

◎籠宮隆之(国語研) 山住賢司, △榎洋一(都立大/国語研) 前川喜久雄(国語研)

はじめに 国立国語研究所・通信総合研究所・東京工業大学では、1999年度より開放的融合研究「話し言葉の言語的・パラ言語的構造の解明に基づく『話し言葉工学』の構築」¹⁾によって『日本語話し言葉コーパス』²⁾を作成し、話し言葉に関する様々な研究を行ってきた。これまでに、収録時に収録作業者が評定した予備的な印象評定値を分析し、印象評定に関わる講演音声の特徴について報告してきた^{3) 4)}。

今回、新たに講演の印象を評定するための尺度を作成し⁵⁾、実際に『日本語話し言葉コーパス』の評定を行った。本発表では、そのデータを分析し、印象評定に関わる主要因の解析結果を報告する。

評定項目 以下の5下位尺度20項目よりなる、講演の印象を評定するための尺度⁵⁾。

好悪	好きな—嫌いな 心地よい—不快な 感じの良い—感じの悪い 親しみやすい—親しみにくい
上手さ	流暢な—たどたどしい 話し慣れた—話し慣れていない なめらかな—しどろもどろな 上手い—下手な
速さ感	速い—遅い スピード感のある—ゆったりした せわしげな—のんきな 落ち着きのない—落ち着きのある
活動性	声の大きい—声の小さい 力強い—弱々しい 元気のある—元気がない 積極的な—消極的な
スタイル	礼儀正しい—無礼な まじめな—ふまじめな 丁寧な—ぞんざいな 上品な—下品な

対象音声 『日本語話し言葉コーパス』のうち、「コア」²⁾と呼ばれるデータセットに対して印象評定を行った。コアを対象とするのは、1) コアには分節ラベルや韻律ラベルが付与される²⁾ため、印象評定結果と音声特徴の比較を行いやすい。2) また、分節ラベルや韻律ラベルに関わる種々の変異現象の解析に印象評定結果が必要となる。——などの理由による。コアには対話音声や朗読音声も含まれるが、印象評定には独話の講演(187講演)を対象とした。

講演の時間的推移に伴う印象の変化を評定するには、講演を分割する必要がある。また、各講演は短いものでも8分以上あり、講演全体に対して印象評定を行なうのは多大な負担がかかる。そこで、以下の方法で各講演から3箇所(3箇所の聴取単位)を切り出した。

- 講演のを、「講演の冒頭」「講演の中盤」「講演の終盤」に分ける。
- この中から意味上のまとまりを持ち、文末表現で終わる箇所を選ぶ。
- この際、200ms以上のポーズまたは明確な文末表現で区切られた発話単位²⁾で切る。
- 「これから○○というタイトルで発表します」などのようなメタ的表現を述べている箇所は外す。
- 10秒以上のポーズを含む箇所は対象としない。
- 1つの聴取単位が1分前後の長さになるようにする。

また、評定者が同一の音声に対し一貫した評定を行えるかを確認するため、同一の聴取単位を2回評定させるものを、コア以外の『日本語話し言葉コーパス』から含めた。

評定者 人材派遣会社を通じて募集した、20代男性・20代女性・50代男性・50代女性それぞれ5名ずつ、計20名。過去に音声学・言語学・心理学に関する仕事に従事したことのないものを募集した。

刺激呈示・回答方法 刺激呈示には計算機を用いた。評定者は計算機のサウンドボードよりヘッドホンを通じて音声を聴取した。刺激音声はランダムに呈示した。

回答も計算機上のプログラムにより行った。刺激音声が終わるとともに評定作業を行なうようにした。評定語もランダムに呈示するようにした。

分析 以上の手続きにより、収集した印象評定データの分析を行う。ただし、今回は評定の一貫性確認のために用いた聴取単位の相関係数が0.5以上と比較的良好であった8名のデータを用いる。

評定データの扱い 各項目につき7段階で評定させた結果に対し、1~7点の得点を与えた。そして、下位尺度に属する項目の得点合計を下位尺度の得点とした。各下位尺度は4項目からなるので、各下位尺度につき4~28点の得点が得られる。今回の分析では、上記の得点の評定者8名の平均を、各聴取単位に対する得点として扱う。

