

議論スキル向上のための発言の自動評価とその効果的なフィードバック

岡本 康佑† 松原 茂樹†† 長尾 確†††

名古屋大学大学院情報科学研究科† 名古屋大学情報連携統括本部††

名古屋大学大学院情報学研究科†††

1. はじめに

大学のセミナーなどの会議において、参加者が互いの発言を評価し発言者にフィードバックすることは、発言者の議論能力の向上に役立つ。しかし、会議中に他者の発言を評価することは容易ではなく、評価者の負担も大きい。

これに対して著者らは、会議中の発言を自動評価し、結果を発言者に提示するシステムを開発している[1]。このシステムでは、評価を発言単位で参加者に提示する。しかし、参加者の発言の特徴は考慮されず、参加者はどの発言スキルを改善すればよいのか分かり難い。このため、参加者の発言全体を分析して、フィードバックする手法が求められる。

分析結果をユーザにフィードバックする手法として、中村らは、事前に設定したルールにより、提示する情報を選択する手法を提案している[2]。また小松らは、特徴を複数検出し、それらを選択的に組み合わせてユーザに提示する手法を提案している[3]。情報の提示方法はいずれの手法も、定型文やテンプレートを用いている。テンプレートを用いることで、ユーザに分かりやすく情報を伝えることができる一方、ユーザに提示する文章などは画一的になりやすい。

本稿では、ユーザに提示する情報を表す文章を、その要素単位ごとに重み付けすることで柔軟に生成する手法を提案する。

2. 発言の自動評価

発言の自動評価システムは、会議における参加者の発言スキルを評価する[1]。会議中に部屋の中央に設置されたマイクから音声を収録し、その音声と会議の議事録に基づいて、参加者の発言の音響的特徴と言語的特徴を分析する。評価する発言スキルを表1に示す。

3. 発言の自動評価に対する総評の生成

総評生成では、参加者全体の発言の平均評価値からのずれの度合いをスコアとして利用する。スコアを比較し、よりスコアの高い要因を参加者が改善すべき発言スキルとみなして、それを

表1：評価する発言スキル

音響的特徴に基づく評価	言語的特徴に基づく評価
声の大きさ	発言の簡潔さ
声の抑揚	発言の一貫性
話の速さ	発言の主題性
発言の流暢さ	
発言のテンポ	

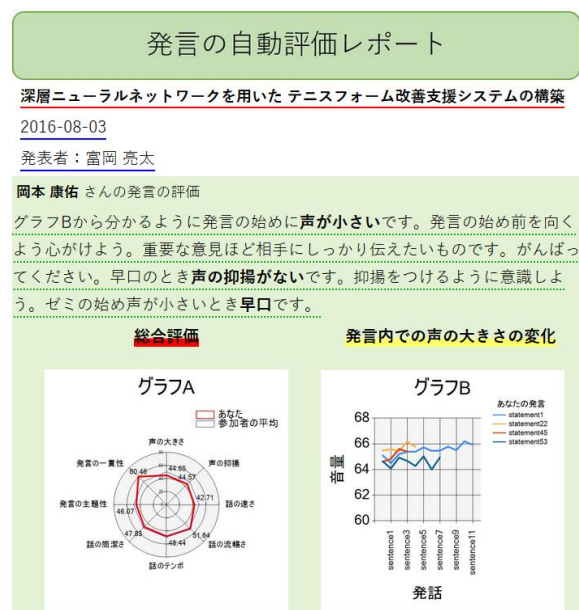


図1：総評の例

促すような総評を生成する。

3.1. 総評の構成

総評は、会議情報、評価を表す文章、および、グラフから構成される。総評の例を図1に示す。会議情報には、会議のタイトル、日時、発表者を記載する。また、総合的な評価、および、提示する文章に対応した個別評価を各々グラフで表示する。

3.2. 総評文章の生成手法

総評文章は、評価項目ごとに文を生成し、それらの文を組み合わせることによって作成する。このとき、それぞれの文要素に重み付けし、総評文の長さの制約のもと全体の重みが最大となる文章を生成する。

3.2.1. 評価項目の細分化

発言の評価の概要は総合評価のグラフで確認できる。しかし、参加者の発言スキル向上のためには、より具体的な要因を提示する必要がある。そのために評価項目を細分化する。具体的

Automatic Evaluation of Statements and Its Effective Feedback for Improving Personal Discussion Skills
 Kosuke Okamoto† Shigeki Matsubara†† Katashi Nagao†††
 † Graduate School of Information Science, Nagoya University
 †† Information and Communications, Nagoya University
 ††† Graduate School of Informatics, Nagoya University

表2：文の種類ごとの文要素

文の種類	文要素
事実文	根拠, 時間的情況, 他評価との関連性, 評価内容, 文末
アドバイス文	時間的情況, 他評価との関連性, アドバイス, 文末
意義文	意義
励起文	励起

