

## 話し言葉の整形作業における削除箇所の自動同定

尾嶋 憲治<sup>†</sup> 河原 達也<sup>†</sup> 秋田 祐哉<sup>†</sup> 内元 清貴<sup>‡</sup>

<sup>†</sup>京都大学 情報学研究科

〒 606-8501 京都市左京区吉田本町

<sup>‡</sup>情報通信研究機構

〒 619-0289 京都府相楽郡精華町光台 3-5

あらまし 『日本語話し言葉コーパス』(CSJ)のコアに対して、講演録の作成を想定して、削除・置換・挿入などの一次整形のアノテーションを付与した。本稿では、この話し言葉の整形作業において削除される箇所を自動推定する手法を提案する。本研究では、削除箇所について IOB 体系でラベリングを行い、このラベルについてのテキストチャンキングの問題とみなし、SVM に基づくチャンカである YamCha を利用する。素性として、同一表現の繰り返しや文節・係り受けといった言語的情報、基本周波数などの韻律的情報などを検討し、また言い直しなどの自己修復部との関係についても調べた。書き起こしと音声認識結果の両方に対して評価実験を行い、書き起こしについては F 値 0.8 が得られた。

## Automatic Detection of Portions to be Deleted during Cleaning Transcripts

Kenji Ojima<sup>†</sup> Tatsuya Kawahara<sup>†</sup> Yuya Akita<sup>†</sup>  
Kiyotaka Uchimoto<sup>‡</sup>

<sup>†</sup>School of Informatics, Kyoto University

Sakyo-ku, Kyoto 606-8501, Japan

<sup>‡</sup>National Institute of Information and Communications Technology

Hikaridai 3-5, Seika-cho, Soraku-gun, Kyoto 619-0289, Japan

**Abstract** We introduce a corpus annotated with editing operations (deletion, insertion, substitution) on faithful speech transcripts for generating document-style texts, and propose a method to automatically detect portions to be deleted during this process. The proposed method is based on an SVM chunker. Features for the SVM include linguistic information such as existence of repeated utterances and syntactic relations as well as prosodic information. We achieved reasonable results not only for the manual transcripts but also for ASR results.

## 1 はじめに

音声認識技術の進展にともない、会議や講演などを対象とした筆記録の作成などの研究が進められている [1]。しかし、このような話し言葉音声の書き起こしは、書き言葉と異なる点が多く、可読性がよいとはいえない。そのため、会議録や講演録の作成を念頭に置いた場合、音声の書き起こしをそのまま利用するのではなく、文の区切りで分割する、フィラーや冗長部分を取り除く、話し言葉特有の表現を適切に置き換える、話中で省略された助詞などの語を補う、などの整形処理を行う必要がある。しかしながら、現在このような整形処理は主に人手で行われており、大きなコストがかかっている。ここで自動的な処理が可能となれば、効率的な筆記録の作成に役立つと考えられる。

これまでに、整形を受ける言語構造の解析や機械的な処理の研究が進められている。話し言葉における繰り返しや言い直しなどの自己修復部の分析・検出については、RIM (Repair Interval Model)[2] が代表的なモデルとして知られており、自己修復部を被修飾部 (RPD)、言いよどみ (DF)、修復部 (RP) の連続した 3 つの区間に分割されると仮定して、その検出を行っている。この分析・検出は、英語だけでなく、日本語を対象としても行われている [3, 5]。また、自己修復部だけでなく、フィラーや言いよどみの検出 [6] も行われている。しかし、これまでの研究では、話し言葉の整形処理との関係は明らかではなかった。統計的機械翻訳の手法を応用して講演の書き起こしから整形を試みる研究 [7] も行われているが、コーパスサイズも小さく、処理の対象を、句点の挿入、助詞の挿入、文体の統一に限定している。

