

フランス地方議会の審議項目のテキストマイニング分析

増 田 正

Text Mining of Deliberation Matters in a Local Council in France

Tadashi MASUDA

要 旨

本研究では、フランス地方議会の機能を把握するため、テキストマイニングの手法を用いて、審議項目に関する内容分析を行う。テキストマイニングのために活用されるソフトウェアとしては様々なものがあるが、今回はKHコーダーを使用した。

フランスは市町村レベルの基礎自治体数が極端に多いため、ある程度の事例数がなければ、典型的な自治体を選び出すことが難しい。そこで、今回は、都市型自治体の事例として、首都圏のイル・ドゥ・フランス地域圏に含まれ、パリの西方に位置するオー・ドゥ・セヌ県のクルブヴォワ市を選択し、審議項目の分析を進めていく。

審議項目とは、ここでは議事日程としてホームページに公開されている議案の一覧のことを指している。分析対象は、2009年の審議項目である。

本稿では、テキストマイニングによる階層的クラスター分析、多次元尺度構成法、共起ネットワーク分析を通じて、フランス地方議会における審議内容の実態を統計的に明らかにしたい。

キーワード：フランス地方議会、地方分権、審議項目、テキストマイニング

Summary

In this study, I try to make an analysis on issues deliberated at a French local council using text-mining approach to describe the council's functions. I use KH Corder, one of common famous software for text mining.

It is difficult to choose a typical local council in France because there are large numbers of basic autonomies at the municipality level. Therefore, I pick up Courbevoie city in Hauts-de-Seine of Ile-de-France region, which lies to the west of Paris, as a case of typical urban cities and

make an analysis of the issues deliberated there.

The deliberated issues mean lists of the bills open to the public as the agenda to the official website. Deliberations in 2009 are to be analyzed in this paper.

This paper aims to clarify statistically the realities of discussions made in a French local council through the hierarchical cluster analysis, the multidimensional scaling method, and the co-occurrence network analysis by text-mining approach.

Key words : French local council, decentralization, deliberation, text-mining

I 本研究の課題と目的

地方分権一括法の施行以来、我が国では本格的な地方分権・地域主権の時代を迎えている。地方の自律性が高まるとの期待が生まれ、自治基本条例の急速な広がりを顕著な例として、制定地方独自の取り組みも活発化している。

一方、地方議会については、改革の始動が遅れたものの、ここに来て自治法の縛りを乗り越えるために、通年議会制の導入、審議の活性化を狙った一般質問における一問一答式の導入、対面式論壇の設置、ひいてはそれらを包括的に規定した議会基本条例の制定例が急速に広がっている。

それにもかかわらず、地方議会の活動については、議会広報、地方紙・全国紙地方版などの情報を除けば、十全に提供されてきたとはいえない状態にある。また、そもそも地方議会の活動を評価する方法としては、国内での比較が主なものであった。

フランスでも、ミッテラン政権下における1982年の地方分権改革にはじまり、シラク政権以降、2003年には憲法が書き換えられたことで、地方分権化の流れは加速している。日仏両国の分権化の取り組みは、規模も方法論も大きく異なるが、地方分権化によって、財政面にしろ、政策面にしろ、自治体の自律度が高まることが期待されている。

本研究では、一つのフランス地方議会を事例として取り上げ、単年度の全議案の内容分析を通じて、議会審議の実態を明らかにしようとするものである。したがって、研究の第一目的は、フランス地方議会（コミューン＝基礎自治体レベル）の実態を把握することである。

これまで、フランス地方議会に関する研究の蓄積は極めて限定されている⁽¹⁾。我が国では、フランスの地方議会が3カ月に一度集会することさえ、知られているとはいえない。いわんや議会の構成にしても、審議の内容にしても同じである。通常、外国の比較研究も国政に焦点を当てる傾向があり、このことはやむを得ないことである。しかし、ここに来て、世界規模でのIT化の進展が、地方議会の研究を可能とする状況をもたらした。現在、世界各地の地方議会情報は、ネットを通じて入手できる。初めて、これを分析するチャンスがもたらされたのである。

