

7L-3 話しことばにおける接続助詞終止とその音声認識への応用

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

日本語話しことばの音声認識のための構文規則の構築について考える。これまで、我々は、音声認識を文節単位で行ってきた。本稿では文節単位で発声された音声に文を組み立てるための規則を検討する。話しことばは、書きことばに比べて言い回しが自由である。しかし、この自由度を反映させた構文規則を音声認識に適用すると、正しいものが認識されにくくなる。話しことばの特徴が扱える規則の構築へ向けて、接続助詞終止を例に、これらの相関関係を考察する。

2 語順の転換

日本語は、語順 / 文節の出現位置の転換が自由だと言われる。確かに、「名詞句 + 助詞」あるいは副詞間の出現位置の転換は自由度が高いが [6]、他の文節の出現位置は必ずしも自由ではない。「御住所とお名前をお願いします」の文節音声認識の結果例を基に文節の出現位置の転換可能性を考察する。音声認識には、CFG の枠組みで定義した規則により次の音素を予測するという手法を実現した、HMM-LR [4] を適用した。この実験の設定は、ビーム幅 100、一本の枝からの枝分かれ 10 である。

御住所と	お名前を	お願いします
go-zyuusho-to	onamae-o	onengai-shi-masu
go-zyuu-ka-to	omo-wa-reru-to	onengai-shi-wa
oshie-cha-to	kima-ru-rashii-yo	onengai-shi-taku

助詞の「と」で文節が終わっているものは、文脈により文頭、文中、または、「そんな風に思われると...」のように文末、どこでも出現可能である。また、連用形で終わっているものも、自由度は下がるが、話しことばでは「お願いしたく思います」「そちらにお願いしたく...」などが可能である。一方、第2文節の候補である、「決まるらしいよ」は、文末表現であり、他の文節が後続するのは、話しことばであっても破格である。また、第3文節の候補である「お願いしは」は、「お願いしはしない」「お願いしはしたが」と使われ文末には使われない。ここでは、語順の転換の自由度が低いと思われるこの文末の文節に着目する。

3 話しことばにおける文末の品詞

文末表現の多様性を把握するため、文末に使われている単語の品詞を言語データベースから検索した [1] [2]。その際、話しことばの特徴が明らかになるよう、話しことばが使われている電話会話と、従来の書きことばに近いキーボード会話との比較を行った。いずれも、国際会議に関する問い合わせをドメインとしている。対象は、キーボード会話 4674 文 (171 会話)、電話会話 6910 文 (151 会話) である。キーボード会

表 1: 言語データベース文末品詞

品詞	キーボード会話		電話会話	
	出現頻度	%	出現頻度	%
助動詞	2679	58	2551	48
補助動詞	367	8	288	5
動詞	18	0	28	1
終助詞	1090	23	1237	23
接続助詞	232	5	712	13
係助詞	10	0	12	0
副助詞	7	0	11	0
格助詞	5	0	37	1
並立助詞	3	0	3	0
間投詞	182	4	403	8
副詞	37	1	39	1
接尾辞	10	0	12	0
数詞	6	0	14	0
固有名詞	2	0	12	0
普通名詞	2	0	14	0
形容詞	4	0	2	0
接続詞	3	0	2	0

話では、句点を一文の終わりとし、電話会話では、一人の話者の意図が、あるひとまとまりになったところを文と認定している。

表 1 から、書きことばでも、話しことばでも、助動詞と終助詞の使用率が高い事が分かる。一方、接続助詞については、キーボード会話では、出現率が 5% に留まっているのに対して、電話会話では、13% と高くなっている。このことから、文末に接続助詞を使うのは、話しことばの特徴だと言えるであろう。

話しことばに関する研究、文献 [5] では、接続助詞終止は省略による不完全なものとして、対象外にしているが、本稿では実データの検索結果を重視し、話しことばに多い接続助詞終止を認める。

4 構文規則の文節音声認識結果への適用

文節を文に組立てる規則を約 50 用意した。言語データベース検索結果を考慮し、文末の品詞としては、助動詞、補助動詞、終助詞、接続助詞などを認めている。接続助詞としては、言語データベース形態素作業マニュアル [7] を基に以下の単語を扱っている:

けれども、けれど、けども、けど、が、と、から、し、なら、ので、んで、のに、て、ちゃ、たら、で、じゃ、だら、ば

これらの組立規則を文節音声認識結果に適用した。その結果、文候補になったものの例を以下に示す。音声認識対象は、国際会議に関する問い合わせ 2 会話 37 文である。37 文中、接続助詞終止のものは 3 文であり、いずれにも「が」が使

表 2: 音声認識接続助詞終止

接続助詞	K-モデル		H-モデル	
	出現頻度	%	出現頻度	%
し	43	57	29	42
が	13	17	15	22
ちゃ	10	13	4	6
と	5	7	3	4
から	3	4	8	12
て	1	1	2	3
なら	0	0	5	7
たら	0	0	3	4

