

例外発言の価値定量化の試み

Trial for Value Quantification from Exceptional Utterances

東京大学大学院工学系研究科

システム量子工学専攻

学籍番号 37-066593

山口広樹

指導教員：大澤幸生 准教授

目次

1	序論	5
1.1	はじめに	5
1.2	論文の構成	8
2	関連研究	9
2.1	ベクトル空間法	9
2.1.1	文書のベクトル化	9
2.1.2	文書間類似度の算出	10
2.2	チャンス発見の二重螺旋プロセス	11
2.3	語の活性度に基づくキーワード抽出法	12
3	提案手法	13
3.1	アプローチ	13
3.1.1	計算機による処理のアプローチ	14
3.1.2	テキスト処理部	15
3.2	入力：会話テキスト 前処理	16
3.2.1	形態素解析	16
3.2.2	類義語処理	16
3.2.3	ユニークタームベクトルの生成	17
3.2.4	発言ベクトルの生成	17
3.3	例外発言とは 定義と抽出手法	18
3.3.1	例外発言の定義	18
3.3.2	例外発言の抽出	18
3.3.3	抽出した例外発言の検証	19
3.4	例外発言の評価手法	20
3.4.1	抽出した例外発言の問題点	20
3.4.2	選択平均情報量	20
3.4.3	選択平均情報量による例外発言の評価	21
3.4.4	選択平均情報量の性質検証	23

3.5	出力：例外発言	26
4	提案システム利用方法	27
4.1	提案システムの利用プロセス	27
4.2	パラメータの設定方法の検証	27
4.2.1	計算範囲の窓 w の値検証	28
4.2.2	閾値 α, β の値検証	32
5	予備実験 オンラインチャット実験	36
5.1	チャットテーマ1：新ゲームの考案	37
5.2	チャットテーマ2：食事をどうするか	38
5.3	チャットテーマ3：簡易型イノベーションゲーム	39
5.4	考察	40
6	提案システムの評価実験	42
6.1	地方自治体の定例会会議録の解析	42
6.1.1	宮崎県定例会	43
6.1.2	大阪府定例会	44
6.1.3	会議録解析の考察	45
6.2	患者インタビューの解析	46
6.2.1	インタビューデータ概要	46
6.2.2	解析環境の設定	46
6.2.3	解析結果	47
6.2.4	解析によって得られた知見	48
6.2.5	他手法との比較	55
6.2.6	コンテキスト変化点抽出の検証	58
6.3	解析結果の評価	60
7	結論	61
	謝辞	62
	参考文献	65

付録	70
A オンラインチャット実験 会話ログ	70
A.1 新ゲーム考案	70
A.2 食事をどうするか	74
A.3 簡易型イノベーションゲーム	77
B 患者インタビューデータ (冒頭部分)	81

目次

1	ベクトル空間 (三次元の場合)	10
2	チャンス発見の二重螺旋プロセス	11
3	ツール起動画面	13
4	アプローチ	14
5	テキスト処理部フローチャート	15
6	会話テキスト-前処理	16
7	選択平均情報量の概念図	20
8	before 情報量	22
9	after 情報量	22
10	before 類似度-評価値 E	24
11	before 類似度-評価値 E (例外発言のみ)	25
12	解析結果出力画面	26
13	システム利用プロセス概要	27
14	組み合わせ発想ゲーム 使用カード	78

表 目 次

1	類義語の例	16
2	抽出した例外発言（一部抜粋）	19
3	評価関数の検証	24
4	例外発言に対する評価値の特徴	25
5	解析結果出力項目	26
6	パラメータ 窓 w の検証 A.1 $w = 10$	29
7	パラメータ 窓 w の検証 A.1 $w = 20$	30
8	パラメータ 窓 w の検証 A.1 $w = 30$	30
9	パラメータ 窓 w の検証 A.2 $w = 10$	30
10	パラメータ 窓 w の検証 A.2 $w = 20$	30
11	パラメータ 窓 w の検証 A.2 $w = 30$	31
12	パラメータ 窓 w の検証 A.3 $w = 10$	31
13	パラメータ 窓 w の検証 A.3 $w = 20$	31
14	パラメータ 窓 w の検証 A.3 $w = 30$	31
15	パラメータ 閾値 α, β の検証 A.1 $\alpha = 0.3, \beta = 0.3$	33
16	パラメータ 閾値 α, β の検証 A.1 $\alpha = 0.2, \beta = 0.3$	33
17	パラメータ 閾値 α, β の検証 A.1 $\alpha = 0.3, \beta = 0.2$	34
18	パラメータ 閾値 α, β の検証 A.2 $\alpha = 0.3, \beta = 0.3$	34
19	パラメータ 閾値 α, β の検証 A.2 $\alpha = 0.2, \beta = 0.3$	34
20	パラメータ 閾値 α, β の検証 A.2 $\alpha = 0.3, \beta = 0.2$	35
21	パラメータ 閾値 α, β の検証 A.3 $\alpha = 0.3, \beta = 0.3$	35
22	パラメータ 閾値 α, β の検証 A.3 $\alpha = 0.2, \beta = 0.3$	35
23	パラメータ 閾値 α, β の検証 A.3 $\alpha = 0.3, \beta = 0.2$	35
24	匿名性による違いの検証 A.2	41
25	匿名性による違いの検証 A.3	41
26	患者インタビューの解析結果	47
27	TFIDF 法 抽出されたキーワード	56
28	IDM 法 抽出されたキーワード	57
29	選択平均情報量のみでの解析結果	58

1 序論

1.1 はじめに

世の中の多くの人が、心の内には主義・主張を持って発言したにも関わらず誰にも相手にしてもらえなかったという経験があるのではないか。筆者はしばしばこういった経験をする。そして胸の内に秘めた主義・主張は解き放たれることなく行き場を失うことになってしまうのである。会話中に生じる感情や閃きは一過性が強く自分自身でさえうまく説明できないことも多く、なんとか言葉にできたとしてもその発言は会話コンテキストに合わないために周りの人に理解されずに排除されてしまうことが多い。会話中に周りに排除されてしまった発言を例外発言と呼ぶこととする。こうした例外発言の裏には、発言者の状況があり、人それぞれの物語が存在する。「例外発言の裏に存在する状況・物語は例外発言に潜在する価値であり、適切に対応することで掘り起こすべきである」という命題を実証することが本研究の目的とするところであり、本論ではその手始めとして例外発言の価値定量化を試み、例外発言からの価値マイニング手法確立のための第一歩としたい。

次に研究背景について述べる。周知の事実として近年、科学技術、特に情報技術が発展してきた。こうして発展してきた技術の恩恵として、人はコンピュータから正確な出力を得られるようになった。しかしその一方で、パーソナルコンピュータを代表とする情報機器への過度な依存や実空間におけるコミュニケーションの希薄化など新たな問題も生じてきていると報告されている [1][2]。また数年前の報告であるため現在では技術的問題の多くは解決されているが、教育における情報機器利用によって基礎学力の低下を招くと問題提起している報告も存在する [3]。

こうした問題の代表例として携帯電話が挙げられる [4][5]。携帯電話の普及によって、電波さえ届けば電話や電子メールで他人とコミュニケーションがとれるのは大変便利であり、もはや生活必需品となっている。さらに、最近の携帯電話には内蔵カメラを代表として多くの機能が搭載されており携帯電話の使用用途が飛躍的に拡大してきた。その一方で、携帯電話の電話帳機能に頼りすぎて携帯電話がないと知人の連絡先が全く分からなくなってしまうなど、いくつかの弊害を引き起こしている。電話帳機能に関しては、知人の電話番号を頭で記憶しておく必要がなくなった

ために、人間の記憶力の低下を招く結果になってしまうのではないかという危惧も生じてくる。また、携帯電話やパーソナルコンピュータ、その他コンピュータが埋め込まれた機器は膨大な数が存在するが、現在の技術では人の感情を満身に認識・表現することが不可能であり、人工知能の分野でも限界があるとされて久しい[6][7]。

人は感情を代表とするようにコンピュータで扱うことが困難な不確定性を内包している。感情を扱う学問として、心理学では一般的に、感情を急激に表れるものと持続するものに分けて考えられている[8]。そうした不確定性を工学的アプローチで扱えるようにしようと、感性工学[9][10]や”Web Intelligence”としてWebとインターネットに人工知能と情報技術(Information Technology, IT)を活用したもの[11][12]として扱うなど工学の分野においても様々な観点から研究されており、本研究もそうした研究の一つであるといえる。ただし、従来の研究とは大きく性質を異にしている。従来の感性を扱った研究では、人の感性を定量化したり[13]、感性情報を利用して文書や画像を分類したり[14]しているものが主流であるのに対し、本論のアプローチでは、感性は人の手によって直接扱い計算機はその計算能力によって人の支援をするという、人と計算機とのインタラクションを通じて解析を行っていく。本論で提案する手法では、計算機は自然言語処理によって重要であると推測される発言を出力することにより、人による感性マイニングの支援をする。

また、本論の最も重要な特徴は例外発言に注目している点である。一見無意味で愚かであると思われる発言には、その真価が未だ発揮されていない主義・主張が込められているとし、そうした発言に焦点をあてることによって埋もれてしまった価値をマイニングすることを目的としている。

例外に注目した関連研究として、例外に注目して知識を得ることを目的とする研究は多数存在し、情報量を用いてデータベースから例外的知識を抽出するもの[15]や例外から得られる情報を基にソフトウェア・プログラムの解析やテストをするもの[16][17][18]などがあり、特にソフトウェアの開発においては必要不可欠なものであるといえる。また、「人と異なる文化・生活習慣をもっている」、「孤立する or 多くの知人を持つ」、「時勢に敏感である」などの特性を持つ人物を”普通でない人(変わり者)”として、シミュレーションによって変わり者の有用性を検証している研究[19]や、アウトライヤー(外れ値)に注目して”なりすまし”を検出したり事前に排

除して統計学に役立てたりする研究 [20][21] も例外に注目した従来研究であるといえる。これらの例外に注目した従来研究ではプログラムやデータベースなどの構造的で主観が含まれないデータを解析対象としているが、本論では、構造は無秩序で流動的であるが率直な主義・主張が顕著に表れると考えられる自然言語による会話を解析対象としている。

また、本研究では人と人との相互作用による感情の揺らぎなど不確定性が高い感情を扱うため、工学的な側面だけでなく、社会行動や社会生活によって影響される社会的な課程における心理の研究を行う社会心理学 [22][23] や、人類の社会的・文化的側面を研究する社会人類学 [24]・文化人類学 [25] などとの関連も大きい分野横断的な研究である。さらに本研究は他の多くの分野への適用が可能であり、コミュニケーションが重要視されている医療・保健分野 [26][27] や教育 [28][29] などの分野へ応用できると考えられる。本論ではその可能性を広げるための第一歩として、地方自治体定例会の会議録と医療データ（患者インタビュー）の解析実験をしている。コミュニケーションは医療・保健分野でも大きく扱われており [30][31]、工学の分野でも医療コミュニケーションを分析・支援する研究 [32][33] が進められている。本研究の今後の展開として、医療・保健分野への応用を考えている。この他にも、システムのステークホルダ（利害関係者）からいかにしてシステム要求を引き出し、分析してまとめるかという技術の集大成である要求工学 [34][35] では、要求獲得を目的とした要求会議をテキストマイニングによって解析する手法 [36][37] も提案されており、こうしたソフトウェア開発現場への適用も可能であると考えられる。

本論文を読んで、「一見無意味で愚かであるがために無価値であると判断されてしまって排除されてしまっている人・発言にもこれからは注意を向けてみよう」と考えていただければ幸いである。

1.2 論文の構成

本論文の構成は、本章が序論、第二章では関連研究として本研究で利用したベクトル空間法とチャンス発見の二重螺旋プロセスの概要を説明し、第三章では提案手法の説明として入力データ、前処理、例外発言の定義と抽出手法、例外発言の評価手法、最後に出力方法について示し、第四章では提案システムの利用方法を示す。続いて第五章ではパラメータ設定方法と提案手法の妥当性確認のための予備実験について示し、第六章では地方自治体定例会会議録と患者インタビューの会話ログデータを対象データとして提案システムの評価実験をする。前処理である形態素解析と類義語処理、ユニークタームベクトルの生成方法、発言ベクトルの生成方法の説明をし、第四章では例外発言の定義と抽出手法を示し、第五章では例外発言の評価手法として提案する選択平均情報量の定義とその性質の検証をし、第六章ではシステムの出力方法を示し、第七章では提案システムの利用方法について論じ、第八章で予備実験としてのオンラインチャット実験の検証をし、第七章で結論を述べ、続いて謝辞、参考文献、最後に付録としてオンラインチャット実験データという構成となっている。

2 関連研究

本論の関連研究として、自然言語処理に関する先行研究を挙げる。自然言語解析では、用いられている単語群をベクトルで表して処理するベクトル空間法 [38] が 1975 年に発表され、自然言語処理の基礎技術として現在でも情報検索 [39] 等に広く利用されている。ベクトル空間法はシンプルな手法であり、それ故に汎用性が高いと考えられる。本研究では構造が無秩序かつ流動的である会話を解析対象とするため、汎用性の高いベクトル空間法を採用した。また従来のテキストマイニング技術には、文書の要約 [40][41] や構造解析 [42][43] を目的としたものが多く、解析結果をいかにして有効利用していくかという方法論を深く追究しているものは僅かである。

解析結果の利用方法までを深く追究している従来研究として、“KeyGraph[44]”を基盤技術として 2000 年に大澤らによって提唱されたチャンス発見学 [45] がある。チャンス発見学では、「チャンスとは意思決定を左右する重要な事象・状況、またはそれらに関する情報である」と定義した上で、人による処理とコンピュータによる処理とを交互に組み合わせて、主観データと環境データとを扱いながら問題の解決策を模索していくチャンス発見の二重螺旋モデル [45] を導入しており、本論のアプローチはこのプロセスに倣っている。またその他にも、人間機械協調システムにおける諸問題を「社会的知性」をキーワードとして新たな人間と機械の協調関係のパラダイムを探る研究 [46] や参加型シミュレーションシステムを利用して交渉モデルを構築する研究 [47]、「分身エージェント」を用いてネットワークコミュニティにおける知識共有を支援する研究 [48] なども進められている。本章では、本論で利用したベクトル空間法、チャンス発見の二重螺旋プロセス、比較手法として用いた活性伝搬法 [49] の概要を説明する。

2.1 ベクトル空間法

前述したように、本研究ではベクトル空間法 [38] を用いて解析を進めていく。本論に入る前に、本論で利用したベクトル空間法の概要について記述する。

2.1.1 文書のベクトル化

ベクトル空間法はその名の通り、対象とする文書をベクトルで表現する手法である。対象とする文書を D 、文書 D に含まれる総単語数を n 、各単語を $t_k (1 \leq k \leq n)$

として，(1) 式のように表わされる．

$$D = (t_1, t_2, \dots, t_n) \quad (1)$$

(1) 式のように文書中に出現する単語の頻度をベクトルとして表すことで，自然言語の文書を計算機で処理することができる．またベクトル化する際には，単語の意味や文書構造等は全く考慮しないため，機械的に文書の処理ができ，且シンプル計算コストも低いという利点がある．本論では，自然言語による会話を対象とするため構造が無秩序で流動的なテキストを処理する必要がある．この条件を満たす手法として，上述したようにシンプルで汎用性が高いと考えられるベクトル空間法を採用した．

2.1.2 文書間類似度の算出

ベクトル空間法は情報検索 (Information Retrieval : IR) のための技術として提案されており，IR のためには，文書間の類似度の算出が重要な要素となる．(1) 式によってベクトル化された 2 つの文書 D_1 ， D_2 間の類似度は (2) 式で表わされるようにベクトルの内積によって算出される．

$$Sim(D_1, D_2) = \frac{D_1 \cdot D_2}{|D_1||D_2|} \quad (2)$$

(2) 式によって算出される類似度は，単語の持つ意味や文書構造に関係なく，文書中の出現単語が似ていれば類似度が高いと見做される．また，ベクトルの内積をベクトルの大きさで割ることによって， $\cos \theta$ を求めてこれを類似度としていることから \cos 類似度と呼ばれ，図 1 はベクトル空間が三次元の場合の例であるが，感覚的な解釈としては，同じ方向を向いているベクトルであるほど類似度が高くなる (同一ベクトルであれば類似度は $\cos 0 = 1$ となり，これが最大値となる) ．

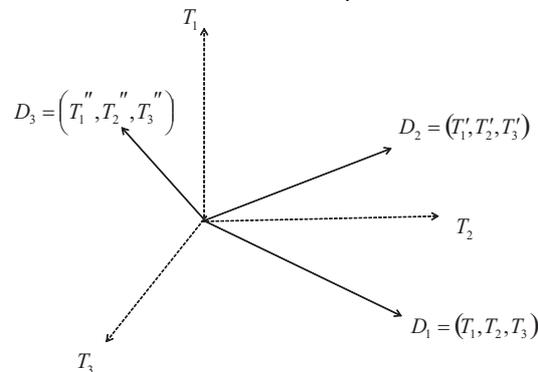


図 1: ベクトル空間 (三次元の場合)

2.2 チャンス発見の二重螺旋プロセス

本研究の解析は、チャンス発見の二重螺旋プロセス [45] に倣う。二重螺旋プロセスは、環境データ（観測結果）を入力として主観的評価によって主体データ（思考内容）を出力する人の螺旋と、主体データ又は環境データを入力として演算処理によって環境データ（解析結果）を出力するコンピュータの螺旋とを相互に実行していくプロセスであり、この相互作用によって意思決定において重要な要素となるチャンスを見つけていくというプロセスである。二重螺旋プロセスの概念図 2 に示す。

