

要約筆記の品質評価の試み

Trial of Quality Evaluation 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として、手話通訳や字幕放送などとともに、要約筆記が用いられている。要約筆記にはPCなどを利用して行なう「パソコン要約筆記」、手書きで行なう「手書き要約筆記」がある。パソコン要約筆記では情報の入力を主にキーボードから行ない、大人数を対象とする際にはプロジェクタ装置を介してモニタ画面をスクリーンに表示する。また、手書き要約筆記には、少人数の情報保障を対象とするノートテイク、大人数を対象とする方法としてOHPやOHC(Over Head Camera)が利用されている。デジタルペンやタブレットを利用する場合もある。いずれにして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ことになる。一般に日本語での発話速度は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。したがって、速記などとは異なり、発話内容のすべてを入力することはできないので、何らかの要約が行なわれることになる。また、要約筆記者による書き間違いや聞き間違いなどが入ることもある。

要約筆記者はさまざまな講演や会議などへの派遣要請を受けて活動することになるが、それとともに技術の向上を目指してさまざまな研修プログラムを受講することで訓練することになる。しかし、個々の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。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ほとんどであり、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、長い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。

本論文では、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ようにすることが目的である。これにより、要約筆記の技術レベルの診断や改善点などが明確にわかるようになり、今後の研修プログラムの指針となる可能性が開ける。また、コンピュータによる自動要約システム[1][2]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について

要約筆記講習会などの研修プログラムで利用している講演内容のテキストと数名の要約筆記者が手書きでリアルタイム要約筆記したテキストを利用した。詳細を表1に示す。講演はNHK教育テレビ番組の「きょうの健康—^{しんしゅつ}滲出性中耳炎」(2008年2月3日放送)である。要約筆記者は7名。要約筆記の模範例も用意した。模範例はあくまで参考であり、目標となる要約率20%とした。全発話の文数は158である。なお、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(「〻」など)、削除記号(訂正線)なども含まれている。手書き要約筆記の場合はPCによる要約筆記とは異なり、二次元的な表現、複数行にわたる括弧記号や行を越えた矢印記号、横方向の波括弧付きの挿入文字などが含まれていることがある。この場合には、計算機可読テキストにする際にはなるべくそのまま残し、大きく編集しないように心がけた。

表1. 要約筆記テキスト

	文字数	文字数/分	要約率(%)
全発話	5059	389.2	
筆記者1	544	41.8	10.8%
筆記者2	519	39.9	10.3%
筆記者3	666	51.2	13.2%
筆記者4	559	43.0	11.0%
筆記者5	631	48.5	12.5%
筆記者6	743	57.2	14.7%
筆記者7	490	37.7	9.7%
模範例	995	76.5	19.7%

3. 要約筆記品質評価システム

要約にはさまざまなレベルがある。前述のリアルタイム手書き要約筆記では、その場における情報保障が目的であり、使い捨てとなるので文体や表現などにはこだわりのない。他のレベルとしては、講演内容のほとんどを筆記する「テープ起こし」がある。この場合は「あー」、「えー」など間投詞、フィラー(ケバ)や冗長表現・重複表現などを削除し、適宜文末処理を行なって読みやすくする。この場合の要約率は90%程度となる。また、概要要約筆記として後日の読み返しにも十分耐えうるレベルのものがある。この場合は丁寧な要約筆記であり、要約率は約40~60%程度になる。PCを利用した二人連携要約筆記などでは、後処理をほどこすだけで十分にこのレベルに到達できる。

本システムでは汎用性とロバスト性をもたせるため、PCや手書きを問わず、また前述のさまざまな要約レベルにも対応できるようにしている。本システムの構成を図1に示す。

テキストアラインメントモジュールは、以前に筆者が作成した、テキストの統計情報と言語情報をもとに、DPマッチング法を利用して対応する文や段落を関連づけるモジュールである(n文対n文の対応付け)[3][4]。これにより、品質評価計算対象範囲を狭く限定することにより、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。なお、極端に要約さ

れたテキストや要約時に言い換え表現が多くなったテキストなどの場合は対応させるテキストどうしの範囲が広がってしまうことになる。その場合は品質評価計算対象範囲が広がることにより品質評価の精度が悪くなる。このような時は人手でより細かく対応させることができれば品質評価の精度を高めることができる。

