

# ファセット分解と公的討議の談話分析

京都大学 鄭 蝦榮<sup>\*1</sup>  
 京都大学 小林潔司<sup>\*2</sup>  
 東京工業大学 羽鳥剛史<sup>\*3</sup>  
 名古屋工業大学 白松 俊<sup>\*4</sup>

By Hayeong JEONG, Kiyoshi KOBAYASHI, Tsuyoshi HATORI, and Shun SHIRAMATSU

近年、公共プロジェクトの構想・計画段階においてパブリック・インボルブメントが導入され、様々な立場のステークホルダーの間でプロジェクトがもたらす影響や効果について議論する公的討議の場が設定されている。本研究では、利害関係者の間でコミュニケーションが実施された公的討議に関する議事録(1次コーパス)を対象として、討議参加者の発話内容の関連や討議参加者間の関心の不一致の程度について客観的に把握するための方法論を提案する。すなわち、公的討議による1次コーパスをファセット分解することにより、発話の意味内容を言語情報化した2次コーパスを作成する。このような2次コーパスを対象として、談話分析手法を用いて討議参加者の発話内容の類似性や討議における対立の構造を定量的に把握することが可能となる。本研究では、現実の公的討議事例を対象として、討議参加者間の関心不一致度の計量化や意見対立の構造を可視化することにより、本研究で提案した方法論の有効性を検討する。

【キーワード】 ファセット分解、パブリック・インボルブメント、談話分析

## 1. はじめに

近年、公共プロジェクトの構想・計画段階において、専門家、市民、企業や行政等、プロジェクトに関わる様々な関係主体による公的討議を行うことの重要性が指摘されている。多くのプロジェクトにおいて、パブリック・インボルブメントが導入されている。その中で有識者委員会や流域委員会等の第3者委員会が設置され、プロジェクトの是非に関わる様々な議論が実施されるようになった。

公共プロジェクトに関わる意思決定において、このような公的討議を設けることの意義は多様であるが、専門家や市民による討議を通じて、プロジェクトの効果や望ましさに関わる判断情報を得ることが期待されている。しかし、多様な利害関心を有する

討議参加者は、公共プロジェクトの効果や費用に対して様々な期待や信念を有しており、討議参加者の間でプロジェクトに対する意見や判断が異なる場合が少くない。公共プロジェクトに対する利害の対立や関心が一致しない場合、公的討議を導入したとしても、討議参加者が自分の望ましいと考える結論を導出しようとする意図的、戦略的発言を繰り返し、実質的なコミュニケーションが成立しない可能性がある。さらに、こうした状況において、公的討議の適切な進め方を誤ると、意見が対立したまま、議論が膠着することが危惧される。

近年の情報技術の進展に伴い、言語学研究の知見を応用し、現実の対話データにおける関連情報を抽出・整理することによって、討議参加者間の円滑な

\*1 工学研究科都市社会工学専攻 075-383-3224 hayeong@hse.gcoe.kyoto-u.ac.jp

\*2 経営管理大学院経営管理講座 075-383-3222 kkoba@psa.mbox.media.kyoto-u.ac.jp

\*3 理工学研究科土木工学専攻 03-5734-3578 hatori@plan.cv.titech.ac.jp

\*4 工学研究科情報工学専攻 052-735-5129 siramatu@nitech.ac.jp

議論を支援するための様々な情報システム<sup>1)</sup>が開発されてきた。しかし、パブリック・インボルブメントに関する既往研究において、これらの知見が十分に生かされているとは言い難く、討議参加者の発話内容についての体系的な整理や参加者間の関心の不一致について十分な分析を行った研究は少ない。このような基礎的検討がなされないまま、パブリック・インボルブメントが実践されている場合も少なくない。公的討議を適切に進めていく上で、討議参加者の発話内容や関心の不一致について的確に把握するとともに、これらの情報に基づいて討議におけるファシリテーションを高度化することが重要である。

本研究では、公共プロジェクトを対象とした公的討議に対して談話分析を実施し、討議参加者のプロジェクトに対する関心構造の差異や意見の不一致性を分析するための方法論を提案する。そのために、討議参加者の発話内容をファセット分解<sup>2)</sup>することにより言語情報化した2次コーパスを作成する。さらに、2次コーパスに基づいて、討議参加者のプロジェクトに対する態度の類型化、討議参加者間の関心の不一致度の計量化、意見対立状況の可視化を試みる。以下、2. では本研究の基本的な考え方を説明する。3. では発話コーパスのファセット分解、4. において、2次コーパスに基づいた談話分析を試みる。

## 2. 本研究の基本的な考え方

### (1) 既往文献と内容分析の重要性

近年の情報技術の進展に伴い、現実の対話データにおける関連情報を抽出・整理することにより、討議参加者間の円滑な討議を支援するための様々な情報システムが開発されている<sup>1)</sup>。既存の研究の多くは、討議に用いられた単語の出現頻度に着目した言語統計分析に終始しており、発話の意味内容を対象とした分析事例はほとんどないのが実情である。

公共プロジェクトを対象とした討議過程においては、討議参加者の間で多様な形態のコミュニケーションが実施される。参加者は、多様な認識体系を有しており、自分の主観的な関心に基づいて、意見を述べ、他者の発話を理解する。このような公的討議を分析するために、記述的な質的データに対する内容分析(content analysis)が求められる<sup>3)</sup>。Stone

は、内容分析を「テクストにおける或る特定の特徴を、体系的にかつ客観的に同定することにより、推論を行う調査技法」と定義している<sup>4)</sup>。内容分析は、例えば速記録のように、現実に実施された討議内容を対象として、対話データに組み込まれた文脈との関連性を維持しながら、発話内容を分析することを目的とする。Krippendorffは内容分析の特徴として、1) 非露見性、2) 非構造性、3) 文脈感応性の3つを挙げている<sup>5)</sup>。第1に、アンケート調査やヒアリング調査においては、調査対象者が調査内容に過度に反応することより、調査結果の妥当性が損なわれる可能性がある。内容分析は調査対象者に気づかれにくい手法であり、調査における露見性の問題を回避できる。第2に、内容分析は、あらかじめ構造化されていないデータを分析対象とする。このため、予期できない事象に対しても分析可能である。第3に、対象事例を事前に構造化しないため、文脈に即応することが可能である。内容分析は調査対象者間のコミュニケーション過程をシンボリックな現象として扱い、発話者による意味解釈を推論することにより、文脈との親和性を維持することができる。

### (2) 談話分析の目的

公共プロジェクトを巡る公的討議においては、多様な価値観や関心を有する参加者の間で議論が展開される。本研究では、公的討議の速記録(1次コーパス)を対象として内容分析の1手法である談話分析を実施し、討議の内容やその構造を明確化する方法論を提案する<sup>5), 6)</sup>。談話とは、一般に「文よりも大きい言語単位で、あるまとまりをもって展開した文の集合」<sup>5)</sup>を意味するが、談話分析は、社会的文脈において「実際に使用されている言語」を用いて分析するという点に特徴がある<sup>7)</sup>。Schiffrinは、談話は、「特定の文脈における文の使用」を意味する発話(utterance)で構成されると定義している<sup>8)</sup>。ここで、文脈(context)とは「言語単位の意味解釈に必要な、その言語単位の外側から来る情報」を表しており<sup>9)</sup>、発話者の知識や信念、表情や身振り、あるいは社会状況や文化等、極めて多岐に渡る様々な情報を含む。談話分析は、個々の発話を独立に切り出して分析するのではなく、発話とそれが生起す

