

制約と統計に基づく自然な発話の解析

伝 康晴

ATR 音声翻訳通信研究所

1 はじめに

近年、音声処理と言語処理の境界領域として、自然な発話 (spontaneous speech) の解析の研究が盛んである。しかし、自然な発話では、言い淀み、言い直し、言い誤り、構成素の欠落などが多く見られ、その機械的処理を困難にしている。このような不適格文を解析する試みはこれまでにもなされているが、その大半は、システムが不適格文に出会った時に解析の失敗から回復するための機構に焦点をあてたものであり [1] [2]、実際の発話の分析に基づいて実用性のある手法を提案した例は少ない。また、実際の発話を対象とした研究 [3] でも、言い直しなど特定の種類の不適格性を扱う専用の手続きについて論じたものばかりで、発話から例えば格構造を得る過程を総合的に扱った研究はない。

本稿では、自然な発話を解析するための一般的な枠組を提示する。本研究では、発話の解析を「(音声発話から書き起こされた) 漢字仮名まじり文を入力とし、その文の深層格構造を出力すること」と定義する。したがって、本枠組は、係り受け構造や係り受け関係の決定など、通常の適格文の処理で必要とされる機能をすべて備えている。言い直しの発見や欠落した構成素の補完などは、専用の手続きによってではなく、一般的な発話解析過程の中で自然に実現されている。この一般的な発話解析の機構は、「制約に基づく解釈」と「統計に基づく選好」という2つのモジュールからなる。本枠組では、自然な発話に見られる不適格性を始めとするさまざまな言語現象を形態・統語・意味に関する制約によってとらえ、発話解析はこの制約に基づく尤度付きの探索によって行なう。おのおのの解釈候補の尤度は、コーパスから抽出した係り受け関係の統計情報に照らし合わせて、その解釈がどのくらいもっともらしいかを計算して決める。

以下、本研究の背景と本枠組の詳細を説明し、本枠組の有効性を実際の発話を解析して検証する。

2 自然な発話の解析

自然な発話の解析を困難にする要因の1つは言い直しである。村上ら [4]によると、ATR 対話データベースにおいて、1割近くの文に言い直しが存在する。こういった言い直しを含む文の解析には、以下のような問題点がある。

1. 言い直しには助詞のような明示的なマークがないのでその発見が困難である。
2. 言い直されている句の範囲を決定するのが困難である。
3. 言い直しの種類(原因)は多様であり、その発見にはさまざまな観点から文をながめる必要がある。例えば、(1)、(2)、(3)はそれぞれ、形態的な言い直し、意味的な言い直し、統語的な言い直しと考えられる。

- (1) この一、通、通訳電話に関するさまざまな、あの一、方面から
 - (2) あ、会議ではもちろん通訳、翻訳も入れますので、
 - (3) あ、クレジットカードをね、あの一、クレジットカードの名前となんかナンバーを
- さらに、言い直しの中には適格文と簡単には区別できないものも存在する。

- (4) きょう、協賛する学会会員の特別割引
- 「きょう」は「協賛する」と言い直されているのであるが、「きょう」がたまたま適正な単語(「今日」)であるため、適格文としての解釈も可能である。佐川ら [5]によると、このような適格文との間で曖昧性がある例は、ATR 対話データベースに含まれる言い直しのうちの約1割を占めている。

ここで見られる言い直しの問題点は、実際のところ、格構造解析の際に生じる問題点と共通している。すなわち、これらは、格要素の係り先の決定、係り語の修飾要素の範囲の決定、格関係の種類の決定と同じ問題である。したがって、言い直しの扱いには一般的な格構造解析と同等の能力が必要であ

