慶應義塾大学学術情報リポジトリ
Keio Associated Repository of Academic resouces

Title	テキストマイニングを用いた集団意思決定の力動の可視化
Sub Title	
Author	村田, 一太郎(Murata, Ichitaro) 高木, 晴夫(Takagi, Haruo)
Publisher	慶應義塾大学大学院経営管理研究科
Publication year	2012
Jtitle	
JaLC DOI	
Abstract	
Notes	修士学位論文. 2012年度経営学 第2798号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40003001-00002012-2798

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

慶應義塾大学大学院経営管理研究科修士課程

学位論文　２０１２年度

論文標題

テキストマイニングを用いた集団意思決定の力動の可視化

指導教員（主査）	髙木　晴夫　教授
副指導教員（副査）	井上　哲浩　教授
副指導教員（副査）	高橋　大志　准教授
副指導教員（副査）	

２０１２年2月２８日提出

学籍番号	８１１３０９７８	氏名	村田　一太郎

論　文　要　旨

所属ゼミ	髙木研究室	学籍番号	81130978	氏名	村田　一太郎

（論文題名）

テキストマイニングを用いた集団意思決定の力動の可視化

（内容の要旨）

設立の目的、構成員数の多少、営利非営利の別を問わず、組織で協働作業を進めるには合議による意思決定は不可避であろう。そして、会議の場を司るリーダーは高効率で知的生産性の高い議論を行い、合意形成を図りたいと考えるに違いない。特に企業における重大な意思決定の場面であれば瞬時に最適な意思決定や対応が求められる例は少なくない。しかしながら、組織は常に会議の合意形成で最高のパフォーマンスが出せるとは限らない。

高効率で知的生産性の高い会議を行うにはどうすれば良いか。この命題に対し、高木 1994、北村 1992 は会議のコミュニケーションメカニズムの解明と集団意思決定における問題点について焦点を当てた。一方、筆者は自ら執行した内部統制構築プロジェクトでの経験から、当該先行研究に興味を抱くと共に集団意思決定における合意形成の態様に重大な関心を抱いていた。

本研究は、先行研究と異なるアプローチを用いて合意形成を目的とした模擬会議の模様を観察した。まず、映像記録した参加者の会話をテキスト化し、テキストマイニングやクラスター分析等で定量化を図った。そして、集団の合意形成が如何にして行われるかを観察・分析し、最終的に力動の可視化方法を模索した。

この結果、会議参加者のコミュニケーションの態様が、会議の成果に影響を与える可能性があるとする先行研究とは異なる観点で発見があった。まず、集団には全体の議論をとり仕切る者が時間経過と共に出現し、この人物が合議結果に重要な影響を与えていること、及びこの人物はテキスト化した文字数ベースでの発言量が上位１位か２位の確率が高いことを確認した。一方、この人物の会議での発言の特徴として、表現の豊かさを示すタイプ・トークン比が発言量に反比例し、参加者のなかでは最も低い傾向を示すことも明らかとなった。そして、この力動について可視化方法を模索した結果、テキストデータから結びつきの強さを示す Jaccard 係数を用いたクラスター分析を応用することによりデンドログラムが描画可能であり、会議参加者の動態理解に有効であることがわかった。

目次

第１章　研究の背景

1-1　問題意識　1

1-2　研究の目的と概要　3

1-3　研究の方法　4

1-4　観察研究の概要　8

第２章　先行研究

2-1　文献研究の概要　10

2-2　組織科学へのテキストマイニング導入　14

2-3　人工知能分野における研究　14

第３章　観察研究の実施

3-1　模擬会議の概要と適格性　19

3-2　アンケートの集計　20

3-3　模擬会議の記録とデータベース作成　21

3-4　予測データの収集　25

第４章　テキストマイニングの実施と観察結果の分析

4-1　模擬会議の実施模様に関する考察　30

4-2　テキストマイニングの実施　33

4-3　クラスター分析　57

第５章　　結論と提言

5-1　導出された結論　64

5-2　仮説検証のための追加調査　68

5-3　インプリケーションと今後の課題　72

参考文献等　73

附録　75

第１章　研究の背景

1-1　　問題意識

企業において自社の競争優位を司るコア・コンピタンスは、組織が内包する「顧客に特定の利益をもたらす技術、スキル、ノウハウの集合」[i]をいうが、このコア・コンピタンスを構成する有形・無形の企業内部に存する何某かを、活性を保ったまま維持し、革新させ続けることはリーダーにとってそれほど容易なことではない。それらが単純作業のようにマニュアル化が可能なものであれば伝承・保持に関するナレッジマネジメントはそれほど困難ではないかもしれない。しかし、企業における競争力の源泉がマニュアル化の可能なものばかりで構成されていることは珍しく、暗黙知や個人の能力・経験など、標準化が困難なものの方が多いというのが筆者の理解である。また、組織はその規模が大きくなればなるほど事業や機能等の単位で細分化されていく例が多くみられ、そこでのリーダーは組織の内外と適切なコミュニケーションを図りつつ、相互理解の下に協働作業を進めていかなければならない。

このような協働作業のマネジメントの大変さを痛切に感じたのは筆者がプロジェクトリーダーとして内部統制の整備に関与したときであった。

組織は環境の変化に応じ、常に最適な業務フローを模索し、進化し続けなければならない。特定の業務に関与する人々のコミュニケーションの影響範囲を「場」と表現したとき、その場で誰かが「内部統制に問題有り」と認識したら即座に解決策を議論し、自発的に対応が図られる必要がある。当然、リーダーは己が任されている場の運営について充分な配意をしなければならないのは言うまでもないが、それだけでは不十分である。このワークロードをうまく機能させるためには、たとえ異なる指示命令系統で動く者同士であろうとステークホルダー全員が組織横断的に業務をとらえている必要がある。そして、常に自らの属する階層よりも上位レイヤーのコンテクストを認識・共有していることが重要であると筆者は考えている。

こうした自身の業務経験から、本研究の基礎となる「共同活動のための創造的コミュニケーション」（高木，1994）及び「グループ意思決定プロセスにおけるコミュニケーションメカニズム」（北村、1992）に筆者は重大な関心と興味を持った。

当該研究は、観察対象となる慶應義塾大学大学院経営管理研究科修士課程（第 15 期、1992 年 4 月入学）の学生、8 名ないし 9 名で構成された５つのグループによる各々約２時間の討議模様を分析対象とし、会議におけるコミュニケーションの構造と行動

---

[i] Hamel, G. & Prahalad, C.K. “The Core Competence of the Corporation”, *Harvard Business Review*, Vol.68.

パターンを分析したものである。

北村 1992 のアプローチは次のとおりである。VTR を用いて会議模様を映像記録し、合意形成に至るまでのグループ各員の発言、及び会議の場がどのように遷移していくかを目視により確認する。次に、各参加者の発話について場の話題単位に区切るとともに、議事進行規則に関する言動（コンテクスト）と、内容に関する言動（コンテント）とに分類・識別する。そして、これらのコンテクスト間、及びコンテクスト・コンテント間の階層性に着目し、コミュニケーションの構造をパターン化した後、最終的にはこのコミュニケーションの構造と、当該グループの成績、すなわち会議での生産性の高低とを関連付けることで議論の参加者が配意すべきポイントについて提言をおこなっている。

筆者自身、2011 年４月に髙木晴夫教授による組織マネジメントの講義を受講し、当該研究を知ったのであるが、20 年前の研究が全く色褪せていないことに対する驚きと、人と人のコミュニケーションの態様にはまだまだ謎が潜んでいるのではないかという漠然とした思いが今日この研究を行う動機となった。

そして、本研究を行うにあたって先行研究に対する付加価値を２つ考えた。

まず、模擬的会議の議事の進行を表現するにあたり、会議参加者の動的なコミュニケーションを定量的に計測し、表現することはできないだろうかと考えた。もう一つは、この表現を実現する際に筆者自身、経験のないテキストマイニング技術を用いることであった。

これらに挑むことで、先行研究における知見に加えて、集団の意思決定の場面における、より高効率で、かつ、創造性の高いコミュニケーションを模索する上での新たな知見を得ること、また企業におけるリーダーの育成において何らかの役に立つインプリケーションが得られれば、というのが本研究における問題意識である。

1-2　研究の目的と概要

本研究の目的は、第一義的には集団の意思決定の場面における、より高効率で、かつ、創造性の高いコミュニケーションをいかに作り出すかという本質論に対して、直接的な解決を試みるものではなく、それらの阻害要因に関する新たな知見を得ることを目的としている。

先行研究がベイトソンの論理階型の理論を用いたアプローチをフレームワークの基礎としているが、本研究ではこれに言及しない。前述の会議開始から終了までの時間の経過を一塊の情報ととらえて分析を行う。そして時間の経過とともに構成員の間でどのような力動がみられるのかという視点で会議を分析し、仮説の立案を行うものである。

実現の方法として、組織マネジメントの研究分野ではあまりツールとして用いられてこなかった「テキストマイニング」の技術を利用することとした。

分析フレームワークとしては、類似度（結びつきの強さ）を示す Jaccard 係数を基礎としてクラスター分析を行う。また、頻度スペクトル(the frequency spectrum)をベースとしたタイプ・トークン比（Type-Token Ratio）採用し、可視化の手法として類似度行列を用いたデンドログラム による描画を試みた。

加えて、今回実施した分析手続を詳細に記録することが、組織マネジメント分野における新たな研究手法開発を導くものであるとの指導教授からの助言により、研究過程での失敗事例の記録、詳細な分析ツールの利用方法及びトラブルシューティングなども記録することとした。

1-3　研究の方法

1-3-1　観察研究の実施

観察対象となる模擬会議は、この研究の基礎となる（北村、1992）と同様に、慶應義塾大学大学院経営管理研究科修士課程（35 期、2012 年 4 月入学）の学生により 8 名ないし 9 名で構成された計 10 グループによる、概ね 1 時間半の討議模様をＶＴＲで録画したものである。この討議は、組織マネジメントにおける会議のコミュニケーションに関する進行構造を自ら体験することを通じ、複数の人々がいかにコンセンサスを形成していくかを理解するための授業として、経営管理研究科の清水勝彦教授の指導のもとで 2012 年 4 月 7 日に実施されたものである。

この模擬会議は、レジナルド・ローズ原作のＴＶドラマで、米国の陪審員制度をモチーフとした、映画「12 人の怒れる男」の開始から約 30 分程度、全体の 1/3 を見た後、その後の映画の展開を個人とグループの二回に分けて予想し、個人の予想とグループでの予想とで、正解との差分がどう変化するかを考察する。このグループ予想の作成という、集団のコンセンサス形成が会議の主たる目的となる。

この映画は、米国の陪審員制度をモチーフとしている。陪審員制度において、有罪・無罪の判定をする際、12 人で構成される陪審員が満場一致の時のみ有罪の判決が下される。この映画のストーリーは次のとおりである。まず、ヘンリーフォンダが演じる主人公であり、陪審員の一人である建築師が、ただ一人、検察の有罪という判断に対して疑義をとなえる。一方、残りの陪審員 11 人全員は、検察の示す有罪と判断した証拠を鵜呑みにし、スラムで父親を殺害したとされる少年を有罪と考えていた。多くの陪審員が早くこの場を終わらせて家路につくことを考えるなか、正義感に溢れる建築師は、検察の示した少年が有罪であるとした根拠について順次反駁を行い突き崩していく。それを見て、最終的には建築師以外の 11 人の陪審員全員が当初の有罪判決を次々と覆していくというストーリーである。

一方、学生は授業のなかでこの映画作られた年代の時代背景や陪審員制度とともに、最終的に陪審員全員が有罪判決を覆すという情報が与えられる。そして、前半約 30 分程度の映画を見たあと、登場する 11 人の陪審員の特徴、立ち振る舞い等から残り 11 人の陪審員がどういう順番で有罪から無罪へと判定を覆していくか各々予想を行う。次いでその結果をベースに、それぞれのグループ内で議論を行い、一つのグループ予想を作り上げる。このコンセンサスを形成する過程が本研究の対象となる。

1-3-2　記録作業の方法

このパートは、観察対象となった模擬会議の記録方法について述べるとともに、当該記録が本研究の目的達成において適切な情報を与えうるかに関する妥当性の検証に資するものである。

まず、収録前日に高木晴夫教授が別紙１の事前説明資料を用いて、オリエンテーションの際に学生向けに説明を実施した。この事前説明の目的は、記録映像が学生個人の成績に反映されないということを学生に伝え、模擬会議における発言に影響を与えないことを担保するための配慮である。

授業実施当日、個人の予測を行うための判断材料となる前半 30 分程度の映画鑑賞は、同日、同時刻に１つの会場に学生全員が集合し、清水勝彦教授のガイダンスの後、映画を見た。この学生が討議会場のグループディスカッションルームを離れている間、記録を実施するスタッフ５名が各々２グループ分担当する形で機材の設営を行う。使用した機材はビデオカメラと IC レコーダーを各グループディスカッションルームに 1 台ずつ運び、設置をおこなった。その際、学生がグループディスカッションルームに入る前に映像と音声の記録を開始することとし、記録者とは一切の接触を断っている。この目的は観察対象者に接触することで、何らかの会議におけるコミュニケーションに影響がもたらされることを排除するためである。記録はすべて機材のハードディスク内に記録された。VTR の記録ファイル形式は High Definition (HD) 画質 (MPEG-4 AVC/H.264 AVCHD 規格）は採用せず、スタンダードモード（MPEG-2 形式）の画質とした。理由は記録時間が約 90 分と長時間に上ることから、ハードディスク容量が HD の容量に収まらない可能性が高く、またそのあとの処理もブルーレイなどの容量でなければ保管もできないという物理限界があること。また、映像や音声ファイルを加工する際に PC での利用を想定する場合は、HD 形式では容量が膨大なボリュームとなってしまうため 10 グループ分の保存ができないことと、MPEG-2 形式でも十分な品質が担保できるとの判断による。結果、各グループとも 90 分程度の映像は 4 分割され、各 2 ギガバイト未満のファイルで保存可能[ii]となった。再生が録画した映像がその全ての会議は、１０個模様を 10 台のＶＴＲと予備的に 10 個のボイスレコーダに収めた。なお、これらのファイルには、記録開始以降のタイムスタンプ（時、分、秒）が同時に記録されている。

[ii] 観察対象とした実験的会議の映像ファイルおよび音声ファイルの計 10 グループ分は、慶應義塾大学経営管理研究科・高木晴夫研究室に保存されている。

1-3-3　アンケートの収集

模擬会議の開催前(別紙２)と、終了直後（別紙３）の二回に分けて、会議参加者全員からアンケートを徴収する。このアンケートは無記名とするが、グループ番号を記載することで、いずれのグループの模擬会議の構成員であったかについては特定可能とする。

このアンケートを徴収する目的は以下のとおりである。

- 事前情報の有無と、その情報がグループのコンセンサスに与える影響
- 模擬会議に参加者する学生の行動に対する目的意識
- グループワークの成果（正答率、満足度、発言数、貢献度など）意識
- 問題意識のズレがグループワークの成果に与える影響
- 満足度はどのように会議に対する行動へ影響をもたらしたか
- 目標をはっきり認識していたか、またそれらの有無の影響

これらを取りまとめて、結論に対する補助的な情報として役立てることとした。

1-3-4　テキストマイニングによる分析

本研究においては、模擬会議でのコミュニケーションを映像で一旦記録し、映像記録されたデータをもとに人手を介して日本語のテキスト化を行い、自然言語処理を試みる。こうした自然言語処理も含めていわゆる「言語学」という分野は、大きく分けて５つの研究分野がある。

図表１．言語学の研究分野

音韻論(phonology)	単に音と文字の関係、子音と母音の組み合わさり方などを明らかにする
形態論(morphology)	単語や品詞といった言語が意味を持つ最小の単位について、その語形変化や並び方などを研究する
構文論(syntax)	語や文節がどのような構造化しているのかを研究する
意味論（semantics)	１つひとつの「単語」がどのような意味をもっているかを研究する
語用論(pragmatics)	ある発話を取り上げ、その発言が文脈のなかでどのような働きをもっているかを研究する

出展：林俊克,2002　Excel で学ぶテキストマイニング入門

これら各分野の研究によりテキストマイニングの根幹を支える技術は、まずテキストを単語の出現頻度情報として数値化し、相関ルールやそれらが共起するルール（あるいは相関しないルール、または共起しないルール）などの確率計算を用いたデータマイニング手法が開発され、次第により複雑なアルゴリズムを用いた形態素解析や構文解析へと発展を遂げてきた。

つまり、テキストマイニングはデータマイニングの延長線上にあって、膨大なテキスト情報のデータの山の中から、仮説を発見することや、それを基に分析対象のモデルを構築するといった目的で技術開発が進められてきている（石井，2002)。

本研究においては、模擬的会議の討議模様を分析可能な議事録形式でデータベース化を図った後、このデータに対してこれらの言語学研究の成果を用いてテキストマイニング技術を試行することで膨大な情報の塊から、時間の経過とともにどのような力動がみられるのかを観察する。

従って、テキストマイニングを実施する上で、本研究では以下の検討をしなければならない。

① 会議情報の抽出

- 如何にしてテキストマイニング技術を用いるのか、また、分析対象となるデータベースの構築方針とその実行方法の検討

② 抽出した情報の解析

- データベース化した情報から、どのような手法で考察し、結果を理解するか解析手法等の検討

③ 解析結果の可視化

- 解析した結果を理解するための、若しくは解析し仮説構築を行うために如何にして表現するかの検討

1-4　観察研究の概要

観察研究のステップについては、当初、次の４段階のアプローチを計画した。

【当初計画した研究ステップ】

① 会議模様の議事録をテキスト化し、10 グループのうち半分の 5 グループでテキストマイニングツールを用いた議論の分析を実施する。

② 次に、その 5 グループのテキストマイニング結果に基づく何らかの発見により、多変量解析などで一般化のための仮説モデルの構築を試みる。

③ そして、10 グループのうち残りの 5 グループでモデルの検定を試みる。

④ プロフィール情報と連結、アンケートの情報などを総合的に分析し、何らかの知見がえられないか検討を行う。

しかしながら、想定外の事態がいくつも発生した。まず、最初のトラブルは、詳細は後述するが、会議模様のテキスト起こしの内容に不備があったことによる。この対応のため大規模な修正作業が発生し、データベース化にプラス約 1 か月を要した。

また、分析対象のテキストデータがテキストマイニングに不向きな形式であった。これは、会議の議事録のように文脈が整理されたテキストデータではなく、同意を示す「うんうん」などの感嘆詞や他者の割り込み発言などが頻発し、時間の経過を正確に表現すればするほど日本語として不完全な文章となることによる。

更に、テキストマイニングソフトの利用にあたり、用意したツールのいくつかの機能が実際に本研究では使えないことが判明した。対策としては別の無償ツールを利用することで対応することとした。

最後に、当初はキーワードを変数とみてモデル化を検討したものの、キーワードとして考えていた陪審員名の発生頻度が大きく異なり、グループの予測値に直接的な因果関係がみられそうもないことも明らかとなった。

これらの事態について、実現可能性の観点を重視し対応を検討した結果、最終的に実行した研究ステップは次のとおりである。

【実行した研究ステップ】

① 会議模様の議事録をテキスト化したDB作成。10グループのうち2グループ分を用いてテキストマイニングソフト[iii]の動作環境の確認とレポート出力を試行実施。ソフトウェア処理のためのデータ処理も各種試行した。

② 検証の結果、同テキストマイニングソフトの機能で不備があることが判明。このため、松村・三浦（2009）のＴＴＭ[iv]を利用し、ソフトウェアで自動化された手続きによることなく、逐次チェックが可能なマニュアル操作で実施する方針に変更。各種テキストマイニングの手法を試行した。また、年齢、性別などの属性情報を付加して多変量解析も試みた。

③ 模擬会議の実績データとテキストマイニング結果の考察を実施。合議により決定されたグループの意思決定結果と個人の決定のデータの関連と、上記②のテキストマイニング結果のデータとの関連を分析した。

④ 力動の分析において使用したフレームワークは、類似度（結びつきの強さ）を示すJaccard係数を応用したクラスター分析である。また、頻度スペクトル(the frequency spectrum)をベースとしたタイプ・トークン比（Type-Token Ratio）採用し、可視化の手法としては類似度行列を用いたデンドログラム[v]を採用した。

⑤ これらの結果を基に、前年度M34期の同一授業での模擬会議データを収集したうえ、追加でアンケート調査を実施し、仮説の検証を実施した。

[iii] MiningAssistant /R.2（株式会社ジャストシステム）

[iv] テキストマイニングツールTTM(Tiny Text Miner)　http://mtmr.jp/ttm/　詳細は第４章で述べる。

[v] デンドログラムの描画は、通常、距離のように値が小さいほど類似度が高いことを示す非類似度行列をベースに描画するが、分析フレームワークでJaccard係数という絶対値が大きい程結びつきが強いことを示す指標を用いたことから、類似度行列で描画を行うという試みを実施した。詳細は４章で述べる。

第２章　先行研究

2-1　　文献研究の概要

本研究では、先行研究（北村,1992）における認知論でのアプローチを用いることなく、組織マネジメント分野での新しい研究アプローチを模索することを一つの研究目的に掲げ、テキストマイニング技術を用いることによって新たな知見を得たいと筆者は考えた。このことから、文献研究に関しては従来のテキストマイニングを用いた分析的手続きを如何に技術的に応用するかという点に焦点を絞ることとした。

テキストマイニングは、データマイニングの延長として大規模なテキストの集合全体を処理することで個々のテキストだけからは得られない新たな情報や知識を発見する目的意識で開発が行われてきた（Frldman、Sanger,　2007)。そして、テキストマイニングは企業にとって単なる気付きを齎すだけではなく、現代においては企業活動の一部となっている。膨大な情報の中から特定の目的に関連する情報をシステムによって機械的に収集し、分析方法を一般化することにより、低コストで一定水準の比較可能な定量情報を入手することが可能となる。まさに、テキストマイニングは知識発見そのものではなく、発見知識に基づく効果的な施策策定に利用されている（池尾、井上，2008)。

既に活用例については多くの文献で紹介されているが、マーケティング分野における先進的な利用が多い。

一例として「テキストマイニングによるマーケティング調査」（上田、黒岩、戸谷、豊田，2005）の例を紹介する。

コールセンターを運営する株式会社ＮＴＴソルコは、電話応対を行った際にオペレータが記録したＣＲＭデータを基に、委託元企業に対する「問い合わせ内容」「クレーム」などを分析し、新たな商品開発のヒントやＣＳの向上に向けた提案にテキストマイニングを利用している。株式会社ティップネスでは、キャンペーンを実施した際に顧客からの施設や接客に関する手書き意見をデータベース化し、店舗ごとの頻出キーワード解析により、店舗運営のヒントを抽出している。

また、テキストマイニングにより抽出した情報に統計処理の技法をミックスしてイメージや感覚的な情報を定量化して示す工夫も例示されている。日経広告研究所のブランド連想モデル化研究会[vi]が日本の携帯電話４社に対するブランドイメージ

[vi] 2000 年 5 月に結成されたブランド連想の基礎研究に取り組んでいた、日経広告研究所ブランド連想分析研究会を前身とし、2004 年により具体的な評価方法の開発のための研究会として改称。（出展：上田、黒岩、戸谷、豊田，2005）

調査を行っているが、ここでは２つの異なる時点での自由回答による調査結果をもとに、頻出度分析を行ったあと、頻度の高いキーワードを基にＰＩＮＳ測定法[vii]を用いて回答者のブランドイメージを定量評価しているとともに、キーワードの出現傾向によって従来では定量化しにくかったもの、例えばブランドの「ユニークさ」を、コレスポンデンス分析により強度を分析することを試みている。この計算方法は以下の式で表される。

$$U_j = 1 + \frac{1}{\log n} \sum_{i=1}^{n} \frac{f_{ij}}{F_j} \cdot \log \frac{f_{ij}}{F_j} \qquad 0 \leq U_j \leq 1$$

但し、Ｆｊ：連想イメージ $_j$ の全頻度、fij：分析対象項目ｉの連想ｊの頻度、n：ブランド数

この方法は、ある連想が、全体でいくつ出現したか（Fj）と個々のブランドでいくつ出現したか（ｆｊ）のみから計算されるため、変数選択による影響を受けないことが特徴で、かつ、連想が最もユニークであるときに値が最大値になるように値の範囲を０から１に変換処理を行い U_j が１に近づく程、連想のユニークさが強く、0に近づく程弱くなるという形で計算結果を表現するものである。

もう一例は、SPSS 社のテキストマイニングツールを用いた分析例で、今まで述べた形態素ベースでの特定キーワードの頻出度による解析から、文法的な「係り受け」すなわち平文中の「語」と「語」の係関係の解析についての分析事例も紹介されている。これは、心情や本音などの感情表現を表す語に対して否定的表現、肯定的表現などの定義を行い、これらの語を含む文脈を機械的に抽出し、分類することでテキストに埋もれた細かな情報をくみ取るといった技術である。

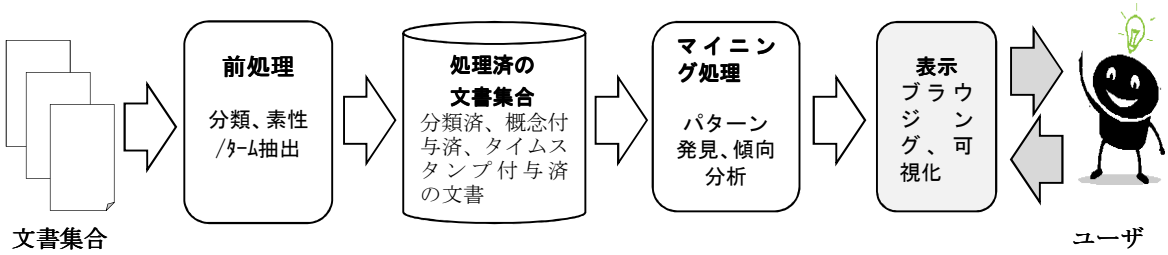
これら多くの書籍を見てみると、テキストマイニングの分析は前述のとおり２つで構成されていることがわかる。一つには、文書集合を形態素に分類し特定キーワードの頻出度から情報を解析するアプローチと、もう一つは前述の特定キーワードのうち感情や特定の背景などの意味を含む語彙に着目し、この頻出度の高いキーワード相互でどのような係り受け構造になっているかを分析するアプローチである。

[vii] ＰＩＮＳ測定法とは、提示した内容に対する回答者の連想イメージが「Positive：好ましい」「Indifference：どちらともいえない」「Negative:好ましくない」という評価をしてもらう方法である。（豊田裕貴，2003）

現在、いくつものテキストマイニングツールが販売されており、それらをソリューションとして提供する企業があるが、基本的な分析のアルゴリズムはその二つのアプローチの組み合わせで構成されており、結果をヒストグラム、クロス集計表、グラフ、デンドログラムなどで可視化の工夫が、そうしたツールの差別化になっていると思われる。

これらツールのアーキテクチャを図式すると以下のとおりとなる。

図表２．テキストマイニングのアーキテクチャ



出展：「テキストマイニングハンドブック」P249

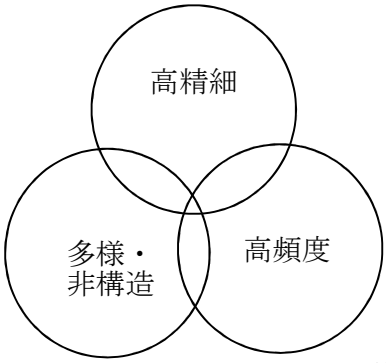
これらの手続きによってツールを利用するユーザは、膨大なテキスト情報から有益なキーワードを識別し、その利用目的に応じた処理をおこない、可視化することでテキストマイニング情報を活用するのである。

近年、ＰＣやスマートフォン市場の拡大で一般の消費者が生活のなかで無意識に大量のデータを生成し続けている。これを分析対象として、大手ＩＴ企業がビッグデータを活用するためのソリューションを相次いで開発している。近年、急速に一般化したソーシャルネットワーク（ＳＮＳ）の膨大な情報から、企業側にとって表面上は顕在化していないが、個の情報発信のなかから今後の自社の発展に有益な情報を抽出するためのツールにならないか、というものである。

一般に「ビッグデータ」という概念は未だ明確な定義がなされているわけではないが、一つの考え方として野村総合研究所ではビッグデータを以下のとおり定義している。

ビッグデータとは、「高精細、高頻度で生成され、多様性に富むデータ」のことを指す。また、これらにはデジタル処理されていない情報も含まれる。

図表３．ビックデータの概念図



高精細：取り扱うデータの解像度が高く個別要素に関するデータであること

高頻度：リアルタイムデータを取り扱うなど、取得・生成頻度が高いデータであること

多様・非構造：取り扱うデータが非構造なものを含む多様なデータであること

※出展：ＮＲＩ「ビッグデータ革命」

このビッグデータは、多様性に富んだデータであり、文章、音声、画像、映像などの非構造化データも大量に含まれている。

近年増加しているデータの８割は、従来型のリレーショナル・データベースが蓄積してきた構造化データとは異なる、文章、音声、画像、映像などの非構造化データである。

これらを収集･分析するには､グーグルが開発したマップリデュース(MapReduce[viii])のような多数のサーバーを活用して分散処理をする必要がある（増田有孝，柿木彰，神田晴彦，2012)。今後テキスト情報のみならず、こうしたビッグデータを対象としたデータマイニング技術も研究され実用化が進められていくことと思われる。

一方、前述のような従来からの研究成果を利用したマーケティングへの展開の現状から離れ、テキストマイニング技術を組織科学分野への応用に向けた研究が進められているのでこれらを次に取り上げたい。

[viii] Jeffrey Dean and Sanjay Ghemawat, 2004

2-2　組織科学へのテキストマイニング導入

先に取り上げた主にマーケティング分野におけるテキストマイニングの利用で発展してきた手法のうち、特に自由記述型アンケートを分析する際に利用するデータマイニングの手法を応用し、組織革新を認知的に研究するために利用した研究がある（喜田，2007)。

これは、アサヒビールの組織革新と経営成果について、組織革新を概念変化という認知現象から追及するのであるが、概念の量的な変化の一つである概念数の変化を、頻出度の高い名詞の数の変化と関連付け、各年次で経営成果に関する変数（シェア、売上高、経常利益）と概念数の関係についての相関分析を実施した。

具体的には、有価証券報告書の「営業の状況」の記述内容をデータベースに取り込み、テキストマイニング技術を用いて形態素分析実施し、概念の変化と紐付ける頻出度の高い名詞を識別するという手法を用いた。

結果、テキストマイニングで抽出した概念数は、経営成果に関する変数と正の相関があることが明らかとなった。この研究の成果は、認知的研究の課題の一つである知識及び認知の可視化の問題について一つの方向性を示すものといえる。

2-3　人工知能分野における研究

テキストマイニングの利用において、マーケティング分野での積極的な利用と、組織革新の認知変化への応用について前述したが、人工知能分野において本研究の目指す方向性に非常に近い、参考となる研究がなされている。

「議論構造の可視化による論点の発見と理解，(松村、加藤、大沢、石塚，2003)」において、議論の論点をすばやく発見し理解するためには、議論構造を可視化してユーザに提示することが有効であるとの立場で、議事録から話題の単位(セグメント)を同定し、さらに同定したセグメント間の関連を調べることにより、議論構造を構造化マップとして可視化するシステムを検討した。

その研究において、話題中の語が伝播するプロセスに着目して、話題の影響力の強さを測るモデルに「影響の普及モデル IDM(Influence Diffusion Model) 」があるが、これを用いて話題の構造化を図ることにより、議論の発展のトリガとなる話題を発見することを試みている。