る様々な文脈との関わりの中で、ことばの仕組みや働きを明らかにすることを目的とする<sup>5), 6)</sup>。

本研究では、公的討議の談話分析の方法として、コーパス言語学における計算論的アプローチを採用する<sup>10) - 15)</sup>。コーパス (corpus) とは集成テキストを表し<sup>12)</sup>、人々がどのように発話しているかを示す大規模なサンプルを意味する。コーパス言語学は、コンピュータを利用した再現可能な計算手法を用いて、大規模なコーパスにおける語や句の用法や意味を分析する方法論である<sup>12), 13), 16)</sup>。羽鳥ら<sup>2)</sup>は、ファセット理論に基づく談話分析手法<sup>17), 18)</sup>を提案し、討議参加者間の関心の不一致や意見の対立を検証する方法論を提案している。さらに、ファセット理論を用いて実際の公的討議のテキスト(速記録)を言語情報化し、討議参加者間の関心の不一致や会話パターンを分析している。しかし、テキストを概念カテゴリー(ファセット)に分解する作業に多大な労力が必要であるため、1回の公的討議テキストのみを対象とした談話分析に留まっている。また、ファセット分解が分析者の裁量に委ねられており、分析手法の再現可能性が大きな課題であった。それに対して、本研究では分析者のファセット分解作業を支援するために、統計的ファセット学習モデルを導入する。コーパス言語学に基づく計算論的アプローチ<sup>10) - 13)</sup>を援用し、討議におけるトピックの抽出や討議参加者間の発話の類似度を算出する方法論を提案する。これにより、大量のコーパスのファセット分解が容易となり、公的討議全体を通じた討議参加者の関心の不一致性や発言内容の動的变化を分析することが可能となった点に本研究の新規性がある。

### (3) 第3者委員会による公的討議の特性

公共プロジェクトの構想・計画段階において専門家、企業、市民等、プロジェクトに関わるステークホルダーによる意見交換を促進するための公的討議の場を、本研究では第3者委員会と呼ぶ。第3者委員会の目的は、公的討議を通じて公共プロジェクトに対する関係者の理解を促進すると共に、討議過程を公開し意思決定過程の透明性を確保することにある。第3者委員会は意思決定に関与するわけではないが、公共プロジェクトに関わる情報の集約と意思決定者の決定に対する国民の信頼の醸成という役割

を担っている<sup>2), 19)-21)</sup>。

第3者委員会は、行政との間に直接的な委託-受託関係を有し、行政に対して説明責任を有している。これと同時に、国民との間にも間接的な委託-受託関係が成立し、委員会メンバーには、社会的に望ましい判断を行政ならびに国民に対して提示することが求められる。すなわち、国民の代表者としてプロジェクトの社会的望ましさ(適切性)に関する判断を提示するとともに、プロジェクトに関連する技術的・専門的事項に対する正確な知識(厳密性)に関する判断を示すことが求められる。つまり、第3者委員会はプロジェクトに関わる厳密性と適切性に関する判断を示すことにより、意思決定のための有用な情報をとりまとめることが期待されている<sup>20)</sup>。

公共プロジェクトに関わるステークホルダーの意見を把握する方法として、アンケート調査、ヒアリング調査等が実施されてきた。これらの手法は、多様な利害関係者の間で分節化された非同期的コミュニケーション<sup>22)</sup>に基づいており、ギデンズの言う「顔の見えないコミットメント(faceless commitments)」の問題が存在する。すなわち、調査者が質問を定式化する段階で、分析する視点をあらかじめ約束事として構造化しており、調査者と被調査者の間で約束事が守られることに対する信頼関係<sup>22)</sup>があらかじめ想定されている。さらに、調査内容の中で、一般の人々が認識していない実用的知識に関する調査者の信念が維持されることが前提となる<sup>24)</sup>。一方、第3者委員会では、多様な利害関係者が一堂に会することにより、双方のコミュニケーションを通して密度の高い社会的対話を実現することが可能である。さらに、第3者委員会における公的討議は、プロジェクトに関わる情報を効果的に集約するとともに、プロジェクトに関わる意思決定の透明性を確保するという意義を持っている。

公共プロジェクトの公的討議において、討議参加者が、公共事業のもたらす効果のうち、ある特定の側面のみに関心を持っている場合が少なくない。そのため、討議参加者が、自己の関心にとって有利な結果を導出するために戦略的な発言を繰り返し、コミュニケーションが成立しない可能性がある。さらに、プロジェクトに関わる専門的・技術的な判断をめぐって討議がなされる場合、専門家や技術者は、理論

的根拠や実証的証拠という「厳密性」を議論の拠とする。しかし、多くの市民は判断の厳密性よりも、自分の関心にとって有用であるか、常識的な内容であるかという判断の「適切性」を重視する。討議者が同一の問題に対しても、厳密性と適切性のいずれかを強調する場合、議論が互いに「すれ違う」可能性がある。第3者委員会では、討議参加者が対象とする問題に対して包括的な討議を行い、検討すべき課題や論点に対して共通認識を獲得することが重要な課題となる。そのためには、討議参加者が議論を踏まえて「討議参加者が当該公共プロジェクト案に対してどのような関心を持っているか」という関心体系を相互認識することが必要である。さらに、議論を重ねることにより、対象プロジェクトが有する課題や論点を包括的に整理し、その内容に対して討議参加者が共通認識を獲得する方向に討議が進展しているかを評価することが必要となる。本研究で提案する談話分析の方法論は、公的討議の包括性や議論の深化過程に関する情報を提供することにより、公的討議の進捗マネジメントをする上で重要な判断情報を提供することを目的としている。

### 3. ファセット分解による言語情報化

#### (1) 発話のファセット分解

本研究では、討議参加者の発話を、参加者がプロジェクトの是非を表明し、その是非について討議する行為と位置づける。討議にあたって、専門的な厳密性に基づいて判断の根拠を提示する場合と、日常的な生活感覚や利害関係に基づいてプロジェクトの適切性に関する判断の結果を述べる場合が存在する。さらに、討議参加者は、戦略的意図や心象的・感情的理由により、相手の発話に対して、肯定的態度あるいは否定的態度を示し、自分の見解を主張することを試みることがある。このような観点から、本研究では、討議参加者の発話行為を「ファセットA(賛否)」、「ファセットB(論証性/非論証性)」、「ファセットC(厳密性/適切性)」、「ファセットD(肯定的態度/否定的態度)」という4つのファセットに基づいて分類する。その上で、討議参加者のそれぞれの発話を4つのファセットの組(ストラクタブル)として言語情報化する。本研究では、公的討議の発話状況を記録した原テキスト群を1次コーパス、ファ