本稿では、整形過程における削除箇所と自己修復部、フィラー、言いよどみ、定型表現との関係进行分析し、講演録の編集において削除される箇所を自動的に推定する手法を提案する。本手法では、SVM (Support Vector Machines) を用いたテキストチャンキングにより削除箇所を自動推定する。そのための素性として種々の特徴を検討するとともに、自己修復部との関係や音声認識結果への頑健性についても調べる。

なお、本手法の分析および実験には『日本語話し言葉コーパス』(CSJ) [8] を利用する。CSJ は、学会講演や模擬講演などのモノログを主な対象として収集・構築されたコーパスである。CSJ に収録され

表 1: 整形要素の頻度

削除	19621	(57.3%)
置換	11640	(34.0%)
挿入	2946	(8.6%)

ている講演のうち、コアと呼ばれる一部の講演に対しては、書き起こしテキストのほかに形態素・係り受け・節単位などの言語的情報や韻律的情報などが付与されている。今回、このコアに対して、講演録作成を想定して一次整形の作業を行い、その編集過程をアノテーションした。

## 2 話し言葉に対する整形処理

会議や講演の音声の書き起こしから会議録を作成する過程においては、主に以下のような順序で整形作業（「整文」とも呼ばれる）が行われる [4]。

1. 無機能語、言いさしの削除（ケバ取り）
2. 助詞の誤用・脱落などの修正・補完、適切な改行、引用符の付与
3. 主語と述語の不一致の修正、倒置の修正、意味不明の場所の補正
4. 話題に直接関係ない部分の削除

後段の処理ほど、意味的・談話的な理解を要することがわかる。逆に、最初のステップにおいては不要・冗長部分の削除のみを行っている。

一次整形（ステップ 1 とステップ 2 の一部）として行われる処理としては、以下の 3 つが挙げられる。

- 削除：言いよどみや重複箇所など、冗長な部分を除く。
- 置換：語の誤りや話し言葉特有の表現などを文書体として適切な表現に置き換える。
- 挿入：発話中で省略された助詞などの語を補う。

## 3 CSJ への一次整形のアノテーションと本研究の対象

本研究に際して、CSJ のコア（188 講演）およびコアに含まれないテストセット（11 講演）に対して、一次整形の箇所を開始タグと終了タグで記述したデータを構築した。

これは私だけじゃなくて周り私の友人もやはり幼馴染みんなかばみんな青山の辺りに住んでおりますからみんな同じことでの前そのキディランドに行って何を買ってもらったって言うともう小学校で自慢し合うようなそういう感じだったんですけども

図 1: 講演の書き起こしの例

これは私だけじゃなくて私の友人もやはり幼馴染みはみんな青山の辺りに住んでおりますからみんな同じことでの前そのキディランドに行って何を買ってもらったって言うともう小学校で自慢し合うような感じだったんです

図 2: 削除箇所を除いた例

これは私だけではなくて私の友人もやはり幼馴染みは皆青山の辺りに住んでおりますから皆同じことでの前そのキディランドに行って何を買ってもらったと言うともう小学校で自慢し合うような感じだったのです

図 3: 削除・置換・挿入を行った例

表 1 に、一次整形における 3 つの要素の頻度を示す。3 つの処理のうちでは削除箇所が最も多く、全体の約 6 割を占めている。また、図 1~3 からわかるように、書き起こしから削除箇所を除く作業だけでも、可読性が十分に向上する。このような理由から、本研究では話し言葉の一次整形において削除される箇所の自動推定について検討を行う。

なお、今回利用したデータには他に、フィラー情報 (F タグ)、言いよどみ (D タグ)、言い直し表現 (R タグ) の各情報が、一次整形の各要素とは独立に付与されている。このうち F タグおよび D タグについては、CSJ に付与されている情報をそのまま利用している。ただし、F タグについては削除が容易であるので本研究の対象としない。

一方 R タグは、自己修復に関する代表的なモデルである RIM [2] を参考に、丸山らが CSJ の独話から収集した言い直し表現の体系 [5] をもとに新たに人手で付与した情報である。この R タグは、言い直し表現の機能的な分類に基づき、R1~R5 の 5 種類に分類される。