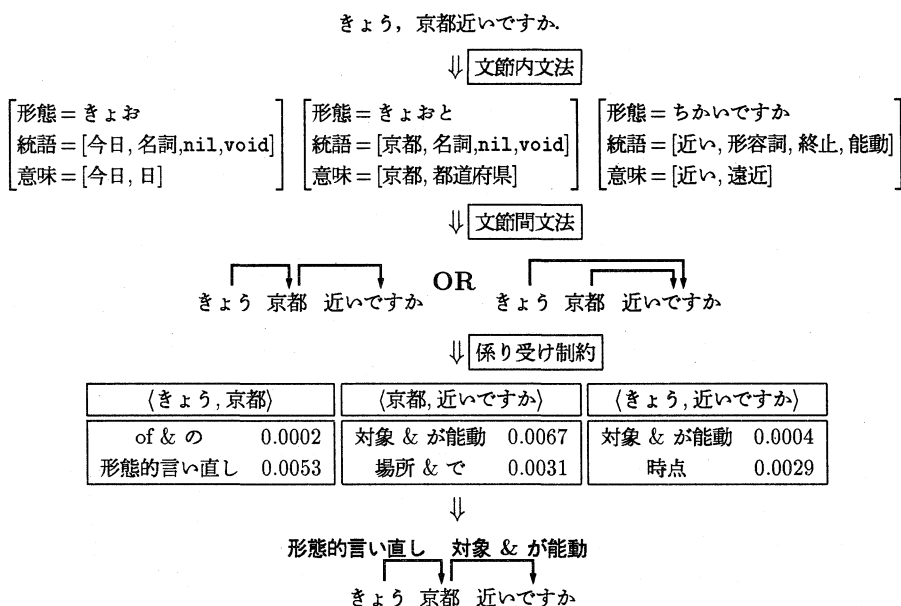


図1 制約に基づく解釈

る。特に、(4)のような例が多数存在することは、言い直しの処理と格構造解析とが独立にはなし得ないことを物語っている。

そこで、本稿で提示する枠組では、言い直しを一種の係り受け関係としてとらえる。例えば、(2)では、「通訳」と「翻訳」との間に「意味的言い直し」という種類の係り受け関係が成り立っていると考える。言い直しを係り受け関係ととらえることによって、言い直しの処理を一般的な格構造解析の過程で自然に実現することができる。なお、ここでは、説明の都合上、言い直しに限定して話をすすめてきたが、言い淀み、言い誤りや助詞の欠落についてもこの枠組で扱える。

3 制約に基づく解釈

本枠組が持つ一般的な格構造解析機能を実現する「制約に基づく解釈」モジュールについて説明する。ここでいう制約とは、文節内文法、文節間文法、および、係り受け関係を規定する記号的な規則のことであり、いずれも形態、統語、意味情報に関する宣言的な規則として記述されている。

入力文が与えられた時の解釈の過程は、図1のようである。まず、入力文は文節内文法をもとに文節

情報列に変換される。おのおのの文節情報は、形態情報(音節列)、統語情報(語彙、品詞、格、態)、意味情報(概念、属性)からなる。次に、この文節情報列が文節間文法をもとに係り受け構造に変換される(この時点では構造的な多義が存在する)。次に、おのおのの係り受け構造に含まれるすべての係り受け関係(係り語と受け語のペア)に対して、係り受け制約をもとに可能な解釈が与えられる。おのおのの解釈には尤度がついており、最終的には、尤度をもっとも大きくなるような係り受け解釈の組が選択される。

係り受け制約はホーン節によって宣言的に記述されており、例えば、格関係を規定する制約や言い直しを規定する制約には(5)や(6)が含まれる。

(5) 係り受け(x, y) \rightarrow 対象(x, y) \wedge が能動(x, y).

(6) 係り受け(x, y) \rightarrow 形態的言い直し(x, y).

ここで、 x, y にはそれぞれ、係り語、受け語の文節情報が入る。

“対象(x, y)”や“が能動(x, y)”といった要素式は選択制限や統語的な制限によって定義されるが、そのような条件を、例えば、詳細な意味マーカを用いて記述するといったことは行なっていない。代わりに、これらの要素式は尤度をともなって成立するも

| 関係 | 係り語 | 受け語 | 頻度 |
|---------|----------------------|-------------------|----|
| が能動 | [ホテル, 名詞, nil, void] | [遠い, 形容詞, 終止, 能動] | 5 |
| を能動 | [英語, 名詞, を, void] | [話す, 動詞, 終止, 能動] | 8 |
| 対象 | [ホテル, 店舗] | [遠い, 遠近] | 5 |
| 対象 | [英語, 言葉] | [話す, 発言] | 8 |
| 形態的言い直し | どおじ | どおじつうやく | 1 |
| 意味的言い直し | [通訳, 翻訳] | [翻訳, 翻訳] | 2 |