そこで、本研究では、テキストマイニングのソフトウェアであるKHコーダーを活用し、2009

年におけるフランス地方議会の審議項目（議案）を分析する。当然、単一地方議会の単年度分析ということで、扱われている分析量は十分ではないものの、外国地方議会研究の最初のステップとして、将来的な我が国地方議会との比較を念頭に、その基礎資料として本研究結果を提示することにしたい。

Ⅱ 先行研究

（１）テキストマイニング

テキストマイニングは、情報科学技術の進展に伴い日進月歩で進歩している分野ではあるが、現段階での研究業績は限られている。テキストマイニングの手法が積極的に活用されている研究領域としては、情報科学、医療分野がある。

政治・行政に関連する最新の論文としては、矢吹初による「平成の大合併の要因分析」がある。これは、総務省のホームページに掲載されている「市町村合併の取り組み事例」を対象に分析を加えているもので、研究ノートとして公開されている⁽²⁾。

国立情報学研究所（CINII）の論文検索では、テキストマイニングで666件がヒットするが、検索語に政治、行政を加えると1件も引っかからなくなる。そこで、1件1件確認してみる。矢吹論文は、新しい順でトップページに現れる。しかし、その後、なかなか関連する論文は見つからない。

次に田垣正晋「市町村障害者基本計画のニーズ調査の自由回答記述に対するKJ法とテキストマイニングのあり方」が見つかるが、タイトルから個別計画と手法に関するものであることがわかる⁽³⁾。

その後、年次をさかのぼればさかのぼるほど、理系の論文しか存在しなくなる。高齢者福祉や医療分野での業績は散見されるものの、地方自治やガバナンスそのものについての業績は見つけることができなかった。

（２）地方議会研究

今度は、テキストマイニングではなく、地方議会研究で先行研究を探してみたい。同検索では、地方議会(すべて)で2,960件がヒットする。これに議事録の検索語を加えると2件ヒットするが、音声認識システムと実践報告であった。内容分析の語を加えると該当する論文は見つからなかった。

地方議会全般の検索で、乙武北斗他による「地方議会議録における政治的カテゴリの自動推定手法の提案（言語理解とコミュニケーション）」が見つかる⁽⁴⁾。しかし、これは方法論的なものであって、内容分析には踏み込んでいない。同論文はキーワードで「地方議会議録」が示されているが、これを含んでいる論文はこれ以外に存在していない。

地方議会に審議を加えてみると13件がヒットする。しかし、審議会、審議官などの関連のないものも含まれており、該当する事例は一つもないことがわかった。検索語を議案に代えてみても同じことで、議案提出に関するもの2件を含む3件しか見出すことができなかった。

(3) 国会+内容分析

地方議会の業績が限られているため、範囲を国会に広げてみる。国会審議の内容分析にまで踏み込んでいるものとしては4件がヒットし、そのうち3件が該当した。新しい順に、三浦まり「国会の準立法活動—女性労働問題をめぐる国会審議の内容分析」(2003)⁽⁵⁾、坂江千寿子「脳死と臓器移植に関する新聞報道の内容分析：見出し文を構成した言葉を中心に」(1998)⁽⁶⁾、岩井奉信「国会における首相演説の内容分析」(1983)⁽⁷⁾である。しかし、これらは、特定領域に限られるか、こちらの期待するテキストマイニングの手法を使ってはいなかった。国会とテキストマイニングで検索してみると、文書検索システム的な業績が1件検出されたのみであった。

以上、(1)(2)(3)の検索結果を検討してみると、地方議会における審議項目(議案)に関する先行研究は存在せず、本稿の試みがユニークなものであることがはっきりする。なお、フランス地方議会で検索した場合、拙稿⁽¹⁾にヒットするのみである。この業績はPDF化され、ネットで公開されている。

