関西大学『社会学部紀要』第33巻第2号，2002，pp.325－373　　ISSN 0287－6817

はいそうですね
―音声の印象の分析のこころみ―

雨宮俊彦・水谷聡秀

"Hai Sou desu ne" : An analysis of voice impressions.

Toshihiko AMEMIYA, Satohide MIZUTANI

Abstract

A student actress (age 22) was asked to speak "Hai Sou desu ne"(Yes, it is.) in several imagined situations. Situations consisted of 11 emotion related scenarios, i.e., no specific emotional response, joy, sadness, anger, cheerful excitement, anxiety, impatience, puzzlement, determined, active and passive. The impressions 11 of tape recorded voices were rated using 28 adjective pairs by 70 students. Results of factor analysis show two major factors, i.e., confidence (potency), pleasure-displeasure (evaluation) and two minor factors, i.e., lightness (activity1), calmness (activity2). In the second part of the paper, correlations between voice impressions and acoustic characteristics of the voices were explored. The 11 voices were digitized and spectral, pitch and pressure analysis done. The pitch values, pressure values and times at phonemic boundaries were measured. Correlations were found between confidence factor and speech speed and pitch changes, between the pleasure-displeasure factor and pressure changes and pitch changes, and between the lightness factor and pitch variations and pressure variations. These correlations were found in specific verbal contents.

Key words : Vocal Emotion Expressions, Voice impressions, Semantic Differential, Confidence, Potency, Evaluation, Activity, Acoustic Analysis, Spectrogram, Pitch, Voice Pressure, Speech Speed, Speech Pause, Speech Rhythm

抄　録

役者志願の22才の女性に、さまざまな状況を想定して「はいそうですね」と発話してもらったものをテープで録音した。想定した状況は、関連する感情でいうと、特定の感情表現なし、喜び、悲しみ、怒り、楽しさ、不安、焦り、戸惑い、断固とした、積極的な、消極的なの11種類である。このテープを70名の大学生にきかせて、その印象を28の形容詞対をつかって5段階で評定してもらった。印象評定の結果を因子分析すると、自信（力量）、友好（評価）のふたつのおおきな因子と、軽快さ（活動性1）、落ち着き（活動性2）のちいさなふたつの因子がえられた。つぎに、11の音声をデジタイズし、スペクトログラム、ピッチ、音圧の時間経過を表示した。そして、音圧、ピッチ、時刻の値を、発話における音素に対応するローマ字表記の区切りで計測し、拍や、単語、文節、文における計測値の平均や変動量などの指標もふくめて、音声の印象の因子得点との相関をもとめた。自信因子と発話速度、ピッチの変化、友好因子と音圧の変化、ピッチの変化、軽快さとピッチ、音圧の変動について、一定の関連がみられた。これらの相関は、「はい」、「ね」などの特定の言語内容についてのみ見い出された。

キーワード：感情の音声表現、声の印象、SD法、自信、力量、評価、活動性、音声分析、スペクトログラム、ピッチ、音圧、発話速度、発話の間、発話のリズム

本論文でもちいた音声資料は神田真澄・五味聖子（1999）、印象評価データは向井秀一・山本雅規（2000）によるものである。また、本研究の雨宮俊彦の部分は、平成13年度の関西大学国内研究員研究費によって行った。

はじめに

すぐれた臨床家であったハリー・スタック・サリヴァンは、みずからの面接技法をのべた「精神医学的面接」(Sullivan, H, S. 1954) の冒頭で、精神医学的面接が音声的コミュニケーションであることを強調している。

「たいていの人には自分がほんとうはそうだということに気づいていないのだが、実は、人間が心底からほんとうに言いたいことの手がかりはたいてい耳経由で届くものだ。……精神医学的面接とはなにより音声的コミュニケーションの問題である。コミュニケーションとはなによりもまず言語的だというおもいこみはきわめて重大な誤りではないだろうか。述べられた命題文のほんとうのところが何であるかをおしえるのは、言語にともなう音である。……面接中に、「実は結婚しています」と言われることもあるだろう。――「むろん、とてもしあわせだね? 」。答えが「はい」だったとする。しかし、この「はい」にはありとあらゆる含みがありうる。葬送行進曲もあればアポロン賛歌もある。この「はい」が実はそれとなくの「いいえ」ということだってありうるのだし、「はい」と「いいえ」の中間のどの位置でもありうる。」(Sullivan, H, S. 1954, pp. 24-26.)

サリヴァンが指摘するように、声の調子は、はなしの内容以上に、話し手の感情状態を的確につたえる。声の調子への敏感さには、かなりの個人差があるらしい。サリヴァンは、分裂病の研究でとくにすぐれた業績をのこした精神医学者だったが、分裂病では、日大グループが音調テストで、分裂病者の家族が声の調子のつたえる感情をどの程度敏感に把握できるかを調査している (井村1984)。調査の結果、分裂病者の家族には音調テストの成績がわるい場合がおおいことがしめされた。これは、分裂病の病因論としてのダブルバインド論を直接うらづけるものではないにせよ、分裂病にかかわるコミュニケーションのゆがみが、話しの内容と声の調子のずれに関連していることは示唆している。いずれにせよ、臨床心理学の分野では、面接における声の調子、心の失調と関連したコミュニケーションのゆがみにおける声の調子など、声の調子が無視できない役割をはたしている。

西欧の古典レトリックは、議論法から、修辞法、文体論、作文技術、記憶術、ノンバーバル・コミュニケーションまでふくんだ総合的な言語技術だが、古典レトリックにおいても、キケロやクインティリアヌスなどのレトリックの達人や教師たちは、声の調子が聴衆への、はなしの印象をつくるうえで重要であることを指摘している。レトリックにおける、ノンバーバル・コミュニケーションの研究をひきついだ応用分野が、パフォーマンス学で

ある。パフォーマンス学でも、表情や空間的配置などとならんで、声の調子が、話し手の感情や人柄をつたえるうえで重要であることが指摘されている（佐藤1995）。

インタフェースでは、80年代から90年代にかけて、これまでの文字中心のインタフェースにたいして、グラフィックスを活用したGUI（Graphical User Interface）が開発の中心をしめてきた。最近になって、インタフェースにたいするユーザーの印象を形成するうえでは、グラフィックスの質をさらにたかめることよりも、音の質や、音声の上手な利用の方が有効であることなど、音声の重要性が指摘されるようになってきた（Reeves, B. and Nass, C. 1996）。とくに、エージェントインタフェースでは、表情や声の調子による感情の伝達が重要な役割をはたしており、ペットロボットなどでは、声のピッチの変化などによって、感情状態をつたえるこころみが試行的になされている。

以上のように、コミュニケーションにおける声の調子の重要性は、臨床心理学、レトリック、インタフェースなど、さまざまな応用分野で、認識されている。

言葉の内容は、基本的には、意識的に、選択され、くみたてられ、顔の表情もある程度までは、意図的に制御することはできるが、「感情が生み出す声の変化は、簡単に隠蔽できるものではない。」（Ekamn, P. 1985, p.88.）と指摘されているように、声の調子の意識期な制御は、よりむつかしい。恐怖や喜び、悲しみ、怒りなどの基本的な感情については、声の調子のみから、比較的高い確率で感情を判断できることもしめされている（Banse, R. and Scherer, K, R. 1996）。声の調子は、感情とともに、発声器官の緊張や唾液の分泌、生理的、心理的なリズムが変化し、これらが音声としてあらわれるもので、声の属性と感情状態の間には、組織的な対応があることが期待される。

声の調子による感情の表出は、コミュニケーションにおいて重要な役割をはたしており、そのしくみには基本的な規則性が期待できる。しかし、声の調子と感情の表出の関連については、基礎的な研究は非常にすくない。声の調子と感情の対応について、もっとも、組織的な研究をしているのは、感情の構成要素説をとなえているScherer, K, Rたちのグループ（Scherer, K, R. 1986, Banse, R. and Scherer, K, R. 1996）である。しかし、感情の構成要素説自体が感情の理論としては、また少数派の新参の理論（遠藤2000）だし、声の属性のパラメーターについても、Banse, R. and Scherer, K, R.（1996）の研究は、ピッチや音圧、速さなどについては、一定の傾向は確認できたが、傾向の確認程度だし、ピッチの輪郭、リズム、声の質などについては、分析しあぐねているといったかんじで、傾向性の把握もまだできていない。インタフェース研究の領域では、音声の属性と感情には、こんな関連があるといったまとめがなされる（Murray,I,R. and Arnott,J,L. 1993、Picard,R,W. 1997）が、

経験的な知見のとりあえずの集約の段階である。

声の調子と感情の関連の研究がおくれているのは、ふたつの困難があるからである。

ひとつは、音声の属性把握のむつかしさである。表情などの視覚的な刺激なら写真にとったりスケッチしたりして詳細な分析ができる。音声の場合は、耳で聞いただけでは、印象のおおざっぱな言語化しかできない。

一般の研究者が、コンピュータをつかって音声を数値化し、ピッチやフォルマント、音圧、時間経過などを詳細に分析できるようになったのは、最近のことである。コンピュータによる音の数値化による分析が普及するまでは、周波数の分析には、サウンド・スペクトログラフなどの高価な専門的な音声分析装置が必要だった。また、城生（2001）で紹介されているように、[ma]と[na]における子音の[m]と[n]をいれかえるMalécotの1956の実験などは、録音テープの試行錯誤的なはりかえといった大変な作業を必要とし、それだけで、一本の論文となっている。（この実験では、母音の[a]の部分が子音によって影響をうけて、それぞれ[^{m}a]、[^{n}a]となるため、子音部分を、はりかえると[n^{m}a]、[m^{n}a]となる。きこえかたは、[n^{m}a]は[ma]として、[m^{n}a]は[na]としてきこえる。鼻音の子音である[m]と[n]の場合、子音から母音へのわたりの影響で、母音部分がかわり、きこえはこの母音部分によってきまる。）