- R1: 発音エラーに伴う言い直し  
(例) コンテキスト {R1 (D いぞ)} || 依存 } モデルを使う
- R2: 単純な繰り返し  
(例) その下の {R2 波線 || 波線 } を付けました
- R3: 語彙的な誤りに伴う言い直し

(F ま) {R5 そんな風に思える |(F あの) ベットボトルの使い方もそういう {R3 風な || 風に } 考えられる } んじゃないかなと思います

図 4: 入れ子状に存在する R タグの例

(例) 独特の {R3 旋律の |(F えー) 旋律を } 形作っている

- R4: 情報不足に伴う言い直し  
(例) 海外に {R4 興味 || 元々興味 } があります
- R5: 別表現への言い換え  
(例) {R5 こちら側のグラフ || 右側のグラフ } は何を表しているかと言います

なお、文の構造により、図 4 のように複数の R タグが入れ子状に存在する場合も見られる。

上で述べた削除箇所および R タグは、それぞれ異なる基準で独立に付与されたラベルである。削除箇所のラベルは言語的特徴を問わず、整形作業をする上で削除すべき箇所に広く付与されている。一方、R タグは言い直し表現と認定された箇所について、整形作業に関係なく付与されている。削除箇所を R タグとの関係で分類すると、5 種類の R タグの被言い直し部 (RPD) と言い直し部 (RP) の品詞や意味の関係も考慮することにより、表 2 に示す 4 つのタイプに分類することができる。

このうち R タグに関する削除箇所には上で述べたような繰り返しの表現が多く見られ、「その他」つまりタイプ 4 に含まれるものには、以下のような種類の定形表現が多く見られた。

- 終助詞 (例) 富士山がよく見えるんですよ
- 文中・文末の冗長表現 (例) その前日ですね
- 文頭の接続詞 (例) で今ちょうど冬の時期で
- 間投的に用いられる表現 (例) 何ですかね

表 2 には、学習データにおける各 R タグと削除箇所との関係も示している。R1,R2,R3 が付与された箇所では、その大多数で被言い直し部 (RPD) が削除箇所に該当している。またこの 3 つに加えて R4 でも、言い直し部 (RP) を削除する場合は、RPD を削除する場合の 1 割に満たないのに対し、R5 が付与された箇所では RP を削除する場合は、RPD の場合の約半数と、その比率が大きく異なっている。

表 2: 各 R タグと削除箇所の関係

		総数	RPD を	RP を
		総数	削除	削除
タイプ 1	R1	1426	1421	113
タイプ 2	R2	1253	1183	69
	R3	1174	1108	57
	R4	961	624	60
タイプ 3	R5	595	250	116
タイプ 4	D タグのみ	2552	—	—
	その他	12127	—	—

検査	名詞	-	F	O	F	O
音	名詞	-	F	O	F	O
に対する	助詞	格助詞	O	O	F	O
反応	動詞	-	O	O	O	O
検査	名詞	-	F	O	F	O
音	名詞	-	O	B	O	B
へ	助詞	格助詞	O	O	O	B
の	助詞	格助詞	O	O	O	O
反応	名詞	-	O	B	O	B
と	助詞	格助詞	O	O	O	O

右から 3 列目と 4 列目は全く同じ形態素列の場合に繰り返すと認定  
右から 1 列目と 2 列目は自立語が同じでかつ付属語の品詞が同じ形態素列の場合に繰り返すと認定

図 5: 表現の重複（繰り返し）の例

## 4 削除箇所同定のアプローチ

本研究では、文章中の削除箇所の推定をテキストチャンキングの問題として扱い、SVMに基づくテキストチャンカである YamCha [9] を利用する。なお、チャンクラベルとして、IOB ラベリングスキームに基づいたラベルを形態素 (CSJ の短単位) ごとに付与する。YamCha における多項式カーネルの次数は 3、解析方向は Right-to-Left とした。

