

## 要約筆記の品質評価の試み

高尾 哲康<sup>†</sup>

富山国際大学現代社会学部<sup>‡</sup>

### 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、要約筆記が用いられている。要約筆記には「PC 要約筆記」、「手書き要約筆記」があり、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。要約筆記者は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する。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[1]。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[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

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。講習テキストには NHK 教育テレビ番組の「きょうの健康—<sup>しんじつ</sup>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(「〃」など)、削除記号(訂正線)などを含めている。手書き要約筆記の場合は PC による要約筆記とは異なり、二次元的な表現や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その意図がわかるような XML タグ付きテキストにした。

### 3. 要約筆記品質評価システム

要約には講演内容のほとんどを筆記する「テープ起こし」(要約率 90%)、丁寧な要約筆記である「概要要約筆記」(要約率 40~60%)がある。PC を利用した 2 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た。本システム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評価対象の要約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関連づけるモジュールである(m 文対 n 文の対応付け)[3]。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。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る(前述の m, n の値が大きいく)。このような場合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。

品質評価計算モジュールは、表記のゆれ(漢字の読みのひらがな・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、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する)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。具体的には主に次のような指針で計算する。

Trial of Quality Evaluation of Summary Transcript

<sup>†</sup>Toyama University of International Studies

Faculty of Modern Society

<sup>‡</sup>Tetsuyasu Takao

・全発話テキストに出現する固有名詞や専門用語、数詞などコストが高い(一般的に出現頻度が低い)単語は、全発話テキストに出現する順序で要約テキストに出現すれば評価は高くなる。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い。

・全発話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テキストに出現しなくても評価はあまり下げない。

・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。

品質評価結果は全体とともにテキストアライメント単位でも出力する。局所的評価(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。要約筆記においては話しこと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[1]。しかし、意味主導で筆記するだけの十分な時間的余裕が少なく、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

評価計算には、テキストを形態素列と見なし、0~1 の重み付き編集単位要素の編集距離(Edit Distance または Levenshtein Distance)を利用した。編集距離は列 A と列 B について、A を編集操作(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。操作数を、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。要約評価は重み付き操作数を比較元テキストの形態素数で割り、数値の範囲を 0~1 に正規化した数値をベースにした。筆者ごとの要約データを本システムに適用した結果を表 1 に示す。筆者 6 は他の筆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く、筆者 4 は内容をよく把握し、コンパクトに要約して書いていることがわかる。筆者 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 1 に示す。本システムによる評価の効果がよく出ている箇所には以下のものがあつた。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(文字列編集距離や形態素編集距離)との差が大きい箇所はよい要約筆

表 1. 要約筆記テキストと要約評価

|     | 文字数  | 文字数/分 | 要約率(%) | 全発話との比較 |        |        | 模範例との比較 |        |        |
|-----|------|-------|--------|---------|--------|--------|---------|--------|--------|
|     |      |       |        | 編集距離    | 要約     | 評価     | 編集距離    | 要約     | 評価     |
| 全発話 | 5059 | 389.2 |        | 文字列     | 形態素    |        | 文字列     | 形態素    |        |
| 筆者1 | 544  | 41.8  | 10.80% | 0.1524  | 0.1245 | 0.2627 | 0.3821  | 0.3317 | 0.4014 |
| 筆者2 | 519  | 39.9  | 10.30% | 0.1587  | 0.1344 | 0.2660 | 0.2813  | 0.2547 | 0.3495 |
| 筆者3 | 666  | 51.2  | 13.20% | 0.2006  | 0.1753 | 0.3401 | 0.3939  | 0.3230 | 0.3975 |
| 筆者4 | 559  | 43.0  | 11.00% | 0.2067  | 0.1938 | 0.3474 | 0.4446  | 0.4401 | 0.5079 |
| 筆者5 | 631  | 48.5  | 12.50% | 0.1539  | 0.1260 | 0.3032 | 0.2951  | 0.2610 | 0.3716 |
| 筆者6 | 743  | 57.2  | 14.70% | 0.1917  | 0.1673 | 0.3281 | 0.3297  | 0.2980 | 0.4132 |
| 筆者7 | 490  | 37.7  | 9.70%  | 0.1672  | 0.1503 | 0.2816 | 0.3340  | 0.2913 | 0.3770 |
| 模範例 | 995  | 76.5  | 19.70% | 0.2283  | 0.2079 | 0.3855 | 1.0000  | 1.0000 | 1.0000 |

記の事例となる。

発話：「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：「中耳炎をほうっておくとこの滲出性中耳炎になる。」

発話：「え一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：「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：「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：「滲出性治療は？」

今後は、失敗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、並列構文→箇条書きなどの評価法を検討していく(XML タグの利用)。

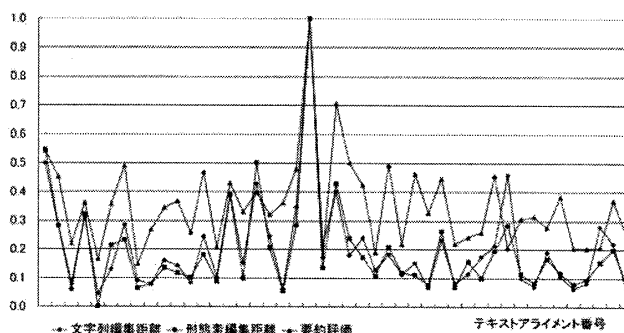


図 1. テキストアライメント単位の要約評価

#### 5. まとめ

本実験でシステムの有効性は確認できたが、まだ改良点も多く、今後の研究が望まれる。

#### 参考文献

- [1] 要約筆記再履修資料、名古屋市登録要約筆者の会編 (2007)
- [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)