2．で具体的にのべるが、最近の音声分析ソフトをもちいれば、ピッチ（F0）、フォルマント周波数（F1、F2など）、音圧などを表示し、分析することができる。以下では、ピッチ（F0）、フォルマント周波数などの音声の属性がどんなものかを、佐藤（2001）にもとづき、日本語の場合について、ごく簡単にのべる。

言葉の音には、有声音と無声音がある。有声音は声帯の振動による音であり、無声音は声帯の振動をともなわない音である。有声音と無声音のちがいは、のどに指をあてて声をだし、振動するか否かでわかる。有声音の代表は、[a][i][u][e][o]の母音である。子音には、有声音と無声音とがある。

まず子音から説明する。[p][t][k]などは、息がそれぞれ唇や歯茎、軟口蓋などの調音点で、せきとめられ一気に放出されることよよってだされる音で、破裂音といわれる。破裂音は、ごく短いスパイクフィルといわれる音である。これは、楽音ににた倍音からなる有声音とはことなる雑音的な音で、無声音の典型である。破裂音に連続して母音がつづくと[b][d][g]の有声音となる。破裂音と後続母音の時間感覚は、連続的だが、人間の聴覚系はこれをあるところで、[p][t][k]か[b][d][g]にカテゴリー化してきいていることが実験的にしめされている。破裂音につづいて、無声性がつよいのは、[s][h]などの摩擦音である。こ

れらは、歯茎や声門のせまいところを息が通過するときにしょうずる音である。摩擦音は、乱流的だが音の共鳴がおきているので、雑音的ななかに楽音的な一定の周波数成分がみられる。歯茎を通過してしょうずる[s]よりも、声門を通過しててしょうずる[h]のほうが、無声性はよわい。[m]や[n]は、唇や歯茎をひらくときにでる息が、はなをとおりぬけるときの音であり、鼻音とよばれる。鼻音では、息が鼻をおりぬけるときの音の共鳴がしょうずるので、摩擦音よりもさらに、雑音成分がよわくなり楽音的な倍音成分がふくまれるようになる。日本語の場合、鼻音は、唇や歯茎をひらくときの息によっているので、子音に分類されるが、言語によっては、声帯の振動によってでる音を鼻腔内で共鳴させる鼻母音もある。以上、今回の実験と関連するあたりの子音を中心に簡略に説明をこころみた。音声学は、生理的な音のだしかたと物理的な音の属性、心理的な聴覚的認知がからんで、なかなかむつかしい。より具体的な説明については、テキストを参照していただきたい。ここでは、ひと口に子音といっても、子音性、あるいは母音性（Sonority）にはさまざまな段階があることを確認しておきたい。[p][t][k]などの破裂音がもっとも子音的である。これらが有声化したのが、濁音の[b][d][g]である。子音でも母音的な共鳴成分をもっているのが、[m]や[n]である。[s][h]などの摩擦音は、破裂音と鼻音の中間的な母音性をもっている。２．で具体的に測定するが、[p][t][k]などの雑音的な破裂音についてピッチをいうことはむつかしいが、有声音の[d]、母音性がたかい鼻音の[n]や、やや母音的な息の音の[h]、より母音性はおちるが歯茎をとおりぬける息の音の[s]については、[a][i][u][e][o]の母音とはべつに、子音のピッチを抽出することが可能である。

[a][i][u][e][o]の母音は、肺からでた息が声帯を振動させ、これが口腔と咽頭のなかで共鳴することによってしょうずる楽音的な音である。声帯のながさは、男性で20～24mm、女性で15～20mm程度である。声帯の振動によってしょうずる音は、基本周波数（F0）とその倍音成分からなる。基本周波数（F0）は、倍音的な周波数成分（順にF1、F2、F3とよばれる）がのるベースを形成し、音の高さの印象であるピッチを規定する。声帯のおおきなひとのF0はひくく、声帯のちいさなひとはF0はたかくなる。F0は、同一のひとでも、言葉の高低のアクセント表現として、また、本論文で問題とする感情表現と関連しても変動する。F0の周波数は、個体差、アクセント、感情表現と関連してかわるが、[a][i][u][e][o]の別は、F1以上のパターンと対応している。方言などにおける母音の変異は、F3も関係してくるが、基本的な[a][i][u][e][o]の区別は、F1とF2のみで可能であることがしめされている。F1とF2は、口腔と咽頭のなかでの共鳴パターンによってかわる。口腔内の共鳴パターンをきめているのは、単純化していうと、舌の位置と口のあけかたである。舌

の位置は、[o]、[a]、[u]、[e]、[i]の順で、前にでる。唇は舌の位置とはほぼ逆に、[i]、[e]、[a]、[u]、[o]の順で前方に突きだされる。[i]、[e]では、舌が前方にでて、唇が後ろにひかれているので、舌の前の部分の口腔がせまくなり、逆に[a]、[u]、[o]では、舌の前の口腔がひろくなる。舌の前の口腔のひろさは、F2を規定しており、佐藤（2001）による成人男性のデータを例にとると、F2は[i]が2226Hz、[e]が1741Hzとたかく、[u]が1112Hz、[a]が1028Hz、[o]が781Hzとひくい。F1は、[a]が648Hz、[e]が524Hz、[o]が447Hz、[i]が360Hz、[u]が322Hの順である。F1は、舌や唇の位置のように外からは確認しにくいが、佐藤（2001）による発生時の声道のレントゲン写真をもとにした図をみると、咽頭の部分が、[a]の場合は非常にせまく、[i]や[u]の場合に広くなっているのを確認できる。F2が口腔の前の部分の空間での共鳴に対応しているのにたいし、F1がより奥の咽頭の部分の空間での共鳴に対応していることがわかる。F1、F2は、F0のうえにのって変動するが、そのパターンによって、[a][i][u][e][o]の母音が、区別されることになる。

ランボーが母音の響きと色彩の関係を歌ったように、母音には、それぞれの感性的な印象がある。Marks, L, E.（1975）は、母音をきいて色彩をかんじる共感覚の報告を整理して、各母音のF1、F2と、色の反対色レベルでの次元の対応を検討している。まず、白-黒の明るさの次元では、[i]、[e]が明るい色、つぎに[a]が明るく、[o]や[u]は暗い色と対応づけられる傾向がある。一般的に周波数の高い音は明るい色と関連づけられる傾向があることが実験的にしめされているが、母音の場合は、F２が母音の明るさの印象と関連していることをMarks, L, E.（1975）は指摘している。これは、[i]、[e]の明るい印象とはぴったりだが、[a]、[o]、[u]の明るさの印象とはあわない。[a]の明るさの印象と、[u]の暗さの印象には、F２だけではなく、F1もかかわっているとかんがえたほうがいいだろう。各母音の明るさの印象は、F2を基本に、F1も加味して判断されるようだ。より微妙になるが、Marks, L, E.（1975）は、赤-緑次元が、F2/F1の比によっていることを指摘している。母音をきいての共感覚の報告での様々な色を赤-緑次元で整理すると、[a]や[o]が赤っぽい色、[u]や[e]が緑ぽい色、[i]が中間といった結果がでている。上にしめしたデータで、F2/F1を計算してみると、[a]が1.6、[o]が1.7、[e]が3.3、[u]が3.5、[i]が6.2となる。[i]が例外だが（Marks, L, E.（1975）は、[i]の場合は、F2と対応した明るさの次元の印象が前面にでてしまったため、F2/F1と赤-緑次元の対応がかくされたのだろうというような推測をしている）、ほかは、F2/F1の比と赤-緑次元がおよそ対応している。F2/F1の比の大きい拡散した[u]や[e]が緑の共感覚を、F2/F1の比の小さいまとまった[a]や[o]が赤の共感覚をしょうじやすいことになる。F2/F1と赤-緑次元の対応があるとしても、なぜかをいうのはむつかしいが、緑は樹木

の葉などのように奥行き感をもって知覚されやすいのにたいし、赤は進出色だしより奥行き感がすくなく知覚されるといった、色彩の空間定位の傾向と母音の拡散、まとまりの印象が対応しているといった説明も、あるいは可能かもしれない。

以上、言語音の周波数成分について説明した。音声をデジタイズし、分析ソフトにかければ、周波数成分に関連した、音声のピッチやF1、F2の推移は、各周波数成分ごとの音の強さや、全体の音圧もふくめて、すぐに表示される。ここで、感情との関連で、音声の属性をとりだそうとするときにむつかしいのは、こうした音素や拍などに対応づけてカテゴリー化されるような言語音の成分に、より連続的で超分節的な音の成分がくわわって時間的に展開されるなかから、どうやって適切な音声の属性をとりだすかである。たとえば、ピッチについて、ある区間の平均値をだすことができるが、どんな区間をとりだすのかがむつかしい。任意の時間間隔を設定して区間を設定することもできるが、言語的な内容との対応がとれなくなってしまう。本論文では、2．で説明するように、音素から文まで、言語音の分節化の各レベルにおうじた区切りをおこなって、各区間の平均やばらつきをもとめた。これは、かなり手間のかかる方法だが、言語的な内容との対応をとるためにはこうするしかない。しかし、そうしても、個々の区間の平均やばらつきだけでは、ピッチのパターンはひろえない。ピッチのパターンについては、言語学のピッチ・アクセント研究でピッチのタイプの分類がこころみられている。しかし、感情表現と関連しては、母親語（Motherlise）におけるほめ言葉や禁止などの表現について、ピッチの上昇、下降などの区別がなされているが（Fernard, A. 1992）、まだ、組織だった記述枠組みは提示されていない。区切りと時間的なパターンに関連するのが、発話のリズムである。発話全体の速さの平均や、個々の間（Pause）を、本論文では測定した。これらは有効な指標となるが、適切な間と不自然な間とはどうちがうかなど、発話のリズムにかんしてはわからないことがおおい。あるリズムをもって、ピッチの高低と、音の強さの変化をもって展開される、発話の変化から、人間は感情などの話者の状態をかんじとっている。研究の現状は、リズムとピッチと強さが、関連して時間的に展開されるパターンをそとがわから計測して、関連する指標をえりわけている段階である。もうひとつ、むつかしい問題は、話者の状態と関連した声の質のあつかいである。人間は、あったかい感じの声とか、つめたいかんじの声、安心感をあたえる声など、発話のリズムやピッチと強さのパターン以外にも、声の質から話者についてのさまざまな判断をしている。感情と声の質の関連については、研究がほとんどなされていない。本論文でも、声の質については、あつかえなかった。