大澤らは、仏の細菌学者パスツールの名言として知られる”Chance favours the prepared mind.”という言葉が示しているように、隠れたチャンスを見出すためには、二重螺旋プロセスによってチャンスを受け入れられる心理状態を作っておく必要があるとしている。本論でもこの考えに倣って、人によるプロセスと計算機によるプロセスとを交互に行うことによって解析を進めていく。

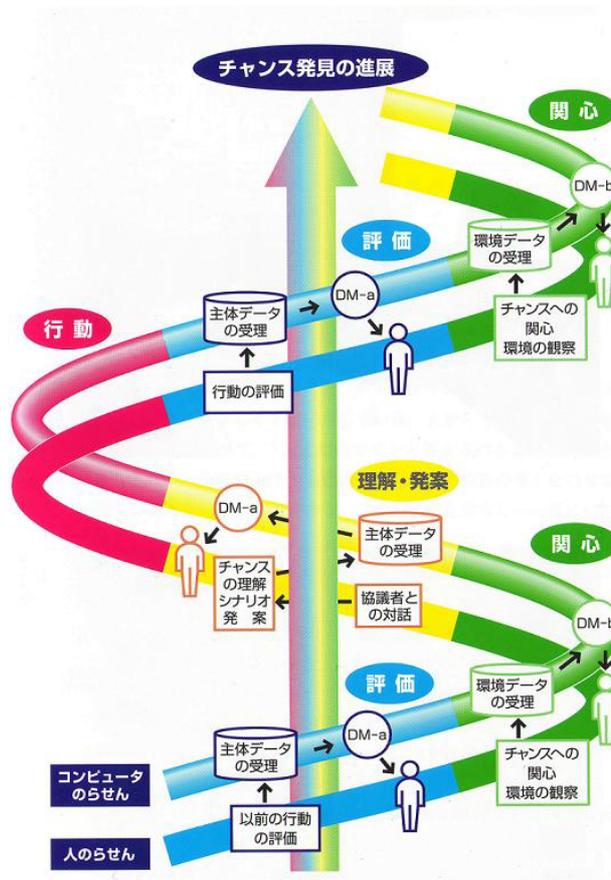


図 2: チャンス発見の二重螺旋プロセス
(画像出典：チャンス発見の情報技術[24] 口絵-1)

2.3 語の活性度に基づくキーワード抽出法

本論では活性伝搬法 (IDM 法) [49] を比較実験として利用する。IDM 法は、一文内での語の共起の統計情報を用い、大量のコーパスを用いずに単独の文書だけでキーワードを抽出する手法である。IDM 法では記憶のメカニズムを禁じた活性伝搬モデル [50][51] を利用してキーワード抽出する手法である。IDM 法で使われている活性伝搬モデルは (3) 式で表わされる。但し、 $A(t)$ は活性回数 t の語の活性値を表すベクトル、 C はネットワークに注入される活性値を表すベクトル I は $A(t-1)$ の活性値を $A(t)$ に伝搬させる単位行列、 R はネットワークの構造を表す伝搬行列であり、 R の i 行 j 列の要素 R_{ij} は語 w_i と語 w_j の関係の強さを表す (対角成分は 0)。また、 γ は活性値の減衰率を表すパラメータ、 α はネットワークが語の活性値に及ぼす影響力の程度を表す伝搬パラメータである。

$$A(t) = C + ((1 - \gamma)I + \alpha R)A(t - 1) \quad (3)$$

IDM 法は、ある語が想起 (活性化) されるとその語に関連する語も想起 (活性化) されるというプライミング効果 (Priming Effect) [52] を文書内の重要語を抽出する手法である。人の認知的側面を考慮して考案された手法であるため、発言の裏に潜む状況・物語を抽出することを目的とする本手法の比較手法として適している。

3 提案手法

本章では，提案手法の説明をする．まず，提案手法における基本アプローチを示し，その後，入力データ，前処理，例外発言の定義と抽出手法，例外発言の評価手法，最後に出力方法の説明を順にしていく．

3.1 アプローチ

提案システム概要

本論では，例外発言を定量化する方法としてテキスト処理技術を応用して会話ログを解析するツールを作成した．作成したツールの起動画面を図3に示す．

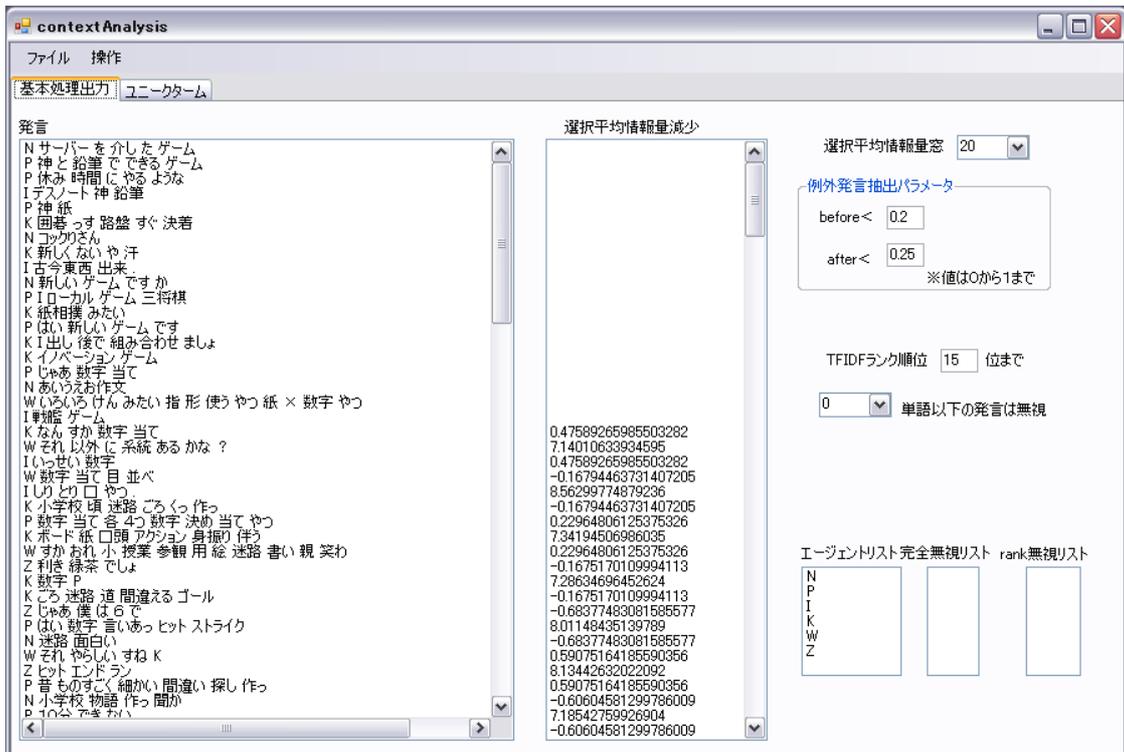


図 3: ツール起動画面

3.1.1 計算機による処理のアプローチ

本論で提案するシステムは二重螺旋プロセスを取り入れて人による処理と計算機による処理とを繰り返すものであるが，ここでは，計算機による処理のアプローチの概要を示す（図4）．また，システム利用方法に関しては4章で論じる．



図4: アプローチ

入力データとして会話ログを用いる．会話における例外発言に注目した理由は，構造が無秩序で流動的であるためコンピュータで扱うのは困難であるが，会話できる人であればだれでも解析対象とすることができると考えられるからである．本論では，3～5人程度の少人数による会話が最も適切であると考えている．多人数インタラクションからのアイデア創発のための環境として，少人数（3～5人）による会話は議論が活発化・収束しやすく建設的であると考えられる．なぜなら，二人の会話ではアイデア創発力に欠け，多人数の会話であると異なる話題が並行的に発生したり意見の食い違いが生じやすいからである．

しかし少人数でのインタラクションの解析に限定してしまうと汎用性が失われてしまう危険が生じることから，より汎用性の高いシステムを目指すために，提案システムは独話（日記等）から多人数による会議までに対応できるように設計する．

3.1.2 テキスト処理部

本研究の核となるテキスト処理部では，前処理として形態素解析，類義語処理，ユニークタームの生成を行った後，発言ベクトルの生成，例外発言の抽出と評価を行っている．図5にテキスト処理部のフローチャートを示す．

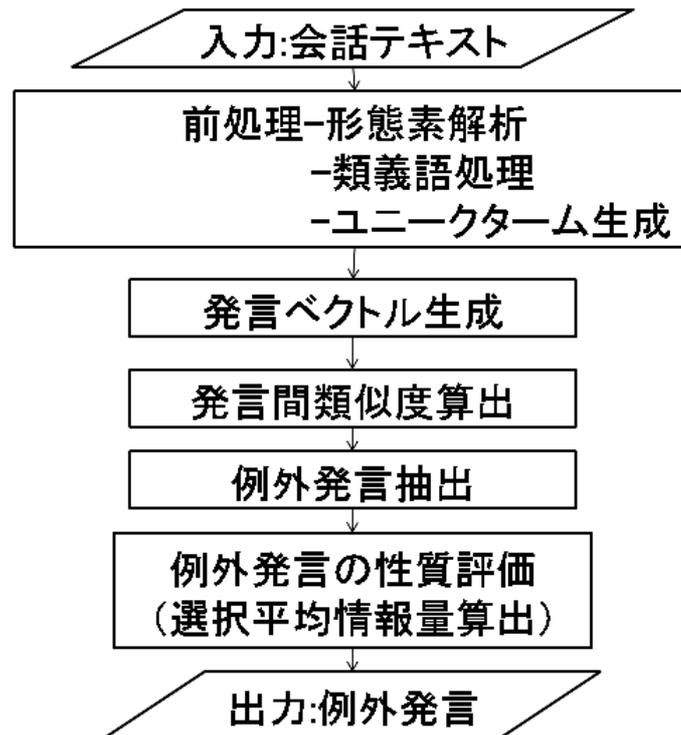


図 5: テキスト処理部フローチャート

テキスト処理部では，会話テキストを入力として例外発言を抽出・評価した後にランク付して出力する．ここでは，例外発言の評価のために選択平均情報量という新たな情報量を定義して評価関数に利用している．次章から各処理の詳細を記述する．

3.2 入力：会話テキスト 前処理

3.2.1 形態素解析

前項のベクトル空間法を適用するためには、会話内のテキストに含まれる単語を前処理で抽出する必要がある。さらに、日本語の文書を対象とする場合には、文を単語ごとに区切る形態素解析の処理が必要になる。本論では、フリーソフトウェアの形態素解析器『茶筌』[53]を用いて形態素解析を行った。入力となる会話テキストは、発言者名と発言内容を一対として一発言としている。図6に入力会話テキストとそれを形態素解析した結果の例を示す。

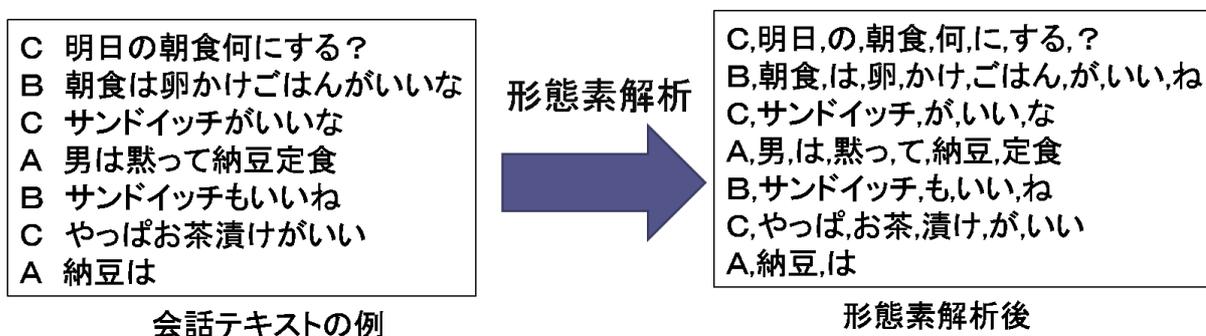


図 6: 会話テキスト-前処理

3.2.2 類義語処理

同義語・類義語を同じ単語であると認識させるために、類義語辞典を用意し、前処理として類義語処理を行った。これによって、異なる表現の使用による不具合を回避する。表??に類義語の例を示す。ここで、漢字表記とひらがな表記も異なる単語として処理されてしまうため、適宜登録しておく必要がある。

表 1: 類義語の例

何故, なぜ, 何で, なんで, どうして
感動, 感激, 感嘆, 感銘, 感心, 感慨
任務, 使命, 勤め, つとめ, 勤務, 役割, 役目
不評, 不評判, 悪評, 醜聞, 汚名, 烙印
文化, 文明, 人文, カルチャー

3.2.3 ユニークタームベクトルの生成

2.1章に記述した空間ベクトル法を利用するための前処理として、文書をベクトル化するための基となるユニークタームベクトルを生成する。ベクトルの各要素に一単語を割り当てて、文書のベクトルを生成する際に各要素の値に該当単語の出現頻度を格納する。簡単な例として、図6に示した短い会話のユニークタームベクトルを挙げる。ここでの出現単語は25種類であるため、生成されるベクトルは25次元となる。

$$\text{UniqueTerm} = \{ \text{明日, の, 朝食, 何, に, する, ?, は, 卵, かけ, ごはん, が, いい, ね, サンドイッチ, な, 男, 黙っ, て, 納豆, 定食, も, やっぱ, お茶, 漬け} \}$$

3.2.4 発言ベクトルの生成

先に生成したユニークタームベクトルに基づいて、各発言から発言ベクトルを生成する。図6に示した会話の発言ベクトルは以下のように生成される。

発言ベクトル, 発言者

$$\begin{aligned} U_0 &= \{1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0\}, C \\ U_1 &= \{0, 0, 1, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0\}, B \\ U_2 &= \{0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0\}, C \\ U_3 &= \{0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 0, 0, 0, 0\}, A \\ U_4 &= \{0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0\}, B \\ U_5 &= \{0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1\}, C \\ U_6 &= \{0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0\}, A \end{aligned}$$

3.3 例外発言とは 定義と抽出手法

3.3.1 例外発言の定義

広辞苑第五版によると，例外とは『通例の原則に当てはまらないこと．一般の原則の適用を受けないこと．また，そのもの』とされている．したがって例外発言の辞書的な意味は『通例の原則に当てはまらない発言』，または『一般の原則を受けない発言』となる．本稿では，通例（または一般）の原則を「会話が成立している」「話の流れが定まっている」として，例外発言を以下のように定義する．

- 前後の文脈とずれている，会話の流れに沿っていない排除されてしまった発言．

この定義に基づいてテキストから例外発言を自動抽出する．まず，前述のベクトル空間法によって，会話内の各発言を (1) 式でベクトル化する．その後，(2) 式の \cos 類似度を用いて，次式のように二つの類似度を算出する．

$$\textit{before} \text{ 類似度 } S_{\textit{before}\cdot n} = \textit{Sim}(U_k, W_{\textit{before}\cdot n}) \quad (4)$$

$$\textit{after} \text{ 類似度 } S_{\textit{after}\cdot n} = \textit{Sim}(U_k, W_{\textit{after}\cdot n}) \quad (5)$$

但し， n は注目発言の発言番号， t は転置行列， $W_{\textit{before}\cdot n}, W_{\textit{after}\cdot n}$ は，注目発言前後の w 個の発言のベクトル総和であり (6) 式で表わさせる．

$$W_{\textit{before}\cdot n} = \sum_{n-w \leq k < n} U_k, \quad W_{\textit{after}\cdot n} = \sum_{n < k \leq n+w} U_k \quad (6)$$

窓 w を設けることによって，計算コストを軽減し，且，会話における話題展開の速さの違いに対応できるよう，ユーザーがツール上で探索的に妥当な値を設定できる．

3.3.2 例外発言の抽出

前項で定義した二つの類似度を評価することで，例外発言の抽出を行う．ここでは，「前後の文脈とずれている」を「前後の発言では使用されていない単語を多く含んでいる」発言であると解釈し，(7) 式に示す条件式のように二つの類似度各々に閾値を設けることで抽出した．

$$S_{\textit{before}\cdot n} < \alpha \quad \cap \quad S_{\textit{after}\cdot n} < \beta \quad (7)$$

(7) 式を満たす発言が例外発言として抽出される．しかしこの抽出条件では，「全く意味が含まれないような発言」や「単純に短い発言」などの本研究の目的に合致しない発言が多く抽出されるという問題が生じてしまうと考えられ，本論ではこの問題の解決策として抽出した例外発言の性質評価を行う．

3.3.3 抽出した例外発言の検証

前節で定義した抽出手法を検証するために、オンラインチャットを利用した予備実験を行った。本実験は「付録 A.1 10 分間で楽しめる新ゲームを考案する」のオンラインチャット実験データを用いた。

対象データは全 128 発言で約 40 分のチャットログデータで、本手法で解析したところ、20 発言が例外発言として抽出された。表 2 に抽出した発言の一部を載せる。ここでのパラメータは、 $w = 20, \alpha = 0.2, \beta = 0.25$ とした。パラメータの設定方法については 4.2 で検証する。

表 2: 抽出した例外発言（一部抜粋）

発言 No.	発言者	発言内容	before 類似度	after 類似度
32	Z	じゃあ僕は 6 で	0.187	0.220
35	W	それやらしいですね > K さん	0.15430	0.140
36	Z	ヒットエンドら ~ ん	0.154	0.14
39	P	でも 10 分じゃできない	0.129	0.220
51	P	マスには乾杯とかばかり	0.139	0.147
53	P	あれしんどいのね かなり	0.113	0.120
60	Z	テニサー自重 $w > W$	0.107	0.139
66	K	あ、身体使う系だ。新しい > Z さん	0.191	0.194
67	W	絵しりとりとか	0.109	0.196
70	K	数独っす	0.176	0.229

