

ポーズ情報を導入した対話文の形態素解析

金子拓也 石崎俊

慶應義塾大学 大学院 政策・メディア研究科
〒252-0816 神奈川県藤沢市遠藤 5322
takuya@sfc.keio.ac.jp ishizaki@sfc.keio.ac.jp

自然で自発的な対話を対象とする対話文理解システムの構築を目指して、その要素技術である対話文の形態素解析の研究を進めている。本稿では、休止単位ごとに形態素解析を行う手法について検討する。その結果、休止単位に基づいて対話テキストを形態素解析する場合には、1. 話者別の接続規則学習が再現率を向上させる。2. 学習に用いるデータが小規模の場合はポーズ情報を学習することで再現率が向上する。3. 話者の対話データをある程度の規模で学習すれば、その話者の新規発話を解析する際に、ポーズ情報を追加学習しなくてもよい。4. 継続時間を用いたポーズ情報の分類は僅かに再現率を向上させるがそのための計算コストは高いことなどが分かった。

Morphological Analysis of Transcripts of Spoken Dialogue Based on Pause Unit

Takuya KANEKO and Shun ISHIZAKI

Graduate School of Media and Governance, Keio University
5322 Endo, Fujisawa, Kanagawa, 252-0816 JAPAN
takuya@sfc.keio.ac.jp ishizaki@sfc.keio.ac.jp

We proposed a morphological analysis method for a dialogue text understanding system which progressively understands utterances allowing for unexpected ones. In this paper, we studied analysis method based on pause unit in several conditions. As a result, we found: the costs that is assigned to each bi-gram of morphemes are different among speakers; pauses are effective in increasing recall of analysis in some cases; if the system has learned enough pausal data, additional learning of pause is unnecessary for it to analyse new utterances; it is not much effective in increasing recall to classify pauses by their duration.

1 はじめに

自然で自発的な対話を対象とする頑健な対話文理解システムの構築を目指している。対話文には言い淀み、言い誤り、フィラーなどの非文要素が頻出し、また、助詞の省略や倒置など統語的に扱いにくい現象が生じる。さらに、対話には書き言葉のように明確な句読法が存在しないため、解析単位の決定が問題となる。解析単位には文などが考えられるが、対話文ではそのような単位は自明ではなく、聞き手が推測しているに過ぎない単位と考えられる。このことは、人間に対話テキストを文単位でセグメンテーションさせると、何らかのマニュアルに基づいていたとしても、作業者や作業日によって揺れが生じる

のが常識であることから明らかである。つまり、文単位を決定するのは言語理解においては高次の判断であるから、低次処理の解析単位にはなり得ない。しかし、ポーズを区切り文字と看做すことにより、人手にしろ機械的にしろ言語理解の低次の処理で解析単位を決定することが可能となる。また、音声情報処理の分野では音声認識の認識単位としてポーズを利用したものが [5][6] ので、ポーズを低次の解析単位とする対話文理解システムの意味解析部は、音声対話理解システムの言語理解部にも親和性があると考えられる。そのような背景と目標に基づいて、本研究はポーズに挟まれた比較的短い区間ごとに形態素解析を行う手法について検討する。

表 1: 対話データ
基本統計

話者数	7名
延べ形態素数	2,882
異なり形態素数	502
語断片数	49
ポーズ出現頻度	901

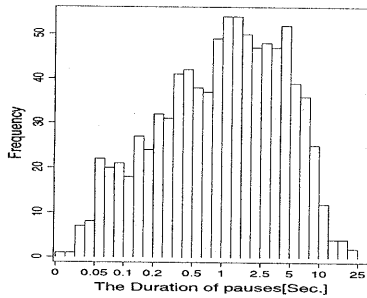


図 1: 自己発話内のポーズ出現頻度分布
(横軸は対数目盛り)

