

要約筆記品質評価システムにおける要約筆記支援機能

Support Function of Quality Evaluation System of Summary Transcript

高尾 哲康

Takao Tetsuyasu

1. はじめに

聴覚障害者や高齢者への情報保障手段のひとつに要約筆記がある。要約筆記には「PC 要約筆記」と「手書き要約筆記」があり、いずれも要約筆記者が講演や番組などを聞き取り、リアルタイムで要約を行ない、キーボードや手書きで入力する。一般に日本語の発話速度は200～400文字/分であり、要約筆記者による入力量はPCの場合で100～200文字/分、手書きの場合で40～80文字/分となっている。要約筆記者は「速く」、「正確に」、「読みやすく」の3原則をもとに、技術の向上を目指してさまざまな研修プログラムで訓練を重ねる。個々の研修プログラムでは要約筆記の品質の尺度として、要約筆記利用者からのフィードバックや意見・要望を受けることが多い[1][2]。これらのフィードバックは個々の事例として受けることが多く、定量的な品質評価を受けることはほとんどなかった。そのため、長期間の研修を経ても要約筆記の品質向上の実感が得られにくくなっていた。これまで筆者らは講演者の発話内容のテキストと要約筆記者が入力したテキストをもとに定量的な評価ができるシステムを試作した[3][4][5]。要約評価となる評価値計算には重み付き編集距離単位（主に形態素基本形と品詞）列の編集距離（Levenshtein Distance）計算に基づく方式を提案した。これにより個々の文や段落など局所的範囲での評価に有効性を確認できた。さらに、要約筆記されたテキストどうしの関連性や要約筆記者それぞれのタイプやくせを見つけ出すために多次元尺度法や潜在意味解析などの評価も行なえるようにした[4]。本論文では、要約評価値計算の過程で得られた情報をもとに要約筆記者が書き下したテキストについて、よりよい要約テキストとなるような書き換え候補を提示し[5]、さらによりよい要約になっていることを確認できる機能を紹介する。これにより、要約筆記者がアドバイスを受けることでよりよい要約ができるようになり、要約筆記者の養成支援になることを目指している。また、コンピュータによる自動要約システム[6]への知見が得られると考える。

2. 品質評価に利用した要約筆記データ

要約筆記講習会の研修プログラムで使用した発話テキスト（エッセイ文（約6分）、Sとする）と要約筆記者6名が手書きでリアルタイム要約筆記したテキスト（K1～K6とする）を利用した。

詳細を表1に示す。表1には、編集距離に基づく要約評価結果も載せた。文字数には句読点や記号(矢印記号「→」、項目を表わす中黒「・」)、繰り返し記号(「〃」など)、削除記号(訂正線)などを含めている。計算機可読テキストにする際には二次元的な表現や複数行にわたる括弧記号などの意図がわかるようにXMLタグ付きテキストにした。

表1. 要約筆記テキストと要約評価値の向上

	文字数	文字数/分	要約率 (%)	要約評価値の変化(全発話との比較)					
				要約筆記原データ	各単位最大1箇所修正		各単位最大2箇所修正		向上率
全発話(S)	1808	226.0							
筆記者1(K1)	545	68.1	30.14%	0.6649	0.7485	(10)	0.7718	(3)	0.1069
筆記者2(K2)	434	54.3	24.00%	0.6635	0.7417	(8)	0.7889	(3)	0.1254
筆記者3(K3)	539	67.4	29.81%	0.5625	0.6867	(8)	0.7549	(6)	0.1924
筆記者4(K4)	665	83.1	36.78%	0.7783	0.8328	(7)	0.8496	(4)	0.0713
筆記者5(K5)	395	49.4	21.85%	0.5101	0.5624	(4)	0.6491	(4)	0.1390
筆記者6(K6)	443	55.4	24.50%	0.4512	0.5454	(6)	0.6067	(6)	0.1555