講演特徴との関係 どのような講演の特徴が印象評定に寄与しているかを分析する。今回の分析では、書き起こしテキスト²⁾より得られる以下の項目を講演特徴として扱う。

- モーラ/秒 — 聴取単位中の総モーラ数 / 聴取単位の発話単位時間長の総計
- モーラ/秒のゆれ — 「モーラ/秒」の分散
- 文節/秒 — 聴取単位中の文節数の総計 / 聴取単位の時間
- ポーズ比 — 聴取単位の発話単位中の総計 / 聴取単位の時間
- ポーズ/秒 — 聴取単位中のポーズ数 / 聴取単位の時間
- フィラー/秒 — 聴取単位中のフィラーの総計 / 聴取単位の時間
- 語断片/秒 — 聴取単位中の語断片の総計 / 聴取単位の時間
- 笑いながら/秒 — 聴取単位中の笑いながらの発話の総計 / 聴取単位の時間
- 笑い/秒 — 聴取単位中の単独の笑いの総計 / 聴取単位の時間

これらの講演特徴を説明変数、印象評定結果得られたそれぞれの下位尺度を目的変数とした重回帰分析を行った。幾つかのモデルを構築し、AIC(赤池情報量基準)や多重共線性の問題を考慮しながら最適なモデルを選択した。この結果、「上手さ」「速さ感」「活動性」に関して比較的良好な結果が得られた(表1)。

この結果から、1)「活動性」および「上手さ」には「ポーズ比」が大きな役割を果たしている。2)「速さ感」には「モーラ/秒」と「ポーズ比」が大きな役割を果た

* Scaling of "Tone of Voice": Method and Results. By KAGOMIYA Takayuki(NIJLA), YAMASUMI Kenji, MAKI Yohich(Tokyo Metropolitan Univ. / NIJLA), MAEKAWA Kikuo(NIJLA)

表1: 各下位尺度を目的変数とした重回帰分析

「上手さ」	p 値	標準化偏回帰係数	寄与率
ポーズ比	$p < 0.001$	-0.611	0.390
言い淀み/秒	$p < 0.001$	-0.280	0.064
モーラ/秒	$p < 0.001$	0.141	0.045
自由度調整済み決定係数 0.496, $p < 0.001$			
「速さ感」	p 値	標準化偏回帰係数	寄与率
モーラ/秒	$p < 0.001$	0.578	0.412
ポーズ比	$p < 0.001$	-0.380	0.223
笑い/秒	$p < 0.001$	-0.104	0.016
言い淀み/秒	$p < 0.001$	0.072	0.009
自由度調整済み決定係数 0.657, $p < 0.001$			
「活動性」	p 値	標準化偏回帰係数	寄与率
ポーズ比	$p < 0.001$	-0.620	0.335
フィラー/秒	$p < 0.001$	-0.319	0.029
文節/秒	$p < 0.001$	0.214	0.023
言い淀み/秒	$p < 0.001$	-0.129	0.015
自由度調整済み決定係数 0.398, $p < 0.001$			

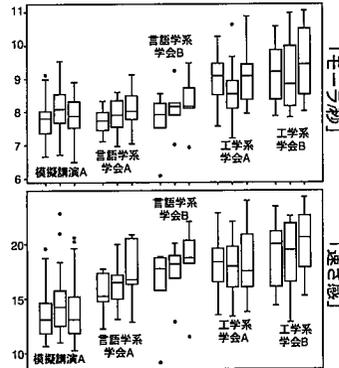


図1: 講演種別の「モーラ/秒」と「速さ感」の対比

している。— ことが分かる。これは、講演の収録時に収集した印象評定の分析³⁾と同様の結果を示している。

「速さ感」に関しては、講演の時間的推移に伴う発話速度の違いも反映されている。図1に、講演種別に、「モーラ/秒」および「速さ感」の得点を掲げる。各講演種別に左より「講演冒頭」「講演中程」「講演終盤」の順に並べる。この図から、講演の時間的推移に伴う「モーラ/秒」の違いに「速さ感」が対応していることが分かる。

