

## 誤認識傾向を考慮した音声認識のための構文規則

7E-1

保坂順子

竹沢寿幸

江原暉将

ATR 自動翻訳電話研究所

## 1 はじめに

音声認識に制約として有効に働く構文規則の構築を検討している。日本語は文節の出現順序が比較的自由であり[1]、かつ文節の省略も頻繁に行なわれる。我々が対象としている話し言葉では、特にその傾向が強い。自然言語処理では格に注目して文の規則を構築することがあるが、文節の出現順序や省略を考慮してすべての規則を列挙した場合、それは、これらの制約を与えないのとほとんど等価となる。音声認識には、格とは別の観点からの制約が必要であろう。本稿では、まず音声認識において誤認識の対象となりやすい要素を明らかにする。さらに、その考察に基づき、制約として働きそうな構文情報を取り入れ、実験によりその有効性を確認する。

## 2 音声認識における誤認識傾向

構文情報を予測として音声認識に適用した場合に、誤認識されやすいものの傾向を調べる。我々は、文節単位の発声を入力としている。そこで、文節の音声認識により得た候補を、文の構文情報を使い絞り込むという二段方式を採用した[2]。

文節の音声認識には、構文情報を音素予測に使っているHMM-LR方式[3]を適用した。この実験の設定は、グローバルビーム幅100、一本の枝からのローカルビーム幅10である。137文を構成する353文節について実験したところ、1位で68.3%、5位までの累積で95.5%の認識率だった。

文の構文情報を音声認識に適用する際に、誤認識されやすい要素を調べる実験は、上記の音声認識の結果得られた文節候補5位までを基に行なった。実験のために、許容率[2]の高い文の規則を用意した。この規則は約80%の許容率だった。実験の結果、文の候補になったもので、発声された入力との文字列のマッチングがとれていないものの例を以下に示す。

- 1: 会議に申し込みたいのですが。
- 1a: kaingi-ni moushiko-mi-tai-N-desu-nga
- 1b: kaingi-ni moushiko-mi-tai-no-desu-ka
- 2: こちらは会議事務局です。
- 2a: kata-wa kaingizimukyoku-desu
- 3: どうもありがとうございました。
- 3a: go-o aringat-ou-gozaime-shi-ta
- 3b: go-mo aringat-ou-gozaime-shi-ta
- 3c: mono aringat-ou-gozaime-shi-ta
- 4: 名前は清水太郎です。
- 4a: namae-e-wa shimizu-taroo-desu
- 4b: nana-e-wa shimizu-taroo-desu-shi
- 5: ご住所とお名前をお願いします。
- 5a: gozyuusho-to onamae-o onengai-shi-masu-shi

文候補1aの“N”は、撥音を表す。準体助詞の「の」は「ん」と認識されやすいが、これらは互換性があり、そのニュアンスの違いを構文規則で表すのは難しい。文候補1bは、疑問文と認識されている。文単位の構文規則では、疑問文、平叙文または命令文などを区別することはできない。

2aは、「緑の制服を来ているかたは 会議事務局です。」とすれば容認性が高くなるであろう。また3cは、「高価なものありがとうございます。」とすれば普通の文になる。

3aと3bでは、数字の5が使われている。これらの文は一見非文に見えるが、トランプをしている状況を考えると、普通の文である。5が出れば、自分は上がれるという時、誰かが5を出してくれれば、「5をありがとうございました。」と言うであろう。また、あと3と5が必要だという時、誰かが3を出してくれ、さらに続いて他の誰かが5を出してくれれば「5もありがとうございました。」と使える。

4aと4bは、方向性を表す格助詞「へ」と方向性のない「です」が共起しているため、非文である。「名前へは」と「7へは」は非文節にも見えるが、「住所へは結構ですが、名前へは振りがなをお願いします。」または、「1から6にはコメントを書きましたが、7へはまだ何も意見を書いていません。」のように使われるのは問題ない。

4bや5aのように文末に使われる接続助詞は、一部のものに限られるであろう。

文候補1a-bは、文の構文規則では扱えないが、文候補2-5は構文規則で扱えようである。

## 3 誤認識傾向への対処

前記の文候補2-5において、以下のような文節が文中で容認性が低いのが目立った。

- 形式名詞を含む文節
- 数詞を含む文節
- 格助詞「へ」を含む文節
- 接続助詞終止の文末文節

今回報告する文の構文規則では、この4点に以下のように対処している。

形式名詞は、必ず修飾を受けるものとして扱う。詳細は、4節で述べる。

数詞は、住所、電話番号、日付及び値段としてのみ使用できる。数詞は音素数が少なく特に誤認識されやすいので[4]強い制約を与える必要がある。我々は、国際会議に関する問合せをドメインとして扱っているため、そのドメインでの運用を考慮し、トランプをする状況などは扱わない。

格助詞「へ」を含む文節は、方向性を持たない述語とは共起を許さない。

接続助詞は、文末でも使われるものと、文中でのみ使われるものとを区別した。これは、ATRで収録されている対話データベース[5]を基に調べた実際の運用傾向と、音声認識における誤認識率の相関関係の考察[6]に基づく。

