

平成21年5月22日現在

研究種目：基盤研究（C）
 研究期間：2006～2008
 課題番号：18520355
 研究課題名（和文） 計量的言語使用研究のためのマルチメディア・コーパスの
 開発に関する基礎的研究
 研究課題名（英文） Development of the Multimedia Corpus for Quantitative Studies
 on Language Use
 研究代表者
 石井 正彦（ISHII MASAHIKO）
 大阪大学・文学研究科・准教授
 研究者番号：10159676

研究成果の概要：国立国語研究所「テレビ放送の語彙調査」のNHK分のデータを用いて、音声文字化テキストと映像・音声とを同期させ、さらに、番組特性や話者属性の情報とも関連させて、単語の検索と同時に、その単語が発話されている映像・音声の再生もできる「テレビ放送のマルチメディア・コーパス」を試作した。いくつかの事例分析により、同コーパスを用いた計量的な言語使用研究の可能性・有効性を確認するとともに、今後の課題を明らかにした。

交付額

(金額単位：円)

	直接経費	間接経費	合計
2006年度	1,800,000	0	1,800,000
2007年度	1,000,000	300,000	1,300,000
2008年度	700,000	210,000	910,000
年度			
年度			
総計	3,500,000	510,000	4,010,000

研究分野：人文学

科研費の分科・細目：言語学・日本語学

キーワード：日本語、コーパス、マルチメディア、言語使用、計量言語学

1. 研究開始当初の背景

(1) 「コーパス」といえば、現在のところ、新聞・雑誌・小説など、印刷媒体を中心とした書きことばのそれが中心であり、「自然な用例」の集積として、文法をはじめとする「言語構造」の研究に利用されることが一般的である。しかし、コーパスを用いた言語研究を、話しことばを基本とする「言語使用」の研究にも拡大・発展させていくためには、言語形式が検索できるというだけではなく、その使用にかかわる各種の情報をも同時に得ることのできるコーパスが必要になる。

(2) ただし、そうした「言語使用にかかわる情報」は、「言語使用」というものをどのように考えるか、そして、そもそも「コーパスを使ってどのような言語使用の研究を行おうとするか」によって、様々に異なってくることが予想される。たとえば、南(1987)が試案としてあげる「談話の十要素」(言語表現そのもの、参加者、話題、コミュニケーションの機能、表現態度、媒体、状況、ネットワーク、文脈、非言語表現)などは、そのような情報の候補になるだろうが、これらの情報をコード化し、コーパスに搭載することは、まだ行われておらず、当然、そうしたコーパ

スを使えばどのような言語使用研究が可能になるかという検討も、すべて今後の課題である。

(3) このように、言語使用研究のためのコーパスがどのような情報を用意すべきかについては、将来の研究にまつところが大きい。ただし、そうした検討には、話しことばの場合、実際の発話場面における映像と音声を参照することが大いに役立つはずである。したがって、当面は、音声文字化テキストから言語形式を検索すると同時に、その発話場面の映像・音声をも参照できる「マルチメディア・コーパス」をつくり、それを使って、どのような「言語使用にかかわる情報」を用意すべきかを検討することが考えられる。

2. 研究の目的

(1) 研究代表者自身が担当者の1人として参加した国立国語研究所「テレビ放送の語彙調査」のデータを利用して、単語の検索と同時に、その単語が発話されている映像・音声の再生もできる「テレビ放送のマルチメディア・コーパス」を試作し、それをういた事例分析を行うことにより、新しい「コーパス言語学的・計量言語学的な言語使用研究」の可能性を探る。

(2) テレビ放送は、言語と映像とによる複合的なメディア（マルチメディア）である。テレビで、音声により発話され、文字により提示される言語は、映像とともにより大きな情報を構成し、送出される。それは、入念に計画・演出された番組の言語においても、生放送の出演者の台本にない発話であっても、本質的に変わりがない。テレビの言語は、どのようなときも、画面に映っている映像との関係の中に存在するのである。したがって、テレビ放送の言語を、映像から切り離し、それ自体で自律的なもの、完結したものとして扱うことは適当でない。それは、語彙調査においても変わらない。