図2 コーパスから抽出した係り受けデータ

のとして扱う。尤度は次節で説明する「統計に基づく選好」モジュールにおいて計算される。例えば、「きょう」と「近い」の関係を「対象」とする解釈の尤度は、「京都」と「近い」の関係を「対象」とする解釈の尤度より低い。また、「きょう」と「京都」の関係を「形態的言い直し」とする解釈の尤度は、それを「of」とする解釈の尤度より高い。このように、尤度を比較することによって、最適な解釈を選択することができる。

4 統計に基づく選好

「制約に基づく解釈」モジュールにおいて作られる解釈候補に対して、その尤度を与えるのが「統計に基づく選好」モジュールである。尤度は、その解釈が生じる頻度確率を計算することによって与える。頻度情報の拠り所として、自然な発話のコーパスからあらかじめ人手で係り受け関係を抽出したデータ(図2)を用いる。係り語 α と受け語 β の間に関係 π が成り立つ尤度、すなわち、要素式 $\pi(\alpha, \beta)$ の尤度 $L(\pi, \alpha, \beta)$ は、次式で与えられる。

$$(7) L(\pi, \alpha, \beta) = \frac{\pi(\alpha, \beta) \text{ の頻度}}{\sum_{p, x, y} p(x, y) \text{ の頻度}}$$

しかし、このままではコーパスの希薄性の問題を避けられないので、 $\pi(\alpha, \beta)$ の頻度を補間してやる必要がある。ここでは、 $\pi(\alpha, \beta)$ に類似した事例の頻度を適当に加味してやることによって、 $\pi(\alpha, \beta)$ の頻度を補間する。例えば、解釈候補“対象(京都, 近い)”の頻度を計算する際に、これと類似した事例“対象(ホテル, 遠い)”の頻度を考慮に入れるということである。ここで注意すべきことは、どの関係に着目しているかによって類似性の判断基準が異なるということである。例えば、「対象」のような深層格関係の場合には候補・事例間の意味的な類似性

が基準になるであろうし、「が能動」のような表層格関係の場合には候補・事例間の統語的な類似性が基準になるであろう。また、「形態的言い直し」のような関係の場合には、言い直しパターンとしての類似性を調べる必要がある(例えば、「きょう」を「京都」で言い直すパターンと、「同時」を「同時通訳」で言い直すパターンは類似している)。本研究では、 $\pi(\alpha, \beta)$ と $\pi(x, y)$ の類似性の判断基準として以下のものを用いている。

π が深層格関係の場合 $\langle \alpha, x \rangle$ と $\langle \beta, y \rangle$ のそれぞれの意味情報に関して、シソーラスに基づく類似度[6]を求めたものの相乗平均。

π が表層格関係の場合 統語情報に関するシソーラスを定義して同様な計算を行なったもの。

π が言い直しの場合 $\langle \alpha, \beta \rangle$ と $\langle x, y \rangle$ のそれぞれの類似度(=言い直しパターンを類別する指標)の差を取ったものの1からの残余。すなわち、
 $1 - \text{abs}(\text{sim}(\alpha, \beta) - \text{sim}(x, y))$ 。

さらに、解釈候補に対して類似度 s を持つ事例の頻度をどのくらいの重みで寄与させるかについて、重み w を s の(3次の)多項式関数として与える。

$$(8) w(s) = a_0 + a_1 s + a_2 s^2 + a_3 s^3$$

ここで、係数 $a_k (k = 0, 1, 2, 3)$ は、コーパスから抽出した係り受けデータを学習事例として、関係名ごとに交差検定によってチューニングしている。

以上より、尤度の定義は次式のようにになる。

$$(9) L(\pi, \alpha, \beta) = \frac{\sum_{x, y} w(s) \times \pi(x, y) \text{ の頻度}}{\sum_{p, x, y} p(x, y) \text{ の頻度}}$$

ただし、 s は $\pi(\alpha, \beta)$ と $\pi(x, y)$ の類似度

5 実験

本枠組の有効性を検証するために小規模な実験を行なった。コーパスとしてATR自動翻訳電話研究所の対話データベース(国際電話の申し込みに関

表1 実験結果

| | クローズ試験 | オープン試験 |
|----|------------|------------|
| 正解 | 84 (75.0%) | 58 (51.8%) |
| 誤り | 23 (20.5%) | 37 (33.0%) |
| 失敗 | 5 (4.5%) | 17 (15.2%) |
| 合計 | 112 (100%) | 112 (100%) |