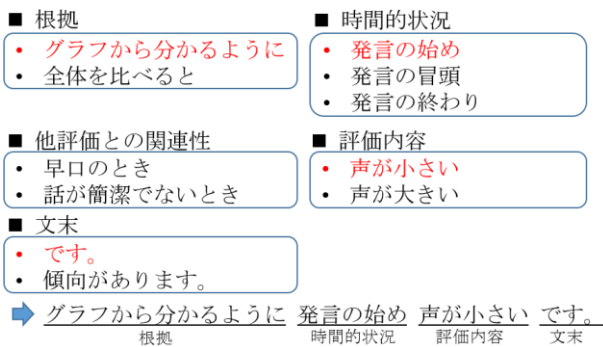


図2：文要素と選択

には、「発言の始めに声が小さい」などの時間的な細分化、「声が大さいとき抑揚がない」などの他評価との関連性を考慮した細分化を行う。細分化された表化項目の評価に基づき文章を生成する。

3.2.2. 総評文章の文の種類

総評の文章は表2に示す4種類から構成される。事実文は、評価項目と評価を記載し、現状を知らせる。アドバイス文は良くない評価項目に対する改善法を記し、発言スキルの向上を促す。意義文は、改善すべき評価指標の重要性を、励起文は改善への励みとなる文を記載することで、参加者の発言スキル改善を動機付ける。

3.2.3. 文要素

文はいくつかの文要素から構成される。事実文の生成例を図2に示す。図2では、根拠、時間的情況、評価、文末から赤字で示された文要素を一つずつ選択し、文を生成している。このときの文要素の選択は、文要素に対する重み付けと、総評文の最大長制約に基づいて行う。

3.2.4. 文要素への重み付けと文生成

総評の文章は、評価値が低い項目について述べるべきであり、そのような文要素の重みを大きくする。また、より評価項目との関連性の高い文要素の重みを同様に大きくする。同じ文要素が繰り返し提示された場合、総評の提示効果が薄れる可能性があるため、過去に提示したことがある文要素の重みを小さくする。

以上より、文要素 e_i の重み $c(e_i)$ は、評価項目と文要素との関連度 $R(e_i)$ 、過去の提示状態を示す値 $P(e_i)$ 、 e_i のスコア $S(e_i)$ を用いて以下のよ

表3：生成された総評へのアンケート結果

アンケート項目	ゼミ番号				合計
	1	2	3	4	
理解しやすいか	8/8	8/8	7/7	6/6	29/29
文章は自然か	5/3	5/3	6/7	5/6	21/29
納得するか	7/8	4/8	5/7	3/6	19/29
改善の意思があるか	8/8	8/8	7/7	6/6	29/29
参考になるか	6/8	6/8	3/7	3/6	18/29

うに表せる。

$$c(e_i) = S(e_i) \times R(e_i) \times P(e_i)$$

総評文の最大長を L_{max} として以下の整数計画問題の解 $\{x_i\}$ を文要素として選択する。

$$\max: \sum_i c(e_i)x_i$$

$$\text{subject to: } \sum_i l(e_i)x_i \leq L_{max}$$

$$x_i = \{0,1\}$$

ただし、 $l(e_i)$ は文要素 e_i の文字列長を示す。

4. 評価実験

著者の研究室で行われたゼミ4件で本システムを運用し、二択のアンケート評価を行った。アンケートの内容とそれに「はい」と回答された割合を表2に示す。「理解しやすさ」「文章の自然さ」「改善の促進」は「はい」の割合が高く、評価結果を正しく参加者にフィードバックできていること、および、総評の参加者による発言スキルの改善を促す効果があることを確認した。一方、参加者の総評に対する納得度は高くなく、信頼性が不十分な文章の生成も観察された。

5. おわりに

本稿では、会議における発言を自動評価し、総評を生成することで、参加者の発言スキルの向上を促す手法を提案し、実験によりその効果を確認した。今後、ユーザの納得度を向上させるための文生成法について検討する予定である。

参考文献

- [1] 岡本, 松原, 長尾, 会議における発言の音響的特徴と言語的特徴に基づく評価, 情報処理学会全国大会, pp. 521-522, 2016.
- [2] 中村, 志垣, 廣森, 山口, 東野, 大衆生活ノウハウの定量化とモデル化によるスマートライフ支援システム, 情報処理学会論文誌, Vol. 56, No. 8, pp. 1621-1633, 2015.
- [3] 小松, 西尾, 向井, 篠原, 電力消費を活用した省エネルギーアドバイスレポートの自動生成システム, 電気学会論文誌, Vol. 134-C, No. 9, pp. 1394-1405, 2014.