以上、音声の属性把握のむつかしさについてのべた、声の調子と感情の関連の研究のむ

つかしさのもうひとつの原因は、声の調子と関連づけるべき感情自体、どんな種類のものがあり、どんな特徴をもっているのか、心理学の研究が、まだ、諸説いりみだれている状態にあるからである。感情研究の現状については、次の論文で、ここでの研究と照合してのべる。

以上をようするに、把握しにくく定式化されていない部分もおおい音声の属性を、まだ定説がないような感情の種類や性質とをむすびつけようとするのだから、どうしても研究は、断片的で、こうしたらこうしましたのレベルで、一般的な理論にまで到達しがたいのはさけがたい。言語学や認知心理学の音声を対象にした研究で、音声の音素へのカテゴリー化やアクセントのタイプなどの、カテゴリー化やタイプなどの一義的で明確な基準との関連での音声の研究がなされるが、怒りだとか不安だのといった明確なカテゴリー化が可能かどうか不明なような基準との関連での研究がほとんどなされてこなかったのは、このためだろう。

本論文で報告する研究も、探索的なものにならざるをえない。音声特性と印象の関係について、一般性のある結果を定式化するための手がかりとして、本論文の分析で見い出されたもっとも重要な結果は、「はい」とか「ね」といった分節化された言葉の内容、あるいは、その言葉がある発話位置によって、音声の特性と印象の関連がことなることである。これについては本論文の最後でふれる。以下はまず、1.では音声の印象評定を因子分析した結果を2.では音声の印象と音声特性の関連について分析した結果をのべる。

1．音声の印象

1.1.音声の録音

11種類のさまざまな状景を想定した上で、役者志望の女性（22歳）に「はい、そうですね」という短文を、想定した状景にあった感情を込めて発声してもらった。十一種のテープは以下にしめす状況を想定して録音された。このテープを、印象評定の実験と音声の特徴の分析の両方でもちいる。十一種の「はいそうですね」には、それぞれ、A. 感情なし、B. 喜び、C. 悲しみ、D. 怒り、E. 楽しさ、F. 不安、G. 焦り、H. 戸惑い、I. 断固とした、J. 積極的な、K. 消極的などの一般的な感情状態をしめすラベルがつけられているが、具体的には以下のような状況を想定したものである。たとえば、ここでの「怒り」は、上司にたいする発話として設定されているので、威嚇するといったニュアンスはふくまれていない。緊張感とともに不快感を表出するといった設定である。また、「楽しさ」も、ほんわ

かした楽しさではなく、ウキウキするような楽しさである。なお、実験でもちいた11種類の声については、http://www2.ipcku.kansai-u.ac.jp/~ame/の、言語技術研究室に、wave形式のファイルでアップしてある。

《十一種類の「はいそうですね」と想定した情景》

①特定の感情表現なし（A）：「はい、そうですね」。

②喜び（B）：結婚式にて、新婦に対して「幸せな家庭を築けるといいですね」の言葉に、幸せいっぱい、うれしそうに応えて「はい、そうですね」。

③悲しみ（C）：葬式にて、死者の近親者に対しての「最後まで精一杯生きておられましたね」の言葉に、悲しみが抑えきれないまま「はい、そうですね」。

④怒り（D）：会社にて、いやな上司に「そんなんだから、いつまでたっても役に立たないんだよ！この役立たず！」と怒鳴られ、頭にきて「はい、そうですね」。

⑤楽しさ（E）：遊園地で大好きな先輩とデート中。「次はジェットコースターに乗ろうか！」と誘われ、楽しさでウキウキしながら「はい、そうですね」。

⑥不安（F）：受験を控えた面接にて、先生に「この成績だと、こっちの学校も受けておいたほうがよさそうだね」と言われ、不安になりながら「はい、そうですね」。

⑦焦り（G）：退社後に大切な待ち合わせをしているのに、仕事が長引いて、間に合うかどうか微妙になっている。急いで帰ろうとしているところへ、普段から雑用を押し付けてくる上司に「こっちの資料にも目を通しておいたほうがいいんじゃないの？」と言われ、急ぎながら失礼ではない態度で「はい、そうですね」。

⑧戸惑い（H）：通りすがりの人に、良く知らない道を尋ねられる。「駅ってこのまままっすぐでしたっけ？」内心どうだったかなと思いつつ、「はい、そうですね」。

⑨断固とした（I）：社外のプレゼンテーションにて、新商品について「これはこの点で改善されてますよね？」と尋ねられ、自信たっぷりに答えて「はい、そうですね」。

⑩積極的な（J）：第一志望の会社の面接が受かって、自分を推してくれた上司が「これから一緒に頑張ろうな」と言われ、ハキハキと希望にあふれて「はい、そうですね」。

⑪消極的な（K）：自分にとって全くやりがいのない仕事を頼まれているのに、「いやー、大役だよ君ー」と言われ、やる気がわかないまま「はい、そうですね」。

1.2. 音声の印象評定

十一種類の「はいそうですね」の印象評定をSD法により、集団でおこなった。条件は、

以下のとおりである。

（1）被験者:講義に出席している大学生70名（男性28名、女性42名）

（2）印象評定の形容詞対

先行研究と、予備研究にもとづき、二十八の形容詞対をもちいた。評定段階は、「非常に（1）――（2）――どちらでもない（3）――（4）――非常に（5）」の5件法である。表1.1.に形容詞対のリストをしめす。被験者には、形容詞対は、表1.1.の順序でしめした。

表1.1.声の印象評定にもちいた二十八の形容詞対

Adj 1　はっきりした――ぼんやりした
Adj 2　落ち着いた――落ち着きのない
Adj 3　激しい――穏やかな
Adj 4　冷たい――暖かい
Adj 5　消極的な――積極的な
Adj 6　とげとげした――丸みのある
Adj 7　強い――弱い
Adj 8　心地よい――不快な
Adj 9　不安な――安心な
Adj 10　鈍い――鋭い
Adj 11　明るい――暗い
Adj 12　はりのある――はりのない
Adj 13　柔らかい――硬い
Adj 14　やさしい――怖い
Adj 15　きれいな――汚い
Adj 16　悲しそうな――うれしそうな
Adj 17　繊細な――粗野な
Adj 18　好意的な――悪意のある
Adj 19　高圧的な――低姿勢な
Adj 20　誠実な――不誠実な
Adj 21　自信のある――自信のない

Adj 22　心のこもった――うわべだけの

Adj 23　にごった――澄んだ

Adj 24　自然な――不自然な

Adj 25　歯切れのよい――歯切れの悪い

Adj 26　前向きな――後ろ向きな

Adj 27　重い――軽い

Adj 28　好きな――嫌いな

（3）調査の手順

①講義受講者に調査用紙を配布し、氏名、性別などの 記入をしてもらう。

②「はいそうですね」という声が十一種類提示されるので、それぞれの音声の印象を形容詞対で、直感的に判断し5件法で評定するようにもとめた。音声刺激の入った調査用カセットテープを、ラジカセでながし、被験者全員に、はっきりききとれることを確認した。

③音声刺激は、AからKの順でながした。ひとつの音声刺激を呈示するごとに、質問紙への記入を求めた。ひとつの刺激は、まず3回くり返して呈示し、しばらく時間を置いて、また3 回くり返して呈示した。

④全員の記述が完了したかを挙手で確認した上で、次の音声刺激に移行した。

⑤以上の作業を11 試行おこない、調査用紙を回収した。

1.3. 音声の印象

70人の被験者について、十一種類の音声を二十八の形容詞対、5段階で評定したデータがえられた。総計、21560のデータとなるはずだが、5件だけ評定値をマークしてないデータがあった。分析は欠損値のみをのぞいておこなった。

