

日本語音声対話文の統計的係り受け解析とその評価

大野 誠寛†

松原 茂樹‡

河口 信夫‡

稲垣 康善†

†名古屋大学大学院工学研究科情報工学専攻

‡名古屋大学情報連携基盤センター/CIAIR

ohno@inagaki.nuie.nagoya-u.ac.jp

1 はじめに

ユーザフレンドリな音声対話システムの実現のために、実対話環境下において利用可能な口バストな音声理解技術が望まれる。現状の音声対話システムの多くは、あらかじめ定められた単語や言い回しで構成された発話文への対応にとどまっておき、より豊かな対話の実現のためには、自然発話の言語構造の解析が不可欠である。一方、対話音声には、省略や倒置など、従来の書き言葉に対する係り受け解析では捕えられない現象を含む発話があり、無視できない頻度で出現する [3]。

このような背景のもと、著者らは、統計情報を用いた係り受け解析手法を提案している [2]。本稿では、この手法の口バスト性について評価する。特に、係り先がない文節や前方に係る係り受け、ポーズをまたぐ係り受けに対する本手法の効果について報告する。

2 話し言葉の統計的係り受け解析

本手法では、形態素解析及び文節まとめ上げが施された文節列を入力とする。入力文節列 $B (= b_1 \cdots b_n)$ の係り受け構造を S とするとき、 $P(S|B)$ の確率値を最大にする係り受け構造 S を求める。

通常、書き言葉に対する係り受け解析手法では、係り受けの非交差性、後方修飾性、係り先の唯一性の3つの性質を、絶対的制約として用いるが、本手法では、倒置やフィラー、言い淀み、言い誤りなどが頻出することに注目し、非交差性のみを、満たすべき性質として係り受け構造を求める。ただし、後方修飾性、及び、係り先の唯一性の充足については、統計情報を反映することにより考慮する。

それぞれの係り受けは独立であると仮定すると、 $P(S|B)$ は以下の式で計算できる。

$$P(S|B) = \prod_{i=1}^n P(i \xrightarrow{rel} j|B) \quad (1)$$

ここで、 $P(i \xrightarrow{rel} j|B)$ は、入力文節列 B が与えられたときに、文節 b_i から b_j への係り受け関係がある確率を表す。最尤の係り受け構造は、式 (1) の確率を最大とする構造であるとして動的計画法を用いて計算する。

次に、 $P(i \xrightarrow{rel} j|B)$ の計算について述べる。まず、係り文節における自立語の原形を h_i 、その品詞を t_i 、係りの種類を r_i とし、受け文節における自立語の原形を

表 1: 係り受け全体の実験結果

項目	係り受け正解率	ターン正解率
正解率	87.0%(21,089/24,250)	70.1%(4,260/6,078)

h_j 、その品詞を t_j とする。また、文節間距離を d_{ij} 、文節間のポーズの数を p_{ij} 、係り文節の位置を l_i とする。ここで、係りの種類とは、係り文節が付属語を伴うときはその付属語の語彙、品詞、活用形であり、そうでない場合は一番最後の形態素の品詞、活用形である。また、係り文節の位置は、その文節が入力ターン内で一番最後の文節か否かを表す。

以上の属性を用いて、確率 $P(i \xrightarrow{rel} j|B)$ を以下のように計算する。

$$P(i \xrightarrow{rel} j|B) = \frac{C(i \rightarrow j, h_i, h_j, t_i, t_j, r_i, d_{ij}, p_{ij}, l_i)}{C(h_i, h_j, t_i, t_j, r_i, d_{ij}, p_{ij}, l_i)} \quad (2)$$

ただし、 C は共起頻度関数である。係り先のない文節はそれ自身に係る (すなわち $i = j$) とみなすことにより、係り先をもたない場合の確率も計算できる。

3 解析実験

本手法の有効性を評価するため、係り受け解析実験を行った。実験には、名古屋大学 CIAIR 車内音声対話コーパス [1] に係り受け解析を与えて構築した音声言語係り受けコーパス [3] を用いた。

3.1 実験の概要

データとして 81 対話における運転者の発話を使用した。その規模は、24,250 文節からなる 6,078 ターンである (平均ターン長は 4.0 文節)。対話ごとに分割し交差検定を行った。すなわち、81 対話におけるある 1 対話をテストセットとし、残りの対話を学習セットとする実験を 81 回繰り返した。なお、本手法では 1 ターンを解析の単位とした。

3.2 実験の結果

係り受け、及び、ターンに対する正解率を求めた。係り受け全体に対する実験結果を表 1 に示す。正解データの係り受けと一致したものが 21,089 個あり、正解率は 87.0% である。このことから、本手法により、自然発話文に対しても、書き言葉を対象とした他の係り受け解析手法 [4] と同等の高い精度で係り受けを抽出できることを確認した。

4 考察

本節では、受け文節のない係り受け、前方の文節への係り受け、及び、ポーズをまたぐ係り受け現象に着目し、

Evaluation of Stochastic Dependency Parsing for Japanese Spoken Language: Tomohiro Ohno, Shigeki Matsubara, Nobuo Kawaguchi and Yasuyoshi Inagaki (Nagoya University)

