

例外発言に潜む新価値

山口広樹, 西原陽子, 大澤幸生

東京大学大学院工学系研究科

〒113-0032 東京都文京区本郷 7-3-1 工 8-514

E-mail : {gutsu, nishihara, ohsawa}@panda.q.t.u-tokyo.ac.jp

Abstract: 本研究では, 潜在的価値を有しているにも関わらず一般に排除されがちな例外人物である「あほ」の価値を表出化するためのコーディネート支援することを目的とする「あほ学」の一つとして, テキスト処理によって例外発言である「あほ発言」を抽出し、「あほ発言」から新価値を掘り起こす支援をすることを目的とする。

1. はじめに

近年科学技術はめまぐるしい発展を遂げている。その成果物であるロボットなどの高性能で便利な技術にばかり頼り, 我々人間は本来の能力を使うことが少なくなってきた。科学技術が出力する正確な結果ばかりを信じ, 例外は多くの人々に無価値であると判断されて排除されてしまう。このように排除されてしまった例外発言を「あほ発言」とする。こうした「あほ発言」の中には, 他人と価値観が異なっているという理由で排除されていて, 既存の価値観を壊して新しい価値観を創るきっかけとなる可能性が含まれている。既存の価値観を壊すことはイノベーションを引き起こす上で絶対不可欠であるため, 「あほ発言」は非常に有益な価値を生み出す可能性がある。しかし, 「あほ発言」は排除されがちであるため, その潜在的価値を表出化するためには, その価値を発見して利用することを支援する必要がある。本研究では会話ログをテキスト処理することでそうした発言を抽出して, その潜在的価値を表出化する。

例外に注目した関連研究としては, 例外を利用してプログラムの解析やテストを行うもの[1][2]などがあるが, いずれも自然言語による会話は解析対象としていない。本研究では, 構造が無秩序で流動的であるが, 率直な意見や主張が含まれると考えられる自然言語による会話を解析の対象とする。人の生活の核となる日常会話を対象とすることで, 日常会話からイノベーションを引き起こすことも可能となると考える。

会話から例外人物を抽出するためには会話コンテキストの抽出が必要となる。コンテキスト理解の手法としては, 形態素解析[3]によって得られた単語群をベクトルで表すベクトル空間法[4]がある。ベクトル空間法は文書の構造などにロバストであると考えられ, 本研究ではこれを採用した。また, コンテキスト抽出のための手法としてキーワード抽出法があるが, 従来研究におけるキーワード抽出法には, 繰り返し言及される語は重要な概念を表すと仮定して出現頻度の高い語をキーワードにする方法[5]や活性伝搬法を利用した方法[6]や語の共起の統計情報に基づくもの[7]などがあるが, これらのキーワード抽出法は自然言語による会話から例外を抽出する手法ではない。本研究では, 会話コンテキストの変化を評価することで例外を抽出し, その価値を掘り起こすプロセスの支援を行う。ここでは, 各発言と会話コンテキストとの関連を, 注目発言と注目発言の前後に使用

された単語群との関連値を評価することで評価し、それを基に発言性質を推定する。こうして得られた発言性質を可視化することで発言性質が容易に理解可能となる。これによって排除されてしまった発言を抽出して、埋もれてしまった例外発言の潜在的価値を掘り起こす支援をする。

また本研究では、計算機による情報処理と人の観経験知や暗黙知などによる直に基づいた感性を交えた処理とを組み合わせるという2重らせんモデル[8]を用いて「あほ発言」のコーディネートを行う。閃きは人固有の能力であるため、人の閃きを計算機による情報処理で支援する。提案手法の評価として2002年度グッドデザインアワード[9]を受賞したデザインを利用して、チャットによる実験を行った。

2. 「あほ発言」とは

2.1 「あほ発言」の抽出方法

提案手法では会話テキストを入力とし、各発言と前後のコンテキストとの関連を評価することで排除されてしまっている例外発言を抽出する。以下では「あほ発言」をすると考えられる人物「あほ」の定義をする。

2.2 「あほ」の特徴と効用仮説

「あほ」を定量的に評価して利用するために、「あほ」の特徴を以下に挙げる。

- 特徴1. 無知・浅学による著しい認識のズレ。
- 特徴2. K・Y（空気読めない 空気読みすぎ-自己主張ができない-）/自己中心的。
- 特徴3. 価値観が常人と著しく異なる。変わり者。

筆者らは以上を「あほ」の特徴として挙げる。本研究では上述の特徴から、「あほ」の効用仮説を立ててそれを実証するというアプローチをする。「あほ」の特徴に基づく効用仮説を以下に示す。

- 効用仮説1. 「あほ」には人を惹きつける魅力があるため、会話コンテキストを変化させるファリテータとなる。
- 効用仮説2. 新たな価値観を導入し、思いもよらない発想の転換を起こさせる。
- 効用仮説3. あほの言動は理解に苦しむ。それ故に周囲の人間に熟考を促し、それがイノベーションにつながる
- 効用仮説4. 「あほ」は予期せぬ行動をするため、周囲に危機感を持たせリスクに対する意識を高める。
- 効用仮説5. 周囲の人に説明させることで、周囲の人の理解が深まりメタ認知を引き起こす。周囲の人間の説明能力向上を促進する。
- 効用仮説6. 突拍子もない発言で場を沸かせて議論を活性化する。または、場を和ませて周囲の人をリラックスさせる。