(3) しかし、「テレビ放送の語彙調査」当時は、単位切り・同語異語判別を施した音声文字化テキストを（実際の音声を伴う）映像と同期させることが技術的に困難であったため、番組ジャンルをはじめとする番組特性や、話し手の職業・性別・年齢等の話者属性などとの関係は探ったものの、個々の単語の使用を映像や音声を参照しながら検討するまでには至らなかった。テレビの単語使用を、言語としての側面だけではなく、映像との関係に規定されている側面をも考慮して明らかにしていくことが、課題として残されたのである。

(4) しかし、近年のコンピュータ技術・映像処理技術の進歩によって、パソコン環境でも、言語と映像とを関連づけた処理が可能になってきた。そこで、国語研究所とNHKの許可を得て、「テレビ放送の語彙調査」のNHK総合テレビ・同教育テレビのデータについて、音声文字化テキストと（実際の音声を含む）映像とを同期させ、単語を検索するとその発話場面の映像と音声再生できる「マルチメディア・コーパス」の作成に着手することにした。

3. 研究の方法

(1) 「テレビ放送の語彙調査」は、国語研究所が、テレビ放送における「単語使用」の実態記述を目標に行った（テレビ放送を対象とするものとしては）初めての本格的な語彙調査である。調査対象（母集団）は、1989年4～6月の3か月間に、全国放送網のキー局である6放送局7チャンネル（NHK総合・NHK教育・日本テレビ・TBS・フジテレビ・テレビ朝日・テレビ東京）が放送した、すべての番組（コマーシャルも含む）の語彙であり、標本は、母集団となる放送を5分の幅をもつ抽出単位に分割し、それらを週・曜日・時間帯・チャンネルごとに等しくなるよう構成した集団から、504分の1の比率で無作為に抽出している。5分間の抽出標本の数は全部で364、総時間数は30時間20分。調査単位は、いわゆる「長い単位」の系列（ほぼ文節に相当）だが、より実質的な単語に焦点を当てるために、補助用言や複合辞の類も独立の単語と認めていない。標本延べ語数は、全体で141,975語、番組本編の音声で103,081語であり、同じく異なり語数は、全体で26,033語、番組本編の音声で17,647語である。本研究では、このうち、音声の延べ語数で、NHK総合13,305語、教育13,631語、計26,936語（各52標本）を用いる。

(2) 文字化テキストと映像との同期といっても、上のデータ量で、単語単位で映像と関連づけるような精密な同期は難しい。そこで、ある見出し語を検索し、検索された単位語の一つを指定すると、その単位語を含む「文」が発話されている映像・音声再生される、という程度の「同期」をめざすことにする。

具体的には、まず、標本の5分間の映像を10秒ごとのクリップに分割し、一方で、文字化テキストを文単位に分割して、一つ一つの文に、それが発話として含まれる映像クリップを対応づける（＝その開始時点と終了時点とを入力する）作業を行う。一つの文が複数の映像クリップにまたがる場合は、その文の開始時点を含むクリップと対応づける。文の途中で10秒間の再生が終わってしまっても、

すぐに続行することは簡単だからである。対応づけの作業には、(株)アイ・ビー・イー社製“IBE Outliner v2.2”を購入・利用する。

(3) 次いで、延べ 26,936 の単位語を標本別・発話順に集めたデータベースに、それぞれの単位語（を含む文）に対応する映像クリップの開始時点と終了時点の情報を付与し、単位語の一つを指定すれば、それを含む（文を含む）映像クリップが再生されるようにする。単位語のデータベースはマイクロソフト社製“Microsoft Office Excel 2003”で作成し、映像クリップの再生には、オープンソースのメディアプレーヤー“MPPlayer”（フリーソフトウェア）を用いる。

(4) データベースには、単位語（出現形）のほかに、見出し語（読み・代表形）と語種・品詞情報、話者情報（性別、年齢、職業）、番組情報（ジャンル、チャンネル、曜日、時間帯、長さ、視聴率）を付与し、検索された単位語をこれらの情報でさらに絞り込んでから映像クリップを再生したり、逆に、これらの情報から（異なる見出し語の）単位語を絞り込んでその映像クリップを再生させたりできるようにする。現段階では、このデータベースを「テレビ放送のマルチメディア・コーパス」の試作版とし、固有のインターフェイスは作成しない。