表 2 に示した結果から、一見無意味で無価値であると思える発言を例外発言として抽出することができたといえる。しかし上表には載っていないが、1,2 単語から成るような短い発言は必然的に類似度が低くなってしまっていて例外発言として抽出されていることが分かった。こうした発言には主義・主張が含まれることは稀であるため、例外発言を抽出する目的に合致しない。この問題を解決するための手法を次章で述べる。

3.4 例外発言の評価手法

3.4.1 抽出した例外発言の問題点

前項では例外発言の抽出方法について述べた。しかし前述のように、前後の発言群との類似度を評価するだけでは単語数の少ない発言は必然的に類似度が低くなって例外発言として抽出されてしまうが、そういった短い発言に主義・主張が含まれることは稀であり、本研究の目的とする発言とは異なる。そこで本論では、1,2単語から成るような短い発言は除外することにした。

しかし例外発言から価値マイニングをするにあたって、「特定の発言をノイズとして除外する」という処理は容易に実行すべき処理ではない。なぜなら、『一見無意味であると思われる発言からも、画期的なアイデアを創発するトリガーとなるうる』という考えが本研究の根底にあるからである。

そこで本研究では、こうした発言を除去するのではなく、抽出した発言の性質を定量的に評価してその重要度に基づいてランク付して出力することでこの問題を解決することとした。評価のために定義した情報量「選択平均情報量」の詳細については次節から述べる。

3.4.2 選択平均情報量

抽出した例外発言の評価を行うために、ここでは新たな情報量を定義し、評価関数に取り入れた。選択平均情報量の概念図を図7に示す。

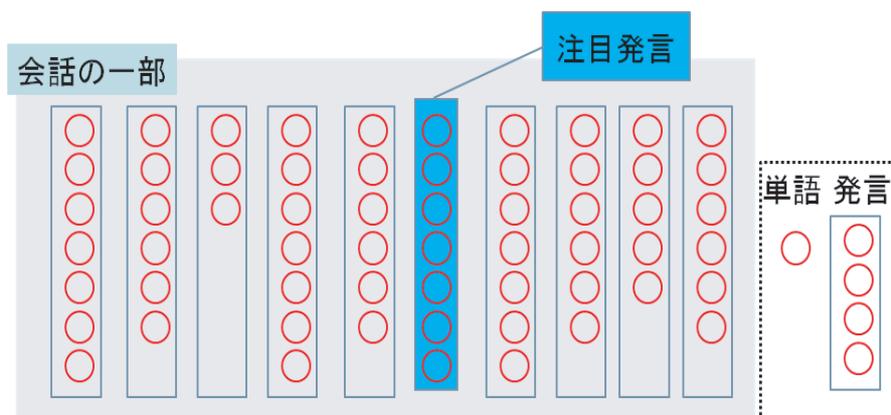


図 7: 選択平均情報量の概念図

選択平均情報量は、会話の一部（発言群，発言集合）において注目発言の持つ情報量を，注目発言に含まれる単語群の情報量の総和として算出したものである．情報量は対象となる集合に与えるインパクト（驚き）の指標となるため，注目発言が発言集合に与えるインパクトを表す．この選択平均情報量は，(8) 式で定義される．ここで， $P_{U_k t_i}$ は発言ベクトル U_k における単語 t_i の生起確率である．

$$E_{AU_k} = \sum_{t_i \in U_k} P_{U_k t_i} I_{At_i} \quad (8)$$

但し， I_{At_i} は発言群 A において単語 t_i の持つ情報量であり，(9) 式で表わされる．ここで， P_{At_i} は発言集合 A における単語 t_i の生起確率を表す．

$$I_{At_i} = -\log P_{At_i} \quad (9)$$

選択平均情報量の性質については 3.4.4 で実例を挙げて論じる．

3.4.3 選択平均情報量による例外発言の評価

前節で定義した選択平均情報量を用いて，抽出された例外発言の評価をする．まず類似度と同様に，before 情報量と after 情報量を以下のように定義する．ここで， $P_{U_k t_i}$ は発言ベクトル U_k における単語 t_i の生起確率， $I_{W_{before \cdot k} t_i}$ ， $I_{W_{after \cdot k} t_i}$ はそれぞれ発言集合 $W_{before \cdot k}$ ， $W_{after \cdot k}$ における単語 t_i の持つ情報量を表す．

$$\text{before 情報量} \quad E_{before U_k} = \sum_{t_i \in U_k} P_{U_k t_i} I_{W_{before \cdot k} t_i} \quad (10)$$

$$\text{after 情報量} \quad E_{after U_k} = \sum_{t_i \in U_k} P_{U_k t_i} I_{W_{after \cdot k} t_i} \quad (11)$$

但し， $W_{before \cdot k}$ ， $W_{after \cdot k}$ は，注目発言前後の w 個（注目発言含む）の発言のベクトル総和であり (12) 式で表わさせる．類似度算出とは相違点があり，前後のベクトルには注目発言ベクトルも含まれることに注意されたい．

$$W_{before \cdot n} = \sum_{n-w < k \leq n} U_k, \quad W_{after \cdot n} = \sum_{n \leq k < n+w} U_k \quad (12)$$

ここで，類似度算出と同様に窓 w を設けることによって，計算コストを軽減し，且，会話における話題展開の速さの違いに対応できるようにするためにユーザーがツール上で探索的に値を設定できるようになっており，類似度算出のために設定した窓 w と同じ値をとる．上式で定義した二つの情報量は，次に示す性質を持つ．

- before 情報量

注目発言以前の会話における注目発言の情報量であり以下の特徴を持つ。

- before 情報量 大 ⇒ 注目発言以前の会話に与える驚きが大きく、新しい話題を挿入している。
- before 情報量 小 ⇒ 注目発言以前の会話に与える驚きが小さく、それまでの会話コンテキストに即している。

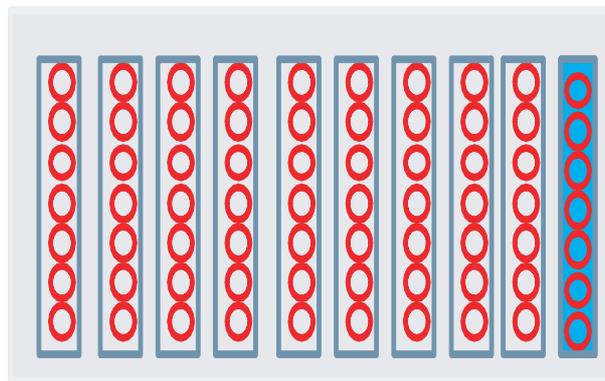


図 8: before 情報量

- after 情報量

注目発言以後の会話における注目発言の情報量であり以下の特徴を持つ。

- after 情報量 大 ⇒ 注目発言以降の会話に与える驚きが大きく、その後の会話とは内容が異なっている。
- after 情報量 小 ⇒ 注目発言以降の会話に与える驚きが小さく、その後の会話と内容が類似している。

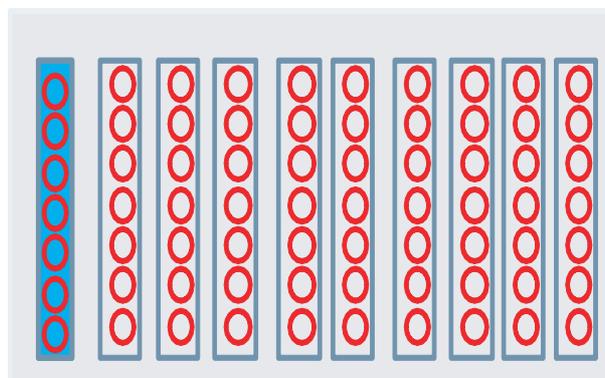


図 9: after 情報量

- 評価関数

次に，上で定義した二つの情報量を利用して例外発言の評価関数を (13) 式で定義する．

$$\text{評価関数 } \bar{E} = E_{\text{before}U_k} - E_{\text{after}U_k} \quad (13)$$

上式のように評価関数は二つの情報量の差分で定義され，以下のような特徴を持つ．

- \bar{E} 大 $\Rightarrow E_{\text{before}U_k}$ 大 $E_{\text{after}U_k}$ 小 \Rightarrow 新たな話題を導入
- \bar{E} 小 $\Rightarrow E_{\text{before}U_k} \div E_{\text{after}U_k} \Rightarrow$ 会話コンテキストに影響小さい

上記の特徴から，評価値が大きいほど有益な意味を含むのではないかと考えられる．そこで，抽出した例外発言をこの値をスコアとしてランク付して出力することで，重要な発言が注目されやすいように設計する．

3.4.4 選択平均情報量の性質検証

本節では，選択平均情報量（以下，単に情報量と表現する）の性質を検証する．検証のために，3.3.3 と同様に付録 A.1 のオンラインチャット実験データを用いた．

ここでは情報量の性質を検証するために，まず全発言に対する評価値を考察する．(13) 式で算出された評価値（前後の情報量の減少量）と，before 類似度との関係を図 10 に示す．

図 10 は before 類似度と評価値 \bar{E} との関係を示している．図中に半透明の三角形で示したように，before 類似度が低いほど評価値の絶対値が大きく値のばらつきがおおきくなる傾向がある．この原因は，発言前の内容との関連性が低い発言は新しい話題を挿入しているため，受け入れられるか排除されるかで大きく発言の性質が変わってしまう不安定な発言であるためだと考えられる．また，評価値が高い発言ほど会話コンテキストを大きく変化させている発言である．

次に，評価値 \bar{E} が高くなった発言の検証をする．評価値 \bar{E} が高くなった発言上位 5 発言を表 3 に挙げる．

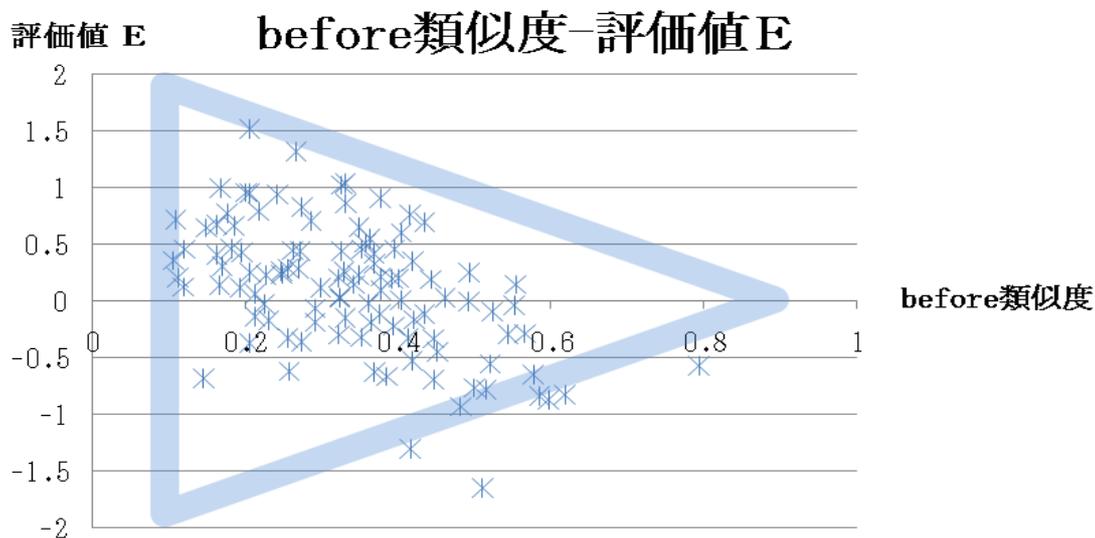


図 10: before 類似度-評価値 E

表 3: 評価関数の検証

ランク	評価値 \bar{E}	発言者	発言内容
1	1.40663030975852	N	身体動作を使うのはおもしろそうであります
2	1.290746483	Z	「利き緑茶」なんてどうでしょう？
3	0.896906507	K	なんかただの確率のゲームのような > Pさん
4	0.861169691	I	オレンジオレンジ多すぎるから色変えました .
5	0.688916778	P	そういえば、サークルの合宿で自作すごろく...

表 3 の結果とチャットにおける会話の流れとの関係を照らし合わせて考察すると、評価値が高くなった発言は会話コンテキストを変化させるファシリテータとなっている発言であることが分かる。こうした発言は会話構造の解析においては重要であると考えられる。しかし本論の目的は例外発言の価値定量化であり、上表の結果から例外発言を考察することはできないが、この結果から、発言の性質（重要度）を評価するための評価関数として妥当なものであるといえる。また、オンラインチャット実験の内容の詳細については 5 章で述べる。

次に、例外発言に対する評価値の検証をする。例外発言のみについての before 類似度と評価値との関係を図 11 に示す。類似度が低い発言を例外発言として抽出しているため、例外発言は図中に半透明の四角で表示したように分布した。ここで注目すべき点として、例外発言は情報量の値にばらつきが大きいことが挙げられる。

前述のように、類似度が小さい発言ほど情報量の値のばらつきが大きくなるが、特に例外発言はその傾向が強く、主義・主張が含まれているであろう発言とそうでない発言が混在しているといえる。例外発言に対する評価値の特徴を表 4 にまとめた。

表 4: 例外発言に対する評価値の特徴

評価値 \bar{E} 大	裏に主義・主張含まれる可能性が大きい発言
評価値 \bar{E} 小	特に意味のない単なる例外発言

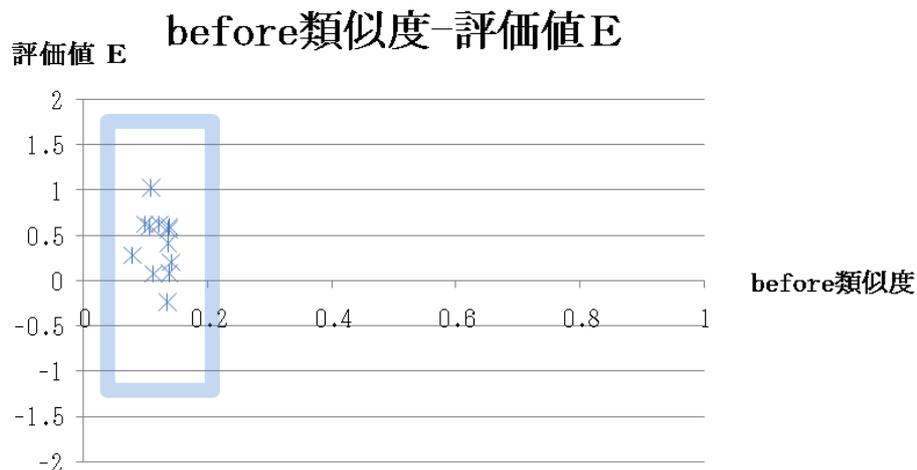


図 11: before 類似度-評価値 E (例外発言のみ)

4 提案システム利用方法

4.1 提案システムの利用プロセス

本章では、提案システムの利用方法について論じる。本研究では二重螺旋プロセスを応用して、図13に示すプロセスで解析を進めていく。まず初めに会話コミュニケーション（もしくは日記などの独話）からテキストデータを取得し、計算機による処理で例外発言が出力される。その後出力結果から「そういえばこんなこともあったな」といったように人のコミュニケーションが生じて更なるテキストデータを得る。再度計算機による処理をし、というように人による処理と計算機による処理を交互に行うことでより深い知識を掘り起こしていく。このプロセスはエンドレスに続けることもできるが、満足な結果が得られたところでプロセスを終了する。

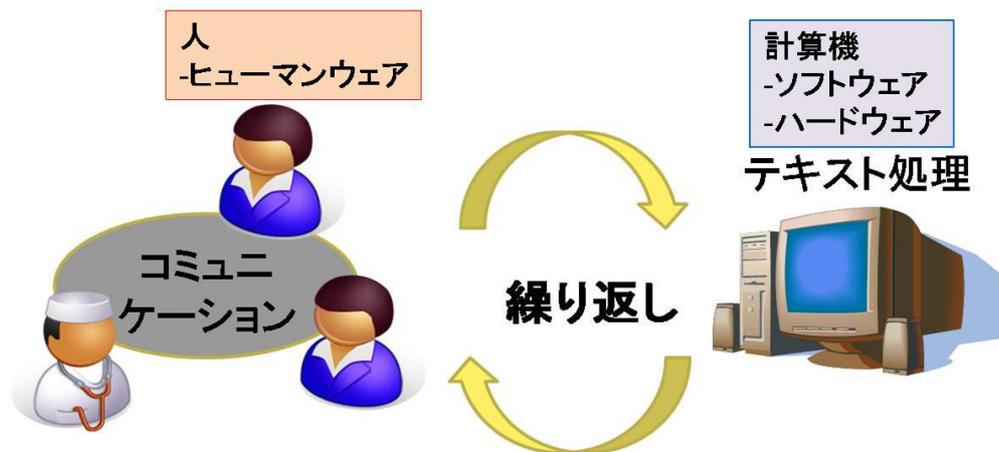


図 13: システム利用プロセス概要

4.2 パラメータの設定方法の検証

本システムでは、ツール上でユーザーが設定できるパラメータとして、計算範囲を決める窓 w と、例外発言抽出のための閾値であり例外発言の定義のために用いる値 α, β がある。ここでは、オンラインチャット実験として行った「付録 A.1 10分間で楽しめる新ゲームを考案する」、「付録 A.2 食事をどうするか」、「付録 14 簡易型イノベーションゲーム」の三つのオンラインチャットログの解析を通してパ

ラメータの検証をする．ここでのオンラインチャット実験は予備実験という位置付けでパラメータの設定のために用いているためチャットデータの詳細には触れず，5章でオンラインチャット実験の内容の解析結果の検証をする．

4.2.1 計算範囲の窓 w の値検証

まず例外発言の抽出と評価値の算出をするときに用いる発言数を決める窓 w の検証をする．検証のために，三つのチャットログデータに対して各々3つの窓を設定して，パラメータの検証をした．また，検証前から自明なこととして，窓 w は，大きいほど会話の大きな流れを考慮することができ，小さいほど計算コストを減らすことができる．この二項目はトレードオフの関係にあるため客観的な最適値は存在しないと考えられる．しかし，本システムを利用にあたって，ユーザーが満足できる妥当なパラメータの設定方法を明示的に示しておく必要がある．パラメータの設定方法の検証を目的として実験を行う．

