

前提投射の実例のツリーバンクによる検索

窪田悠介 (筑波大学)、峯島宏次 (お茶の水女子大学)

1. はじめに

- 理論言語学研究でのコーパスの利用
- 「深い」自然言語処理研究のための言語資源の構築

→ どちらもある程度精緻な情報がついたコーパスが必要

例:

- 統語構造アノテーション (各種ツリーバンク)
- 意味情報アノテーション (FrameNetなど)
- 談話情報アノテーション (照応や談話構造など)

1.1 目的

- **前提投射**の実例をNPCMJツリーバンクから採取する
 - ツリーバンクの言語研究への活用の例を示す
 - 他のタイプの言語資源と比べたときのツリーバンクの利点
 - 他のタイプの言語資源と組み合わせる可能性

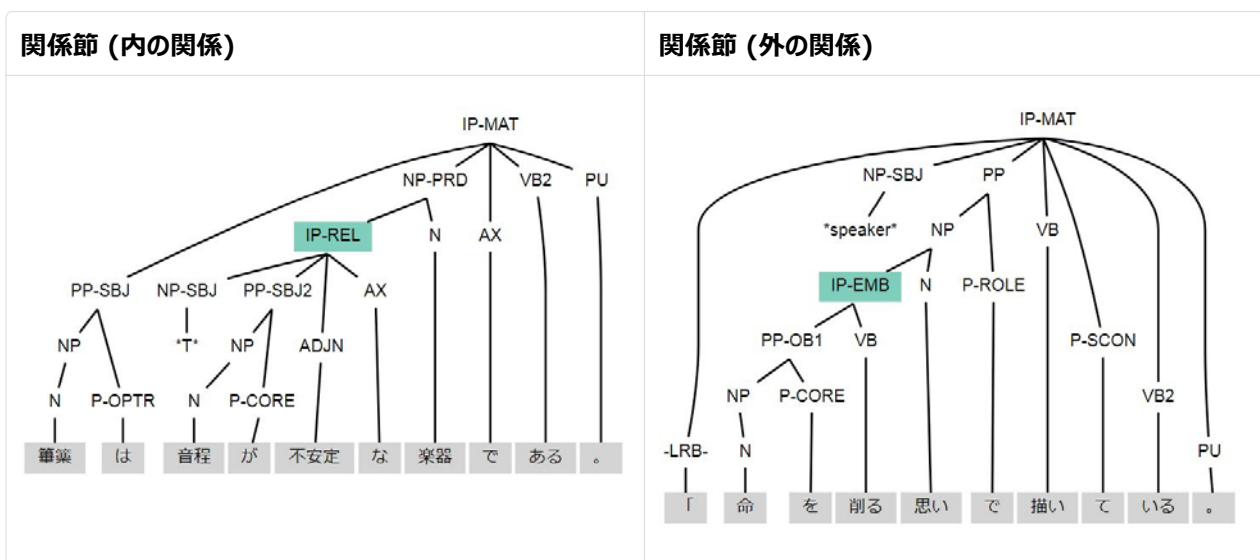
2. NINJAL Parsed Corpus of Modern Japanese (NPCMJ)

国立国語研究所で開発中の句構造ツリーバンク

- 現在2万文 (内、**実例1万文**) をwebで公開中 (プロジェクト終了時 (2021年度末) には6万文)
- インターフェース: <http://npcmj.ninjal.ac.jp/>

2.1 NPCMJの特徴

- 言語研究のためのリソースとして設計されている
- 文法関係、ゼロ代名詞などの細かなアノテーション



2.2 言語資源としてのNPCMJ

ジャンル	文数
新聞記事	4323
Wikipedia	2745
聖書	1652
その他実テキスト	1919
教科書例文	6051
辞書項目	3419
合計	20109

3. 前提投射

(1) ジョンが来た。

(2) a. ジョンも来たの? (疑問文)

b. ジョンも来たらしい。(モーダル)

c. ジョンも来たなら、さぞ賑やかだっただろう。(条件節)

d. ジョンも来たという噂が街に流れた。(埋め込み)

- (2)a-dは(1)を前提とする (Karttunen & Peters 1979, Kadmon 2001)。

4. 前提投射の実例の検索

4.1 探したい構造:

[疑問文 ... **Xも** ...]

[... **Xも** ... モーダル]

[[... **Xも** ... たら/なら] ...]