Ⅲ 分析手法

(1) 分析対象

a 自治体

フランス地方議会の審議項目について単語の頻出度という観点から内容分析するにあたり、最初に分析対象を設定する。フランス地方議会といっても、基礎自治体レベルであるコミューン(市町村)に限定したとしても、3万6千を超える数が存在している。なお、広域的レベルには、県(département)と地域圏(région)がある。今回は、基礎自治体レベルを対象とし、パリ近郊の都市型自治体を選択し、事例研究する。事例の選択にあたって、基礎自治体の典型例や模範例を抽出することは最初から意図していない。あえていうならば、首都圏の都市型自治体における地方議会研究ということになる。

クルブヴォワ市の人口は、国立統計研究所(L'INSEE)によれば、84,974人(2007年)であり、面積は4.17㎢しかなく、パリと同程度の超過密型都市である。首都圏のイル・ドゥ・フランス地域圏において、人口規模では、パリを除く自治体の第8位にランクしている。なお、人口5万人以上の自治体は同地域圏で39を数える。本研究の目的は、この自治体の特徴を詳細に記述することではないので、人口と面積以外のデータについては省略する。

同市議会の構成は、人口10万人以下のカテゴリーに属するため、市議会議員は市長を含めて

53人である。議会構成については、別稿に譲りたい。

b 審議項目（議案）

議会の審議項目を内容分析するにあたり、継続的に分析が可能であり、なるべく新しいものを利用することとした。フランス地方議会は、3か月に1回は集会する必要がある、最近では、会議の招集や議案などが、公式ホームページ上で、積極的に公開されるようになってきている。それぞれの自治体の行政能力や情報政策が異なるため、すべての自治体で公開されているわけではないが、少なくとも、情報公開の度合いは高まってきている。

今回、2009年（1月～12月）のクルブヴォワ市議会審議項目（事前に公開された議案, *ordre du jour*）を分析対象とする。単年度としたのは、主として研究予算及び研究期間上の制約によっているが、追跡して分析することが可能である。

審議項目を示すものとして、本研究で分析された、項目として整理された議案のほかに、当然、議事録 (*compte-rendu*) も対象とすることができる。審議結果を反映している議事録は、純粹に審議内容を示すという面では優れていることは疑いが無い。しかし、今回は、議案として項目別に整理されているという点で、ネット上において事前に公開されている議案を活用することにした。

その各年度の状況をみると、公開の形式については、HTML文書を標準とするが、紙媒体をスキャンしてPDF化したものも含まれている。最近では、議事録はPDF、議案はHTMLの形式であることが多い。これらをテキストマイニングソフトを使って内容分析するには、テキストファイルとして形式を統一する必要がある。2009年度の議会は、①2009年3月4日、②3月25日、③3月30日、④6月29日、⑤7月8日、⑥10月12日、⑦12月14日の7回招集されている。なお、③については、紙媒体の資料をスキャンしたPDFであり、他の6回分とは体裁が異なっていた。

それらの審議項目(議案)については、以下の同市の公式ホームページからダウンロードし、ファイルを手にした。

<http://www.ville-courbevoie.fr/maire/conseil/archives-conseil/>

※2007年分からファイルが公開されており、07年には7回分、08年には8回分存在する。

(2) KHコーダーによる解析

テキストマイニング用のソフトウェアには様々なものがあるが、本研究では、KHコーダーを使用する。KHコーダーは、日本語による内容分析ツールとして設計されている。したがって、第一にフランス語を日本語に翻訳したデータを解析することが自然である。フランス語を直接解析するには、日本語を想定している以上、もともとの品詞の分類に難がある。データそれ自体も全角が望ましいとされている。しかし、実際に試してみると、半角のアルファベットであれ、解析には耐えうるということがわかったため、フランス語の文章を直接解析した結果についても、日本語

の結果とともに提示したい。

作業手順としては、フランス語の審議項目（議案）を翻訳し、日本語にする。それをKHコーダーを使って内容分析し、使用頻度の一覧表を作成し、次いで、それぞれの単語間の関係を視覚的にとらえるために、三種類の手法で図示する。次に、フランス語の文書と同じ方法で解析し、日本語版の結果と比較する。