まず，各パラメータでの実験結果を表 6-14 に示す．ここでは例外発言抽出のための閾値はデフォルト値としている $\alpha = 0.2$ $\beta = 0.3$ に固定しておき，抽出した例外発言のうち上位 10 位までを表示した．

また本実験では，ハンドルネームを使って行っており匿名性の強い状態から実験を始める．したがって，実空間における顔を合わせた状態でのコミュニケーションとは環境が異なっている．また，チャットが進む過程で人物をある程度特定することもできる．故に，各参加者に関する先行情報が少ない状況で実験が始まりその後実験が進んでいくと匿名性が弱まるため，匿名性が強い場合と弱い場合の両方のシチュエーションについての解析ができる．

また，妥当な窓の値を求める上で解析結果を評価する指標が必要となる．ここでは，実際に本システムを利用するシチュエーションにおいては多人数によるアンケートなどの評価は現実的でないと判断して，解析者（システムのオペレーター，ここでは著者本人）の主観によって評価する．格チャットにおけるパラメータに関する考察を次に述べる．

まず A.1 のチャットデータは，抽出した発言数がやや多いが上位発言の内容を考慮して窓 $w = 20$ が適当であるとした．次に，A.2 のチャットデータは，抽出された発言数を考慮して窓 $w = 20$ が適当であるとした．最後に，A.3 のチャットデータは，テーマがやや難解であったためか話題の展開が緩やかであるため窓を大きく

とる必要があり，窓 $w = 30$ が適当であるとした．

また全解析結果に共通して，窓 w は大きいほど注目発言前後の会話コンテキストとの類似度が小さくなり，抽出される例外発言は比較的多くなる傾向がある．以上から窓 w の設定方法としては，「解析対象となる会話の話題展開の早さに応じて解析者が探索的に設定する必要があるといえる．

以上のように本論では解析者が探索時に設定する必要があるとしたが，話題展開の速さを自動的に判別できればパラメータを自動設定できるとも考えられる．これはテキストセグメンテーションの応用によって可能である．テキストセグメンテーションの先行研究としては，教師無しデータからの分割手法である Text Tiling アルゴリズム [54][55] や隠れマルコフモデル (HMM) を用いて話題分割するもの [56] など多数存在する．また，本論で定義した選択平均情報量によって話題が展開するトリガーとなった発言を抽出することもできる (後述，6.2.5 参照) ため，選択平均情報量を利用して再帰的にパラメータを自動設定できると考えられる．しかし本論ではこの検証を行うまでは至らなかったため，今後の課題の一つとしておく．

表 6: パラメータ 窓 w の検証 A.1 $w = 10$

ランク	発言者	発言内容	評価値
1	Z	種田のフォームははガチ $w > I$ さん	0.827
2	Z	人生的な意味で連敗中なので	0.818
3	Z	「利き緑茶」なんてどうでしょう?	0.802
4	K	ジャッジがいるよな。 $> W$ さん	0.426
5	Z	テニサー自重 $w > W$	0.400
6	P	それやらしいですね $> K$ さん	0.346
7	W	今度じの字間違った～	0.219
8	P	マスには乾杯とかばかり	0.212
以上 8 発言			

表 7: パラメータ 窓 w の検証 A.1 $w = 20$

ランク	発言者	発言内容	評価値
1	K	なんかただの確率のゲームのような > P さん	0.897
2	P	でも 10 分じゃできない	0.578
3	Z	種田のフォームははガチ $w > I$ さん	0.565
4	P	クリリンのことがあああああ	0.558
5	P	何度記憶を亡くしたことが	0.527
6	P	やっぱみんなで遊ぼうよ	0.502
7	Z	テニサー自重 $w > W$	0.387
8	W	今度じの字間違った～	0.247
9	W	蛇眼の力をなめるなよ!	0.231
10	Z	人生的な意味で連敗中なので	0.218
ほか 10 発言			

表 8: パラメータ 窓 w の検証 A.1 $w = 30$

ランク	発言者	発言内容	評価値
1	K	あ、身体使う系だ。新しい > Z さん	0.797
2	P	そういえば、サークルの合宿で自作すごろくやりましたよ	0.768
3	P	何度記憶を亡くしたことが	0.633
4	K	物語は厳しいっすね。面白いけど。 > N さん	0.546
5	N	リアル鬼ごっこ。90 人で 10 人を追いかける	0.425
6	Z	やっぱみんなで遊ぼうよ	0.404
7	P	男飲むとか、1 年飲むとか、全員飲むとか	0.277
8	P	でも 10 分じゃできない	0.185
以上 8 発言			

表 9: パラメータ 窓 w の検証 A.2 $w = 10$

ランク	発言者	発言内容	評価値
1	Y	バンビ説再浮上	0.294
以上 1 発言			

表 10: パラメータ 窓 w の検証 A.2 $w = 20$

ランク	発言者	発言内容	評価値
1	Y	バンビ説再浮上	0.652
2	X	そろそろ行きましょう	0.524
3	T	とりあえず生で	0.259
4	Y	留守に何する	0.222
5	X	飲むなら軽く。	0.0192
以上 5 発言			

表 11: パラメータ 窓 w の検証 A.2 $w = 30$

ランク	発言者	発言内容	評価値
1	T	そういうことか!	0.675
2	T	じゃんけん!!	0.482
3	T	食べる方に一票	0.320
4	Y	バンビ説再浮上	0.262
5	T	どうしたいかな~と	0.127
6	Y	知らないけど	0.120
7	X	飲むなら軽く。	0.0745
8	K	留守に何する	-0.321
9	Y	はい多数決とろう	-0.428
以上 9 発言			

表 12: パラメータ 窓 w の検証 A.3 $w = 10$

ランク	発言者	発言内容	評価値
1	DAN	1イモ?	1.63
2	DAN	やっぱこれだな	0.724
以上 2 発言			

表 13: パラメータ 窓 w の検証 A.3 $w = 20$

ランク	発言者	発言内容	評価値
1	M	停電による医療事故防止機能付き EV	0.696
2	DAN	やっぱこれだな	0.303
3	DAN	そか電気自動車か	-0.306
以上 3 発言			

表 14: パラメータ 窓 w の検証 A.3 $w = 30$

ランク	発言者	発言内容	評価値
1	DAN	旅しながら広めて感じ j	0.392
2	DAN	やっぱこれだな	0.276
3	DAN	そか電気自動車か	-0.337
以上 3 発言			

4.2.2 閾値 α, β の値検証

例外発言抽出のための閾値である α, β の値を検証する．ここでも3つのオンラインチャット実験データを用いて，それぞれ3パターンのパラメータ設定で解析実験することで検証を行った．表 15-23 に解析結果の上位10発言までを示す．ここで，窓 w は前節で適値だと判断した値に設定した．

6つの解析結果を見たところ，閾値 α, β を変化させても結果に大きな違いはないことが分かった．窓 w を変化させると評価値に大きく影響して解析結果に大きな違いが見られたが，閾値を変化させても抽出される発言数が変わるだけで本質的な違いはない．また，窓 w の値によって類似度の計算範囲も変わるため，パラメータ設定としては窓 w のほうが圧倒的に重要度が高い．

二つの閾値の設定方法は本研究の目的から「排除されてしまっている例外発言であるにも関わらずコンテキストに影響を与えている発言」を抽出するために，before 類似度よりも after 類似度が大きくなる発言がそれに該当すると考えられるため， α よりも閾値 β の値をやや大きく，もしくは等しい値に設定するのが良いと考えられる．これは，コンテキストに影響を与えている発言は after 類似度が before 類似度よりも大きくなるからである．

総じて，閾値 α, β の値は窓 w ほど重要ではないが，例外発言であるがコンテキストを変化させている発言を抽出するためには α よりも β をやや大きく（または等しく）設定するのが良いと考えられ，ツールのデフォルト値を $\alpha = 0.2, \beta = 0.3$ とすることとした．また，より綿密な解析を行うためには3つのパラメータを数パターンの組み合わせで解析して，適したパラメータを探しつつ解析結果を検討していくと良い．

以上の考慮した上で，各チャットの解析におけるパラメータは次のように設定した．

A.1 のチャットデータは $w = 20, \alpha = 0.2, \beta = 0.3$ ，A.2 のチャットデータは $w = 20, \alpha = 0.2, \beta = 0.3$ ，A.3 のチャットデータは $w = 30, \alpha = 0.3, \beta = 0.3$ として各チャット内容の解析を行うこととした．

表 15: パラメータ 閾値 α, β の検証 A.1 $\alpha = 0.3, \beta = 0.3$

ランク	発言者	発言内容	評価値
1	K	なんかただの確率の ゲームのような > Pさん	0.897
2	P	でも10分じゃできない	0.578
3	Z	種田のフォームははガチw > Iさん	0.515
4	P	クリリンのことかあああああ	0.558
5	P	何度記憶を亡くしたことが	0.527
6	P	やっぱみんなで遊ぼうよ	0.502
7	Z	テニサー自重w > W	0.400
8	W	今度じの字間違った～	0.219
9	W	蛇眼の力をなめるなよ!	0.231
10	Z	人生的な意味で連敗中なので	0.218
ほか 24 発言			

表 16: パラメータ 閾値 α, β の検証 A.1 $\alpha = 0.2, \beta = 0.3$

ランク	発言者	発言内容	評価値
1	K	なんかただの確率の ゲームのような > Pさん	0.897
2	P	でも10分じゃできない	0.578
3	Z	種田のフォームははガチw > Iさん	0.516
4	P	クリリンのことかあああああ	0.558
5	P	何度記憶を亡くしたことが	0.527
6	P	やっぱみんなで遊ぼうよ	0.502
7	Z	テニサー自重w > W	0.400
8	W	今度じの字間違った～	0.219
9	W	蛇眼の力をなめるなよ!	0.231
10	Z	人生的な意味で連敗中なので	0.218
ほか 10 発言			

表 17: パラメータ 閾値 α, β の検証 A.1 $\alpha = 0.3, \beta = 0.2$

ランク	発言者	発言内容	評価値
1	P	クリリンのことかあああああ	0.558
2	Z	種田のフォームははガチ w > I さん	0.515
3	Z	テニサー自重 w > W	0.400
4	W	今度じの字間違った～	0.219
5	Z	人生的な意味で連敗中なので	0.218
6	W	でもこれ 10 分では記憶無くなるまで 楽しめませんね～ > P さん	0.180
7	N	リアル鬼ごっこ . 90 人で 10 人を追いかける	0.0603
8	P	マスには乾杯とかばかり	0.0138
9	K	あ、身体使う系だ。新しい > Z さん	0.00154
10	K	いや、自分のペースでゆっくり 楽しめるのがいいな～ > W さん	-0.0381
ほか 10 発言			

表 18: パラメータ 閾値 α, β の検証 A.2 $\alpha = 0.3, \beta = 0.3$

ランク	発言者	発言内容	評価値
1	Y	バンビ説再浮上	0.652
2	T	いや、たべようよ	0.645
3	T	そろそろ行きましょう	0.524
4	T	食べる方に一票	0.342
5	T	じゃんけん !!	0.330
6	Y	とりあえず生で	0.259
7	K	留守に何する	0.222
8	T	どうしたいかな～と	0.113
9	G	いかんせん空腹です	0.0963
10	X	飲むなら軽く。	0.0192
ほか 8 発言			

表 19: パラメータ 閾値 α, β の検証 A.2 $\alpha = 0.2, \beta = 0.3$

ランク	発言者	発言内容	評価値
1	Y	バンビ説再浮上	0.652
2	T	そろそろ行きましょう	0.524
3	Y	とりあえず生で	0.259
4	K	留守に何する	0.2222
5	X	飲むなら軽く。	0.0192
以上 5 発言			

表 20: パラメータ 閾値 α, β の検証 A.2 $\alpha = 0.3, \beta = 0.2$

ランク	発言者	発言内容	評価値
1	Y	バンビ説再浮上	0.652
2	T	そろそろ行きましょう	0.524
3	X	飲むなら軽く。	0.0192
4	Y	知らないけど	-0.121
5	G	ぼくは八海山	-0.196
6	Y	はい多数決とろう	-0.672
以上 6 発言			

表 21: パラメータ 閾値 α, β の検証 A.3 $\alpha = 0.3, \beta = 0.3$

ランク	発言者	発言内容	評価値
1	DAN	旅しながら広めて感じ j	0.392
2	DAN	やっぱこれだな	0.275761803
3	DAN	それだな。市民ネットワーク	-0.0134
4	DAN	j がないよ	-0.170422312
5	DAN	そか電気自動車か	-0.33662885
6	DAN	広めて感じ j 広めてく感じ	-0.458
以上 6 発言			

表 22: パラメータ 閾値 α, β の検証 A.3 $\alpha = 0.2, \beta = 0.3$

ランク	発言者	発言内容	評価値
1	DAN	旅しながら広めて感じ j	0.392
2	DAN	やっぱこれだな	0.276
3	DAN	そか電気自動車か	-0.337
以上 3 発言			

表 23: パラメータ 閾値 α, β の検証 A.3 $\alpha = 0.3, \beta = 0.2$

ランク	発言者	発言内容	評価値
1	DAN	やっぱこれだな	0.276
2	DAN	そか電気自動車か	-0.337
以上 2 発言			

5 予備実験 オンラインチャット実験

パラメータの設定のためにも利用したオンラインチャット実験の解析をする。ここでは、前章で求めたパラメータを使って解析実験を行った。各実験の参加者は、大学教員と大学院生と研究生を含む3~5名で行われ、30分程度を目安に行われたものである。付録A.1~A.3にチャットの会話ログを載せた。

本実験では、被験者は同室でチャットを行った。互いに顔を合わせられる状態で実験を行うことによって、より実際の会話に近い環境を作った。またチャット内ではハンドルネームを使っていて、実験の始めの時点では発言者が参加者の中の誰であるかを判別できない状態であるが、会話の中で徐々に判別できるようになっている。これによって、匿名性の高い状況での会話と匿名性低い状況での会話との違いも同時に検証することができる。

まず各チャット内容の概要を示す。

- テーマ1：新ゲームの考案

「10分間で楽しめる新しいゲームを考案する」というテーマで行われたチャット。序盤では、幼少時代の記憶から複数のゲームを挙げてそれを組み合わせようという流れで進んで行き、ここで多くのアイデアが出される。その中から身体を使ったジェスチャーゲームが面白いという話になり、ジェスチャーゲームのルールを考えるという流れになる。本実験でテーマの結論は出なかったが、何か楽しそうなゲームが作れそうだという雰囲気を実験を終えた。

- テーマ2：食事をどうするか

「研究室ミーティング後に何処に食事をしに行こうか」という日常会話でしばしば話題となるテーマで行われたチャット。まずはお酒を飲みに行くか食事メインにするかの話から始まる。その後、どちらにするかをじゃんけんをして決めようとなるが、チャットでは

- テーマ3：簡易型イノベーションゲーム

組み合わせ発想ゲームであるイノベーションゲームの一部を利用して簡易型イノベーションゲームとした。ここでは、グッドデザインアワードという賞を受賞した「2001年充電の旅」と「医療事故市民オンブズマン・メディオ」という2つのデザインを組み合わせる新しいデザインを生み出すことをテーマ

としてチャットを行った．テーマがやや難解であったために比較的発言数が少なく，話題の展開が遅い．テーマの答えを思い付けなくて苦し紛れに答えたアイデアが，一見無意味であるが実は新しいコンセプトが含まれていることに被験者の一人が気づき，他のアイデアもでてこなかったため，このアイデアが本チャットの暫定的な結論となった．

5.1 チャットテーマ1：新ゲームの考案

テーマ 10分で楽しめる新しいゲームを考案する

前章の考察から，パラメータは $w = 20, \alpha = 0.2, \beta = 0.3$ とした．解析結果は表 16 に示したものになる．抽出した発言から想起できるアイデアを以下に挙げる．

- チャットでの会話では「アニメやドラマなどの名シーンを再現してそれを見て何を再現しているか当てる」というゲームが最も良いと考えられる案とされており，ここに必要なルールとしてジャッジの方法や再現するお題の決め方などにも言及されているが，会話の中で言及されているものの話題として取り上げられなかった．しかし新しいゲームを考案する上では重要な要素であり，注目されていないが重要だと思われる発言を抽出できたといえる．
- ランキング第 13 位の発言で出てきた「リアル鬼ごっこ」というアイデアは特に周りに注目されることがなかったが，あるテレビ番組で街中で鬼ごっこをするというものがあり，この発言者 N はこの番組を見たことがあって自分も参加してみたいという気持ちを持っているのではないかと考えられる．
- ランキング 3 位に「種田のフォームはガチ」という発言があるが，発言者本人にインタビューしてみたところ，この発言者 Z は元高校球児で野球選手のものまねが流行ったことがあってその時の楽しかった思い出を共有できないかと考えた，ということであり，この思い出と基にして何か良いアイデアがないかと意図された発言であった．チャットの中でこうした思い出が掘り起こされて利用されることはなかったが，本解析によって発言者の物語の一部を垣間見ることができたといえる．

5.2 チャットテーマ2：食事をどうするか

テーマ 研究室ミーティング後『どこに食事に行くか』

前章の考察から，パラメータは $w = 20, \alpha = 0.2, \beta = 0.3$ とした．解析結果は表 19 に示したものになる．抽出した発言から掘り起こした知見を以下に挙げる．この実験では何か新しいアイデアを考えようというものではないため，新しいアイデアの創発に役立つ知見ではなく主に発言の裏にある状況を挙げた．