表 3: 品詞とポーズの接続確率

ポーズの前に出現する品詞	確率 [%]	ポーズの後に出現する品詞	確率 [%]
感動詞	88.0	名詞-一般	22.6
名詞-接尾-特殊	76.4	感動詞	15.5
語断片	71.4	動詞-自立	7.3
動詞-非自立	66.6	形容詞-自立	5.7
助詞-終助詞	64.4	接続詞	4.6
フィルター	60.1	副詞-一般	4.5
名詞-代名詞-一般	50.0	副詞-助詞類接続	3.3
副詞-助詞類接続	42.4	語断片	3.3
接続詞	33.8	名詞-代名詞-一般	2.5
名詞-接尾-助数詞	33.3	助動詞	2.2

表 2: 対話データ例

音声時刻	音声開始終了 フラグ	音声表記	発音表記出現形	読み表記出現形	見出し語表記出現形	見出し語表記基本形	品詞	活用型	活用形
4.089	S	*init*							
4.954	E	e-Qto	エト	エト	えと	えと	フィルター		
6.894	E	kono'	コノ	コノ	この	この	連体詞		
7.079	E	hito	ヒト	ヒト	人	人	名詞-一般		
7.469	E	wa-	ワ	ハ	は	は	助詞-係助詞		
8.176	S	0.707							
8.619	E	kuro'me	クロメ	クロメ	黒目	黒目	名詞-一般		
8.833	S	qa	ガ	ガ	が	が	助詞-格助詞-一般		
9.398	E	hikakuteki	ヒカクテキ	ヒカクテキ	比較的	比較的	副詞-一般		
9.750	S	ni-	ニ	ニ	に	に	助詞-副詞化		
10.053	E	0.303							
10.352	S	o'-ku	オーク	オオク	多く	多い	形容詞-自立	形容詞・アウオ段	連用テ接続
10.771	E	te-	テ	テ	て	て	助詞-接続助詞		
11.571	S	0.800							

2 対話データ

(社)電子情報技術産業協会(旧称:(社)日本電子工業振興協会)対話理解技術専門委員会が公開するマルチモーダル対話コーパス [1][2][3][4] の顔課題のうち FM1,FM2,FM3,FS1 を対象として、音声並びに形態素情報のタギングを行い、これを対話データとして使用した。表 1 は対話データの基本的な統計情報である。話者数は同一人物を 1 名と数える。本対話データは FNO,FMA,FTS,FKO,MNN,MYS,MKS という話者識別子を持つ 7 名のデータから構成されている。なお、話者識別子の先頭の F は女性を、M は男性を表している。次に、図 2 に対話データ例を示す。第 1 項は音声時刻、第 2 項は単語の開始(S)/終了(E)フラグ、第 3 項は音声表記若しくはポーズ継続時間、第 4 項は発音表記の出現形、第 5 項は読み表記の出現形、第 6 項は見出し語表記の出現形、第 7 項は見出し語表記の基本形、第 8 項は品詞名、第 9 項は活用型、第 10 項は活用形を表している。最後に本データの特徴記述として自己発話内のポーズ出現頻度分布を図 1 に、品詞とポーズの接続傾向を表 3 にまとめる。ここで自己発話内のポーズとは同一話者の発話内で計測したポーズのことであり、つ

まり対話相手の発話とは無関係に計測したポーズであることを意味する。品詞とポーズの接続傾向をさらに詳細に調べれば、格助詞の後に 2 秒未満のポーズが接続する確率は 20%、格助詞の後に 2 秒以上のポーズが接続する確率は 5%、2 秒未満のポーズの後に格助詞が接続する確率は 1%、2 秒以上のポーズの後に格助詞が接続する確率は 0%等、ポーズ継続時間に関係する傾向も観察することができる。

3 手法

3.1 解析テキストと解析条件

解析用テキストは対話データに基づいて機械的に作成する。この際ポーズ情報を形態素に変換する。ポーズ継続時間を基準としてポーズ情報を種類の異なる複数の形態素に変換する場合には、任意のポーズ継続時間を閾値として指定する。ポーズを 2 分類する場合には閾値 (T1) を、ポーズを 3 分類する場合には閾値のペア (T1,T2) を指定する。図 2 の見出し語表記テキストでは閾値が指定されていないのでポーズ情報を分類しない。従って□は全てのポーズを表している。図 2 の読み表記テキストでは閾値