セットを用いて言語情報化した発話テキスト群を2次コーパスと呼ぶ。ファセットの詳細に関しては3.(2)で改めて言及する。また、発話によっては、自分の意見を明示的に表明しない場合がありうる。そのため、ファセットのカテゴリー分類において、上述の2値的分類だけでなく「判別が不可能(中立)」という第3のカテゴリーを設定した。

2次コーパスを作成するためには、討議参加者の発話内容をファセットの組み合わせとして表現することが必要となる。しかし、膨大な1次コーパスを論理的な整合性を維持しつつ、ファセット分解を行うことは容易ではない。そこで、本研究では、統計的ファセット学習モデルを用いて、1次コーパスのファセット分解1次案を作成する。このように作成したファセット分解1次案は、分析者によるファセット分解の労力を低減するとともに、ファセット分解における論理的一貫性を維持しつつ2次コーパスを確定するための判断情報として有用である。

本研究で提案する統計的ファセット学習モデルは、教師を持つ学習(supervised learning)により大規模な1次コーパスにおける各発話をファセット分解するパターン認識手法である。一般に、パターン認識とは、文字や映像や音声等の入力データから、その特徴を取り出し、標準的パターンに基づいて各データのカテゴリーを判別する方法を意味する。本研究では、Vapnikら<sup>23)-25)</sup>が提案した統計的機械学習手法の1つであるSVM(Support Vector Machine)手法を用いてファセット分解を試みる。SVMの詳細は参考文献23)-25)に譲る。SVMは2値分類器であり、サンプルの識別境界は1つに限られる。本研究では、それぞれのファセット $X \in \{A, B, C, D\}$ について3つのカテゴリーが存在するため、2個の境界超平面 $y_h^X(U_i), (h=1,2)$ を用いる。境界超平面は、コーパスに含まれる単語に対して重みづけした高次元の判別関数で定義され、各発話から識別境界までの距離を最大にするように決定される。判別関数は非線形のKernel関数で表現されるが、本研究では関数近似の柔軟性に優れるd次元のPolynomial関数を用いた。以上で求めたSVMモデルは大規模であり推計結果に関しては割愛する。

SVMの適用にあたって、1次コーパスに含まれる発話 $U_i (i=1, \dots, m)$ を、ファセット $X \in \{A, B, C, D\}$

表1 ファセットの分類

ファセット	ファセットの定義	ファセットの分類		
		X(1)	X(-1)	X(0)
		賛成	反対	その他 中立 挨拶 質問 議論進行にかか わる発話 X(1)とX(-1)両方 に当てはまらないと判断される すべての発話は X(0)に分類
ファセットA	発話内容から「発話の目的」を把握するためのファセット:当該公共プロジェクトの計画樹立および施行について賛成を表明する発話であるか反対を表明する発話であるかを確認することによって、発話者が発話行為の目的を理解する。			
ファセットB	発話内容から「論証のための情報源の伝達可否」を把握するためのファセット:発話者が自分の見解を主張するために聞き手に情報源を提示するかどうか論証性の有無を把握する。そして、聞き手と自分の見解を共有化するための発話者の情報伝達の努力水準を理解する。	論証性 科学的論証 数学的論証 心理的論証	非論証性 論証不全	
ファセットC	発話内容から「論証の情報源の属性」を把握するためのファセット:発話者が当該公共プロジェクトについて理解する際に判断材料として用いる情報を確認する。情報の属性を、専門的な厳密性と日常的な適切性として区別して発話者の関心体系を理解する。	厳密性 統計的事実 科学的事実 経験的事実 契約事実	適切性 心理的事実 (不安・信頼・感想) 感応概念 (価値・規範・信念) 公共性・倫理性・公平性・正統性	
ファセットD	発話内容から「相手の発話に対する態度」を把握するためのファセット:相手の発話に対する発話者の態度を確認することによって、相手の情報源に対する信頼や合意を理解する。	肯定的表現 尊重, 激励, 美化, お願い, 可能性, 共通性, 同意・合意・譲歩・妥協を主に表す発話	否定的表現 回避, 強制, 要求, 疑い, 不信, 説教, 訓戒, 軽視, 無視, 否定, 攻撃, 不可能性, 制限, 停止を主に表す発話	

を用いてカテゴリ  $F^X(U_i) \in \{X(-1), X(0), X(1)\}$

のいずれかに分類する。  $F^X(U_i)$  は対応関係であり

$$F^X : U_i \mapsto \{X(-1), X(0), X(1)\} \quad (1) \\ (X \in \{A, B, C, D\})$$

と定義される。さらに、 $n$  次元空間上 ( $n$  は 1 次コーパスに含まれる単語総数) に分布している  $m$  個の発話  $U_i (i=1, \dots, m)$  のそれぞれに対して、ファセット  $X \in \{A, B, C, D\}$  に関する 2 種類の Kernel 関数  $y_1^X(U_i), y_2^X(U_i)$  を用いて、1)  $y_1^X(U_i) > 0$  の場合  $F^X(U_i) = X(1)$ , 2)  $y_1^X(U_i) \leq 0$ , かつ  $y_2^X(U_i) > 0$  の場合  $F^X(U_i) = X(0)$ , 3)  $y_2^X(U_i) \leq 0$  が成立する場合、 $F^X(U_i) = X(-1)$  と設定する。さらに、発話  $U_i$  に対して、発話ファセットベクトル  $\mathbf{G}_i$  を、4 つのファセット  $X \in \{A, B, C, D\}$  のそれぞれのカテゴリ  $y \in \{X(-1), X(0), X(1)\}$  に該当するか否かを表す 0-1 変数 ( $\delta_i^y \in \{0, 1\}$ ) を用いて

$$\mathbf{G}_i = \{\delta_i^{A(-1)}, \delta_i^{A(0)}, \delta_i^{A(1)}, \delta_i^{B(-1)}, \delta_i^{B(0)}, \delta_i^{B(1)}, \\ \delta_i^{C(-1)}, \delta_i^{C(0)}, \delta_i^{C(1)}, \delta_i^{D(-1)}, \delta_i^{D(0)}, \delta_i^{D(1)}\} \quad (2)$$

と定義する。このようにして、すべての発話  $U_i (i=1, \dots, m)$  に対して定義された発話ファセットベクトル  $\mathbf{G}_i$  の集合を 2 次コーパスと呼ぶ。

## (2) 分析事例の概要

本研究では、Y 川水系流域委員会における討議内容を分析事例としてとりあげる。同委員会の討議内容は、1 次コーパス(委員会における発話内容をそのまま記載した速記録)としてウェブ上で公開されている。同委員会は、平成 13 年に Y 川水系における「河川整備計画」について専門家の意見を聞くことを目的として設置された。本委員会の他、4 つの地域別部会、5 つのテーマ別部会、5 つのワーキンググループ (WG)、3 つのダムサブワーキンググループ (ダムサブ WG) に分かれており、すべての部会、WG を含めると、合計 400 回以上の会議が実施されている。これらの会議の討議内容は速記録によって記録されており、速記録では発話者の固有名詞と発話内容が発話の順番に沿って明記されている。