- ランキング1位の「バンビ説再浮上」という発言は，発言者本人にインタビューしたところ，最初のほうに一度出てきた「バンビ」という定食屋がしばらくしてから再び出てきた時の発言で，もともとここに行きたかったためにうまく誘導できないかという意図があった．しかしその後回文をつかった言葉遊びが始まったために，この発言は放置されることとなってしまった．
- ランキング3位の「とりあえず生だ」という発言は，本人へのインタビューによると，発言者 Y は飲みに行きたかったにも関わらず，会話の流れが食事に行くという方向へ向かっていたために飲みに行くという方向に会話を展開させようとした発言であった．
- ランキング4位の「留守になにする」という発言は，前後コンテキストに沿わないわけではなく，回文を使った言葉遊びの中で出てきた発言である．この発言の内容自体には意味が含まれていないため例外発言として抽出されたと考えられる．
- ランキング5位の「飲むなら軽く」という発言は，飲みメインか食事メインかで意見が分かれていたため，どちらの意見も取り入れることができるようにと意図された発言であると考えられる．
- チャット実験による会話の結果として，このチャット後に「バンビ」という店に食事に行ったため，ランキング1位の「バンビ説再浮上」という発言は，会話の中では影響が少なかったが，結果としてチャット参加者の心を動かしたのではないかと考えられる．

- オンラインチャット上でじゃんけんをしようという話がでてくるが、やはり難しかったため多数決をとることになっている。ここから、オンラインチャットにじゃんけんができる機能があると良いという新たなニーズを得ることができた。

5.3 チャットテーマ3：簡易型イノベーションゲーム

テーマ 二つのアイデアを組み合わせる新しいアイデアを創発する

前章の考察から、パラメータは $w = 30, \alpha = 0.3, \beta = 0.3$ とした。解析結果は表 21 に示したものになる。抽出した発言から想起できる知見を以下に挙げる。

- ランキング1位の「旅しながら広めて感じj」という発言は、「2008年医療事故市民オンブズマン・メディオの旅」という名前を組み合わせただけの一見無意味なアイデアに対して、少し視点を変えれば実は良いアイデアではないかということに気づいたことを報告する発言であるが、タイピングミスによってうまく周りに伝わらなかった発言であった。その直後に訂正しているため少し反応があり、この反応によって高い評価を得たと考えられる。
- ランキング2位の「それだな。市民ネットワーク」という発言は、発言者本人によると、前の発言で出てきた「市民ネットワーク」というキーワードが新しい発想につながるのではないかと漠然と感じて発言されたものであるが、その後の会話では結局誰にも触れられずに終わってしまった。しかしこのキーワードに注目して再度議論することで新たなアイデアが出てくる可能性があると考えられる。
- 例外発言として抽出された発言は全て同一の被験者の発言であったが、ここには、本実験では参加者が少なく匿名性が弱かったために、被験者の先行情報が暗黙知として存在していたと考えられる。
- テーマがやや難解であったのが原因で、実験時間に対してチャットの発言数が少なくなってしまう、その中でも比較的多く発言していた DAN という参加者の発言のみが例外発言として抽出されたという解釈も可能である。

- 本来のイノベーションゲームは、企業家が基礎カードを基にしてアイデアを出し、そのアイデアを投資家もしくは消費者がゲーム内での擬似的なお金を用いて評価するものであるが、本実験では評価者がいないため被験者間で競うという要素はない。それ故に本来のイノベーションゲームよりも自由な発想が起こりやすい環境を作ることができたと考えられる。

5.4 考察

3つのオンラインチャット実験の解析結果から、本手法によって、会話の中で大きく注目されることはなかったが裏に情況・物語を含んでいる発言を抽出することができたといえる。しかし、本実験は、各テーマについて繰り返しの実験を行っておらず発言者本人へのインタビューをする程度で終わってしまっているため、本システムの利用法の検証はできなかった。さらに、得られた知見の中には情報のフィードバックをした後に再度会話を発生させることで何らかの価値を付加できると考えられるものも多く、前述した繰り返しプロセスが必要であることが確認できた。

また、チャットのようなくだけた会話では、省略が起こっていたり、共有知や暗黙知であるために発言されることなく会話ログデータとして残らないといった現象が頻繁におこる。このため、会話ログのみでは情報欠落があると考えられる。こうした問題の解決方法として、欠落部分へアノテーションを付加するによって情報を補完することが考えられるが、アノテーション付加は作業量は膨大で、さらに発話者本人へのインタビューを必要となることも多いため、多大な労力を費やすことができる状況でなければこの解決策は実現性が低い。

次に、チャットにおける匿名性の違いによる影響を考察する。例として、「A.2 食事をどうするか」の解析結果のランキングと発言 No と発言者のみを表 24 に挙げる。また、このチャット実験では全 139 発言であった。

このチャット実験では各参加者の識別ができ始めている。したがって表 24 から、匿名性が低くなってきた会話の中盤から例外発言が多く抽出されるようになったことが分かった。この背景には被験者間の人間関係があると思われ、普段の生活から突拍子もない発言が多い人物の発言が例外発言として抽出されることが多いという傾向にあった。特にチャット参加者が少なく匿名性の低かった A.3 の実験ではこの傾向が強く見られ（表 5.4）、例外発言として抽出された発言は全て同一人物のものであった。ここから、ある程度気心の知れた間柄における会話では特定の人物の発言が排除されやすくなることがあるというコミュニケーションリスクがあることが分かった。また、こうした関係の人物間のコミュニケーションでは共有知や暗黙知が多数存在しており、会話ログデータでは観測できない情報が多く存在すると考えられる。

表 24: 匿名性による違いの検証 A.2

ランク	発言 No	発言者
1	66	Y
2	77	T
3	79	Y
4	75	K
5	59	X

表 25: 匿名性による違いの検証 A.3

ランク	発言 No	発言者
1	30	DAN
2	51	DAN
3	37	DAN
4	33	DAN
5	45	DAN

6 提案システムの評価実験

地方自治体定例会会議録と患者インタビューのデータの解析実験によって提案システムの評価をおこなった。地方自治体定例会会議録の解析では、抽出される例外発言の数とその発言者の内訳から会話参加者間における人間関係を知るための重要な情報をえることができた。また患者インタビューの解析では、患者の持っている関心事・問題意識に言及されているにも関わらず、周りの人（ここでは記者2人）に注目されずに排除されてしまっている関心事・問題意識を掘り起こすことができることが分かった。

6.1 地方自治体の定例会会議録の解析

本システムの検証実験として、地方自治体の定例会の会議録の解析をした。ここでは、テレビ出演などの芸能活動をしていた有名人が知事になった（俗に言うタレント議員）という特殊な事例として宮崎県と大阪府の二つの都道府県を対象として解析実験した。解析の結果、議長や副議長の発言が例外発言として抽出されることがしばしばあったが、司会として進行をしている非常に形式的な発言であってそこには状況・物語が存在せず本研究の目的に合致しないため、議長や副議長の発言は例外発言としては抽出しないよう除外した。

本評価実験の目的は例外発言に潜む状況・物語を掘り起こすことではなく、抽出した例外発言数と例外発言が抽出された発言者の内訳から「会話の活性度と会話参加者間における人間関係を知るための情報を得られるか」を評価することが目的である。本実験では、二つの自治体の会議録に対しそれぞれタレント議員が就任する前、直後、数か月後の3か月分の定例会の会議録データの解析をした。入力データとなる会議録には一度の発言が長いものが多かったため、「続いて」や「次に」、「最後に」といった明確にトピックが変わる言葉で発言を区切って同一人物の連続する発言として解析した。

6.1.1 宮崎県定例会

平成 18 年 11 月と平成 19 年 2 月と平成 20 年 6 月の定例会の会議録の解析をした。ここで、平成 19 年 1 月 23 日に東国原英夫知事が就任していることに注意する。またパラメータは、全て $w = 10, \alpha = 0.2, \beta = 0.25$ に統一して設定した。解析結果を以下に示す。

平成 18 年 11 月 全 155 発言。例外発言は一つも抽出されず。

平成 19 年 2 月 全 221 発言。3 人の議員がそれぞれ一発言ずつ、計 3 発言が例外発言として抽出された。

平成 20 年 6 月 全 265 発言。東国原知事の 1 発言を含む 9 発言が例外発言として抽出された。

東国原知事の就任前の会議録では例外発言が抽出されなかったが、就任後には少数だが例外発言が抽出されるようになった。それだけでなく、会議における全体の発言数も増えている。したがって東国原知事の就任によって、形式的に行われていた感情のあまり含まれない無機質な会議が、発言に感情が含まれる白熱した会議になったのではないかと推測できる。さらに、東国原知事の発言が例外発言として抽出されることが極僅かであることから、東国原知事が地元議員に受け入れられているであろうと考えられる。実際のところ、東国原知事は 90% を超える非常に高い支持率を維持しており、宮崎県民からの期待は大きく信頼されていることから、本解析結果は妥当であると言える。

6.1.2 大阪府定例会

平成 19 年 12 月と平成 20 年 5 月と平成 20 年 9 月の定例会の会議録の解析をした。ここで、平成 20 年 2 月 6 日に橋下徹知事が就任していることに注意する。またパラメータは、全て $w = 20, \alpha = 0.2, \beta = 0.25$ に統一して設定した。解析結果を以下に示す。

平成 19 年 12 月 全 298 発言。太田前知事の 1 発言を含む全 5 発言が例外発言として抽出された。

平成 20 年 5 月 全 348 発言。橋下知事の 8 発言を含む全 18 発言が例外発言として抽出された。

平成 20 年 9 月 全 278 発言。例外発言は一つも抽出されず。

橋下知事の就任前の会議でも例外発言が抽出されたが、就任直後の会議では多くの例外発言が抽出され、橋下知事に対する期待や反発心が強いのではないかと推測される。しかし、就任数か月後の会議では例外発言が一つも抽出されないという結果になっており、「橋下知事が受け入れられた」、もしくは「橋下知事が府議会のルールに沿うようになった」というような推測ができる。また、会議における全発言数に注目すると、平成 20 年 5 月に行われた定例会では発言数が大きく増えたが、平成 0 年 9 月の定例会では橋下知事就任前よりも少なくなっている。例外発言が抽出されなくなっていることを同時に考慮すると、橋下知事が期待外れで会議が盛り上がりなくなってしまうのではないかと考えることができる。また、ここでは三回分の会議録のみの解析結果であるため、この結果だけで結論を下すことは不可能であり、より詳細に解析を進めるためには他月の会議録の解析と合わせて関係者へのインタビューが必要となる。また、東国原宮崎県知事と比べるとまだ就任して日が浅いため、今後の定例会会議録の解析を進めることで、橋下大阪府知事と地元議員や大阪府民との関係が見えてくると考えられる。

6.1.3 会議録解析の考察

本システムの評価実験として、宮崎県と大阪府の定例会の会議録を各3カ月ずつ解析した。2都道府県の解析結果を比較する。まず宮崎県の東国原知事は、就任前には例外発言が抽出されていなかったが就任後からは抽出されるようになったことから、形式的に行われていただけの定例会を活性化させており、且、知事自身の発言が例外発言として抽出されることが少なかったため地元議員に受け入れられていたと考えられる。一方大阪府の橋下知事は、就任初期には比較的多くの発言が例外発言として抽出されており、議論を活性化させたという解釈もできるが、知事本人の例外発言が多くあまり快く受け入れられてはいないのではないかと考えられる。さらに、就任数か月後の定例会では例外発言が一つも抽出されなかった。これは、「知事が地元議員に受け入れられるようになった」、「知事が議会のルールに合わせるようになって形式的な会議になった」などの理由が考えられるが、今回の解析だけでは結論を出すことができない。しかし、橋下知事とその周りの環境に大きな変化があったのではないかと推測できる。但し、より詳細に解析するためには、関係者や専門家へのさらなるインタビューを行う必要がある。

以上のように、本実験により「議論の活性度を測る」、「コミュニケーションにおける人間関係を推測する」の2つの目的を達成するための有力な情報を得ることができ、コミュニティの分析にも利用できることが分かった。但し、より詳細な解析のためには会話データの収集を始めとして解析対象の周辺情報の収集として関係者へのインタビュー等が必要となり、相応の労力を要してしまうという問題点も存在する。

6.2 患者インタビューの解析

6.2.1 インタビューデータ概要

患者に対するインタビューの解析実験を行った．まず解析対象としたインタビューデータの概要を以下に挙げる，付録 B にインタビューの冒頭部分を載せる．

- 被験者は心臓ペースメーカーの植え込み手術を受けた患者さんに対するインタビュー．
- 被験者は「ICD 友の会」という患者会に積極的に参加しており，エンジニアとして働いている人物．
- 患者 1 人と記者 (インタビュアー) 2 人によるインタビューの会話ログであり，全 365 発言で 1 時間 30 分から 2 時間ほど行われた．
- 基本的には一人の連続した発言を 1 発言としたが，一度に複数のトピックに言及している場合は適宜区切って同一人物の複数発言であるとし，前後の他の人の発言を対象として評価値を算出した．

6.2.2 解析環境の設定

適当な解析環境を設定するために，まずいくつかのパラメータで解析実験を行ったところ，以下のような環境に設定した．

- 記者 1 はインタビューの進行を促すための発言が多いため例外発言としては抽出しない．
- インタビューであるため，インタビュアーが事前に質問項目を用意していたと考えられ，そのため話の展開が早いため窓は小さい値が良く，ここでは窓 $w = 10$ が適値であるとした．
- 例外発言抽出のための閾値は $\alpha = 0.2, \beta = 0.25$ とした．

6.2.3 解析結果

抽出した例外発言を表 26 に示す。例外発言は全 21 発言抽出され、その内 11 発言が患者の発言、10 発言が記者 2 の発言であった。ランキング上位の 15 発言と、ランクは低いと重要であると判断した 2 発言について、ランク・発言者・発言内容・評価値を表 26 に示した。

表 26: 患者インタビューの解析結果

ランク	発言者	発言内容	評価値
1	患者	ありましたね。 「なんで」という感じは、言葉としてはそれしかないですね、 「なんで」という感じですね。	1.15
2	患者	採用されないかもしれないですね。やめておきます。 この話はブログには一切書いていません。 ブログに書けないことはいっぱいあります。	1.01
3	患者	退院はそれからあと 1 週間です。	0.980
4	記者 2	誰も分からないです	0.876
5	患者	今日取材を受けたということはブログに書いてよろしいのでしょうか。	0.868
6	記者 2	そこはご自宅から近いのですか。	0.814
7	記者 2	データは事前にありますしね。	0.800
8	患者	気がついたら写真に撮って掲示しています。	0.712
9	患者	最後は自己リスクです。自己リスクでどこまで理解できるかと。 理解できない人はやらなければいいだけの話で、 理解できる人にはきちんと真実を伝えてほしいと思っています。	0.493
10	患者	退少し寒かったのどうとうとしながら、少し寝たかもしれません。	0.417
11	記者 2	それでも 3~4 日くらいですか。	0.415
12	患者	じっとしていることがないものですから。	0.408
13	記者 2	もっぱら発信ばかりという感じですか。	0.404
14	患者	今日は時間の無駄になったかもしれません。	0.342
15	患者	ポーンとこの手帳を渡されて、退院と帰ってきて、 これを開けて読まないといけないと。	0.203
19	患者	ICD は誤作動があるから本当に大変だと思います。	0.0134
20	患者	大したことは書いていなくても裸のところが分かりますね。	0.0103
ほか 4 発言			

6.2.4 解析によって得られた知見

表 26 の結果から、会話によるコミュニケーションに含まれる二つのリスクを発見することができた。この二つのリスクを以下に挙げる。

1. 発言自体から想起できるリスク

発言者がある関心・問題意識（リスク）について言及しているにも関わらず、周りに排除されてしまうことでその問題に焦点が当てられることなく会話が次の話題へ展開してしまったり、発言者の主義・主張が置き去りにされてしまっている。このように排除されてしまった発言の中で言及されている関心・問題意識。

2. 例外発言の内訳から見てくる人間関係のコミュニケーションリスク

会話参加者間の人間関係によって起こっている、重要視されるべき人物の発言が軽視されてしまっているというコミュニケーションに含まれるリスク。

これらのリスクを発見するためには4章で述べたシステム利用プロセスをきちんと実行する必要があり、発言自体から想起できるリスクを掘り起こすためには抽出された発言の前後の文脈を考慮する必要があり、コミュニケーションリスクは解析結果の抽出された発言の発言者の内訳と会話内容全体の雰囲気を読み取る必要が生じる。ここでは、解析対象とする会話の背景を把握するために会話参加者本人へのインタビューが必要となる。本論では本人へのインタビューはしていないため、これは今後の課題の一つとしておく。本論では、抽出発言から想起できる患者の状況・物語を次のように考察した。

本手法では抽出した発言に注目してその発言前後の会話を考慮して例外発言の裏に潜む状況・物語を考察することで、1. で挙げたリスクを掘り起こすことができる。ここでは、抽出した例外発言の中でも特に重要であると判断した患者の発言に関して、考察を発言前後の会話を挙げながら考察する。

- ありましたね。「なんで」という感じは、言葉としてはそれしかないですね、「なんで」という感じですね。

~~~~実データ引用ここから~~~~

.....

記者1：逆にこれまでずっとスポーツも万能でずっと健康で過ごしていらっしやった野口さんにとっ ては、なぜ僕にこういうことが起こるのかとか、運命ではないですが、そういう思いはなかったのでしょうか。なぜわざわざ僕にみたいな。

抽出発言

記者1：でもそこから、そういう思いもあるけれども、今差し迫った治療のほうに行こうという事は割とすぐに決められたということですか。

.....