3. 要約筆記品質評価システム

要約には「テープ起こし」(要約率 90%) や丁寧な要約筆記である「概要要約筆記」(要約率 40~60%) などがある。PC を利用した 2 人連携要約筆記では、後編集処理をほどこすだけでも十分概要要約筆記のレベルに到達できる。本システムでは、PC や手書きを問わず、さまざまな要約レベルにも対応できるようにしている。本システムの概要、システム構成を図1に示す。

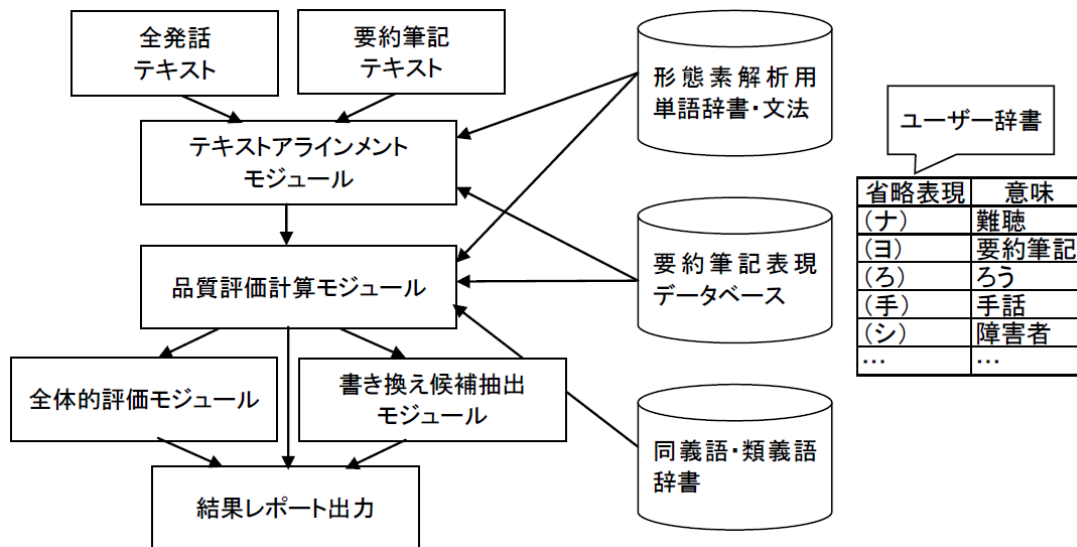


図1. システム構成図

本システムのメインはテキストアライメントモジュールと品質評価計算モジュールから構成される。テキストアライメントモジュールは全発話テキストと要約筆記テキストを入力とし、統計情報と言語情報をもとに、動的計画法を利用して対応する文や段落を対応づけるモジュールである（m文対n文）[7]。アライメント単位ごとに発話文と要約筆記文のペアが作成される。これにより品質評価計算対象範囲を狭くすることにより、後段の品質評価計算モジュールにおける評価計算精度を高めることができる。品質評価計算モジュールは、表記ゆれや要約筆記特有の省略表現などを吸収して正規化した形態素解析結果の形態素列に対し、単語コスト、品詞コスト、単語間接続コスト、重複出現コスト（出現のたびに単調減少）を統計処理することにより、要約の品質評価の計算を行なう[3]。さらにテキスト全体およびアライメント単位で形態素列のN-gramに基づいた多次元尺度法と潜在意味解析を行なう。多次元尺度法は、個体間の類似度をそれぞれの個体から得られる多変量の数値をもとに個体間距離を計算し、類似性の高いものが近くに配置するようにする方式である。潜在意味解析は単語の持つ多義性と多様性に対処した統計的技法であり、多変量ゆえに高次元となる単語文書行列を低次元に圧縮して近似的に表現する。これにより関連性の強い単語どうしが強調され、文書間の意味的類似度が明確になる。この処理には、統計解析用ソフトRとそのライブラリ群を利用した[4]。

要約候補文の抽出は、要約評価値を算出するマトリクス（表2）において、最右下のセルから最左上のセルまで評価値が最も小さくなる方向（上方、左方、左上方のいずれか）に順次たどることで発話文と要約筆記文との対応セルを求める。次に、発話文と要約筆記文との対応関係のうち相互にマッチしないもの（前後のセル間で評価値の差が大きい場合）を抽出する。表2では、

