

事後アンケートによる放置発言の考察

A Consideration of Ignored Utterances Based on Post-Questionnaires

山口 広樹^{*1}
Hiroki Yamaguchi大澤 幸生^{*1}
Yukio Ohsawa西原陽子^{*1}
Yoko Nishihara^{*1} 東京大学大学院工学系研究科システム創成学専攻^{*1}
Dept. of Systems Innovation, School of Engineering, the Univ. of Tokyo.

According to subjective evaluation from results of questionnaires we consider features and the extraction method of Ignored Utterances(IUs) in a conversation. Here, evaluating “contribution” and “intensity of feelings” for each utterance, We regard IU, which have low “contribution” despite high “intensity of feelings”, as important utterance. Through online-chat experiments, we discuss the features of IUs and the validity of the extraction method.

1. はじめに

著者らは、放置発言に注目して発言の裏に潜む発言者の状況を掘り起こすことを目的とするによる会話解析手法[1]を提案しているが、[1]では解析事例は1例のみであり、抽出した放置発言を定量的に検証したものではない。本論では、事後アンケートを用いて抽出した発言を評価し、その特徴をまとめる。

本論では、放置発言の評価実験として、オンラインチャットを用いた。チャット実験では、共通の背景を持つ少人数のグループ(ここでは、同研究室に所属する大学生および大学院生)を被験者として、共通する関心事(ここでは研究室での生活)をテーマとした。これによって、日常生活で最も頻繁に行われていると考えられる、ある特定グループ内の会話における放置発言の特徴を検証する。

本チャット実験における事後アンケートは、各回のチャット後に被験者全員に対して行う。ここでは、評価項目として「話題貢献度」と「思い入れ」の2項目を設定し、それぞれ、抽出した放置発言の妥当性と重要性(本来注目されるべき発言であるかどうか)を検証する。本研究は、「会話内で埋もれてしまったシナリオを掘り起こす」ことを目的とする。また、こうして掘り起こされたシナリオは、集団の共有知や発言者の固有知を内包しているものも多く含まれることを実験によって明らかにし、放置発言の評価値の利用によって、重要発言を高スコアを持つ発言として抽出できることを示す。

また、放置発言に含まれる共有知や固有知を通じて、特定の集団の特徴の理解を助け、集団の所属者によるメタ認知[Suwa 08]を促進できる。

関連研究

本研究ではテキストマイニングの技術としてベクトル空間法[Salton 75]を採用し、cos類似度[Salton 83]を応用することで発言とその前後コンテキスト間の関係を算出している。

また、従来のテキストマイニング技術の応用分野は、要約[Hammouda 04]や文書クラスタリング[Leuski 01]を目的とした研究や、コミュニケーションを促進する場づくりの研究[Nishimoto 98]等が進められてきた。

しかしながら従来の研究では、会話の傾向等を解析することが主な目的となっており、会話の内容自体にまで踏み込んだ研究は未だ多くはない。本研究では、会話の内容にまで深く踏み込むことで、より実践的なシステムの実現を目指す。そのため、本論ではチャット実験での会話内容の一部を抜粋し、その文脈

を示した上で放置発言の考察を行う。

2. 放置発言の抽出手法と評価手法

本論では放置発言を“会話内で周囲の会話参加者から放置されている発言”と定義する。ここでは発言に含まれる単語のみを計算対象とし、各発言とその前後の会話との類似度を計算し、その類似度が低い発言を放置発言として抽出する。

2.1 抽出手法

まず、対象テキストデータを形態素解析[茶筌]し、そこで得られた単語群を(1)式の単語ベクトルに変換し、同様に各発言における出現単語をベクトルで表したものを(2)式の発言ベクトルで表す。ここで、 w_k は k 番目に出現した単語の出現頻度であり、 T は転置ベクトルを示す。