~~~~引用ここまで~~~~

ここでは、「なぜ自分がこんな病気に」という、重篤な患者であれば誰しもが感じるであろう疑問に関して言及しているが、インタビュアーはその疑問に関して深く質問することなく直後に次の話題へと展開させてしまっている。ここでの会話内容から察するに、この患者は「マラソン大会に参加するほどスポーツ好きであって健康的な生活を送っていたにも関わらず何故そんな自分が心臓疾患になってしまったのか」という強い思いがあったと考えられる。

- 採用されないかもしれないですね。やめておきます。この話はブログには一切書いていません。ブログに書けないことはいっぱいあります。

~~~~実データ引用ここから~~~~

.....

患者：今日取材を受けたということはブログに書いてよろしいのでしょうか。

記者2：具体的な出版物はまだ出版しますと言っていないので、もしかすると本にならないかもしれませんが。

抽出発言

記者2：これは最初にお話ししたかもしれませんが、バックにN社が協力してくださっていて、いろいろ情報をいただいたり、最終的にはこの本を買っていただいて、そういう面でもご協力をいただくことになっています。

.....

~~~~引用ここまで~~~~

患者は自身のブログを公開しており、患者のライフワークの一つとなっている。このため「ブログにインタビューの内容を書いても大丈夫か」という疑問は記者にとっては重要ではなかったが患者にとっては重要なことであったと考えられ、「ブログに書けないことはいっぱいあります。」という発言からブログに関することでもっと話したかったことがあったのではないかと考えられる。

- 最後は自己リスクです。自己リスクでどこまで理解できるかと。理解できない人はやらなければいいだけの話で、理解できる人にはきちんと真実を伝えてほしいと思っています。

~~~~実データ引用ここから~~~~

.....

患者：先生方は「体脂肪計を測らなくても死にはせん」と言いますが、違うのです。やはりわれわれ、アスリートとは言いませんが、運動する人にとっては自分をいかに管理できるかということに、今より最新技術は使いたいのですから、それをいきなり使ってはいけないのではなくて、使うことを前提として何を注意すればいいかという方向に持って行ってほしいと思います。難しさは分かると言いながら、そちらにもっと振ってくれないかなということをおわたしがいちばん思っています。この間、実は友の会の中でもこの体脂肪計の話をしたら先生は笑っていました。抽出発言 . あとは患者の視点に立った制度設計の話があって、これは少しいやらしい話です。IH だとか、ここは、わたしは自信がありません。今のペースメーカーでいちばん大きなところは M 社ですね。

.....

~~~~引用ここまで~~~~

スポーツ好きで自身の体の状態を知っておくために体脂肪計を使って測定したいけれども、ペースメーカーに悪影響を及ぼす可能性があるため使用禁止とされてしまった。しかし、自分は技術に関する理解力があるためきちんとした説明をして欲しいという強い思いがあったのだと考えられる！患者に対して技術に関する説明をしても理解できない患者も多いが、この患者のように技術に関する理解力を持っている患者に対してはきちんとした説明をするような配慮が必要。また、患者にはこうした説明を受ける権利がある」といった主張が含まれると推測される。しかしここでも、患者の強い主張・問題意識について言及されているのも関わらず、特に触れられることなく会話が次の話題へ進んでしまっている。

- ポーンとこの手帳を渡されて、退院と帰ってきて、これを開けて読まないといけないと。

~~~~実データ引用ここから~~~~

.....

患者：この最後に書いてあります。ここに書いてあって申し込みがここに書いてありますが、これを渡されるだけで説明がないのです。抽出発言。たぶんこれも何か縛りがかかっているのでしょうか。先生も含めて病院は患者会のことについては手を出しません。だから掲示とかの協力はするけれども、入会の支援はしません。

.....

~~~~引用ここまで~~~~

患者が参加している患者会に関する話題で、「患者会は患者にとってとても有益なものであるにも関わらず、病院は患者会への入会の支援をしてくれない」という、患者会のメンバーであるからこそ気付いている問題を指摘している発言である。しかし、インタビュアーはそれほど重要な問題ではないと判断してしまってこの問題に触れることはなかった。

- ICD は誤作動があるから本当に大変だと思います。

~~~~実データ引用ここから~~~~

.....

記者2：あまり悩みばかりの本でも読むと怖くなってきたりして。先ほどうつ病の方とおっしゃいましたが、アメリカとかはもっと心臓病の方が多いので本当にうつ病が大きな問題です。今実はアメリカの情報も少し集めようとしています。この方はサイコセラピストですが、循環器病院ですずっとやっていらして、そういう心のケアをやっていらっしゃる方です。後ろのほうにICDは特別にこういうこともあるという項目がありますが。

抽出発言

記者2：これとは別にこの先生が関係している大学では抗うつ薬の臨床試験がありますね。そのいろいろなグループで、アメリカは心筋梗塞が多いですが、心臓病になった方が発症したうつ病にどの薬がいいか。そういうグループの臨床試験を行っているという論文がこの間出ていました。心臓の治療でメンタルな面はこれから問題になると思います。

.....

~~~~引用ここまで~~~~

一つ前の発言でうつ病の話題がでてきてICDでは心のケアが大きな問題だという記者2の発言の後の患者の発言であって、患者はICDについてもっと話したいこと

があったのではないかと考えられるが、記者2はこれを単なる相槌として受け取ってすぐに次の話題へと展開してしまっている。

- 体験談は各メーカーにもありますね。あれはものすごく役に立ちます。大したことは書いていなくても裸のところが分かりますね。

~~~~実データ引用ここから~~~~

.....

記者1：これはもともとはイギリスでやっていることでLという団体がありますが、そこは別にがんだけではなくて50疾患ぐらいがあります。これから患者になったばかりの人、診断を受けたばかりの人、家族とかがそのネットで実際に患者さんが語っておられるのを見たり読んだりして役に立ったことが多いということで、イギリスではけっこう評価されているサイトなので、それを日本でもつくりたいかと思っています。

患者： 抽出発言 あれはものすごくいいですね。これでしゃべっているとちょっと雰囲気伝わってくると思いますね。

記者1：わたし自身はまだ行っていませんが、実際に今インタビューを集めていますが、これは自ら語ってもよいと言ってもらえる人のところにしか行けないので、ある意味、自分で何か言いたい、社会に対しても何か言いたいという方が多いのですが、それにしても1回話を聞くと5時間ぐらいかかったりして、毎回毎回すごいと聞いています。これを今2年目ですが3年で一通りのことをやろうとしています。

.....

~~~~引用ここまで~~~~

闘病の体験談に関する話題が出てきたところでの患者の発言であり、「ものすごくいい」という発言からも推測できるように、患者はやや興奮ぎみでたくさん話したいことがありそうだが、記者1がすぐに自分の仕事に関する発言でこの話題を終わらせてしまっている。しかしここから「体験談は非常に役に立つが、イギリスのように容易に入手できるようなところが日本にはまだないから欲しい」という患者の要望が抽出できる。

以上が発言自体から想起できるリスクの一部である。本解析によって、患者が持つ多数の問題意識を抽出することができた。こうした結果と筆者自身の経験から、患

者は周囲の環境に含まれるリスクに対するセンシング能力に優れていると考えられる。なぜなら、患者ほど自身の体と身の回りの環境の体への影響について熟考する人物はいないからである。医療のエキスパートである医師や看護師でさえも気付かないような問題に患者が気付くことも少なくない。こうした患者の発言に注目することで有益な情報を得ることができることが本実験によって明らかになった。

また、この患者インタビューには例外発言の内訳から見えてくるコミュニケーションリスクも発見できた。全 21 発言抽出されその内 11 発言が患者の発言であり例外発言の過半数が患者の発言であったという解析結果から、医療サービスを受ける側であるため最も重要視されるべき患者の発言が軽視されてしまっているという事実が浮き彫りになった。本インタビューでは、インタビュアーの 2 人が記者であったことから、事前にインタビューの筋書きがあつて患者がその流れに反した発言をした時に半ば無理やり話題を事前の筋書きに戻していたためにこうした結果になったのではないかと考えられる。このリスクを表す象徴的な発言として「勝手なことばかりしゃべってお時間を取らせてしまいました。」と患者がインタビューの最後に発言している。この発言から患者はもっと話したいことがあったのではないかと推測でき、例外発言として抽出された発言にも丁寧に対応していれば、記者の事前の筋書きとは違った方向にも話題が広がって、より有益な情報を患者から引き出すことができたのではないかと考えられる。本手法では、これらのように潜在価値を持っているにも関わらず排除されてしまった例外発言を定量化することで、潜在価値を發揮できなかった発言から新たな価値観・問題意識を掘り起こすきっかけを作ることができると分かった。

本実験の結論として、本手法による例外発言の定量化による評価を行うことで、前述したような患者の状況・物語が潜んでいる可能性が高いと推測できる発言を抽出できたと言える。また、この解析結果を基にして更にインタビューを行うことで患者が潜在的に持っている価値を引き出すことができると考えられる。しかし、本論ではこの段階までには至らず、本研究の今後の課題の一つとなる。したがって本実験によって、例外発言の価値定量化の一部が成功したと分かった。

6.2.5 他手法との比較

本手法と他手法との比較実験を行った．本手法は会話データから発言単位で抽出するものであるが，ここでは，発言単位ではなく，単語単位でキーワードを抽出する手法である TF-IDF 法 [59] と活性伝搬法 (IDM 法) [49] を比較検証のために利用した．

- TF-IDF 法との比較

TF-IDF (Term Frequency and Inverse Document Frequency) 法は，式 14 に示すように文書内の単語の重要度を評価して，文書を特徴付けるキーワードを抽出する手法であり，情報検索 (Information Retrieval: IR) に利用されている手法である．この評価関数にはいくつかのパターンが存在するが，本論では次式を用いた．但し， t は任意の単語， $TF(t)$ は文書中の単語 t の出現数， $DF(t)$ は単語 t が出現する文書数である．

$$TFIDF(t) = TF(t) \times \log \frac{1}{DF(t)} \quad (14)$$

ここで， DF 値を算出するためにはコーパスが必要となる．しかし本データの解析に適切なコーパスが手に入らなかったため，本論では解析対象データ自体をコーパスとして用いた．但し，一発言をドキュメントとして DF 値を算出した．TFIDF 法を用いて抽出した単語の上位 10 位を表 27 に示す．また，ここでも同様の類義語処理を前処理として行ったデータを解析した．TFIDF 法は文書を要約するために重要となるキーワードを抽出する手法であるため，「ペースメーカー」や「病院」，「自分」といった会話の大筋を表す単語が抽出された．しかし，本手法で注目した発言に含まれる「なんで」や「ブログ」，「自己リスク」といった言葉は重要語としては抽出されなかった．逆に，本手法による解析では「自分」というキーワードは抽出しなかった．しかしこの単語は患者が一人称を示すためによく使う言葉であってこの単語の裏に状況・物語が含まれてはいないため，本手法の目的とするキーワードではない．以上の結果から，TFIDF 法は一般的な文書の要約には有用であるが会

表 27: TFIDF 法 抽出されたキーワード

| ランク | 抽出単語 |
|-----|---------|
| 1 | 自分 |
| 2 | そういう |
| 3 | ペースメーカー |
| 4 | 病院 |
| 5 | いろいろ |
| 6 | 分かる |
| 7 | 行く |
| 8 | ぐらい |
| 9 | 多い |
| 10 | こういう |

話の解析には適していない。本手法では発言の裏に潜んでいる状況・物語を掘り起こすことができるため、より深い知見が得られることが分かった。

- IDM 法との比較

IDM 法は、ある発言とそれに対する返答の関係によって伝搬していく単語の活性度から重要語を抽出することができ、本論で対象としている会話の分析にも適している。こうした特徴から、本手法との比較手法に適していると考えられる。IDM 法の概要は 2.3 に示したものである。IDM 法によって抽出されたキーワードを表 28 に示す。

IDM 法では、会話の主題となる「ペースメーカー」という単語の他に、本手法でも重要であるとされた「ブログ」や「仕事」といった単語もキーワードとして抽出されている。したがって、IDM 法では会話中の話題転換のトリガーとなったようなキーワードを抽出することができ、本手法では、会話の中ではその潜在価値を発揮することができなかつたが焦点を当ててさらに議論すれば新たな知見を得ることができるであろう発言を抽出することができる。ここで、IDM 法によって抽出できる話題の転換点となったような発言（キーワード）は、会話内で注目されて既に何らかの価値を発揮したものであり会話参加者であれば話の流れとして覚えていると考えられるものであると言える。一方本論では、イノベーションを引き起こすような発言はこのような発言ではなく、注目されることはなかつたが裏に発言者の状況・物語

表 28: IDM 法 抽出されたキーワード

| ランク | 抽出単語 |
|-----|---------|
| 1 | ペースメーカー |
| 2 | ブログ |
| 3 | 自分 |
| 4 | アブレーション |
| 5 | 病院 |
| 6 | 入る |
| 7 | 症状 |
| 8 | 説明 |
| 9 | 仕事 |
| 10 | 友の会 |

が潜んでいる発言であるとする立場に立っている。したがって、従来手法は既存の価値観が形成された過程を解析することには役に立つが、抜本的な意識革新（イノベーション）を直接的に引き起こすことはできないと考えられる。こうした観点から考察すると、本手法の最大の特徴である、未だ潜在価値が発揮されていない例外に注目することがイノベーションを引き起こすために必要不可欠な要素となる。実験では、発言者本人でさえ明確には気付いていなかったような価値観を掘り起こすことができたため、本手法は人の隠れた感性・価値観を掘り起こすための有効な手段であると確認できた。

6.2.6 コンテキスト変化点抽出の検証

本論で定義した選択平均情報量を例外発言に限定せずに全ての発言に用いると、コンテキスト変化点を抽出できると考えられる。これを確かめるために、パラメータを $w = 50, \alpha = \beta = 1$ と設定して解析実験を行った。ここで、窓 w の値を大きく設定することで各発言の前後関係をより広く考慮することができるようになる。このとき、 w を大きくすると評価値が計算できる発言が少なくなってしまうことに注意する必要がある。解析結果の評価上位 10 発言を表 29 に示す。ここでも記者 1 の発言は前の解析と同様の理由により抽出しなかった。

表 29: 選択平均情報量のみでの解析結果

| ランク | 発言者 | 発言内容 |
|-----|------|---|
| 1 | 患者 | 5 に入院しています。14 日が植え込み手術ですから 9 日間ありました。 |
| 2 | 記者 2 | アブレーションはけっこう今は簡単にできるのですか。 |
| 3 | 患者 | 焼くというのはそれほど大したことはないかと思ったら意外と痛いのです。痛いというのは本来は熱いだけですが、熱さを痛みで感じるのです。 |
| 4 | 記者 2 | もうデータも事前にありますし。 |
| 5 | 患者 | 不整脈があるからです。
ペースメーカーはそういう刺激を与えてくれる機能もありますが、早いうちにずっと心電図をつけているのと一緒なので ... |
| 6 | 記者 2 | 肥大のほうは、今は特に問題とされていないのですか。 |
| 7 | 患者 | 荒れますね。わたしのところも 1 回荒れかけました。 |
| 8 | 記者 2 | 入れたあとのアブレーションはそれなりに効果はありますか。 |
| 9 | 患者 | 結果的にあのときは 3 日、4 日ぐらいの話です。 |
| 10 | 患者 | 鼻をつまむのを忘れます。 |

表 29 の結果を評価するために、6.2.5 で比較手法として利用した IDM 法で抽出した結果 (表 28) をここでも利用する。IDM 法で抽出された「アブレーション」は抽出発言に含まれており、「ブログ」や「病院」といったキーワードに関しては、これと関係する発言が抽出された。また、「自分」というキーワードは本手法では注目すべき単語として抽出しなかった。しかし「自分」というキーワードは、自分に関する話を始めたところであると考えられるため、コンテキスト変化を促進した単語であるといえるが、これは発言者 (ここでは患者) の口癖のようなものであって重要

な単語であるとは一概に言えない．また「ペースメーカー」や「友の会」という単語は，会話内容を要約するためには重要であるがコンテキスト変化点としてはそれほど重要ではない．一方本手法の結果(表 29)では，入退院の話に話題が展開する発言や「アブレーション」という手術に関する話題の導入部分が抽出でき，コンテキスト変化点の抽出手法としても利用できることが確認できた．

ここで，コンテキスト変化点の抽出を目的の一つとするためには，本手法ではいくつかの課題が存在する．ここでは窓を設けているために前後 50 発言以外は全く考慮していないため，比較的話題展開が早い会話でないと解析精度が低くなってしまおうと考えられる．ここで解析対象としたデータはインタビュー形式であったために話題展開が早いため，本手法でも精度良くコンテキスト変化点が抽出できたと言える．ここから，本手法でのコンテキスト変化点抽出の限界をみえてきた．また本手法では，例外発言の定量化のために注目発言前後の複数発言が存在しなければ計算することができない．したがって，解析対象全体におけるコンテキストを抽出することは困難であり，文書の序盤と終盤の数発言の定量化を行うことができない．但し，窓 w を可変値として自動で最適値を設定することができれば，計算可能範囲を大幅に拡張することができ，且話題展開が緩やかな会話の解析も可能となると考えられ，本アルゴリズムは改良の余地があると言える．本実験により，本研究の今後の課題の一つが明らかになった．

6.3 解析結果の評価

4.1 に示したように，本システムでは解析結果を発言者本人もしくは関係者にフィードバックして新たな会話ログデータを得てさらに解析をする，といったことの繰り返しでより深い知識を追究する必要がある．

しかし本実験では，抽出された例外発言からその発言の裏に潜む状況・物語を解析者が推測するというプロセスで解析を進めた．このため，得られた知見の評価が不十分であり，本論での解析は未完成となっている．本手法によって得られた結果の評価方法として，発言者本人や会話参加者，もしくは関連分野の専門家（例：患者インタビューであれば医師などの医療関係者）へのアンケートによって定量的に評価できると考えられるが，本論ではそうした実験段階までには至らなかったため，得られた知見の評価手法の確立は本研究の今後の課題の一つとなる．