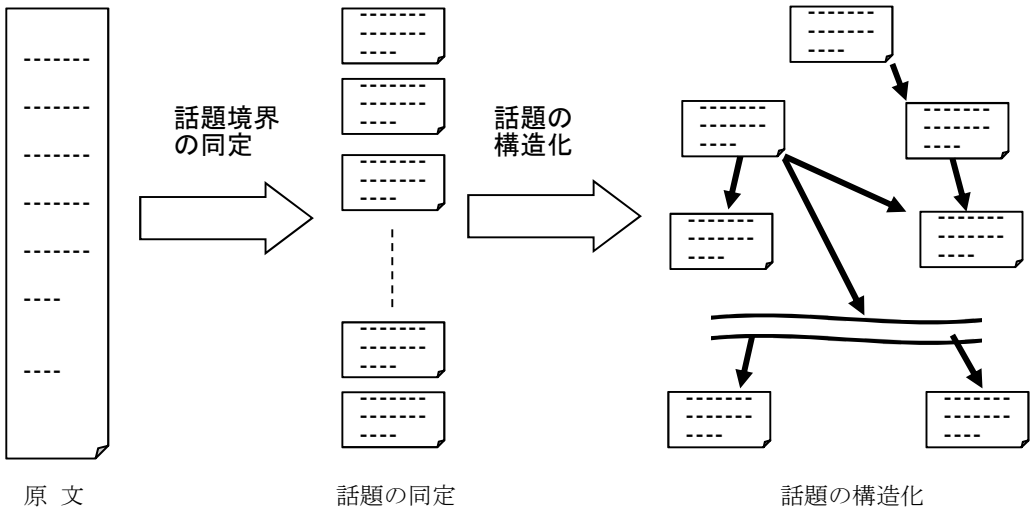
ここでの問題意識は、組織におけるミーティングや会議などでナレッジマネジメン

トの目的から作成される議事録が、多くの場合、長い議事録を読むだけの時間と労力がないことで実際に作成されても埋もれてしまうことが多いことにある。

この解決方法として「議論を発展させるトリガとなるような話題」を興味深い話題と考え、議事録における議論の流れを構造化して視覚的に表示することにより、議論全体の流れを読者が把握することを支援する、というものであった。

この研究では、話題と話題の関係がわかるように以下の一連の手続きで構造化を行っている。

図表４．議論構造の可視化のプロセス



出展：「議論構造の可視化による論点の発見と理解，（松村、加藤、大沢、石塚，2003)」

そして、最終的に議論構造の可視化システム DSV(Discussion Structure Visualiser)というシステムを提案するものである。

DSV は次の５つのステップで構成される。

Step1）形態素解析

入力テキストが日本語の場合、文章は分かち書きされていないので、形態素解析ツールにより必要な品詞だけを抜き出す。DSV はデフォルトでは名詞だけを残すが、オプションにより動詞、形容詞、副詞、未知語も選択可能。

Step2）話題境界の同定

与えられたウインドウサイズと切り出すセグメントの数に基づいて、入力テキストをセグメントに分割する。

Step3）話題の構造化

与えられた閾値に基づいて、セグメント間の関連を特徴ベクトルに基づいて計算し、議事録を構造化する。

Step4）IDM

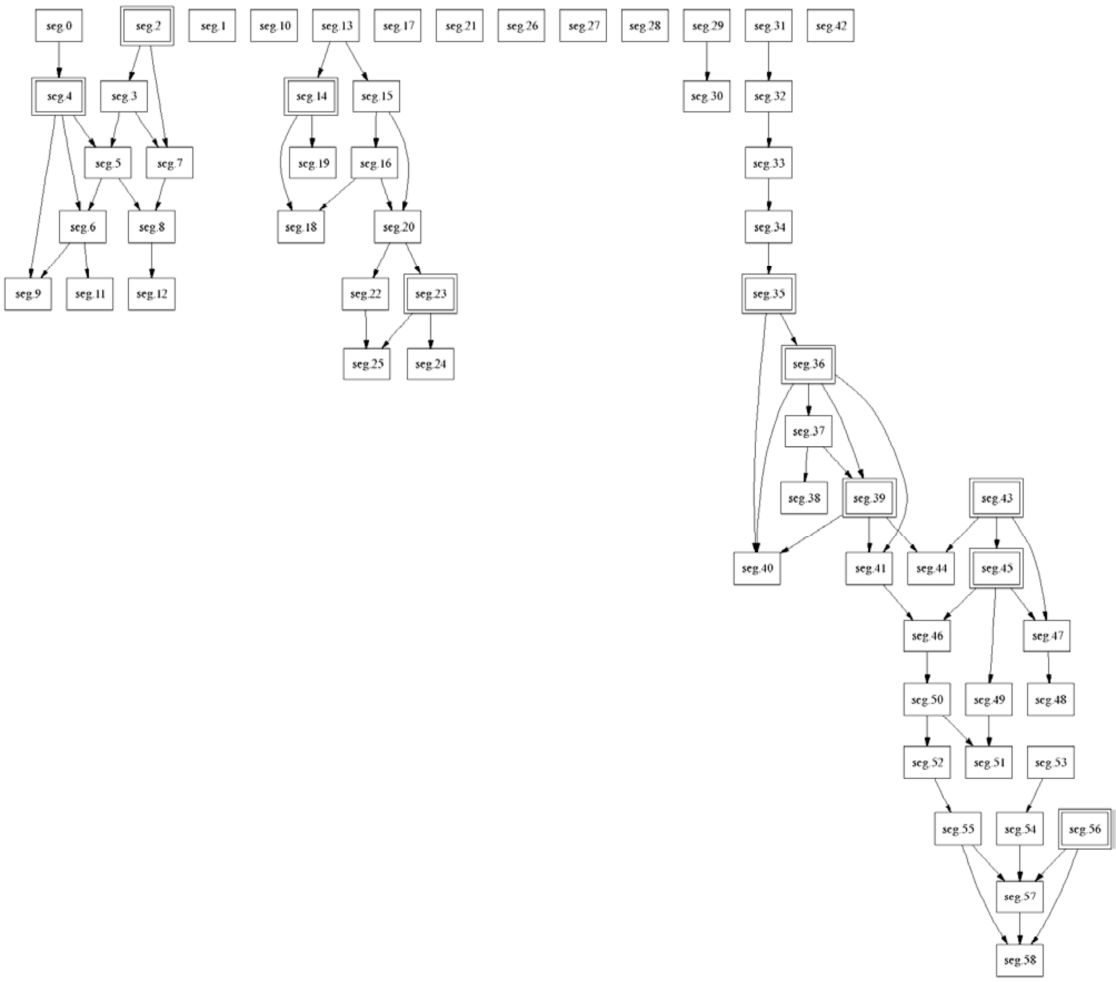
構造化された議事録に IDM を適用し、影響力の大きいセグメントを同定する。

Step5）可視化

構造化された議事録を可視化し、ユーザが議論の構造を把握できるようにする。また、このときに媒介影響量の大きいセグメントを目立たせ、ユーザの迅速な論点の発見と理解を支援する。そして、話題構造を可視化した図を構造化マップ[ix]と呼ぶ。（図表５参照）

[ix] 構造化マップは AT&T 研究所が開発して無償で配布しているグラフ描画ソフトウェア Graphviz を用いている。http://www.research.att.com/sw/tools/graphviz/

図表５．構造化マップの例



出展：「議論構造の可視化による論点の発見と理解，（松村、加藤、大沢、石塚，2003）」

これらの手続きにより、まず、議事録から話題の単位（セグメント）を同定し、さらに同定したセグメント間の関連を調査する。そして、次に議論構造を構造化マップとして可視化するシステムを構築し、構造化マップが表す話題の一連の流れを明らかにする。こうして結論にたどり着くまでの合意形成に至るプロセスについて読み手が理解することを支援するというものである。

この構造化マップを用いた可視化の手法は従来にないアプローチであり、発話の影響度を指数化する特徴ベクトルの算出など表現は非常に斬新な切り口である。

しかしながら、この手法に用いられた各指数等の算出は簡単ではない。本研究の過程で筆者はＴＴＭの構成ツールのバグ報告で松村氏にコンタクトする機会を得た。論文を読むだけではこれらの指数等を算出する具体的な方法が理解できなかったため、各研究で計算に何を利用したのかを伺ったところ、松村氏自身が必要の都度データ形状にあわせてプログラミングを行い、数値を作成したとのことであった。また、使用するプログラム言語もその都度選定しているとのことで誰でもが再現性をもって実施できる方法とは言えない高難度の手法であることがわかった。

本研究においては、組織科学におけるテキストマイニングの応用が研究目的のひとつである。従って、特殊なプログラム、あるいは高価なプログラムを使用して分析を行うのではなく、汎用性の高い分析手法を模索することも本研究の重要な課題であると考え、可能な限り低コストで汎用性のある方法論の模索に取り組むこととした。

第３章 観察研究の実施

3-1　　模擬会議の概要と適格性

本研究の観察対象となる模擬会議は、前述のとおり慶應義塾大学大学院経営管理研究科修士課程（35 期、2012 年 4 月入学）の学生 78 名により、経営管理研究科の清水勝彦教授の指導のもとで 2012 年 4 月 7 日に実施された。この授業は、組織マネジメントの授業として実施されたものであり、会議のコミュニケーションに関する進行構造を自ら体験することを通じ、複数の人々がいかにコンセンサスを形成していくかを理解するために計画されている。

授業では、前述のとおり映画「12 人の怒れる男」を題材として進められ、映画の最初 30 分間が放映されたあと映像は止められる。この映画のストーリーは、スラム街の少年が父親を殺害した容疑で逮捕され、裁判所において検察の示した有罪の証拠をもとに 12 人の陪審員達が有罪か否かを議論していく過程が描かれている。主人公である建築師が被告を有罪とするには疑問が残ると主張する。そして、検察の示した証拠の数々に対して自身が思った疑問点とその論証を示すことで最終的には残りの 11 人の陪審員全員が有罪から無罪へと判定を覆していくこととなる。しかしながら、映画の最初３０分間の映像の中に正解を特定することが可能となる情報はない。あくまで演じられた役柄上の個性、家族構成、職業等の情報をもとに、学生は建築師以外 11 人の陪審員がいかなる順番で有罪判定を覆していくのかを予測することとなる。

このとき、はじめに学生個人での予測を行い、その後、予め決められた１０のグループ､即ち８名又は９名で構成された集団で各々１つの予測値を合議決定する｡なお、この意思決定に与えられた時間は９０分間である。

最終的に各グループにおいて「個人毎の予想値の誤差」や「個人予想の誤差の平均値」に対して「グループで決定した予想値の誤差」を比較し、誤差の絶対値の大きさがどのように変化するかを考察する。すなわち、この模擬会議においてはグループ予想の作成という集団のコンセンサス形成が会議の主たる目的となっている。

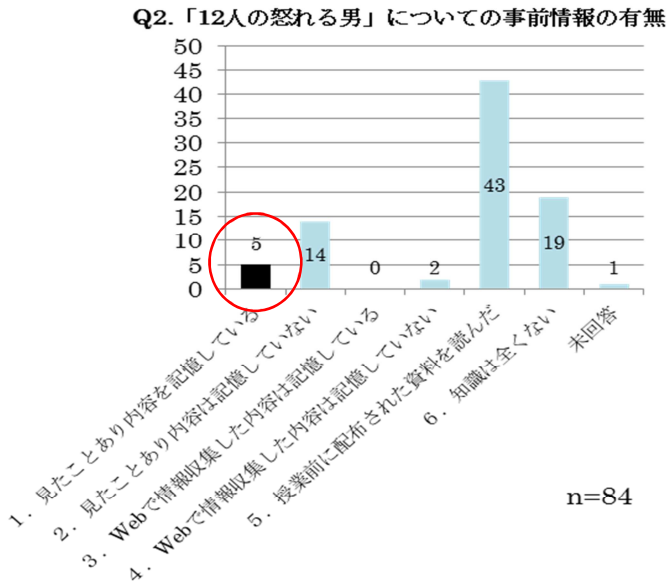
なお、冒頭述べた授業の目的を達成するため、この授業は学生が入学した直後に実施するよう計画されており、同じグループの各個人においてなんらかの人間関係が生じないよう配慮されている。これは本研究の適格性を担保する上で重要なポイントである。

3-2　　アンケートの集計

観察研究の分析に資するため、事前に各グループ室に別紙２、別紙３のアンケート用紙を配布して模擬討議の開始前と終了後にそれぞれ参加者全員の解答を得た。この集計結果は別紙８、別紙９のとおりである。

このアンケートにおいて、特に筆者が着目したポイントはQ2「12 人の怒れる男」についての事前情報の有無である。アンケートからこの事前情報の有無に関して集計した結果は図表６のとおりである。

図表６．映画に関する事前情報の有無



回収したアンケートより、赤丸の 5＋14＝19 名の参加者がこの授業を受講する以前に映画をみたと回答していることがわかる。また、このうち５名の参加者が「(映画を) 見たことがあり、内容を記憶している」と回答している。本授業の開始に先立ち、清水教授から「以前にこの映画を観たことのある者はグループ討議では発言を控えて参加する」旨、指示がなされている。しかし、グループでの討議においてこれらの発言がグループの意思決定になんらかの影響を与える可能性は否定できず、本研究においてこれらの参加者のデータはノイズとなる可能性が高いと想定される。

従って、録画記録された映像をもとに、これらの事前知識有りの者が誰であるかを特定し、データの取り扱いについては別途判断することとした。

3-3　模擬会議の記録とデータベース作成

3-3-1　模擬会議の記録

本研究において収録された記録映像の録画時間は各々約 90 分間である。この映像は、別紙５のビデオカメラ（SONY HXR-NX70J）を用いて MPEG2 形式の映像ファイルとして録画した。この記録は全部で１０グループ分あり、これを図表７の形式でデータベース化を行うこととした。

図表７．データベースの構造

通番	経過時間	経過時間累積	発言者	発言内容
1	0:15:06	8H	品川氏	そうすると最初に銀行員から見ていくのはいかがですか。
2	0:15:15	8D	渋谷氏	はい、いいと思います。
3	0:15:20	8A	大崎氏	順番をたとえば A から順に。基本的にみんなでカテゴライズされたのが、A からの順番で見ていくと、時系列がわかりそうな気がするのですが。

使用したソフトは、Microsoft Excel 2010 である。Excel を利用した理由は、時間の経過と発言者及びコメントの一覧性が高いこと、最終的に各種の分析ツールで処理する際に CSV 形式で出力する必要があること、そしてタイムスタンプや発言内容の補正処理等の利便性の高さが故である。

また、Excel97 以降のバージョンでは１つのセルに入力できる文字数の上限が 32,767 文字まで拡張されている。本模擬会議の形式からして、一人ひとりが発言する時間は限定的であり、オーバーフローするリスクも想定されないこともデータベースソフトを利用していない理由の一つである。

現在、存在するテキストマイニングツールの多くは CSV 形式で保存されたデータベースを用いて分析を行うことを想定している。このため、テキスト化を行う際に MS-EXCEL を利用し、必要となった都度 CSV ファイル形式に変換を行うこととした。また、属性情報などの分析の際、異なるデータベースとのリレーションが必要なときはマッチング・キーとして発言者名を使用することとした。後段で触れるが、各種の頻

度解析や、単位時間当たりのキーワード出現頻度などの計算も、このデータベースに列情報を追加し、Excel の関数をフル活用して算出した。

3-3-2　データベース作成

次に、映像ファイルから文字列への変換（テキスト化）方法について述べる。

研究計画の策定段階において、当初は録画映像の音声ファイルを基に音声自動識別技術[x]を利用することで、テキスト化については機械処理することを目論んでいた。製造元に問い合わせ、デモなどの過程を経て技術仕様を確認した。しかし、入手した会議模様の音声ファイルに記録されている音声データにおいて、参加者の最小発話時間が相当短く（音声自動認識のためには一つの発言で連続した発話が最低 10 秒以上、最適時間は 20 秒必要であった）、発話者の識別が困難であることが分かった。このため、機械処理を断念することとなり、非効率ながらもやむなくテープ起こし業者[xi]を利用することとした。このテープ起こし業者への依頼であるが、引き受け先を探すことに大変苦心した。理由は、９０分の映像×１０グループということで合計９００分もの長時間の書き起こしであること。また、更に条件が悪いことに会議スタイルが雑談形式のチャット状態であることから、音声や発言者の識別が著しく困難であることによる。このため、委託先へ示した入力フォーマットは図表８のとおりであるが、発言者が不明の場合は「？」を、また、発言内容で識別できなかったコメントには「〇〇〇」と記述するように依頼した。

図表８．テキスト書き起こし業者への委託フォーマット

発話時間	発言者	発言内容
0:15:06	品川氏	そうすると最初に銀行員から見ていくのはいかがですか。
0:15:15	？	はい、〇〇〇と思います。
0:15:20	大崎氏	順番をたとえば A から順に。基本的にみんなでカテゴライズされたのが、A からの順番で見ていくと、時系列がわかりそうな気がするのですが。

[x] 株式会社ボイススピリッツは、Panasonic の社内ベンチャー企業であり、独自の日本語の音声認識技術を用いた音声識別ソフト、ソリューションを販売している。URL : http://psuf.panasonic.co.jp/voicespirits/

[xi] 委託先：アステープ起こしタイナーズ http://www.tapeokoshi.net/

実際の発注方法であるが、原稿を送ったのち、４分割されたファイルの１つ分だけ入力完了後、一旦送付をしてもらい、品質確認を実施してから本発注とした。また、納品スケジュールについては映像ファイルを送付、到着後、２グループずつ２週間ごと、つまりそれぞれ 13 日で作成開始と納品を同時進行で進め、最終納品日は２ヶ月後とした。実際、委託先がテキスト化に要した期間は、記録 DVD の現地到着が 2012/5/19 で、最終納品日が 2012/7/20 と予定どおりの納品であった。

3-3-3　テキスト原稿の補正手順

次の作業工程は、逐次納品された原始データを筆者が記録 DVD と照合を行い、業者が識別できなかった発言者や発言内容について繰り返し映像ファイルを目視することで可能な限り不明箇所の補正を図っていった。具体的な進め方は以下のとおりである。

（１）原稿の印刷

・送付されてきた書き起こし原稿を A4 版の用紙の印刷設定を、横長、両面印刷、1 印刷ページ当たり２ページに設定して印刷する。

（２）チェック作業

・映像ファイルと、プリントした生の原稿の記載内容を目視で確認する。

（３）テキストの修正作業

①ワークファイル作成

・書き起こし業者から送付されてきた完成原稿（Excel ファイル）を開き、「【完成原稿】G##修正」というファイル名で保存。

②レコード通番の付与

・A 列に１列挿入し、3 行目からレコード通番を 1 番から最後のレコードまで通番を付与

③レコードの加筆修正、削除

・会議の開始前後の不要箇所や、誤記入、未判明箇所の補正、「(口々に話す)」及びデータベースに不要な情報を修正、削除

3-3-4　テキスト原稿の補正実施結果

作業開始後、委託先からは２週間ごとに２グループずつファイルが概ね予定どおり到着し、順調に修正作業は進められているかにみえた。しかし、最終納品直後、実は重大なトラブルが発生していることに気付くこととなる。

前述したとおり、９０分の映像ファイルはJPEG２フォーマットの制限から４分割されている。このため、書き起こしをしてもらったテキストファイルは図表７のフォーマットにあるとおり、発話時間はWindows Media Playerでファイルを再生したときに画面に表示される「X：XX：XX」という経過時間の記録である。一方、ファイルは４分割されていることから各ファイルの出だしは必ず０：００：００からのスタートとなる。委託先から送付されたファイルは、タイムスタンプに空欄は一つもなく、このため「？」や「○○○」の補正に集中するあまり、タイムスタンプのズレについて筆者は気付かなかったのであるが、累積時間を作るために連結を行ってみたところ、ほぼ全てのグループで「秒」表示部分「０：０▲：●●」の黒丸●部分（一部「分」▲部分）の記録にズレが生じていることを発見した。

単位時間当たりの発言量などは重要な情報である。やむなく筆者はこの補正作業を開始することとした。作業手順は不明箇所の修正同様、打ち出した委託先原稿の紙にWindows Media Playerでファイルを目視により確認しつつ、タイムスタンプのチェックを図るという非効率極まりない作業であった。

そして、更に一グループあたり数千レコードの議事録のタイムスタンプを確認したあとは、データベースの時間の記録を修正入力する必要がある。当初、１セルずつ文字修正により更新してみたが、１時間やっても一人のマンパワーでは1/10も処理できないことからExcel関数でタイムスタンプを分解するツールを作り「秒」だけを部分入力すれば修正が完了する環境を開発して作業を進めた。

しかしながら、この作業を完了するまでに８月一か月を費消してしまい、分析作業の時間を圧迫することとなった。これらの作業を経て１０グループ分の基本データベースを得た。なお、トラブルに関するテープ起し業者とのやりとりは別紙１２参照。

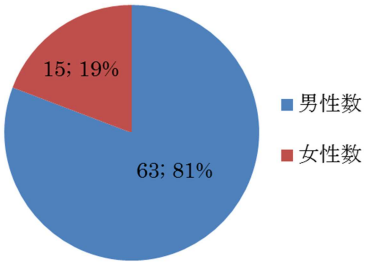
3-4　　予測データの収集

授業終了後、各グループ室から種々のデータを回収し、以下のとおり分析を行った。

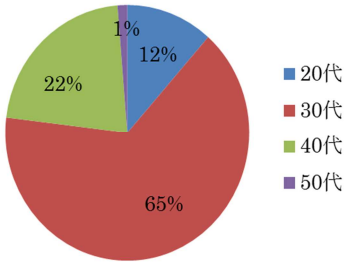
3-4-1　属性情報等

模擬会議の参加者 78 名構成要素は次のとおりである。

図表９．参加者の性別



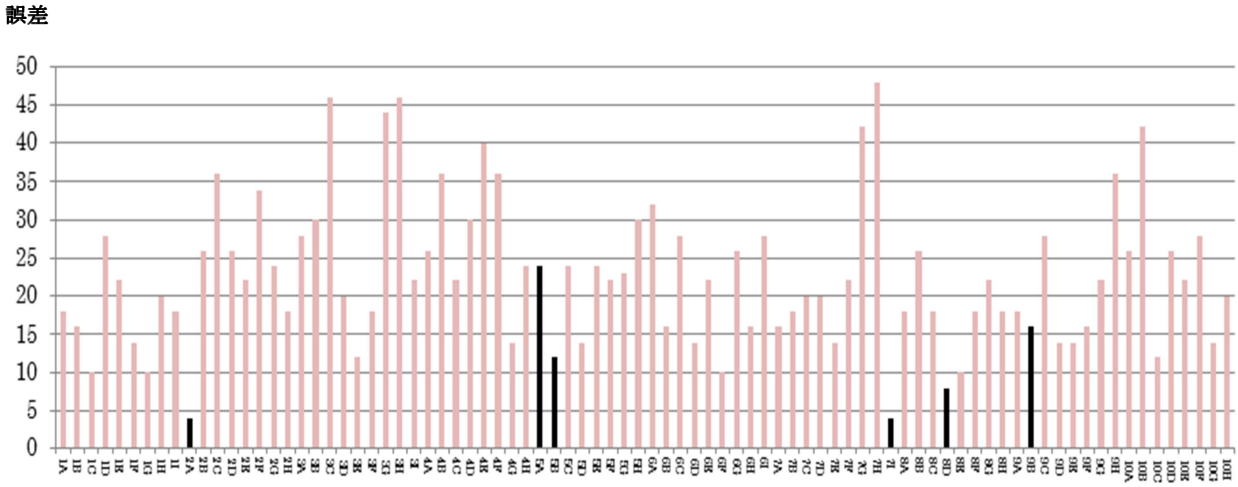
図表１０.参加者の年齢構成



3-4-2　個人の予測値の観察

各グループの記録を集計した結果、授業に参加した学生の個人別の予測と正解との誤差の絶対値の合計（以下、個人誤差とする）は図表１１のとおりとなった。

図表１１．個人別予測結果（個人誤差）



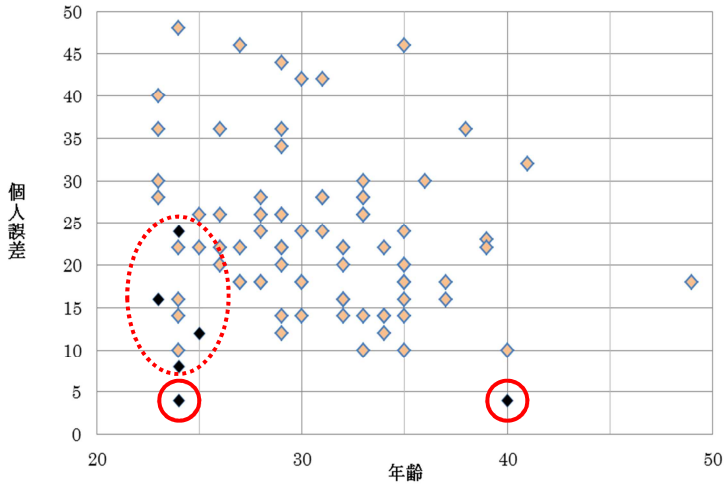
統計的にみると、誤差の全体平均は 22.6 ポイント、最小値は 4.0 ポイント、最大値は 48.0 ポイントと 12 倍の開きがある。また、中央値は 22.0 ポイントで分散は 9.6 ポイントとなっていた。

このとき、注意すべき事項がある。図表１１において、グラフ内で凡例が黒ベタとなっているサンプルが６名分（参加者：２Ａ、５Ａ、５Ｂ、７Ｉ、８Ｄ、９Ｂ）あることに気づく。この６名は、録画した映像記録から「事前に１２人の怒れる男の映画を観たことがある」と筆者が識別した参加者（以下、既知の参加者とする）である。前述の図表６において既知の参加者は自己申告では５名であったが映像をもとにプラス１名が既知の参加者であると識別した。この既知の参加者６名の誤差を見てみると、うち２名（参加者：２Ａ、７Ｉ）は参加者全体の最小値である 4.0 ポイントの個人誤差となっている。この結果は平均値、中央値及び分散からみても誤差が異常に小さいと思われる。従って、事前情報により個人予測値にバイアスがかかっている可能性が高いと言わざるを得ない。このことから、以降の分析に配意するとともに、グループの平均誤差の算定から除外することとした。

この結果、誤差の全体平均は 23.5 ポイント（＋0.9 ポイント悪化）、最小値は 10.0 ポイント（＋6.0 ポイント悪化）、最大値は 48.0 ポイント（変化なし）、中央値は 22.0 ポイント（変化なし）及び分散は 9.6 ポイント（△0.4 ポイント減少）となった。

また、個人別予測結果の誤差の分布について年齢の属性情報を付加すると以下の図表１２が描画できた。図表１１同様に既知の参加者についてはドットを黒ベタで描画している。

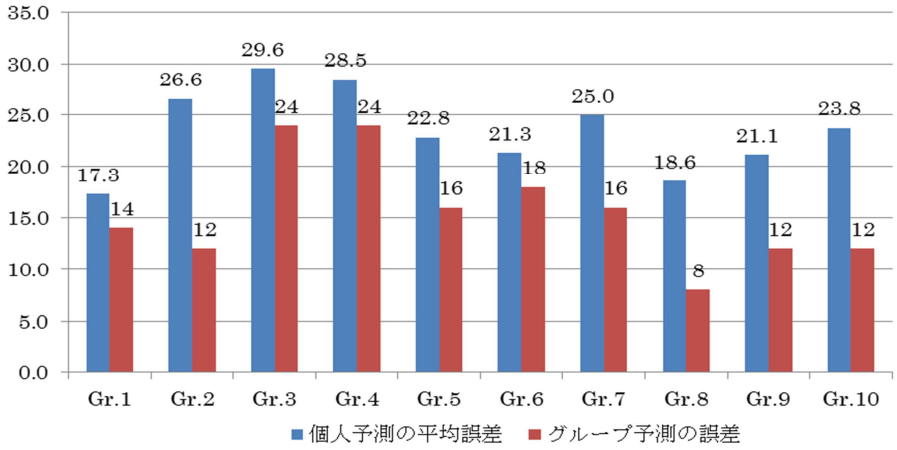
図表１２．個人別予測結果の誤差と年齢の関係



3-4-3　グループ別予測結果の観察

次に会議の結果得られた各グループの予測誤差がどのようであったかを観察する。各々のグループ予測値及び当該予測と誤差の絶対値の合計（以下、グループ予測の誤差とする）を集計した（別紙６、別紙７参照）。この結果得られたグループ毎の個人平均誤差とグループ予測の誤差を比較したものが以下の図表１３である。

図表１３.各グループ個人予測の平均誤差とグループ予測の誤差



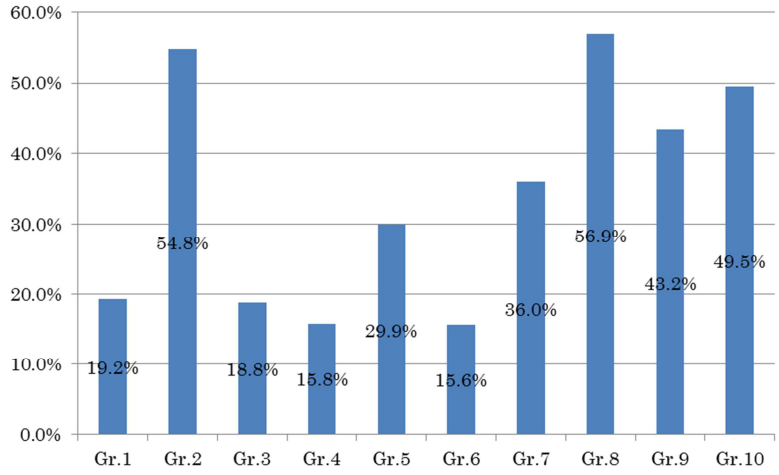
この図表１３の数値を基に個人予測の平均誤差とグループ決定の誤差とを比較すると、以下の図表１４が得られる。各々のグループの個人予測の平均誤差からグループ決定の誤差を引いたものを誤差の改善ポイントとし、個人予測の平均誤差とのウエイトを改善率とした。

図表１４.合議による個人予測の平均誤差の改善率

Gr.#	個人予測の平均誤差	グループ予測の誤差	誤差の改善ポイント	改善率	改善順位
Gr.1	17.3	14	3.3	19.2%	7
Gr.2	26.6	12	14.6	54.8%	2
Gr.3	29.6	24	5.6	18.8%	8
Gr.4	28.5	24	4.5	15.8%	9
Gr.5	22.8	16	6.8	29.9%	6
Gr.6	21.3	18	3.3	15.6%	10
Gr.7	25.0	16	9.0	36.0%	5
Gr.8	18.6	8	10.6	56.9%	1
Gr.9	21.1	12	9.1	43.2%	4
Gr.10	23.8	12	11.8	49.5%	3

各グループ別の改善率をグラフにしたものが図表１５である。最も改善率が高かったのはグループ８であり 56.9%の改善が図られた。次いでグループ２が 54.8%、グループ１０が 49.5%であり上位３グループでは誤差が約半分となった。一方、グループ６は 15.6%、グループ４は 15.8%、グループ３は 18.8%と２割に満たない改善結果のグループもみられた。

図表１５．グループ別改善率



観察の結果、グループ皆で議論を行うことで改善結果に大小の差がでているものの、全てのグループで平均誤差が改善するという結果が得られた。

また、各グループで最良の予測（個人予測の誤差が最も小さかった人の予測）の値とグループ予測の誤差の集計値を比較したものが以下の図表１６である。

図表１６．個人予測の誤差の最小値とグループ予測の誤差の比較

Gr.#	個人予測の誤差最小値(①)	グループ予測の誤差(②)	③＝②−①	if ③＞0= 価値創造？
Gr.1	10.0	14	−4.0	
Gr.2	18.0	12	6.0	>0
Gr.3	12.0	24	−12.0	
Gr.4	14.0	24	−10.0	
Gr.5	14.0	16	−2.0	
Gr.6	10.0	18	−8.0	
Gr.7	14.0	16	−2.0	
Gr.8	10.0	8	2.0	>0
Gr.9	14.0	12	2.0	>0
Gr.10	12.0	12	0.0	

この図表１６を注視して欲しい。ここでグループ２、グループ８及びグループ９において列③の値が正であることが確認できる。これは、グループで最も正解に近い予測をした人の予測よりも更に全体で議論した方が良い予測値（誤差が小さい）が得られたことを示している。

この結果は、協働作業による価値創造を示すと思われ、会議での議論によって予測精度が向上したと考える以外に合理的な説明ができないと思われる。先行研究の北村 1992 及び高木 1994 においては、これらの改善がおこるメカニズムについて言及をしてはいない。

しかしながら、髙木晴夫教授によると本授業の経験上、グループ決定が改善するのは約８割で、更に約２割の確率で図表１６に示したグループ構成員の最小誤差の値を超える誤差の改善を示すグループ決定がみられるという。