本研究では、ウェブ上で公開されている 236 個の速記録の中から、「専門家と市民との意見交換会」と「専門家と行政との討議」が実施された一連の合計 14 回の会議を分析対象としてとりあげた。この会議の討議参加者は、1) 座長、2) 専門家、3) 賛成派市民、4) 反対派市民、5) プロジェクト実施機関である行政の 5 つのグループに大別できる。分析対象とする会議では、ダムに関する主題がとりあげられて

いる。14個の電子速記録における討議参加者の発話数は、合計8831単位である。なお、句読点単位を1回の発話単位として定義している。

### (3) 発話のファセット分解

パブリック・インボルブメントの運用を担当した経験のある建設コンサルタントに勤務する4名の技術者が各発話の予備的ファセット分解を試みた。発話ファセット分解の客観性を可能な限り確保するために、同一の発話に対して2名の調査者がファセット分解を試みた。その際、表-1に示す4つのファセットを用いて討議参加者の発話のフレーム構造を規定した。同表においては、各ファセットの要素とその典型的な表現事例を示している。

ファセットAは、当該プロジェクトに対する賛否に関する意見を表している。すなわち、ファセットA(1)は、プロジェクトに対する賛成意見を表している。典型的な表現例としては、プロジェクトの長所についての説明や市民の要求等が該当する。ファセットA(-1)は、プロジェクトに対する反対意見を表しており、プロジェクトの短所についての説明や他の代替案の提案等が該当する。ファセットBは、発話の論証性を表しており、発話の論拠を示した場合はファセットB(1)に、示さなかった場合はファセットB(-1)に分類される。また、ファセットCは、発話の論拠の内容を表しており、ファセットC(1)は、厳密性に基づく論拠を示した場合、ファセットC(-1)は、適切性に基づく論拠を示した場合を表している。ファセットC(1)の典型的表現事例としては、シミュレーションや実験データを用いた説明等が挙げられる。一方、ファセットC(-1)の典型的表現事例として、地域の経験談や自分の不安感等の心理状態を主張する場合が挙げられる。最後に、ファセットDは、発話における肯定的態度および否定的態度を表している。ファセットD(1)は、肯定的態度を表しており、発話の形容詞や副詞において、他者や物事に対する肯定的表現や希望が述べられる場合が該当する。ファセットD(-1)は、プロジェクトに対する否定的態度を表しており、その典型的表現として、発話の形容詞や副詞において、他者や物事に対する批判的表現や否定的意見が述べられる場合が該当する。この結果、討議参加者間の意見の対立を、ファセットA, Dの

カテゴリーの違いとして表現する。また、討議参加者が発話の論拠を示しているかどうかを、ファセットBにより確認できる。厳密性、適切性に関わる議論のすれ違いはファセットCの違いにより表現される。なお、カテゴリー分類が困難な場合に対処するため、 $X(0), X \in \{A, B, C, D\}$ という中立カテゴリーを設定した。

上記の方法に従って、討議参加者の発話を4つのファセットに分解した例を以下に示す。

【例文1】○○町のような貧弱な財政事情の中でダムしか生きる道がないんです。→A(1)B(-1)C(-1)D(-1)

【例文2】日本のダムで、ダム周辺にいろいろな施設をつくって繁栄させます、地域の活性化を試しますといってつくったダムはたくさんありますけれども成功したためしはほとんどありません。→A(-1)B(1)C(1)D(-1)

以上の例文は、本事例における「ダム建設による地域振興に関する発話」である。例文1)について、ファセットAに関しては「ダムしか~ない」という表現より、ダム建設の賛成意見A(1)として分類される。ファセットBに関して「貧弱な財政事情」という表現から、日常的な関心事、また心理的不安感の表出と判断され、適切性B(-1)に分解される。ファセットCに関して、再び「貧弱な財政事情」という表現から心理的な論拠C(-1)を提示したものと判断される。ファセットDに関して「~しか生きる道がない」という表現から、否定的態度D(-1)を示すものと分解される。以上のことより、この発話は、A(1)B(-1)C(-1)D(-1)というファセットに分解される。次に、例文2)は、ダム建設の経験的事実をその論拠として用いて、その結果について悲観的意見を述べるとともに、ダム建設に反対している。この発話は、A(-1)B(1)C(1)D(-1)と分解される。

2次コーパスの作成にあたっては、まず、調査者がファセット分解表(表-1)に基づいて、速記録における各発話の予備的ファセット分解を試みる。その際、ファセット分解の客観性を確保するために、複数の調査者が分解作業を行い、調査者の間で一致したパターンのみを統計的機械学習における学習サンプルデータとして用いる。このようなサンプル数は各ファセットに対して、それぞれ34,65,59,52%であり、ファセットAを除いて過半数のサンプルが学習

表-2 SVM によるファセット分解結果

総発話 8831	ファセットA		ファセットB		ファセットC		ファセットD		
	賛/反		論/非論		厳/適		肯/否		
	数	%	数	%	数	%	数	%	
調査者	○	6130	35	5447	31	3715	21	2088	12
	●	5839	33	4214	24	7252	41	9716	55
	△	5693	32	8045	45	6739	38	5902	33
SVM	○	313	4	2292	26	904	10	452	5
	●	49	1	1793	20	2142	24	2959	34
	△	8469	95	4746	54	5785	66	5420	61
最終版	○	3532	40	2914	33	2031	23	1236	14
	●	3267	37	2296	26	3533	40	4946	56
	△	2782	23	3621	41	3267	37	2649	30
精度	48		66		55		50		
F値	35		54		44		34		

注) ○はカテゴリーX(1), ●はカテゴリーX(-1), △は判別不能(中立)カテゴリーX(0)を表す。

サンプルとして選ばれた。本研究では、同一の討議内容について議論された一連の 14 個の速記録すべてを対象として、上記の手順により学習サンプルデータを作成した。その上で、学習サンプルデータを教師データとして位置づけ、SVM 手法を用いて発話データのファセット分解を行うための Kernel 関数を推計した。最後に、14 個の速記録に対する調査者による分解結果と SVM による分解結果との一致率を求め、SVM モデルの判別精度を検証した。その際、Kernel 関数を K-fold Cross Validation 法を用いて推計した。すなわち、発話コーパスの全体を K 個のサブグループに分解し、そのうち  $K - 1$  個のグループの属する発話コーパスに対して、調査者が作成した予備的ファセット分解の結果を教師データとして SVM モデルを学習させた。その上で、残りの 1 個のグループに属する発話コーパスに対して SVM モデルによる 2 次コーパスを機械的に作成する。ただし、 $K = 10$  に設定する。このような試行を教師データとして用いる発話コーパスの集合を逐次変えることにより、合計 K 回の試行により機械判別を行った。発話  $U_i (i=1, \dots, m)$  のそれぞれに対してファセット  $X \in \{A, B, C, D\}$  のカテゴリー分類の結果が  $y \in \{X(-1), X(0), X(1)\}$  に判別されるか否かを示すダミー変数  $\delta_i^y$  を定義する。その上で、1) 教師データと再現データのファセット分解の結果がともに  $\delta_i^y = 1$  となる回数を  $\alpha$ , 2) 教師データが  $\delta_i^y = 1$  と再現データが  $\delta_i^y = 0$  となる回数を  $\beta$ , 3) 教師データが  $\delta_i^y = 0$ , 再現データが  $\delta_i^y = 1$  となる回数を  $\gamma$ ,