見出し語表記テキスト
□えとこの人は□ □黒目が比較的□ □多くて□ □で髪のが□

読み表記テキスト [T1=0.316]
□エトコノヒトハ□ □クロメガヒカクテキニ□ ○オオクテ□ □デカミノケガ□

発音表記テキスト [T1=0.316, T2=0.794]
□エトコノヒトワ△ △クロメガヒカクテキニ○ ○オオクテ□ □デカミノケガ△

図 2: 解析テキスト例
(T1, T2 はポーズ分類の閾値)
(○ > T1 ≥ □)
(○ > T1 ≥ △ > T2 ≥ □)

表 4: 解析条件

解析条件	閾数	学習セット	テストセット
1	0	本人以外の全話者のデータ	本人のデータ 100%
2	0	全話者のデータ 100%	本人のデータ 100%
3	0	本人のデータ 100%	本人のデータ 100%
4	1	本人のデータ 100%	本人のデータ 100%
5	0	本人のデータ初めの 80% (但しテストセットに含まれるポーズ以外の接続規則を学習)	本人のデータ終りの 20%
6	0	本人のデータ 100% (条件 5 のテストセットに含まれるポーズ情報も学習)	本人のデータ終りの 20%
7	1	本人のデータ初めの 80% (但しテストセットに含まれるポーズ以外の接続規則を学習)	本人のデータ終りの 20%
8	1	本人のデータ 100% (条件 7 のテストセットに含まれるポーズ情報も学習)	本人のデータ終りの 20%

T1 = 0.316 が指定されているので、○が 0.316 秒未満のポーズ、□が 0.316 秒以上のポーズを表す。図 2 の発音表記テキストでは閾値のペア T1 = 0.316, T2 = 0.794 が指定されているので○が 0.316 秒未満のポーズ、△が 0.316 秒以上 0.798 秒未満のポーズ、□が 0.798 秒以上のポーズを表す。ここで解析用テキストの各行 (2 つのポーズに挟まれた区間) を休止単位と呼ぶことにする。また、閾値を指定しなかった場合を閾数 0 の場合、閾値をひとつ指定した場合を閾数 1 の場合、閾値をふたつ指定した場合を閾数 2 の場合と呼ぶ。本研究では閾数、学習セット、テストセットの組合せで 8 つの解析条件を設計した。各解析条件は表 4 の通りである。

3.2 接続規則の学習

3.1 項で設計した解析条件ごとに対話データから品詞接続のコストを計算して接続規則を作成する¹。コストは式 1 により計算した。ここで、 $N(s_i)$ は品詞 s_i が出現する回数、 $N(s_i, s_{i+1})$ は品詞 s_i の直後に品詞 s_{i+1} が出現する回数、 $C(s_i, s_{i+1})$ が品詞 s_i の直後に品詞 s_{i+1} が出現するコストである。なお、通常は式 1 の対数をとってコストとするが、本研究では対話データが小規模であり、予備実験で再現率が低下する傾向が観察されたため、対数を取らない。

$$C(s_i, s_{i+1}) = \frac{N(s_i)}{N(s_i, s_{i+1})} \quad (1)$$

3.3 再現率と適合率

再現率とは解析対象発話全体の正しい解析による形態素と解析システムを用いた場合の形態素が一

致する回数を、正しい解析による形態素の数で割った値である。適合率とは正しい解析による形態素と解析システムを用いた場合の形態素が一致する回数を、解析システムを用いた場合の形態素の数で割った値である。言い換えれば式 2、式 3 のようになる。

$n(M_c)$ = 正しい解析による形態素の数

$n(M_s)$ = 解析システムを用いた場合の形態素の数

$n(M_c \cap M_s)$ = 解析結果における正しい形態素の数

$$\text{再現率} = \frac{n(M_c \cap M_s)}{n(M_c)} \quad (2)$$

$$\text{適合率} = \frac{n(M_c \cap M_s)}{n(M_s)} \quad (3)$$

本研究では解析結果における正しい形態素の数の計算に際して、形態素の形態素構造を 4 項組 (見出し語基本形、品詞、活用型、活用形) で定義する²。休止単位ごとに形態素解析を行い、そのつど解析結果の第 1 候補に存在する正しい形態素の数を加算していくことにより、再現率と適合率を計算する。実際には再現率を式 4 により、適合率を式 5 により計算している。