なぜ、このようなことが起こるのであろうか。非常に興味深い結果であるが、メカニズムの解明は非常に難しい。収集した事実関係から推測した仮説は後述とする。

第４章　テキストマイニングの実施と観察結果の分析

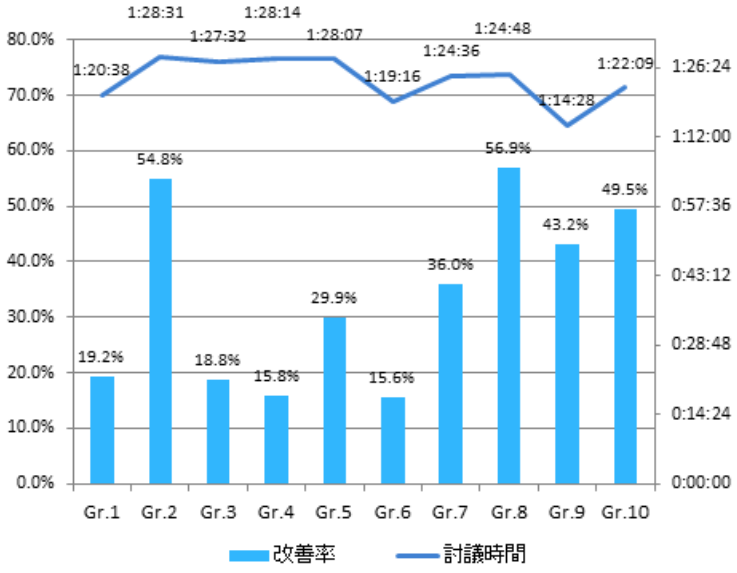
4-1　　模擬会議の実施模様に関する考察

映像記録、並びに書き起こしたテキストデータベースを基に各グループがどのように討議を進めたかについて順に整理する。

4-1-1　グループ討議時間

授業として計画されていたグループ討議の所要時間は１時間 30 分である。各グループが実際にグループ決定を導き出すまでに要した正味の討議時間を映像とテキストデータから計算し、討議時間と会議の生産性に関係があるかをみるため、図表１５の改善率とともにグラフ化したものが図表１７である。

図表１７．グループ別討議時間と改善率



最短のグループ９と最長のグループ２では会議時間に 18.8%の差がみられるが、改善率が 20%未満であった４つのグループ（グループ１、グループ３、グループ４及びグループ６）が必ずしも討議時間が短かったとはいえず、また約３割から５割以上の改善のみられた残りのグループにおいても討議時間が一様に長かったわけでもないことが解る。

4-1-2　各グループにおける討論の特徴

本研究は、会議における参加者のコミュニケーション構造と行動の特徴から集団の意思決定に与える影響のメカニズム解明を目的としている。先に述べたとおり、各グループの構成員同士はいずれも入学するまで面識はなく、かつ、知り合って直後の授業であることから未だ人間関係の構築ができていない。また、構成員に年齢や経験の差はあるもののこれらの情報は一切事前に与えられず、同期の学生という身分において上下関係はない。このため、お互いに誰がリーダーシップをとるか、討論を進めつつ手さぐりで議論が繰り広げられていくこととなる。このことが、本研究における観察対象としての妥当性を担保するものであることをあらためて記しておく。

筆者が全グループの映像記録を観た結果、会議終了までのプロセスにおける各グループのグループ討論の特徴等を取りまとめたものが図表１８である。

まず、司会進行役の有無については、グループ内で指名する場合、または自ら進行を買っている場合など形式的に識別できたグループは６グループ存在した。残りのグループでは、司会としての定義はされていないものの相槌も含めて発言頻度が多かったり、一回あたりの発言量が多いと感じられたりする場面も多く存在した。これらは次項のテキストマイニングの実施結果にて述べたい。

次に、各個人が予測した順位を書き出すなどしてグループ内の構成員間の平均値の順位を算定していたグループは４グループ存在した。このうち３グループにおいては、当該平均値をグループ決定の際にまずは優先し、最終的には議論による調整を図っていたが、残りの７グループはいずれも議論が一致するまで話し合いを行っている。また、合意形成の方法論としてすべてに適用されていた訳ではないが、多数決が用いられたグループが３グループ存在した。

個人の順位付けの結果を全員で共有する場面は全グループでみられた。

合意形成の過程で１２人の陪審員を３つの分類に区分けして順位付けを行う際に議論をし易くする工夫、すなわち有罪判決を覆すタイミングが比較的早いと思われる陪審員と、比較的遅いと思われる陪審員、その中間と思われる陪審員の３つに分ける作業（北村 1992 では「グルーピング作業」と定義）が、グループ２とグループ７及びグループ９を除く７つのグループにおいてこれらの作業が確認され、残りの３つのグループのうち、グループ７は陪審員を「論理的」か「感情的」か、という２分類で議論を進めていたが、板書上は中間にどちらとも言えない陪審員も記されており、広義には３つに分類するグルーピング作業を８つのグループが実施していたといえる。

図表１８．グループ間の議論の進め方等一覧

区分		G1	G2	G3	G4	G5	G6	G7	G8	G9	G10
改善順位		７位	２位	８位	９位	６位	１０位	５位	１位	４位	３位
司会進行役有無		×	○	○	○	○	○	○	×	×	×
			2B氏 →2D氏	3A氏	4B氏 →4G氏	5D氏	6A氏 →6H氏	7B氏			
構成員間の平均値算定		×	○	×	○	○	×	○	×	×	×
意思決定の方法		議論による	平均値優先	議論による	議論による	平均値優先	議論による	平均値優先	議論による	議論による	議論による
多数決での決定有無		×	○	×	×	×	○	×	○	×	×
グループ決定の分類学		3分類	N/A	3分類	3分類	3分類	3分類	2分類	3分類	N/A	3分類
		論理的、多数派、感情的	-	流され派、理由派、感情派	有罪寄り、中立的、無罪寄り	翻意時期：早い、中間、遅い	同情、どっち付かず、敵対	論理的⇔感情的	翻意時期：早い、中間、遅い	-	有罪寄り、中立的、無罪寄り
		有罪意識：強、中、弱マトリクス		価値観、偏見、同情	翻意時期：早い、中間、遅い						
個人決定の情報共有	時期	開始0-5分	開始0-30分	開始0-10分	開始0-20分	開始0-10分	開始0-10分	開始0-10分	開始0-10分	開始0-10分	開始0-10分
	方法	個人順位を読み上げ	PC投入 平均順位を板書	個人順位を板書	個人順位を板書	個人、平均順位とも板書	個人順位を板書	個人順位を読み上げ	個人順位を板書	個人順位を板書	個人順位を板書
	板書有無	×	○	×	○	○	○	×	○	○	○
	発表順序	特になし	特になし	特になし	特になし	特になし	特になし	乖離値を付けた人から理由を発表	乖離値を付けた人から理由を発表	特になし	特になし
		場の議論となった陪審員毎に	場の議論となった陪審員毎に	場の議論となった陪審員毎に	場の議論となった陪審員毎に	場の議論となった陪審員毎に	場の議論となった陪審員毎に			場の議論となった陪審員毎に	場の議論となった陪審員毎に
事前情報の有無		**N/A**	**○**	**N/A**	**N/A**	**○**	**N/A**	**○**	**○**	**○**	**N/A**
			2A氏			**5A氏、5B氏**		**7I氏**	**8D氏**	**9B氏**	

4-2　テキストマイニングの実施

本研究の主たる目的は、映像で記録した参加者の会話をテキスト化した後、テキストマイニングや定量化手法の適用を試みることで合意形成を目的とした模擬会議の模様を観察し、集団の合意形成が如何にして行われるかを観察・分析することにある。そして、この力動について可視化するための方法を模索するというのがテーマであった。この実現のため、筆者は多くの TRY&ERROR を繰り返した。本来であれば割愛すべきところであるが、後進のためにまずは失敗のプロセスから記すこととした。

4-2-1　テキストマイニングツールの試行

研究計画の作成段階において、筆者はテキストマイニングツールの借用について「TRUE TELLER®」を有する株式会社野村総合研究所、「見える化エンジン™」を有する株式会社プラスアルファ・コンサルティング及び「TRUSTIA/R.2」を有する株式会社ジャストシステムに協力要請を行った。テキストマイニングツールとしては最高峰の TRUE TELLER®は１ライセンス 600-900 万円、見える化エンジンは ASP で月額最低でも 10-15 万円/１分析セグメントで 10 グループ分だと単純計算で 10 倍、TRUSTIA/R.2 は基本 72 万円、アカデミックでも 36 万円といずれも非常に高額なソフトウェアである。研究計画を説明した時点では分析対象のテキストデータベースが手元になく、具体的な動作の検討や費用の相談などが出来ない状況であった。このため、分析データが入手した後、具体的な相談ができるときに再度コンタクトをとることとなった。

しかしながら、その後、研究計画の説明のために副査の井上哲浩教授及び高橋大志准教授の下を訪ねたところ、井上教授からは「高額なソフトを一発勝負で貸してもらったとしてもテキストマイニングは何度も条件を変えて分析を行う必要がある。手元にあるソフトウェアでなければ研究を完遂するのは難しいのではないか」、更に高橋准教授からは「テキストマイニングは「気付き」を得る技術。結局、道具は所詮気付きのヒントを得るための道具であり、何度も実際の映像を見ることで知見が得られることとなる。」という助言を受けた。加えて、幸運にも株式会社ジャストシステムより有期にて TRUSTIA/R.2 の無償貸与を得たことにより、まずは無料で入手可能なソフトウェアを基に分析を行うこととした。

この結果、はじめにテキストマイニングで使用したソフトウェアは TRUSTIA/R.2、無償のソフトウェアとして入手可能な茶筅 ver.2.3.3[xii]、エディター用には既に利用

[xii] 奈良先端科学技術大学院情報科学研究所自然言語処理学講座（松本研究室）が公開しているサイト http://chasen-legacy.sourceforge.jp/より入手。サイト内にある chasen for windows より、

している Microsoft Excel2010 及び Word2010 を使用した。それぞれインストールを終え、テキストデータが少しずつ出来上がっていくところで順次、茶筅による形態素解析や TRUSTIA/R.2 の使用を試みていった。

ここで、改めてテキストマイニングという技術を考える。

既に述べたとおり、テキストマイニング自体は特に新しい技術ではなく、既に多くの企業において現場でその成果が生かされている技術である。このテキストマイニングには基本的に３つの要素が存在する。第一に、情報の抽出。第二に、情報の解析。そして第三に、解析した結果の可視化である。

まず、一番目の情報の抽出で躓きが発生した。本研究における分析対象は、実際の人と人の会話を、時間の経過を意識しつつ、正確に書き起こした特殊な日本語の文書であった。テキストマイニングソフトを利用して機械処理を行うには、ノイズを除去したり、機械処理可能な状態に補正したりする前処理が必要となる。つまり、会話であるが故の割り込み、頷きなどの感嘆詞の存在など、日本語文法上で完結した文章にはなっていない部分がたくさん存在するため、単純に CSV 化したデータをテキストマイニングソフトに流し込むだけではツールは分析してくれないのである。

これはマーケティングなどで行われるアンケート、コールセンターに寄せられた情報、メールなどのある程度完成した日本語を対象としたテキストマイニングとは根本から異なることを意味する。実際、これらの困難を筆者は想定していなかった。最終的には断念することとなるのだが、最初、発言者が発言を終了するまでに割り込んで入った「頷き」や「どもり」あるいは「割り込み発言」などを手作業で補正し、ツールにかけることを試みた。しかし、わずか３０分程度のテキストを補正するだけでも丸一日必要で、これを１０グループ分全て行うのはマンパワー的に限界があると悟り作業をやむなく中断した。

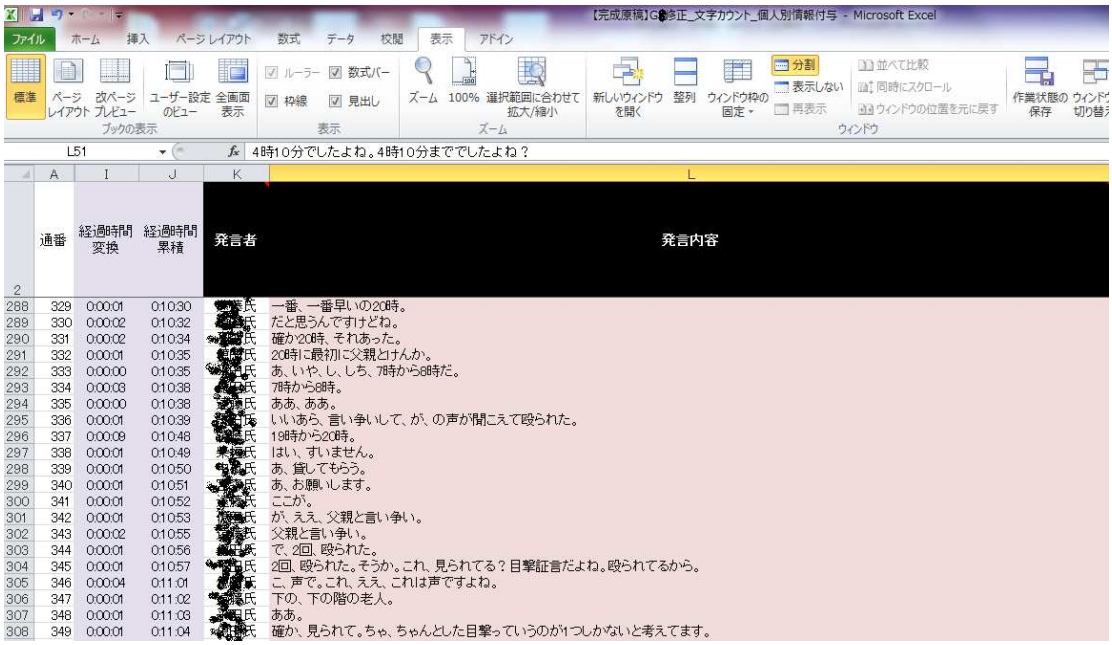
その後、本研究の対象である非定型文のテキストデータベースでは構文解析までを自動で行うことは無理があると判断し、頻度解析に重点を移すことに方針を変更した。そして、データベース上のテキストで不揃いであったキーワードとなる陪審員の名前統一と発言者の解明を可能な限り実施した後、茶筅を用いた形態素解析を行い、ノイズになる部分を排除した状態にすればエッセンスのみをテキストマイニングソフトが拾うと想定されることから、第一段階ではこの方法を試すこととした。

ChaSen-2.4.2-1-ipadic-sjis-2.7.0.EXE をダウンロード。

4-2-2　データベースの作成

最初に作成した Excel データベースは図表１９のとおりである。

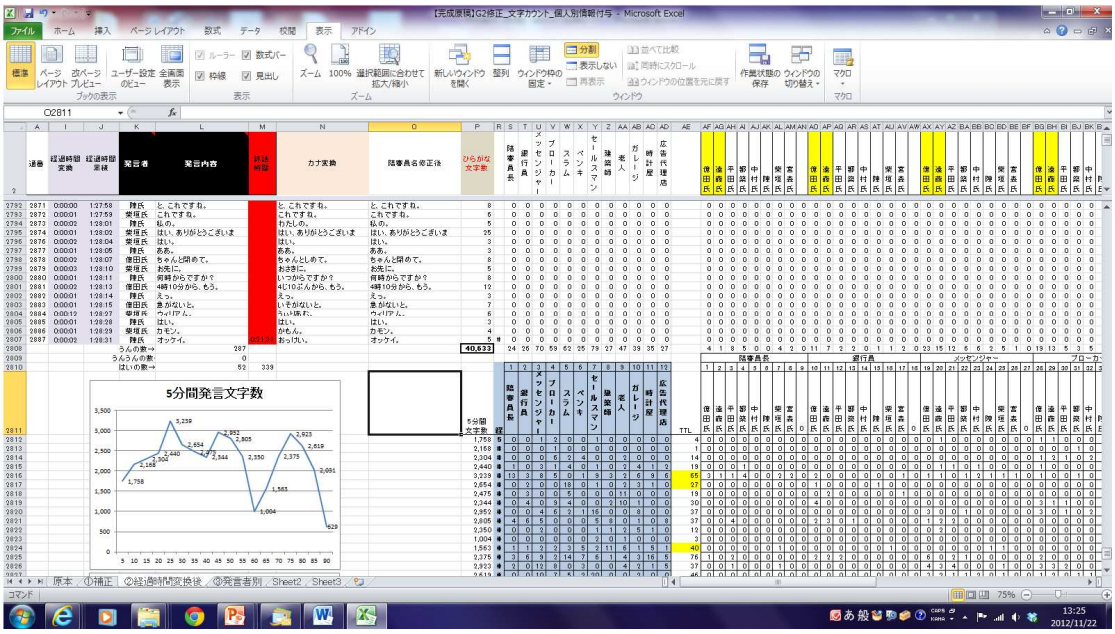
図表１９.最初に作った Excel データベースの画面ショット



	通番	経過時間変換	経過時間累積	発言者	発言内容
288	329	0:00:01	01:0:30	[redacted]氏	一番、一番早いの20時。
289	330	0:00:02	01:0:32	[redacted]氏	だと思うんですけどね。
290	331	0:00:02	01:0:34	[redacted]氏	確か20時、それあった。
291	332	0:00:01	01:0:35	[redacted]氏	20時に最初に父親とけんか。
292	333	0:00:00	01:0:35	[redacted]氏	あ、いや、し、しち、7時から8時だ。
293	334	0:00:03	01:0:38	[redacted]氏	7時から8時。
294	335	0:00:00	01:0:38	[redacted]氏	ああ、ああ。
295	336	0:00:01	01:0:39	[redacted]氏	いいあら、言い争いして、が、の声が聞こえて殴られた。
296	337	0:00:09	01:0:48	[redacted]氏	19時から20時。
297	338	0:00:01	01:0:49	[redacted]氏	はい、すいません。
298	339	0:00:01	01:0:50	[redacted]氏	あ、貸してもらう。
299	340	0:00:01	01:0:51	[redacted]氏	あ、お願いします。
300	341	0:00:01	01:0:52	[redacted]氏	ここが。
301	342	0:00:01	01:0:53	[redacted]氏	が、ええ、父親と言い争い。
302	343	0:00:02	01:0:55	[redacted]氏	父親と言い争い。
303	344	0:00:01	01:0:56	[redacted]氏	で、2回、殴られた。
304	345	0:00:01	01:0:57	[redacted]氏	2回、殴られた。そうか。これ、見られてる？目撃証言だよね。殴られてるから。
305	346	0:00:04	01:1:01	[redacted]氏	こ、声で。これ、ええ、これは声ですよね。
306	347	0:00:01	01:1:02	[redacted]氏	下の、下の階の老人。
307	348	0:00:01	01:1:03	[redacted]氏	ああ。
308	349	0:00:01	01:1:04	[redacted]氏	確か、見られて。ちゃ、ちゃんとした目撃っていうのが1つしかないと考えてます。

その後、キーワード検索や個人別の発言量などを経過時間５分間毎の発言量などを算出できるように変化させ、最終的には次のような体裁のデータベースとなった。

図表２０．最終的な Excel データベースの画面ショット

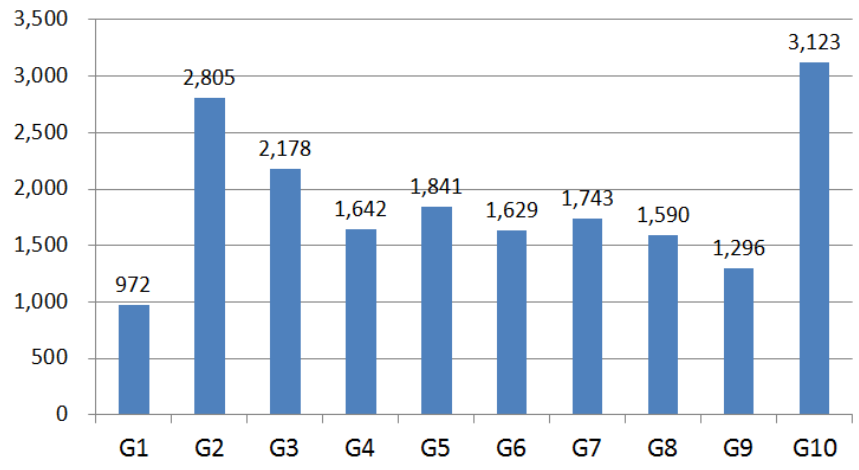


このExcelデータベースは磁気ディスクに記録され、髙木晴夫研究室に保管されている。

4-2-3　レコード数のカウント

次に全体感を掴むため、分析対象のテキストデータが全体でどれ位のボリュームになるかレコード数を指標として調査した。各グループのデータベースに通番を付与し、グループ討議の最初と最後の雑談部分を識別したあとグループ１からグループ１０までの全てのレコード数（記録された発言の行数）を合計すると、延べ18,819レコードとなった。グループ別のレコード数は図表２１のとおりである。

図表２１．グループ別の発言レコード数



結果、グループ１が972レコードと最小値であった。一方、最大値のグループ１０は3,123レコード、二番目に発言レコード数の多いグループ２も2,805レコードであり共にグループ１と比較すると３倍以上の開きがある。この差は一体何から生じるのであろうか。映像とデータベースを比較観察すると大きく２つの要因があることが判明した。

一つ目の要因は、これら発言レコード数の多いグループの特徴は「うん」「うんうん」などの「頷き役」が居るか居ないか、この差が原因の一つであった。実際に「うん」「うんうん」「はい」の３つの相槌の件数をカウントしてみると、グループ１は僅か47回にとどまるところ、グループ２では339回、そしてグループ１０では実に10

倍以上の 491 回を数えた。

次いで二番目の要因であるが、これは定量的に計測することが困難であるため筆者の主観的定性評価となってしまうが、映像の観察の結果分かったことにレコード数の多いグループの議論の進め方には前の発言者が発言を終了しないうちに発言を開始する「割り込み発言」が比較的多くみられるという印象があった。

4-2-4　TRUSTIA/R.2 の試行、限界

この出来上がったデータベースのテキストデータをもとに、TRUSTIA/R.2 のテキストマイニングモジュールである MiningAssistant でキーワードの頻出度解析を行うこととする。

キーワードとして筆者がまず着目したのは１２名の陪審員名が同じくらいの頻度で会話に登場しているかどうかである。この理由は、筆者が想像した「改善率の高いグループ討議」は 12 名の陪審員が翻意する順位について、「12 人の陪審員の何れに対しても、同じくらいの時間をかけて慎重にグループで討議が行われることが、より良い予測結果につながるのではないか？」という仮説に基づく。

それ故、最初の作業は陪審員名の平仄合わせである。頻度カウントの際、各々が発した不揃いな呼び名でテキスト化されている陪審員名を統一する必要がある。このため、図表２２のとおりデータベースのテキストを修正した。

図表２２．陪審員名の呼称統一の平仄一覧

正式な陪審員名		統一した陪審員名
陪審員長	→	陪審員長
銀行員	→	銀行員
メッセンジャー会社オーナー	→	メッセンジャー
株式ブローカー	→	ブローカー
スラムから来た男	→	スラム
ペンキ職人	→	ペンキ
セールスマン	→	セールスマン
建築師	→	建築師
老人	→	老人
ガレージの主人	→	ガレージ
時計屋	→	時計屋
広告代理店社員	→	広告代理店

次に、データベースの元データから助詞、助動詞、記号、句読点などの不要なノイズを除去する作業である。これには形態素解析ツール茶筅を利用した。茶筅により識別される品詞区分は図表２３のとおり。分析の対象に応じてこれらの品詞がキーワードとして利用可能かどうかを判断し、不要なものを除去する。この手続きにはExcelとWordを利用した。

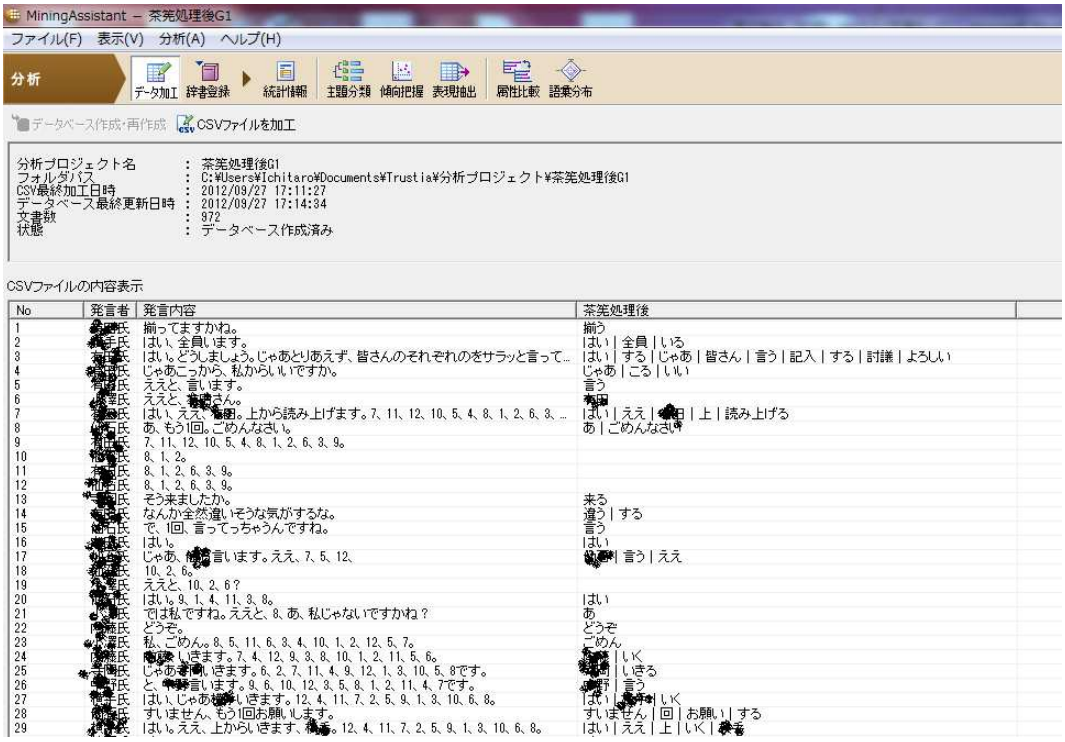
図表２３．茶筅により識別される品詞区分

品詞区分	詳細区分	キーワードとして適切かどうか
記号		キーワードとして意味を持ちそうにない
形容詞	自立	キーワードとして絶対に必要
	接尾	
	非自立	
助詞	格助詞－一般	キーワードとして意味を持ちそうにない
	格助詞－引用	
	格助詞－連語	
	係助詞	
	接続助詞	
	副詞化	
	副助詞	
	副助詞－並立助詞－終助詞	
	並立助詞	
	連体化	
助動詞		キーワードとして意味を持ちそうにない
接頭詞	名詞接続	キーワードとして意味を持ちそうにない
動詞	自立	キーワードともとれる表現を多く含んでいる（判断が分かれるところ）
	接尾	キーワードとして意味を持ちそうにない
	非自立	キーワードとして採用した表現もあるが、解釈が難しい語を多く含んでいる
副詞	一般	キーワードとして採用した表現もあるが、解釈が難しい語を多く含んでいる
	助詞類接続	キーワードともとれる表現を多く含んでいる（判断が分かれるところ）
名詞	サ変接続	キーワードとして絶対に必要
	一般	
	形容動詞語幹	
	固有名詞－組織	
	数	キーワードとして採用した表現もあるが、解釈が難しい語を多く含んでいる
	接尾－一般	
	接尾－助数詞	
	接尾－助動詞語幹	
	接尾－特殊	
	代名詞－一般	
	非自立－一般	
	非自立－副詞可能	
	副詞可能	
連体詞		キーワードとして採用した表現もあるが、解釈が難しい語を多く含んでいる

（出展：Excelで学ぶテキストマイニング入門）

これらの手作業を経た後、CSV 形式で保存し、TRUSTIA/R.2 のテキストマイニングモジュールである MiningAssistant で読み込んだ結果が図表２４である。

図表２４．MiningAssistant の CSV ファイル読み取り後の画面ショット

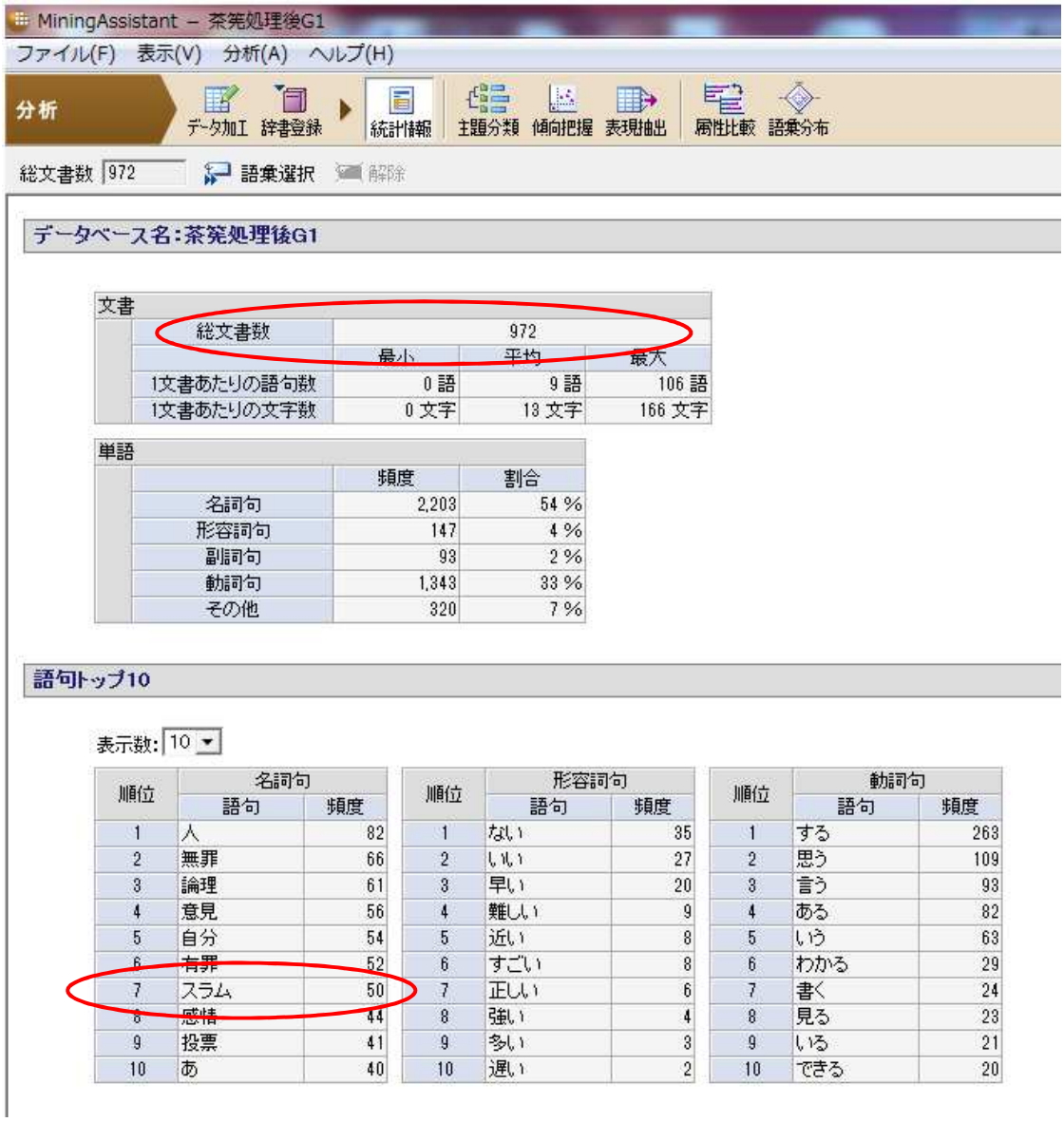


この後、テキストマイニングのエンジンが入っている MiningAssistant で統計処理をしたあとの画面が図表２５である。

表中にある赤い丸印で囲った部分を見てほしい。総文書数 972 が先に述べた「図表２１．グループ別の発言レコード数」におけるグループ１のレコード数 972 と合致しており、適正にインプットされていることが確認できる。

一方、MiningAssistant で統計処理を行うと文書内の各品詞別の単語頻出度を集計したサマリが自動表示される仕様となっている。これは、４つのサイズで集計が可能であり、上位 10,20,30,50 の区分で集計表示が可能である。図表２５においてトップ 10 までを見る限り、キーワードである陪審員名は「スラム」一つだけしか表示されておらず、使用頻度 50 回、名詞句のなかでは第７位のランクであることがわかる。

図表２５．MiningAssistant による統計処理後の画面ショット



グループ１において、名詞句のキーワードとして１２人の陪審員全員がどのような頻度で語られていたのかを調べてみる。

プルダウンメニューで表示数を変更し、１２名全員の陪審員が表示されるまで「語句トップ 10」を順次「語句トップ 20」「語句トップ 30」「語句トップ 50」と拡大した。結果、スラム 50 回、陪審員長 39 回、建築師 36 回、老人 34 回、ペンキ職人 32 回、株式ブローカー31 回、ガレージ 30 回、時計屋 26 回、メッセンジャー25 回、銀行員 24 回、セールスマン 19 回、そして広告代理店 11 回であった。