われていた。現在、文節音声認識には、2つの異なる HMM (Hidden Markov Model) 音韻モデル、K-モデルと H-モデルが用意されているが、ここでは K-モデルを使っている。

1: もしもし
1a: moshimoshi
1b: oshie-ru-shi
1c: mo-chi-mashi-te

2: そちらは会議事務局ですか
2a: sochira-wa kaingizimukyoku-desu-ka
2b: sochira-wa kaingizimukyoku-desu-kara
2c: kochira-wa kaingizimukyoku-desu-nga

3: それでは登録用紙をお送り致します
3a: soredewa tourokuyoushi-o ookuri-ita-shi-masu
3b: soredewa tourokuyoushi-mo ookuri-ita-shi-masu-shi
3c: soredewa tourokuyoushi-o ookuri-ita-shi-masu-to

接続助詞終止の文の中で、1b, 1c, 3b, 3c は、容認性が低い文である。

5 音声認識における接続助詞終止

ここでは、文脈、意味などを使わず、カテゴリーの細分化により文の容認性を高めることを試みる。そこで、音声認識において誤認識の要因となり易い単語を別に扱うことを考える。文節音声認識の結果第 5 位までに入った文末文節候補 185 を対象に、接続助詞終止を調べた。ここでは、HMM を使った認識傾向を調べるため、K-モデルと H-モデルを使った結果を比較する。K-モデルを使った場合 75 文節 (41%)、H-モデルを使った場合 69 文節 (37%) で、文末に接続助詞が使われていた。その内訳を表 2 に示す。

両モデルとも、「し、が、ちゃ」などの出現率が高い。これらの接続助詞は、音節数の少ない単語なので誤認識の要因になり易く [3] 出現率が高くなっていると思われる。特に「し」では、母音の無声化が作用していると思われる。一方、音節数の多い「けれども」「けれど」などは、誤認識の対象にはなっていない。

6 言語データベースにおける接続助詞終止

キーボード会話、電話会話で使われている接続助詞の内訳を表 3 に示す。

キーボード会話では、「が」の使用が 85% と顕著である。書きことばだけを対象とする場合には、接続助詞終止には、「が」を認めるだけでよいであろう。しかし、話しことばでは、「が」と並んで「けれども」とそれに準じる単語の出現率も高くなっている。また、「ので」の使用率も、書きことばと比較してかなり高い。

前述の文献 [5] では基本的に接続助詞終止を対象外にしているが、いくつかの接続助詞終止は、例外的に認定することを明記している:

表 3: データベース接続助詞終止

接続助詞	キーボード会話		電話会話	
	出現頻度	%	出現頻度	%
が	197	85	274	38
ので	11	5	96	13
て	8	3	23	3
で	0	0	1	0
から	6	3	14	2
けれども	5	2	212	30
けれど	1	0	18	3
けど	1	0	12	2
けども	0	0	37	5
し	2	1	10	1
もので	1	0	0	0
たら	0	0	5	1
と	0	0	2	0
ば	0	0	2	0
んで	0	0	5	1
ながら	0	0	1	0

- それだけで理由を表すもの: ... だし, ... だから
- 曖昧、婉曲な終止を表すもの: ... だけど, ... したりして, ... しなきゃ

しかし、表 3 から明かなように「し、から、けど、て」などの出現率はあまり高くない。従って、これらの接続助詞終止を許さない組立規則でも覆い率はそれほど低くならない。特に「し」は前章の表 2 に示すとおり音声認識において誤認識の要因となり易いので検討を要する。

7 おわりに

話しことばの文末に使われる接続助詞について検討した。音声認識への適用を目的とした場合、接続助詞終止に使われる単語と、文中でのみ使われる単語との分離が有効であることが分かった。

国際会議に関する問い合わせをドメインとする言語データベース検索結果は、これまでの話しことば研究とのくいちがいを示唆している。話しことばの特徴を明示するにはさらに他のドメインについても調査する必要がある。

謝辞

本研究の機会を与えて下さると共に適切な助言を頂いた ATR 自動翻訳電話研究所榎松明社長、同データ処理研究室森元暹室長に感謝します。

参考文献

- [1] 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
- [2] 橋本、小倉、江原、森元 1990: 「対話データベースを用いた各種言語現象の検索」, 第 40 回情報処理学会全国大会
- [3] 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
- [4] 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
- [5] 国立国語研究所 1960: 話しことばの文型 (1), 秀英出版
- [6] 久野 晴 1973: 日本文法研究, 大修館書店
- [7] 篠崎、水野、小倉、吉本 1989: 形態素情報利用解説書, ATR テクニカルレポート TR-I-0077