7 結論

本論では、例外発言を「前後の文脈とずれている、会話の流れに沿っていない排除されてしまった発言」とであると定義し、会話データから例外発言を抽出・評価（定量化）するシステムを提案した。例外発言の定量化手法として選択平均情報量を定義して評価関数に利用し、「裏に状況・物語が含まれると推測される例外発言」とそうでないと推測される発言とをランキング形式で出力することによって注目されるべき発言がより注目されやすくなるように設計した。

予備実験としてオンラインチャット実験を行い、パラメータの設定方法を検証すると同時に、本手法によって例外発言の裏に潜む発言者の状況・物語を抽出するための有力な情報を得ることができることを確認した。また、地方自治体定例会の会議録の解析実験によって「会話コミュニケーションにおける人間関係を知る」ための有力な情報が得られることを示し、患者インタビューの解析実験によって「例外発言の裏に秘められている発言者の状況・物語」を掘り起こすことができると実証し、医療・保健分野への適用可能性を示した。

これらの実験を通して、周りから排除されてしまっている例外発言の中には、発言の裏に発言者の状況・物語を含むものが多く存在し、このような例外発言に注目することが今までにない新しい概念を導入するきっかけとなる可能性を持っており、適切なプロセスを経ることでこうした価値を掘り起こすことができることを示した。但し、このプロセスにおいて発言者本人へのインタビューの必要が生じる。

本論では例外発言の価値定量化を試みて、本研究の目的である「例外発言の裏に存在する状況・物語は例外発言に潜在する価値であり、適切に対応することで掘り起こすべきである」ことを実証するための第一歩とした。したがって、本研究は本論で完結するものではなく、例外発言の価値定量化手法も含めて、利用プロセスと得られた知見の評価手法等々を今後改良・発案していく。それと同時に、本論でも適用可能性を示した医療・保健分野などの様々な分野への応用を試みることを本研究の今後の課題とする、として本論の結びとする。

謝辞

本論文を作成するに当たり非常に多くの方にお世話になりました。まず第一に感謝の意を表したい方として、指導教員である大澤幸生先生を挙げさせていただきます。私が本大学院に入学する前の段階で、例外発言に注目してその研究に着手するきっかけを創って下さいました。本研究の本を辿ると、『「あほ」の価値をもっと発揮することはできないか』という問題提起でした。この問題を大澤先生から頂き、情報技術でもっと人間味のあることができないものかと常々思案していた私の意欲を奮い立たせ、且、日頃から持っていた「一見奇妙な意見でももっと取り入れられるべきものがあるのに」という思いとがシンクロして、「これこそ自分がやりたいと思っていた研究だ」ということに気付かせていただきました。また修士1年の夏に私に重篤な病気が見つかって手術やりハビリのために一年間の休学をしていた時には我が事のように心配していただき、入院中に時間を持て余していた時には入院中でも研究が進められるようにとノートPCの貸与をしていただいたり研究に関する議論をしていただいたりしたおかげで、入院中にも思考を止めることなく過ごすことができました。さらに復学後にも熱い指導をしていただき、研究室主催のワークショップのオーガナイザを務めさせていただいたりポート大会に誘っていただいたりして、退院後に「間違いなく復帰ができた」という自信を持つことができました。また、私が強く興味を持った医療に関する研究のために医療データを解析に利用できるようにしていただき、さらに本学医科学研究所で助教をされている田中祐次先生を紹介していただきました。

次に感謝の意を表したい方は、東京大学大学院工学系研究科大澤研究室助教の西原陽子先生と東京大学医科学研究所助教の田中祐次先生です。西原先生には、アルゴリズムに関して議論していただいたり参考文献を紹介していただいたりごとにアドバイスをいただきました。田中先生には、被験者になっていただける患者さんを紹介していただいたり医者ならではの視点からアドバイスをいただいたりと大変お世話になりました。また、患者インタビューに答えて下さってインタビューデータの解析をさせていただきました「日本心臓ペースメーカー友の会」福岡県支部北九州茶話会世話人の野口修司氏には、研究に協力していただけて大変感謝しております。加えて、貴重なプログラムの提供を快諾して下さった大阪大学大学院経

菅学研究科准教授の松村真宏先生にも大変お世話になりました。

また、大澤研究室の学生の方々への感謝も忘れられません。特に、同期として大学院に入学し、昨年度修了して今や社会人として精力的に働いている前田雄佐氏、前川知英氏、生井洋氏、長谷川賢氏、そして木村曜女史。半年間という短い期間しか研究室での生活を共にできなかったのは残念でしたが、それにも関わらず彼らのおかげで充実した毎日を過ごすことができました。その中でも特に前川知英氏は、自身の修士研究と就職活動で多忙にも関わらず入院中の私のところへ足繁く面会に来ていただいて、私がつらい闘病生活を乗り越えることができた大きな精神的支えになりました。加えて、今年一緒に修了できるであろう同研究室の修士2年の高市暁広氏、田中瑛一氏、田中良祐氏、安武貴世志氏、修士1年の岩永卓也氏、岡本憲典氏、木村諒史氏、小林正典氏、卒論生の北風勝人氏、高橋祐一氏、元研究生の吉田隆之氏、研究生の宮本勝男氏、博士課程の社会人学生の池田尚司氏、久代紀之氏、郷田慎一氏、堀江健一氏、中村潤氏、武田俊之氏、筑波大学大澤研究室卒業生でNECにお勤めの前野義晴氏、それから東京工業大学寺野研究室の越山修氏、慶應義塾大学政策・メディア研究科の伊藤貴一氏等々、非常にたくさんの方々にはアドバイスをいただいたり時には激しく議論したりと、切磋琢磨できる良い関係を築くことができ感謝しています。また、入院中に目が不自由で文字が読めなかった私に本を読み聞かせて下さるなど、様々な形でサポートして下さった秘書の足立佳奈女史にも大変お世話になりました。

最後になりましたが、常に暖かい目で見守ってくれている家族（父、母、姉、親戚の皆様、それからチワワのモカとミルク）にも、感謝を忘れることができません。特に母には、私の入院中に不慣れな東京で生活して毎日病院に面会に来ていただいたりと、どれだけ感謝をしても足りません。かく言う現在も謝辞を書きながらも涙をこらえきれないほどです。

こうして改めてお世話になった方々を挙げてみると本当に多くの方々にお世話になりました。ここでは書ききれませんが、陰ながら支えて下さった、大橋弘忠教授、吉村忍教授、菅野太郎准教授、昨年度退職された元システム創成学科事務

室の神崎典子女史等々，システム量子工学（システム創成学）専攻の教職員の方々にも大変お世話になりました．さらに，東京大学付属病院の医師・看護師の方々，特に東京大学医学部付属病院で私のために尽力して下さった執刀医で現在横浜市立大学大学院医学研究科教授の川原信隆先生，担当医の矢向今日子先生，田中將太先生，甲賀智之先生を始めとして多くの医師・看護師等の東大病院の医療従事者の方々に命を救っていただき，この感謝は言葉では表すことができません．また，入院中に知り合い同室で共に闘病生活を送った東京大学大学院生の山崎光弘氏，写真家の平山和充氏，茨城県にお住まいの東野亮氏，ソニーにお勤めの田中仁朗氏等々，たくさんの患者仲間とご家族との出会いもあり，こうした出会いも全て私の糧となりました．感謝の言葉が絶えません．

皆々様，心より感謝申し上げます．そして，今後ともよろしくお願い致します．

2009年2月

山口広樹

参考文献

- [1] 黒須俊夫, モバイル・インターネットの進展と青少年の情報行動 (1), 群馬大学社会情報学部研究論集, 第 15 巻, pp.279-304, 2008.
- [2] 参考 web ページ URL
<http://www2.cc.niigata-u.ac.jp/c40804/dzemi2000/kc799.html>
- [3] 社団法人私立大学情報教育協会, 私立大学教員による情報機器を利用した授業改善に関する調査の報告, 平成十三年度報告, 2002.
- [4] 深谷昌志監修, 電子メディアの中の高校生, モノグラフ・高校生 VOL.63, ベネッセ教育研究所, 2001.
- [5] 木立るり子, 五十嵐世津子, 一戸とも子, ほか 12 名, 看護系学生 (生徒) における携帯電話の利用とメリット・デメリットへの意識, 弘前大学医学部保健学科紀要, 第 6 巻, pp.65-76, 2007.
- [6] テリー・ウィノグラード, フェルナンド・フローレス著, 平賀譲訳, コンピュータと認知を理解する: 人工知能の限界と新しい設計理念, 産業図書, 1989.
- [7] サム・ウィリアムズ著, 本田成親訳, 人工知能のパラドックス: コンピュータ世界の夢と現実, 工学図書, 2004.
- [8] 戸田正直, 感情, 認知科学選書 24, 東京大学出版会, 1992.
- [9] 長沢伸也, 感性工学の基礎と現状, ファジィ学会誌, Vol.10, No.4, pp.647-661, 1998.
- [10] 目良和也, 市村匠, 相沢輝昭, 山下利之, 語の好感度に基づく自然言語からの情緒生起手法, 人工知能学会論文誌, 17 巻, 3 号 A, 2002.
- [11] Ning Zhong, Y. Yao, Web Intelligence(WI), Proceedings 24th Annual Computer Software and Applications Conference. COMPSAC2000, 2000.
- [12] 武田英明, 知性のネットワークとしての WWW, 人工知能学会誌, Vol.17, No.3, pp.346-351, 2002.
- [13] 志良堂哲也, 永山格, バッティングフォームの感性定量化分析, 電子情報通信学会技術研究報告 HIP, ヒューマン情報処理, Vol.99, Issue.451, pp.37-42, 1999.
- [14] 中山紀男, 大倉典子, 感性情報を用いたテキスト分類手法, 電子情報通史学会総合大会後援論文集, Vol.2003, p.269, 2003.

- [15] 鈴木英之進, 志村正道, 情報理論的手法を用いたデータベースからの例外的知識の発見, 人工知能学会誌, Vol.12, No.2, pp.305-312, 1997.
- [16] 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.
- [17] Flaviu Cristian, Exception Handling and Sftware Fault Tolerance, IEEE TRANSSACTION ON COMPUTERS, Vol.C-31, No.6, 1982.
- [18] Claus Hagen, and Gustavo Alonso, Exception Handling in Workflow Management Systems, IEEE TRANSSACTION ON ENGINEERING, Vol.26, No.10, 2000.
- [19] 前川知英, 大澤幸生, 普通でない人の化学～変わり者は社会・組織でいかに振舞うのか?～, 東京大学大学院工学系研究科システム量子工学専攻 修士論文, 2007.
- [20] 山西健司, データ・テキストマイニングの最新動向 - 外れ値検出と評判分析を例に -, 応用数理, vol.12, No.4, pp.341-356, 2002.
- [21] C.C. Aggarwal and P.S. Yu, Outlier Detection for High Dimensional Data, Proc. ACM SIGMOD 2001, May, Santa Barbara, pp.37-46, 2001.
- [22] 井上隆二, 山下富美代, 図解雑学 社会心理学 (図解雑学シリーズ), ナツメ社, 2000.
- [23] 新版 災害と日本人 巨大地震の社会心理, 時事通信社, 1995.
- [24] 松山章子, 健康開発と社会人類学・医療人類学, 長崎大学熱帯医学研究所共同研究報告書, Vol.18, page 33, 2006.
- [25] 川田順造, 文化人類学とは何か, 日本文化人類学会, Vol.71, No.3, pp.311-346, 2006.
- [26] 小倉能理子, ほか5名, 看護学生のコミュニケーションの実態 教育系学生との比較, 弘前大学医学部保健学科紀要, 第6巻, pp.31-39, 2007.
- [27] 斎藤久美子, ほか5名, 弘前大学医学部保健学科紀要, 第6巻, pp.41-46, 2007.
- [28] 保田卓, 教育におけるコミュニケーションの象徴的一般化:「ライフコース」メディアと「媒介可能/媒介不能」コード (II-7 教育とコミュニケーション), 日本教育社会学会大会発表要旨集録, No.55, pp.140-141, 2003.

- [29] 糸魚川幸宏, 場面における教育コミュニケーションの特質分析, 情報処理学会第46回(平成5年前期)全国大会, pp.1-33, 1993.
- [30] 住大恭康, 医師・患者コミュニケーションの諸相, 医事学研究, Vol.20 pp.1-24, 2005.
- [31] 山田隆文, 歯科医療コミュニケーション-インフォームドコンセプトの先にあるもの-, 明倫歯科保険技工学雑誌, vol.4, issue.1, pp.109, 2001.
- [32] 南部美砂子, 原田悦子, 須藤智, 重森雅嘉, 内田香織, 医療現場におけるリスク共有コミュニケーション-看護師を中心とした対話データの収集と分析-, 認知科学, Vol.13, No.1, pp.62-79, 2006.
- [33] 神山祐一, 平野靖, 梶田将司, 間瀬健二, 勝山貴美子, 山内一信, 話題構造の可視化による医師-患者コミュニケーション支援手法, 第7回看護情報研究会論文集, pp.67-70, 2006.
- [34] 大西淳, 要求工学, 共立出版, 2002.
- [35] 大西淳, 要求工学の動向, ソフトウェア工学, Vol.1137-10, pp.65-72, 2002.
- [36] 土井晃一, 要求獲得オフライン法での質的要求の抽出法, 情報処理学会論文誌, Vol.40, No.11, pp.4012-4020, 1999.
- [37] 土井晃一, 要求獲得会議を分析することによるユーザ指向要求獲得法, 情報処理学会論文誌, Vol.44, No.1, pp.48-57, 2003.
- [38] G. Salton, A. Wong, and C. S. Yang, A Vector Space Model for Automatic Indexing, Communication of the ACM, Vol.18, No.11, pp613-620, 1975.
- [39] G. Salton and M.J. McGill, Introduction to Modern Information Retrieval, McGraw-Hill, 1983.
- [40] R. Barzilay, M. Elhadad, Using lexical chains for text summarization, Advances in Automatic Text Summarization, pp.1-12, The MIT Press, London, 1999.
- [41] 相良直樹, 砂山渡, 谷内田正彦, サブピックを考慮した重要文抽出による報知的要約生成, 電子情報通信学会論文誌 Vol.J90-D No.2 pp.427-440,2007.
- [42] 住田一男, 知野哲郎, 小野顕司, 三池誠司, 文書構造解析に基づく自動抄録生成と検索提示機能としての評価, 電子情報通信学会論文誌 D Vol.J78-D2, No.3, pp.511-519, 1995.

- [43] 横山憲司, 難波英嗣, 奥村学, Support Vector Machine を用いた談話構造解析. 情報処理学会自然言語処理研究会 NL-155, pp. 193-200, 2003.
- [44] 大澤幸生, N. E. Benson, 谷内田正彦, KeyGraph : 語の共起グラフの分割・統合によるキーワード抽出, 電子情報通信学会論文誌, J82-D-1, No.2, pp.391-400, 1999.
- [45] 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.
- [46] 中島宏ほか 6 名, 人間-機械協調システムにおける社会的知性, 人工知能学会論文誌, 19 巻, 3 号, 2004.
- [47] 鳥居大祐, 石田亨, ブスケフランソワ, 参加型アプローチによる交渉モデルの獲得, 人工知能学会論文誌, 21 巻, 3 号 H, 2006.
- [48] 平田高志, 村上晴美, 西田豊明, 連想表現と分身エージェントを用いたコミュニティにおける知識共有支援, 人工知能学会論文誌, 16 巻, 2 号, pp.225-233, 2001.
- [49] 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 17 巻 4 号 F, 2002 年.
- [50] M.R. Quillian, Semantic Memory, Semantic information processing, MIT Press, pp.227-270, 1968.
- [51] A.M. Collins and E.F. Loftus, A Spreading Activation Theory of Semantic Processing, Psychological Review, 82, pp.407-428, 1975.
- [52] R.F. Lorch, Priming and searching processes in semantic memory, A test of three models of spreading activation, Journal of Verbal Learning and Verbal Behavior, 21, pp.468-492, 1982.
- [53] 形態素解析ソフト 『茶筌 ver.2.3.3』
<http://chasen-legacy.sourceforge.jp/> .
- [54] M.A.Hearst, TextTiling: Segmenting Text into multi-paragraph Subtopic Passages, Computational Linguistics, Vol.3, Issue.1, pp.33-64, 1997.
- [55] 別所克人, クラスタ内変動最小基準に基づくテキストセグメンテーション, 情報処理学会論文誌, 47(3), pp.957-967, 2006.
- [56] 但馬康宏, 北出大蔵, ほか 4 名, 発話を出力単位とする HMM による対話に対する話題分割, 情報処理学会研究報告, 自然言語処理研究会報告, Vol.2007, issue76, pp.7-10, 2007.

- [57] 宮崎県議会会議録検索システム
http://www.kaigiroku.net/pref_miyazaki/gikai.html/ .
- [58] 大阪府議会会議録検索システム
<http://www.pref.osaka.jp/gikai/discuss/index.html/> .
- [59] Stephen Robertson, Understanding Inverse Document Frequency, Journal of Documentation 60 no. 5, pp.503-520,1997.

付録

A オンラインチャット実験 会話ログ

チャット実験の会話ログ．発言者名と発言内容を表示．

A.1 新ゲーム考案

『10分間で楽しめる新しいゲームを考案する』というテーマで行ったチャット実験の会話ログ

全 128 発言．40 分程度．

N サーバーを介してやるゲームですか？

P いや、神と鉛筆でできるようなゲームです

P 休み時間にやるような

I デスノート？> 神と鉛筆

P あ、神 紙

K 囲碁っすね。6路盤ならすぐに決着がつきます。

N コックリさん

K あ、新しくないや(汗)

I 古今東西ならここでも出来ますよ。

N 新しいゲームを考えるんですか

P Iありますよね。ローカルなゲーム。三将棋とか

K 紙相撲みたいな？

P はい。新しいゲームです

K いっぱい出して、後で組み合わせましょう。

K イノベーションゲーム的に

P じゃあ。数字当て

N あいうえお作文

W いろいろあるよね、じゃんけんみたいに指の形とか使うやつとか、紙に ×とか数字書くやつとか

I 戦艦ゲーム

K なんすかそれ> 数字当て？

W それ以外に系統あるかな？

I いっせいのーで(数字)とか

W 数字当てもあるし、4目並べとかも

I しりとりとか、口だけでやるやつもある。

K 俺、小学校の頃迷路すごろくっていうの作ってたよ

P 数字当ては、二人で各4つ数字を決めておいて、それを当ててくやつ

K ボードか紙か口頭かアクション(身振り)を伴うものか

W なんすかそれ?おれ小1のとき、授業参観用の絵で俺一人だけ迷路書いていて親に笑われました

Z 「利き緑茶」なんてどうでしょう?