4) 双方が同時に  $\delta_i^y = 0$  となる回数を  $\lambda$  とすれば、判別精度を表す精度、F 値<sup>26)</sup>は

$$\text{精度 } A = \frac{\alpha + \lambda}{\alpha + \beta + \gamma + \lambda} \quad (3a)$$

$$F \text{ 値 } F = \frac{2P \times R}{P + R} \quad (3b)$$

と定義できる。ただし、 $P = \frac{\beta}{\alpha + \beta}$ ,  $R = \frac{\alpha}{\alpha + \gamma}$  である。

#### (4) ファセット分解の結果

表-2 に SVM によりファセット分解した結果を示している。調査者の欄は複数の調査者によるファセット分解の結果を集計している。このため、ファセット分解数が、総発話数を上回っている。当該整備計画に対する賛成発話が約 35%, 反対が 33%, 論拠に基づいた発話が 31%とそうでない発話が 24%, 厳密的論証に基づいた発話が 21%, 適切的論証が 41%, 肯定的発話が 12%, 否定的発話が 55%である。過半数の発話者が当該プロジェクトに対する適切的論証を通じて、否定的意見を述べていると判断される。この結果は、討議場において発話権が付与された討議参加者のみに該当するものであり、討議参加者全体の関心として一般化できないことに注意する必要がある。一方、SVM の欄は、SVM を用いた判別結果を表している。賛成発話が 4%, 反対が 1%, 論拠に基づいた発話が 26%, そうでない発話が 20%, 厳密的論証が 10%, 適切的論証が 24%, 肯定的発話が 5%, 否定的発話が 34%となる。式(3a)に示すファセット分解の精度は 48%~66%程度である。一般に、0.6~0.7 以上の精度を得ることが望ましい<sup>27)</sup>とされる。しかし、本研究では 3 つのカテゴリーに分解することもあり、SVM を用いた場合、調査者のファセット分解より中立カテゴリー X(0)に分類される結果が多い。その結果、精度、F 値とも上記の値より若干低くなっている。そこで、最終的に分析者により、SVM の結果を参照しながら、最終的なファセット分解を確定した。その結果は、同表の最終版の欄に示すとおりである。SVM を導入することにより、分析者による最終判断のための素材情報を与えることが可能となり、2 次コーパスを確定する労力を大幅に削減できたことは事実である。なお、SVM を用いた場合、ファセット A の分解精度が悪い。その原因としては、発話内容に賛否に関わる表現が明示的に用い

られない、あるいは、意図的に曖昧な表現が用いられる場合が少なくないことがあげられる。このように高度に文脈に依存する発話のファセット分解を、発話単位の情報に基づいて判断することには限界があり、分析者の最終判断に頼らざるを得ない部分がある。可能な限りファセット分解の客観性を確保するためには、討議における会話パターンの分析結果(4. (2)参照)に基づいて分解案の修正を試みるというフィードバック作業のシステム化が必要である。

## 4. 公的討議の談話分析

### (1) 討議参加者の空間配置

ファセット分解を行った 2 次コーパスを基にして談話分析を実施し、1) 討議参加者間の関心の不一致や意見の対立、2) 討論過程における会話パターンの変化を視覚的に表現する。まず、2 次コーパス全体を対象として、多次元尺度構成法(Multidimensional Scaling: MDS)<sup>28),29)</sup>を用いて、討議参加者間の関心の類似性を明示化する。MDS は、複数の変数間の相関(類似度データ)に基づいて、諸変数をある次元の空間布置に配置し、変数間の関係を空間上の距離によって表現する方法である。ここで、変数間の類似度が高いほど距離が小さくなり、類似度が低いほど距離が大きくなる。

MDS は非計量多次元尺度構成法に属し、討議参加者間の非類似度に関する順序情報に基づいて  $Z$  人の参加者の配置を決定する。参加者  $p_i \in (p_1, \dots, p_z)$  の中で、討議参加者  $p_i$  と参加者  $p_j$  の関心の非類似度  $dsim(p_i, p_j)$  を、2 次元の空間における  $p_i$  と  $p_j$  の距離  $dis(p_i, p_j)$  として表す。MDS は討議参加者間の非類似度が  $dsim(p_i, p_j) < dsim(p_i, p_k)$  であれば、空間上の距離は  $dis(p_i, p_j) < dis(p_i, p_k)$  という弱単調関係になるように各参加者の配置を決定する。つまり、発話の非類似度が大きいほど空間上で遠く、非類似度が小さいほど空間上で近くに配置する。

まず、参加者  $p_k$  が発話した 2 次コーパスに対して発話ファセットベクトルを構成するダミー変数  $\delta'_j (y = A(-1), A(0), A(1), \dots, D(-1), D(0), D(1))$  が出現する相対頻度  $W_y^{pk}$  を、

$$W_y^{pk} = \frac{\sum_{j \in \Omega_{pk}} \delta'_j}{\sum_{j \in \Omega_{pk}} \sum_{y=A(-1)}^{D(1)} \delta'_j} \quad (4)$$

と定義する。ただし、 $\Omega_{pk}$  は、参加者  $p_k$  の発言ファセットベクトルの集合である。ファセット行列  $S$  を

$$S = \begin{bmatrix} \mathbf{F}_{p1} \\ \vdots \\ \mathbf{F}_{pz} \end{bmatrix} = \begin{bmatrix} W_{A(-1)}^{p1} & \cdots & W_{D(1)}^{p1} \\ \vdots & \ddots & \vdots \\ W_{A(-1)}^{pz} & \cdots & W_{D(1)}^{pz} \end{bmatrix} \quad (5)$$

と定義する。その上で、2 人の討議参加者のファセットベクトルの類似度を、コサイン角度距離<sup>30)</sup>

$$\begin{aligned} sim(p_i, p_j) &= \cos(\mathbf{F}_{pi}, \mathbf{F}_{pj}) = \frac{\mathbf{F}_{pi} \cdot \mathbf{F}_{pj}}{|\mathbf{F}_{pi}| \cdot |\mathbf{F}_{pj}|} \\ &= \frac{\sum_{y=A(-1)}^{D(1)} W_y^{pi} W_y^{pj}}{\sqrt{\sum_{y=A(-1)}^{D(1)} (W_y^{pi})^2} \sqrt{\sum_{y=A(-1)}^{D(1)} (W_y^{pj})^2}} \end{aligned} \quad (6)$$