$$\text{単語ベクトル } W = T(w_1, w_2, \dots, w_n) \quad (1)$$

$$\text{発言ベクトル } U = T(w_1, w_2, \dots, w_n) \quad (2)$$

次に、生成した単語ベクトルと発言ベクトルを基に、発言と発言前後のコンテキストとの関連値を(3)式で算出する。本研究でのコンテキストの定義を文章の前後の脈絡とする。(3)式で、 i は注目発言の発言番号、 $W_{before-i}$ は注目発言前までの発言ベクトルの総和、 $W_{after-i}$ は注目発言後の発言ベクトルの総和、 w は計算範囲窓を示す。

$$\begin{aligned} I_{before-i} &= W_{before-i}^T \cdot U \\ I_{after-i} &= U^T \cdot W_{after-i} \end{aligned} \quad \left(\begin{array}{l} W_{before-i}^T = \sum_{i-w \leq k < i} U_k \\ W_{after-i}^T = \sum_{i < k \leq i+w} U_k \end{array} \right) \quad (3)$$

(3)式で求めた発言ベクトルと前後のコンテキストとの関連値から発言性質を推定する。(4)式に示す条件式を満たす発言、つまり前後のコンテキストとの関連値が共に低いものを放置発言として抽出する。ここで α, β は、それぞれ before 類似度、after 類似度の閾値であり、ユーザーが設定可能な変数ある。

$$I_{before-i} < \alpha \cap I_{after-i} < \beta \quad (4)$$

2.2 評価手法

ここでは、選択平均情報量という新たな情報量を定義し、それを用いて放置発言の評価を試みる。選択平均情報量は(5)式で定義される。

$$E_{AU_k} = \sum_{t_i \in U_k} p_{U_k t_i} I_{A t_i} \quad (5)$$

$$(I_{A t_i} = -\log p_{A t_i})$$

但し、 $\left. \begin{array}{l} t_i : i\text{番目の単語の出現頻度} \\ U_k : \text{注目発言ベクトル} \\ p_{U_k t_i} : U_k\text{における単語}t_i\text{の生起確率} \\ p_{A t_i} : \text{集合}A\text{における単語}t_i\text{の生起確率} \end{array} \right\}$

選択平均情報量は、注目発言が発言群に与えるインパクトを表す。ここでは類似度と同様に、(6)式のように before 情報量と after 情報量を定義する。但し、計算範囲窓 W は類似度算出時と同じ値を用いる。

$$E_{W_{before-k} U_k} = \sum_{t_i \in U_k} p_{U_k t_i} I_{W_{before-k} t_i} \quad (6)$$

$$E_{W_{after-k} U_k} = \sum_{t_i \in U_k} p_{U_k t_i} I_{W_{after-k} t_i}$$

(6)式の二つの情報量はそれぞれ、注目発言の前までの会話に与えたインパクトの程度と、注目発言の後に与えたインパクトの程度を表している。算出した二つの情報量を基に放置発言の性質を評価するために、(7)式のように評価関数を定義する。

$$\bar{E} = \left| E_{W_{before-k} U_k} - E_{W_{after-k} U_k} \right| \quad (7)$$

(7)式の評価関数によって算出された放置発言の評価値は、表1のような性質を持つ。

表1 評価値の性質

評価値 \bar{E} 大	コンテキストに多少影響を与えており、主張・物語が含まれる可能性が大きい放置発言
評価値 \bar{E} 小	コンテキストに影響を与えておらず、主張・物語が含まれる可能性が小さい放置発言

3. 事後アンケート

放置発言の評価として、事後アンケートを行った。本アンケートは、チャット参加者全員に対して実施し、会話ログからランダムに抽出した発言に対し、各「話題貢献度」と「思い入れ」の2項目について5段階評価を行った。各項目の定義を以下に挙げる。

【話題貢献度】

評価すべき発言が会話内で話題として取り上げられているかどうかの程度。

【思い入れ】

評価すべき発言内と言外(言いたくも言えなかったこと、敢えて言及する必要のなかったこと等)に含まれる思いの強さ。

ここで、各項目の評価基準は表2のように設定した。

表2 アンケート評価基準

話題貢献度	
1	放置された
2	ほとんど貢献なし
3	どちらともいえない。
4	多少話題に取り入れられた
5	話題に取り入れられた
思い入れ	
1	思い入れはない
2	大して思い入れはない
3	どちらとも言えない。分からない
4	多少思い入れがある
5	思い入れがある