	喜び	楽しさ	断固	積極的	なし	焦り	消極的	不安	悲しみ	戸惑い	怒り	全平均	女-男	性差	交互作用
はっきりした	1	7	8	9	5	6	2	-8	-5	-6	7	3.5	0.0		
積極的な	2	8	5	8	1	2	0	-8	-6	-6	3	3.2	0.3	***	
強い	-2	7	7	7	4	3	4	-7	-6	-4	9	3.4	0.1		
安心な	1	0	2	5	-1	-1	-1	-7	-8	-5	-4	2.7	0.2	**	
鋭い	-2	5	5	6	1	2	3	-6	-3	-5	5	3.2	-0.2	*	
明るい	4	8	4	7	-1	4	0	-8	-8	-2	1	3.1	0.2	**	**
はりのある	8	-6	5	-5	3	8	6	5	3	-8	-1	3.3	0.0		
高圧的な	-6	0	3	0	3	1	4	-4	-5	-2	8	3.0	-0.1		
自信のある	-2	6	6	7	4	-1	3	-6	-7	-6	4	3.1	0.2	*	
歯切れのよい	0	5	5	8	3	-1	2	-8	-6	-7	1	3.1	0.2	**	**
前向きな	1	6	4	8	1	-1	0	-7	-7	-5	-2	3.0	0.2	*	
丸みのある	6	-2	-3	0	-3	-3	-5	1	-1	2	-9	2.7	0.1		**
暖かい	5	3	-2	3	-5	0	-5	-5	-2	-1	-7	2.7	0.2	***	**
心地よい	0	-3	0	4	-4	-2	-3	-5	-7	-4	-8	2.4	0.1		***
きれいな	3	0	1	4	1	0	-1	-2	-3	-1	-6	2.9	0.2	***	*
繊細な	4	-4	-1	1	0	-1	-4	1	4	-1	-9	2.8	0.1		
好意的な	5	5	1	6	-1	1	-2	-2	0	-1	-7	3.1	0.3	***	*
誠実な	3	2	4	6	1	0	-3	-3	1	-2	-8	3.0	0.2	***	
心のこもった	2	3	-1	5	-5	-2	-4	-6	2	-4	-8	2.7	0.3	***	*
澄んだ	2	3	1	4	-1	0	-2	-4	-3	-3	-4	2.9	0.2	**	
自然な	1	0	2	4	-5	-3	-2	-5	-2	-3	-5	2.7	0.2	**	**
好きな	0	0	0	5	-5	-3	-3	-6	-6	-4	-8	2.4	0.3	***	
柔らかい	6	0	-3	-1	-5	-3	-3	-1	-2	0	-6	2.7	0.2	**	
やさしい	5	1	-1	2	-3	0	-2	-3	-4	0	-6	2.8	0.2	*	
うれしそうな	5	7	1	5	-1	1	0	-4	-9	-1	0	3.1	0.1		
軽い	4	6	1	2	-4	0	1	-4	-5	-2	2	3.0	0.0		
落ち着いた	-1	-9	3	0	7	-1	1	3	-5	2	-7	2.9	0.2	***	
穏やかな	5	-7	-3	-3	2	-4	-1	5	-3	5	-7	2.8	0.0		*

図1.1.十一種の音声についての被験者による評定平均値

(十一の音声に対する二十八の形容詞対の評定平均値は、5段階の評定得点の平均値から3を引いて、5倍したものを小数点以下を四捨五入してしめした。全平均は、5段階の評定値のままの平均である。女－男は、5段階の評定値について、女性の平均値から男性の平均値をひいたものである。性差は男女差の有意性検定、交互作用は性差と音声の差の交互作用の有意性検定の結果である。***が0.1パーセント水準で有意差あり、**が1パーセント水準で有意差あり、*が5パーセント水準で有意差ありである。)

1.3.1.二十八形容詞対の評定平均

図1.1.に、70人の被験者による十一種の音声についての二十八の形容詞対の評定平均値をしめした。評定平均値をそのまま表示したのでは、よめない表になるので、以下の処理をおこなっている。

・十一種の音声は、1.3.2.で報告する因子分析の因子得点の結果を参考に、にた印象の

ものがグループになるようにならびかえた。

・二十八の形容詞対も、1.3.2.で報告する因子分析の因子負荷量の結果を参考に、意味のにた形容詞対がおなじグループになるようにならびかえた。また、おなじグループのなかで、因子負荷量が正負逆の形容詞対については、全体として肯定的な意味になるように、項目を逆転した。図1.1.には、必要な逆転をくわえた結果を、値がおおきくなるほうの形容詞とともにしめした。

・形容詞対の評定平均値は、5段階の評定得点の平均値から3を引いて、5倍したものを小数点以下を四捨五入してしめした。こうすると、最低の1が－10、どちらでもないの3が0、最高の5が10となり、－10から＋10までの評定平均値の表示がえられる。

・－10から＋10までの評定平均値で、絶対値が5以上の項目を枠でかこむ。これは、5段階評定で、2～4の、中間的な判断の項目をのぞいたものに相当する。

以上の処理をおこなった結果が、図1.1.の十一種の音声についての二十八の形容詞対の評定平均値である。

二十八の形容詞対の十一種の音声について評定の全平均が、全平均である。全平均は、5段階の評定値のままの平均である。女－男は、5段階の評定値について、女性の平均値から男性の平均値をひいたものである。性差は男女差の有意性検定、交互作用は性差と音声の差の交互作用の有意性検定の結果である。

評定平均で枠でかこまれた項目をみると、あまり特徴のない声と、特徴がくっきりした声があることがわかる。焦りの声は、2つの形容詞対のみが、消極的の声は3つの形容詞対のみが、枠でかこまれいているだけである。これらは、今回もちいた形容詞対では、特徴があまりとらえられないような声である。ぎゃくに、怒りなどは、17もの形容詞対が枠にかこまれていて特徴のある声であることがわかる。あとは、個々の声の印象を、図1.5、図1.6にしめした因子分析の結果よりも、よりローデータにちかいかたちで、表から確認することができる。形容詞対についてみると、「澄んだ」は、枠が0、「きれいな」と「繊細な」は枠が1である。これらの形容詞対は、今回もちいた11の音声を区別するうえでは、あまり有効でなかったといえる。

形容詞ごとの評定平均値をみると、値が3.0の平均より、0.3以上おおきいのが、「はっきりした」(3.5)、「はりのある」(3.3) のふたつである。逆に、値が3.0の平均より、0.3以上ちいさいのが、「安心な」(2.7)、「丸みのある」(2.7)、「暖かい」(2.7)、「心地よい」(2.4)、「心のこもった」(2.7)、「自然な」(2.7)、「好きな」(2.4)、「柔らかい」(2.7) のやっつである。十一の音声は、すべて同一話者による「はいそうですね」だが、全体として声の印

象は、はっきりしてめりはりはあるが、やや人工的で心地よくないものとしてうけとめられている。実際に声をきいてみると、芝居の練習をしていたひとだけあって、めりはりがあるが、やや芝居がかったという印象はうける。ただ、これが、話者の音声の個性によるのか、音の録音の設定のしかたによるのか、実験の設定にあるのかは、別の話者や、べつの音の記録方法、べつの実験事態の設定をしてみないとわからない。

女－男は、5段階の評定値について、女性の平均値から男性の平均値をひいたものである。性差は男女差の有意性検定である。鋭いと高圧的なの二項目をのぞいて、女性の評定平均のほうがたかい傾向がある。全体として、今回の実験の音声にたいしては、女性のほうが、男声よりも、よりよい印象をもつ傾向にある。これが、女性の評価の傾向なのか、同性、異性のちがいかは、男声の声もつかって実験してみないとわからない。男女で、評定に1パーセント水準で有意差がでたのは、「積極的」、「安心な」、「明るい」、「歯切れのよい」、「暖かい」、「きれいな」、「好意的な」、「誠実な」、「心のこもった」、「澄んだ」、「自然な」、「好きな」、「柔らかい」、「落ち着いた」の14形容詞対である。

1.3.2.因子分析の結果

二十八の形容詞対を変数として、70人の被験者×11の音声を770の個体として、因子の抽出を、主因子法による因子分析をおこなう。図1.2.に固有値のスクリープロットをしめした。

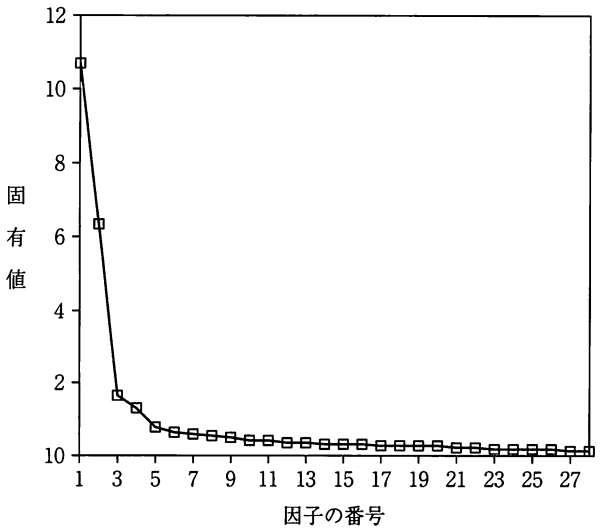

図1.2.固有値のスクリープロット

第一因子と第二因子がとくにおおきく、第三因子と第四因子は変量１個から２個分しかない。あとは、変量１個分よりちいさい。二因子としてもよいが、ややこまかく印象構造をみるために四因子として、斜交回転をおこなう。回転方法は、プロマックス回転による。図１.３.に、因子負荷量のパターン行列をしめした。

図１.３.では因子負荷量の絶対値が0.5以上の形容詞対を、下線でしめした。各因子の内容をみると、第一因子は、ぼんやりした、消極的な、弱い、不安な、鈍い、暗い、はりのない、低姿勢な、自信のない、にごった、歯切れの悪い、後ろ向きな、の12変量が関連している。このままだと、自信のなさの因子となるが、全体を逆転して、自信の因子とする。さきほどの対になるほうの形容詞を列挙すると、はっきりした、積極的な、強い、安心な、鋭い、明るい、はりのある、高圧的な、自信のある、澄んだ、歯切れのよい、前向きなとなる。第二因子は、冷たい、不快な、汚い、粗野な、悪意のある、不誠実な、うわべだけ