を用いて定義する。討議参加者を空間上に配置する際に、非類似度が高いほど遠くに、小さいほど近くに位置させるために、式 (6) に示す類似度指標に対して、逆余弦アークコサイン角度距離を用いて非類似度を

$$dsim(p_i, p_j) = \cos^{-1}\{sim_i(p_i, p_j)\} \quad (7)$$

と定義する。さらに、MDS を用いて 2 次元空間上に配置した討議参加者間の空間距離をマトリクス  $D$

$$D = \begin{pmatrix} dis(p_1, p_1) & \cdots & dis(p_1, p_z) \\ \vdots & \ddots & \vdots \\ dis(p_z, p_1) & \cdots & dis(p_z, p_z) \end{pmatrix} \quad (8)$$

を用いて定義する。MDS では、討議参加者を  $dis(p_i, p_j)$  と非類似度  $dsim(p_i, p_j)$  の差の 2 乗和を最小化するように 2 次元空間上に参加者の配置を決定する。非類似度と点間距離の不適合度を表すストレス(Stress)を用いて、 $dis(p_i, p_j)$  と  $dsim(p_i, p_j)$  の差の 2 乗和を最小化する座標値を決定する。

$$Stress = \sqrt{\frac{\sum_{i=1}^{z-1} \sum_{j=i+1}^z (dis(p_i, p_j) - dsim(p_i, p_j))^2}{\sum_{i=1}^{z-1} \sum_{j=i+1}^z (dis(p_i, p_j) - \bar{dis})^2}} \quad (10)$$

ただし、 $\bar{dis}$  はすべての討議参加者間における距離の 2 乗平均値であり、次式で定義される。

$$\bar{dis} = \sqrt{\frac{\sum_{i=1}^{z-1} \sum_{j=i+1}^z (dis(p_i, p_j))^2}{C_2}} \quad (11)$$

ストレス(Stress)を最小化する配置を求めるために、2 次元空間上の各参加者の位置を少しづつ改善し、逐次ストレスを減少させ、ストレスを最小にするような 2 次元上の空間配置を決定する。図-1 は、対象と

する 14 個の速記録(1 次コーパス)に登場する討議参加者(専門家 49 名, 行政 27 名, 市民 5 名)を, ファセットベクトルの類似度に基づいて 2 次元空間上に配置した結果を表している。同図において参加者間の距離のみが意味をもっており, 参加者の空間座標には特別の意味はない。ここで, ◆は市民を, \*は専門家を, □は行政をそれぞれ表す。討議参加者は互いに似通った関心を示すいくつかのグループに分類される。◆ラベルを持つ参加者は, 互いに近くに位置している。そして, ◆ラベルを有する参加者を核として, その他の討議参加者が空間配置されている。特に, その上部には□ラベルが, 下部には\*ラベルが多く分布している。すなわち, 専門家たちと行政の間で関心の不一致が存在している。

図-2において, 1) 座長( $\otimes$ ), 2) 専門家(\*), 3) (ダム事業に対する) 賛成派市民(◆), 4) 反対派市民( $\diamond$ ), 5) 行政(□)の 5 グループに属する参加者の中で, 特に発話数が多かった代表的な討議参加者 5 人をとりあげ, 選択した 5 名の関心の類似性を確認する。代表者 5 人の配置を表した図-2 から座長と専門家の距離に比べて, 座長と市民や行政の距離が比較的に遠いことが分かる。上述したように, 空間上の距離が離れるほど, 討議参加者の関心体系が異なっていることを意味する。座長と一部の専門家の関心体系は似通っているが, これらのグループと行政あるいは市民とは, プロジェクトに対して互いに異なる関心を示している。しかし, 公的討議の開始時点において, 討議参加者の関心体系が異なっていること自体は問題ではない。むしろ, 多様な意見を持っているメンバーが公的討議に参加することは望ましいことであるともいえる。問題は, 公的討議を通じて価値観や意見の異なる討議参加者の関心体系にどのような変化が発生しているかである。そこで, 以下では, 図-2 に示している討議参加者に再び着目するとともに, 討議の蓄積により, これらの代表的参加者の関心体系や発言内容が, 時間とともにどのように変化したかを分析する。

## (2) ファセット分解に基づく会話パターン

座長, 専門家, 賛成派市民, 反対派市民, 行政代表 5 人の討議参加者の発話ファセットの時間的变化パターンを分析するために, 14 回の公的討議により

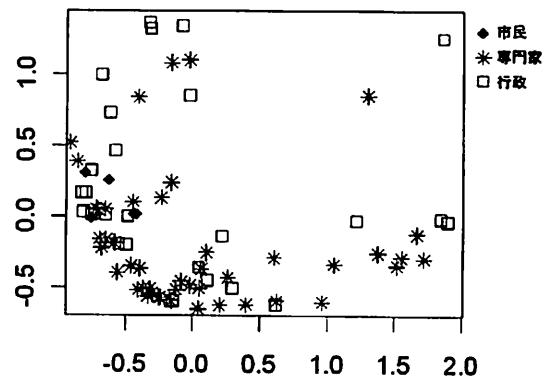


図-1 発話の類似度

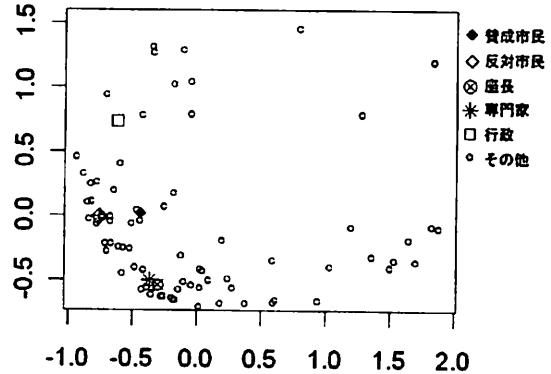


図-2 代表 5 人の位置づけ

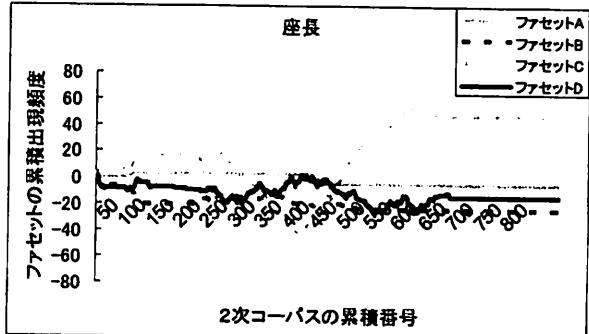


図-3 ファセットパターンの変化(座長)

得られた 1 次コーパスを時間の経過の順に並べ, 発話コーパスの時間的シーケンスを作成した。さらに, 各代表的討議参加者が利用したファセットカテゴリーの累積出現頻度を算定した。図-3, 図-4, 図-5, 図-6, 図-7 は, それぞれ代表的討議参加者の発話コーパスにおけるファセットカテゴリーの累積出現頻度の推移状態を表したものである。ここで, 横軸は経過時間順に序列化した発話コーパスの累積番号であり, 縦軸は各発話のファセット要素の累積出現頻度である。ファセット  $X$  のカテゴリー  $X(l)$  の出現頻度が増加すれば, グラフが経過時間に関して増加関数となる。 $X(-l)$  の出現頻度が多くなければ減少関数