事後アンケートは、次項のオンラインチャット実験で行う。ここでは各チャット後に、放置発言とその他の発言とをランダムに抽出して、評価用の発言セットを作成し、各発言について、前後の文脈を会話ログを参照にしながら評価を行うという形式でアンケート評価を行う。ここでは、アンケート用のHTMLファイルとCSVファイルを自動作成するプログラムを用いる。

4. オンラインチャット実験

明確な答えが存在しない参加者共通の関心事に関することをテーマとして与えてオンラインチャットを行った。日常会話でテーマとなるような簡単な話題をテーマとして与え、その会話ログを対象データとして解析し、さらに事後アンケートによって放置発言の性質評価を試みた。

アンケート結果の解析としては図1に示すようなグラフによる可視化を用いた、2軸(話題貢献度と思い入れ)のグラフ上に円グラフをマッピングし、円グラフには放置発言とその他の発言との比率を表示した。

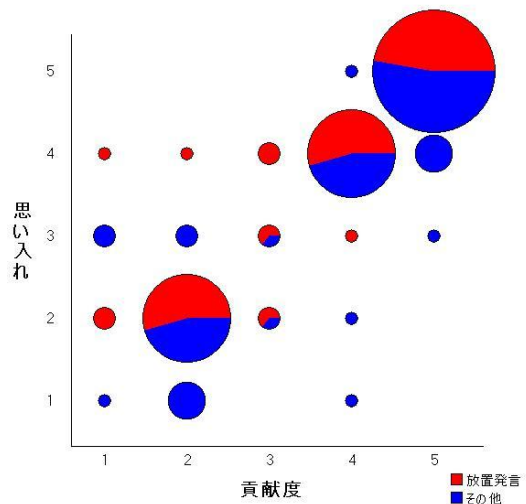


図1 アンケート可視化手法

この可視化によって、放置発言の 2 軸における傾向を見ることが出来る。図 1 は一回分のアンケート結果を可視化したものであるが、この場合は、赤い部分(放置発言)が図の上部に多く存在し、比較的思い入れが高いことが分かる。

5. 結果と考察

5.1 アンケート結果-可視化結果

オンラインチャット実験は全 7 回実施され、事後アンケートでは、延べ 20 名、全 77 発言、253 回答(内、被放置発言 112、その他発言 141)が得られた。全ての回答を可視化した結果を図 2 に示す。

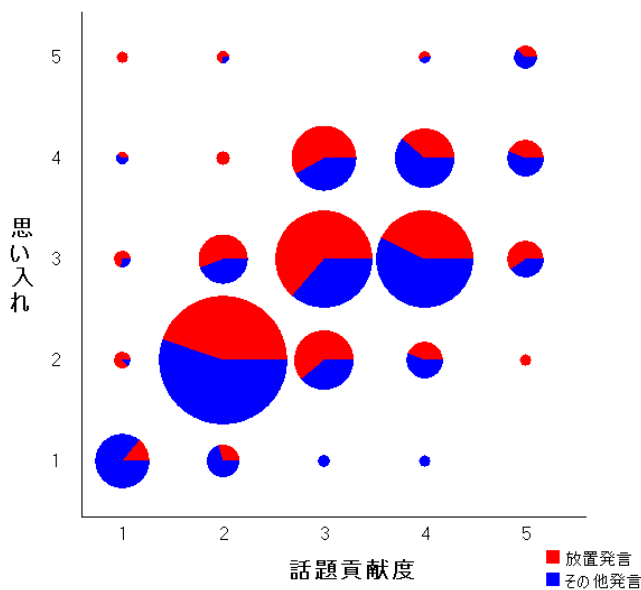


図 2 アンケート可視化結果

図 2 より、放置発言の評価を示す赤い部分は、図の上部に多く分布しており、思い入れが強いと評価された発言が多いことが分かる。放置発言とその他発言の 2 クラス間で t 検定を行ったところ、平均の差に $P < 0.05$ で有意差が認められた(表 3)。