重回帰分析で良好な結果が出なかったもののうち、「スタイル」に関して講演種別で差が見られた。図2に、講演種別の「スタイル」の得点を掲げる。模擬講演と学会では、模擬講演の方が「スタイル」の得点が低い。これは、模擬講演の収録の際に講演者をできるだけリラックスさせるようにしたことを反映していると考えられる。特に模擬講演Dの得点が低いのは、模擬講演Dの講演者は学生・大学院生が中心であり、また模擬講演の収録スタッフも20代の者が中心のため、よりリラックスして収録が行われたためであると考えられる。

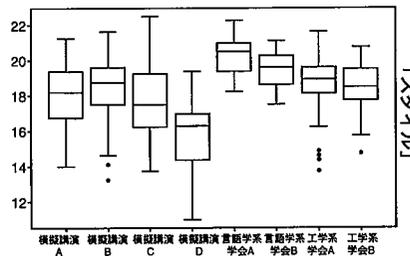


図2: 講演種別による「スタイル」得点の変化

各印象項目間の関係 講演の収録時に評定した印象評定の分析では、講演の特徴と印象評定項目の他に、それぞれの印象評定項目間での間にもさまざまな因果関係が見られた⁴⁾。今回収集した各下位尺度間の相関係数を表2に掲げる。「好悪」と「上手さ」との間や、「上手さ」と「活動性」との間、「速さ感」と「活動性」との間に、強い相関関係が見られる。

また、先行研究⁴⁾では、「発話速度感」が速くなれば「高評価」になるが、ある限界を超えると「低評価」になることを明らかにした。今回のデータにおける「速さ感」とその他の尺度との関係を示す散布図を図に掲げる。各図ともにx軸が「速さ感」であり、y軸がそれぞれの下位尺度である。「速さ感」に対するそれ以外の下位尺度の回帰を、1次式から3次式まで計算し、それぞれの回帰式に対してAICを計算したところ、2次式が最もAICが低かった。よって、図には1次式と2次式による回帰式を掲げた。

この図から、どの下位尺度も「速さ感」が中間(16)付近までは上昇するが、それを境に上昇が止まる、もしくは下降していることが分かる。これは先行研究⁴⁾の結果、および内田の研究⁶⁾での単文による合成音声を用いた結果に類似した結果を示している。

今後の課題 1) 分析のパラメータに韻律ラベルを用いる。2) 印象評定項目間の関係について因果関係を探る。— など、更に研究を進めたい。

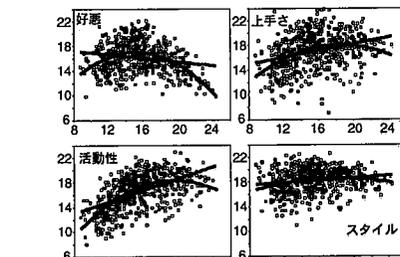


図3: 「速さ感」(横軸) とその他の下位尺度 (縦軸) との相関

表2: 各下位尺度間の相関係数

	Fac. 1	Fac. 2	Fac. 3	Fac. 4	Fac. 5
Fac.1 (好悪)	1.000				
Fac.2 (上手さ)	0.656	1.000			
Fac.3 (速さ感)	-0.194	0.297	1.000		
Fac.4 (活動性)	0.444	0.668	0.528	1.000	
Fac.5 (スタイル)	0.419	0.411	0.157	0.163	1.000

参考文献

- 1) 古井, 前川, 井佐原: “科学技術振興調整費開放的融合研究推進制度— 大規模コーパスに基づく「話し言葉工学」の構築 —”, 日本音響学会誌, 56, 11 (2000).
- 2) 前川, 籠宮, 小磯, 小椋, 菊池: “日本語話し言葉コーパスの設計”, 音声研究, 4, 2 (2000).
- 3) 籠宮, 山住, 旗, 前川: “発話速度の違いが印象評定に及ぼす影響”, 日本音響学会 2002 年春季研究発表会講演論文集 (2002).
- 4) 山住, 籠宮, 旗, 前川: “自発音声コーパスにおける講演の音声特徴と印象との関係について”, 人工知能学会研究会資料 SIG-SLUD-A202 (2002).
- 5) 山住, 籠宮, 旗, 前川: “講演音声の特徴を捉える評定尺度の構築”, 日本音響学会 2003 年秋季研究発表会講演論文集 (2003).
- 6) 内田: “音声の発話速度の制御がピッチ感及び話者の性格印象に与える影響”, 日本音響学会誌, 56, 6 (2000).