となる。各ファセットに対して、カテゴリー  $X(1)$  と  $X(-1)$  の出現頻度のバランスがとれている場合、累積出現頻度は横軸の近傍で推移することになる。

座長のカテゴリー累積出現頻度の時間的変化を図-3 に示している。座長は、分析対象とした 14 回の討議の内 8 回の討議に参加して合計 849 単位の発話を行っている。ファセット分解の結果、1) ファセット A(賛/反)に関しては、ほとんど意見を表明していない、2) ファセット B(論/非論)に関しては、非論証的発話が比較的に多い、3) ファセット C(厳/適)に関しては、会議全体を通じて、厳密的論拠を示すことが多い。しかし、4) ファセット C に関して、累積的出現頻度が急増した局面が存在する。ある特定の主題に関して、適切的論拠に関する発話を多用したことが読み取れる。5) ファセット D(肯/否)に関しては、全体を通じて否定的なカテゴリーを用いることが多い。

一方、図-3～図-7 に示すように、他の代表的討議参加者のカテゴリーの累積的出現頻度は、時間的に安定しており、座長のカテゴリー変化とは異なった変化パターンを示している。例えば、専門家は 6 回の討議に参加し合計 444 単位の発話を、行政は 1 回の討議で 151 発話を、賛成市民は、1 回の討議で 143 単位の発話を、反対市民は 1 回の討議で 228 単位の発話を行っている。座長と専門家を除いて、他の討議参加者は、1 回のみの討議に単発的な参加をしていることが多い。自分の価値判断を述べたとしても、時間的な制約が多いため、相手の価値判断に関する理解の程度も低く、討議者間コミュニケーションを通じた相互理解、共通認識を得ることは容易ではないと考えられる。また、専門家や行政の場合、中立カテゴリーを多用する傾向が強いが、反対市民や賛成市民は同じファセットカテゴリーを繰り返し利用する傾向が見られる。

以上の分析事例より、ファセットカテゴリーの累積出現頻度は簡単な指標ではあるが、討議参加者の発話意図や発話内容の時間的変化過程を簡単に追跡することが可能であり、各々の討議参加者の微妙な関心の変化を追跡することが可能であることが理解できる。さらに、討議参加者間の関心の不一致や意見の対立を確認し、これらの事象が生起するプロセスやその問題の原因を調べることが可能となる。

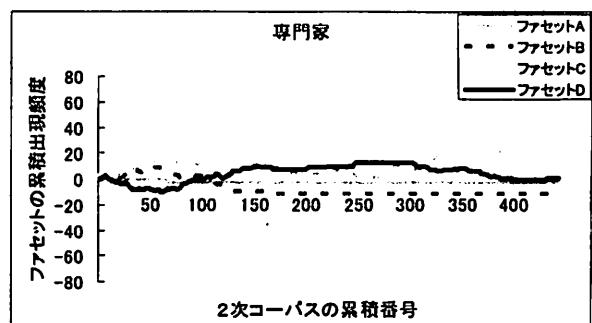


図-4 ファセットパターンの変化(専門家)

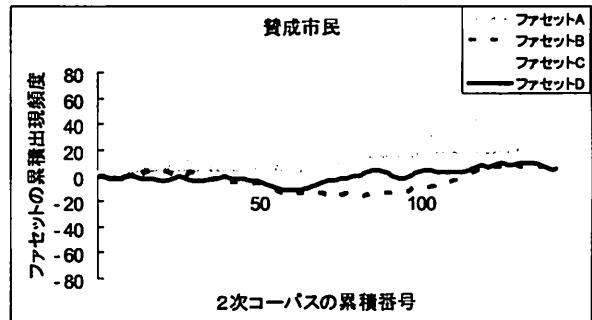


図-5 ファセットパターンの変化(賛成市民)

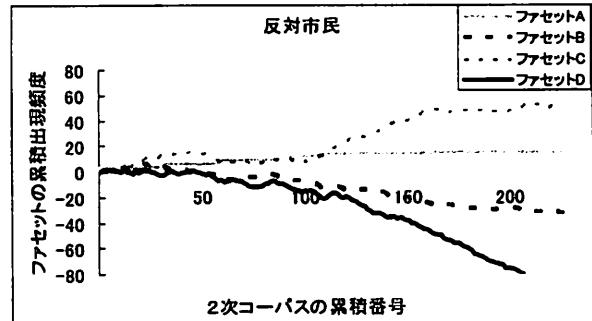


図-6 ファセットパターンの変化(反対市民)

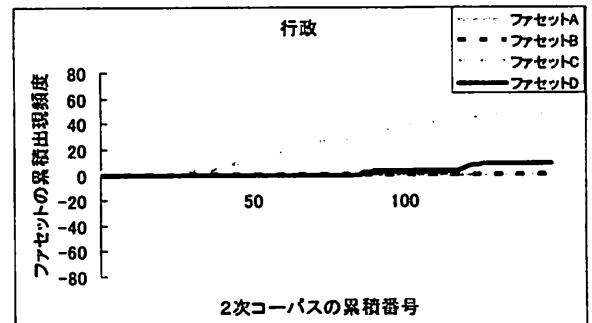


図-7 ファセットパターンの変化(行政)

## 5. おわりに

本研究では、公共プロジェクトを対象とした公的討議に対してファセット分解と談話分析を実施し、討議参加者の関心の違いや発話内容の時間的変化を分析する方法論を提案した。さらに、現実の公的討

議事例を対象として、討議参加者の間で関心の不一致問題や意見対立問題が存在することを確認した。公的討議においては、討議参加者が互いに納得可能な望ましい代替案を探り、社会的な合意形成を導く上で、ファシリテーション技術の高度化が重要な課題である。その場合、討議参加者の間においてコンフリクトが生じている原因とパターンを明確にした上で、誤解を排除し、相互理解を促進することが求められる。本研究で提案した談話分析の方法論は、公的討議の包括性や議論の深化過程に関する情報を提供することを目的としており、公的討議の進捗マネジメントをする上で重要な判断情報を提供しうると考える。しかし、今後に残されたいいくつかの検討事項や発展の方向性を指摘できる。第1に本研究で提案した方法論の有効性は、言語情報化に用いるファセット分解の概念フレームに多くを依存している。実証分析の蓄積により、ファセット分解の概念フレームの有効性に関して継続的に分析することが必要である。第2に、ファセットの分解精度の向上が望まれる。発話単位のSVMのみに基づいて、高度に文脈に依存した発話のファセット分解を行うことには限界がある。討議全体を通じた会話パターンの分析結果を用いて、発話のファセット分解を修正する方法論を開発することが必要である。第3に、発話内容は、発話者の特性や討議場の文脈に依存する。観測不可能な個人の特性や場の文脈などの潜在変数を考慮したファセット遷移確率モデルを開発することにより、ファセット分解の精度を向上させることが必要である。第4に、ファセット分解による言語情報化の過程の中で、多くの情報が捨象されている。公的討議に関するファセット以外の情報を有するデータベースを構築することにより、「討議における発話の長さや発話のターン（発言権）、参加者間の対立関係」などの特性を分析できる方法論を開発することが必要である。第5に、公的討議のあり方にに関する規範的研究が必要である。すなわち、公的討議における規範的ルールを開発し、公的討議パターンの望ましさに関する評価モデルを開発することが求められる。このような公的討議の評価を通じて、ファシリテーション技術の向上を図ることが可能となる。