		1	2	3	4
Adj 1	はっきりした―ぼんやりした	0.822	0.058	0.007	-0.007
Adj 2	落ち着いた―落ち着きのない	0.105	0.055	0.339	0.834
Adj 3	激しい―穏やかな	0.491	0.109	0.217	-0.550
Adj 4	冷たい―暖かい	0.103	-0.539	0.482	0.232
Adj 5	消極的な―積極的な	-0.713	-0.211	0.129	0.086
Adj 6	とげとげした―丸みのある	0.478	-0.384	0.390	-0.091
Adj 7	強い―弱い	0.924	-0.164	0.023	0.022
Adj 8	心地よい―不快な	0.199	0.574	-0.111	0.224
Adj 9	不安な―安心な	-0.531	-0.262	0.167	-0.266
Adj 10	鈍い―鋭い	-0.765	-0.034	-0.102	0.121
Adj 11	明るい―暗い	0.602	0.071	-0.526	-0.145
Adj 12	はりのある―はりのない	0.853	0.035	-0.069	-0.070
Adj 13	柔らかい―硬い	-0.363	0.246	-0.525	0.006
Adj 14	やさしい―怖い	-0.094	0.337	-0.568	0.064
Adj 15	きれいな―汚い	0.134	0.570	-0.088	0.236
Adj 16	悲しそうな―うれしそうな	-0.479	0.023	0.627	0.069
Adj 17	繊細な―粗野な	-0.382	0.781	0.271	0.121
Adj 18	好意的な―悪意のある	0.012	0.821	-0.132	-0.112
Adj 19	高圧的な―低姿勢な	0.670	-0.540	0.117	0.191
Adj 20	誠実な―不誠実な	0.071	0.948	0.202	0.049
Adj 21	自信のある―自信のない	0.856	-0.012	0.008	0.169
Adj 22	心のこもった―うわべだけの	0.027	0.979	0.143	-0.217
Adj 23	にごった―澄んだ	-0.257	-0.595	0.112	0.038
Adj 24	自然な―不自然な	0.158	0.702	0.090	0.016
Adj 25	歯切れのよい―歯切れの悪い	0.757	0.199	0.001	0.106
Adj 26	前向きな―後ろ向きな	0.649	0.335	-0.147	0.052
Adj 27	重い―軽い	-0.260	0.159	0.637	0.259
Adj 28	好きな―嫌いな	0.228	0.654	-0.085	0.097

図１.３.因子負荷量のパターン行列

の、不自然な、嫌いなで、非友好の因子である。これも、因子を逆転して友好の因子とする。友好因子の形容詞は、暖かい、心地よい、きれいな、繊細な、好意的な、誠実な、心のこもった、自然な、好きなである。第三因子と関係するのは、明るい、柔らかい、やさしい、うれしそうな、軽い、の6変量である。これを、軽快さの因子とする。第四因子は、落ち着きのない、激しい、だが、これも逆転して、落ち着いた、穏やかなの、落ち着きの因子とする。Adj 6の「とげとげした－丸みのある」は、どの因子にもかかわらない。Adj 11の明るいは、自信の因子と軽快さの因子の両方にかかわっている。このふたつ以外は、すべて、ひとつの因子にのみ絶対値が0.5以上の負荷をもつので、およそ単純構造が達成されているといえる。

図1.4.に、因子を逆転したあとの因子間相関をしめした。自信の因子は、他の因子とあまり相関がなくほぼ独立の因子である。これにたいし、友好、軽快、落ち着きの三因子はたがいに正の相関関係にある。とくに、友好因子と軽快因子は0.66のたかい相関をしめしている。落ち着き因子も、友好因子、軽快因子の両方と、0.4弱の相関をしめしている。

	1.自信	2.友好	3.軽快さ	4.落ち着き
1.自信	1	0.18	0.17	-0.09
2.友好	0.18	1	0.66	0.39
3.軽快さ	0.17	0.66	1	0.38
4.落ち着き	-0.09	0.39	0.38	1

図1.4.因子間相関

図1.5.に、自信因子の因子得点と友好因子の因子得点を、各音声ごとに男女別に平均したものをプロットした結果をしめした。アルファベットのAからKまでが声の種類で、大文字が男子の平均、小文字が女子の平均である。矢印でつないだのは、男女の平均値に有意な差があった場合である。焦り、消極的な、戸惑い、断固とした、の友好の因子の平均値について、1パーセント水準での有意差がみられた。自信については、有意差はみられなかった。図1.1.の形容詞対ごとの平均値の比較でも観察したが、焦りや、消極的、戸惑い、断固としたといった、基本的な感情との関連のない複雑で不定形な感情と関連した音声について、男女差がみられている。図1.1.で、評定者の性別と声の種類との交互作用がしょうじたのは、友好因子にぞくする形容詞がおおい。暖かい、心地のよい、きれいな、好意的な、心のこもった、自然な、である。図1.1.で交互作用がしめされた自信因子と関連した形容詞は、明るいと歯切れのよいの二つだけである。自信因子で評定者による性差がみられなかったことは、友好度の評定とくらべて、音声の自信の印象が、より

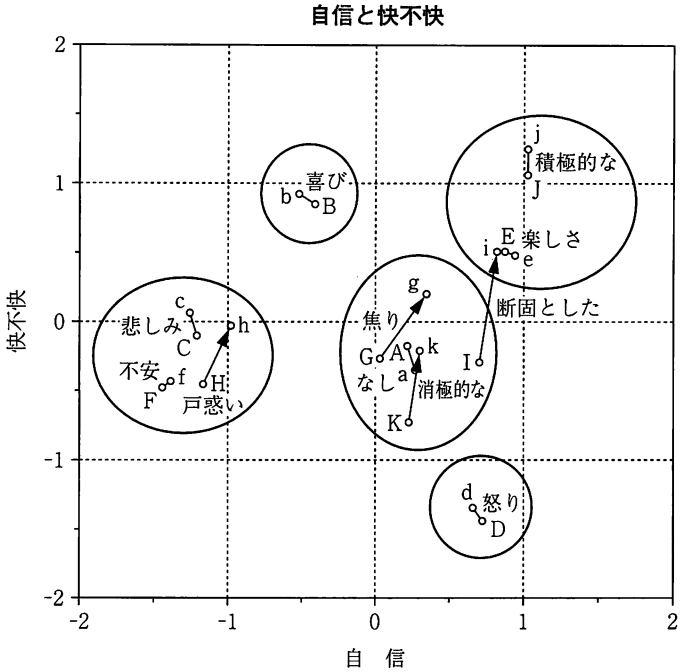

図１.５.第一因子と第二因子の因子得点のグラフ

単純な音声属性と心理過程によっている可能性を示唆している。図１.５.では、ちかいところにマップされた音声をグループ化してみた。周辺には、楽しさ、喜び、悲しみ、不安、怒りといった、基本的な感情と関連した音声が位置している。ごくおおざっぱにいうと、自信+・友好＋→楽しさ・積極的な、自信＋・友好－→怒り、自信－・友好＋→喜び、自信－・友好ニュートラル→悲しみ・不安・戸惑い、といった、自信と友好度による感情の分化がみられる。友好度は、感情の次元分析で一般的な、快・不快の次元に相当する。自信は、力量の次元と関連している。ここでの次元は、音声の印象についてのものなので、力量が、自信あり-自信なしの次元として、第一次元にきている。

図１.６.に、軽快さと落ち着きの因子の因子得点のグラフをしめした。軽快さと落ち着きの因子は、形容詞対１個から２個程度のおもみしかないが、ここでも基本感情に関連した音声が周辺に位置している。軽快さ＋落ち着き＋→喜び、軽快さ＋落ち着き－→楽しさ、軽快さ－・落ち着き＋→不安・感情なし、軽快さ－・落ち着き－→悲しみ・怒りといった分化がみられる。ここで、喜びと楽しさの分化は、状況設定をみたほうがいいかもしれない。喜びは、結婚式での祝福にこたえての「はいそうですね」で、楽しさはデート中ジェットコースターにのろうかというさそいへのこたえとしての「はいそうですね」である。

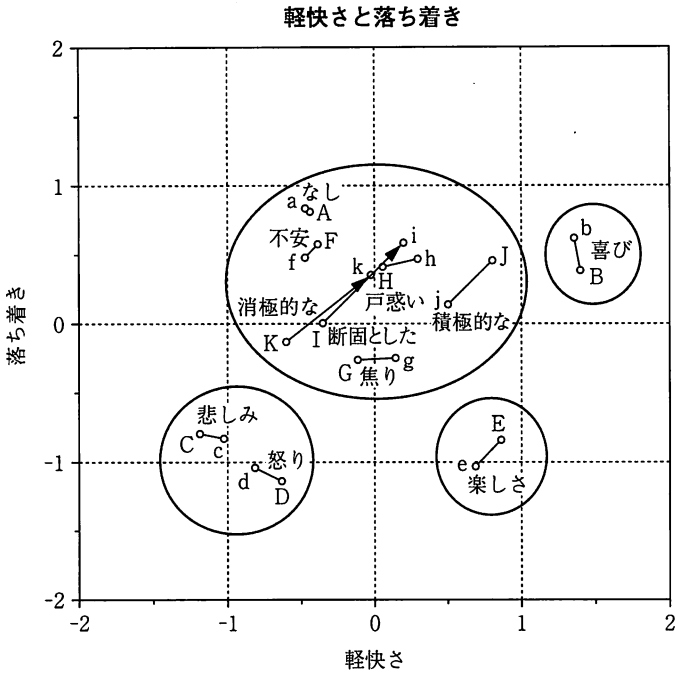

図1.6.第三因子と第四因子の因子得点のグラフ

落ち着きのあるなしの差は状況設定と対応している。消極的なと断固としたでは、軽快さ、落ち着きの両方の次元で、男女の評定に有意差がある。女性のほうが、音声を、男性よりも、軽快で、落ち着いたものとして、より好意的にうけとる傾向がある。

2．音声の「物理的特性」と「印象に関する因子」との関連

2.1.目的

われわれ人間は、音声の物理的な情報を聴覚器官でとらえ、その器官でとらえた情報を脳におくり、音声の「高さ」や「大きさ」、「持続時間」などの心的特性を認識することができる。また、研究者は、音声の物理的な情報から、音声解析の装置やソフトウェアをつかうことによって、うえであげた3つの心的特性とまったく同じではないが、それらに対応している物理的特性を得ることができる。音声にふくまれている感情や印象を感じ取ったりするとき、うえであげた心的特性によって感じ取りかたが変わってくるであろう。こういったことを、心的特性と対応のある物理的特性をつかって研究できないであろうか。