K 数字は0~9とか? > Pさん

K すごろくが迷路になってます。道を間違えるとゴールできません > Wさん

Z じゃあ僕は6で

P はい。数字を言いあって、ヒットとかストライクとか

N 迷路面白そうです

W それやらしいですね > Kさん

Z ヒットエンドら~ん

P 昔、ものすごく細かい間違い探し作ったことがあります。

N 小学校の時は、物語を作って人に聞かせてました

P でも10分じゃできない

K なんかただの確率のゲームのような > Pさん

W いや、数字当ては頭使いますよ!

P そうですよ。でもそこに子供ながらの純粋さが生きるわけですよ

K 緑茶を用意するのがめんどいっすね > 利き緑茶

P これは、、イエモンですね という感じでしょうか

W あれ?おれのところで流行った数字当てと違うのかな?

K え~。性格は出そうですけど…。 > Wさん

K 物語は厳しいっすね。面白いけど。 > Nさん

Z 「すっぼん」「ゲームボーイ」「竹刀」を使い、完結する物語を作って披露しあうゲーム

P そういえば、サークルの合宿で自作すごろくやりましたよ

K 予め考えとかないと無理っすね。

P マスには乾杯とかばかり

W それやりました! > Pさん

P あれしんどいのね かなり

W 男飲むとか、1年飲むとか、全員飲むとか

P 何度記憶を亡くしたとか

K 出た目の数だけ飲むとかやったことある(汗) > Pさん

P じゃなくてえ 10分で楽しめる健全な新しいゲームを考えましょう

Z 10分間じっと正座 合図の笛と同時に50m走 足のしびれに耐えて一番でゴールした人の勝ち

W でもこれ10分では記憶無くなるまで楽しめませんね～>Pさん
Z テニサー自重w>W
W それやったら10分間ぐるぐるまわって50m走とかどうですか?>Zさん
N 死ぬって>W
K 研究室でやるんだったら、ランキングとかつくと面白いですね。何回か連続で負けると罰ゲームがあってもいいかも
W ばれたか>Zさん
P やっぱここは、小学生の心に戻って、有意義な休み時間を夢見ましょう
K あ、身体使う系だ。新しい>Zさん
W 絵でしりとりとか
Z 人生的な意味で連敗中なので
Z 甘んじてバツゲームをうけます
K 数独とかもいいっすね。
W それ小学生やります?>Kさん
P やっぱみんなで遊ぼうよ
P 世界のナベアツゲームgとか
W 覚えてしりとり!
P 3の倍数と3のつく数字で「あほ」になります
K あれ? 小学生前提なんすか?>Wさん
N 10分でテンション上げるのしんどそう>ナベアツ
W 自然とテンションあがるのがいいっすね～>にしさん
Z やべ、テンション上がってきたッ!!!
W やっぱ動きある感じのゲームがよくないっすか?
P リズムゲームとかあったよね
I 名前忘れたけど、テーマに対してリアクションが違うやつが負けるげーむがあった。
N リアル鬼ごっこ。90人で10人を追いかける
P やっぱりコンパゲームだけど
K いや、自分のペースでゆっくり楽しめるのがいいな～>Wさん
I テーマ「野球」に対してピッチャーの真似するかバッターの真似するかで分かれたり。
I で、みんな同じリアクションならテーマ出したやつが巻け。
W やりましたそれ!>Iさん
Z 名シーン再現ゲーム。誰が一番リアルに、漫画、スポーツ、演説等の名シーンを再現できたか競う。「きもていーーー!!!」
P クリリンのことかあああああ
Z 種田のフォームははガチw>Iさん
W 蛇眼の力をなめるなよ!
N 身体動作を使うのはおもしろそうであります

I 種田なんかをテーマにしたやついないよ。
W じゃの時まちがった～
W 今度じの字間違った～
W そうですね！> Nさん
I オレンジ多すぎるから色変えました。
K どうやって競うの？ お題出す感じかな。> 名シーン再現
Z オレンジ多すぎるから色変えました。
P どっきりしかけて、リアクション王を決定する。。。身体使いますよ
W 逆にペア対抗とかにして、自分のペアが何を再現しているのかを当てるとか。
W んで、間違ったらサドンデス!!> Kさん
W 赤にしよ～
K ジャッジがいるよな。> Wさん
W それはやる人があらかじめ紙にお題(答え)を書いておけばいいんじゃないっすか？> Kさん
P お題は誰が決めるの？
W お題イズフリーダム
P お題のブルース？
K ん～。どうだろ。やる人とお題出す人は別の方がいいんじゃないか？> Wさん
K まあ、見る人が採点する形式ならどっちでもいいか
P じゃあやっぱ相手チームだ
P お題だすの
W まあ勝ち負けより
W ずっとやる方が楽しくないっすか？
P そろそろ時間ですね
K 確かに。> Wさん
N そっすね> Wさん
K じゃあ、適当にお題出してって、やる人はどんどんやってくのがいいか
K で、ミスったら交代とか。そんな感じかな
P 続きをまた後でやりましょう。今回のフィードバック付で
W そうっすね。一回ごとにターンチェンジとかでもいいと思います。> Kさん
P それでは
W お疲れーライス。
Z だが断る
K お疲れ様でした
K まとまりませんでしたね。。
<終了>

A.2 食事をどうするか

研究室のミーティング後『どこに食事に行くか』という雑談.
全 139 発言 . 30 分程度 .

- T 僕肉食べたいです
T ジンギスカンどこ？
X バンビにしますか。
N 焼酎なら明日に残りませんよ
Y 確かに焼酎は残らないから、それでいきましょう
G モンゴルです
X 後樂園のジンギスカン？
K 呑みたいという重いよりも食欲の方が勝ってます G 後樂園のジンギスカン知りません
N ジンギスカン食べれるのか
T 別に連続バンビでもいいですよ。。
G 肉ならなんでもよいです。バンビでもせんごくでも。
X チンギスハンじゃないほうのジンギスカンですよ
Y ジンギスカンないよね バンビ
Y あ そうか チンギスハンか
T あ、
G あ、
Y おろ
T 今さらですが、バイクなんで酒は全く飲めないです
T すいません
G 残念です
X バイクじゃなくても飲まないくせに
Y バイクはエンジンをストップして引けば歩行者よ T
T いっつも、本郷でごはん食べようとする
T 迷うんですよね。。
G じゃあ、僕が代りに乗って帰りますので、Tさんはガンガン飲んでください！
Y 本郷では、よくまぐろ市場行って海鮮丼
T いや、たべようよ
T 海鮮丼いいですねー
K 自分もたまに行きます まぐろ市場
Y いや、呑もうよ
T マグロ市場のスタンプがたまってきました

K 飲みと食べが拮抗してますね
G 飲みで
G いや、食べで
Y じゃあじゃんけんできめよう
K スタンプなくしましたw
T 最初はゲー !!
N 東京のマグロはおいしいですね
T じゃんけん !!
N ばー
K チョキ
G だんしんゲー !
Y はい
Y チョキ
T あいこでした
Y あーいこーで
K あいこでー !!
G すとっきんゲー
Y はい多数決とろう
X パー
T 食べる方に一票
Y 呑み 1
G 帰って、部屋の隅で膝をかかえて泣く、に一票
Y えと、今何人いる ?
K 食べも呑みも今日いけないです ;; すいません~
X 1
G 2
Y さぁあん !?
T 4
Y 5
K ちと厳しいッス。申しわけなし
N 今日は帰ります
Y あれ ? 6 担当のあほは ?
T 実はおれも帰って自炊する、と揺れてます
Y いま家には米粒一つ残っていないんですわ
G 帰って、全てに絶望する、と揺れています
T それはおきのどくです
G 夢も希望も、一粒も残ってません
T A は ?

K モンゴルさん大変ですね
Y A はどれ？
X 何が？
Y 名前ヌルとか可能なんだ
X 僕が A ですよ。
T どうしたいかな～と
X 飲むなら軽く。
Y じゃあ根津の別荘でパーティ
G バンビで何時まであいてるんですか？
K べ、別荘？
T A のいえ？
G 僕は、谷中の別荘のほうがいいです。
T ばんびはわからない。。
Y バンビ説 再浮上 G あ、やっぱり根津の別荘がいいです
Y 結構遅くまでやってるんじゃない？
X 回文になってないよ。T
Y 知らないけど
G モンゴルゴンモ
T 回文？
Y A はこえれから回文しばりね
X わ
G こばやしヤバ子
X か
T 回文てなに？
G えいいち？いいえ
T あ！
T そういうことか！
K 留守に何する
T えいいちのいえ
G 空腹です
T そろそろ行きましょう
G いかんせん空腹です
Y とりあえず生で
G ぼくは八海山
G お通しはカットで
Y 里いく？
Y もしくは朝倉
G 里てなんですか？

G 朝倉でなんですか？
Y 根津の飲み屋
K お通しカット好きだなー
G 食べメインがいいです。すいません。
Y お通しだけで3杯はいける
K それってお通しがすごいんですか？ww
Y よし、じゃあ間にとって食べメインでさくら水産
T 家で卵かけごはん食べます
G ぼくはさく水でおKです
G あ、でもTさんが行かないなら行かないです
Y ごはんセットがあるじゃないか
K さくすいクーポンつかえば1杯無料ですよ
G たしかに
Y そんなのあるんだ クーポン
G こころで決めましょう
Y だな
G 行くのは何人ですか？
Y さくらかバンビでOK？
X 1
G 2
T バンビなら3
Y 3
Y あ4
G ばんび 行きましょう
T 行こう
Y バンビ行こうか ビールもあるし
Y よし、終わり
T おつかれさまでした
G 失礼します
K おつかれさまです
<終了>

A.3 簡易型イノベーションゲーム

大澤研究室考案の「イノベーションゲーム」という組み合わせ発想ゲームの一部
分(二つのアイデアから新たなアイデアを創発する)を「簡易型イノベーションゲー
ム」としてオンラインチャットで行ったもの。使用した二つのアイデアは図14に示

したものである。

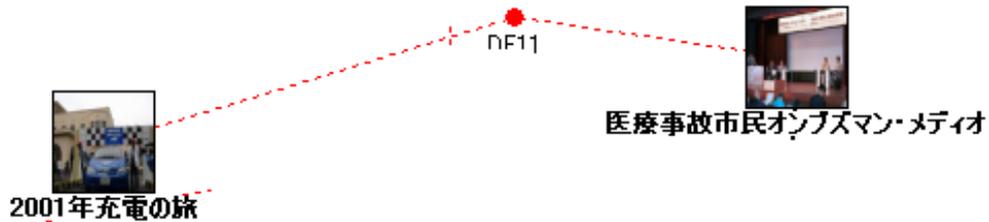


図 14: 組み合わせ発想ゲーム 使用カード

全 84 発言 . 60 分程度 .

D ダミーノードとはなんですか？

M これ真面目に答えていいの？

DAN はい じゃあ M さん答えちゃって

M 小職もあまりよく理解していませんが、ダミーノードとは、keyGraph に仮想のノードを設定することです。

DAN うん、もしかしたらそこに隠れたノードがあるかもってやつ

DAN このテーマの意図としては、

M 仮想ノードを設定することにより、例えば黒幕を見つけるだとかいう技法だと思います。前野さんが完成させたものですね。

DAN 二つの技術から発想できるのをみつけちゃおうってこと

DAN ダミーノードって何かについてはだいたいわかった？

D M さん、おイモさん、ありがとうございました。

M これだけじゃ、わかんないんじゃない？

DAN 1 イモ？

M 1 イモって何ですか？

DAN 今何人参加してる？

動物 人じゃないです

DAN 1 はただのうちのまちがい

DAN じゃあ何人と何匹？

M 了解

D 背番号は 11 です

M 二刀流です。

DAN そろそろ本題に入りたいんだが

M まだ、本題じゃなかったの？

D "2008" 年医療事故市民オンブズマン・メディアの旅

DAN 充電の旅と医療オンブズマンの間にきそうなコンセプトとかない？
DAN 中身がないよ D
D すいません。適当すぎました。
DAN あれ？意外といいかも 2008 年医療事故市民オンブズマン・メディアの旅
D すいません。いきなり結論を出してしまいました。
DAN 旅しながら広めて感じ j
DAN 広めて感じ j 広めてく感じ
M なるほどね~いいかも 旅しながら広めて感じ
DAN j がないよ
D そうです。医療事故市民オンブズマンメディアのこと知りませんでした。
DAN でもこれって、病院への過剰な訴訟とかにつながってるんだよね
M 市民ネットワークというのが何かキーワードになりそうだね。
DAN それだな。市民ネットワーク
M でもこういうきっかけが重要なんじゃないかな でもこれって、病院への過剰な訴訟とかにつながってるんだよね
D 市民「コンセンネットワーク」ってなんですか？
DAN インターネットじゃなくて実際に旅先で逢うってのがいいな
DAN コンセントだから、とくかく繋がればいいんじゃない？
M ちなみにこの 2001 年充電の旅というのは、実は自分も関わりがあります。自動車評論家の館内さん主催の日本 EV クラブという団体があって、その団体は車と地球環境問題について考えるというような団体なんです。自分も所属していて自作 EV 作りました。
DAN EV？ EV て何？卑猥な響き
M Electrical Vehicle ようは電気自動車のこと。
DAN そか電気自動車か
D え？結局、Ero Voice の略でいいんですよね？
DAN でも”Ero”って英語で通じるのか？
M DV じゃなくて、EV はそろそろ置いといて、2つのノードをつなげるバシッとしたノードないかな。
DAN 2008 年医療事故市民オンブズマン・メディアの旅
D EV で全国の病院を回って、コミュニケーション&情報収集の旅
DAN やっぱこれだな
D どうしても「旅」という発想から離れられないですね
DAN それはオレの今後の研究に使えるようなデータだな D
D Ero Voice 学のすすめ
DAN すすめられない。いただけない
M Ero は置いといて、ちょっと真面目に書きます。車は、自動車業界という巨大な組織があります。医療においても白い巨塔といわれる強力な組織があります。それ

らに対する市民の訴えというか反攻みたいなものが見えますね。
DAN 反骨精神。反逆の美学
M この二つを結びつけたら、どんなビジネスが生まれますかね。
DAN 旅しながらひたすら訴訟起こして病院から金ふんだくる
M 停電による医療事故防止機能付き EV
M DAN と武蔵の漫才になってるよ。
DAN 車じゃなくて普通の発電機でいいじゃん
DAN どっちがぼけでどっちがつつこみ？
動物 医療事故防止機能付き Ero Voice
DAN また Ero に戻っちゃったよ
M EV 付き医療防止機能
DAN 無理やりだな
D 巨大自動車業界で EV 開発にいそしむ明石屋さんまと、メディアで全国の被害者
をとりまとめる大竹しのぶが恋に落ちる月9ドラマ。
DAN そういえば昔。イノベーションゲームのことを頭のいいこじつけだと言ってた
奴がいたけど これじゃあ頭の悪いこじつけだな
M なかなか斬新だと思うよ D 巨大自動車業界
DAN なんでさんまが自動車業界？
M 市民団体統括株式会社 市民活動の支援やメンバー集めをしたりする。会費で運
営。
D EV に乗ってブルペンからマウンドに向かうダルビッシュ有。ドラゴン桜2で、オ
ンブズマンメディアに参加する役を演じる、妻サエコ。
M とところでオンブズマン・メディアとは何？
DAN そろそろ終わりにしたいんで、結論を
M 違う方向に展開しようと思ったんだけど。
DAN オンブズマンは 国民に代わって行政活動を監視し、また国民からの行政機関
に対する苦情を処理する者。スウェーデンに始まる。行政監査専門員。
M 結論：バシッとくるダミーノードは見つからなかったけど、これ方法とてもイノ
ベティブな感じがするよ。
DAN じゃあさ。こうしよう。今日の総括を D にしてもらおう
M これ方法 この方法
D 支配人 D さん、さようなら
<終了>

B 患者インタビューデータ（冒頭部分）

患者インタビューの冒頭部分．全部で1時間半から2時間程度．全362発言．

記者1：最初、すごく大ざっぱに言うと2005年の健康診断と書かれていましたので、今から3年くらい前の会社の健康診断で指摘されたところから気づかれて、今言われたように実際にペースメーカーの埋め込みをされたのが2006年だから2年前。

患者：そうです。

記者1：その後現在に至っているということでございます。

患者：そういう意味ではまだ新米です。

記者1：もう少しで丸2年。その間、あとでたぶんマラソンの話がいっぱい出てくるとは思いますが、一貫して続けておられるということですね。

患者：そうですね。それを一つの目標にやってきました。

記者2：これはずっとやっていたのですか。

患者：もともと走っていましたが、2005年に引かかる前に自分では年を取ったのか、最近は調子悪いと思っていたくらいです。それがこの間そこそこ走れたので、この病気が原因だったということがよく分かりました。ペースメーカーはどちらかというと心房粗動のものです。

記者1：そうしましたら書いていただいた順番に重複するところもあると思いますが、伺ってまいります。

患者：うそを書いているところは直してください（笑）。

以下省略