$n(PU)$ = 休止単位の数 (PU : *Pause Unit*)

$n_i(M_c \cap M_s)$ = i 番目の休止単位の解析結果が含む正しい形態素の数

$$\text{再現率} = \frac{\sum_{i=1}^{n(PU)} n_i(M_c \cap M_s)}{n(M_c)} \quad (4)$$

$$\text{適合率} = \frac{\sum_{i=1}^{n(PU)} n_i(M_c \cap M_s)}{n(M_s)} \quad (5)$$

これ以降の図表では再現率 (Recall) を R、適合率 (Precision) を P と記述する場合がある。休止単位の解析結果が含む正しい形態素の数は図 3 のようにカウントする。カウントに際して順序関係は考慮し

¹ 辞書は茶筌 2.02 版に標準のものをほぼそのまま使用している。つまり辞書は学習させない。

² これ以降特に断りが無い限り、形態素と形態素構造を区別しない。

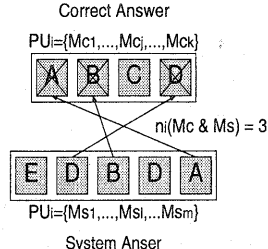


図3:休止単位の解析結果が含む正しい形態素の数の数え方、 $n_i(M_c \cap M_s) = 3$ の場合

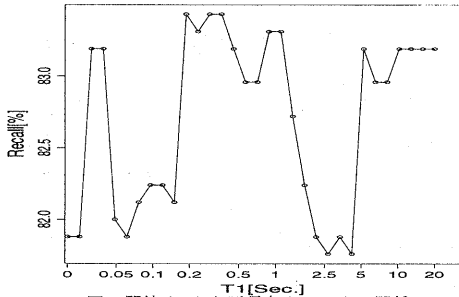


図4:閾値 (T1) と再現率 (Recall) の関係 (話者 FKO の発音表記テキストを解析した場合)

ていない。なぜなら順序が異なっても休止単位の内に正しい形態素が存在するほうが、形態素タグの人手修正が容易になるからである。しかし、生理的な理由から休止単位の長さは短い場合が多いので、図3のように順序が交差するような現象を実際に観察したことはない。また、語断片に関する情報は学習課程で接続規則に含まれるが、辞書には含まない。つまり、語断片に関しては意図的に解析誤りが生じる仕組みになっている。語断片はセットを有限に確定できないので辞書に含むのが妥当でないと判断したからである。

3.4 閾値の設定

図4はポーズ継続時間を用いてポーズ情報を2分類したときの閾値と再現率の関係である。この事例では閾値が0.2秒、0.316秒、0.398秒のときに再現率が最高の83.43%になることが分かる。また、閾値の設定次第ではポーズ情報を分類しなかった場合よりも再現率が下回る場合があることも分かる。例えば閾値が0秒の場合には再現率が81.88%であるのに対して、閾値が3.162秒、5.102秒の場合には再現率が81.76%と、僅かだが下回っている。本研究では解析条件(表4)の閾数が1の場合、各学習セットの再現率を最大にするポーズ継続時間を閾値として設定した。

4 結果と考察

解析には奈良先端科学技術大学院大学松本研究室が公開する茶笑2.02版[7]を使用した。解析結果を表5に示す。また、図5に解析条件1,2,3,4の再現率平均値を、図6に解析条件5,6,7,8の再現率平均値を、それぞれ示す。

4.1 話者ごとの接続規則の違い

図5で解析条件1,2を解析条件3,4と比較すると、全表記(見出し語表記、読み表記、発音表記)において後者の再現率が高いことが分かる。例えば解析条件3は解析条件2よりも漢字表記で2.29、読み表記で4.27、発音表記で5.67ほど再現率が高い。接続規則以外の解析条件が全て同じなので、話者別学習が再現率を向上させるという事実は、話者ごとに形態素の接続規則が大きく異なる場合があることを意味する。そこで話者間の品詞接続確率の差を調べたのが表6であり、差を絶対値で表している。実際、話者間で品詞接続確率に大きな差がある場合が多いことが分かる。特にある品詞に対するポーズの後接しやすさには、話者によって大きな差がある場合が多いことが分かる。従って、話者別に接続規則を学習させることにより再現率が向上する。