SVM に与える素性として、前後 3 形態素の単語情報 (表層表現、読み、品詞情報、活用の種類、活用形) やポーズの有無、フィルター情報、表現の重複 (繰り返し) の有無、および自己修復部の情報 (R タグ) を用いた。また上記に加えて、書き起こしデータに対する実験では文節境界と係り受け情報、さらに基本周波数の情報も用いた。表現の重複 (繰り返し) の有無は、図 5 のように、ある形態素列に対して、その前後 5 形態素および 10 形態素以内に、単語情報 (表層、読み、品詞) が全く同じ形態素列、および単語情報が同じ自立語でかつその後続く品詞が同じ形態素が含まれる形態素列を抽出し、文頭よりの表現と文末側の表現にそれぞれ F ラベルと B ラベルを独立に与えた。

書き起こしでの実験においては、品詞情報や文節、係り受けの情報はコーパスに付与された正解を用いており、フィルターは F タグが付与された感動詞としている。一方、音声認識結果での実験における削除箇所および R タグのラベルは、書き起こしデータに付与されたラベルを自動的に移行したものをと、その範囲が適切になるよう人手で修正したものである。また、認識結果において品詞が感動詞と認定されたものをフィルターとしている。

## 5 評価実験

### 5.1 RIM に基づく実験

まず、提案手法との性能の比較を目的として、自己修復部の検出における従来手法として代表的な RIM[2] に基づいた実験を行った。RIM では、自己修復部は (RPD | DF | RP) の連続した 3 つの区間に分割されるが、表 2 を考慮した上で、このうち RPD と DF を削除箇所とみなして評価を行う。

RIM ではまず DF の区間を定め、その後 RPD、RP の区間を推定していることから、本実験でも同様の手順で行った。フィルターまたは言いよどみであると判定された連続する形態素列を DF の候補とし、その前後 5 形態素に対しパターンマッチングを行うことにより、同一または読みが類似するものを RPD、RP として推定した。ここで推定された RPD と DF の区間を連続して一つの削除箇所とみなすが、削除箇所の定義よりフィルターについては削除箇所から除外している。

書き起こし・音声認識結果の両方に対して本実験を行った。書き起こしデータに対しては、音声認識テストセットのうち人手による形態素区切りが与えられている 19 講演を評価に使い、音声認識結果に対しても、同じ 19 講演を評価に用いている。なお、音声認識の単語認識精度は約 70% である。

実験結果を表 3 に示す。自己修復部を含まない削除箇所が多数を占めていることから、再現率は非常に低い結果となった。また、R タグを含む自己修復部に対しても低い結果となっている。これは、R タグが付与されている部分のうち、DF となりうるフィルターや言いよどみを持つものが全体の約 1/4 程度しかないことに起因している。

表 3: RIM による削除箇所推定精度

評価データ		範囲が一致		
		再現率	適合率	F 値
書き起こし	削除箇所全体	47/1874 (2.5%)	47/86 (54.7%)	0.048
	R タグを含む箇所	44/ 565 (7.8%)	44/86 (51.2%)	0.135
音声認識結果 (削除箇所全体)		20/1599 (1.3%)	20/75 (26.7%)	0.024

  

評価データ		始端が一致			範囲が重複		
		再現率	適合率	F 値	再現率	適合率	F 値
書き起こし	削除箇所全体	3.0%	65.1%	0.057	3.3%	81.4%	0.064
	R タグを含む箇所	9.4%	61.6%	0.163	10.1%	75.6%	0.178
(19 講演)							
音声認識結果 (削除箇所全体)		1.7%	36.0%	0.032	1.9%	42.7%	0.036

## 5.2 書き起こしデータに対する評価

本節では、書き起こしを入力とし、形態素解析や係り受け解析、R タグ認定の精度がそれぞれ正しい場合の、削除箇所の推定精度の評価を行った。

まずはじめに、単語情報やポーズ情報、繰り返しの情報のほかに、文節や係り受けなどの言語的情報を素性として用いた実験を行った。前節の実験と同一の書き起こし 19 講演を評価に使い、これを除くコアのデータ (169 講演) を学習セットとした。形態素や係り受け、R タグについては、人手で付与あるいは修正したものを用いた。