言語学の分野では、イントネーションの研究について、音声の物理的特性をつかった研

究が行われること（杉藤,1999; 城生,2001）は多いのだが、心理学の分野では、物理的特性をつかった研究そのものが少なく、この論文のテーマになっている、音声にともなう感情や印象の研究についても、音声の物理的特性をつかった研究が行われることは少ないようである。しかし、音声認識などが盛んに研究されている情報工学の分野では、音声の物理的特性をつかって、音声にともなう感情や印象の研究が盛んになりつつある（森山・斎藤・小沢,1999; 向井・目加田・春日・松本・小池,2000など)。これらのレビューは次の論文で行われる。

以降、物理的特性については本論文においてよく出てくるため特性と略することがあるが、心的特性については混乱を避けるため略さずに用いる。

われわれは、音声にともなう感情や印象について、物理的特性との関連から研究を進めていくのだが、研究のはじめの段階として、１．の実験で役者志願の22歳の女性１名によって感情のこめられた「はいそうですね」という11種類の音声を用いることにした。感情と印象の両者について研究するのは１つの論文では収まりきれないので、音声にふくまれる感情についての検討はつぎの論文で行うことにし、ここでは音声から受け取られる印象について検討する。印象について研究するといっても、１．のSD法で用いられた印象に関する28個の形容詞対との関連について検討するのは、非常に労力を必要とするため、因子分析によって集約された４つの因子との関連について検討することにした。したがって、１．で用いられた11種類の音声から、音声解析のソフトウェアをつかって取り出された物理的特性のうち、どの特性あるいはどの特性の組み合わせ（以降、特性の組み合わせも単に特性とよぶことがある）が、１．の因子分析によってえられた「自信因子」と「友好因子」、「軽快さ因子」、「落ち着き因子」にたいしてどのように影響しているのかを検討することになる。

音声の特性やその特性の組み合わせにはさまざまあるが、ここでは、ピッチや音圧、持続時間という特性と、それらの特性の組み合わせ（それらの特性から派生された特性）を測定する。特性の組み合わせとしては、ピッチの変化に関わる特性や、音圧の変化に関わる特性を求める。変化に関わる特性には、上昇や下降の大きさ（変化量）という特性や、単なる変化の大きさ（変動量）という特性、上昇や下降の緩急（変化率）という特性、緩急の大きさ（変動率）という特性がある。

ここでは、音声の文全体についてのみ、以上であげた持続時間という特性（１つ）やピッチの変化に関する特性（４つ)、音圧の変化に関する特性（４つ）の９つを測定するのではなく、ローマ字や仮名文字、単語、文節で区分された単位からも、同じく９つの特性

を測定する。また、文全体については、ピッチや音圧そのものの程度（平均）やばらつき（標準偏差）といった特性をも求める。

もう少し具体的に述べると、測定された特性には、仮名文字の単位においては、たとえば、「は」の部分での持続時間やピッチの緩急の大きさ（変動率）、音圧の上昇や下降の大きさ（変化量）など9つの特性があり、ほかのすべての仮名文字でも同じく9つの特性がある。単語単位においては、たとえば、「はい」の部分での持続時間やピッチの変化の大きさ（変動量）など9つの特性があり、ほかのすべての単語でも同じく9つの特性がある。また、ローマ字の単位でも、文節の単位においても、うえと同じことがいえる。測定されたほかの特性については、方法に詳細を述べているので、そちらを参考にされたい。

まとめると、本研究の目的は、以上であげたようなさまざまな音声の特性が4つの因子にたいしてどのように影響しているのかを検討することである。

2.2. 音声の視覚的表現と音声解析ソフトの紹介

本研究の方法について述べるまえに、ここでは、音やその物理的特性の表現について簡単に説明し、それらを表現するソフトウェア（以降、ソフトと略す）について紹介する。

2.2.1. 音声の波形による表現

音は物理的には、振幅（音圧）の時間的な変動によって表現され、紙面上やモニタ上で、横軸に時間、縦軸に振幅がとられ、波形で表現されている（図2.2.1.）。音に関するさまざまなソフトには、音が波形で表現される機能がごく基本的に備わっている。

音の波形を見慣れた人は、波形を見ることによって、ある程度、それがどのような音であるかを視覚的に判断することができるだろう。人は聴覚的に、音がどのくらい大きくて、どのくらい複雑なのかを判断できるが、視覚的に表現された波形を見ることでも、ある程度、音の大きさや複雑さを判断できる。しかし、人は、聴覚的に音が高いのか低いのかを判断できても、視覚的に表現された波形を見ただけでは、音の高低を判断することは難しいだろうし、また、聴覚的にさまざまな母音や子音を聞き分けることができても、視覚的に表現された波形を見てもそれらを判断することも難しいだろう。これらを解決する手段として、2.2.2で述べるスペクトログラムの表現がある。

たとえば、図2.2.1では音声を波形で表現しているが、この図を見ることで、「はい」と「そうですね」の区間がおおまかに分かるが、母音や子音を区分するのはスペクトログラムと比べて少し難しい。また、「はい」や「そう」、「です」の始めの部分では音圧が大

きくなっており、音圧と音の大きさには完全な線形関係ではないので断定できないが、人には音が大きく聞こえていると見当できる。しかし、音の高さがどうなっているかを見当するのは困難である。

図２.２.１. 音声の波形表現

この図は１．の実験で使用された「戸惑い」の音声波形であり、横軸が時間で縦軸が振幅を表わしている。また、図の下にあるローマ字は実際の音声と対応して筆者が書き込んだものである。

２.２.２. 音声のフーリエ変換による表現

もっとも単純な音の一つには正弦波がある。こういった音なら、人は波形から視覚的に音の高さを判断できるが、自然界に存在する音はたいてい複雑な音であり、そう簡単には判断できない。人はそういった複雑な音の波形を見ることで、音の大きさを判断できても、音の高さを判断するのは困難であり、さまざまな母音や子音を見分けることはかなり困難である。そこで、音をフーリエの定理に基づいて分解し、スペクトログラムという図で表現することがある。人は、波形を見ることよりも、この図を見ることによって、容易に音の高さを判断でき、母音や子音を見分けることができるのである。

フーリエの定理とは、Moore（1989）の解説をかりて説明すると、任意の複雑な波形が（一定の制限のもとで）特定の周波数、振幅、位相をもつ一連の正弦波に分解できるというものである。周波数とは同一波形が毎秒繰り返される回数であり（１Hz=１サイクル／秒)、振幅とは平均圧力からの圧力変化量であり、位相とは固定した時刻からの波の進み具合を表すサイクル上の位置である。

このように分解されたものを周波数ごとに表現した図はスペクトログラムとよばれる。ちなみに、フーリエ分解は数値的に行なわれるが、人間の聴覚器官のうち蝸牛でも、フーリエ分解のような波形の分解が行なわれている。

スペクトログラムを描画する機能は、音声解析ソフトにはたいてい備わっている。マイクロソフト社製のWindows95以降のOSでつかえるソフトには、NTTアドバンステクノロジ製の「SP 4 win」やアニモ社製の「SUGI Speech Analyzer」がある。とくに、後者のソフト制作者の一人である杉藤（1999）は、マニュアルでソフトの操作だけではなく、日本語のイントネーションの分析などについて簡単に紹介している。

図2.2.2.では、音声が狭帯域スペクトログラムで表現されているが、この図では波形よりも母音や子音を区分しやすく、あばら骨のような模様では母音か有声子音であると見当でき、雑音のようなものが入っている色の濃い箇所では無声子音であると見当できる。また、「そう」の始めから終わりにかけて、高い周波数成分が減り、低い周波数成分が増えていっているように見えるが、このことから音が徐々に低くなっていると見当できる。

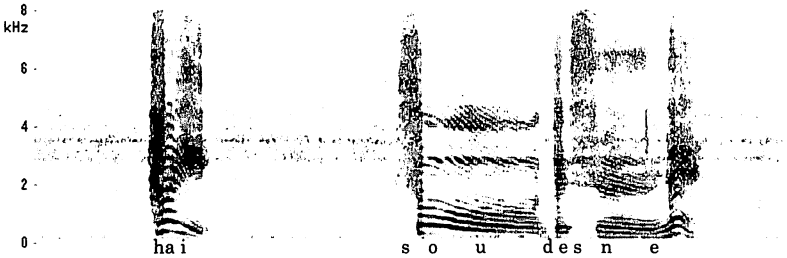

図2.2.2. 音声のスペクトログラム表現
この図は、SUGI Speech Analyzerをつかって、「戸惑い」の音声を狭帯域スペクトログラムで表現したものであり、横軸が時間で縦軸が周波数成分を表わしている。色が濃いところほどその部分の周波数成分が強いことを意味している。図の下にあるローマ字は実際の音声と対応して筆者が書き込んだものである。

2.2.3. 音の高さの表現

音の高さというのは、Moore（1989）によると、音を音階上に順序づけるときの聴覚的な特性として定義される、主観的な特性である。そのため、複雑な音の高さを計測するとき、被験者にその音とおなじ高さにきこえるように正弦波の周波数を調整させることによって音の高さが調べられ、その正弦波の周波数が複雑な音の高さの測度として使われる。

音声解析ソフトには、被験者の判断によるのではなく、数値的な処理によって、音の高さを抽出する機能があり、うえであげた2つのソフトにはその機能が備わっている。以降、実際に人間が主観的にとらえた音の高さと区別するために、数値的な処理によって抽出された音の高さをピッチとよぶ。斎藤・中田（1981）によると、ピッチの抽出とは、なんらかの方法によって音声波形もしくはその生成の原因である音源波形の周期性の度合いと、その周期を抽出することであるが、確実で正確なピッチの自動抽出法はまだ確立されていない。

図2.2.3.にSUGI Speech Analyzerで、ノイズ閾値を10%に、有効周波数を20Hzから2000Hzに設定して抽出したピッチを示したが、狭帯域スペクトログラムでは見当しかできなかったピッチの変化がこの図ではよく分かる。たとえば、「そう」の始めから終わりにかけて、音が徐々に低くなり、「ね」の半ばから急に音が高くなって低くなっているこ