本研究では、効用仮説 1,2,3 に焦点を当てて例外発言を利用して効用仮説を実証し、さらにその効用をより効果的に作用させる支援をする。

2.3 「あほ発言」の抽出

本稿では、「あほ発言」を抽出するための手法として、会話のコンテキストと各発言との関連値を算出する。発言とコンテキストとの関連値を評価するために会話データを定量化する。まず、形態素解析によって得られた出現単語群を(1)式の単語ベクトルで表し、同様に一つの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 は転置である。

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

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

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は文章の前後の脈絡とする。(3)式で、 W_{before} は注目発言前までの発言ベクトルの総和であり W_{after} は注目発言後の発言ベクトルの総和である。

$$\begin{aligned} I_{before} &= W_{before}^T \cdot U \\ I_{after} &= U^T \cdot W_{after} \end{aligned} \quad (3)$$

(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。発言性質は以下のように分類する。

発言性質1.	I_{before} と I_{after} とともに小	新価値を導入したが排除されている。
発言性質2.	I_{before} と I_{after} とともに大	会話の流れに乗っているだけである。
発言性質3.	$I_{before} \ll I_{after}$	新たな話題・価値観を導入し受け入れられている。
発言性質4.	$I_{before} \gg I_{after}$	あるトピックの結論となっている。

「あほ発言」は周囲に排除されてしまっている発言なので、発言性質 1 であると推定された発言であると考えられる。さらに発言に対する聞き手の反応として、会話参加者毎に発言ベクトルとの関連値を算出することで、「発言に強く反応している」、「発言に反応せず聞いていない」などの発言者への対応を評価する。これによって、どの発言が排除されているか明確になる。

3. 発言性質の可視化

2.2 で行ったコンテスト変化の評価による発言性質の抽出結果を図 1 に示すように可視化する。発言性質の可視化には、図 1 の右部に示すようにスピーカーやヘッドホンを使って会話参加者の状態が容易に理解できるようにした。ここでは、発言者のみではなく発言に対する聞き手の反応も同時に可視化する。

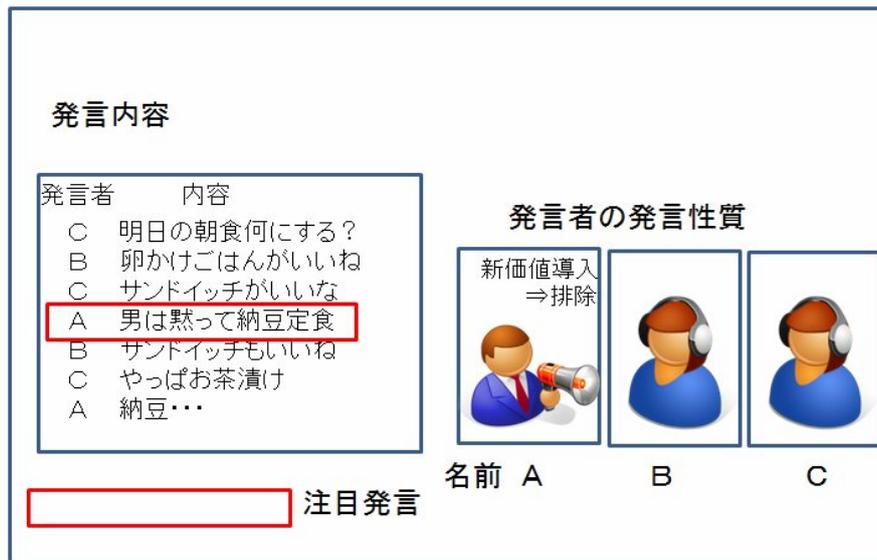


図 1：発言性質の可視化

4. 「あほ発言」のコーディネート方法

提案システムを用いて「あほ発言」をコーディネートする方法を説明する。提案システムはオペレータ（コーディネータ）に使用してもらい、コンテキストの理解はオペレータの主観で行う。オペレータは潜在的価値がありそうな発言を発見したらそれを次の議題に挙げる。これを繰り返すことで新たな価値を創造していく。このように、本システムの使用は計算機による処理と人による処理とを交互に繰り返すことで「あほ発言」をコーディネートしていくことが可能となると考えられる。

5. オンラインチャットを用いた実験

5.1 実験概要

本研究では、問題解決型のテーマを与えてそのオンラインチャット実験を行った。解のないテーマを与えて解を模索していく上での「あほ」の効用に注目した。

チャットに参加する被験者は大学院生と社会人で 1 回につき 3～5 人とし、時間は 30 分程度とした。今回の実験では、ある 2 つのデザインからアイデアを創発するというテーマでチャットを行った。実験に利用した 2 つのデザインには、2002 年度グッドデザインアワード[7]を新領域デザイン部門で受賞した「2001 年充電の旅」と「医療事故市民オンブズマン・メディオ」を利用した。