削除箇所の推定結果を表 4 に示す。繰り返しの素性を加えることにより、再現率の向上が見られた。また、文節や係り受け情報などの言語的情報を合わせて利用することで、より再現率が向上した。さらに R タグを利用した場合と利用しない場合の性能の差は F 値で 0.03 程度であるが、人手で付与されている R タグの自動推定の困難さを考慮すると、比較的頑健に自動推定できる素性のみでも高い性能が期待できるといえる。

次に、基本周波数 (F0) を素性として利用した実験を行った。ここで利用した F0 は、5 ミリ秒ごとに値が自動記録されている CSJ のデータの中から、各形態素の末尾に該当する値をとっている。同じコアのデータ 169 講演で学習を行っているが、CSJ では韻律情報はコアのみに含まれるため、学習データは同一であるが、評価データについては音声認識テストセットのうちコアに含まれる 8 講演のみを用いている。削除箇所の推定結果を表 5 に示す。F0 の導入により適合率の向上が見られ、推定精度が向上している。このことから、F0 も削除箇所の推定に有効な情報であることが確認された。

## 5.3 自己修復のパターン (R タグ) との関係

本節では、削除箇所タイプと削除箇所推定精度との関係を示す。本実験では、前節で用いた学習セット 169 講演およびテストセット 19 講演の書き起こしデータから、各 R タグについて、削除箇所のうち、R タグの付与された箇所と重なる部分のみを残し、同様に YamCha を用いて削除箇所の自動推定実験を行った。また、R タグの付与された範囲と重ならない削除箇所についても、その削除箇所のみを残して同様に実験を行った。ただし、R1 タグについては、被言い直し部 (RPD) が語断片など特殊な品詞情報が付与されているため、品詞情報および活用の種類・活用形を素性に用いなかった。

削除箇所の推定結果を表 6 に示す。3 章で示したタイプ 1,2 については、繰り返し等の素性による効果が見られたほか、R タグなど他の素性を加えても性能の向上が見られた。特に R2 については、繰り返しの素性の効果が大きく見られた。一方、R5 については、繰り返しの素性による効果は見られず、また R タグの利用に関わらず精度が非常に低い。これは、他のタイプの R タグでは被言い直し部 (RPD) が削除されることが多く、言い直し部 (RP) が削除される場合がほとんどないのに対し、R5 タグに該当する箇所では、RP が削除箇所となる場合も多く、削除すべきかどうかの判断には意味的な包含関係を考慮する必要があることが原因である。

## 5.4 音声認識結果に対する評価

本節では、書き起こしで学習したモデルを音声認識結果に対して適用した場合の性能を調べる。前節

表 4: 言語的情報を用いた書き起こし (19 講演) に対する削除箇所推定精度

素性	範囲が一致			範囲が重複		
	再現率	適合率	F 値	再現率	適合率	F 値
単語情報のみ	1239/1874 (66.1%)	1239/1460 (84.9%)	0.743	74.6%	95.3%	0.837
+ポーズ・フィラー	1270/1874 (67.8%)	1270/1495 (84.9%)	0.754	75.9%	95.7%	0.845
+ポーズ・フィラー・繰返し	1308/1874 (69.8%)	1308/1573 (83.2%)	0.759	78.4%	93.2%	0.852
+ポーズ・フィラー・繰返し・文節・係り受け	1412/1874 (75.3%)	1412/1657 (85.2%)	0.800	83.2%	93.9%	0.883
+ポーズ・フィラー・繰返し・文節・係り受け・R タグ	1495/1874 (79.8%)	1495/1726 (86.6%)	0.831	87.0%	94.3%	0.904