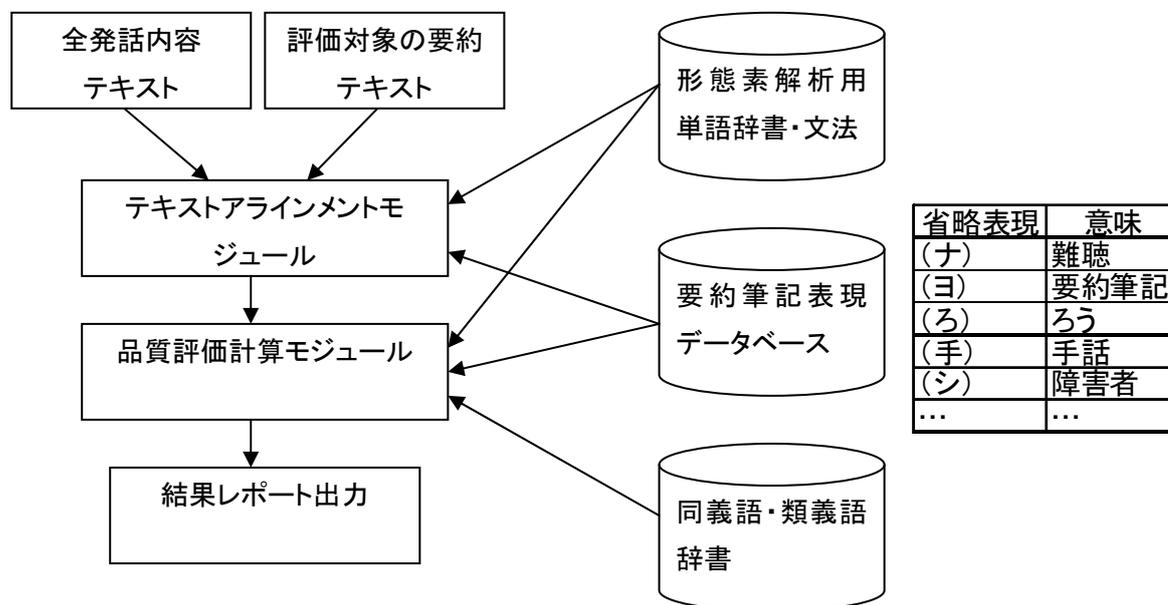


図1. 要約筆記品質評価システム

要約筆記表現データベースは、主に手書き要約筆記で使われる省略表現を収録したものである。品質評価計算モジュールは、表記のゆれ（漢字のひらがな・カタカナ表記など）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テキストに対し、単語コスト（頻度に相当）、品詞コスト（固有名詞や専門用語、数詞、名詞、動詞などは高くなり、助詞・助動詞などの機能語や感嘆詞・間投詞、活用語尾などは低くなる）、単語間接続コスト（単語どうしのつながりの程度を表わす）、重複出現コスト（何度も出てくると少しずつ低くなる）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。具体的には次のような指針で計算する。

- 全発話内容テキストと要約筆記テキストが同一であれば満点となる。全発話内容テキストの代わりに要約筆記模範例のテキストを指定してもよい。その場合は模範例を目標とみなしてどれだけ十分に要約できているかを判定することになる。模範例の品質を越える要約筆記テキストの場合もあるが、その場合でも同一と同じとみなす。
- 全発話内容テキストに出現する固有名詞や専門用語、数詞、名詞、動詞など、コストが高い（一般的に出現頻度が低い）単語は、全発話内容テキストに出現する順序で要約筆記テキストに出現すれば評価は高くなる。たとえ構文表現の都合などで出現順序が入れ替わっても出現位置のずれが小さければ評価の低下は小さくなる。
- 全発話内容テキストに出現する感嘆詞・間投詞や助詞・助動詞、活用語尾などは要約筆記テキストに出現しなくても評価はあまり下げないが、出現しても評価はあまり高くない。
- 全発話内容テキストに含まれる重複表現のうち、2番目以降が要約筆記テキストに出現し

なくても評価は下げないが、出現してもそれほど高くはならない。

- 同義語や類義語による言い換えや、言い回しの言い換え表現（二重否定→肯定など）はなるべく認めるようにして、評価は下げない。
- 書き忘れや書き間違いなどがあればその分だけ評価が下がる。
- 品質評価結果は全体とともにテキストアラインメント単位でも出力する。これにより、局所的な評価（最小で1文対1文）が得られるので、要約筆記のスキル向上に役立つことになる。