する参加者と事務局の対話)を用い、そのうちの10対話から人手で4168件の係り受けデータを抽出した。実験は、10対話を統計データとしそのうちの2対話を試験データとしたクローズ試験と、同じ2対話を試験データとし残りの8対話を統計データに用いたオープン試験を行なった。現在のところ、文節内処理は人手で行なっており(入力文は文節情報列)、1文節発話は試験データから除いている。

結果を表1に示す。「正解」は1発話中のすべての係り受け構造と係り受け解釈が人手でつけたものと完全に一致した場合、「誤り」は係り受け解釈もしくは係り受け構造に誤りがある場合、「失敗」は解がでなかった場合である。ここでは、言い直しの検出だけでなく通常の連体格や連用格の解釈が正しいことも求めており、評価基準がかなり厳しくなっていることを考えると、この段階としてはまずまずの結果と言える。クローズ試験での失敗は、述語欠落を含む文であり、現時点では文節間文法がこれに対応していない。オープン試験で増えた失敗は、統計データが少ないため前節の補間法を用いても尤度が0になる解釈しか見つからなかった文である。

「誤り」として多く見られるのは、(10)のような文で「会議」の係り先が「主旨」(正解は「テーマ」)と解釈されて言い直しの修飾部分に誤って含まれてしまう場合と、(11)のような文で「オーバーヘッドプロジェクト」と「スライド」の間の係り受け関係が「意味的言い直し」(正解は「連言」)と解釈されて適格な部分が言い直しとして誤認されてしまう場合である。

- (10) 今回の会議の主旨といますか、テーマと
いったようなものについて、少しおうかが
いしたいんですが。
- (11) えーっとオーバーヘッドプロジェクトと
2インチ×2インチのスライドと使えるよ
うになっています。

前者は構造的多義の解消の誤りであり、等位構造などでもよく見られる。現時点では構造に関する選好(格要素はなるべく近い所に係るとか、等位構造では似た構造のものが等位するとか)がないので、今後こういった選好の導入を検討する必要がある。後者は係り受け関係の多義の解消の誤りであり、連体格関係の「の」や任意格の「で」などでもよく見られる。特に、オープン試験では、「対象」格が提題化されて「は」になったものを正しく「対象」と解釈できない例が目立った。これは統計データが少ないために分布に偏りがあったためと思われる。

6 おわりに

本稿では、自然な発話を解析するための一般的な枠組を提示し、言い淀みや言い直しを含む実際の発話に対してその有効性を検討した。その結果、まだ不十分な点が数多くあるものの、現段階としてはまずまずの結果を得た。

今後は、係り受けデータの量を増やすとともに、現状では扱っていない述部欠落を含む文や係り受けがねじれた文などに適応対象を広げてゆきたい。

謝 辞

係り受けデータの作成に協力していただいた神戸大学大学院文化科学研究科の高木一広、金城由美子の両氏に感謝致します。

参考文献

- [1] Weichedel, R.M. and Sondheimer, N.K.: Metarules as a Basis for Processing Ill-Formed Input, *Computational Linguistics*, Vol.9, No.3-4, pp.161-177 (1983).
- [2] Mellish, C.: Some Chart-Based Technique for Parsing Ill-Formed Input, *Proceedings of the 27th Annual Meeting of ACL*, pp.102-109 (1989).
- [3] 佐川, 大西, 杉江: 自己修復を含む日本語不適格文の分析とその計算機による理解手法に関する考察, *情報処理学会論文誌*, Vol.35, No.1, pp.46-52 (1994).
- [4] 村上, 嵯峨山: 自由発話音声認識における音響的および言語的な問題点の検討, *電子情報通信学会技術研究報告*, SP91-100, pp.71-78 (1991).
- [5] 佐川, 大西, 杉江: 大規模コーパスに基づいた日本語自己修復文の分析, *情報処理学会研究報告*, 94-NL-100, pp.73-80 (1994).
- [6] Sumita, E. and Iida, H.: Example-Based Transfer of Japanese Adnominal Particles into English, *IEICE Transactions on Information and Systems*, Vol. E75-D, No.4, pp.585-594 (1992).