表 5: 基本周波数を用いた書き起こし (8 講演) に対する削除箇所推定精度

素性	範囲が一致			範囲が重複		
	再現率	適合率	F 値	再現率	適合率	F 値
単語情報のみ	573/826 (69.4%)	573/659 (86.9%)	0.772	76.5%	95.0%	0.848
+ポーズ・フィラー・繰返し	612/826 (74.1%)	612/724 (84.5%)	0.790	81.7%	93.2%	0.871
+ポーズ・フィラー・繰返し・F0	611/826 (74.0%)	611/706 (86.5%)	0.798	80.6%	94.2%	0.869
+ポーズ・フィラー・繰返し・F0・R タグ	697/826 (84.4%)	697/784 (88.9%)	0.866	89.7%	95.0%	0.923

で述べた書き起こしデータ 169 講演で学習を行ったモデルを、同一の 19 講演の音声認識結果に適用し評価を行った。なお、音声認識の単語認識精度は約 70%である。

評価結果を表 7 に示す。表 4 と表 7 を比較すると、書き起こしの場合に比べて、再現率、適合率ともほぼ半分に低下している。この低下幅は単語認識精度の 2 乗にほぼ相当する。特に、認識誤りによって言いどみとなった箇所では本来削除箇所とならないところでも、削除箇所と誤検出される例が多く見られた。ただし、本実験においても繰り返しの素性の導入により再現率の向上が見られている。繰り返しの素性は、単語の表層表現に基づいて定められていることから、音声認識結果に対しても頑健で効果的といえる。

## 6 おわりに

本稿では、一次整形のアノテーションを行ったコーパスを紹介し、これを用いて整形過程において削除される箇所を自動推定する手法を提案した。本手法では SVM による機械学習により削除箇所を自動推

定する。素性として、ポーズやフィラー、さらに表現の繰り返しの情報などを用いた。その結果、書き起こしに対して F 値で 0.8 前後の性能が得られた。一方、音声認識結果に対してはかなり精度が低下した。

今後の課題として、音声認識時における信頼度などの情報の利用が挙げられる。予備実験において、これを単に素性として利用した場合にその効果はほとんど見られなかったことから、他の情報と複合的に利用する必要があると考えられる。

## 謝辞

本研究の一部は、総務省 SCOPE「音声認識技術を用いた会議録及び字幕の作成支援システム」の支援により行われた。

## 参考文献

- [1] 河原達也. 筆記録作成のための話し言葉処理技術. 電子情報通信学会技術研究報告, SP2006-120, NLC2006-64 (SLP-64-36), pp. 209-214, 2006.