図表２６．「語句トップ３０」＋「語句トップ５０」抜粋の画面ショット

語句トップ30

表示数: 30

順位	名詞句 語句	名詞句 頻度	順位	形容詞句 語句	形容詞句 頻度	順位	動詞句 語句	動詞句 頻度
1	人	82	1	ない	35	1	する	263
2	無罪	66	2	いい	27	2	思う	109
3	論理	61	3	早い	20	3	言う	93
4	意見	56	4	難しい	9	4	ある	82
5	自分	54	5	近い	8	5	いう	63
6	有罪	52	6	すごい	8	6	わかる	29
7	スラム	50	7	正しい	6	7	書く	24
8	感情	44	8	強い	4	8	見る	23
9	投票	41	9	多い	3	9	いる	21
10	あ	40	10	遅い	2	10	できる	20
11	陪審	39	11	よい	2	11	行く	18
12	建築	36	12	弱い	2	12	入る	18
13	老人	34	13	良い	2	13	変わる	17
14	ペンキ	32	14	近しい	2	14	違う	16
15	ブローカー	31	15	やすい	2	15	変える	15
16	話	30	16	よろしい	1	16	やる	15
17	ガレージ	30	17	いかつい	1	17	える	12
18	主張	28	18	理屈っぽ	1	18	分ける	12
19	意識	26	19	うまい	1	19	しれる	11
20	時計	26	20	やばい	1	20	いく	11
21	メッセンジャー	25	21	ものすごい	1	21	決める	11
22	最初	25	22	重い	1	22	なる	11
23	銀行	24	23	うれしい	1	23	持つ	10
24	あと	24	24	大きい	1	24	聞く	10
25	議論	22	25	疑わしい	1	25	言える	9
26	感じ	21	26	低い	1	26	入れる	9
27	最後	20	27	細かい	1	27	出る	9
28	理屈	19	28	遠い	1	28	まとめる	9
29	セールスマン	19	29	面白い	1	29	みる	8
30	無記名	18	30	おもしろい	1	30	くる	8
中　略								
40	順番	13	40			40	思える	5
41	場	12	41			41	決まる	5
42	時点	12	42			42	ねる	5
43	広告	11	43			43	比べる	5
44	手	9	44			44	投げる	4
45	順位	9	45			45	話し合う	4

続いて、データの正確性を担保するため MiningAssistant の統計処理結果について検証を行う。グループ１の Excel データベースにおいて「=LEN()」「=SUBSTITUTE()」

及び「=COUNTIF()」を使用し、１レコード毎の発言者の発言から 12 名の陪審員名が何回発言されているかをカウントした結果、図表２６にある数値との一致を確認した。

図表２７．グループ１の Excel データベース画面ショットと集計

S86　=(LEN($O86)-LEN(SUBSTITUTE($O86,S$2,"")))/LEN(S$2)

	A	I	J	K	L	O	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD
1	■グループ1	9	10	11	12	15	19	20	21	22	23	24	25	26	27	28	29	30
2	通番	経過時間変換	経過時間累積	発言者	発言内容	陪審員名修正後	陪審員長	銀行員	メッセンジャー	ブローカー	スラム	ベンキ	セールスマン	建築師	老人	ガレージ	時計屋	広告代理店
84	185	0:00:01	0:07:37	[redacted]氏	セールスマンもそうですね。	セールスマンもそうですね。	0	0	0	0	0	0	1	0	0	0	0	0
85	186	0:00:02	0:07:39	[redacted]氏	うん。	うん。	0	0	0	0	0	0	0	0	0	0	0	0
86	187	0:00:01	0:07:40	[redacted]氏	あと、1番の陪審員長が意外	あと、1番の陪審員長が意外	1	0	0	0	0	0	0	0	0	0	0	0
87	188	0:00:03	0:07:43	[redacted]氏	ああ、そうですね。	ああ、そうですね。	0	0	0	0	0	0	0	0	0	0	0	0
88	189	0:00:01	0:07:44	[redacted]氏	投げやりな、まとめるだけで。	投げやりな、まとめるだけで。	0	0	0	0	0	0	0	0	0	0	0	0
89	190	0:00:04	0:07:48	[redacted]氏	うん、あ、そっか。まとめるだけ	うん、あ、そっか。まとめるだけ	0	0	0	0	0	0	0	0	0	0	0	0
90	191	0:00:04	0:07:52	[redacted]氏	ですね、あはははは。〇〇だっ	ですね、あはははは。〇〇だっ	0	0	0	0	0	0	0	0	0	0	0	0
91	192	0:00:04	0:07:56	[redacted]氏	だから、どちらでもいいがまと	だから、どちらでもいいがまと	0	0	0	0	0	0	0	0	0	0	0	0
92	193	0:00:02	0:07:58	[redacted]氏	ああ、そうですね。	ああ、そうですね。	0	0	0	0	0	0	0	0	0	0	0	0
93	194	0:00:01	0:07:59	[redacted]氏	そうですね。	そうですね。	0	0	0	0	0	0	0	0	0	0	0	0
94	195	0:00:01	0:08:00	[redacted]氏	どちらでもいいがまとめる、どこ	どちらでもいいがまとめる、どこ	0	0	0	0	0	0	0	0	0	0	0	0
95	196	0:00:02	0:08:02	[redacted]氏	いや、いいよ、当事者意識なし	いや、いいよ、当事者意識なし	0	0	0	0	0	0	0	0	0	0	0	0
96	197	0:00:04	0:08:06	[redacted]氏	セールスマン？	セールスマン？	0	0	0	0	0	0	1	0	0	0	0	0
97	198	0:00:01	0:08:07	[redacted]氏	セールスマン。	セールスマン。	0	0	0	0	0	0	1	0	0	0	0	0
98	199	0:00:01	0:08:08	[redacted]氏	セールスマン、そうですね。	セールスマン、そうですね。	0	0	0	0	0	0	1	0	0	0	0	0
99	200	0:00:01	0:08:09	[redacted]氏	じゃないや、セールスマン。	じゃないや、セールスマン。	0	0	0	0	0	0	1	0	0	0	0	0
100	201	0:00:04	0:08:13	[redacted]氏	どちらでもいいって、でも建築	どちらでもいいって、でも建築	0	0	0	0	0	0	0	1	0	0	0	0
101	202	0:00:09	0:08:22	[redacted]氏	は、おも、	は、おも、	0	0	0	0	0	0	0	0	0	0	0	0
102	203	0:00:01	0:08:23	[redacted]氏	私、感情が。	私、感情が。	0	0	0	0	0	0	0	0	0	0	0	0
103	204	0:00:01	0:08:24	[redacted]氏	それはでも、思い込みであの、	それはでも、思い込みであの、	0	0	1	0	0	0	0	0	0	0	0	0
104	205	0:00:04	0:08:28	[redacted]氏	ガレージの主人。	ガレージ。	0	0	0	0	0	0	0	0	0	1	0	0
105	206	0:00:01	0:08:29	[redacted]氏	ガレージの主人、この2人はも	ガレージ、この2人はもう頭が、	0	0	0	0	0	0	0	0	0	1	0	0

			1	2	3	4	5	6	7	8	9	10	11	12
5分間文字数	行数	経過時間	陪審員長	銀行員	メッセンジャー	ブローカー	スラム	ベンキ	セールスマン	建築師	老人	ガレージ	時計屋	広告代理店
1,362	48	5	0	0	0	0	0	0	0	0	0	0	0	0
1,978	126	10	1	0	2	2	0	0	6	4	0	3	0	0
2,287	202	15	0	0	2	2	1	0	0	7	2	0	0	0
1,672	278	20	10	3	1	3	1	5	4	2	3	7	4	1
2,058	342	25	3	0	1	0	4	1	0	1	6	1	1	2
2,265	403	30	6	0	0	7	1	1	0	1	0	0	0	0
1,976	469	35	1	0	4	5	0	0	6	0	1	4	0	0
1,877	512	40	1	8	0	7	0	2	0	0	1	0	3	1
1,922	553	45	4	6	1	2	6	1	0	4	4	1	1	0
2,019	592	50	0	0	0	0	2	11	0	2	3	0	4	0
1,583	628	55	1	0	0	0	0	3	1	4	0	0	3	1
1,971	681	60	0	0	2	0	12	0	1	3	7	1	0	0
2,459	756	65	0	0	0	0	11	1	0	1	3	0	3	0
2,073	833	70	0	0	9	0	7	0	0	0	0	12	0	0
1,724	899	75	2	5	2	2	0	4	0	4	0	0	4	0
2,024	962	80	9	2	1	1	5	3	1	3	4	1	3	6
195	974	85	1	0	0	0	0	0	0	0	0	0	0	0
31,445		TTL	39	24	25	31	50	32	19	36	34	30	26	11

データの整合性について確認がとれたことから、このあとMiningAssistantを利用して本格的に各グループの討論模様を分析することとし、分析用に茶筅、Excel及びWordで前処理を実施後、各グループのCSVファイルを作成した。MiningAssistantはマーケティングツールとして開発されているため便利な機能がいくつも用意されている。これらを順次試していくこととした。

ところが、テストを開始すると思いもかけない仕様の制約で、使用したかった機能のいくつか利用できないことが判明した。

例えば、キーワードに関する頻度分析の結果をビジュアライズするため、正多角形のレーダーチャートが用意されていた。ところが、ソフトウェアが用意している多角形の頂点が表示する各キーワードの項目数の上限が10個までという制約があった。今回の陪審員をキーワードにすると項目は12個必要である。項目設定画面で12個の登録が可能であったので、なんの疑問も持たずに描画してみると何故かグラフィック上は10個までしか表示されないのである（図表２８参照）。これは明らかに設計と開発での齟齬によるソフトのバグである。

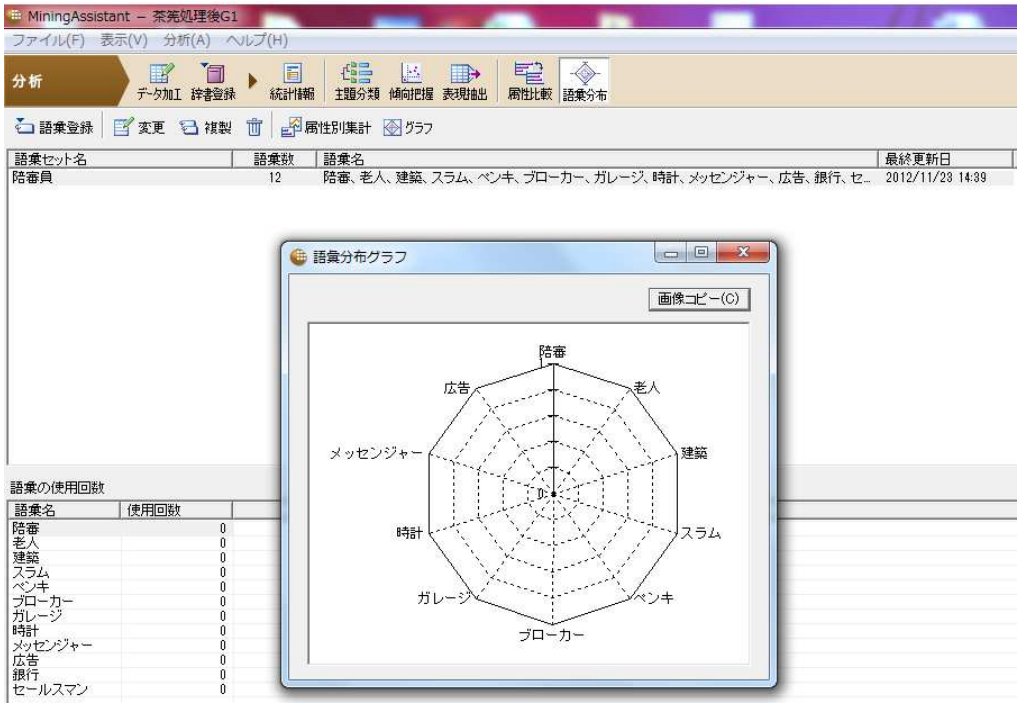
また、MiningAssistantではアンケートを整理するテンプレートの機能を持っている。所定の体裁でアンケートを作れば、eメールで自動受付し、分類、集計さらに美しい体裁で表示する機能を持っている。一見とても便利な仕様であるように思われるが、分析対象のデータのインプットは必ずeメールで収集しなければならない仕様となっていた。このため、ペーパーベースで回収したアンケートの集計に本ソフトを利用しようと思ってもCSV形式でデータベース化したデータを取り込むことができない。何故このような仕様としたのか首を傾げてしまう仕様であった。

一方、単に本ソフトに問題があるということだけではなく、ツールとインプットの親和性という面で本研究の分析対象のテキストが特殊であるという問題もある。

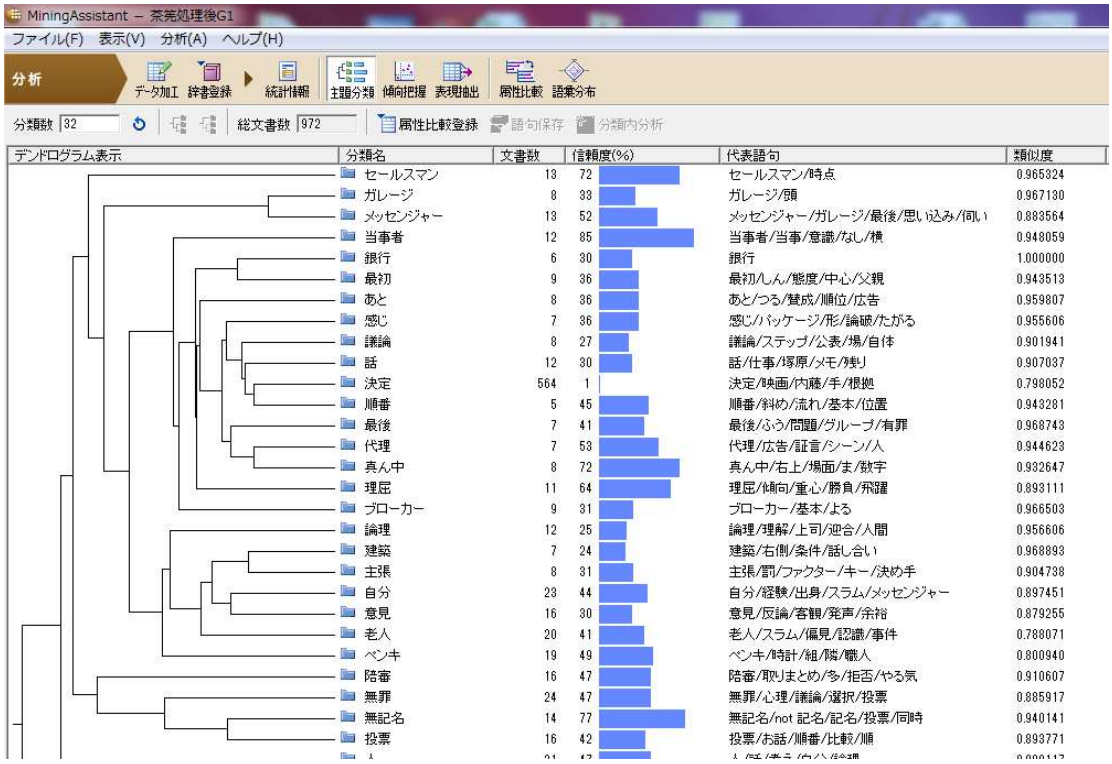
改めて言うまでもないが今回の分析対象のデータが形態素解析によってある程度ノイズとなる助詞類を排除しているものの不完全な日本語の文章である。本ソフトの基本機能である係り受け分析や、クラスター分析により類似度などを試みたところで機械処理結果は意味不明になってしまう（図表２９参照）。あくまでMiningAssistantが得意な分析対象は、メールやアンケートなどのように１つの文章として完結している必要があり、文法上ある程度整理されている文章でなければならないのである。

これらの結果、機械的に構文解析を済まそうと思わなければテキストマイニングの専用ソフトを用いる必要は必ずしもないと判断し、TRUSTIA/R.2のテキストマイニングモジュールMiningAssistantの利用は中止することとした。

図表２８．MiningAssistant のグラフ表示の画面ショット



図表２９．MiningAssistant のクラスター分析の画面ショット



4-2-5　TTM（TinyTextMiner)の利用

TRUSTIA/R.2 の利用を中断したのち、このあとのツールをどうするかを検討した。形態素解析ツールの茶筅は、新聞記事を対象として形態素解析を実施した場合の精度は 99%以上であり、話し言葉を対象とした場合でも若干精度は落ちるが十分実用的な精度（松村、三浦，2009）である。しかし、茶筅処理後は品詞分解したデータを、そのあとの工程で利用可能な状態にするため、更に何らかのツールを使用して加工する必要がある。筆者は、身近で使い慣れた Excel や Word を利用し、必要に応じてマクロを組みながら処理を進めたが、１０グループ分に対して同じ作業を順番に行う必要があり、とても効率的とは言えない状況であった。

その後、修士論文の中間報告の際に今までの顛末を報告すると、井上教授よりテキストマイニングの前処理に関する無償ツール TTM（TinyTextMiner)[xiii]を利用してはどうかとアドバイスを得た。TTM は、従来、計算機科学分野の研究者が中心となって開発や利用を行っていたテキストマイニング技術を、広く人文・社会科学分野の利用者にも普及させたいとのコンセプトで開発された無償ソフトウェアである（松村、三浦，2009)。

TTM を利用するメリットは、茶筅を利用して形態素解析を行う際、事前事後に行う手作業の全てを TTM が担ってくれることにある。つまり、頻度解析で集計のキーとなるキーワードの統一処理、助詞類の不要語削除、必要となる品詞の選択ができることに加え、出力フォーマットが 6 種類用意されており、出現頻度と出現件数別に語のタグ別集計、語×タグのクロス集計や語×語のクロス集計などを自動で処理してくれる。そして、TTM の手引き書も用意されているほか、これを用いて研究された論文も多数存在する。10 グループ分のサンプルを同一処理しなければならないため、筆者は迷わずこれを利用することとした。

TTM は基本的に集計機能を持つソフトウェアであって、TTM 単体では機能しない。このため、テキストマイニングを行う上で前処理に必要な２つの構成要素である形態素解析ソフトウェアの MeCab[xiv]と構文解析ソフトウェア CaboCha[xv]は別にインストールが必要となる。

[xiii] TTM のダウンロードは http://mtmr.jp/ttm/参照

[xiv] MeCab は京都大学情報学研究科と日本電信電話株式会社コミュニケーション科学基礎研究所の協働研究ユニットプロジェクトを通じて開発された形態素解析エンジンである。http://mecab.sourceforge.net

[xv] CaboCha は奈良先端科学技術大学院大学及び工藤 拓氏により開発された日本語係り受け解析エンジンである。http://chasen.org/~taku/software/cabocha/

なお、インストールするPC及びOSの環境によって正常動作しない場合があるので注意が必要である。実際、筆者はこのTTMの利用環境を構築する際に松村氏のガイダンスに沿ってインストールを試みたのであるが、試験データを流してみたところ頻度解析の結果が誤っていること、また、数回利用を繰り返すと結果がどんどん変化してしまうというバグを発見した。

基本的な話ではあるが、研究に利用するツールが正常に動作するかを確認する手続きは非常に重要である。筆者にとって他にツールの選択の余地がなかったこともあり、大阪大学の松村真宏准教授、並びに関西学院大学の三浦麻子教授に連絡をとることとし、TTMの利用継続を前提に不具合の原因を究明することとした。

原因を究明するためにはさまざまな情報が必要となる。テストに用いたPC、環境検証データ並びにTTMのエラーとなった処理結果等を報告した。松村准教授、三浦教授と数度のやり取りを経て、筆者はWindows7環境においてPCが64bitか32bitかによって適合するMeCabのリリースバージョンが異なるというバグがあり、その組み合わせによってはTTMの計算結果が誤った値を示すことを発見した。Windows7で64bitのPCが正常に動作するMeCabのリリースバージョンはmecab-0.994.exe[xvi]であった。

これらの結果、TTMで正しい計算結果を示す環境を構築することができたのであるが、後進の参考までに松村氏とのメール交信記録を別紙１３に追録する。

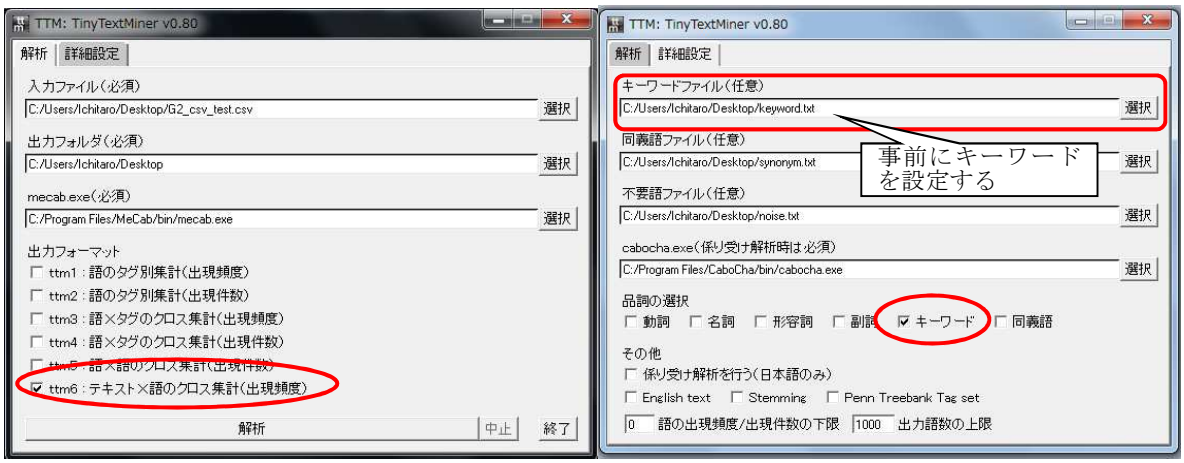
余談になるが、松村准教授と三浦教授には大変感謝している。一面識もない私の、それも深夜の面倒な長文メールと大量データのCSVファイルの処理を一晩で処理して返信をいただいた。松村准教授は井上哲浩教授を良くご存じでおられたことが幸いしたのだと思うが、深夜にも関わらずTTMの検証作業を即座に実施していただいた。研究者魂とはこういったことなのかと深い感動を覚えた。この迅速なご対応をいただけなければ研究の遂行は正直いって相当難しかったであろう。

[xvi] mecab-0.994.exeのダウンロードサイト
http://sourceforge.net/projects/mecab/files/mecab-win32/0.98/mecab-0.98.exe/download

4-2-6　キーワードの頻出度

グループ討議の模様を定量的に比較するため、TRUSTIA/R.2 の利用を中止する前に図表２６の語句トップ XX を表示する機能を利用し、頻出度上位のキーワードを調べた。しかし、１０グループの表示結果を比較すると陪審員名以外に共通項は見受けられなかったため、陪審員名をキーワードとして頻出度を調べることとした。TTM を起動し、データベースから１０グループ分各々約 1.5 時間分のテキストデータを順次処理していった。TTM の設定方法等は図表３０のとおりである。

図表３０．TTM の設定方法と出力結果



この結果として、キーワード頻出度をまとめたものが図表３１である。集計の結果、各グループ討論において各陪審員の頻出度は平均的とは言えない結果となった。また、最多と最少で２倍から３倍の開きがあった。

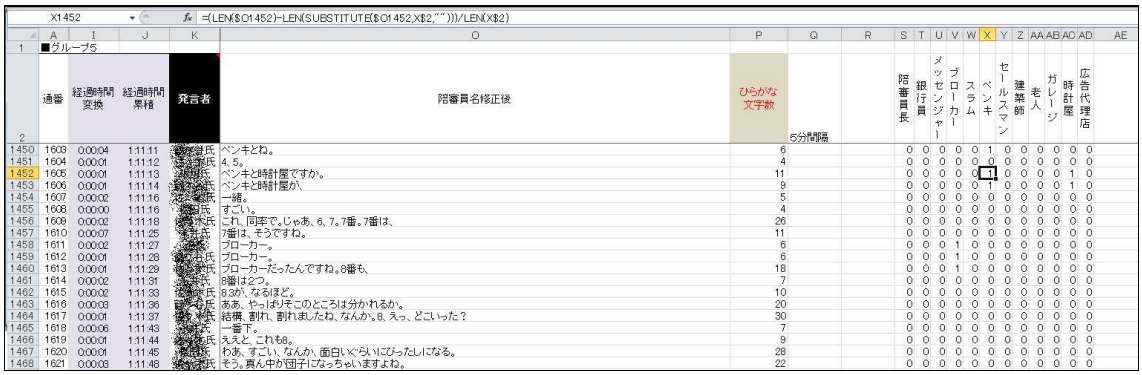
図表３１．陪審員名の出現頻度

次に、フロー情報、すなわち単位時間毎に区切った場合、キーワードがどのように発せられていたのかを調査した。データベースをもとに５分間隔での陪審員名頻出度を調べるものである。作り上げたデータベース[xvii]は図表３２のとおりである。

この方法により全グループの推移を集計し、層グラフにしたものが図表３３である。

唯一確認できる点は、いくつか例外はあるものの、開始直後に発生した一つ目の山が個人決定順位を、また会議終了１５分前くらいの時点で発生した山がグループ決定順位を、グループ全員で確認作業を行ったために築かれたものであることは識別できるが、図表１５のグループ改善率と比較しても特徴的な事象は確認できなかった。

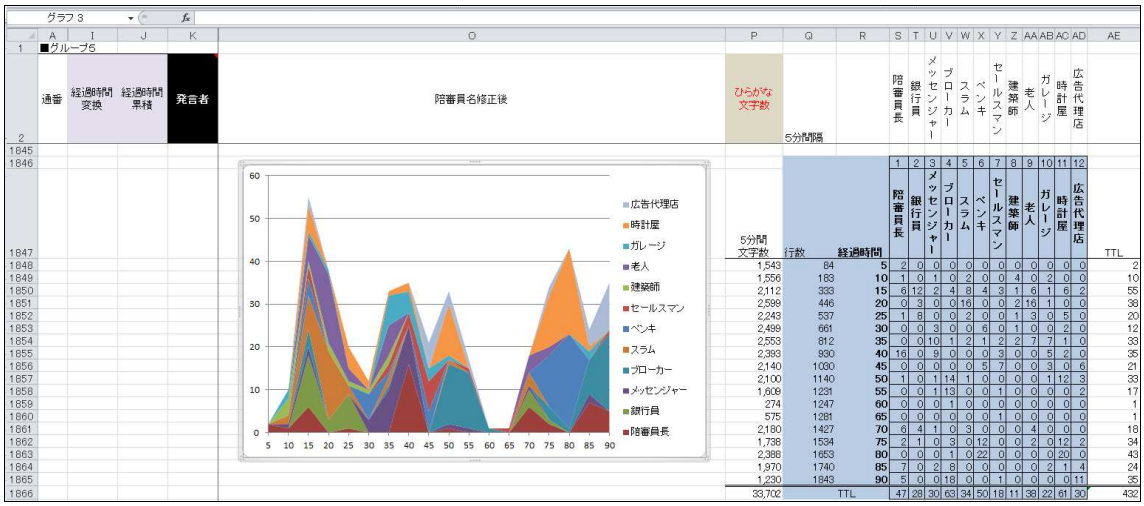
図表３２．５分間隔のキーワード頻出度集計用データベース



X1 452　=(LEN($O1 452)−LEN(SUBSTITUTE($O1 452,X$2,"")))/LEN(X$2)

■グループ5

通番	経過時間変換	経過時間累積	発言者	陪審員名修正後	ひらがな文字数	5分間隔	陪審員長	銀行員	メッセンジャー	ブローカー	スラム	ペンキ	セールスマン	建築師	老人	ガレージ	時計屋	広告代理店
1450	1603	0:00:04	1:11:11	ペンキとね。	6		0	0	0	0	0	1	0	0	0	0	0	0
1451	1604	0:00:01	1:11:12	4, 5。	4		0	0	0	0	0	0	0	0	0	0	0	0
1452	1605	0:00:01	1:11:13	ペンキと時計屋ですか。	11		0	0	0	0	0	1	0	0	0	0	1	0
1453	1606	0:00:01	1:11:14	ペンキと時計屋が、	9		0	0	0	0	0	1	0	0	0	0	1	0
1454	1607	0:00:02	1:11:16	一緒。	5		0	0	0	0	0	0	0	0	0	0	0	0
1455	1608	0:00:00	1:11:16	すごい。	4		0	0	0	0	0	0	0	0	0	0	0	0
1456	1609	0:00:02	1:11:18	これ、同率で。じゃあ、6、7。7番。7番は、	26		0	0	0	0	0	0	0	0	0	0	0	0
1457	1610	0:00:07	1:11:25	7番は、そうですね。	11		0	0	0	0	0	0	0	0	0	0	0	0
1458	1611	0:00:02	1:11:27	ブローカー。	6		0	0	0	1	0	0	0	0	0	0	0	0
1459	1612	0:00:01	1:11:28	ブローカー。	6		0	0	0	1	0	0	0	0	0	0	0	0
1460	1613	0:00:01	1:11:29	ブローカーだったんですね。8番も、	18		0	0	0	1	0	0	0	0	0	0	0	0
1461	1614	0:00:02	1:11:31	8番は2つ。	7		0	0	0	0	0	0	0	0	0	0	0	0
1462	1615	0:00:02	1:11:33	8.3が、なるほど。	10		0	0	0	0	0	0	0	0	0	0	0	0
1463	1616	0:00:03	1:11:36	ああ、やっぱりそこのところは分かれるか。	20		0	0	0	0	0	0	0	0	0	0	0	0
1464	1617	0:00:01	1:11:37	結構、割れ、割れましたね、なんか。8、えっ、どこいった？	30		0	0	0	0	0	0	0	0	0	0	0	0
1465	1618	0:00:06	1:11:43	一番下。	7		0	0	0	0	0	0	0	0	0	0	0	0
1466	1619	0:00:01	1:11:44	ええと、これも6。	9		0	0	0	0	0	0	0	0	0	0	0	0
1467	1620	0:00:01	1:11:45	わあ、すごい、なんか、面白いぐらいにひったしになる。	28		0	0	0	0	0	0	0	0	0	0	0	0
1468	1621	0:00:08	1:11:48	そう。真ん中が団子になっちゃいますよね。	22		0	0	0	0	0	0	0	0	0	0	0	0