4. 研究成果

(1) この試作版マルチメディア・コーパスを使うと、「どのような話し手が、どのような番組の、どのような場面で、どのような聞き手に向かって、どのような単語を、どのような文脈の中で、どのような音声（口調）で、どのような非言語行動（視線・表情・身振りなど）を伴いながら、発話したのか」といった検討が、そこに収められた実例から効率的に行える（とくに、下線部分の情報は、実際の映像・音声を再生・確認することで、具体的に把握できる）ようになり、テレビ放送における単語使用のより詳しい分析が可能になる。本研究では、いくつかの事例分析を行い、同コーパスを用いることによって可能となる新しい計量的な言語使用研究の可能性を探った。以下に、その一例を報告する。

(2) 思考動詞「思う」が「～と思う」という形式で「一人称主語の思考態度を表明する」際の言語使用

①方法

テレビ放送の語彙調査で、「思う」は、使用度数 633、使用順位 20 位で、動詞では、「する」「なる」「いう」「ある」に次いで、よく使われている。試作版には、そのうちの 163

例（NHK総合 84 例、教育 79 例）が収められている。

まず、見出し語「思う」を検索し、その K W I C を表示させて、引用節「～と」をとり、一人称主語の述語となっているものを絞り込む。これにより、108 例が得られた。

「思う」は、他の思考動詞と同様、スル形式をとって一人称主語の述語となる時、発話行為時現在の話し手の思考活動を表すが、宮崎(2001)によれば、それは「話し手がどのような思考的態度や立場をとるかということ」を他者（聞き手）に向けて表明する」ムード性（態度表明性）を表すことを基本とするという。そして、この「(～と) 思う」の表明する話し手の思考態度や立場は、その引用文によって、まずは、以下のように、「評価的態度」「認識的態度」「行為志向的態度」の 3 種に大別されるとする（例文は、試作版マルチメディア・コーパスから）。

<1> 評価的態度（を表す文）

- a) ダイガクノ ジュケンセード ジタイオ
カエナキヤイケナイト オモイマス。
b) コガイデ ゼンシンウンドーオ スルト
ユー コトワ ヒジョーニ タイセツダ
ト オモイマス。

<2> 認識的態度（を表す文）

- c) ソーユー コトオデスネ ツネニ イシ
キシテイタダケレバデスネ ジコモ ス
コシワ ヘルト オモイマス。
d) タンボノ イネノ クキニ ハイマツワ
リハイマツワリシテ カラミツイテイル
ツルクサ ソーユー オソラク ドーサ
オ マタ エンジタンダロート オモイ
マス。

<3> 行為志向的態度（を表す文）

- e) エー キョーワ マター ウタオ ヒト
ツ ミナサント イッシュヨニ ウタッテ
ミタイト オモイマス。
f) ソレデワ ヒトツネツ アノ コチラデ
ミナサンガタニ サイカクニンオ シテ
イタダキタイト オモイマス。

そこで、K W I C から得られた 108 例の（一人称主語の述語である）「(～と) 思う」を、上記 3 種の態度の下に分類するとともに、それぞれの発話の映像・音声を再生し、まずは、どのような話し手が、どのような聞き手に向かって、いずれの態度を表明しているかを調べた。

「話し手」の分類については、その職業と番組中での役割との二つの観点から行うことができる。役割の方が単語使用に直接的にかかわっているとも考えられるが、たとえば、トーク番組の司会をアナウンサーが務めたときとタレントが務めたとき、あるいは、そのゲストがタレントであるときと専門家で

あるとき、さらには一般人であるときなどでは、単語使用に違いが生じることも想像される。また、アナウンサーは番組によってその話し方を変えるのに対して、タレントはあまり変えない可能性があるとの報告もあり、職業と役割のどちらが単語使用により深く関与しているのか、簡単にはいえない。

ただし、実際の番組を検討すると、アナウンス番組でアナウンスするのはアナウンサーであり、解説・講義番組で講師を務めるのは専門家であるというように、職業と役割とはある程度対応している。そこで、ここでは、分類があまり細かくなならないことと判別のしやすさを理由に、職業において話者を分類することにする。得られたのは、以下の4分類である。

- ・アナウンサー類……アナウンサーのほか、キャスター、レポーター、司会者などを含む。
- ・タレント類……いわゆるテレビ・タレントのほか、俳優、声優、コメディアン、漫才師、落語家、歌手などを含む。
- ・専門家類……大学教授、医師、科学者、評論家、芸術家、小説家、政治家、プロ野球解説者など。
- ・一般人類……テレビに出演することが通常でない一般の人々。