5.2 実験結果・考察

チャット終了後に本手法を用いて発言の評価・可視化を行った。発言数は 100 で抽出された例外発言は 18 であった。抽出された例外発言の一部を表 1 に載せる。表 1 から分かるように、返事などの一単語の短い文も「あほ」発言として抽出されてしまっているため、コーディネータはこういった新価値が含まれていないであろう発言を除いて「あほ発言」を探さなければならない。発言に含まれる単語数も考慮できるようにアルゴリズムを改良する必要がある。

実験では、潜在的価値がありそうな発言として、発言者Bの「2008年医療事故市民オンブズマン・メディオの旅」という発言に注目した。このアイデアは2つのデザインのタイトルを組み合わせただけという安易なものと捉えられたため、「あほ発言」であり無価値であると見做されて排除されてしまったと考えられる。本手法によって抽出したこの発言をチャット終了後に再検討することで、「旅をしながら医療事故への関心を直接的なコミュニケーションで広めていく」といった新たなデザインコンセプトを生み出すことができた。これは、「あほ発言」として埋もれてしまった例外発言に潜んでいた新価値を掘り起こすことができたといえる。但し、抽出された例外発言の中には再検討をしてもやはり無価値であろうと判断される発言も多いため、コーディネータの磨かれたセンスが必要であり、提案システムはそのセンスを発揮するための支援をするシステムである。つまり、計算機のみで新価値を発見することはできないため、計算機はツールとして発想支援をするものであり、実際に新価値を発見することは人の閃きによるものである。実験結果から、排除されてしまった発言を本手法によって抽出して再検討することで、本来であれば埋もれてしまった価値を掘り起こす支援ができることが分かった。

表1 実験結果 抽出された例外発言

発言番号	発言者	発言内容	I_{before}	I_{after}
3	A	はい	0	0
13	A	1イモ?	1	4
19	C	了解	0	1
21	B	二刀流です	3	5
44	A	E V?	0	1
53	A	やっぱこれだな	4	6
56	C	了解	1	0
68	B	2008年医療事故市民オンブズマン・メディオの旅	5	7
69	A	無理やりだな	0	1
78	B	オンブズマンは	5	1

6. おわりに

本稿では、一般的には排除されてしまいがちな例外人物「あほ」に注目してその潜在的価値を見出すことで新価値を創造することを目的とした「あほ学」の一つとして、「あほ発言」からの価値創造を支援するためのシステムを提案した。提案システムでは、会話における発言の性質の評価に基づく視覚化によって埋もれてしまった価値を発見して、オペレータがその価値を検討して再度議論の場を設けることで吟味していった。つまり、計算機による情報処理と人による感性を交えた処理とを交互に行うことによって「あほ発言」の持つ潜在的価値を掘り起こす。「あほ発言」の発見には、計算機による処理だけでなく人の感性による処理を交えることが不可欠であると考えられる。実験の結果、システムによって例外発言を抽出し、新たなアイデアを創ることができた。しかし、返事などの一単語のみの発言も抽出されてしまっていたので、単語数を考慮できるようにアルゴリズムを改良する必要がある。

また、会話における発話データのみでは、省略や隠語の使用等によって解析に必要な情報が不足していると考えられる。今後の課題として、このような不足していると考えられる情報を補完していく。また、「あほ」がいる場合といない場合を比較することで、チームパフォーマンスへの「あほ」の影響についての検証をしていき、同時に提案手法を利用することで掘り起こした新価値の評価を行っていく。

参考文献

- [1]Saurabh Sinha, Mary J. Harrold, Analysis and Testing of Programs with Exception Handling Constructs, IEEE TRANSSACTION ON SOFTWARE ENGINEERING, Vol.26 No.9, pp.849-871, 2000.
- [2]Liskov.B.H, Snvder.A, Exception Handling CLU, IEEE TRANSSACTION ON SOFTWARE ENGINEERING, Volume SE-5, Issue:6, pp.846-558, 1979.
- [3] 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.
- [4] Salton. G, A. Wang, C. S. Yang, A Vector Model for Automatic Indexing, Communication of the ACM, Vol.18, No.11, pp.613-620, 1975.
- [5] H. P. Luhn, A Statistical Approach to the Mechanized Encoding and Searching of Literary Information, IBM Journal of Research and Development, Vol.1, No.4, pp.709-317, 1957.
- [6]松村真宏, 大澤幸生, 石塚満: 語の活性度に基づくキーワード抽出法, 人工知能学会論文誌 第17巻4号, pp.398-406, 2002.
- [7]松尾豊, 石塚満, 語の共起の統計情報に基づく文書からのキーワード抽出アルゴリズム, 人工知能学会論文誌, 第17巻3号, pp.217, 2002.
- [8]大澤幸生, チャンス発見の情報技術, 東京電機大学出版局, 2003.
- [9]グッドデザインアワード, <http://www.g-mark.org/>.