- ・「急性中耳炎は」（発話）と文頭（要約筆記）
- ・「で起こる。」（発話）と「で」（要約筆記）

が該当する。この際、直前のセルの評価値との差が大きいもの、発話文内の形態素コストの大きいものから優先的に要約候補文を提示する。なお、候補文の提示における修正箇所は自立語を含む文節単位とし、文章としての形態素間のつながりが保たれるようにした。

表2. 評価値計算と要約候補抽出

i \ j			急性	中耳炎	は	風邪	が	原因	で	起こる	。
		コスト値	3135	4000	10	2914	10	2326	10	2206	188
	コスト値	0.000	0.156	0.356	0.356	0.502	0.502	0.618	0.618	0.728	0.738
カゼ	2914	0.145	0.302	0.501	0.502	0.356	0.357	0.473	0.473	0.583	0.593
が	10	0.146	0.302	0.502	0.502	0.357	0.356	0.472	0.473	0.583	0.592
原因	2326	0.262	0.418	0.617	0.618	0.473	0.472	0.356	0.357	0.467	0.476
で	10	0.262	0.419	0.618	0.618	0.473	0.473	0.357	0.356	0.466	0.476

$$E_{i,j} = \min(E_{i-1,j} + C_{i-1}/C, \quad E_{i,j-1} + C_{j-1}/C, \quad E_{i-1,j-1} + A)$$

$$A = \begin{cases} 0 & : i-1 \text{ と } j-1 \text{ の位置の形態素がマッチ} \\ & \text{(表記基本形、品詞、同義語) した場合} \\ (C_{i-1} + C_{j-1})/C & : \text{上記以外 (C:コスト値の総和)} \end{cases}$$

4. 実験結果

筆記者ごとの要約筆記文を本システムに適用した結果を表1の右側に示す。各筆記者が書き下したままの要約筆記文（要約筆記原データ）の要約評価値とともに、アラインメント単位ごとの要約候補文の提示にしたがい、1箇所および2箇所の修正をほどこした場合の要約評価値の向上率も示した。要約評価値向上率の変化を図2に示した。要約筆記原データの要約評価値が低い場合（要約筆記に十分習熟していない場合）ほど向上率が高くなることわかる。要約候補文提示による効果は、要約筆記の初心者ほど効果が高いことになり、学習効果や訓練効果が大きいことわかる。