和訳作業が入ることで、用語は必ずしも一対一対応しないことになるが、大まかな傾向を把握することが目的であるため、一つの単語を一つの訳語に固定し、重複を避ける類の用語の調整と振り分けは行わないこととする。

データはテキストファイル化し、改行や空白を修正する。審議項目（議案）ごとに並べ、新しいものは改めて改行する。

IV 分析結果

(1) 単語の使用頻度

a 日本語

図1は頻度7以上の頻出語を抜き出したものである。多いものから順に、料金(23)、協会(22)、予算(21)、協定(21)、認可(18)、会計(18)、公共(17)、変更(16)、締結(16)、サービス(16)、施設(15)と続いている。当然のことながら、金銭関係の頻出語が目立っている。この種のカテゴリーとしては、補助(14)、購入(9)、財産(8)、費用(7)、価格(5)、有料(5)と続いている。また、地方自治体の組織（都市・広域行政を含む）に関するものも目立っている。上記のほかに、地区(12)、地方(10)、市議会(9)等である。また、市町村が日常的に管理

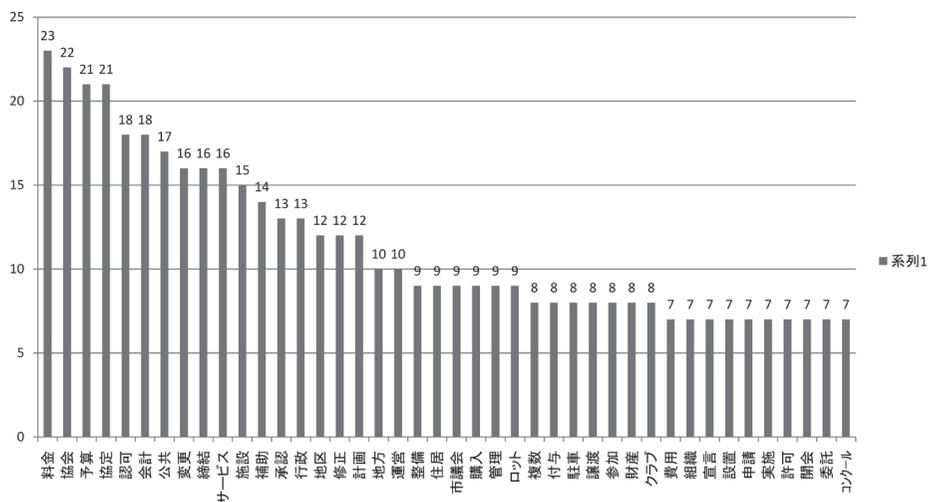


図1 日本語議案における頻出語

する駐車場などの施設・道路に関する行政的手続きに関連した単語が目につく。

b フランス語

図2はフランス語版である。日本語の結果と異なる部分では、ville (25) があるが、自治体(市)そのものを示す名称であるため、和訳には登場していない。tarifs (22), Mise (19), autorisation (17), budget (16), convention (16), premier (15), Courbevoie (14), approbation (13), compte (13), signature (13), compter (12), public (12), rue (11), などと続いている。日本語とフランス語の結果は、統一されていない印象を受ける。例えば、協会は日本語では22回だが、フランス語のassociationは8回しかない。また、compteとcompterが混在したり、複数形と単数形を区別するなど、日本語の内容分析ソフトゆえの限界もある。したがって、フランス語版の分析結果は、あくまで参考として補足的に活用するものにとどめたい。このことは、以下の分析でも共通している。