その一方で左右の分布に注目すると、赤い領域は一様に分布していることが分かる。ここで、抽出されるべき放置発言は話題貢献度が低いものであると考えられるため、放置発言の抽出手法の妥当性として、今回の実験では良い結果が得られず、放置発言の抽出手法の妥当性は高いとはいえないことが分かった。

しかしながら、「会話内で埋もれてしまったシナリオを掘り起こす」という本研究の目的を再度見据えると、達成すべき目標は、「高精度に放置発言を抽出すること」ではなく、「抽出した放置発言を利用して、計算機で人の感性を扱う支援をすること」であるとなる。したがって、「話題貢献度」が低いと評価されるより、「思い入れ」が強いと評価される発言を抽出することが重要であるといえる。

本論での実験結果から、放置発言には会話参加者の「思い入れ」が強い発言が多く含まれることが示された。次節では、これらの発言を体系的に扱うべく、放置発言の特性で 3 種類に分類する。

5.2 放置発言の種類

実験で得られた 32 個の放置発言から、放置発言には次に示す 3 種類があるという仮説がたてられた。

- (1) 参加者間で共有されている認識に関する発言。「言わずもがな」な発言。共有知。
- (2) 発言者個人の特有な経験に基づいた、その他の参加者には理解が難しい内容の発言。固有知。
- (3) (周知であるないに関らず)単に事実のみに言及した発言で、それほど重要ではないと考えられる発言。

以下、各種類の例を挙げて 3 種類の放置発言を説明する。ここでは、3 つの放置発言をその前後の文脈と共に紹介する。

(1) 共有知

例 1 「代わりに MR. Y でお許しを」

```
-----データ引用1-----
Y「火器をつかうには申請が必要だとか」
Z「じゃあ報知機鳴ったらみんなでダッシュだな」
F「けど B は置いておきます」
Z「F 君も残ろう」
F「代わりに MR. Y でお許しを」
Z「Y「私めがのこののでございますか」」
-----データ引用ここまで-----
```

ここでは、会話には参加していない Y という人物が登場する発言が放置発言となっている。実験後のインタビューによると、この Y 氏は「潜在能力が高く期待できる人材だが、それを未だ発揮できていないように思われ、今後に期待している。」というように評価が参加者の共通の認識として存在し、Y 氏への関心の高さが感じられた。したがってこの放置発言は、単に放置されてしまった発言ではなく、参加者共通の認識のために、敢えて言及されることはなかったものであると考えられる。

(2) 固有知

例 2 「あの指向性スピーカー使いませんか?w」

```
-----データ引用1-----
K「忘年会でアニソン披露してもらおう！」
M「先生の評価があがるかもね」
K「うなぎのぼりですね、絶対w」
K「やはり Y はネ申」
M「ステージも用意してあげなきゃ」
S「あの指向性スピーカー使いませんか?w」
K「Pに Y ビーム」
M「それはいい。そして Y は伝説となる。」
-----データ引用ここまで-----
```

この放置発言では「指向性スピーカー」がキーワードとなっている。この発言はコンテキストを変化させてはいるが、明確にトピックとして取り上げられることはなかった発言である。

発言者の S 氏へのインタビューによると、「指向性スピーカーの有効な利用方法を模索していた」という私情がこの放置発言の背景に存在していることが分かった。

この発言の利用方法の一つとして、ここでは議論されなかった「指向性スピーカーの有効利用方法」に関して、改めて議論するきっかけとすることが考えられる。これによって従来よりも多くの知を結集し、画期的なアイデアが生み出される可能性を高めることができる。発言者の固有知を含む放置発言は、発言者以外には理解されていないために放置されてしまった発言で

あるため、本システムによってそれを再考する機会を創ることが可能となる。

(3) 単なる事実

例3 「どうりでまだ残ってる」

-----データ引用1-----

Ta 「時間かかっただけあって、昨日のおでんはおいしかったですね」

To 「うまかったけど、あまり量はたべれなかった」

To 「水分多いとすぐに腹一杯になってしまう」

N 「どうりでまだ残ってる」

To 「今日中に食べないと腐っちゃうよね>おでん」

-----データ引用ここまで-----

この放置発言は、単におでんがまだ残っているという事実を述べている発言であり、「残っているおでんを早く食べなければならぬ」という意思決定につながる発言ではある。しかしながら、仮にこの発言がなかったとしても同じ意思決定に至ると考えられるため、それほど重要な発言ではないといえる。