なお、本研究の遂行にあたって、科学研究費補助

金・研究活動スタート支援【課題名：住民コンサー  
ンの可視化のためのコーパス分析。課題番号：  
21860048、研究代表：鄭蝦榮】の助成を頂いた。こ  
こに記して感謝の意を表す。

### 【参考文献】

- 1) 白松俊、駒谷和範、橋田浩一、尾形哲也、奥乃博：ゲーム理論に基づく参照結束性のモデル化と日本語・英語の大規模コーパスを用いた統計的検証、自然言語処理、Vol.14(5), pp.199-239, 2007.
- 2) 羽鳥剛史、川除隆広、小林潔司、夏目卓生、藤崎英司：ファセット理論に基づく公的討論過程のプロトコル分析、土木計画学研究・論文集、No.23, pp. 91-102, 2006.
- 3) Krippendorff, K.: *Content Analysis: An Introduction to Its Methodology*, Sage Publication, Inc., 1980. (三上俊治、椎野信雄、橋元良明訳：メッセージ分析の技法－「内容分析」への招待、勁草書房, 1989.)
- 4) Stone, P.J., Dunphy, D.C., Smith, M.S. and Ogilvie, D.M.: *The General Inquirer: A Computer Approach to Content Analysis*, MIT Press, 1966.
- 5) 橋内武、ディスコース—談話の織りなす世界、くろしお出版, 2003.
- 6) Schiffrin, D., Tannen, D., and Hamilton, H.E.: *The Handbook of Discourse Analysis*, Blackwell Publishers, 2003.
- 7) 小泉保：入門語用論研究－理論と応用－、研究社, 2001.
- 8) Schiffrin, D.: *Approaches to Discourse*, Blackwell, 1994.
- 9) 田窪行則、西山祐司、三藤博、亀山恵、片桐恭弘：談話と文脈、岩波書店, 2004.
- 10) Francis, N.: Problems of Assembling and Computerizing large Corpora, In: Johansson, S.(ed.): *Computer Corpora in English Language Research*, Norwegian Computing Centre for the Humanities, Bergen, pp.7-24, 1982.
- 11) Sinclair, J.: *Corpus, Concordance, Collocation*, Oxford University Press, 1991.
- 12) Stubbs, M.: *Words and Phrases: Corpus Studies of Lexical Semantics*, Blackwell, 2002.(南出康世, 石

- 川慎一郎監訳: コーパス語彙意味論—語から句へ  
一, 研究社, 2006.)
- 13) Wang, S.-P.: Corpus-based approaches and discourse analysis in relation to reduplication and repetition, *Journal of Pragmatics*, Vol.37, pp.505-540, 2005.
  - 14) McCarthy, M.: *Spoken Language and Applied Linguistics*, Cambridge University Press, 1998.
  - 15) Sinclair, J.: The search for units of meaning, *Textus*, Vol.9, pp.75-106, 1996.
  - 16) 大津由紀雄, 池内正幸, 今西典子, 水光雅則編: 言語研究入門—生成文法を学ぶ人のために, 研究社, 2002.
  - 17) Ericsson, K.A. and Simon, H.A.: *Protocol Analysis: Verbal Reports as Data*, MIT Press, 1993.
  - 18) 海保博之, 原田悦子: 談話分析入門: 発話データから何を読むか, 新曜社, 1993.
  - 19) Lupia, A. and McCubbins, M.D.: *The Democratic Dilemma; Can Citizens Learn What They Need To Know?*, Cambridge University Press, 1998.
  - 20) 羽鳥剛史, 鄭蝦榮, 小林潔司: 第3者委員会の公開と信頼形成への影響, 土木学会論文集D, Vol.64, pp.148-169, 2008.
  - 21) Jeong, H., Hatori, T., and Kobayashi, K.: "Discourse Analysis of Public Debates: A Corpus-based Approach", IEEE, SMC, 2007.
  - 22) 石井和平: 社会情報学, 情報技術と社会の共変, 学術出版社, 2007.
  - 23) 工藤拓, 松本欲治: Support Vector Machine を用いた Chunk 同定, 自然言語処理, Vol.9, pp.3-22,
  - 24) 若木裕美, 正田備也, 高須淳宏, 安達淳: 検索語の曖昧性解消のためのトピック指向単語抽出および単語クラスタリング, 情報処理学会論文誌: データベース, Vol.47, No.SIG19 (TOD32), pp.72-85, 2006.
  - 25) 大久保崇: 韻律変化による文章装飾機能を備えた音声ディクテーションシステム, 早稲田大学, 2004.
  - 26) Ong, B.S.: Towards Automatic Music Structural Analysis: Identifying Characteristic Within - Song Excerpts in Popular Music, Universitat Pompeu Fabra Barcelona, 2005.
  - 27) Qian, G., Sural, S., Gu, Y., and Pramanik, S.: Similarity between Euclidean and cosine angle distance for nearest neighbor queries, Proceedings of the 19th Annual ACM Symposium on Applied Computing, 2004.
  - 28) Kruskal, J.B.: Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis, *Psychometrika*, Vol.29, pp.1-29, 1964.
  - 29) Kruskal, J.B.: Nonmetric multidimensional scaling: a numerical method, *Psychometrika*, Vol.29, pp. 115-129, 1964.
  - 30) Koreniusa, T., Laurikkala, J., and Juhola, M.: On principal component analysis, cosine and Euclidean measures in information retrieval, *Information Sciences*, Vol.177, pp.4893-4905, 2007.

## Facet Decomposition and Discourse Analysis of Public Debate

By Hayeong JEONG, Kiyoshi KOBAYASHI, Tsuyoshi HATORI, Shun SHIRAMATSU

Recently, public involvement in the early stage of conceptual plan for public project has been implemented in many public projects and public debate is set up for discussion among diverse stakeholders from all walks of life about the project effects and impacts. This study proposes a new methodology to investigate the level of discourse similarity and interest dissonance among participants in public debate objectively using first corpus of debate minutes. The level of discourse similarity and interest dissonance can be measured quantitatively by using discourse analysis of second corpus which is derived by Facet decomposition of first corpus. Finally, the applicability of the proposed methodology is discussed by applying to practical cases of public debate, in order to measure the interest dissonance among participants and visualize the conflict structure.