4.2 ポーズ情報導入の効果

図6はポーズ情報の学習効果を比較している。解析条件5を6と、若しくは解析条件7を8と比較すると、全表記において後者の再現率が高いことが分かる。解析条件6は解析条件5よりも漢字表記で2.84、読み表記で14.18、発音表記で14.35ほど再現率が高い。また、解析条件8は解析条件7よりも漢字表記で5.41、読み表記で19.12、発音表記で16.95ほど再現率が高い。解析条件5と7はテストセットに含まれるポーズ以外の接続規則を学習した場合であり、解析条件6と8は解析条件5と7のテストセットが含むポーズ情報も学習した場合である。両者の違いはポーズ情報を学習したか否かだけなので、ポーズ情報を学習することにより再現率が向上するといえる。ただしこの言明は本人の新規の発話を解析する際に、その発話のポーズ情報を追加学習する必要があるということと同じではない。なぜならば再現率が向上するのは当該話者の発話デー

表 5: 解析結果の再現率 (R) と適合率 (P)(見出し語表記のみ抜粋)

解析条件 1: 閾数 0、学習セットは本人以外の全話者のデータ、テストセットは本人のデータ 100%									
		FNO	FMA	FTS	FKO	MNN	MYS	MKS	平均值
見出し語	R	91.26	97.06	84.98	85.46	80.62	89.50	83.51	87.48
	M	P	88.68	97.06	82.51	83.86	78.64	87.75	80.20
解析条件 2: 閾数 0、全話者のデータを 100%学習、本人のデータを 100%解析									
		FNO	FMA	FTS	FKO	MNN	MYS	MKS	平均值
見出し語	R	93.20	97.06	93.66	93.68	91.67	93.50	93.81	93.80
	M	P	90.57	97.06	92.02	93.02	88.91	93.03	92.23
解析条件 3: 閾数 0、本人のデータを 100%学習、本人のデータを 100%解析									
		FNO	FMA	FTS	FKO	MNN	MYS	MKS	平均值
見出し語	R	99.03	97.06	93.88	94.52	94.57	96.00	97.59	96.09
	M	P	99.03	97.06	91.84	94.18	92.42	96.00	97.26
解析条件 4: 閾数 1、本人のデータを 100%学習、本人のデータを 100%解析									
		FNO	FMA	FTS	FKO	MNN	MYS	MKS	平均值
見出し語	R	99.03	97.06	94.77	96.07	95.54	96.50	98.28	96.75
	M	P	99.03	97.06	93.42	95.72	93.19	96.50	98.28
解析条件 5: 閾数 0、学習セットは本人のデータ初めから 80%、テストセットは本人のデータ終りの 20% (但しテストセットに含まれるボーズ以外の品詞接続確率を学習した場合)									
		FNO	FMA	FTS	FKO	MNN	MYS	MKS	平均值
見出し語	R	100.00	83.33	92.22	95.48	89.32	90.48	95.00	92.50
	M	P	100.00	83.33	89.25	96.02	90.20	90.48	96.61
解析条件 6: 閾数 0、本人のデータを 100%学習、本人のデータの終り 20%を解析 (解析条件 5 のテストセットに含まれるボーズ情報も学習した場合)									
		FNO	FMA	FTS	FKO	MNN	MYS	MKS	平均值
見出し語	R	100.00	100.00	92.78	97.18	91.26	92.86	95.00	95.34
	M	P	100.00	100.00	89.78	98.29	93.07	92.86	96.61
解析条件 7: 閾数 1、学習セットは本人のデータ初めから 80%、テストセットは本人のデータ終りの 20% (但しテストセットに含まれるボーズ以外の品詞接続確率を学習した場合)									
		FNO	FMA	FTS	FKO	MNN	MYS	MKS	平均值
見出し語	R	90.91	83.33	91.67	95.48	88.35	90.48	95.00	90.75
	M	P	90.91	83.33	88.71	96.02	89.22	90.48	96.61
解析条件 8: 閾数 1、本人のデータを 100%学習、本人のデータの終り 20%を解析 (解析条件 7 のテストセットに含まれるボーズ情報も学習した場合)									
		FNO	FMA	FTS	FKO	MNN	MYS	MKS	平均值
見出し語	R	100.00	100.00	93.89	97.74	91.26	95.24	95.00	96.16
	M	P	100.00	100.00	91.85	98.30	93.07	95.24	96.61