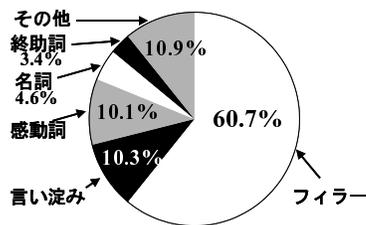


図 1: 最終文節以外で受け文節がない係り受けの内訳

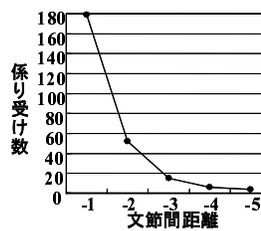


図 2: 前方の文節への係り受けの文節間距離

表 2: 受け文節のない係り受けの解析結果 (ポーズの直前でなく、Filler・言い淀み以外)

適合率	60.4%(996/1,650)
再現率	69.5%(996/1,434)

表 3: 前方の文節への係り受けの解析結果

適合率	60.5%(49/81)
再現率	19.1%(49/256)

表 4: ポーズをまたぐ係り受けの解析結果

適合率	6.5%(34/521)
再現率	37.0%(34/92)

自然な話し言葉に対する本手法のロバスト性について、前節の実験結果をもとに考察する。

4.1 受け文節のない係り受け

通常の書き言葉では、最終文節以外は 1 つの受け文節をもつが、話し言葉では、Filler や言い淀みなど、必ずしもすべての文節に受け文節が存在するとは限らない。実験で使用したコーパスでは、全文節の 51.1% に相当する 12,384 文節には受け文節がなく、そのうちポーズの直前に位置しないものは 4,937 文節であった。その内訳を図 1 に示す。その約 7 割が Filler 及び言い淀みである。これらについては受け文節がないとしてコーパスが作成されており、その特定は難しくない。そこで残りの 3 割の係り受けに対する実験結果を表 2 に示す。それらを除く 1434 文節に対して 996 文節正解しており、それらの特定についても高い正解率を示している。

4.2 前方の文節への係り受け

本手法では、倒置関係の同定を可能にするために、後方修飾性を絶対的な性質として定めなかったが、係り受け解析において、倒置、すなわち、前方修飾性を許すと、係り先を特定するための探索空間が約 2 倍に広がるため、正しい解析の実現が困難になる。前方の文節への係り受けは 256 個存在し、必ずしも無視できる数ではないが、割合としては全体のわずか 1% 程度に過ぎず、そのような文節を特定することの意義は必ずしも明らかではない。

前方の文節への係り受けに関する結果を表 3 に示す。再現性については必ずしも高いとはいえないものの、適合率は 60% を超えている。これは、前方への係り受けを許すことにより、解析精度が上昇することを意味しており、ひいては本手法の倒置現象へのロバスト性を示している。このような良好な結果が得られた理由として、倒置の出現に関して以下の 2 つの傾向が存在することが挙げられる。

一つは、文節の位置に関する出現傾向である。すなわち、倒置となる係り受け関係における係り文節の多くは、発話ターンの最終文節に出現する (倒置全体の 85.2%) ことを考慮し、係り受け確率を計算する式 (2) の属性として係り文節の位置を導入したことの効果が現れている。実際、前方への係り受けであるとして同定された 81 個のうち、係り文節が発話ターンの最後に位置している係り受けの適合率は 75.0% であった。

もう一つは、文節間距離に関する出現傾向である。前方の文節への係り受けの文節間距離は、全体の 90.2% が -1 あるいは、-2 のいずれかであり (図 2 参照)、それは、係り受け確率の計算においても d_{ij} が負の値をとることを許すことにより、反映されている。解析結果では、倒置のうち、文節間距離が -2 以上の係り受けの適合率は 61.0% であった。

4.3 ポーズをまたぐ係り受け

話し言葉では、書き言葉でいう文に相当する文法単位を特定するのは容易ではない。ポーズがそのような単位の境界を形成している可能性は高いが、そうでない場合も少なくないので、本実験ではターンを解析単位としていた。実験データには 92 個のポーズをまたぐ係り受けが存在した。

ポーズをまたぐ係り受けに関する実験結果を表 4 に示す。適合率は著しく低い結果となった。この理由として、そもそもそのような係り受けの出現頻度が多くない (全体の 0.4%) ことと、ポーズをまたぐ場合の文法的特徴が必ずしも明らかでなく、それを確率の計算式に導入していないことが挙げられる。その一方で、37.0% の再現性を実現しており、ある程度の効果は見られた。

5 おわりに

本稿では、著者らが提案する統計的係り受け解析手法のロバスト性について評価した。その結果、特に、前方の文節への係り受けや受け文節のない係り受けを含む発話に対する本手法の有効性を確認した。現在、係り受けコーパス大規模化作業を進めており、今後、その解析精度向上への効果について調査する予定である。

参考文献

- [1] Kawaguchi, N., Matsubara, S., Takeda, K. and Itakura, F.: Multimedia Data Collection of In-Car Speech Communication, *Proc. of 7th Eurospeech*, pp. 2027-2030 (2001).
- [2] Matsubara, S., Murase, T., Kawaguchi, N. and Inagaki, Y.: Stochastic Dependency Parsing of Spontaneous Japanese Spoken Language, *Proc. of 19th COLING*, Vol.1, pp. 640-645 (2002).
- [3] 大野 誠寛, 松原 茂樹, 河口 信夫, 稲垣 康善: 統計的構文解析器を用いた音声言語係り受けコーパスの構築, 言語処理学会第 9 回年次大会論文集 (2003).
- [4] 内元 清貴, 関根 聡, 井佐原 均: 最大エントロピー法に基づくモデルを用いた日本語係り受け解析, 情報処理学会論文誌, Vol.40, No.9, pp.3397-3407 (1999).