とが分かる。

また、ドットの連なりから外れているドットがところどころにあるが、人間の発声する音声は、音の高さが突然欠けたように短い間だけ低くなったりしないので、それらはノイズだと考えられる。音声の録音時に雑音が混入していたり、発話者の発声が粗雑であったりするとノイズが多くなる。ところで、ピッチを抽出するとき、たいていの場合はノイズと考えられるドットがいくつも散在しており、そういったドットを取り除くためにノイズ除去が行われるのだが、ノイズ閾値を小さい値に設定すればそれらが残ることが多く、大きい値に設定すれば必要なピッチも消すことになるので注意せねばならない。

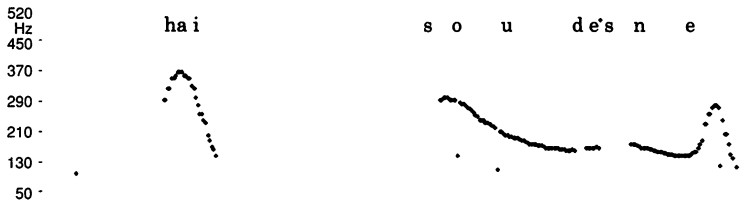

図2.2.3. 抽出されたピッチの表現

この図は、SUGI Speech Analyzerをつかって、「戸惑い」の音声から抽出されたピッチを表現したものであり、横軸が時間で縦軸がピッチを表わしている。連続的なドットの上にあるローマ字は実際の音声と対応して筆者が書き込んだものである。

2.3. 方法

2.3.1. スペクトログラム表現とピッチの抽出

11種類の音声の物理的特性と4因子との関連を分析するにあたって、16000Hz、16bitでサンプリングされた各音声データから、音声言語研究所アニモ社製のSUGI Speech Analyzer Version 1.07をつかって、音声波形、ピッチと狭帯域スペクトログラム、音圧をそれぞれグラフで描画した（図2.3.1.参照）。スペクトログラムはソフトの初期設定になっていたハミング窓をもちい、狭帯域スペクトログラムで表された。ピッチはノイズ閾値が10%で、有効周波数が20Hzから2000Hzのあいだで抽出された。

2.3.2.「ローマ字による区分」と「その始点と終点の時刻・ピッチ・音圧の記述」

分析者は狭帯域スペクトログラムを見ながら音声聴取をおこない、各音声において、音素に対応する、各ローマ字（h,a,i,s,o,u,d,e,s,u,n,e）の発声と、吐息やポーズ（発話前、発話中、発話後）の計15区間を定めた。その際の問題点は「3.1. ローマ字や仮名文字よる音声区分の難しさ」で述べる。

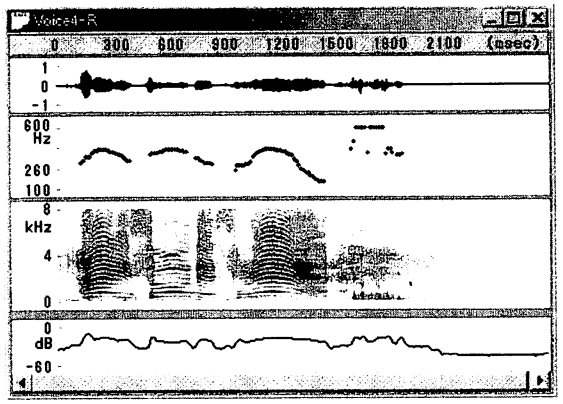

図2.3.1. 音声解析ソフトによる音声の視覚的表現
この図は、「怒り」の音声について、上から順に、音声波形とピッチ、スペクトログラム、音圧をグラフで描画した。

ローマ字や吐息、ポーズの始点、それらの終点が、録音されはじめてからどのくらいの時間が経った時刻なのか、各区間の「始点の時刻（ts)」と「終点の時刻（te)」を1ミリ秒単位で記述した。また、各区間の「始点の音圧（as)」と「終点の音圧（ae)」を1dB単位で記述した。さらに、各区間の始点と終点において1Hz単位でピッチの記述を試みたが、各区間の始点と終点においては、ピッチが抽出されていないこともあるので、各区間のなかで最初に抽出されたピッチを「代替始点のピッチ（ps)」、ピッチの抽出された時刻を「代替始点の時刻（ts2)」とし、各区間のなかで最後に抽出されたピッチを「代替終点のピッチ（pe)」、ピッチの抽出された時刻を「代替終点の時刻（te2)」とし、それぞれピッチに関しては1Hz単位で、時刻に関しては1ミリ秒単位で記述した。したがって、各音声において、15区間の始点と終点それぞれの時刻と音圧、ピッチ、ピッチを抽出した時刻の4つの値が記述された。ただし、音声が観測されなかった区間は、さきの4つの値を空白のままにした。

ここで、ローマ字単位での15区間の表記法について述べる。ローマ字は12個あるが、「s」と「u」、「e」のローマ字は2つ存在する。ローマ字の右に数字を添え、各ローマ字(h,a,i,s,o,u,d,e,s,u,n,e）の区間を「h1」、「a1」、「i1」、「s1」、「o1」、「u1」、「d1」、「e1」、「s2」、「u2」、「n1」、「e2」として記号であらわした。つぎに、発話前の吐息の区間を「sp1」として、「はい」と「そうですね」の間にあるポーズの区間を「sp2」として、発話後の吐息の区間を「sp3」として記号であらわした。

つぎに、ローマ字単位の15区間について、それぞれ9つの音声特性変数（以降、単に変数と述べることがある）の値をもとめる。持続時間、ピッチの変化量・変動量・変化率・

変動率、音圧の変化量・変動量・変化率・変動率の9変数である。仮名文字単位、単語単位、文節単位で区切った区間についても同様に9変数の値をもとめる。文全体については、ローマ字単位の区間の9変数をもとに、文全体でのそれらの基礎統計量（平均値と標準偏差）などを算出する。以下、ローマ字単位、仮名文字単位、単語単位、文節単位、文全体の順に説明する。

2.3.3. ローマ字単位の音声特性と4因子との関連を求める方法

各音声において、すべてのローマ字の発声区間と吐息やポーズの区間の持続時間（⊿t）が「終点の時刻（te）－始点の時刻（ts）」で求められた。つぎに、各区間で以上に述べた変数が求められた。まず、ピッチの変化量（⊿p）が「代替終点のピッチ（pe）－代替始点のピッチ（ps）」で求められ、音圧の変化量（⊿a）が「終点の音圧（ae）－始点の音圧（as）」で求められた。また、ピッチの変化率や変動率を求めるために、代替持続時間（⊿t2）が「代替終点の時刻（te2）－代替始点の時刻（ts2）」で求められた。さらに、ピッチの変化率（⊿p/⊿t2）が「ピッチの変化量／代替持続時間」で求められ、音圧の変化率（⊿a/⊿t）が「音圧の変化量／持続時間」で求められた。なお、ピッチの上がり下がりの度合い、音圧の上がり下がりの度合いが、どのくらいかを表わす指標として、ピッチの変動量（|⊿p|）が「｜ピッチの変化量｜」（ピッチの変化量の絶対値）で、音圧の変動量（|⊿a|）が「｜音圧の変化量｜」で求められ、ピッチの変動率（|⊿p/⊿t2|）が「｜ピッチの変化率｜」（ピッチの変化率の絶対値）で、音圧の変動率（|⊿a/⊿t|）が「｜音圧の変化率｜」で求められた。

本論文においては、以上の変数、それから以下に出てくる変数のすべてを総称して音声特性変数とよぶことにする。

このようにして求められた9つの音声特性変数（⊿t,⊿p,|⊿p|,⊿a,|⊿a|,⊿p/⊿t2,|⊿p/⊿t2|,⊿a/⊿t,|⊿a/⊿t|）が4因子にどのくらい関連しているのかを検討する。分析方法としては、11種類の音声を対象（個体）として、ローマ字単位の15区間について、9つの各音声特性変数と4つの各因子得点の平均値（70人の被験者）とのピアソンの積率相関係数を求めた。ここで、9つの音声特性変数はローマ字単位の各区間についてのものだが、音声にふくまれる印象に関する因子得点は、「はいそうですね」の発話全体についてのものである。これは、2.3.4.から2.3.6.で述べる、仮名文字単位、単語単位、文節単位の場合も同様である。

さいごに、各区間の変数の表記法について説明する。たとえば、「s1」の区間における、

ピッチの変化量は⊿p[s 1]、音圧の変動率は|⊿a/⊿t|[s 1]などと、ローマ字を特定して音声特性変数をあらわすときには、音声特性変数の右側にローマ字を[　]で括ってあらわした。ほかの音声特性変数の場合も同様である。

2.3.4. 仮名文字単位の音声特性と4因子との関連を求める方法

各音声において、すべての仮名（は、い、そ、う、で、す、ね）の発声区間と吐息やポーズの区間（計10区間）が区分されるように、2.3.2.で区分されたローマ字の区間をつかって、それらの区間が定められた。たとえば、「は」では「h」と「a」の組み合わせでその区間が定められており、「い」では「i」だけでその区間が定められている。

つぎに、各仮名文字の音声特性変数が求められるが、たとえば、「は」では、「は」の持続時間（⊿t[は]）は「a」の終点の時刻から「h」の始点の時刻をさしひいた値とされ、「は」のピッチの変化量（⊿p[は]）は「a」の代替終点のピッチから「h」の代替始点のピッチをさしひいた値とされ、「は」の音圧の変化量（⊿a[は]）は「a」の終点の音圧から「h」の始点の音圧をさしひいた値とされる。ただし、ピッチの変化量に関しては「h」にピッチが抽出されていない場合があり、その場合には「a」の代替始点のピッチが代わりに使われる。