表 6: 各タイプごとの削除箇所推定精度

			範囲が一致			始端が一致			範囲が重複		
			再現率	精度	F 値	再現率	精度	F 値	再現率	精度	F 値
タイプ 1	R1	表層・読み	16.2%	51.2%	0.246	16.9%	51.2%	0.257	17.6%	60.5%	0.273
		+ポーズ (P)・フィラー (F)・繰返し	77.2%	77.8%	0.775	79.4%	80.0%	0.797	84.6%	86.9%	0.856
		+P・F・繰返し・文節・係受け・R タグ	97.8%	96.4%	0.964	99.3%	96.4%	0.978	99.3%	99.3%	0.993
タイプ 2	R2	単語情報のみ	18.3%	32.6%	0.234	25.6%	45.7%	0.328	62.2%	89.1%	0.733
		+ポーズ・フィラー・繰返し	57.3%	68.1%	0.623	65.9%	78.3%	0.715	80.5%	97.1%	0.880
		+P・F・繰返し・文節・係受け・R タグ	89.0%	90.1%	0.896	92.7%	93.8%	0.933	95.1%	97.5%	0.963
	R3	単語情報のみ	19.3%	43.2%	0.267	22.1%	49.4%	0.305	43.6%	92.6%	0.593
		+ポーズ・フィラー・繰返し	40.9%	59.2%	0.484	49.2%	71.2%	0.582	69.6%	97.6%	0.813
		+P・F・繰返し・文節・係受け・R タグ	76.2%	83.1%	0.795	86.2%	94.0%	0.899	93.9%	98.6%	0.960
	R4	単語情報のみ	25.8%	67.6%	0.373	26.8%	70.3%	0.388	38.5%	97.3%	0.552
		+ポーズ・フィラー・繰返し	40.2%	66.1%	0.500	47.4%	78.0%	0.590	60.4%	94.9%	0.738
		+P・F・繰返し・文節・係受け・R タグ	64.9%	80.8%	0.720	70.1%	87.2%	0.777	75.0%	93.6%	0.833
タイプ 3	R5	単語情報のみ	22.4%	55.0%	0.319	26.5%	65.0%	0.377	37.8%	92.5%	0.536
		+ポーズ・フィラー・繰返し	24.5%	43.6%	0.314	32.7%	58.2%	0.418	44.9%	83.6%	0.584
		+P・F・繰返し・文節・係受け・R タグ	33.7%	50.0%	0.402	42.9%	63.6%	0.512	58.2%	90.9%	0.709
タイプ 4	Rなし	単語情報のみ	78.5%	88.8%	0.833	80.2%	90.7%	0.851	82.9%	95.2%	0.886
		+ポーズ・フィラー・繰返し	79.2%	85.2%	0.821	80.5%	86.6%	0.835	84.1%	94.4%	0.889
		+P・F・繰返し・文節・係受け・R タグ	81.5%	89.0%	0.851	83.0%	90.7%	0.867	85.9%	94.0%	0.898

表 7: 音声認識結果 (19 講演) での削除箇所推定精度

素性	範囲が一致		
	再現率	適合率	F 値
単語情報のみ	605/1873 (32.3%)	605/1398 (43.3%)	0.370
+ポーズ・フィラー	600/1873 (32.0%)	600/1428 (42.0%)	0.364
+ポーズ・フィラー・繰返し	641/1873 (34.2%)	641/1501 (42.7%)	0.380

  

素性	始端が一致			範囲が重複		
	再現率	適合率	F 値	再現率	適合率	F 値
単語情報のみ	33.5%	44.4%	0.382	58.2%	78.8%	0.670
+ポーズ・フィラー	33.4%	44.3%	0.377	58.4%	78.0%	0.668
+ポーズ・フィラー・繰返し	36.1%	44.6%	0.399	60.9%	78.1%	0.684

- [2] C. Nakatani and J. Hirschberg. A speech-first model for repair identification and correction. In *Proc. ACL*, 1993.
- [3] 下岡和也, 河原達也, 内元清貴, 井佐原均. 『日本語話し言葉コーパス』における自己修復部 (D タグ) の自動検出および修正に関する検討. 情報処理学会研究報告, 2005-NL-167-14, 2005-SLP-56-14, pp. 95-100, 2005.
- [4] 日本速記協会編. 発言記録作成標準. 2007.
- [5] 丸山岳彦, 佐野真一郎. 自発的な話し言葉に現れる言い直し表現の機能的分析. 言語処理学会第 13 回年次大会, pp. 1026-1029, 2007.
- [6] 浅原正幸, 松本裕治. 形態素解析とチャンキングの組み合わせによるフィラー/言い直し検出. 言語処理学会第 9 回年次大会, pp. 651-654, 2003.
- [7] 下岡和也, 河原達也, 奥乃博. 講演の書き起こしに対する統計的手法を用いた文体の整形. 情報処理学会研究報告, 自然言語処理 (149-12), 音声言語情報処理 (41-3), pp. 81-88, 2002.
- [8] 古井貞熙, 前川喜久雄, 井佐原均. 科学技術振興調整費開放の融合研究制度: 大規模コーパスに基づく『話し言葉工学』の構築. 日本音響学会誌, Vol. 56, No. 11, pp. 752-755, 2000.
- [9] T. Kudo and Y. Matsumoto. Chunking with Support Vector Machines. In *Proc. NAACL*, 2001.