[... [... **Xも** ...]と/ことを/のを V]

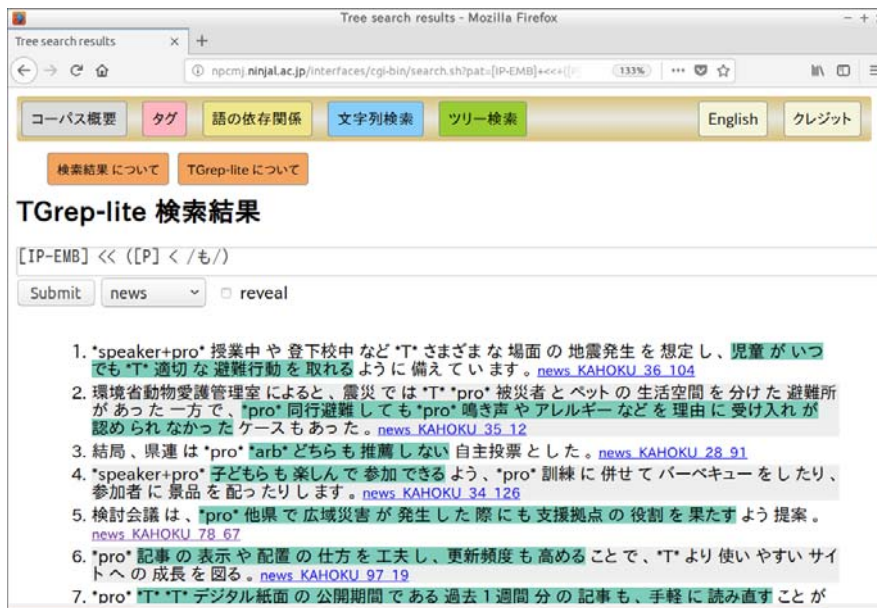
→ ツリーバンクを使うと構造を指定して検索できる。

4.2 手法

- 利用するコーパス: NPCMJ web公開版 (実例1万文)
- 調査対象:
 - 前提トリガー: 「も」「まで」「再び」「さらに」(累加的意味を表す取り立て詞・副詞)
 - 埋め込み環境: 埋め込み節、条件節、文末モーダル
- 検索式の例: `[IP-EMB] << ([P] < /も/)`
- 調査方法:

- 窪田と峯島でまず別々に検索結果を分類
- 結果を付き合わせて調整

4.3 検索画面



4.4 結果 (概要)

タイプ	文数
yes	60
local	6
factive	23
factive, local	7
ambig	11
n/a	281
total	388

- yes: 前提投射の例
- local: ローカルな環境で前提が満たされている例
- factive: 叙実的環境に埋め込まれているため投射している例
- factive, local: 叙実的環境に埋め込まれているが、ローカルな環境で前提が満たされている例
- ambig: 統語的、または語彙的曖昧性により、前提投射が起こっているかの判別が不能な例
- n/a: 「も」や「まで」の多義性などの理由で、前提投射に該当しない例

内訳:

- yesの内訳は、「も」44件、「まで」1件、「再び」7件、「さらに」8件だった (合計60件)。
- n/aの内訳は、「も」201件、「まで」74件、「再び」1件、「さらに」5件だった (合計281件)。

データの公開:

以下のURLに結果のデータを公開してある。

<http://www.u.tsukuba.ac.jp/~kubota.yusuke.fn/papers/l sj-presup-data.csv>

4.5 結果 (例文)

4.5.1 yesの例

埋め込み節

パターン: [... [... **Xも** ...] COMP V]

(3) 受賞を「能を支える東北の多くの愛好家仲間に代わってお受けしたい」と喜び、これから**も**能楽の振興に尽力することを誓う。 (37_news_KAHOKU_73;K201401010A0F50XX00001;JP)

(4) 検討会議は、他県で広域災害が発生した際にも**支援拠点の役割を果たすよう提案。**
(67_news_KAHOKU_78;K201401010A0FZ0XX00002;JP)

パターン: [... [... **さらに** ...] COMP V]

(5) 14年は誘致実現へ向け、土台を**さらに**固める大事な年になる**見通しだ。**
(8_news_KAHOKU_89;K201401040A0F70XX00001;JP)

条件節

パターン: [[... **Xも** ... たら/なら] ...]

(6) ベテラン農家**も**加われば、長年の経験で培われた技術を共有できる。
(38_news_KAHOKU_1106;K201401010A0M106B00004;JP)

文末モーダル

パターン: [... **Xも** ... モーダル]

(7) 各地の震災遺構の保存や活用について**も**新しいアイデアが生まれる**かもしれない**
(99_news_KAHOKU_74;K201401110A0T20XX00001;JP)

パターン : [...さらに...モーダル]

(8) このような中の下の所得の国は新興国として経済的に**さらに**発展するでしょう

(56_ted_talk_7;HansRosling_2010S;43;JP)

4.5.2 時間軸上の変化を表す埋め込み節

(9) 臨済宗が**再び**活気に満ち溢れる**ようになる**のは、北宋代であり、[...] (9_wikipedia_KYOTO_2;BDS_00002;9;JP)

(10) 当初は平日のみだったが、2006年3月18日のダイヤ改正から**土曜・休日にも**運転される**ようになった**。

(46_wikipedia_KYOTO_22;RLW_00004;60;JP)

(11) 沿岸部は入院ベッドが**今も減ったままだ**。(14_news_KAHOKU_93;K201403050A0T10XX00201;JP)

4.5.3 localの例

パターン : [... [... Xも ...] COMP V] (埋め込み節)

(12) 彼らの夢は自転車を買う事であり**後になってバイクも買う事**です (40_ted_talk_7;HansRosling_2010S;32;JP)