グラフ 3

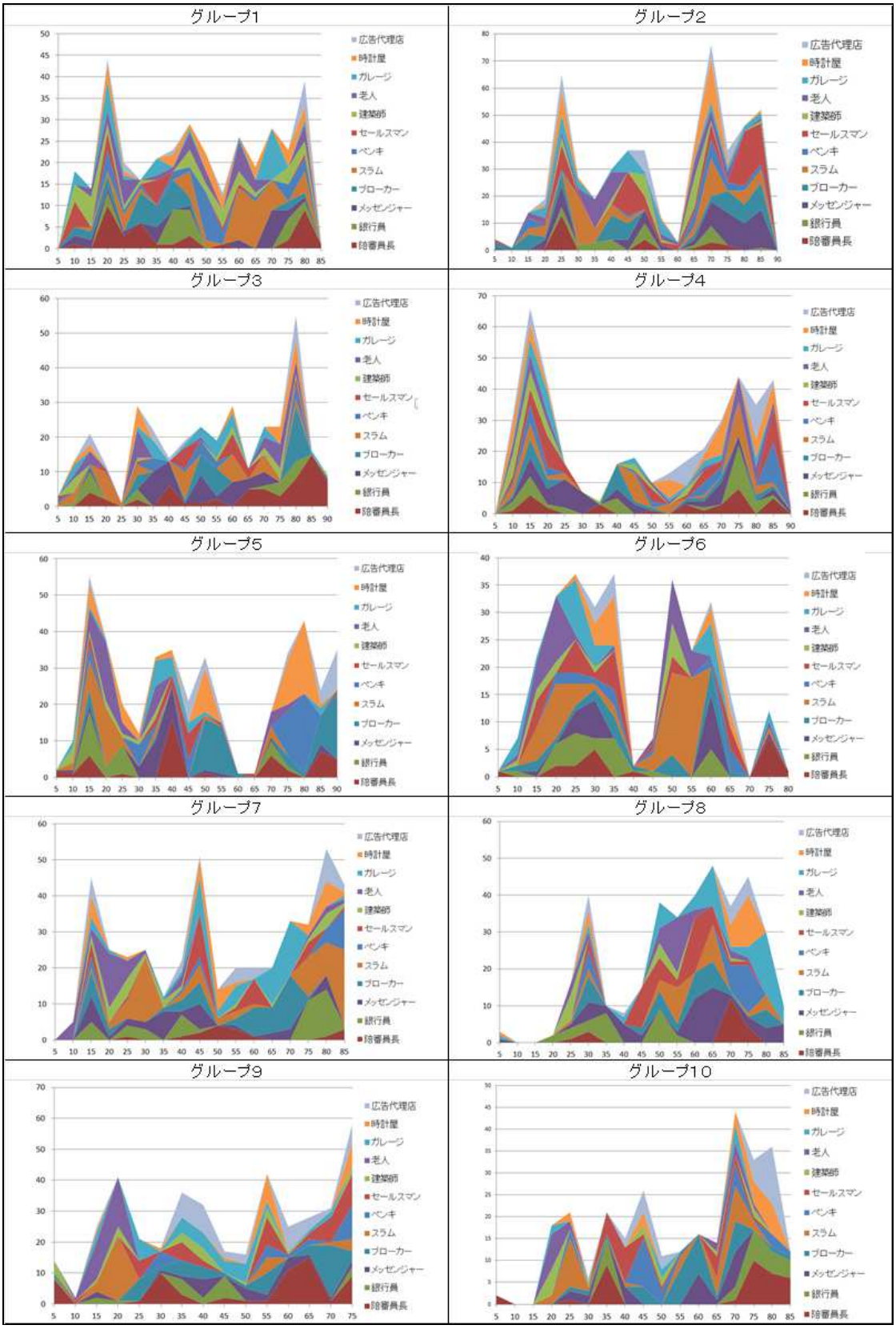
■グループ5

	5分間文字数	行数	経過時間	1 陪審員長	2 銀行員	3 メッセンジャー	4 ブローカー	5 スラム	6 ペンキ	7 セールスマン	8 建築師	9 老人	10 ガレージ	11 時計屋	12 広告代理店	TTL
1848	1,543	84	5	2	0	0	0	0	0	0	0	0	0	0	0	2
1849	1,556	183	10	1	0	1	0	2	0	0	4	0	2	0	0	10
1850	2,112	333	15	6	12	2	4	8	4	3	1	6	1	6	2	55
1851	2,599	446	20	0	3	0	0	16	0	0	2	16	1	0	0	38
1852	2,243	537	25	1	8	0	0	2	0	0	1	3	0	5	0	20
1853	2,499	661	30	0	0	3	0	0	6	0	1	0	0	2	0	12
1854	2,553	812	35	0	0	10	1	2	1	2	2	7	7	1	0	33
1855	2,393	930	40	16	0	9	0	0	0	3	0	0	5	2	0	35
1856	2,140	1090	45	0	0	0	0	0	5	7	0	0	3	0	6	21
1857	2,100	1140	50	1	0	1	14	1	0	0	0	0	1	12	3	33
1858	1,609	1231	55	0	0	1	13	0	0	1	0	0	0	0	2	17
1859	274	1247	60	0	0	0	1	0	0	0	0	0	0	0	0	1
1860	575	1281	65	0	0	0	0	0	0	1	0	0	0	0	0	1
1861	2,180	1427	70	6	4	1	0	3	0	0	0	4	0	0	0	18
1862	1,738	1534	75	2	1	0	3	0	12	0	0	2	0	12	2	34
1863	2,388	1653	80	0	0	0	1	0	22	0	0	0	0	20	0	43
1864	1,970	1740	85	7	0	2	8	0	0	0	0	0	2	1	4	24
1865	1,230	1843	90	5	0	0	18	0	0	1	0	0	0	0	11	35
1866	33,702	TTL		47	28	30	63	34	50	18	11	38	22	61	30	432

---

xvii　このスプレッドシートでは、数千件という膨大な行数であり、かつ、グループ毎にレコード数が異なることから、誤りなく範囲設定し頻出度を自動集計させるため、Excel 内でセルＱ列に「経過時間フラグ」を立てると自動で集計範囲が調整されるように Excel 関数を組み上げている。

図表３３.単位時間（５分間隔）のキーワード頻出推移



4-2-7　個人の発言量と Type-Token Ratio の算出

これまで、グループ毎の会議模様を外観的に俯瞰するために一つの定量評価手法としてキーワードの頻出度に着目し、会議全体での頻出度と一定間隔（５分）毎の頻出度推移をみてきた。しかしながら、グループの改善に結びつく何らかの特徴については発見できなかった。本節では、この模擬会議が「明示された特定の目的達成のための対話」である特徴に鑑み、各グループの構成メンバーに着目して定量分析を実施したのでその結果を述べることとしたい。

そもそも本模擬会議の目的は、８名～９名で構成されたグループにおいて与えられた１．５時間の間に一つのグループ決定を合意するというものである。グループを構成する一人一人は、何れも入学して初めて顔を合わせる者同士であり、年齢や社会人経験などの差こそあれ、基本的には一学生として対等の条件である。しかし、記録された映像を見ると、図表１８にまとめたとおりグループ毎に司会者を決めたり、平均値を出したり、また３つのグループに分類したりといった行動がみられ、結果として全てのグループで一つの意思決定が図られている。映像を見た限りでは、必ずしも司会者が場の仕切りを全て行っているようにはみえないが、何れかの者によるリーダーシップが集団意思決定に作用しているものと思われるふしがあった。これを定量的に炙り出すにはどうすれば良いか。

熟慮の末、次なる調査は「発言量が多い者がリーダーシップをとっているのではないか。また、集団の意思決定に何らかの影響を及ぼす特徴をもっているのではないか。」という仮説を検証することとした。

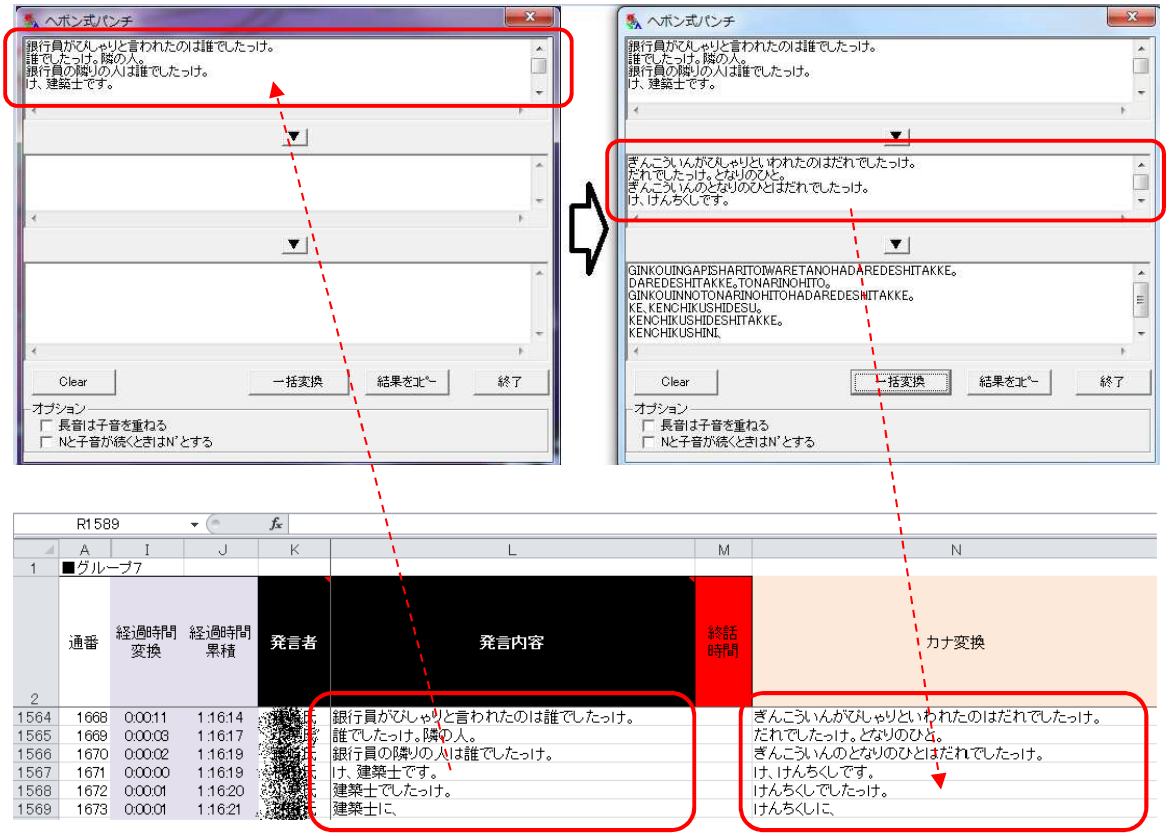
方法は、テキストデータから割り出せる量の基準値としては一番微細な表現単位である「かな」「英文字」「数字」で各メンバーの発言量を定量評価するのである。しかし、これには一つ問題があった。書き起こしたデータベースは「漢字交じり」である。これを全て「かな」に変換する必要がある。

筆者が調査したところ、「ヘボン式パンチ[xviii]」という漢字混じりの文章をヘボン式ローマ字に変換する無償のソフトウェアがあった。このソフトウェアは、2 段階の変換ステップがあり、最終的にローマ字へ変換する前処理で「漢字混じり文」を全て「かな」にできる機能がある。検証の結果、数千レコードを一括処理することは困難であったが、800 レコード程度は一括変換できたので十分実用に耐えるものであった。このソフトはインストール不要で、ダウンロードしてきたファイルを解凍するだけで利用可能である。具体的なやり方は次のとおりである。

[xviii] ヘボン式パンチ：Copyright(C)2006 丸干商店が著作権を有するソフトウェアである
http://hp.vector.co.jp/authors/VA031375/index.htm

データベースの「漢字かな混じり」発言内容を範囲選択し、Excel コマンドの Copy で PC のコンソールにそのテキストデータを取り込む。次に起動した「ヘボン式パンチ」の一番上の枠内ブランクに取り込んだテキストデータを Paste する。「一括変換」ボタンを押下すると、中段の枠内に「かな」変換後の文章が表示される。これを再度 Copy & Paste し、Excel データベースに戻すという単純な手順を繰り返す。ヘボン式パンチの具体的な使用イメージは図表３４のとおりである。

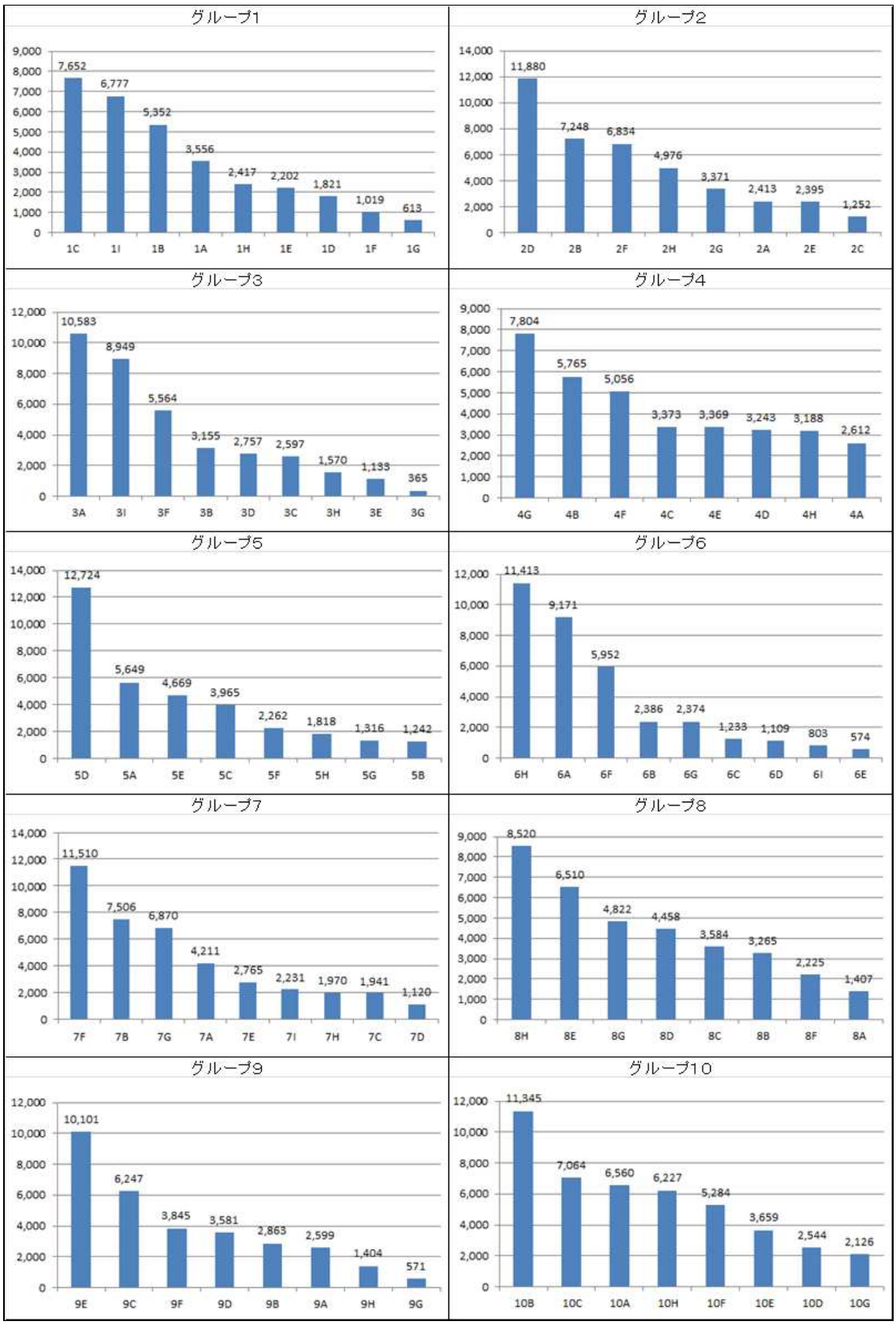
図表３４．ヘボン式パンチの使用イメージ図



全てのグループを集計し、一覧にしたものが図表３５である。

発言量の多寡をみると、個人別では最も発言文字数が多かったのは５グループの５D 氏で 12,724 文字。最も少なかったのは３グループの３G 氏で 365 文字と差は約 35 倍あった。また、１グループと３グループを除く８つのグループにおいて、発言量第１位の人と２位以下の人との差が非常に大きい傾向がある。実に７つのグループで上位１位の人の発言量がグループ全会話の 25%超を占めていた。

図表３５．個人別発言文字数



更に上位１位と２位の発言量を合計して全体の発言量と比較すると、最大値は６グループの 58.8%、最小値は４グループの 39.4%、平均は 48.3%、中央値は 47.4%であった。各グループの人数は８名または９名であることからすると、発言量上位１位と２位でグループの全会話の約半分、グループによっては６割近くを占めていたことが明らかとなった。観察の結果、発言量が多い人が司会者であるとは限らないのであるが、全体の議論を方向付ける上で発言量の多寡がリーダーシップの面で他の構成員に対して何らかの力学的な影響を与えている可能性は否定できないと思われるため、更に定量的な指標を調査していくこととした。

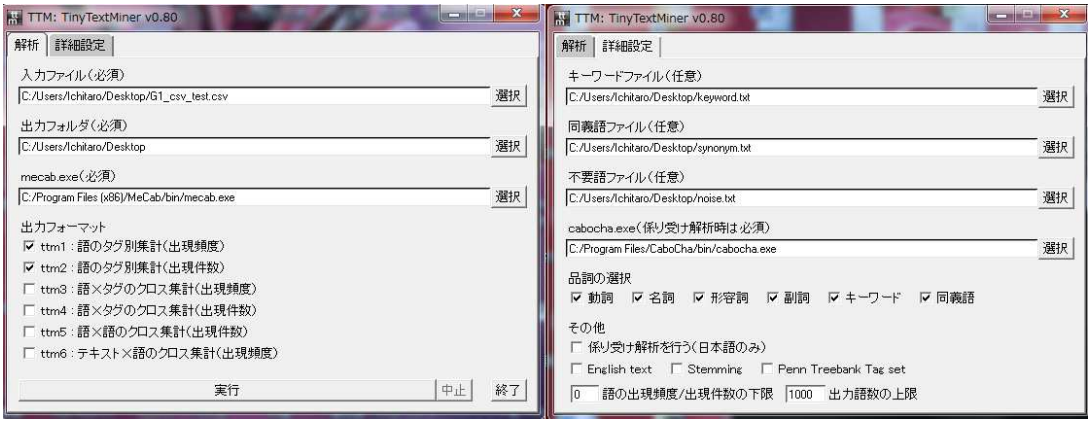
次に調査したのは、発言者の語彙の豊かさを示す指標であるタイプ・トークン比(Type-Token-Ratio、ＴＴＲ)である。この指標は使用された総単語を「延べ語数：N」としたとき、使用された単語数を「異なり語数：V」として比率を求めたものである。ＴＴＲは以下の式で表される。

$$\mathrm{TTR} = \frac{V}{N}$$

但し、Ｎ：延べ語数、Ｖ：異なり語数

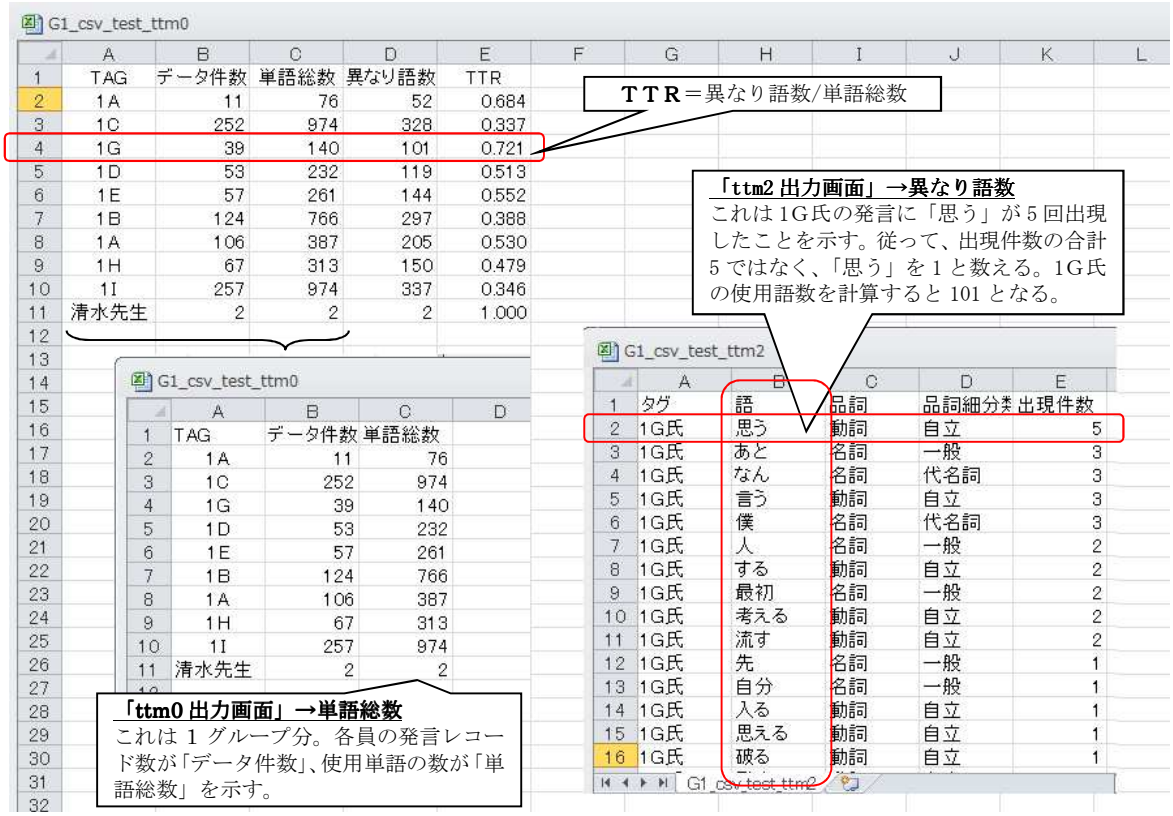
このＴＴＲを計算するには、TTM の出力ファイル ttm0（タグ、データ件数、単語総数）と ttm2（タグ、語、品詞、品詞細分類、出現件数）の出力機能を利用して、延べ語数Ｎと異なり語数Ｖを計算する。TTM の設定は図表３６のとおり。

図表３６．ＴＴＲ出力時の TTM 設定方法



TTM の出力ファイル ttm0 と ttm2 の画面イメージと、ＴＴＲを計算する方法は図表３７のとおりとなる。

図表３７．TTM による出力ファイル ttm0 と ttm2 とＴＴＲ計算方法



これらを集計した結果が図表３８である。

図表３８はＴＴＲの値をもとに降順ソートしているが、図表３５の個人別発言量（発言文字数ベース）と照らしてみると、発言量の上位１位の人が最もＴＴＲの低い結果を示した例が７つのグループで、また上位２位の人が最も低い結果を示した例が２つのグループでみられている。発言量の上位３位までをハイライトしてあるが、これらの結果をみる限り発言量とＴＴＲの値は概ね反比例する傾向があるものと思われる。

図表３８．タイプ・トークン比の算定結果

グループ1

TAG	データ件数	単語総数	異なり語数	TTR
1F	39	140	101	0.721
1G	11	76	52	0.684
1E	57	261	144	0.552
1A	106	387	205	0.530
1D	53	232	119	0.513
1H	67	313	150	0.479
1B	124	766	297	0.388
1I	257	974	337	0.346
1C	252	974	328	0.337

グループ2

TAG	データ件数	単語総数	異なり語数	TTR
2C	105	138	87	0.630
2E	192	271	143	0.528
2A	271	234	120	0.513
2G	268	387	191	0.494
2H	413	492	201	0.409
2F	416	887	304	0.343
2B	465	893	295	0.330
2D	625	1422	455	0.320

グループ3

TAG	データ件数	単語総数	異なり語数	TTR
3G	42	26	21	0.808
3B	231	347	191	0.550
3H	201	130	67	0.515
3C	90	292	147	0.503
3D	189	348	175	0.503
3E	47	130	58	0.446
3F	293	674	290	0.430
3I	505	1178	373	0.317
3A	515	1214	359	0.296

グループ4

TAG	データ件数	単語総数	異なり語数	TTR
4A	95	316	161	0.509
4D	226	439	200	0.456
4H	248	337	146	0.433
4C	208	426	177	0.415
4E	86	399	150	0.376
4G	229	995	357	0.359
4F	230	637	226	0.355
4B	305	683	223	0.327

グループ5

TAG	データ件数	単語総数	異なり語数	TTR
5G	81	162	101	0.623
5B	78	137	78	0.569
5F	136	264	149	0.564
5H	81	219	111	0.507
5C	241	498	200	0.402
5A	349	693	227	0.328
5E	289	577	164	0.284
5D	580	1513	348	0.230

グループ6

TAG	データ件数	単語総数	異なり語数	TTR
6E	22	64	45	0.703
6I	44	106	71	0.670
6D	78	123	77	0.626
6C	66	169	99	0.586
6B	84	300	142	0.473
6G	104	357	135	0.378
6A	446	1225	443	0.362
6F	275	750	257	0.343
6H	341	1449	376	0.259

グループ7

TAG	データ件数	単語総数	異なり語数	TTR
7H	75	223	127	0.570
7E	159	326	182	0.558
7D	36	141	72	0.511
7I	95	274	129	0.471
7A	155	519	219	0.422
7B	442	758	284	0.375
7G	283	835	302	0.362
7C	85	263	83	0.316
7F	374	1522	441	0.290

グループ8

TAG	データ件数	単語総数	異なり語数	TTR
8A	67	209	116	0.555
8B	126	468	236	0.504
8C	146	486	231	0.475
8F	67	282	133	0.472
8G	200	593	262	0.442
8D	191	532	221	0.415
8H	278	1056	323	0.306
8E	292	794	231	0.291

グループ9

TAG	データ件数	単語総数	異なり語数	TTR
9G	31	66	48	0.727
9H	60	192	109	0.568
9B	187	336	172	0.512
9F	157	479	220	0.459
9A	127	325	144	0.443
9C	244	779	269	0.345
9E	337	1263	395	0.313
9D	143	440	106	0.241

グループ10

TAG	データ件数	単語総数	異なり語数	TTR
10G	192	244	115	0.471
10E	205	434	199	0.459
10D	145	312	137	0.439
10F	320	677	288	0.425
10C	439	829	298	0.359
10A	574	752	266	0.354
10H	447	794	262	0.330
10B	667	1485	328	0.221

発言量1位の人
発言量2位の人
発言量3位の人

4-3　　クラスター分析

4-3-1　Jaccard 係数の導出

次なる仮説として、グループ内において合意形成の過程で「人と人との見えない結びつき（紐帯）」ができているのではないかと考えた。

一般に、クラスター分析は似ているものを同じグループとして分類する統計的分析手法である。クラスター分析にはいくつかの方法があるが、本研究においては階層的クラスター分析の手法を採用することとした。この階層的クラスター分析を行うためには、階層構造が想定される二軸を設定するとともに分析に用いるための指標が必要となる。前述のとおり、人と人との紐帯を明らかにすることが目的であるため、二軸としてはグループのメンバー８名又は９名をとり、分析の指標としては共起確率を示す Jaccard 係数を用いることとした。

Jaccard 係数とは以下に示す計算式で算出される。

$$\text{Jaccard 係数}\quad J(X;Y) = \frac{|X \cap Y|}{|X| + |Y|}$$

但し、Ｘ：参加者ｘの発話件数、Ｙ：参加者ｙの発話件数、Ｘ∩Ｙ：参加者ｘ→参加者ｙのコメントチェーン発生件数

Jaccard 係数は、テキストマイニングで出現する語と語の共起の強さを測る指標として用いられることが多い。今回は、グループ討議の参加者全ての発言を分母とし、分子には参加者ｘ、参加者ｙが連続してコメントチェーンを作った件数として、特定人物間の会話の共起発生率を計算した。算定方法は図表３９のとおりである。

図表３９．Jaccard 係数の算定方法（グループ１の例を抜粋）

	発言	TAG	組み合わせ
4	揃ってますかね。	1A	
5	はい、全員います。	1I	1A→1I
6	はい。どうしましょう。じゃあとりあえず、皆さ	1A	1I→1A
7	じゃあこっから、私からいいですか。	1A	1A→1A
8	ええと、言います。	1A	1A→1A
9	ええと、[illegible]さん。	1B	1A→1B
10	はい、ええ[illegible]上から読み上げます。7	1A	1B→1A
11	あ、もう1回。ごめんなさい。	1B	1A→1B
12	7、11、12、10、5、4、8、1、2、6、3、9、	1A	1B→1A
13	8、1、2。	1C	1A→1C
14	8、1、2、6、3、9。	1A	1C→1A
15	8、1、2、6、3、9。	1C	1A→1C
16	そう来ましたか。	1E	1C→1E
17	なんか全然違いそうな気がするな。	1A	1E→1A
18	で、1回、言ってっちゃうんですね。	1C	1A→1C
19	はい。	1A	1C→1A
20	じゃあ、[illegible]言います。ええ、7、5、12、	1C	1A→1C
21	10、2、6。	1C	1C→1C
22	ええと、10、2、6？	1B	1C→1B
23	はい。9、1、4、11、3、8。	1C	1B→1C
24	では私ですね。ええと、8、あ、私じゃない	1B	1C→1B
25	どうぞ。	1F	1B→1F

■個人別発言

TAG	データ件数	単語総数
1G	11	7
1C	252	84
1F	39	8
1D	53	29
1E	57	27
1B	124	65
1A	106	30
1H	67	22
1I	257	73
TTL	966	345

■組み合わせ　　着

発	1G	1C	1F	1D	1E	1B	1A	1H	1I
1G		1G→1C	1G→1F	1G→1D	1G→1E	1G→1B	1G→1A	1G→1H	1G→1I
1C	1C→1G		1C→1F	1C→1D	1C→1E	1C→1B	1C→1A	1C→1H	1C→1I
1F	1F→1G	1F→1C		1F→1D	1F→1E	1F→1B	1F→1A	1F→1H	1F→1I
1D	1D→1G	1D→1C	1D→1F		1D→1E	1D→1B	1D→1A	1D→1H	1D→1I
1E	1E→1G	1E→1C	1E→1F	1E→1D		1E→1B	1E→1A	1E→1H	1E→1I
1B	1B→1G	1B→1C	1B→1F	1B→1D	1B→1E		1B→1A	1B→1H	1B→1I
1A	1A→1G	1A→1C	1A→1F	1A→1D	1A→1E	1A→1B		1A→1H	1A→1I
1H	1H→1G	1H→1C	1H→1F	1H→1D	1H→1E	1H→1B	1H→1A		1H→1I
1I	1I→1G	1I→1C	1I→1F	1I→1D	1I→1E	1I→1B	1I→1A	1I→1H	

これを集計したものが図表４０のグループ１の Jaccard 係数集計結果である。

図表４０を見ると明らかであるが、発信と着信で組み合わせが２通りできることから、Jaccard 係数は２通りの組み合わせが存在する。よって結果は発信着信の入れ替えで結果が異なるため非対称行列となっている。

図表４０．グループ１の参加者の Jaccard 係数

■｜X｜+｜着　　・・・上記「組み合わせ」の各個人データ件数の合計を算定

発	1G	1C	1F	1D	1E	1B	1A	1H	1I
1G		263	50	64	68	135	117	78	268
1C	263		291	305	309	376	358	319	509
1F	50	291		92	96	163	145	106	296
1D	64	305	92		110	177	159	120	310
1E	68	309	96	110		181	163	124	314
1B	135	376	163	177	181		230	191	381
1A	117	358	145	159	163	230		173	363
1H	78	319	106	120	124	191	173		324
1I	268	509	296	310	314	381	363	324	

■｜X∩Y｜着　　・・・左記の列「○」において、上記「組み合わせ」の発生頻度を算定

発	1G	1C	1F	1D	1E	1B	1A	1H	1I
1G		2	3	1	1	1	1	0	2
1C	3		6	9	12	42	40	23	112
1F	1	7		2	3	7	3	2	14
1D	1	13	1		4	7	9	4	14
1E	2	8	1	3		7	8	5	23
1B	1	38	9	6	6		14	10	37
1A	0	39	2	4	4	14		5	35
1H	1	28	6	4	2	6	4		13
1I	2	111	11	23	25	38	24	15	

発言者「1C」→次の発言者「1I」の組み合わせが最も強い共起確率

■Jaccard J(X;Y)＝｜X∩Y｜／(｜X｜＋｜Y｜)

	1G	1C	1F	1D	1E	1B	1A	1H	1I
1G		0.008	0.060	0.016	0.015	0.007	0.009	0.000	0.007
1C	0.011		0.021	0.030	0.039	0.112	0.112	0.072	0.220
1F	0.020	0.024		0.022	0.031	0.043	0.021	0.019	0.047
1D	0.016	0.043	0.011		0.036	0.040	0.057	0.033	0.045
1E	0.029	0.026	0.010	0.027		0.039	0.049	0.040	0.073
1B	0.007	0.101	0.055	0.034	0.033		0.061	0.052	0.097
1A	0.000	0.109	0.014	0.025	0.025	0.061		0.029	0.096
1H	0.013	0.088	0.057	0.033	0.016	0.031	0.023		0.040
1I	0.007	0.218	0.097	0.074	0.080	0.100	0.066	0.046	

発言者「1I」→次の発言者「1C」の組み合わせが二番目に強い共起確率

全グループを集計したものが、図表４１である。これを図表３５の個人別発言量と比較すると、グループ１０を除く９グループで発言量第一位と第二位の組み合わせが

最も強い共起件数となっており、グループ１０では第一位と第三位の組み合わせが最も強い共起件数となっていた。

図表４１．Jaccard 係数一覧

グループ1

	1G	1C	1F	1D	1E	1B	1A	1H	1I
1G		0.008	0.060	0.016	0.015	0.007	0.009	0.000	0.007
1C	0.011		0.021	0.030	0.039	0.112	0.112	0.072	0.220
1F	0.020	0.024		0.022	0.031	0.043	0.021	0.019	0.047
1D	0.016	0.043	0.011		0.036	0.040	0.057	0.033	0.045
1E	0.029	0.026	0.010	0.027		0.039	0.049	0.040	0.073
1B	0.007	0.101	0.055	0.034	0.033		0.061	0.052	0.097
1A	0.000	0.109	0.014	0.025	0.025	0.061		0.029	0.096
1H	0.013	0.088	0.057	0.033	0.016	0.031	0.023		0.040
1I	0.007	0.218	0.037	0.074	0.080	0.100	0.066	0.046	