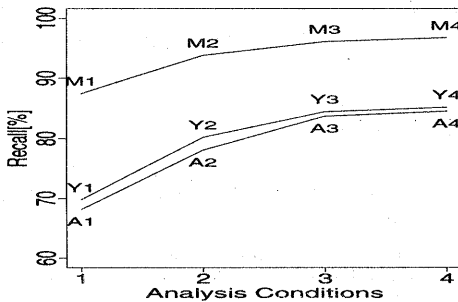


図 5: 解析条件 1,2,3,4 の再現率平均値 (M:見出し語表記,Y:読み表記,A:発音表記)

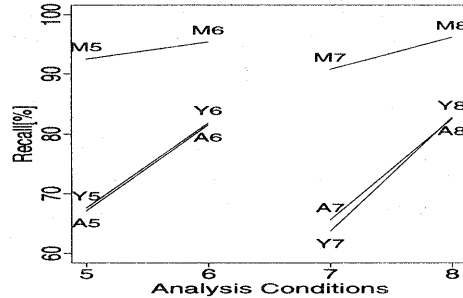


図 6: 解析条件 5,6,7,8 の再現率平均値 (M:見出し語表記,Y:読み表記,A:発音表記)

表 6: 話者間の品詞接続確率の差 (FTS は女性説明者,FKO は女性回答者,MNN は 男性説明者)

品詞 S_i	品詞 S_{i+1}	FTS [%]	FKO [%]	差	品詞 S_i	品詞 S_{i+1}	FTS [%]	MNN [%]	差
接続詞	名詞-一般	9.1	81.8	72.7	感動詞	ボーズ	1.4	80.6	79.2
助詞-副助詞	ボーズ	1.7	66.7	65.0	助詞-終助詞	ボーズ	4.2	64.0	59.8
名詞-数	名詞-接尾-助数詞	25.0	80.0	55.0	名詞-形容動詞語幹	助動詞	2.7	50.0	47.3
名詞-接尾-一般	ボーズ	33.3	80.0	46.7	助詞-副/並立/終助詞	助詞-終助詞	33.3	76.5	43.2
名詞-形容動詞語幹	助動詞	2.7	36.4	33.7	名詞-非自立-一般	ボーズ	5.3	40.0	34.7
語断片	ボーズ	86.7	50.0	36.7	助詞-格助詞-連語	ボーズ	2.9	33.3	30.4
助詞-副/並立/終助詞	ボーズ	42.9	10.0	32.9	連体詞	名詞-非自立-一般	22.2	50.0	27.8
名詞-非自立-副詞可能	助詞-格助詞-一般	64.3	33.3	31.0	助詞-格助詞-一般	ボーズ	31.4	7.1	24.3
名詞-非自立-助動詞語幹	助動詞	64.3	33.3	31.0	助動詞	ボーズ	6.7	30.8	24.1
副詞-助詞類接続	ボーズ	3.7	33.3	29.6	助詞-副助詞	ボーズ	1.7	25.0	23.3

タが少ない場合であり、例えば解析条件5と6の読み表記と発音表記では話者FNAが33.33から100.00へ、話者FNOが77.27から90.91へ、話者MYSが71.43から80.95へと向上しているが、当該話者の発話データがある程度の規模を持つ場合(話者FTS,FKO)には小幅にしか向上しないからである。この事実から、当該話者の発話データがある程度の規模で学習しておけば、新規の発話を解析する際に、その発話のポーズ情報を追加学習するしなくてもよいということ推測できる。