試作版マルチメディア・コーパスにおいて、話者の職業が特定できたものは、26,936語中、25,281語であり、その内訳は、以下のようになる。

職業	延べ語数	比率(%)
アナウンサー類	7830	31.0
タレント類	4464	17.6
専門家類	10587	41.9
一般人類	2400	9.5
計	25281	100.0

これを、「テレビ放送の語彙調査」全体における内訳と比べると、タレント類の発話が少なく、専門家類の発話が多い。これは、NHKの特徴であり、試作版の利用においても注意を要する。

一方、「聞き手」については、ここでは、試作版マルチメディア・コーパスで検索した映像をもとに、「話し手が『(~と)思う』を発話した時点で、その視線を向けている相手」とみなすことにする。テレビ放送では、最終的な聞き手は、もちろん、視聴者である。しかし、実際の発話場面で話し手に同席する者がいる場合には、話し手は、視聴者だけで

なく、同席者をも聞き手として発話することになる。このとき、視聴者も同席者もともに聞き手なのであるが、話し手は、「直接の聞き手」が視聴者であるか同席者であるかによって、言語使用を変えることが多い。そこで、話し手が、視聴者・同席者いずれを直接の聞き手として発話しているのかを、話し手の視線がどちらに向けられているかによって、判別することにしたのである。なお、発話に伴う視線には「説得」などの機能が生じるといわれるが、ここでは、そうした非言語行動としての視線の機能には、立ち入らないことにする。あくまで、直接の聞き手を示す非言語行動としてのみ視線を扱うということである。

②結果

調査の結果は、以下のとおり。表中、“→”の左が話し手、右が直接の聞き手であり、()内の数値は、話し手の視線が映像に映っていないが、話し相手としての聞き手が同定できる場合の数である(外数)。

<1> 評価的態度 (31例)	
専門家類→同席者	10 (6)
一般人類→同席者	9
専門家類→視聴者	3
タレント類→同席者	1
アナウンサー類→視聴者	1
その他	1

<2> 識的態度 (41例)	
専門家類→同席者	13 (5)
一般人類→同席者	7
専門家類→視聴者	6 (5)
アナウンサー類→同席者	3
タレント類→同席者	1
アナウンサー類→視聴者	1

<3> 行為志向的態度 (36例)	
アナウンサー類→視聴者	11
専門家類→視聴者	6 (2)
専門家類→同席者	4 (2)
一般人類→同席者	3
タレント類→同席者	2
タレント類→視聴者	1
その他	5

これをみると、まず、評価的態度を表明する「(~と)思う」は、専門家類や一般人類が同席者に向かって発話した例が多数を占める。認識的態度も、同じように、専門家類や一般人類が同席者に向かって発話した例が多いが、専門家類が視聴者に向かって発話した例も少なくなく、その点で、評価的態度と異なっている。行為志向的態度は、アナウンサー類および専門家類が視聴者に向かっ

て発話した例が多く、前二者の態度ときわだった違いを見せている。このように、テレビ放送における「(〜と) 思う」の使用には、いずれの態度でも専門家類の発話が多く、タレント類の発話が少ないが、評価的・認識的態度の場合は一般人類の発話も多く、また、行為志向的態度の場合はアナウンサー類の発話が最も多くなる、という話し手の違いのほかに、評価的→認識的→行為志向的の順に視聴者に向かっての発話の割合が増えていく、という(直接の)聞き手の違いが見出せる。

③考察

以上のように、試作版マルチメディア・コーパスを使って、一人称主語の述語「(〜と) 思う」が発話される場面の映像を検索すると、評価的・認識的態度では専門家類や一般人類の、行為志向的態度ではアナウンサー類の発話が多く、また、評価的→認識的→行為志向的の順に視聴者に向かっての発話の割合が増えていくという傾向が観察される(とくに、後者の傾向は、映像を参照することによって、はじめて確認できるものである)。これは、テレビ放送で、アナウンサー類・タレント類・専門家類・一般人類が、それぞれ、どのような役割を担っているか、そして、それぞれの役割に照らして、どのような聞き手に向かってどのような発話が許されているか、ということの反映ではないかと考えられる。南(1987)の「談話の十要素」でいえば、参加者と(話し手・聞き手が構成する)ネットワークとに關係するテレビ放送のあり方の反映であろう。