グループ2

	2G	2D	2C	2F	2E	2B	2H	2A
2G		0.082	0.013	0.069	0.041	0.063	0.075	0.045
2D	0.076		0.044	0.105	0.055	0.168	0.091	0.079
2C	0.021	0.047		0.019	0.010	0.058	0.017	0.013
2F	0.064	0.100	0.025		0.033	0.074	0.156	0.048
2E	0.059	0.047	0.024	0.039		0.040	0.040	0.095
2B	0.076	0.161	0.035	0.092	0.046		0.059	0.057
2H	0.054	0.114	0.025	0.130	0.051	0.059		0.070
2A	0.045	0.074	0.035	0.041	0.086	0.063	0.070	

グループ3

	3I	3C	3A	3D	3G	3B	3E	3F	3H
3I		0.040	0.160	0.095	0.011	0.091	0.014	0.119	0.075
3C	0.049		0.040	0.018	0.000	0.022	0.015	0.010	0.041
3A	0.151	0.043		0.088	0.023	0.076	0.052	0.100	0.091
3D	0.092	0.029	0.081		0.009	0.038	0.004	0.037	0.038
3G	0.020	0.000	0.025	0.013		0.015	0.000	0.009	0.029
3B	0.092	0.022	0.076	0.038	0.018		0.007	0.092	0.053
3E	0.014	0.015	0.050	0.000	0.000	0.014		0.006	0.004
3F	0.122	0.013	0.121	0.025	0.012	0.088	0.000		0.038
3H	0.072	0.027	0.066	0.049	0.045	0.058	0.008	0.067	

グループ4

	4F	4B	4A	4G	4D	4C	4H	4E
4F		0.082	0.049	0.074	0.103	0.087	0.073	0.032
4B	0.084		0.063	0.114	0.096	0.086	0.112	0.036
4A	0.055	0.065		0.031	0.037	0.023	0.044	0.028
4G	0.092	0.124	0.034		0.033	0.057	0.101	0.048
4D	0.090	0.104	0.034	0.048		0.104	0.078	0.045
4C	0.071	0.092	0.026	0.055	0.118		0.081	0.024
4H	0.073	0.096	0.052	0.107	0.074	0.077		0.057
4E	0.035	0.033	0.028	0.057	0.042	0.041	0.039	

グループ5

	5F	5D	5H	5A	5G	5B	5C	5E
5F		0.087	0.028	0.049	0.009	0.014	0.053	0.038
5D	0.096		0.071	0.152	0.079	0.062	0.112	0.148
5H	0.018	0.062		0.042	0.000	0.006	0.028	0.019
5A	0.039	0.161	0.037		0.016	0.049	0.098	0.121
5G	0.000	0.080	0.006	0.016		0.006	0.025	0.030
5B	0.009	0.073	0.006	0.037	0.006		0.016	0.011
5C	0.061	0.127	0.016	0.086	0.019	0.016		0.083
5E	0.035	0.129	0.011	0.143	0.035	0.014	0.089	

グループ6

	6B	6D	6A	6E	6C	6F	6I	6H	6G
6B		0.019	0.036	0.000	0.027	0.072	0.008	0.038	0.011
6D	0.006		0.048	0.020	0.028	0.040	0.115	0.024	0.000
6A	0.057	0.046		0.009	0.035	0.146	0.035	0.203	0.069
6E	0.000	0.010	0.021		0.000	0.010	0.000	0.014	0.008
6C	0.020	0.014	0.041	0.000		0.059	0.027	0.022	0.024
6F	0.050	0.054	0.129	0.020	0.067		0.003	0.097	0.061
6I	0.008	0.115	0.033	0.000	0.027	0.003		0.016	0.000
6H	0.026	0.019	0.221	0.019	0.022	0.093	0.010		0.056
6G	0.037	0.011	0.056	0.000	0.012	0.055	0.000	0.063	

グループ7

	7G	7H	7B	7I	7E	7A	7F	7D	7C
7G		0.011	0.117	0.053	0.093	0.078	0.125	0.013	0.024
7H	0.017		0.093	0.018	0.004	0.004	0.011	0.009	0.038
7B	0.097	0.091		0.065	0.063	0.092	0.179	0.040	0.044
7I	0.063	0.012	0.063		0.016	0.024	0.038	0.000	0.017
7E	0.081	0.026	0.062	0.035		0.029	0.086	0.015	0.045
7A	0.080	0.004	0.094	0.012	0.041		0.070	0.010	0.017
7F	0.139	0.018	0.162	0.043	0.088	0.079		0.017	0.050
7D	0.009	0.000	0.036	0.000	0.005	0.010	0.022		0.017
7C	0.033	0.044	0.042	0.011	0.045	0.017	0.044	0.000	

グループ8

	8F	8D	8B	8H	8G	8C	8A	8E
8F		0.035	0.016	0.052	0.030	0.038	0.052	0.025
8D	0.043		0.091	0.083	0.092	0.030	0.023	0.110
8B	0.026	0.076		0.069	0.080	0.033	0.026	0.062
8H	0.043	0.081	0.047		0.096	0.083	0.043	0.158
8G	0.019	0.095	0.064	0.115		0.061	0.034	0.077
8C	0.042	0.050	0.044	0.045	0.075		0.042	0.098
8A	0.030	0.031	0.031	0.049	0.022	0.038		0.039
8E	0.042	0.108	0.077	0.147	0.071	0.087	0.033	

グループ9

	9F	9A	9G	9C	9H	9B	9E	9D
9F		0.046	0.011	0.097	0.023	0.049	0.109	0.057
9A	0.035		0.025	0.105	0.016	0.070	0.075	0.033
9G	0.000	0.032		0.040	0.011	0.005	0.022	0.017
9C	0.082	0.067	0.029		0.036	0.084	0.164	0.062
9H	0.041	0.021	0.022	0.026		0.028	0.045	0.034
9B	0.052	0.067	0.018	0.079	0.032		0.105	0.106
9E	0.119	0.097	0.022	0.126	0.038	0.120		0.090
9D	0.057	0.041	0.006	0.072	0.059	0.091	0.083	

グループ10

	10A	10G	10D	10C	10E	10H	10F	10B
10A		0.057	0.064	0.079	0.068	0.071	0.055	0.138
10G	0.051		0.021	0.041	0.025	0.025	0.021	0.043
10D	0.074	0.015		0.022	0.014	0.015	0.019	0.049
10C	0.092	0.046	0.021		0.028	0.060	0.067	0.099
10E	0.059	0.020	0.011	0.037		0.040	0.048	0.057
10H	0.070	0.017	0.019	0.074	0.032		0.073	0.094
10F	0.054	0.023	0.024	0.057	0.038	0.078		0.066
10B	0.136	0.044	0.052	0.098	0.064	0.093	0.061	

4-3-2　個人決定とグループ決定の誤差比較

グループ予測と発言量上位の人の個人予測を再度検証してみる。図表４２及び図表４３がその取りまとめ結果である。

会議の合意決定されたグループ予測と、図表３５に示されたグループ内の発言量が上位１位又は２位の人の個人予測とを比較したとき、その誤差がグループ内の誰よりも小さいケースが６つのグループ[xix]すなわち６０％であった。言い方をかえると１０のグループのうち６つのグループでの合意決定は、そのグループの上位１位又は２位の発言量の人の個人決定に酷似しているということである。

更に、上記（1）の残りの４つのグループ予測と当該グループの発言量上位１位又は２位の人の個人予測とを比較すると、誤差の大小ではなく、絶対値の完全一致の個数、若しくは±１以内の値が最も多いケースが２つのグループ[xx]で認められた。つまり、発言量上位１位又は２位の人の予測がそのまま若しくは±１の範囲でグループ決定に採用されているケースが残りの２０％でみられた。

残りの２グループはこれらのケースに該当していなかった。

これらの結果は、上位１位と２位の人が全体決定に深く関与していることを示しているものと思われる。

[xix] このカテゴリに該当した６つのグループは図表４１の「仮説適否」欄に「◎」が付与されている。具体的にはグループ１、グループ２、グループ５、グループ６、グループ８及びグループ９である。

[xx] このカテゴリに該当した２つのグループは図表４１の「仮説適否」欄に「○」が付与されている。具体的にはグループ４及びグループ７である。

図表４２．グループ決定と個人決定の分析（1/2）

TAG	個人決定の誤差													グループ決定と個人決定の誤差の絶対値															コメント（誤差最小の人）
	1	2	3	4	5	6	7	8	9	10	11	12	合計	1	2	3	4	5	6	7	8	9	10	11	12	合計	誤差最小者	仮説適否	
1A	1	6	0	1	2	2	1	0	0	4	1	0	18	2	5	0	3	3	1	2	0	1	5	1	1	24		◎	1C(発言量1位)と1I(同2位)が最小
1B	0	0	1	5	0	2	3	0	0	2	1	2	16	1	1	1	1	1	1	0	0	1	1	1	1	10			
1C	1	0	0	1	1	0	2	0	2	1	1	1	10	2	1	0	3	0	1	1	0	1	0	1	0	10	○		
1D	2	0	2	3	1	5	2	0	5	2	0	6	28	3	1	2	1	0	6	1	0	4	1	0	5	24			
1E	2	3	5	0	1	3	5	0	1	0	1	1	22	3	4	5	4	2	4	2	0	0	1	1	0	26			
1F	1	1	0	2	0	2	3	0	0	1	1	3	14	2	2	0	2	1	3	0	0	1	0	1	2	14			
1G	1	1	2	1	0	1	1	0	0	1	0	2	10	0	0	2	5	1	0	2	0	1	0	0	1	12			
1H	1	2	1	1	1	3	3	0	3	2	0	3	20	2	3	1	5	0	4	0	0	2	3	0	2	22			
1I	4	1	1	4	1	1	2	0	1	0	2	1	18	3	2	1	0	0	0	1	0	0	1	2	0	10	○		
2A	1	0	0	1	0	0	0	0	0	1	0	1	4	2	2	2	1	1	0	2	0	0	1	1	0	12	n/a	◎	2A(既知)を除くと、2B(発言量2位)が最小で次が2D(同1位)
2B	4	3	2	2	2	0	4	0	1	2	4	2	26	3	1	0	2	1	0	2	0	1	0	3	1	14	○		
2C	1	1	5	8	1	2	2	0	10	1	4	1	36	2	3	3	8	2	2	4	0	10	1	3	2	40			
2D	0	3	3	0	4	3	3	0	2	2	2	4	26	1	1	1	0	3	3	1	0	2	0	1	3	16			
2E	2	1	5	2	1	3	1	0	2	2	1	2	22	3	3	3	2	2	3	1	0	2	0	0	3	22			
2F	1	1	9	6	9	4	0	0	0	1	2	1	34	2	1	7	6	8	4	2	0	0	1	1	0	32			
2G	3	1	5	1	1	3	5	0	1	1	2	1	24	2	1	3	1	2	3	3	0	1	1	1	0	18			
2H	3	1	2	1	1	1	3	0	1	1	3	1	18	4	3	0	1	2	1	5	0	1	3	2	0	22			
3A	2	2	0	7	1	4	2	0	3	1	4	2	28	6	2	2	3	0	2	3	0	2	10	4	2	36		×	3D(発言量5位)が最小、次が3E(同8位)
3B	3	3	1	7	6	1	3	0	1	2	2	1	30	7	3	1	3	7	1	4	0	0	11	2	1	40			
3C	3	6	8	8	1	1	1	0	7	2	6	3	46	7	6	6	4	0	1	2	0	6	11	6	3	52			
3D	4	0	3	4	1	0	3	0	2	1	1	1	20	0	0	1	0	0	2	4	0	1	10	1	1	20	○		
3E	2	2	0	3	0	2	1	0	0	1	1	0	12	2	2	2	1	1	4	0	0	1	10	1	0	24			
3F	1	1	3	1	1	5	1	0	3	0	1	1	18	5	1	1	5	0	3	0	0	2	9	1	1	28			
3G	2	7	5	8	1	2	2	0	6	4	5	2	44	2	7	3	4	0	4	1	0	5	5	5	2	38			
3H	0	1	10	8	6	1	5	0	4	1	6	4	46	4	1	8	4	7	1	6	0	3	10	6	4	54			
3I	3	2	6	1	1	1	2	0	0	2	3	1	22	1	2	4	3	2	3	3	0	1	11	3	1	34			
4A	0	1	0	5	1	1	4	0	3	0	5	6	26	0	2	5	5	1	3	2	0	1	2	3	2	26		○	4C(発言量4位)が最小で次が4B(同2位) ※4Bのグループ決定完全一致数5個は最多
4B	4	3	6	0	4	6	4	0	1	2	2	4	36	4	0	1	0	4	2	6	0	1	0	0	0	18			
4C	1	1	3	1	1	6	1	0	1	1	2	4	22	1	2	2	1	1	2	1	0	1	1	0	0	12	○		
4D	2	2	10	1	1	1	2	0	6	1	2	2	30	2	1	5	1	1	5	0	0	4	1	0	2	22			
4E	3	2	7	4	3	3	3	0	6	2	2	5	40	3	1	2	4	3	1	1	0	4	0	4	1	24			
4F	6	2	4	0	2	4	1	0	5	2	5	5	36	6	1	1	0	2	0	3	0	3	0	3	1	20			
4G	2	0	3	1	0	1	4	0	0	2	0	1	14	2	3	2	1	0	3	2	0	2	0	2	3	20			
4H	0	0	9	2	1	0	4	0	2	2	3	1	24	0	3	4	2	1	4	2	0	0	0	1	5	22			
5A	3	2	0	5	1	3	0	0	2	1	6	1	24	3	1	1	1	0	4	3	0	1	1	4	1	20		◎	5D(発言量1位)が最小、次が5G(同7位)
5B	2	0	0	4	0	2	1	0	0	1	2	0	12	2	1	1	0	1	1	2	0	1	1	0	0	10			
5C	2	1	3	3	1	4	4	0	1	2	1	2	24	2	0	2	1	0	5	1	0	0	0	1	2	14			
5D	2	0	1	3	1	1	3	0	1	2	0	0	14	2	1	0	1	0	2	0	0	0	0	2	0	8	○		
5E	4	1	2	8	1	1	1	0	3	1	2	0	24	4	2	1	4	2	2	2	0	2	1	4	0	24			
5F	3	1	4	1	1	0	2	0	1	2	3	4	22	3	0	3	3	0	1	1	0	0	0	1	4	16			
5G	2	1	1	6	1	1	3	0	1	2	5	0	23	2	0	0	2	0	2	0	0	0	0	3	0	9			
5H	3	3	3	3	1	3	5	0	5	1	2	1	30	3	2	2	1	2	2	2	0	4	1	0	1	20			

発言量1位の人
発言量2位の人
映画を以前に観たことがあり、かつ、結果の記憶が有る人

図表４３．グループ決定と個人決定の分析（2/2）

TAG	個人決定の誤差													グループ決定と個人決定の誤差の絶対値															コメント（誤差最小の人）
	1	2	3	4	5	6	7	8	9	10	11	12	合計	1	2	3	4	5	6	7	8	9	10	11	12	合計	誤差最小者	仮説適否	
6A	3	2	1	7	1	3	3	0	4	2	4	2	32	7	1	1	5	0	2	2	0	3	1	3	1	26			
6B	2	0	0	2	1	2	3	0	2	1	1	2	16	6	1	2	0	0	1	2	0	1	0	2	1	16			
6C	3	3	0	7	3	1	3	0	1	1	5	1	28	7	2	2	5	4	0	2	0	0	0	4	2	28			
6D	0	2	0	1	3	2	2	0	0	1	1	2	14	4	1	2	1	4	3	1	0	1	0	0	1	18			
6E	2	4	1	1	1	4	0	0	1	2	1	5	22	6	5	1	3	0	3	1	0	0	3	0	2	24		◎	6H(発言量1位)と6F(同3位)が最小で次が6B(同4位)
6F	1	1	1	1	1	0	0	0	1	2	1	1	10	3	0	1	1	0	1	1	0	0	1	0	2	10	○		
6G	3	2	3	1	1	2	5	0	2	1	3	3	26	7	1	1	1	0	1	4	0	1	0	2	0	18			
6H	3	2	3	1	1	0	1	0	0	2	1	2	16	1	1	1	1	2	1	0	0	1	1	0	1	10	○		
6I	4	0	6	4	1	4	3	0	1	1	4	0	28	0	1	4	2	0	3	4	0	0	0	3	3	20			
7A	2	1	1	2	0	2	3	0	0	2	1	2	16	2	1	1	2	4	2	1	0	0	0	1	2	16			
7B	0	3	2	0	1	1	0	0	1	2	5	3	18	0	1	0	0	3	1	2	0	1	0	5	1	14			
7C	4	0	3	5	1	1	3	0	0	1	1	1	20	4	2	1	5	3	1	1	0	0	1	1	3	22			
7D	2	1	2	0	1	3	0	0	1	2	4	4	20	2	1	0	0	5	3	2	0	1	0	4	0	18			
7E	1	2	1	2	1	0	3	0	0	2	1	1	14	1	0	1	2	3	0	1	0	0	0	1	3	12	○	○	7E(発言量5位)が最小で次が7F(同1位)と7B(同2位) ※7Bのグループ決定完全一致数5個は最多。7Bと7Fの絶対値誤差±1は9個で最多
7F	4	0	4	1	3	1	2	0	0	1	1	5	22	4	2	2	1	1	1	0	0	0	1	1	1	14			
7G	3	4	4	1	1	3	4	0	10	6	3	3	42	3	6	2	1	5	3	2	0	10	8	3	1	44			
7H	1	0	10	3	9	3	4	0	4	6	6	2	48	1	2	8	3	5	3	6	0	4	8	6	6	52			
7I	0	2	0	0	0	1	1	0	0	0	0	0	4	0	4	2	0	4	1	3	0	0	2	0	4	20	n/a		
8A	1	1	3	1	1	1	4	0	1	2	2	1	18	0	0	3	0	1	1	2	0	1	1	1	0	10			
8B	4	3	2	4	0	3	4	0	0	2	1	3	26	3	4	2	3	0	3	2	0	0	1	0	2	20			
8C	2	1	2	2	1	1	5	0	1	1	1	1	18	1	0	2	1	1	1	3	0	1	0	0	0	10			
8D	1	1	1	1	0	0	1	0	0	2	1	0	8	0	0	1	0	0	0	1	0	0	1	0	1	4	n/a	◎	8D(既知)を除くと、8E(発言量2位)が最小で次が8C(同5位)と8A(同8位) ※8Eのグループ決定完全一致数6個は最多。8Eの絶対値誤差±1は11個、8Hは同9個
8E	1	0	3	0	0	0	3	0	0	2	0	1	10	0	1	3	1	0	0	1	0	0	1	1	0	8	○		
8F	2	1	7	1	0	2	1	0	0	1	3	0	18	3	0	7	2	0	2	3	0	0	0	2	1	20			
8G	1	3	1	3	0	2	3	0	3	2	2	2	22	2	2	1	2	0	2	1	0	3	1	1	1	16			
8H	1	1	3	0	1	2	3	0	1	2	1	3	18	0	0	3	1	1	2	1	0	1	1	0	2	12			
9A	2	0	1	1	0	1	2	0	0	2	4	5	18	1	1	2	1	0	1	1	0	0	0	3	4	14			
9B	2	1	3	1	0	1	1	0	0	2	3	2	16	1	0	0	1	0	1	2	0	0	0	2	3	10	n/a		
9C	4	1	9	1	1	3	1	0	3	1	2	2	28	5	0	6	1	1	3	2	0	3	1	1	1	24			
9D	0	0	3	4	0	0	3	0	0	1	0	3	14	1	1	0	4	0	0	0	0	0	1	1	4	12			
9E	1	1	3	1	0	0	4	0	0	2	1	1	14	0	0	0	1	0	0	1	0	0	0	0	0	2	○	◎	9E(発言量1位)が最小、次が9F(同3位)
9F	4	0	3	0	0	1	3	0	0	2	2	1	16	3	1	0	0	0	1	0	0	0	0	1	0	6			
9G	1	3	3	0	0	2	3	0	4	2	0	4	22	0	2	0	0	0	2	0	0	4	0	1	3	12			
9H	4	5	7	7	0	2	0	0	0	4	5	2	36	5	6	4	7	0	2	3	0	0	6	4	3	40			
10A	4	1	2	6	0	1	4	0	0	2	5	1	26	3	1	1	5	0	2	3	0	0	0	5	2	22			
10B	1	1	7	8	1	5	1	0	5	2	6	5	42	2	1	6	7	1	2	0	0	5	0	6	2	32			
10C	1	1	0	1	2	4	1	0	1	1	0	0	12	0	1	1	0	2	7	0	0	1	1	0	3	16			
10D	3	1	1	5	0	4	5	0	0	3	4	0	26	2	1	0	4	0	1	4	0	0	5	4	3	24		×	10H(発言量4位)が最小、次が10G(同8位)
10E	2	5	7	0	1	3	0	0	1	2	0	1	22	1	5	6	1	1	0	1	0	1	0	0	2	18			
10F	1	1	1	8	3	4	0	0	0	2	4	4	28	2	1	0	7	3	1	1	0	0	0	4	1	20			
10G	2	2	2	2	0	1	1	0	0	2	0	2	14	1	2	1	1	0	4	0	0	0	0	0	5	14			
10H	0	1	2	2	1	5	0	0	1	2	2	4	20	1	1	1	1	1	2	1	0	1	0	2	1	12	○		

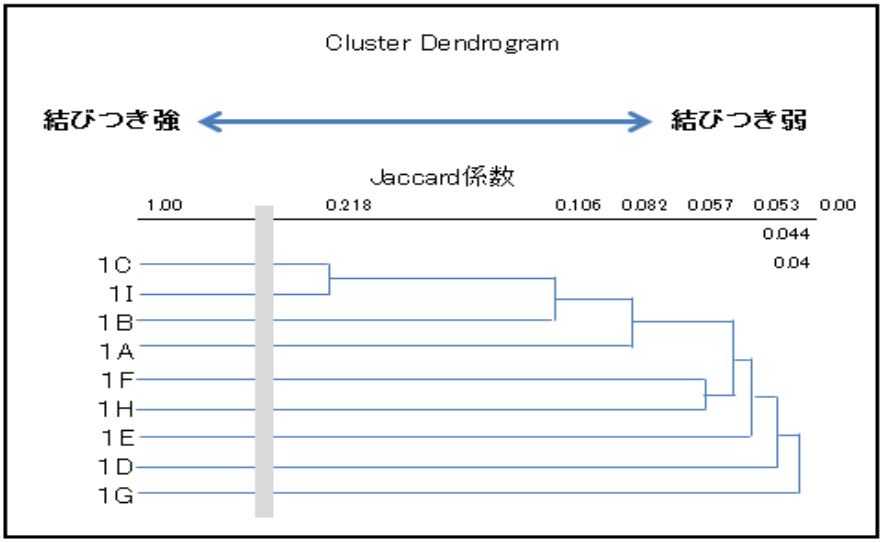
発言量1位の人
発言量2位の人
映画を以前に観たことがあり、かつ、結果の記憶が有る人

4-3-3　デンドログラムの描画

階層的クラスター分析を実施した結果は、可視化の方法として多くの場合デンドログラムを用いる。階層的クラスター分析では、通常、データから距離行列を基にコーフェン行列を求めてデンドログラムを描くというプロセスを経る。また、このときにクラスター間の距離をどのように求めるかで、最近隣法、最遠隣法、群平均法、メディアン法 、重心法、ウォード法などの手法が用いられ、様々な統計処理ソフトウェアで自動処理が可能となっている。

しかしながら、今回の行列は Jaccard 係数を用いたことから、これらのソフトウェアで単純に描画することができない。Jaccard 係数は 0 から 1 までの数値で表示され、数値が大きいほど結びつきが強いことを示す指標である。このため、今回作図するデンドログラムにおいて最初に作られる葉と葉の結節、すなわち１つ目の階層がとる値は既存ソフトウェアのクラスター分析で作図されるように最小値をとるのではなく、最大値を最初に結節する値としてとるためである。一例として、図表４０の結果をデンドログラムにすると図表４４のようになる。これは正にグループ１で繰り広げられた模擬討議の力動を表しているものと推察される。

図表４４．グループ１のデンドログラム



＜所見＞
①1Cと1Iの結びつきが最も強い
②次いで、(1C＆1I)と1Bの結びつきが強い
③次いで、(1C＆1I＆1B)と1Aの結びつきが強い
④次いで、1Fと1Hの結びつきが強い
⑤次いで、(1C＆1I＆1B＆1A)と(1F＆1H)の結びつきが強い
⑥次いで、(1C＆1I＆1B＆1A＆1F＆1H)と1Dの結びつきが強い
⑦最終的に、(1C＆1I＆1B＆1A＆1F＆1H＆1D)と1Gで一つのクラスターを形成

第５章　結論と提言

5-1　　導出された結論

5-1-1　問題意識の確認と考察

テキスト化したデータを基に、文字列の量や使用単語の量などに着目すると、グループ内で個人毎の発言量に大きく差があること、また、その発言量の多い人の使用する日本語表現が他のメンバーと比較して使用語彙が少ない傾向があることが明らかとなった。これらは各々のグループ討議において、集団の意思決定に何らかの影響を与えているものと思われる。

発言量が多いということは、一般的な表現でいうところの「声が大きい人」とイメージが重なる。そして声が大きい人というのは通常は発言力、すなわちその人の影響範囲において何らかのパワーを持っていることが想像される。一方、これらの人は発言量が多いのに、使用語彙が少ないという結果がＴＴＲの算定から導出されている。これは同じ言葉を繰り返し発言していることを意味している。これら発言量の多い人の行動が、集団の意思決定に何らかの影響を与えてはいないであろうか。

集団において一つの意思決定を図るとして、立場に上下関係がない者同士でこれを行うとした場合、意見を主張するものが一人しか居ないか、あるいは全員が同じ意見である場合を除き、必ず誰かが妥協しなければ最終的に合意形成に至らない可能性がある。従って、これらには妥協を促すための制約条件が設けられている必要がある。

ＭＢＡプログラムの授業の一つであることを考えれば、模擬会議といえどもゲームの結果、自らのグループの予測結果が他に勝るよう各個人にとっての動機付けが働くとの想定に異論はないであろう。しかし、グループ予測の良し悪しが個人の成績に結びつくことはない。予測値の良否の前に各人にとって最悪の結果は９０分間の時間内にグループ決定が導出できないことである。

つまり、今回は第一に時間内で結論を導出すること、第二に良い予測結果を導出すること、この二つがこの集団が互いの意見を摺合せ、合意をするための重要なコンテクストとなっている。

一方、集団の意思決定を図るために議論を行うということを行動科学的に労力の消費と考えたとき、当たり前であるが各人の意見が一致していることが最も労力の消費が少ないが、それを望むことには無理がある。意見が異なる場合、自らの意見を主張

するか、相手の意見を受け入れることとなるが、比較論で言えば、

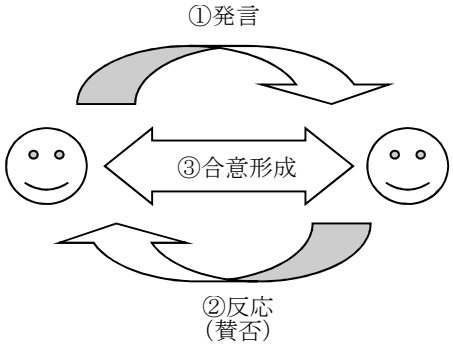
相手を説き伏せる労力　＞　相手の意見に妥協する労力

とはならないであろうか。

もし、発言をしたときに自分の意見が少数派であれば、賢明な者は自身の意見を引く方向に行動するかもしれない。最も争いが少なく済むことが最も経済的であるからである。一方、意見の孤立は避けたいと考えるかもしれない。なぜなら必ず少数意見は自らがそのように主張する理由を問われるからである。

こうした心理面を慮ると、その場を仕切る合理主義者が居たとして、意識的に行動するかどうかは分からないが同一グループで自分と意見の合う他の者が見つかれば非常に好都合だと考えるかもしれない。なぜなら、図表４５に示すとおり集団の意思決定は少なくとも二者の会話があってはじめて成立するため、発信者がいて受信者がいるだけでは合意形成ができたことにはならないからである。

図表４５．会話の合意形成モデル



会議の場においては、最低でも自分以外に同意してくれる者が一人以上いることで、少数派の妥協を促す方向に場が流れていく可能性がある。映像を見ただけでは判断がつかないが、何れかの者がグループ内で繰り返しされる討議を通じて「人間関係の紐帯を作っていた」とは考えられないであろうか。これらの観点で、客観的に判明した事実を整理してみたい。

5-1-2　確認された事実

（１）確認された事実１：グループ予測と発言量上位の人の個人予測の相関に関する事実①

会議の合意決定されたグループ予測と、図表３５に示されたグループ内の発言量が上位１位又は２位の人の個人予測とを比較したとき、その誤差が外のグループ内の誰よりも小さいケースが６０％であった。言い方をかえると１０のグループのうち６つのグループでの合意決定は、そのグループの上位１位又は２位の発言量の人の個人決定に酷似しているということである。

（２）確認された事実２：グループ予測と発言量上位の人の個人予測の相関に関する事実②

上記（１）の残りの４つのグループ予測と当該グループの発言量上位１位又は２位の人の個人予測とを比較すると、誤差の大小ではなく、絶対値の完全一致の個数、若しくは±１以内の値が最も多いケースが２つのグループ で認められた。つまり、発言量上位１位又は２位の人の予測がそのまま若しくは±１の範囲でグループ決定に採用されているケースが残りの２０％でみられた。

（３）確認された事実３：発言量上位の人の語彙表現に関する事実

タイプ・トークン比率を算出すると、発言量上位１位の人が最もＴＴＲの低い結果を示した例が７０％、また上位２位の人が最も低い結果を示した例が２０％であった。

（４）確認された事実４：発言量上位１位と２位の結びつきに関する事実

Jaccard 係数を算出してクラスター分析を行った結果、９０％の確率で発言量上位１位と２位の人同士の組み合わせが最も強い結節度であるとの結果が得られた。

（５）確認された事実５：グループ内での価値創造に関する事実

図表１６に示したグループ２、グループ８及びグループ９において散見された事象で、グループの中で個人予測の誤差が最も小さい（最も正解に近い予測をした）人の予測よりも全体で議論したあとの予測値の誤差が小さくなる事象（グループ討議の価

値創造？）について、事前アンケート及び実写映像の分析により明らかとなった事実は当該３つグループには、いずれも映画を以前に見たことがある既知の参加者がいた。

5-1-3　結論

これらの結果を総合的に判断すると、各グループの発言量の上位者、特に１位又は２位の人が集団の中で会議を進めていく過程で頻繁に言葉のキャッチボールを繰り返すことでお互いの影響力の確認を行い、人間関係の紐帯を作あげていく。これらは集団において階層クラスターを形作り、結節力の強い者同士が意見を発すると、結節力の弱い者に対しては妥協を促す形での力動が働いていたのではないかという仮説が浮かび上がる。

この仮説を裏付ける証左として、今回の Jaccard 係数が示す人と人との結びつきの強さ、すなわちクラスター分析の結果がこれを示しているものと推察される。そして、合意形成の過程で生じたグループ内の力動の階層構造を可視化すると図表４４に示したデンドログラムのような形になっていたのではないか、というのが本研究で明らかとなった事実である。

また、グループによる協働作業によって得られた価値創造（グループ予測の誤差が個人予測の最低値を上回る結果が齎されたこと）については、そのグループの中によりハイスコアを示した映画の事前情報について既知であった参加者がいて、発言量１位又は２位の議論の主導者たちがこれらの既知のメンバーの個人予測結果の値を意識しつつグループ予測に反映したことが要因にあったかもしれない。事前アンケートと映像のモニタリング結果からはこういった推測を否定できない結果が得られているからである。

5-2　仮説検証のための追加調査

5-2-1　仮説検証の方法

これまでの仮説を検証するため、筆者は自らが属する M34 期の同じ授業を検証してみることとした。具体的な手続きは次の２つの調査を実施する。

まず、一つ目の調査は、図表４６に示すアンケートを M34 期の全員を対象として収集することとした。このアンケートの実施目的は次の 2 点である。

①録画記録等がなく、定量的にグループ内で声が大きかったものを特定することができないため、アンケートを用いて同一グループ内で発言量が多かったと思う上位３名に投票してもらい、声の大きかった上位二人を割り出す

②ノイズの排除のため、映画を事前に見たことのある既知の参加者を特定する

次に、２つ目の調査は１２あるグループの代表者から、グループ討議で使用した実物の集計表を入手し、当該グループの各個人予測の誤差、そのグループの平均誤差、及びグループ予測の誤差の３つを求める。

この結果、アンケート①を基に特定した各グループの発言量上位２名のうち、どちらかの予測がグループ予測と近似していれば仮説はそれなりに信憑性があるものと裏付けられることになる。（当然②の結果は考慮される）