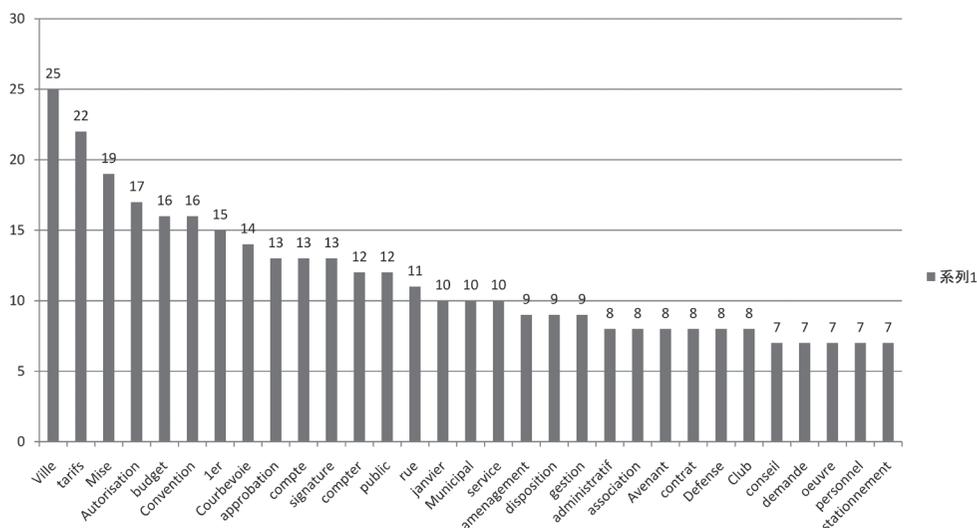


図2 フランス語議案における頻出語

(2) 階層的クラスター分析

a 日本語

図3は、一定数以上の頻出語を8分割したクラスターである。クラスター数は分析者が自由に設定することができる。ここでは、大規模なクラスターの生成を避けるため、各クラスターの構成要素が20以下となるように8分割を選択した。

独立性が高く、少数語で連結している単位としては、①「デファンス」と「地区」、②「認可」と「締結」、③議事運営上の用語である「開会」と「宣言」がある。これらは、相互に関連性が高く、セットで登場していることを示している。

stationnementが近接しており、市の駐車場サービスに関する結びつきであることがわかる。同じく3単語では、rue, acquisition, logementのクラスターがあるが、住居・まちづくり関連の結びつきであるといえる。

4単語のクラスターには、①premier, compter, tarifs, janvier=新年度会計関連、②budget, compte, administrative, approbation= 予算・承認関連がある。③municipal, personnel, mise, disposition=市の人事関連があることが読み取れる。5単語のautorisation, signature, marche, œvre, marchesは、行政手続関連であり、実質的には複数・単数を統合すれば4単語である。最後に、一番大きなクラスターには、15が分類されている。

(3) 多次元尺度構成法 (MDS)

a 日本語

この項目では、多次元尺度構成法を用いた分析結果である図5（2次元）と図6（3次元）を一括して説明する。統計的な説明は省略するが、これによって、主要次元における各単語間の散らばり具合が、プロットの集合体として表現できる。先のクラスターは1次元に再編された関係を示していたが、ここでは2次元（平面図）と3次元（立体図）として、視覚可能な方法で表現する。

図5では、各データが2次元に展開されている。階層的クラスター分析同様に、出現回数の下位の単語は除外している。プロット（点）の位置が互いに近ければ、それらは類似している特徴を有していることを示している。例えば、次元1では、左端の「開会」と右端の「管理」は対照をなしている。次元2では、上端の「譲渡」と下端の「デファンス」や「付与」がそれぞれにあたる。次元1・次元2の両方に対照的な位置関係にあるのは、対角に近いもの同士である。例えば、右上VS左下の関係は、「実施」と「修正」、その逆が「参加」と「許可」である。図3のデンドログラムにおいては、「開会」と「実施」が最も離れた位置に置かれており、これはデンドログラムでも矛盾していないように見えるが、非常に近接していても、プロットの距離が離れていることがある。例えば、

