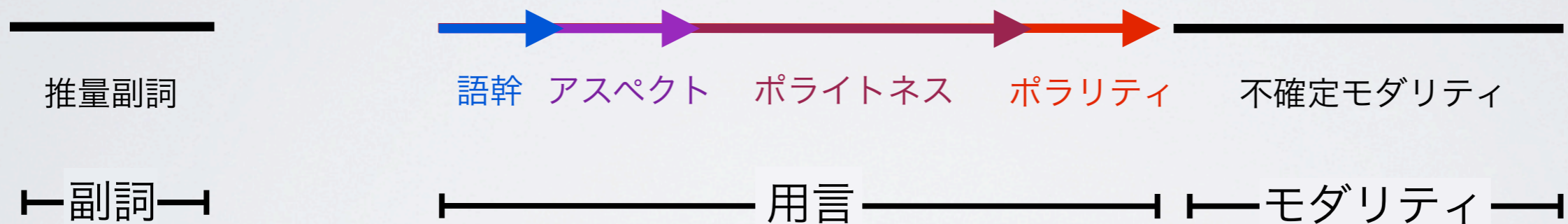
1. 背景 (1)

- 第2言語としての日本語学習者が作文を書く上で、命題よりも命題に対して自分の態度を表すモダリティの述べ方が困難である
 - ➡日本語はSOV型・膠着言語であるため、命題を述べてモダリティを付加しようとするとき、用言に接続するアスペクト、テンス、ポラリティに次いでモダリティ表現が並ぶ順位を把握する必要がある

1. 背景 (2)

- 用言にモダリティ表現が接続する例：

たぶん先生は覚えていらっしやらないかもしれないが、...

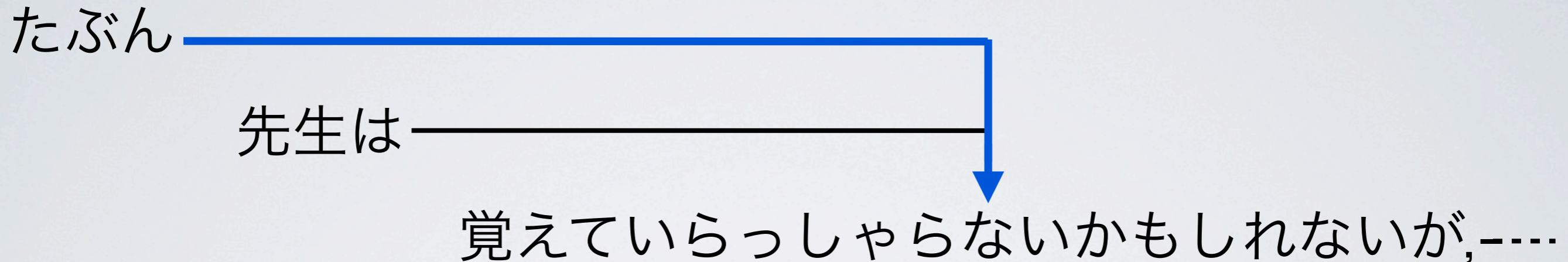


- 「たぶん」のような推量を表す副詞は文末に現れるモダリティを予測する力を持ち、それと共起する

2. 目的

本研究ではコーパスを用いてモダリティ表現と副詞・動詞の係り受け関係およびモダリティ表現と動詞に続くアスペクト， テンス， ポラリティなどの文法カテゴリーとの共起情報から日本語学習者に役立つ参考資源を提供する

3. 副詞-用言-モダリティ表現の抽出 (1)



(副詞,用言,モダリティ) →

(たぶん,覚えていらっしやらない,かもしれない)

- CaboChaで係り受け解析を行い, 抽出する

3. 副詞-用言-モダリティ表現の抽出 (2)

- モダリティ表現は一つ以上の文節からなるものが多数ある
 - ➡例：「に|違いない」「とは|いえない」など
 - ➡副詞の係り先のみではなく、その前後の文節も確認する必要がある
- 副詞-用言-モダリティ表現を文の構文構造から抽出する
- モダリティ表現または副詞-モダリティ表現のみよりも副詞-用言-モダリティ表現はある程度まとまったフレーズになっているため、学習者にわかりやすい形として提供できる

4. 用言とモダリティ表現

- モダリティ表現の分類を副詞との共起のみではなく、用言とでも行う
 - 用言を形成する形態素に対してアスペクト，ボイス，テンス，ポラリティ，ポライトネスの種類を判定する
- ➡です→ポライトネス， た→テンス， てい→アスペクト， られ→ボイスなど

5. データ

- モダリティの使用は文章における談話によって異なってくるため、それぞれの学習目的に沿ったドメインおよびジャンルに制限する：
 - 論文作成：科学技術論文（医学，言語処理など）
 - 一般文章作成：『現代日本語書き言葉均衡コーパス』
 - 新聞記事作成：毎日新聞
 - 百科事典記事作文：Wikipedia
 - ...など

今後の課題

- 推量副詞と用言によるモダリティ表現の分類
- モダリティ表現と「つつじ」機能表現辞書との統合
- モダリティ表現（または機能表現）を認識する構文解析の使用

デモ

<http://wombat.ryu.titech.ac.jp/natsume>（開発中）