放置発言の評価値

実験で抽出した放置発言を前述した3種類に分類し、放置発言の重要度を測る評価値を検証する。評価値を種類別に検証した結果を表3に示す。

表3 放置発言の種類別評価値

発言種類	発言数	平均値	標準偏差
(1)共有知	11	0.49	0.30
(2)固有知	17	0.48	0.28
(3)単なる事実	4	0.16	0.15

表3から、種類(1)と(2)は(3)に比べて平均値・分散値ともに高いことが分かる。t検定を行ったところ、(1)と(3)、(2)と(3)において、平均値有意差が認められた。したがって、評価値を利用して、共有知や固有知を含む可能性の高い放置発言を重要発言として抽出できるといえる。

また、これらの放置発言から会話参加者の共有知や固有知を表出化することによって、特定の集団の特徴理解を支援し、その結果、集団の所属者のメタ認知を促進できるといえる。

6. おわりに

本論では、事後アンケートによる放置発言の性質の考察を行った。アンケート結果の可視化により、放置発言はその他の発言に比べて「思い入れ」が強いことが示され、本来注目されるべきである発言であるにも関わらず注目されなかった発言が多く存在することが分かった。

また、ここでは放置発言の3つの種類を明らかにし、重要な発言とそうではない発言とを計算処理で分類できることを示した。重要であると考えられる集団の共有知と発言者の固有知と、さほど重要ではないと考えられる単なる事実に言及した発言とを計算処理によって判別できる可能性を示した。

本システムによって放置発言から得られる共有知や固有知を用いることで、特定の集団の特徴理解を支援することができ、集団の所属者のメタ認知の促進につながる。さらに、特に注目されることのなかった発言者の固有知に注目する機会を提供

することで、画期的なアイデアが創られる可能性を高めることができる。

参考文献

[Barzilay 99]Regina Barzilay and M. Elhadad, Using lexical chains for text summarization, Advances in Automatic Text Summarization, pp.1-12, The MIT Press, London, 1999.

[Hammouda 04]Hammouda K.M. and Kamel M.S., Efficient Phrse-Based Document Indexing for Web Document Clustering, IEEE Trans. on Knowlege and Data Engineering, Vol.16, Issue 10, pp.1279-1296, 2004.

[Leuski 01]Anton Leuski, Evaluating Document Clustering for Interactive Information Retrieval, Proc. 2001 ACM International Conference on Information and Knowledge Management, pp.33-40. Atlanta, Georgia, USA, Nov. 2001.

[Nishimoto 98]Kazushi Nishimoto, Yasuyuki Sumi and Kenji Mase, Enhancement of Creative Aspects of a Daily Conversation with a Topic Development Agent, Coordination of Technology for Creative Applications, Splinger Berlin/Heidelberg, Vol.1364, pp.63-76, 1998.

[Salton 75]G. Salton, A. Wang, C. S. Yang, A Vector Model for Automatic Indexing, Communication of the ACM, Vol.18, No.11, pp.613-620, 1975.

[Salton 83] G. Salton and M.J. McGill, Introduction to Modern Information Retrieval, McGraw-Hill, 1983.

[Suwa 08] Masaki Suwa, A Cognitive Model of Acquiring Embodied Expertise Through Meta-cognitive Verbalization, Inormation and Medea Technologies, Vol. 3, No. 2, p.399-408, 2008.

[茶筌] 松本裕治, 北内啓, 山下達雄, 平野善隆, 松田寛, 高岡一馬, 浅原正幸, 日本語形態素解析システム『茶筌』version2.2.1 使用説明書, 2000.

[山口 2009] 山口広樹, 大澤幸生, 西原陽子, 患者インタビューの解析～患者の心を知る技術～, 第23回人工知能学会全国大会, 香川, 2009年6月.