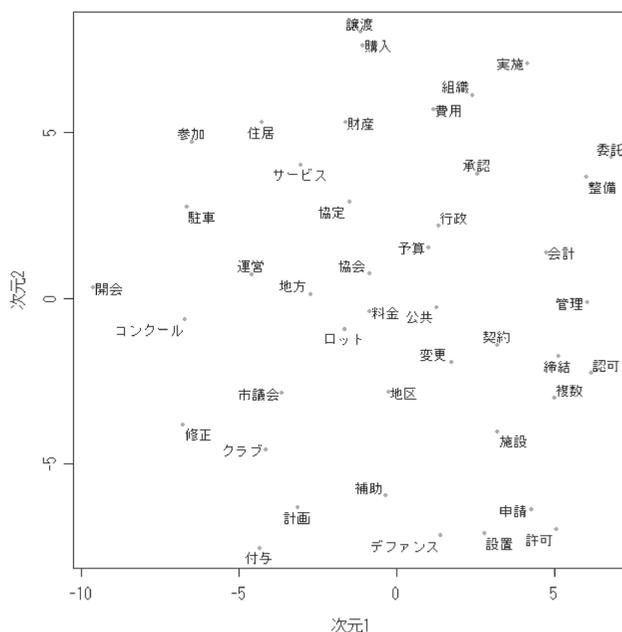


図5 日本語・多次元尺度構成法（2次元）

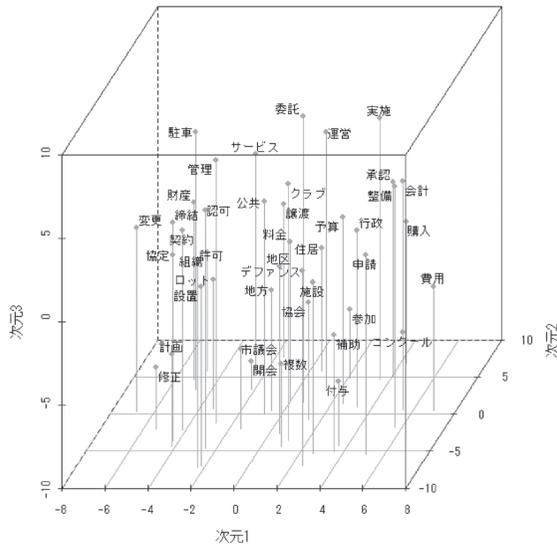


図6 日本語・多次元尺度構成法（3次元）

「ロット」と「譲渡」は次元1ではほとんど同じ位置だが、次元2では上下に離れている。

図6は3次元での展開である。次元3は高さとして表現されている。次元3に着目すれば、「委託」「実施」「駐車」などが上方に、「付与」「開会」などが下方に位置することがわかる。

b フランス語

図7は、日本語の図5に相当する2次元の配置である。解釈の仕方はaと同じである。次元1では、左端のacquisitionと右端のservicesが対照的な位置にある。次元2では、marcheとouvertureが離れている。次元1・次元2を考慮すれば、右上と左下の対角線的な関係がoeuvreとpartenariatであり、左上と右下の対角線的な関係がsociétéとIndemnitéである。図4のデンドログラムにおいては、tarifsとdemandeが最も離れた位置に置かれているが、それは2軸