4.5.4 factiveの例

(13) 私が**再び福岡に向かった**のは、所事務所で再生スキームの概略が決まってから数日後のことである。(埋め込み節)

(2_book_excerpt-30;JP)

(14) スマートフォンやタブレット端末から**も利用できるため**、外出先で思いついたことがあればその場で書き込める。(埋め込み節)

(29_news_KAHOKU_87;K201402230A0T10XX00001;JP)

4.5.5 factive + local の例

(15) 仙台市都心部周囲には広瀬川や青葉山などの自然があり、また都心部にも街路樹などの緑が多いことから、「杜の都」との別名をもつ。(埋め込み節) (3_wikipedia_Sendai_City;JP)

(16) ヘプバーンは母国語として英語とオランダ語を、**さらに**フランス語、イタリア語、スペイン語、ドイツ語を流暢に操ることができた**ため**である。(埋め込み節) (236_wikipedia_Audrey_Hepburn;JP)

4.5.6 曖昧性のため判定が困難な例

統語的曖昧性

(17) 「あなたがた**も**去ろうとするのか」(文末モーダル) (507_bible_new;b.JOH.6.67;JP)

(18) 同時に、警報を国から自治体へ、その先の市民へ、地震で寸断されることなく伝達する通信システムの開発と整備**も**進める**必要がある** (81_news_KAHOKU_63;K201401110A0T10XX00001;JP)

語彙的曖昧性

(19) しかし、それでも、国語の苦手な子には、静かに本と向きあう時間が苦痛になること**も**考えておかなくてはなりません。

(5_book_excerpt-28;JP)

4.5.7 「も」のゆるい例

(20) 仙台市は、大企業の支店の存在に経済を支えられている側面が強く、また**工業も**発展していない**ため**、経済関係者の中から「支店経済都市」「支社都市」と呼ばれている。(217_wikipedia_Sendai_City;JP)

(21) 記事の表示や配置の仕方を工夫し、**更新頻度も**高める**こと**で、より使いやすいサイトへの成長を図る。

(19_news_KAHOKU_97;K201402230A0T20XX00001;JP)

(22) 近畿日本鉄道にも**近鉄奈良線があるため**、「JR奈良線」と案内されることもある。

(13_wikipedia_KYOTO_22;RLW_00004)

(23) フェイスブックは皆さんもご承知の通り急速に拡張している通信媒体です
(18_ted_talk_8;StefanaBroadbent_2009G;JP)

4.5.8 多重の埋込み

(24) オランダは第一次世界大戦では中立国であり、**再び起ころうとしていた**世界大戦でも中立を保ち、ドイツからの侵略を免れることができ**と思われていた**ためである。(46_wikipedia_Audrey_Hepburn;JP)

(25) 管理責任者の五十嵐桂樹さん(62)は「震災の記憶を薄れないようにし、**再び災害があった場合**の対応について考える**きっかけにしてもらえれば**」と話す。(21_news_KAHOKU_39;K201403130A0T10XX00001;JP)

4.6 考察、今後の課題

考察

- 「よくなる」(時間的変化)のようなものが前提投射を引き起こす。
- 多重の埋込み
- 累加の「も」といわゆる詠嘆の「も」の間の境界は明確か? (cf. 4.5.7節、沼田(1995, 2009))
 - 累加の「も」の前提命題も、必ずしも文自体に現れている述語文字通りではない可能性がある。
 - 例えば、(22)では、「近畿日本鉄道」とJRが対比されており、「(「JRに近鉄奈良線がある」ではなく)「JRに奈良線がある」というのが前提命題となっている。

今後の課題

- 他の環境や他のトリガー、他の「投射的意味」(Potts(2005)のCIなど)を調べる
 - 環境: 疑問文、否定、引用節
 - トリガー: 叙実動詞、分裂文、所有格名詞
 - 他の「投射的意味」: 非制限的関係節、評価副詞、程度副詞、尊敬語
- 生起頻度が低いトリガーの検索方法
 - BCCWJを機械的に全文パースして検索する
 - 単純な文の埋め込み関係ならば、機械的にパースしただけのデータでもそれなりの精度が期待できる

参考文献

Kadmon, Nirit (2001) *Formal Pragmatics*. Blackwell. // Karttunen, Lauri and Stanley Peters (1979) Conventional implicature. In C.-K. Oh and D. Dinnens, eds., *Syntax and Semantics*, vol. 11, 1-56. Academic Press. // Potts, Christopher (2005) *The Logic of Conventional Implicatures*. Oxford: Oxford University Press. // 沼田善子 (1995) 「現代日本語の「も」」『「も」の言語学』ひつじ書房. // 沼田善子 (2009) 『現代日本語とりたて詞の研究』ひつじ書房.

謝辞

- 本研究には以下のコーパスを利用しました。国立国語研究所『NINJAL Parsed Corpus of Modern Japanese』<http://npcmj.ninjal.ac.jp/interfaces/> (2018年9月30日確認)
- 本研究はJSPS科研費 18K00523, 15H03210 の助成、および国立国語研究所「統語・意味解析コーパスと言語研究」プロジェクトの支援を受けています。