## 4 形式名詞

「こと」や「もの」など、名詞の一種である形式名詞について考察する。形式名詞は、実質的な意味がうすく、連体修飾語を受けてのみ用いられる[7]、と言われる。これらは、確かに、例文6や7のように連体修飾されることが多い。

6: 昨日言ったことは取り消したい。

7: 値段が高いものは必ずしも質がいいわけではない。

一方、例文8や9のように、修飾を受けないこともある。

8: それは、ことですね。

9: ものは確かだ。

そこで、「こと」や「もの」には、二種類あると定義してみる。一つは例文6や7のように使われる形式名詞であり、もう一つは例文8や9のように使われる普通名詞である。しかし、これは、文の解析には役に立つが、制約という観点からは、意味のない区別になる。

「こと」、「もの」さらに「方(ほう、かた)」が、実際にはどの程度連体修飾を受けて使われているかを調べるため、対話データベースを検索した。検索対象は、国際会議に関する問い合わせをドメインとしている対話のうち書き言葉に近いキーボード会話3318文(113会話)と、話し言葉である電話会話7351文(96会話)である。

表1: こと, もの, 方

	キーボード会話		電話会話	
	出現頻度	%	出現頻度	%
連体修飾あり	358	63	774	40
連体修飾なし	214	37	1145	60
文頭	0	0	0	0
合計	572	100	1919	100

表1から、連体修飾を受けることを形式名詞の特徴にすると、書き言葉における被覆率は63%になるが、我々が対象としている話し言葉においては40%にしかならないことが分かる。さらに、連体修飾されないものについて調べたところ、ほとんどのものが、名詞に助詞の「の」が後続したものであることによって修飾されていることが分かった。また、動詞に助詞の「と」が後続し、さらに「の」が後続したものに修飾されるもの、「どんな」や「その」などの連体詞に修飾されるものなどが若干あった。何の修飾も受けないものは、キーボード会話に次の一例だけあった。

恐らく教授にことの次第を伝えることが良いと思います。

対話データベースでは、「こと」、「もの」と「方(ほう、かた)」を含む文節が2491例あったが、そのうち修飾されないものは一例だけだった。形式名詞は、必ず修飾されるという制約を加えても、その被覆率は99%を越える。

## 5 構文規則

形式名詞、数詞、格助詞「へ」、接続助詞終止を考慮して2節で用意した許容率の高い文の構文規則を改良した。これは、改良前の規則の許容率が80%だったのと比較すると、同じ実験対象に対して、約30%の許容率に下がった。文節内規則と、文の構文規則の内訳を表2に示す。

音声認識の誤認識傾向を考慮して改良した構文規則がどの程度制約として働くか調べた。制約を使わず、文節候補に対

表2: 構文規則の規模とパーブレキシティー

	文節内規則	文の規則
規則数	1,973	471
終端数	744	133
パーブレキシティー	3.57 / 音素	99.7 / 文節

してすべての組合せをその尤度の順に並べた場合と、構文規則を制約として使った場合との比較を表3に示す。

表3: 制約としての構文規則

ランク	累積認識率 (%)	
	制約なし	制約あり
1	37.2	70.1
2	60.6	76.6
3	68.6	83.9
4	70.1	83.9
5	73.7	83.9

この構文規則では、以下の6種類を文として認めている。

1. 「はい」「いいえ」などの感動詞を扱うもの。
2. 「それでは、登録用紙をお送り下さい」のように、接続詞が文頭にあり、述部で終るもの。
3. 接続詞が文頭にあり、述部で終るもののうち、述語が方向性を持たないもの。
4. 接続詞がなく、述部で終るもの。
5. 接続詞がなく、述部で終るもののうち、述語が方向性を持たないもの。
6. 例外を含むもの。

表3から、構文規則の制約としての有効性が分かる。特に、第1位ではその効果が顕著である。

## 6 おわりに

音声認識誤り傾向を考慮した、文節に基づく構文規則を構築し、その制約としての有効性を確認した。

現在、構文規則では、接続詞の文中での使用は考慮していない。話し言葉では、「それでは、登録用紙をお送り下さい。」の文節の順番を変えて、「登録用紙を、それでは、お送り下さい。」と言うこともある。また、文は述語文節で終ることを前提としているが、「テーブルの上にあったリンゴを食べたのは誰?」「わたし。」のように名詞文節で文を終らせることもある。さらに構文規則を改良するために、接続詞の割り込みや文末で使われる名詞文節などについて、その実際の使用を対話データベースを基に調べる予定である。

## 参考文献

- [1] 久野 暁 1973: 日本文法研究, 大修館書店
- [2] 竹沢, 保坂, 北, 森元, 江原 1990: “構文規則を用いた音声認識”, 信学技報 (12月講演予定)
- [3] 北, 坂野, 保坂, 川端 1989: “SL-TRANSにおける文節音声認識”, 第39回情報処理学会全国大会
- [4] 保坂, 小暮, 小倉 1990: “音声認識のための連鎖制約としての文法”, 第40回情報処理学会全国大会
- [5] 江原, 小倉, 森元 1990: “電話対話データベースの構築”, 第40回情報処理学会全国大会
- [6] 保坂, 竹沢, 江原 1990: “話し言葉における接続助詞終止とその音声認識への応用”, 第41回情報処理学会全国大会
- [7] 小川, 林 1982, 1988: 日本語教育事典, 大修館書店