図表４６．アンケート実施結果イメージ

5-2-2　追加調査の実施結果

追加調査は、2012/11/11 から 2012/11/18 までの期間で実施した。入手した情報の取りまとめは次のとおりである。

まず、12 グループ全ての個人予測及びグループ予測結果を入手し、集計した結果が別紙１０、別紙１１である。また、図表４６のアンケートを実施した結果、回答者は 92 名中 62 名で 67.4%と一定の評価が可能な水準の回収率を得た。アンケート結果の集計結果は図表４７と図表４８のそれぞれ①欄「アンケート結果」のとおりとなった。

発言量の多い人をどのように特定したかについては、各グループで発言量が多かったと投票された人の各得点を、１位３ポイント、２位２ポイント、３位１ポイントと換算して集計した。結果は図表４７と図表４８の「①アンケート結果」欄の「発言量」の数値のとおりとなっている。

次に、この発言量の数値が大きいものから順に１位、２位としてハイライトをかけ、次に別紙１０、別紙１１の集計過程で得られたグループ予測と個人予測を比較したものを「②グループ決定と個人決定の誤差の絶対値」欄にインプットした。そして「③誤差最小者」欄には、②の「合計」欄が最も小さい値の人＝グループ予測に近似している人として「○」を付与している。

判定結果欄は次のとおり◎、○、×の３分類を行った。

第一に、◎のものは仮説どおりの結果を示している。これは発言量上位１位又は２位の人の個人予測とグループ予測を比較したときに、誤差の絶対値がグループ内で最も小さい結果が得られたことを示す。つまり、発言量上位１位又は２位の人の個人予測とグループ予測が近似していたことを意味するものである。

第二に、○のものは発言量上位１位又は２位の人の個人予測よりもグループ予測に近似していた人がほかに居るが、発言量上位１位又は２位のものの個人予測とグループ予測を比較し、１２個ある予測値の絶対値が完全一致するか、若しくは±１以内である個数がグループ内で１位であるケースを示している。この区分を作った理由は、予測値が完全一致、若しくは±１であれば、当該発言量上位者にとって「自分の主張が受け入れられた個数が最大であること」で相当の充足感が得られているという推測による。

第三に、×のものは上記以外のケースを示している。

結果は、◎が 8/12（66.7%）、○が 2/12（16.7%）、×が 2/12（16.7%）であった。

図表４７．追加アンケート結果と考察（1/2）

CD	性別	①アンケート結果				②グループ決定と個人決定の誤差の絶対値													③	判定結果	コメント
		1位	2位	3位	発言量	陪審員	銀行員	メッセンジャー	ブローカー	スラム	ペンキ	セールスマン	建築師	老人	ガレージ	時計屋	広告代理店	合計	誤差最小者		
1A	M			1	**1**	0	2	0	0	5	1	0	0	1	0	2	1	12		◎	1Eと1Hが発言量上位2名であり、③列の誤差最小者に該当
1B	M			1	**1**	4	3	5	0	4	3	0	0	0	1	1	1	22			
1C	M				**0**	0	1	1	1	3	1	1	0	0	1	2	1	12			
1D	F		1	1	**3**	3	2	0	0	5	2	0	0	1	0	1	2	16			
1E	M	2	2		**10**	1	1	0	0	3	3	0	0	0	0	0	0	8	○		
1F	M	1	1	1	**6**	4	4	5	2	4	1	7	0	0	2	2	3	34			
1G	F				**0**	1	8	2	3	5	1	0	0	3	1	5	3	32			
1H	M	2	1	1	**9**	0	1	1	0	0	0	1	0	0	0	3	2	8	○		
2A	M			1	**1**	0	2	2	4	0	3	1	0	6	1	2	0	21		◎	2Hと2Eが発言量上位2名であり、うち2Eが③列の誤差最小者に該当
2B	M			1	**1**	3	2	0	1	0	3	3	0	2	0	2	2	18			
2C	M		1	3	**5**	0	3	0	4	2	2	1	0	0	1	3	0	16			
2D	M				**0**	2	4	0	0	1	3	5	0	1	5	6	1	28			
2E	F	3	2	1	**14**	1	2	1	0	2	1	0	0	0	1	0	0	8	○		
2F	M				**0**	2	1	2	0	0	0	2	0	0	0	1	0	8	○		
2G	M				**0**	0	0	4	3	0	0	1	0	0	0	0	2	10			
2H	M	3	3		**15**	3	2	1	2	0	1	0	0	0	1	1	3	14			
3A	M				**0**	3	4	5	1	1	0	4	0	1	1	1	1	22		◎	3Fと3Gが発言量上位2名であり、うち3Fが③列の誤差最小者に該当
3B	F			1	**1**	1	4	3	0	0	4	2	0	0	4	0	2	20			
3C	M		2	1	**5**	1	4	6	2	0	7	3	0	0	1	0	2	26			
3D	M		1	2	**4**	1	1	1	1	0	2	4	0	1	1	3	5	20			
3E	F				**0**	3	3	2	1	2	2	3	0	1	1	3	1	22			
3F	M	4	1	1	**15**	0	5	0	1	0	0	1	0	3	0	1	1	12	○		
3G	M	3	2	1	**14**	1	5	1	1	4	4	2	0	4	1	4	3	30			
3H	M		1	1	**3**	1	2	5	1	0	0	6	0	0	1	0	0	16			
4A	M		3	1	**7**	2	3	0	1	1	3	3	0	3	0	5	3	24		○	4Dと4Aが発言量上位2名。③列の誤差最小者に該当しないが、4Dの誤差をみると±1以下の個数が9/12個で最多
4B	M				**0**	2	0	0	8	1	4	0	0	7	0	1	3	26			
4C	M				**0**	2	4	1	2	1	0	0	0	1	3	4	0	18			
4D	M	3		1	**10**	**1**	**0**	**1**	3	**1**	**0**	**1**	**0**	**1**	**1**	3	4	16			
4E	M				**0**	2	3	7	2	0	0	0	0	7	1	2	6	30			
4F	M	1	1	1	**6**	2	3	0	3	6	6	1	0	3	7	2	3	36			
4G	F			1	**1**	3	**1**	**0**	**0**	**0**	3	2	**0**	**0**	**0**	3	2	14	○		
4H	M				**0**	**0**	**1**	3	2	**0**	**1**	3	**0**	**0**	**0**	3	**1**	14	○		
5A	M	1	2	1	**8**	2	1	3	7	1	2	2	0	0	2	4	6	30		◎	5Dと5Aが発言量上位2名であり、5Dが③列の誤差最小者に該当
5B	M				**0**	4	2	1	4	2	2	4	0	2	0	4	3	28			
5C	M			2	**2**	2	1	0	1	1	3	3	0	2	1	1	1	16			
5D	F	5	1		**17**	4	2	1	0	1	1	0	0	2	0	2	1	14	○		
5E	F				**0**	1	1	6	2	3	1	4	0	2	0	0	2	22			
5F	M			1	**1**	2	1	5	1	1	2	0	0	2	6	2	4	26			
5G	M		2	2	**6**	6	3	2	2	0	1	0	0	1	1	4	5	25			
5H	M		1		**2**	2	1	3	1	1	2	1	0	0	1	2	0	14	○		
6A	M				**0**	1	0	3	2	1	1	1	0	1	3	0	1	14		◎	6Cと6Bが発言量上位2名であり、6Cが③列の誤差最小者に該当
6B	M	1	2		**7**	2	1	2	5	1	1	2	0	0	0	3	1	18			
6C	M	4	1		**14**	0	0	1	2	1	0	1	0	1	2	2	2	12	○		
6D	M		2	2	**6**	1	0	1	5	0	2	2	0	0	2	1	0	14			
6E	F				**0**	7	5	2	3	1	2	4	0	1	1	7	3	36			
6F	M				**0**	1	0	2	3	0	1	2	0	0	0	3	2	14			
6G	F			1	**1**	4	1	4	4	0	2	3	0	0	1	4	3	26			
6H	M			2	**2**	5	1	0	1	1	1	1	0	2	1	0	5	18			

図表４８．追加アンケート結果と考察（2/2）

CD	性別	①アンケート結果				②グループ決定と個人決定の誤差の絶対値													③	判定結果	コメント
		1位	2位	3位	発言量	陪審員	銀行員	メッセンジャー	ブローカー	スラム	ペンキ	セールスマン	建築師	老人	ガレージ	時計屋	広告代理店	合計	誤差最小者		
7A	F		1	2	**4**	1	0	1	1	0	1	1	0	0	2	1	0	8		◎	7Eと7Bが発言量上位2名であり、7Eが③列の誤差最小者に該当
7B	M	2	1		**8**	3	2	0	1	2	3	1	0	1	1	6	0	20			
7C	M		1	1	**3**	1	1	3	6	1	1	3	0	4	0	0	2	22			
7D	M		2		**4**	3	1	1	0	1	3	1	0	2	0	1	3	16			
7E	M	4		1	**13**	1	0	2	1	0	0	1	0	2	1	2	0	10	○		
7F	M				**0**	2	1	4	2	1	6	2	0	2	0	1	1	22			
7G	M		1	2	**4**	5	0	2	3	7	6	2	0	1	5	4	5	40			
7H	M				**0**	3	1	2	2	2	3	3	0	1	2	1	0	20			
7I	F				**0**	2	2	4	3	0	4	4	0	1	2	2	0	24			
8A	F		1		**2**	2	0	0	5	1	0	2	0	1	0	3	2	16		×	8Eと8Bが発言量上位2名であるが、③列の誤差最小者に該当せず。個別予測値も2番目に多いに過ぎない。
8B	F	1	1	2	**7**	2	1	3	1	1	1	1	0	2	1	2	1	16			
8C	M			1	**1**	2	1	0	6	0	0	4	0	0	4	3	0	20			
8D	M				**0**	2	1	0	0	1	3	1	0	0	0	0	2	10	○		
8E	M	3	1		**11**	4	3	2	7	1	2	2	0	3	0	1	3	28			
8F	M				**0**													-			
8G	M	1	1	1	**6**	4	2	4	1	1	1	3	0	1	1	0	2	20			
8H	M		1	1	**3**	1	1	1	2	1	4	1	0	4	2	2	5	24			
9A	M	1		2	**5**	2	1	1	4	2	1	3	0	1	1	1	1	18		◎	9Dと9A、9Gが発言量上位2名であり、9Dが③列の誤差最小者に該当
9B	F				**0**	2	8	4	2	1	0	5	0	0	0	1	1	24			
9C	M				**0**	1	0	0	1	2	2	1	0	2	1	1	1	12			
9D	M	2	1		**8**	1	1	1	0	1	1	1	0	0	0	1	3	10	○		
9E	F				**0**	4	1	0	5	7	2	0	0	2	4	1	4	30			
9F	M				**0**													-			
9G	M		2	1	**5**	4	0	5	0	2	4	1	0	2	2	0	0	20			
9H	M				**0**	2	0	1	1	3	1	3	0	0	3	1	1	16			
10A	M				**0**	3	4	4	0	2	1	1	0	0	0	0	1	16		○	10Fと10Dが発言量上位2名であるが、③列の誤差最小者に該当しないが10Fと10Bの誤差をみると±1以下の個数が9/12個で最多タイ
10B	M				**0**	2	2	0	0	0	0	1	0	0	2	0	1	8	○		
10C	M		1		**2**	2	2	4	3	1	4	2	0	1	1	0	2	22			
10D	M	1	1	3	**8**	3	5	3	0	0	0	1	0	2	0	2	2	18			
10E	F				**0**	4	0	5	1	0	0	0	0	0	0	0	0	10			
10F	M	3	2		**13**	3	5	3	1	1	1	1	0	1	1	1	0	18			
10G	F	1	1	2	**7**	0	0	6	3	0	1	0	0	0	1	0	3	14			
10H	M	1	1	1	**6**	0	1	5	1	2	1	2	0	1	6	2	1	22			
11A	M	3	1	1	**12**	2	4	2	4	1	2	1	0	1	1	1	1	20		◎	11Aと11Gが発言量上位2名であり、11Gが③列の誤差最小者に該当
11B	M				**0**	1	3	1	1	1	3	1	0	2	0	5	2	20			
11C	M				**0**	5	4	5	3	1	4	2	0	4	3	1	2	34			
11D	F				**0**	0	4	0	2	1	0	3	0	1	1	2	0	14			
11E	M				**0**	1	1	1	7	0	2	2	0	0	0	1	3	18			
11F	M		1	2	**4**	3	2	5	5	1	1	2	0	1	0	0	2	22			
11G	M	2	3		**12**	2	2	0	0	0	1	1	0	0	0	2	2	10	○		
11H	M			2	**2**	1	1	0	2	0	1	2	0	4	0	2	1	14			
12A	M			1	**1**	5	5	1	0	1	1	4	0	1	1	2	3	24		×	12Bと12Fが発言量上位2名であるが、③列の誤差最小者に該当しない個別予測値も2番目に多いに過ぎない。
12B	M	4			**12**	5	1	3	2	2	1	1	0	2	2	7	2	28			
12C	M				**0**	2	1	0	1	2	3	0	0	2	0	2	1	14			
12D	M				**0**	1	1	2	1	2	1	1	0	2	0	1	0	12	○		
12E	F				**0**	5	1	9	1	0	0	1	0	1	1	4	1	24			
12F	M		3	1	**7**	0	3	1	6	1	3	2	0	0	1	1	2	20			
12G	F		1	2	**4**	2	1	2	1	2	5	2	0	2	0	2	1	20			
12H	M				**0**	5	0	4	1	3	0	1	0	2	1	1	2	20			

5-3　インプリケーションと今後の課題

これらの得られた知見を実際の組織運営に反映させることを考えてみる。

本研究の目的を振り返ると、第一義的には集団の意思決定の場面における、より高効率で、かつ、創造性の高いコミュニケーションをいかに作り出すかという本質論に対して、直接的な解決を試みるものではなく、それらの阻害要因に関する新たな知見を得ることを目的としていた。

今回の結果を鑑みると、創造性の高いコミュニケーションを作るのは発言量の多い、すなわち声の大きなリーダーが優秀であったときに、その集団の成果が大きくなるという結果が得られたこととなる。これは、至極当然のことかもしれない。集団の意思決定における成果はリーダーの資質に依存するということであり、会議の運営をより効率的にすることとは基本的にあまり関係がないことを意味している。

また違った見方をすると、ある集団を見回してみたときに、その集団の中で誰の声が大きいかを調査し、加えてその声の大きな人の能力がわかれば、凡そその集団の生み出す結論の出来映えが予測可能であるとも言える結果が得られた。

しからば、組織においては優秀なリーダーの資質ある人の声が、より大きくなるように育成を図るべきである。何れが優秀か、この見定めが出来かねるということであれば、候補者全てにおいてこれらの訓練を積ませていく必要がある。

しかし、一方ではグループ内に予測能力の高い成績上位者がいるにもかかわらず、グループ予測が向上しなかった例も散見されている。これは筆者の推測になるが、その人は自身の効用最大化の方向が、良く言えば集団優先、悪く言えば体力温存＝妥協の方向に志向されていたのではないだろうか。これは組織にとっては由々しき問題である。このような行動をした人が声を大きくするためのモチベーションを如何にしてつけるかが課題となろう。

最後になるが、本研究はまだ時間的制約から踏み込めていない分析対象が山のようにある。また、大変な稼働を要するが本研究で得られる映像記録や授業データ等はとても大きな発見をもたらす宝の山であり、更にサンプルを増やしていくことが望まれる。本研究に興味をもってくれた諸君が、更なる探求を重ねてくれることを筆者は切に願っている。

以　上

参考文献

- 池尾恭一，井上哲浩，（2008）．「戦略的データマイニング~アスクルの事例で学ぶ」
- 石井哲，(2002)．「テキストマイニング活用法」
- 石田貞夫，加藤千恵子，劉晨，石村友二郎，（2010）．「多変量解析によるデータマイニング」
- 石田基広，（2008）．「R によるテキストマイニング入門」
- 上田隆穂，黒岩祥汰，戸谷圭子，豊田裕貴，（2005）．「テキストマイニングによるマーケティング調査」
- 上村龍太郎，吉田文彦，山下俊恵，（2007）．「社会知能システム入門」
- 北村則龍，（1992）．「グループ意思決定プロセスにおけるコミュニケーションメカニズム」慶應義塾大学大学院経営管理研究科修士課程学位論文
- 喜田昌樹，（2007）．「組織革新の認知的研究」
- 喜田昌樹，（2008）．「テキストマイニング入門：経営研究での活用法」
- 金 明哲，（2009）．「テキストデータの統計科学入門」
- 佐藤良明訳,(1982).「精神と自然」Gregory Bateson, Mind and Nature：A Necessary Unity, 1972
- 佐藤良明訳，(1990)．「精神の生態学」Gregory Bateson, Steps to an Ecology of Mind, 1972
- 高木晴夫，（1994）．「共同活動のための創造的コミュニケーション」，慶應経営論集　第 11 巻第 2 号 pp. 43-62.
- 竹内郁郎，(1973).「社会的コミュニケーションの構造：現代の社会とコミュニケーション 1」
- 辻井潤一 監訳，(2010). The Text Mining Handbook: Advanced Approaches to Analyzing Unstructured Data Ronen Feldman and James Sanger (Bar-Ilan

University and ABS Ventures) Cambridge, England: Cambridge University Press, 2007.

- 豊田裕貴（2003）．「ブランド自由連想分析による類似化・差別化ポイントの尺度化：新たなブランドイメージ分析の構築」『日経広告研究所報』207 号
- 林　俊克（2002）．「Excel で学ぶテキストマイニング入門」
- 増田有孝，柿木彰，神田晴彦，（2012）．「ビッグデータ革命」
- 松村真宏，三浦麻子，（2009）．「人文・社会科学のためのテキストマイニング」
- Jeffrey Dean and Sanjay Ghemawat, (2004). MapReduce: Simplified Data Processing on Large Clusters

  http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/ja//archive/mapreduce-osdi04.pdf
- Matsumura, N, Kato, Y., & Ohsawa, Y. (2003). Visualization of discussion structure for discovering and understanding discussion points. *Journal of Japan Society*, (Idm). Retrieved from http://sciencelinks.jp/j-east/article/200308/000020030803A0122135.php
- Matsumura, N, & Ohsawa, Y. (2002). Automatic indexing based on term activity. *Transactions of the*, 398-406. Retrieved from http://adsabs.harvard.edu/abs/2002TJSAI..17..398M
- Matsumura, Naohiro, Corporation, T., Ohsawa, Y., & Ishizuka, M. (n.d.). Influence Diffusion Model in Text-Based Communication, 259-267.
- Medelyan, O. (2009). Human-competitive automatic topic indexing. *Machine Learning*, *1994*(July), 287-290. The University of Waikato. Retrieved from http://researchcommons.waikato.ac.nz/handle/10289/3513

附録

別紙１	Ｍ３５学生に対する説明資料	76
別紙２	アンケート（グループ討議開始前）	77
別紙３	アンケート（グループ討議終了後）	79
別紙４	グループ討議卓上の注意喚起用紙	81
別紙５	使用機材リスト	82
別紙６	Ｍ３５の「１２人の怒れる男」観察研究の集計結果（１）	83
別紙７	Ｍ３５の「１２人の怒れる男」観察研究の集計結果（２）	84
別紙８	グループ討議開始前アンケートの集計結果（全体）	85
別紙９	グループ討議終了後アンケートの集計結果（全体）	86
別紙１０	Ｍ３４の「１２人の怒れる男」観察研究の集計結果（１）	87
別紙１１	Ｍ３４の「１２人の怒れる男」観察研究の集計結果（２）	88
別紙１２	テープ起こし業者とのやりとりの記録	89
別紙１３	ＴＴＭのバグに関する松村教授との対応模様	91

別紙１　Ｍ３５学生に対する説明資料

2012/4/6

MBA１年生に研究協力のお願い

髙木晴夫研究室

高木研究室では集団のコミュニケーションがどのように進行するかについて研究しています。この研究は、チーム活動の効率化や、会議の生産性向上に資するものです。

この入学合宿の「組織マネジメント」（清水教授担当）で行われる「１２人の怒れる男」の授業のグループ討議について、ぜひ録音録画をさせていただきたくお願いいたします。この研究企画については、授業担当の清水教授、学習指導の田中教授と坂爪教授から事前了承を得ています。

この研究に先行する成果は論文「協働活動のための創造的コミュニケーション」にまとめられ、みなさんの授業で配布されているものです。今回の新しい研究では、現代の新しいコミュニケーション解析技術（テキストマイニングなど）を応用するものです。

録音録画は高木研究室で分析するためだけに用い、清水先生に提供することは絶対にありません。授業の成績とはまったく関係しません。また個人の発言など個人情報の機密取り扱いは塾の規則をまもっておこないます。

具体的には、以下のとおりの実施です。

１．実施日時

2012 年 4 月 7 日（土曜日）　13:00～16:15　（の間のグループ討議について）

２．実施方法

①機器の設置

・各グループ室（１２室）へ、ビデオカメラ及び IC レコーダーを各１台配備。

・全体講義終了の前に研究スタッフがスイッチを入れ記録を開始しておく。

②会議中

・カメラと IC レコーダーは勝手に回っています。そのまま放っておいて下さい。

・座席表の記入をお願いします。

※討議終了後、「集計表」をお借りしますので机上へ１部置いておいて下さい。

③機器の回収

・グループ討議終了後、研究スタッフが機器のスイッチを止め、機器を回収します。

以　上

別紙２　　アンケート（グループ討議開始前）

2012 年４月７日

「１２人の怒れる男」グループ討議開始前アンケート

本アンケートは、集団のコミュニケーションに係る研究の一環として行うものです。「MBA１年生に研究協力のお願い」に記載したとおり、アンケートの内容は皆さんの成績に一切関係ありません。思うままにお答えください。

回収はグループごとに行いますが、ここから個人を特定することはありません。

以下の設問で該当する番号を一つ選択し、○で囲んでください。

１．性別を教えてください

１．男性　　２．女性

２　「12 人の怒れる男」について、事前情報を得ていましたか

１．「12 人の怒れる男」を見たことがあり、内容を記憶している

２．「12 人の怒れる男」を見たことがあるが、内容は記憶していない

３．Web 等で情報を収集したことがあり、その内容を記憶している

４．Web 等で情報を収集したことがあるが、その内容を記憶している

５．授業前に事前に配布された資料を読んだ

６．知識は全くない

３　あなた個人が導き出した回答に対する自信の度合いを教えてください。

１．非常に自信がある　　２．自信がある　　３．自信がない　　４．非常に自信がない

４　このグループ討議において解決すべき問題は何だと思いますか。一つだけ選んでください。

１．メンバーの考え方を理解すること

２．個人の解答より良い（正解の多い）グループ回答をつくること

３．制限時間内にグループのコンセンサスを形成すること

４．決定したルールに従って討議を進めること

５．参加者が多くの発言をするなど、活発な議論を行うこと

６．これ以外の問題（　　）

５　グループで討議することが成果物の質を向上させると考えますか？

今回の課題に限定せずにお答えください。

１．非常にそう思う

２．そう思う

３．そう思わない

４．全くそう思わない

６　グループで討議することは、ビジネスを学ぶうえで有効に働くと思いますか？

今回の課題に限定せずにお答えください。

１．非常にそう思う

２．そう思う

３．そう思わない

４．全くそう思わない

以上、回答にご協力いただき、ありがとうございました。

別紙３　アンケート（グループ討議終了後）

2012 年４月７日

「１２人の怒れる男」グループ討議終了後アンケート

グループ討議お疲れ様でした。引き続き、簡単なアンケートへのご協力をお願いいたします。
以下の設問で該当する番号を**一つ**選択し、〇で囲んでください。

１　あなたのグループワークにおける貢献度を教えてください。

１．非常に貢献した

２．貢献した

３．貢献できなかった

４．全く貢献できなかった

２　グループで導き出した解答に対する満足度合いを教えてください。

１．非常に満足である

２．満足である

３．不満である

４．大変不満である

３　グループ討議全体に対する満足度合いを教えてください。「活発な討議ができた」「他メンバーの知見から学ぶことができた」あるいは「良い発言ができた」など、討議全体に対して感じたことを鑑みて、お答えください。

１．非常に満足である

２．満足である

３．不満である

４．大変不満である

４　グループ討議を終えて、あなた個人の解答は変わりましたか？

１．変わらなかった

２．変わった

５　あなたは今回のグループワークにおいて、自らリーダーシップを発揮できたと思いますか？

１．非常に発揮できた

２．発揮できた

３．発揮できなかった

４．全く発揮できなかった

６　質問５において３．または４．と回答した方は、以下の質問もお答えください。あなた以外の他者のリーダーシップはグループ討議において有効だったと思いますか？

１．非常にそう思う

２．そう思う

３．そう思わない

４．全くそう思わない

以上、回答にご協力いただき、ありがとうございました。

別紙４　グループ討議卓上の注意喚起用紙

ご協力ありがとうございます

機材の不具合を発見したら
Please Call　村田（０９０－ＸＸＸＸ－ＸＸＸＸ）
釜口（０９０－ＸＸＸＸ－ＸＸＸＸ）

お伝えしたとおり、録音録画は高木研究室で分析するためだけに用い、清水先生に提供することは絶対にありません。**授業の成績とはまったく関係しません**。また個人の発言など個人情報の機密取り扱いは塾の規則をまもっておこないます。

グループ討議後、全体討議が始まる前にアンケートを回収します。**グループ代表者はアンケート票をまとめて、**高木晴夫ゼミメンバーにご提出ください。

また、**授業終了後に「集計表」を１部コピーしていただき、**高木ゼミメンバーにお渡しください。ゼミメンバーは、コピー室または大教室におります。

別紙５　**使用機材リスト**

①　ビデオカメラ　　SONY HXR-NX70J　×　１２台

慶應義塾大学湘南藤沢メディアセンター（神奈川県藤沢市遠藤 5322）より借用

②　三脚　×　１２脚

ＫＢＳ学内（高木ゼミ、河野ゼミ、学事）及びＭ３４学生の個人所有を借用

③　ＩＣレコーダー　×　１２個

ＳＦＣより６個及びＭ３４学生の個人所有を借用

別紙６　　Ｍ３５の「１２人の怒れる男」観察研究の集計結果（１）

G#	個人データ					グループ内の個人データ集計値					グループの決定				改善効果	改善率	改善率順位
	P#	性別	系別	年代	個人誤差	誤差平均	分散	誤差最小値	誤差最大値	誤差中央値	誤差	誤差最小値	誤差最大値	誤差中央値			
						①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩=①-⑥	⑪=⑩/①	⑫
1	1A	M	文	40	18	17.3	5.7	10.0	28.0	18.0	14	8	24	15	3.3	19.2%	7
	1B	F	文	20	16												
	1C	M	文	40	10												
	1D	M	文	30	28												
	1E	M	文	30	22												
	1F	M	理	30	14												
	1G	M	文	20	10												
	1H	M	文	40	20												
	1I	M	文	40	18												
2						26.6	6.4	18.0	36.0	26.0	12				14.6	54.8%	2
	2B	M	理	30	26												
	2C	M	理	20	36												
	2D	M	理	30	26												
	2E	M	文	30	22												
	2F	M	文	30	34												
	2G	F	文	30	24												
	2H	F	文	50	18												
3	3A	M	文	30	28	29.6	13.0	12.0	46.0	28.0	24				5.6	18.8%	8
	3B	M	理	30	30												
	3C	M	文	40	46												
	3D	M	文	30	20												
	3E	M	文	30	12												
	3F	M	文	30	18												
	3G	F	文	30	44												
	3H	F	理	30	46												
	3I	M	文	30	22												
4	4A	F	文	30	26	28.5	8.7	14.0	40.0	28.0	24				4.5	15.8%	9
	4B	M	文	30	36												
	4C	M	文	30	22												
	4D	F	理	40	30												
	4E	M	文	20	40												
	4F	M	文	40	36												
	4G	M	理	40	14												
	4H	M	理	40	24												
5						22.8	5.2	14.0	30.0	23.5	16				6.8	29.9%	6
	5C	F	文	30	24												
	5D	M	文	30	14												
	5E	M	文	30	24												
	5F	M	文	30	22												
	5G	M	文	40	23												
	5H	M	文	20	30												
6	6A	F	理	40	32	21.3	7.6	10.0	32.0	22.0	18				3.3	15.6%	*10*
	6B	M	文	40	16												
	6C	M	文	30	28												
	6D	M	文	30	14												
	6E	M	文	30	22												
	6F	M	理	30	10												
	6G	M	理	30	26												
	6H	M	理	30	16												
	6I	M	理	30	28												
7	7A	M	理	40	16	25.0	12.7	14.0	48.0	20.0	16				9.0	36.0%	5
	7B	M	理	30	18												
	7C	M	文	40	20												
	7D	F	文	30	20												
	7E	M	文	30	14												
	7F	F	文	20	22												
	7G	M	理	30	42												
	7H	M	理	20	48												
8	8A	F	文	30	18	18.6	4.9	10.0	26.0	18.0	8				10.6	56.9%	*1*
	8B	M	文	30	26												
	8C	F	理	30	18												
	8E	M	文	40	10												
	8F	M	文	40	18												
	8G	M	文	30	22												
	8H	M	文	30	18												
9	9A	F	文	30	18	21.1	8.2	14.0	36.0	18.0	12				9.1	43.2%	4
	9C	M	文	20	28												
	9D	M	文	30	14												
	9E	M	文	30	14												
	9F	M	文	30	16												
	9G	M	文	40	22												
	9H	M	文	30	36												
10	10A	M	文	30	26	23.8	9.3	12.0	42.0	24.0	12				11.8	49.5%	3
	10B	M	文	30	42												
	10C	M	文	30	12												
	10D	M	文	30	26												
	10E	M	理	30	22												
	10F	M	文	30	28												
	10G	F	文	20	14												
	10H	M	文	30	20												
Gr.平均					23.5	**23.5**	**8.2**	**12.8**	**36.4**	**22.6**	**15.6**				**7.9**	**34.0%**	–
M35個人					23.5	**23.5**	**9.2**	**10.0**	**48.0**	**22.0**	–				–	–	–

別紙７

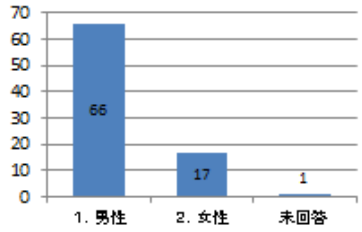
Ｍ３５の「１２人の怒れる男」観察研究の集計結果（２）

区分	グループ決定順位												グループ決定順位の誤差												誤差合計
陪審員	陪審員	銀行員	メッセンジャー	ブローカー	スラム	ベンキ	セールスマン	建築師	老人	ガレージ	時計屋	広告代理店	陪審員	銀行員	メッセンジャー	ブローカー	スラム	ベンキ	セールスマン	建築師	老人	ガレージ	時計屋	広告代理店	
Gr.1	9	6	12	7	2	5	10	1	3	11	4	8	1	1	0	4	1	1	3	0	1	1	0	1	14
Gr.2	7	3	10	11	4	6	9	1	2	12	5	8	1	2	2	0	1	0	2	0	0	2	1	1	12
Gr.3	12	5	10	7	2	8	6	1	3	1	4	9	4	0	2	4	1	2	1	0	1	9	0	0	24
Gr.4	8	2	7	11	3	10	9	1	4	12	6	5	0	3	5	0	0	4	2	0	2	2	2	4	24
Gr.5	8	4	11	7	2	5	10	1	3	12	6	9	0	1	1	4	1	1	3	0	1	2	2	0	16
Gr.6	12	4	10	9	2	7	8	1	3	11	5	6	4	1	2	2	1	1	1	0	1	1	1	3	18
Gr.7	8	3	10	11	7	6	9	1	2	12	4	5	0	2	2	0	4	0	2	0	0	2	0	4	16
Gr.8	7	4	12	10	3	6	9	1	2	11	5	8	1	1	0	1	0	0	2	0	0	1	1	1	8
Gr.9	7	4	9	11	3	6	10	1	2	12	5	8	1	1	3	0	0	0	3	0	0	2	1	1	12
Gr.10	7	5	11	10	3	9	8	1	2	12	4	6	1	0	1	1	0	3	1	0	0	2	0	3	12
最頻値	**7**	**4**	**10**	**11**	**2**	**6**	**9**	**1**	**2**	**12**	**4**	**8**													
最頻値個数	4	4	4	4	4	4	4	10	5	6	4	4													
割合	40%	40%	40%	40%	40%	40%	40%	100%	50%	60%	40%	40%													