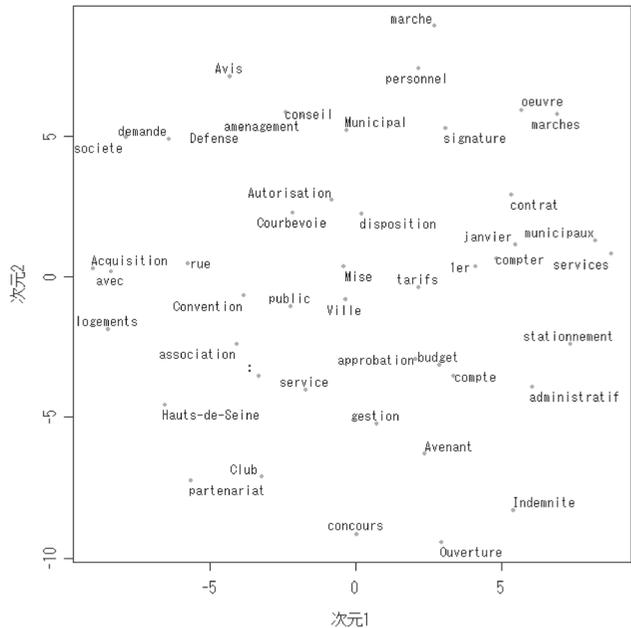


図7 フランス語・多次元尺度構成法（2次元）

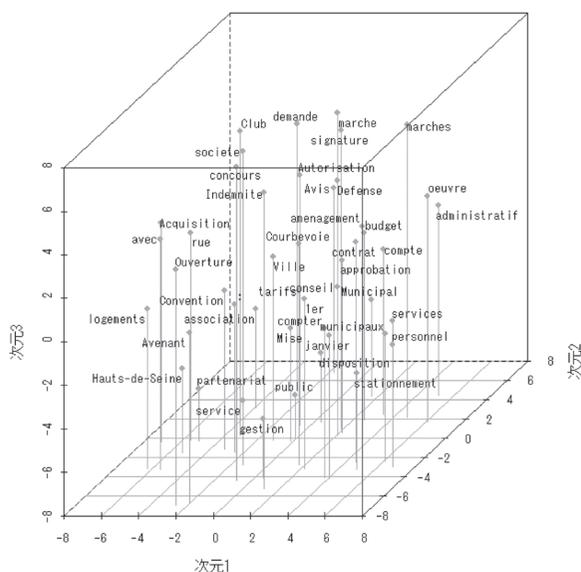


図8 フランス語・多次元尺度構成法（3次元）

の間では次元1による差であることが示唆される。しかし、概ね同一クラスターの構成要素であれば、それほど離れることにはならない。

図8は、日本語の図6に相当する3次元の配置である。前例と同様に次元3に着目すれば、marche, signature, demande, clubなどが上方に、gestion, service, publicなどが下方に位置している。

(4) 共起ネットワーク

a 日本語

図9は単語の総体的な関係を共起ネットワークとして図示したものである。それぞれの構成要素数 (n) とエッジ数 (e) は分析者が設定できる。なお、エッジとは連結線である。それぞれの関係が近ければ近くに置かれ、線で連結される。また、どの単語を媒介してそれぞれの結びついているかが直感的に理解できる。

あまり多数の単語を抜き出してしまうと相互の関係がわかりにくくなる。また、いたずらにエッジを増やしすぎても関係が錯綜しているようにしか見えなくなる。図9は、結局、4つの小グループと1つの大グループとして整理されている。小さなグループについては、とくに説明を要しないが、27単語から構成される大グループには、媒介している中心的な単語が観察できる。例えば、「ロット」にはエッジが5あり、「契約」「変更」「譲渡」「購入」「組織」などをそれ自体が結び付けていることがわかる。ロット (lot) とは購入の単位であり、市民生活に関わる契約に関連したまとめりであることが示唆される。同じように、「委託」やそれに近接する「公共」にはエッ