要約筆記においては「ことばを追うな。意味を追え」[5]と言われるように、話しことばを単に文字化するのではなく、利用者が読みやすいように意味主導型の要約筆記へと発展することが望まれている。しかし、リアルタイム要約筆記においては意味主導で筆記するだけの十分な時間的・労働的余裕が少ないので、結局は話された内容に含まれる単語を中心に筆記することが多くなる。そのため、意味処理を行わない本システムでも活用性は広いと考える。

4. 実験結果

表1の要約筆記データを本システムに適用した結果を表2に示す。なお、テキストを単なる文字列と見なしてその編集距離(Edit Distance または Levenshtein Distance)を計算したもの、テキストを形態素解析結果の分かち書きテキストと見なしてその編集距離を計算したものも同時に算出した。編集距離とは文字列Aと文字列Bについて、Aを編集操作（文字の削除、挿入、置換）してBにするときの必要最低限の操作数のことである。他の尺度と比較するために、文字列編集距離の場合は比較元テキストの文字列長、形態素編集距離の場合は比較元テキストの形態素数で割り、数値の範囲を0~1に正規化した数値とした。0に近ければ元のテキストとの相違が多く、1に近ければ相違が少なくなる。要約評価の場合は前述の各種コストを総合的に計算したものとなる。

表2. 要約評価例

	全発話内容との比較			模範例との比較		
	編集距離		要約評価	編集距離		要約評価
	文字列	形態素		文字列	形態素	
筆記者 1	0.1524	0.1245	0.2627	0.3821	0.3317	0.4014
筆記者 2	0.1587	0.1344	0.2660	0.2813	0.2547	0.3495
筆記者 3	0.2006	0.1753	0.3401	0.3939	0.3230	0.3975
筆記者 4	0.2067	0.1938	0.3474	0.4446	0.4401	0.5079
筆記者 5	0.1539	0.1260	0.3032	0.2951	0.2610	0.3716
筆記者 6	0.1917	0.1673	0.3281	0.3297	0.2980	0.4132
筆記者 7	0.1672	0.1503	0.2816	0.3340	0.2913	0.3770
模範例	0.2283	0.2079	0.3855	1.0000	1.0000	1.0000

模範例どうしを要約評価した場合は内容が同一なので編集距離、要約評価のいずれも 1.0000

となる。表1の要約率と表2を見てみると、筆記者6は他の筆記者に比較してよく書けているのに要約評価としてはそれほど高くなっていないことがわかる。また、筆記者4は内容をよく把握し、かつコンパクトに要約して書いていることもわかる。

表2のうち、筆記者4についてテキストアライメント単位で品質評価計算した結果を図2に示す。編集距離、要約評価のいずれも1.0になっているテキストアライメント部分は、発話内容と同一の文、発話内容：「声をかけても振り向かない」、要約内容：「声をかけてもふりむかない、」となっていた。また、要約評価より文字列編集距離、形態素編集距離が高い箇所がいくつかあった。これを以下に示す。

- 発話内容：「滲出性中耳炎。」
要約内容：「滲出性は」
→専門用語「中耳炎」を専門用語とみなしているためにコストを高くしている。しかし、要約テキスト内では既出なので、この時点での要約評価計算の際にはその分コストを下げてはいるが、重複出現コストの計算面での扱いが不十分であった。
- 発話内容：「このチューブは、入れっぱなしにしておくんですか。」
要約内容：「チューブは入れっぱなし」
→会話的な言い回し処理の解析が不十分であった。
- 発話内容：「そして最後に、マクロライド薬というのがありますね。」
要約内容：「・マクロライド薬」
→会話的な言い回しとともに、単語「最後」、「ある」のコストを考慮しすぎた。