別紙８

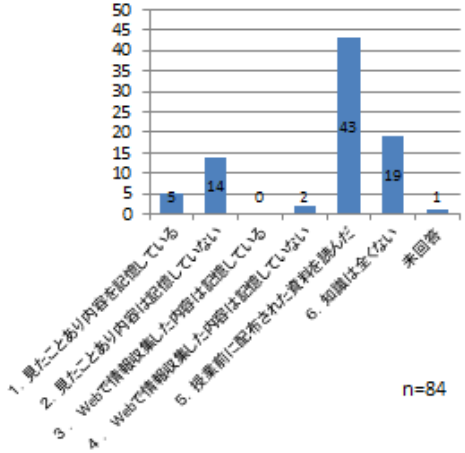
「12人の怒れる男」グループ討議開始前アンケートの集計結果（全体）

Q1.性別



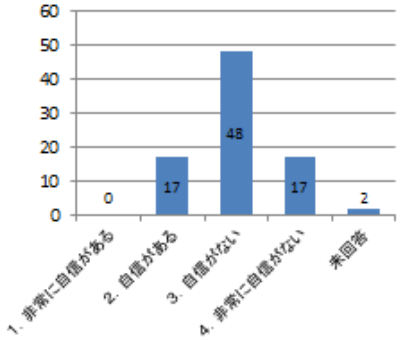
n=84

Q2.「12人の怒れる男」についての事前情報の有無



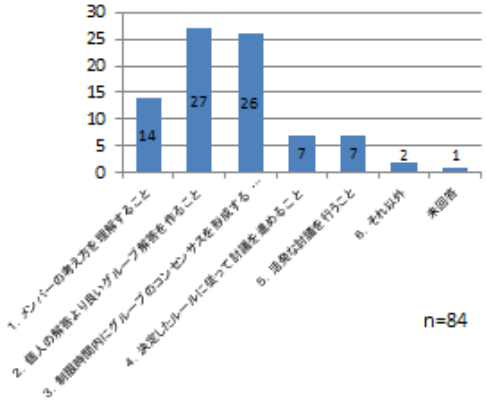
n=84

Q3.あなた個人が導き出した回答に対する自信の度合い



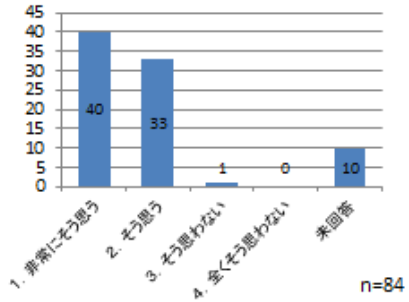
n=84

Q4.このグループ討議において解決すべき課題は何か



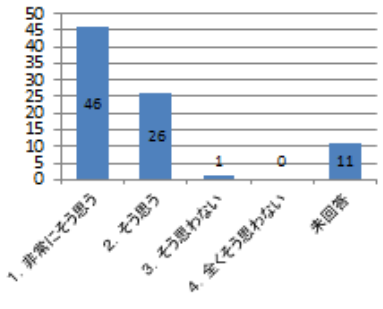
n=84

Q5.グループで討論することが成果物の質を向上させると思うか



n=84

Q6.グループで討論することは、ビジネスを学ぶ上で有効に働くと思うか

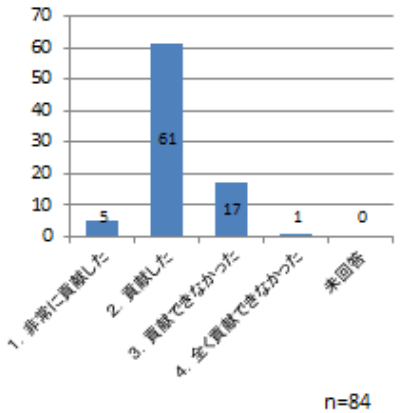


n=84

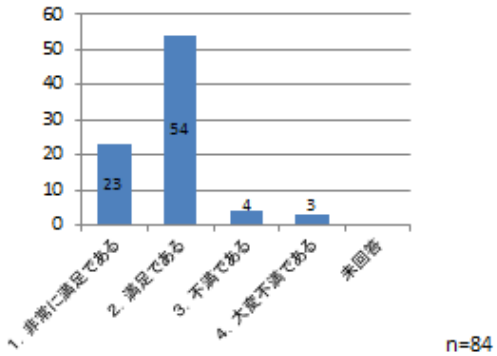
別紙９

「１２人の怒れる男」グループ討議終了後アンケートの集計結果（全体）

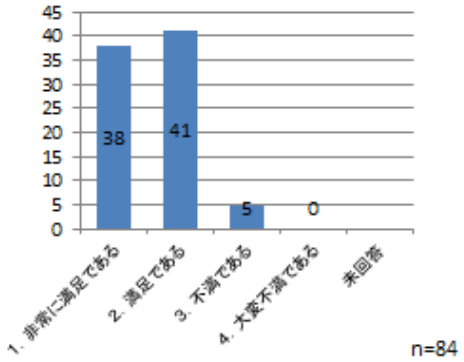
Q1.あなたのグループワークにおける貢献度は



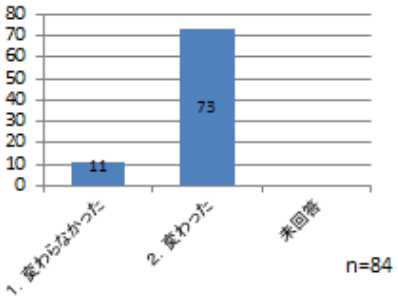
Q2.グループの解答に対する満足度は



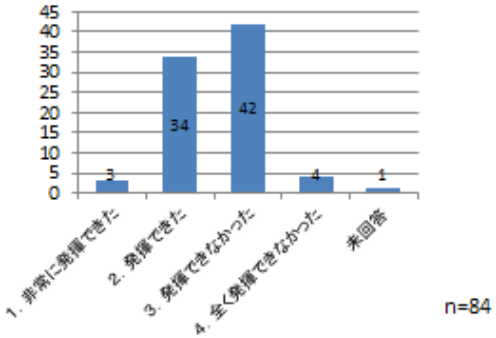
Q3.グループ討議全体に対する満足度



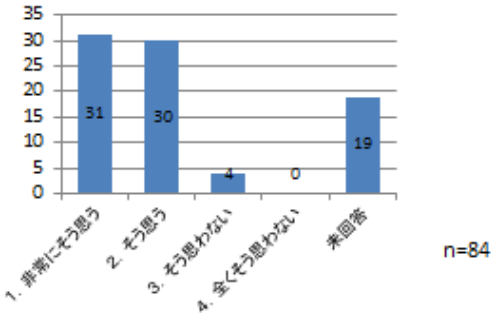
Q4.グループ討議後、あなたの個人の解答が変わったか



Q5.あなたはグループワークでリーダーシップを発揮できたか



Q6.Q5において3又は4と答えた人のみ回答。他者のリーダーシップは有効だったと思うか



別紙１０　　Ｍ３４の「１２人の怒れる男」観察研究の集計結果（１）

G#	個人データ			グループ内の個人データ集計値					グループの決定				改善効果	改善率	改善率順位
	P#	性別	個人誤差	誤差平均 ①	分散 ②	誤差最小値 ③	誤差最大値 ④	誤差中央値 ⑤	誤差 ⑥	誤差最小値 ⑦	誤差最大値 ⑧	誤差中央値 ⑨	⑩-①-⑥	⑪-⑩/①	⑫
1	1A	M	16	21.8	8.9	10.0	36.0	20.0	18				3.8	17.2%	9
	1B	M	24												
	1C	M	10												
	1D	F	16												
	1E	M	16												
	1F	M	32												
	1G	F	36												
	1H	M	24												
2	2A	M	31	20.7	6.4	14.0	31.0	20.0	14				6.7	32.4%	7
	2B	M	14												
	2C	M													
	2D	M	28												
	2E	F	16												
	2F	M	20												
	2G	M	16												
	2H	M	20												
3	3A	M	16	24.0	5.7	16.0	34.0	24.0	24				0.0	0.0%	11
	3B	F	26												
	3C	M	20												
	3D	M	26												
	3E	F	22												
	3F	M	28												
	3G	M	34												
	3H	M	20												
4	4A	M	22	24.0	11.2	12.0	42.0	20.0	12				12.0	50.0%	*2*
	4B	M	34												
	4C	M	18												
	4D	M	16												
	4E	M	34												
	4F	M	42												
	4G	F	12												
	4H	M	14												
5	5A	M	32	21.1	7.6	12.0	32.0	18.5	20				1.1	5.3%	10
	5B	M	16												
	5C	M	16												
	5D	F	16												
	5E	F	24												
	5F	M	32												
	5G	M	21												
	5H	M	12												
6	6A	M	14	23.8	6.1	14.0	34.0	24.0	16				7.8	32.6%	6
	6B	M	26												
	6C	M	18												
	6D	M	24												
	6E	F	34												
	6F	M	24												
	6G	F	28												
	6H	M	22												
7	7A	F		22.9	10.0	12.0	40.0	22.0	16	4	24	15	6.9	30.0%	8
	7B	M	22												
	7C	M	32												
	7D	M	12												
	7E	M	14												
	7F	M	22												
	7G	M	40												
	7H	M	18												
	7I	F													
8	8A	F	20	19.7	4.8	**14.0**	28.0	20.0	**4**				15.7	79.7%	*1*
	8B	F	14												
	8C	M	20												
	8D	M	14												
	8E	M	28												
	8F	M	–												
	8G	M	22												
	8H	M	20												
9	9A	M	26	23.0	9.0	10.0	36.0	25.0	14				9.0	39.1%	5
	9B	F	26												
	9C	M	16												
	9D	M	10												
	9E	F	36												
	9F	M	–												
	9G	M	24												
	9H	M													
10	10A	M	10	15.4	3.2	10.0	18.0	16.0	16				−0.6	−3.7%	12
	10B	M	18												
	10C	M	18												
	10D	M	16												
	10E	F	12												
	10F	M	16												
	10G	F													
	10H	M	18												
11	11A	M	20	24.8	7.5	**16.0**	40.0	23.0	**14**				10.8	43.4%	*3*
	11B	M	20												
	11C	M	40												
	11D	F	24												
	11E	M	26												
	11F	M	30												
	11G	M	16												
	11H	M	22												
12	12A	M	22	23.5	4.0	**18.0**	30.0	23.0	**14**				9.5	40.4%	4
	12B	M	30												
	12C	M	24												
	12D	M	20												
	12E	F	28												
	12F	M	18												
	12G	F	24												
	12H	M	22												
Gr.平均			22.1	**22.0**	**7.0**	**13.2**	**33.4**	**21.3**	**15.2**				**6.9**	**30.6%**	–
M34個人			22.1	**22.1**	**7.4**	**10.0**	**42.0**	**21.5**	–				–	–	–

(誤差:22)事前情報有り

(誤差:12)事前情報有り

(誤差:24)事前情報有り

(誤差:20)事前情報有り

(誤差:14)事前情報有り

価値創造？！
グループ内最小値:14 > グループ決定誤差:4

価値創造？！
グループ内最小値:16 > グループ決定誤差:14

価値創造？！
グループ内最小値:18 > グループ決定誤差:14

別紙１１

Ｍ３４の「１２人の怒れる男」観察研究の集計結果（２）

区分	グループ決定順位												グループ決定順位の誤差												誤差合計
陪審員	陪審員	銀行員	メッセンジャー	ブローカー	スラム	ベンキ	セールスマン	建築師	老人	ガレージ	時計屋	広告代理店	陪審員	銀行員	メッセンジャー	ブローカー	スラム	ベンキ	セールスマン	建築師	老人	ガレージ	時計屋	広告代理店	sum
Gr.1	8	4	11	9	7	5	10	1	2	12	3	6	0	1	1	2	4	1	3	0	0	2	1	3	18
Gr.2	6	4	12	9	2	7	10	1	3	11	5	8	2	1	0	2	1	1	3	0	1	1	1	1	14
Gr.3	9	8	4	11	2	5	10	1	3	12	6	7	1	3	8	0	1	1	3	0	1	2	2	2	24
Gr.4	8	5	12	10	3	4	7	1	2	11	9	6	0	0	0	1	0	2	0	0	0	1	5	3	12
Gr.5	10	3	9	12	2	5	8	1	4	11	7	6	2	2	3	1	1	1	1	0	2	1	3	3	20
Gr.6	11	4	8	10	3	6	9	1	2	12	5	7	3	1	4	1	0	0	2	0	0	2	1	2	16
Gr.7	9	3	10	11	2	5	8	1	4	12	6	7	1	2	2	0	1	1	1	0	2	2	2	2	16
Gr.8	8	4	12	10	3	6	7	1	2	11	5	9	0	1	0	1	0	0	0	0	0	1	1	0	4
Gr.9	8	3	10	11	4	5	9	1	2	12	6	7	0	2	2	0	1	1	2	0	0	2	2	2	14
Gr.10	10	7	6	11	3	5	9	1	2	12	4	8	2	2	6	0	0	1	2	0	0	2	0	1	16
Gr.11	7	6	10	11	3	8	9	1	2	12	4	5	1	1	2	0	0	2	2	0	0	2	0	4	14
Gr.12	7	3	11	9	5	6	10	1	2	12	4	8	1	2	1	2	2	0	3	0	0	2	0	1	14
最頻値	**8**	**4**	**12**	**11**	**3**	**5**	**10**	**1**	**2**	**12**	**5**	**7**													
最頻値個数	4	4	3	5	5	6	4	12	8	8	3	4													
割合	33%	33%	25%	42%	42%	50%	33%	100%	67%	67%	25%	33%													

別紙１２【テープ起こし業者とのやりとりの記録】

●発：村田→着：タイナーズ西山氏

タイナーズ
西山様

お世話になっております。
暑い日が続いておりますが、如何お過ごしでいらっしゃいますか？

先日は、筆耕作業をありがとうございました。

一点、終わった話で恐縮ですが、今回の作業で確認させていただきたいことがございまして連絡申し上げました。

DVD から書き起こしをしていただく際、タイムスタンプを入れていただいておりますが、使用されたソフトウェアは何をお使いになられたでしょうか？

チェックしていたときは文字ばかり気にしておりまして、今週から時間の分析を始めたところ私が Windows Media Player で DVD のタイムスタンプを出して、原稿のタイムスタンプを突合しますと、発話時刻の秒刻みが結構規則性のない誤差となっていて、そのままでは分析に使えない状況となっております。

特に誤差が酷いのはグループ8で、全体に渡って誤差が不規則。
出だしだけ悪くて、途中から合って、後半また誤差が広がるグループ1。

こんな感じで、作業の処理者の問題という気がしなくもないのですが、一応、どのようなアプリを使って作業されたのか、グループによって作業環境が大きく違うのか、もしかしたら私の環境の方が問題があるのか、この辺の分析のため実際の作業環境をお教えいただければと思います。

宜しくお願い致します。

村田

●発：タイナーズ西山氏→着：村田

【村田様へ】ご連絡ありがとうございます。

村田様

こんにちは。お世話になっております。

タイナーズの西山でございます。

ご丁寧なご連絡をいただき、誠にありがとうございます。

お盆を過ぎましたが、まだまだ暑い日が続きそうでございますね。

ご多忙のことと存じますが、どうぞご自愛くださいませ。

さて、ご丁寧なご連絡をいただき、誠にありがとうございます。

私どもで作成させていただいた際は、メディアプレイヤーを用いておりました。

厳密には、映像から音声のみを抽出していったん作成し、その後、その映像にて加筆修正およびタイム

スタンプを施してまいりました。そのいずれ もメディアプレイヤーを用いておりますので、現在、村田様がご使用の環境と同じでございます。

ご連絡をいただきました状況に鑑みますと、思い当たる節がございます。

映像の場合、再生のプレイヤーやPCのスペック(Windows7 か Vista か、core2duo か Celeron かなど)の違いによって、映像と時間表記がずれることが多くございます。

これは、容量が大きいからこそ見られる現象で、とくにお手元の映像はファイル 1 つの容量がかなりの大きさ(1G 超)でございましたので、その影響を多分に受けている可能性があり、おそらく、理由はここにあると存じます。

お手数をお掛けし申し訳ございませんが、なにとぞご確認のほど、よろしくお願いいたします。

ありがとうございます。

--
インタビュー・対談の原稿をご希望なら↓
http://www.tapeokoshi.net/interview.html

翻訳もまた、承っています↓
http://www.tiners-p.com/trans/
==========
「明日テープ起こしを発注したいなあ」
のアステープ起こしタイナーズです。
検索は 「タイナーズ」で！
ご連絡先一覧
WEB: http://www.tapeokoshi.net/
TEL: 073-435-0581
FAX: 073-435-0582
e-mail : master@tapeokoshi.net
営業時間:日、祝除く 10 時～18 時
==========

別紙13　TTM のバグに関する松村教授との対応模様

差出人: "Naohiro Matsumura" <matumura@econ.osaka-u.ac.jp>
宛先: "Ichitaro Murata(KBS_M34)" <ichitaro.murata@a6.keio.jp>
CC: <asarin@kwansei.ac.jp>
件名: Re: （ご報告）TTM のバグと思しき事象につきまして
日時: 2012 年 10 月 21 日 12:58

村田さま

松村です。
貴重なご意見ありがとうございます。
アドバイスに従って TTM のウェブサイトも
MeCab 0.994 へのリンクに張り替えました。

IDM（媒介影響量）や構造化マップについては、
TTM を開発する以前から取り組んでいたので、
Perl や C++や Java や R などの言語
（そのときの気分によって変わってます・・・）
で自分でプログラムを組んで分析してました。
可視化には Graphviz というツールをつかっています。

http://www.graphviz.org

TTM はあくまで前処理だけなので、
その後の処理には別のアプローチを使う必要があります。
統計解析だと Excel や R でいいのですが、
媒介影響量や構造化マップのようなちょっと変わったことをしようとすると
自分でプログラムを組む必要があります。
僕のプログラムはデータに応じてカスタマイズするので
汎用的な形にはなっておらず、すぐには公開できない状況です。

井上先生は大変お世話になっておりますので
ぜひ何とか恩返しがしたいところではありますが、
すぐにはお力になれず申し訳ありません。
どうぞよろしくお伝え下さい。

松村真宏

2012/10/20 Ichitaro Murata(KBS_M34) <ichitaro.murata@a6.keio.jp>:
> 松村先生
>
> ご連絡ありがとうございます。
>
> ちなみにもう一言補足ですが、、、Windows7 の 32bit 環境でもエラーが出てますし、
> １回だけ mecab が停止しただけなので、検証データを持ってない人には気付かない
> かもしれません。ですが、誤差は windows7 の 64bit 環境より小さいですが、完全に
> 一致はしませんでした。
> また、エラー「macab を停止します！」の頻度は繰り返すほど回数が多くなり、検査
> 結果も飛躍的に精度が低下していきました。
>
> 「64bit 環境では」でなくて「windows7 環境では mecab0.98 は使えない」と思った方が
> 良いと思われます。
>
> 繰り返し行って制度が落ちて行った結果 log がありますので、必要ならお申し付け
> ください。・・・といっても、不具合が起こるのが分っているで不要ですね。
>
>
> あと、別件になりますが、厚かましくも可能であればベースでのお願いです。
> 少しアドバイスがいただけたら嬉しいのですが、、、
>
> 全然別の話ですが、先生の研究論文をトレースしておりまして、お送りしました
> テストデータのような議論を「テキストによるコミュニケーションにおける影響の普及モデル」

> での分析や可視化の試み、また「構造化マップ」に表したいと思っております。
>
> 書籍や論文を読んでみたのですが、少し手続きが簡略化されて記載してある？ため
> 媒介影響量の作り方、実際にウインドウを切り分ける手続き、可視化ツールの使い方
> など、なかなか難しくてトレースができないでおります。
>
> もし、先生が学生に指導された際の doc 等で、これらの手ほどきをされておられるような
> ものが閲覧可能であれば、サイト等をお教えいただけませんでしょうか？
>
> 前後致しますが、松村先生のことは指導教授の一人である井上哲浩教授よりご案内
> をいただきました。
>
>　　井上教授紹介サイト
>　　http://www.kbs.keio.ac.jp/faculty/inoue_a.html
>
>
> あくまで可能であればレベルでのお願いですので、NG でしたらその旨おっしゃって下さい。
>
> 以上、
>
> 本当にこのたびはお世話になりました。
>
> ありがとうございました。
>
>
> 村田一太郎
>
>
>
>
>
> -----Original Message----- From: Naohiro Matsumura
> Sent: Sunday, October 21, 2012 11:54 AM
> To: Ichitaro Murata(KBS_M34)
>
> Cc: asarin@kwansei.ac.jp
> Subject: Re: （ご報告）TTM のバグと思しき事象につきまして
>
> 村田さま
>
> 松村です。
> ご連絡ありがとうございます。無事に動いて良かったです。
>
> ところで、こちらで検証した MeCab のバージョンは
> http://mtmr.jp/ttm/
> からリンクを張っているバージョン 0.98 です。
> また、MeCab のバージョンによっては TTM と連携できなかったので、
> バージョン 0.98 を使うことを推奨しています。
>
> ということは、MeCab のバージョン 0.98 が Windows 64bit に
> 対応してないのかもと思ったのですが、やっぱりそのようでした。
>
> http://rmecab.jp/wiki/index.php?plugin=attach&refer=SoftArchive&openfile=README_RMeCab.txt
>
> Windows 64bit 上で TTM を使うときには
> MeCab-0.994 を使うように Web ページを書き換えておきます。
>
> Windows 版も Mac 上で開発しているため、
> 今回の不備にはまったく気づきませんでした。
> 貴重なご指摘をいただき、ありがとうございました。
>
> 松村
>
> 2012/10/20 Ichitaro Murata(KBS_M34) <ichitaro.murata@a6.keio.jp>:
>>

>> 松村先生
>> 三浦先生
>>
>> 早速のご対応深謝申し上げます。
>> 余りにも素早い対応をいただき、感激いたしました。ありがとうございます。
>>
>> さて、本題ですが、、、上手くいきました。
>>
>> 問題は mecab でした。
>>
>> 以下、顛末のご報告です。
>>
>> ※今、別のロケーションで 32bit の win7 環境にて一晩中検証データを作っておりましたが、
>>　途中で眠りに落ちてしまい、このような時間になってしまいました。すみません。
>>
>> 原因を再度調査しようと、一旦、環境を win7 の 32bit 環境に作り直して、検証用データを
>> 入れたところ、正しい結果は表示されませんでしたが、誤差が小さくなりました。
>>
>> 当初、気にしていなかったのですが、mecab が途中で 1 回だけ動作不能になるエラーが発生。
>> そういえば 64bit 版 PC では数回同じようなエラーメッセージが出て都度、問題解決をネットで
>> 検索して対応の方にクリック（windows の処理です）して、TTM の動作を継続していたことを
>> 思い出しました。
>>
>> もう一度やると、mecab の停止エラーが更に頻発し、結果、どんどんキーワードの検出率が
>> 低下していきました。
>>
>> →事実関係：mecab のエラーの都度、検出不能データが増える結果
>>
>>
>> TTM 側に問題があるのではなく、「mecab」になんらかの不具合があるように思われる結果が
>> 出てまいりました。
>> よって、両先生の著作「人文・社会科学のためのテキストマイニング」のなかでご紹介されて
>> いる mecab のダウンロードサイトと違うところでバージョンを変えてみることを考えました。
>>
>> 当初は、本のなかでは確か mecab-0.80 だったかと記憶しておりますが、ご紹介のサイトでは
>> http://mtmr.jp/ttm/より、
>>
>> http://sourceforge.net/projects/mecab/files/mecab-win32/0.98/mecab-0.98.exe/download
>> ということで「mecab-0.98」がダウンロードできたため、てっきり最新バージョンと信じきって
>> おりました。
>>
>> あらためて深夜にググってみましたところ、2012 年度もバージョンアップされておりまして、
>> latest バージョンは mecab-0.994 となっていましたのでこれを入手。
>>
>> ◆正常動作したバージョン「mecab-0.994.exe」
>>
>> http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html#install-windows サイトの
>> Binary package for MS-Windows にありましたダウンロードファイル
>> 「mecab-0.994.exe」に変更して設定変更を行ったところ正常に動作しました。
>>
>> ▼誤動作したバージョン「mecab-0.98.exe」
>> http://mtmr.jp/ttm/より、
>>
>> http://sourceforge.net/projects/mecab/files/mecab-win32/0.98/mecab-0.98.exe/download
>> ※こちらは上記の ver. 0.994 のサイトに更新された方がよいように思われます。
>>
>>
>> これらを 32bit の PC で設定しなおしたところ、お送り頂いたデータと同じものができました。
>> 更に、検証用データを 9 ファイル分作って、チェックしたところ全て excel で検算した結果と
>> 完全に一致いたしました。
>> お送りしたテストデータは１０００レコード程度でしたが、４０００レコード近いものもあり、
>> これも完全一致いたしました。
>>
>> 大変お騒がせいたしましたが、TTM でのエラーなく数値が取れることが判明し、共起確率
>> などの算定にこの環境が使えるのが確認できました。

>>
>> このたびは、本当にお世話になりました。感謝感謝であります。
>>
>> おかげさまで今日は眠れそうです。
>>
>> 末筆ながら、両先生の今後の益々のご活躍を祈念しております。
>> ありがとうございました。
>>
>> 村田一太郎拝
>>
>>
>>
>> 慶應義塾大学大学院 経営管理研究科
>> 髙木晴夫研究室
>>
>> 村田 一太郎
>> e-mail: ichitaro.murata@a6.keio.jp
>> mobile: 090-6942-XXXX
>>
>>
>> -----Original Message----- From: Naohiro Matsumura
>> Sent: Sunday, October 21, 2012 12:46 AM
>> To: Ichitaro Murata(KBS)
>> Cc: asarin@kwansei.ac.jp
>> Subject: Re: （ご報告）TTM のバグと思しき事象につきまして
>>
>>
>> 村田さま
>>
>> 松村です。
>> ご連絡いただき、ありがとうございます。
>> お送りいただいたデータとキーワードファイルを使って
>> 同様の条件で分析したところ、欠損なく抽出されました。
>> 念のため、分析結果のファイルをお送りします。
>>
>> 今のところ村田さまの環境でデータの欠損が起こった原因が分からないのですが、
>> こちらの環境が Win7 32bit を Mac OSX の VMWare 上で動かした環境なので、
>> ひょっとすると Win7 64bit による不具合なのかもしれません。
>> 僕のところには Win7 64bit がなく、確認できずにすみません。
>>
>> もし可能であれば Win7 32bit でご確認いただけますでしょうか。
>> 手数をお掛けいたしますが、よろしくお願い致します。
>>
>> 松村
>>
>> 2012/10/20 Ichitaro Murata(KBS) <ichitaro.murata@a6.keio.jp>:
>>>
>>>
>>> 松村先生
>>> 三浦先生
>>>
>>> はじめまして。
>>>
>>> 私、慶應義塾大学大学院経営管理研究科の修士２年の村田一太郎と申します。
>>> 不躾なメールをお送りしまして申し訳ございません。
>>>
>>> 現在、修士論文作成中で、先生方の執筆された人文・社会科学のためのテキストマイニング
>>> を拝見し、参考にさせていただいております。素晴らしい本をありがとうございます。
>>>
>>> さて、標記でございますが、TTM を利用して会議の分析を実施しておりましたところ、使い方等
>>> が悪いのか、キーワード設定を行い、頻度解析を実施しましたところ、どうもおかしな結果が
>>> 出力されているように感じられ、EXCEL で検算をしてみました。
>>>
>>> 結果はやはり、不具合があるようです。
>>>

>>> 恐れ入りますが、添付の Excel ファイルをご覧願えますでしょうか？
>>>
>>> #1: G1_test_data2.csv・・・TTM に入れた元データになります。処理後が#2 です。
>>> #2: G1_test_data2_ttm6.xlsx・・・これは、TTM で出力された CSV ファイルを excel 形式で保管したものです。
>>> #3: G1_test_data.xlsx・・・これは、TTM で処理する前の元のデータで、キーワードの頻出度を Excel 関数で検算したものです。#2 と比較してくださいませ。
>>> #4: keyword.txt・・・これは TTM に設定しましたキーワードです。
>>>
>>>
>>>
>>> 頭からチェックしたところ、おかしなところがいくつかありました。
>>> サンプルですが、#2 と#3 のファイルでおかしな部分の一部、通番 305~318 までをマーキングしてみました。
>>>
>>> TTM で添付のキーワード設定を行い、キーワードのみを ttm6 にて出力してみましたところ、
>>> エクセルで検索してみた検査結果と異なる結果がでております。
>>> 黄色の反転部分が「本文にキーワードが存在するが、TTM ではカウントされなかった」部分に
>>> なっております。
>>>
>>> 再現性について検証すべく、どういう表示規則があるのか悩んでみたのですが、どうも私に
>>> はわかりませんでした。
>>>
>>> 研究の大詰めにきて、このようなことがわかり、若干途方に暮れているところでありまして、
>>> もし、私の使い方及び環境などに問題があることが分かればとても助かります。
>>>
>>> 現在の環境は、Windows 7 、64bit のパソコンに MS オフィスを使っており、http://mtmr.jp/ttm/より
>>> ダウンロードしたツールを利用して分析を行っております。
>>>
>>> ご多忙のところ恐縮でございますが、何卒、ご見識をお聞かせくださいませ。
>>> 宜しくお願い申し上げます。
>>>
>>> 以上
>>>
>>>
>>>
>>> 慶應義塾大学大学院 経営管理研究科
>>> 髙木晴夫研究室
>>>
>>> 村田 一太郎
>>> e-mail: ichitaro.murata@a6.keio.jp
>>> mobile: 090-6942-XXXX
>>
>>
>> --
>> Naohiro Matsumura, Ph.D.
>> Associate Professor at Osaka University
>> Visiting Scholar at Stanford University
>> matumura@econ.osaka-u.ac.jp
>> http://mtmr.jp
>
> --
> Naohiro Matsumura, Ph.D.
> Associate Professor at Osaka University
> Visiting Scholar at Stanford University
> matumura@econ.osaka-u.ac.jp
> http://mtmr.jp
--
Naohiro Matsumura, Ph.D.
Associate Professor at Osaka University
Visiting Scholar at Stanford University
matumura@econ.osaka-u.ac.jp
http://mtmr.jp

謝辞

この頁を書くことはないのかもしれないと、一体何度思ったことであろう。それ程に苦しんだ研究であった。実際、私の 2012 年はこの研究のために費やしたといっても過言でないように思う。

思えば、髙木晴夫研究室を志望し、本研究に取り組むきっかけとなったのは KBS 入学直後にあったこの「１２人の怒れる男」を我々に指導された髙木教授の授業に心から引きつけられたことによる。20 年前の研究を更に自らの手で進めたい。そう深く心に念じたのを昨日の事のように覚えている。そして今、卒業を目前にその最初の授業の教材をここに握りしめこの頁を記している。非常に感慨深いものがある。そして、ここまで私がたどり着けたのには幾つもの幸運があった。

まず、何をおいても高木晴夫教授の温かくも厳しいご指導なくして今この時はなかったであろう。どこまで研究を進められるかを的確に見極め、迷走する私に何時も的確な道標を示して下さった。心から感謝申し上げる次第である。

次に、今回の研究の肝となるテキストマイニングに関するアドバイス並びにツールでトラブルを抱えた私に TTM を教えて下さった井上哲浩教授と、データ分析に関するヒントと高額なテキスト書き起こしの予算を提供して下さった高橋大志准教授の両名のご指導なくして、やはりこの時を迎えることはできなかったであろう。

髙木ゼミの廣田高敏さん、釜口祥子さん、高橋美寿さん、下村友里さん、田口裕大さん、本当にありがとう。この一年は、どれだけ皆に支えていただいたかわからない。特に M35 期の授業風景の映像撮影に下田まで出向いて下さった髙木教授、釜口祥子さん、下村友里さん、田口裕大さんのご協力なしにはとてもこの大がかりな研究は実現できなかった。

更に忘れてはならないのは、今回の調査を快諾して下さった清水勝彦教授、並びに M35 期の皆さん、追跡調査や不足機材の調達に協力してくれた M34 期の皆さん、KBS 学事の皆さん、本当にありがとう。

大阪大学の松村真宏准教授、並びに関西学院大学の三浦麻子教授には TTM のバグ発見時に大変迅速な対応をいただき、研究者魂これぞという深い感動をいただいた。

SFC 湘南藤沢キャンパスのメディアセンターの方々にも機材貸出で特別なご配慮をいただいた。まだ他にも書ききれていないお世話になった方々がいると思うが何卒ご容赦願いたい。

最後になるが、このような我儘を許してくれた家族、両親に心から感謝を述べるとともに、研究がいくらかでも後進の役に立つことを願って結びの言葉としたい。

2012 年 1 月吉日