4.3 ポーズ情報を分類する意義

図5で解析条件3と4の再現率平均値を比較すると、ポーズ分類学習が全ての表記法において再現率の向上に寄与していることが分かる。その限りでは全ての表記法においてポーズを分類することに意義があるといえる。しかし、全ての表記法において再現率は1未満の向上しかしておらず、再現率の向上度合いが小さいという意味で、ポーズを分類する意義の程度は小さいといえる。また、各話者・各表記法ごとの最適な閾値を決定するためには高い計算コスト(1人・1表記につき閾数1の場合で30分程度、閾数2の場合で5時間程度)を要するので、ポーズを分類することにより得られるメリットがそのコストと比較して充分に見合っているとはいえない。また、幾つまで閾数を増やすことに意味があるかということは今のところ不明である。閾数2で本人を含む全話者の発話、若しくは本人のみの発話を100%学習させた場合には、閾数1の場合と比べて全ての表記法において再現率が若干向上することが分かっているが、学習量を減らした場合については調査していない。前述の通り閾値を設定するために要するコストは無視できないのであるから、閾数は0に設定するのが現実的であろう。

4.4 表記法と再現率の関係

4.1項、4.2項、4.3項で述べた通り、話者別学習／ポーズ情報の導入／ポーズ情報の分類はそれぞれ全ての表記法において再現率の向上に寄与するのであるが、各表記法間でその向上幅を比較すると、話者別学習(解析条件3を解析条件2と比較)により漢字表記で2.29、読み表記で4.27、発音表記で5.67ほど再現率が向上する。ポーズ情報の導入(解析条

件6を解析条件5と比較)により漢字表記で2.84、読み表記で14.18、発音表記で14.35ほど再現率が向上する。ポーズ情報の分類(解析条件4を3と比較)により漢字表記で0.66、読み表記で0.75、発音表記で0.83ほど再現率が向上するなど、再現率を向上させるこれらの手法が発音表記 > 読み表記 > 見出し語表記の順で効果があることが分かる。

5 おわりに

自然で自発的な対話を対象とする対話文理解システムの要素技術として、休止単位に基づく対話文の形態素解析手法について検討した。その結果、休止単位に基づいて対話テキストを形態素解析する場合には、1. 話者別に接続規則を学習させることにより再現率が向上する。2. 学習セットが小規模の場合にはポーズ情報を学習することにより再現率が向上する。3. 当該話者の発話データがある程度の規模で学習しておけば、当該ドメインにおける本人の新規発話を解析する際に、その発話のポーズ情報を追加学習しなくてもよい。4. 継続時間によるポーズの分類は僅かに再現率を向上させるが、閾値を設定する際のコストが高いので分類しないのが現実的。5. これらの手法が発音表記 > 読み表記 > 見出し語表記の順で影響することが分かった。

謝辞

音声情報処理分野におけるポーズ関連の先行研究を紹介して頂いた、中川聖一氏、竹澤寿幸氏に感謝いたします。

参考文献

- [1] 金子拓也, 石崎俊: マルチモーダル対話コーパスのタグ付きデータ部分公開に向けて, 信学技法, TL2000-9(00-7).
- [2] Takuya Kaneko and Shun Ishizaki. 1999. The Multi-Modal Dialogue Corpus. In Proc. of the 2nd ICCS, pp.1010-1013.
- [3] 金子拓也, 石崎俊: マルチモーダル対話コーパスの構築, ~マルチモーダルデータのタギングについて~, 信学技法, TL99-3(99-5).
- [4] 金子拓也, 石崎俊: マルチモーダル対話コーパスの構築について, 信学技法, NLC97-64, PRMU97-266(98-03).
- [5] 竹澤寿幸, 森元運, 部分木に基づく構文規則と前終端記号バイグラムを併用する対話音声認識手法, 信学論D-II, Vol. J79-D-II, No. 12, pp. 2078-2085 (96-12).
- [6] 伊藤克彦, 秋葉友良, 上條俊一, 田中和世: 休止を区切りとした対話処理, 情処研報, 95-SLP-7-21, 1995.
- [7] <http://cl.aist-nara.ac.jp/lab/nlt/chasen.html>