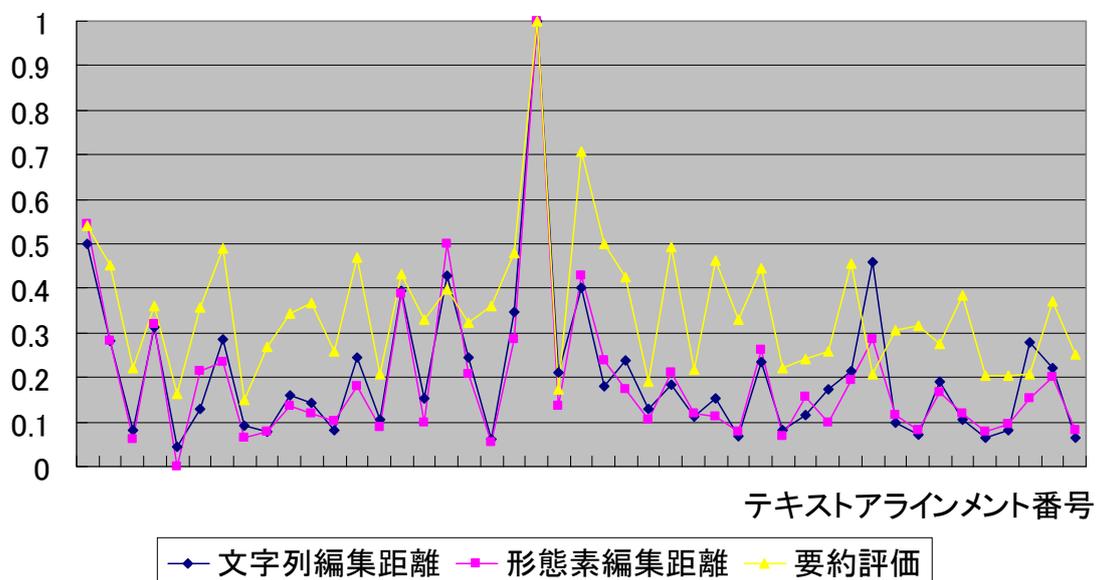


図2. 筆記者4のテキストアライメント単位ごとの要約評価

本システムによる評価の効果がよく出ているテキストアライメント箇所には以下のものがあった。いずれも、単なる文字列や形態素列の編集距離に比べると高い評価となっている。

- 発話内容：「あの一、中耳炎の症状が長引いて、慢性化した場合、まあ滲出性中耳炎になっ

ているということがあります。」

要約内容：「中耳炎をほうっておくところの滲出性中耳炎になる。」

- 発話内容：「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。」

要約内容：「気になる時は受診を。」

→「症状」、「耳鼻咽喉科」は既出なので重複出現コスト下げ効果が出ている。

- 発話内容：「ではこの滲出性中耳炎の治療はどのようにしていくんでしょうか。」

要約内容：「滲出性治療は」

以上のように、よく要約している箇所については要約評価が高くなっていることで本システムの有効性が確認できた。要約評価と他の尺度との差が大きい箇所はよい要約筆記の事例となる。つまり、要約評価と文字列・形態素編集距離との差が極大となる時が最高品質の要約となる。

今後は、うまく評価できなかった箇所についての分析を進めるとともに、コスト計算手法やパラメータの最適化などを行なう。また、要約率をコスト計算の閾値に取り込むことにより、要約率に応じた要約レベルの指針にもなりえる。さらに、テキストアラインメント精度や形態素解析精度のさらなる向上もはかっていく必要がある。

5. まとめ

「高齢者・聴覚障害者からみた字幕表示のあり方に関する研究」[6]でも述べたように、発話内容をできるだけ文字化して提示するやり方から、文字を読む側の立場を考慮してよりコンパクトに内容を十分に伝えることができるものとするのが求められている。リアルタイム字幕放送などでは、話者の発言内容をなるべく多く文字化することに力を入れているため、字幕テキストが長くなって読みきれないことが多くなっている。そのため、内容がつかみにくくなってしまっている。また、分野によっては会話そのものを楽しみたいとの要望もあることから、話者のニュアンスをも何らかの形で取り込む方法を模索する必要がある。本論文で本システムの有用性は確認できたが、まだまだ改良点も多く残されており、今後の研究が望まれる。

最後に、要約筆記データの提供をいただいた要約筆記奉仕員の方々に感謝する。

参考文献

- (1) Inderjeet Mani 著、奥村 学・難波英嗣・植田禎子訳、自動要約、共立出版(2003)
- (2) 特集 テキスト自動要約、情報処理、Vol.43、No.12 (2002)
- (3) 高尾哲康、対訳テキストコーパスからの対訳複合語の自動抽出、山形県立米沢女子短期大学紀要、pp.31-41, (1996)
- (4) 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)
- (5) 要約筆記再履修資料、名古屋市登録要約筆記者の会編 (2007)
- (6) 香取淳子、高尾哲康、高齢者・聴覚障害者からみた字幕表示のあり方に関する研究、ユニバーサル財団研究報告、2007