しかし、一方では、話し手の視線の先が同席者なのか視聴者なのか、はっきりせず、どちらともつかないような例も見られる。そのような場合には、話し手の視線が、カメラを見るでもなく、また、同席者を見るでもないといった状態になったり、発話の途中で、カメラから同席者に視線を移したり、といった非言語行動が観察された。また、その口調(パラ言語)においても、視聴者向けなのか同席者向けなのか、あいまいな例も見受けられた。これらは、おそらく、視聴者と同席者と同じように(直接の聞き手として)遇しようとする話し手の意向の現れであり、テレビ放送における話し手の役割と自らの意向との調整活動ともみなせるものである。

いずれにせよ、テレビ放送における「(〜と) 思う」の使用には、話し手が誰を直接の聞き手として発話(態度表明)するかというネットワークの面での、テレビ放送のあり方を反映した明確な傾向差と、それと話し手の意向との微妙な食い違いを反映した中間的な使用とが、ともに見出される。

(3) 今後の課題

以上の結果は、NHK(総合・教育)のみをデータとした試作版マルチメディア・コーパスの偏りを反映したものかもしれない。とくに、専門家類の発話の多さとタレント類の発話の少なさは、NHKにおけるタレント類の発話の少なさによる可能性が高い。今後、他のチャンネルのデータを追加するなどして、このコーパスの規模を拡大し、本格的な「テレビ放送のマルチメディア・コーパス」を構築していくことが課題となる。

また、この分析では、「(〜と) 思う」の音声上の特徴を分析することはできなかった。マルチメディア・コーパスの特徴は、映像だけでなく、音声も再生できるところにあるから、それをういた分析が可能なコーパスにしていくことも今後の課題である。ただし、この方面には、「日本語話し言葉コーパス」(国立国語研究所(2006))というお手本があるので、それを参考にしていきたい。映像面に関しては、もちろん、文字化テキストと映像との同期が粗いという問題がある。これについては、用いるソフトウェアなども検討して、より精度の高い同期をめざす必要がある。

このほか、試作版には、まだまだ不十分な点が多いが、「(〜と) 思う」のわずかな分析からも、映像を考慮した言語使用研究の可能性を確認することができたものと思う。そして、重要なことは、こうしたマルチメディア・コーパスを使って、文字化テキストと映像・音声とを同時に検索し、言語使用の研究に有用な情報を見出していくことである。映像から得られる非言語行動情報、音声から得られるパラ言語情報などは、その候補であり、それらをコード化してコーパスに付与するとともに、それらから(文字化テキストの)言語項目を検索することも考える必要がある。これについては、上の「日本語話し言葉コーパス」のほか、ATRメディア情報科学研究所の「インタラクション・コーパス」(坊農(2008))、フランス語ジェスチャーのデータベース“Geste et Parole”(モンルドン(2001))などが参考になる。

(引用文献)

- 国立国語研究所(2006)『日本語話し言葉コーパスの構築法』国立国語研究所
坊農真弓(2008)『日本語会話における言語・非言語表現の動的構造に関する研究』ひつじ書房
南不二男(1987)「談話行動論」国立国語研究所『談話行動の諸相—座談資料の分析—』三省堂
宮崎和人(2001)「動詞『思う』のモーダルな用法について」『現代日本語研究』8、大阪大学大学院文学研究科日本語学講座
モンルドン、ジャック(2001)「フランス語ジ

エスチャーのデータベース《Geste et Parole》』『日本語とフランス語—音声と非言語行動—』国立国語研究所

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計2件)

- ① 石井正彦「テレビ放送のマルチメディア・コーパス—映像・音声を利用した計量的言語使用研究の可能性—」『阪大日本語研究』21、1～20頁、2009年、査読無
- ② 石井正彦「語彙・計量研究」『日本語科学』24、117～124頁、2008年、査読無

[学会発表] (計1件)

- ① 石井正彦「テレビ放送のマルチメディア・コーパス」計量国語学会、2008年9月20日、武庫川女子大学

6. 研究組織

(1) 研究代表者

石井 正彦 (ISHII MASAHIKO)
大阪大学・文学研究科・准教授
研究者番号：10159676

(2) 研究分担者

なし

(3) 連携研究者

なし