

# 正規化を指向した機能動詞表現 の述部の言い換え

研究者

泉朋子 今村賢治 菊井玄一郎 藤田篤 佐藤理史  
日本電信電話株式会社 NTTサイバースペース研究所  
名古屋大学大学院工学研究科

発表者 江口晃

# はじめに

- **背景**

- 言語表現の意味を扱う際、意味的に同等な表現は集約したい
- 日本語においては機能動詞構造により、述部の表現が多様である
- 機能動詞構文を最も単純な言語表現にまとめて、述部の表現を縮退させる

- **研究の目的**

- 日本語の多様なLVCを単純な形へと言い換えることによりマイニング等の再現率向上が期待できる

# 機能動詞構文

- 機能動詞構文(Light Verb Construction;LVC)
  - 「動作性名詞+助詞+機能動詞」の構造を持つ
  - 動作性名詞が動詞の意味役割を果たす
  - 機能動詞がアスペクト/ヴォイス/ムードといった機能語の役割を果たす
- LVCの言い換え
  - LVCは「動詞+助動詞」という単純な形で言い換えることができる
  - 例:
    - 開発を続けている → 開発している
    - 締め出しをくろう → 締め出される

# 対象とするLVC

- 対象とする機能動詞

- 村木(1991)に記載されている143組の「格助詞+機能動詞」と、新聞記事19年分から得られた出現頻度上位40組の「格助詞+機能動詞」の、重複を除いた計160種類を使用する

- 下記の構造を正規化対象とした

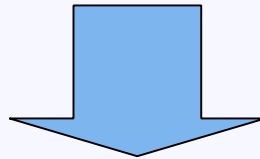
- 動作性名詞+格助詞+機能動詞

- 例： 拡充をはかる、注意を促す

# 制限言語の基本方針

- 制限言語の基本方針

- a. 制限言語は最も単純な「動詞+助動詞」の表現にする
- b. 正規化後の表現パターンを最小限にする
- c. 正規化前後で表現する事態の事実関係が変化しないようにする



- 動作性名詞を動詞化させる

例：変更 → 変更する

- 機能動詞が表す機能語の意味を助動詞に言い換える

# 機能動詞の意味役割と言い換え表現(1)

- 機能動詞の言い換え

- 160種類の機能動詞に対して、それぞれの機能表現に対応した助動詞を割り当てる

- 対応する助動詞が見つからない場合は、機能動詞間で同義のものをまとめ、出現頻度の一番高い機能動詞を制限言語にする

- 機能動詞が表すアスペクト表現は、アスペクト情報を欠落させ単純時制にまとめる

例： 会議を始める → 会議する

# 機能動詞の意味役割と言い換え表現(2)

表1:LVC正規化のための制限言語

制限言語	保持される機能表現	対象のLV	例
する	n.a.	106種類	供給を実施する → 供給する
される	ヴォイス(受動)	14種類	足止めをかう → 足止めされる
させる	ヴォイス(使役)	9種類	混乱をもたらす → 混乱させる
させられる	ヴォイス(受動+使役)	1種類	惨敗を喫した → 惨敗させられた
できる	ムード(可能)	4種類	判断ができた → 判断できた
しようとする	ムード(意思)	5種類	拡大を目指す → 拡大しようとする
させようとする	ヴォイス+ムード(使役+意思)	1種類	強化を促す → 強化させようとする
をやめる	アスペクト(終結相)	2種類	調査を打ち切る → 調査をやめる
となる	ムード(原因・結果)	5種類	削減につながった → 削減となった
がある	ムード(非動作主性)	6種類	不足が見られた → 不足があった

# 対象機能動詞と正規化ルール

- 正規化対象外

- 動作性名詞の意味に依存して、機能動詞/本動詞と変化するもの

- 例： 信頼を置く(=信頼する) 機能動詞

- 電話を置く(≠電話する) 本動詞

- 機能動詞が表すアスペクト表現と単純時制が表すアスペクト情報が対立する場合

- 例： A氏の殺害に取り掛かった [始動+過去]

- A氏を殺害した [過去(完了)]

# 自動正規化と評価(1)

- 評価方法

- 毎日新聞2003年1年分(392,865文)をテストデータとし、文末から「サ変名詞+格助詞+機能動詞」のLVCを抽出
- 正規化された10,005文からランダムに300文選ぶ
- 対象とする述部に焦点を当て、正規化条件のa~cが守られているか、また統語的に誤りがないかを検討する

# 自動正規化と評価(2)

表2: 述部自動正規化の精度

評価結果	割合(頻度)	例
正解	94.2%(261)	結論に達した → 結論した
複合名詞エラー	3.9%(11)	国交正常化実現しようとする
ヴォイスエラー	0.7%(2)	お祭りされる
IDIOMエラー	0.4%(1)	布石を打つ → 布石している
生成統語エラー	0.4%(1)	加速させようとしれている
意味エラー	0.4%(1)	偽造・変造に当たる ≠ 偽造・変造する

# 自動正規化と評価(3)

- 評価結果

- 誤りの多くはサ変名詞を含む複合名詞が動詞化されてしまったものであった

- サ変名詞を含む複合名詞が動詞化を分解する必要がある

- 例: 国交正常化を実現しようとする

- 正規化した述部を元の文に戻し一文としての精度を検討した結果、精度が53.0%と低下した

- 例: 条約比准と国内法の整備に取り組む

- 条約比准と国内法の整備する

- (→条約比准し、国内法を整備する)

# まとめ

- **研究の成果**

- 冗長なアスペクト表現を簡略化させた一方向の言い換えにより、より多くの述部をまとめることができた
- 約94%の精度で述部を単純な形に言い換えることができた
- マイニング等の精度を下げずに再現率向上を図る技術につながると考えられる

- **今後の課題**

- アプリケーションに応じたレベルでの、言い換えから正規化の自動変換ができるような技術の構築