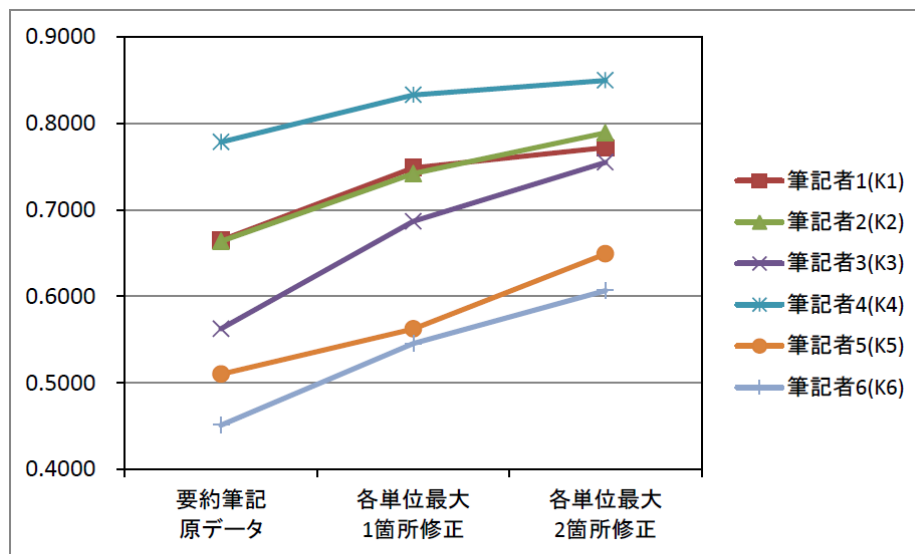


図2. 要約候補選択時の要約評価値の変化

要約筆記文修正候補の例を下記に示す。要約筆記文に対して、下線部分が候補として提示された（下線部分が第1候補、二重下線部分が第2候補）。

発話文：「だれかに自分の話を聞いてもらうことを望んでいます。」

要約文：「聞いてもらうことを望む。」

候補文1：「自分の話を聞いてもらうことを望む。」

候補文2：「だれかに 自分の話を聞いてもらうことを望む。」

発話文：「えー少しでも気になる症状があれば、ぜひ一度、耳鼻咽喉科を受診なさってください。」

要約文：「気になる時は受診を。」

候補文1：「気になる時は耳鼻咽喉科を受診を。」

候補文2：「気になる症状時は耳鼻咽喉科を受診を。」

要約筆記文修正候補失敗例を下記に示す。二重否定文や形容動詞・副詞の誤用などがある。

発話：「人の話を聞くのは、そう簡単なことではないのです。」

要約：「人の話を聞くのは難しい」

候補1：「人の話を聞くのは簡単なこと」(×)

候補2：「人の話を聞くのは簡単なこと ではない」

発話：「話をよく聞いてくれる人が意外に少ないことに気がつくはずです。」

要約：「そんなに多くはいない。」

候補1：「そんなに多くはいない、気がつくはず。」

候補2：「そんなに多く意外にはいない、気がつくはず。」(△)

次に向上率が高かった K3 と K6 について、改良による効果を多次元尺度法にて図 3 に示した(例えば、K3 は要約筆記原データのプロットを示し、K31 は原データに対して 1 箇所を修正、K32 は 2 箇所を修正した後のプロットを示す)。1 箇所の修正で発話文 S に大きく近づき、2 箇所の修正でさらに近づくことが明確になった。

さらに、2 箇所の修正前後の各要約筆記者の類似度の潜在意味解析による結果を図 4 に示した。これも 2 箇所の修正後は発話文 S の近傍によりまとまりつつあることがわかる。

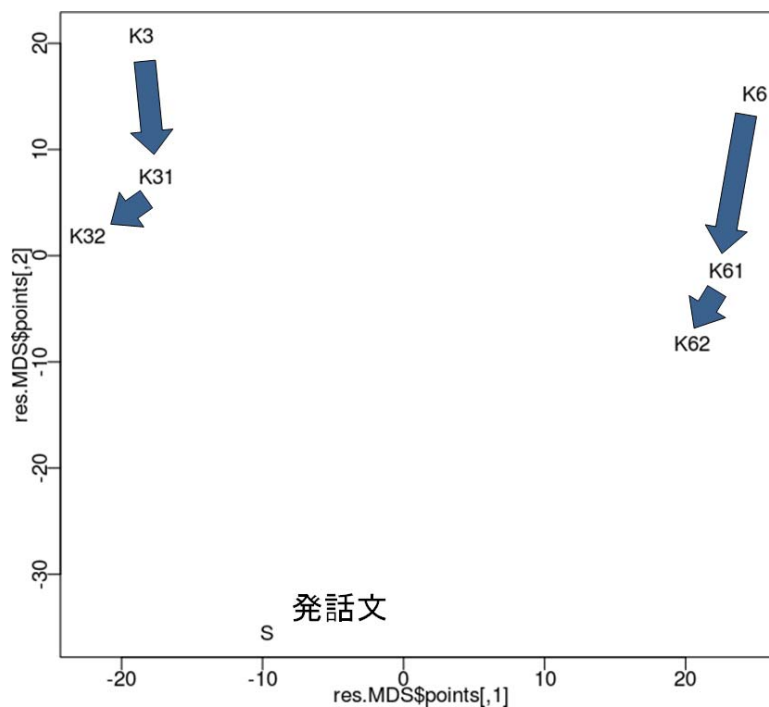


図 3. 多次元尺度法による解析結果 (1 箇所および 2 箇所修正時の変化)