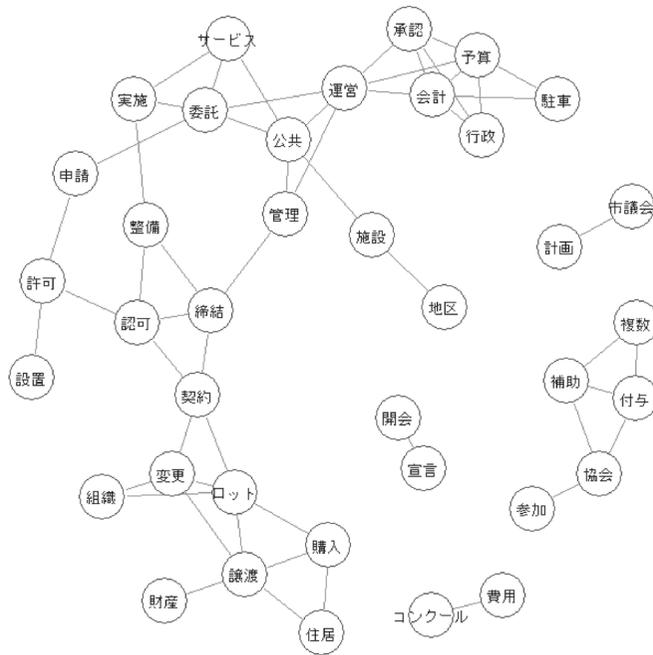


図9 日本語・共起ネットワーク

ジが5つずつある。「委託」は「申請」「実施」「サービス」「公共」「管理」のまとまりを媒介し、「公共」は「管理」「委託」「運営」「サービス」「委託」の中心となっている。これらの意味を読み取ることは容易である。

大きな塊は、①予算・決算関係の承認、②住居・生活関連の購入・譲渡・契約、③公共サービスの委託・実施・管理、④申請・許可・認可等の行政手続関連に概ね大別されるものと考えられる。これに4つの小グループの語彙のまとまりがある。

b フランス語

図10は共起ネットワークのフランス語版である。分析の主眼と意味については、前項と同じである。

図9との違いに着目すれば、5単語以下の小グループ数が6つと多くなり、大グループがやや小さいサイズの中グループ（16単語＋8単語）に分割・再編成されていることである。フランス語版のn（総単語数）が大きいいため、やや込み入っている印象を受ける。小グループについては、とくに媒介する単語があるわけではなく、表現されたままの結びつきを示している。①コンクールの費用、②市の人事、③契約関連、④不動産、⑤許認可、⑥関係団体とのパートナーシップ、である。

図中右端に展開されている8単語の中グループは、aménagementが4つ、avis, conseil,

係を把握することができた。

第3に、多次元尺度法を用いて、2次元と3次元の散布図を作成し、前述のクラスターによって示された距離関係がどの次元での差によって生じているのかを検討した。

第4に、共起ネットワークを通じて、単語のグルーピングと媒介する単語の探索を行った。

研究上の課題としては、事例選択した自治体の分析しかおこなっていないこと、試論として単年度（2009年度）しか分析していないこと、具体的な審議内容（議事録）に踏み込んでいないことなどがあげられる。したがって、研究成果としては必ずしも十分なものであるとはいえないが、まだ緒に就いたばかりの比較地方議会研究の礎石として、分析結果を公表する。

今後、分析対象を拡大し、手法を洗練化されることで、さらなる追加分析を実現させていきたい。

（ますだ ただし・高崎経済大学地域政策学部教授）

註

- (1) 拙稿「フランス地方議会に関する調査報告—Courbevoie市議会を例として—」『地域政策研究』11-1, 2008, 101-103.
- (2) 矢吹初「平成の大合併の要因分析—テキストマイニングによる質的分析—」『青山経済論集』62巻1号, 2010, 107-134.
- (3) 田垣正晋「市町村障害者基本計画のニーズ調査の自由記述回答に対するKJ法とテキストマイニングの併用のあり方」『社会問題研究』58（通号137）, 2009, 71-86.
- (4) 乙武北斗他「地方議会会議録における政治的カテゴリの自動推計手法の提案（言語理解とコミュニケーション）」『電子情報通信学会技術研究報告』2010年7月22日, 110（142）, 7-12
- (5) 三浦まり「国会の準立法活動—女性労働問題をめぐる国会審議の内容分析」『レヴァイアサン』32, 2003, 50-71
- (6) 坂江千寿子「脳死と臓器移植に関する新聞報道の内容分析—見出し文を構成した言葉を中止に」『茨城県立医療大学紀要』3, 1998, 7-15
- (7) 岩井奉信「国会における首相演説の内容分析」『慶応義塾大学新聞研究所年報』20, 1983, 41-67.

付記) 本稿は、平成21年度高崎経済大学特別研究助成金（「地方議会の政策形成に関する比較研究」）による助成を受けている。