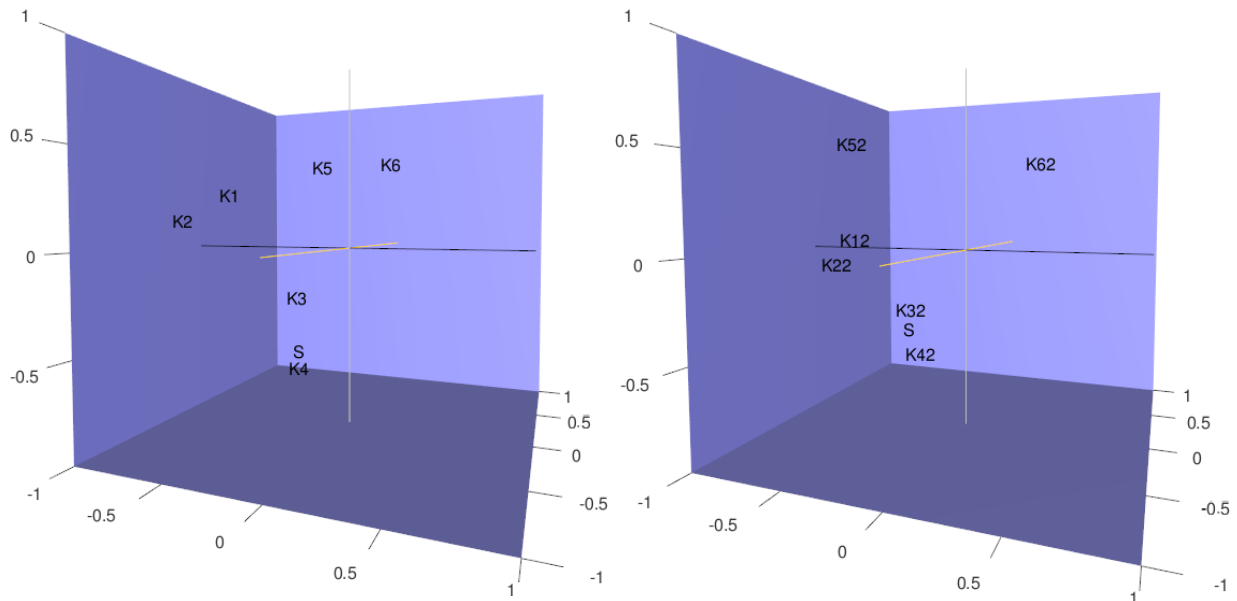


図4. 潜在意味解析による解析結果(右側は2箇所修正時の解析結果)

5. まとめ

本実験から要約文の品質向上のためのアドバイス機能(改良候補提示、改良した場合の全体的評価)の効果が確認できた。今後は、さまざまな要約筆記データを収集し、要約評価精度の向上や失敗箇所についての分析を進めていくとともに、

- ・「急性中耳炎は風邪が原因」→「原因はかぜ」
- ・「話すことより聞くことが重要」→「聞くことが話すことより重要」
- ・「2つの耳と1つの口」→「1つの口、2つの耳」
- ・「人の話を聞くのは簡単なことではない」→「人の話を聞くのは難しい」

などのように、柔軟な文節配置に対応するために構文情報を利用した評価計算の改良などを行なっていく。また、複数人での要約筆記文をマージしてよりよい要約筆記文を構築するなどの応用も検討していく。

参考文献

- [1]要約筆記再履修資料、名古屋市登録要約筆記者の会編(2007)
- [2]話しことばの要約、三宅初穂、全国要約筆記問題研究会(2012)
- [3]高尾哲康、要約筆記品質評価システムの改良、FIT2011、3Q-5、(2011)
- [4]高尾哲康、要約筆記品質評価システムにおける書き手のタイプ判別、IPSJ74 全国大会、3F-4、(2012)
- [5]高尾哲康、要約筆記品質評価システムにおける要約候補文提示機能、FIT2012、2M-6、(2012)
- [6]特集 テキスト自動要約、情報処理、Vol.43、No.12(2002)
- [7]高尾哲康、対訳テキストコーパスからの対訳語情報の自動抽出、情報処理学会自然言語処理研究会、115-8、(1996)