また、ピッチの変動量はピッチの変化量の絶対値として、音圧の変動量は音圧の変化量の絶対値として求められる。ピッチの変化率はその変化量が「⊿t 2 [は]」で除され、音圧の変化率はその変化量が「⊿t[は]」で除されて求められる。また、ピッチの変動率はピッチの変化率の絶対値として、音圧の変動率は音圧の変化率の絶対値として求められる。「い」では、1つの「i」だけで「い」の区間が求められるため、2.3.3.と同じく「i」だけで音声特性変数が求められる。

このようにして求められた音声特性変数（⊿t,⊿p,|⊿p|,⊿a,|⊿a|,⊿p/⊿t 2 ,|⊿p/⊿t 2 |,⊿a/⊿t,|⊿a/⊿t|）が4因子にどのくらい関連しているのかを検討するため、10区間について音声特性変数と因子得点の平均値とのピアソンの積率相関係数が求められた。

さいごに、変数の表記法であるが、「は」の区間において、ピッチの変化量を⊿p[は]という記号であらわし、音圧の変動率を|⊿a/⊿t|[は]という記号であらわすというように、仮名文字を特定して音声特性変数をあらわすときには、音声特性変数の右側に仮名文字を[ ]で括って示す。

2.3.5. 単語単位の音声特性と4因子との関連を求める方法

各音声において、すべての単語（はい、そう、です、ね）の発声区間と吐息やポーズの区間（計7区間）が区分されるように、2.3.2.で区分されたローマ字の区間をつかって、それらの区間が定められた。たとえば、「はい」では「h」と「a」、「i」の組み合わせでその区間が定められている。

つぎに、各単語の音声特性変数が求められるが、たとえば、「はい」では、「はい」の持続時間（⊿t[はい]）は「i」の終点の時刻から「h」の始点をさしひいた値とされ、「はい」のピッチの変化量（⊿p[はい]）は「i」の代替終点のピッチから「h」の代替始点のピッチをさしひいた値とされ、「はい」の音圧の変化量（⊿a[はい]）は「i」の終点の音圧から「h」の始点の音圧をさしひいた値とされる。ただし、ピッチの変化量に関しては「h」にピッチが抽出されていない場合は、「a」の代替始点のピッチが代わりに使われ、もし「a」にピッチが抽出されていない場合は、「i」の始点のピッチが使われる。

また、ピッチの変動量はピッチの変化量の絶対値として、音圧の変動量は音圧の変化量の絶対値として求められる。ピッチの変化率はその変化量が「⊿t2[はい]」で除され、音圧の変化率はその変化量が「⊿t[はい]」で除されて求められる。また、ピッチの変動率はピッチの変化率の絶対値として、音圧の変動率は音圧の変化率の絶対値として求められる。「ね」では、1つの仮名文字である「ね」だけで「ね」の区間が求められるため、2.3.4.と同じく「ね」だけで音声特性変数が求められる。

このようにして求められた音声特性変数（⊿t,⊿p,|⊿p|,⊿a,|⊿a|,⊿p/⊿t2,|⊿p/⊿t2|,⊿a/⊿t,|⊿a/⊿t|）が4因子にどのくらい関連しているのかを検討するため、7区間について音声特性変数と因子得点の平均値とのピアソンの積率相関係数が求められた。

さいごに、変数の表記法であるが、「はい」の区間において、ピッチの変化量を⊿p[はい]という記号であらわし、音圧の変動率を|⊿a/⊿t|[はい]という記号であらわすというように、単語を特定して音声特性変数をあらわすときには、音声特性変数の右側に単語を[ ]で括って示す。

2.3.6. 文節単位の音声特性と4因子との関連を求める方法

各音声において、すべての文節（はい、そうですね）の発声区間と吐息やポーズの区間（計5区間）が区分されるように、2.3.2.で区分されたローマ字の区間をつかって、2.3.4.や2.3.5.と同様な方法で、それらの区間が定められた。

つぎに、各文節の音声特性変数が求められるが、これらもまた、2.3.4.や2.3.5.と

同様な方法で、音声特性変数が求められる。

このようにして求められた音声特性変数（⊿t,⊿p,|⊿p|,⊿a,|⊿a|,⊿p/⊿t 2 ,|⊿p/⊿t 2 |,⊿a/⊿t,|⊿a/⊿t|）が4因子にどのくらい関連しているのかを検討するため、5区間について音声特性変数と因子得点の平均値とのピアソンの積率相関係数が求められた。

さいごに、変数の表記法であるが、「そうですね」の区間においては、ピッチの変化量を⊿p[そうですね]という記号であらわし、音圧の変動率を|⊿a/⊿t|[そうですね]という記号であらわす。文節を特定して音声特性変数をあらわすときには、音声特性変数の右側に文節を[　]で括って示す。

2.3.7. 文全体の音声特性と4因子との関連を求める方法

「はいそうですね」の発声区間（計1区間）で、音声ごとに、音声特性変数を算出した。文全体の音声特性変数については、新たな変数を導入した。これは、「はいそうですね」という音声全体について、ローマ字の15区間における、ピッチをあらわすps,peの平均値と標準偏差、音圧をあらわすas,aeの平均値と標準偏差を求めたものである。また、変化量に関する変数としては、ローマ字単位の15区間、仮名文字単位の10区間、単語単位の7区間における、⊿t、⊿p、⊿a、⊿p/⊿t 2 、|⊿p/⊿t 2 |、⊿a/⊿t、|⊿a/⊿t|　の値について、「はいそうですね」という音声全体での、それぞれの平均値と標準偏差を算出しこれをもちいた。

このようにして求められた音声特性変数が4因子にどのくらい関連しているのかを検討するため、音声特性変数と因子得点の平均値とのピアソンの積率相関係数が求められた。

2.3.8. 文節、単語、仮名文字、ローマ字の単位による区分表

4つの単位によって「はいそうですね」を区切ったわけだが、文全体をくわえて、区切られたローマ字や仮名文字、単語、文節の区間をまとめて表2.3.1.に示した。この表は文の区分表としての役目を果たすだけではなく、音声特性変数をセルに書き込んで、すべての単位をとおして変数の値を見渡すための役目を果たすこともできるし、音声特性変数と因子との相関係数を書き込んで、すべての単位をとおして相関を見渡すための役目を果たすこともできる。たとえば、区分表のうえに「ピッチの変化量」と書き記し、それぞれのセルにピッチの変化量の値を書き込めば、ピッチの変化量をすべての単位をとおして見渡すことができ、区分表のうえに「自信因子とピッチの変化量との相関係数」と書き記し、それぞれのセルにその相関係数を書き込めば、相関係数をすべての単位をとおして見渡す

ことができる。もちろん、相関係数ではなく散布図を示してもよい。

表2.3.1.「はい、そうですね」の文節・単語・仮名文字・ローマ字の区分表

文	sp1 はい sp2 そうですね sp3													
文節	sp1	はい			sp2	そうですね								sp3
単語						そう			です				ね	
仮名文字		は		い		そ		う	で		す			
ローマ字		h	a	(i)		s	o	(u)	d	e	s	u	n	e

2.4. 結果・考察

2.4.1. 音声特性の基礎統計量と音声の観測されなかった区間

11種類の音声刺激の特徴を概観するためのデータとして、文全体について、ピッチと音圧の基礎統計量、ローマ字単位で得られた変化量に関する変数の基礎統計量を表2.4.1.と表2.4.2.に示す。この部分の考察については、ここでの分析の流れから外れるので本論文では省くことにし、先行研究と比較して次の論文で詳しく述べることにする。

表2.4.1. 文全体における音声特性（平均値）

音声	ピッチ	音圧	⊿t	⊿p	⊿p/⊿t2	⊿a	⊿a/⊿t
感情なし	260.43	-23.25	121.50	-6.00	-0.08	0.50	0.02
喜び	254.81	-24.54	133.00	-23.38	-0.23	0.08	0.01
悲しみ	287.88	-26.71	191.42	-7.00	-0.04	0.25	0.00
怒り	357.28	-25.38	166.92	-22.56	0.06	-0.75	0.03
楽しさ	421.70	-21.46	151.92	-27.40	-0.19	-0.08	0.01
不安	232.78	-24.31	251.38	-23.78	-0.18	0.00	0.02
焦り	362.63	-22.63	222.50	7.00	0.17	-0.42	0.01
戸惑い	219.83	-25.73	283.46	-38.33	-0.03	0.08	0.02
断固とした	215.44	-24.33	113.33	6.38	0.07	-0.50	0.03
積極的な	303.31	-24.50	103.69	31.38	0.61	-0.54	-0.01
消極的な	236.94	-25.08	110.42	-17.88	-0.07	0.17	0.04

表2.4.2. 文全体における音声特性（標準偏差）

音声	ピッチ	音圧	⊿t	⊿p	⊿p/⊿t2	⊿a	⊿a/⊿t
感情なし	68.96	4.67	62.71	23.34	0.30	7.44	0.09
喜び	94.48	5.76	109.37	52.32	0.46	9.51	0.09
悲しみ	42.20	4.95	356.35	34.55	0.46	8.08	0.09
怒り	98.91	6.43	169.67	81.44	0.64	8.90	0.11
楽しさ	99.00	5.69	100.36	65.41	0.65	8.14	0.07
不安	64.56	4.53	491.53	33.07	0.27	6.06	0.07
焦り	95.09	7.27	366.70	69.35	0.90	10.82	0.15
戸惑い	80.69	7.20	327.06	90.84	0.55	10.67	0.11
断固とした	41.30	6.69	125.14	51.88	0.78	11.18	0.17
積極的な	92.00	6.95	146.22	104.73	2.10	9.89	0.16
消極的な	40.62	5.50	116.85	47.50	0.60	8.75	0.14

ほかに音声刺激の特徴を知る手段として、ローマ字に対応する音声が観測されなかった区間について述べておく必要があるだろう。たとえば、「す」の区間にふくまれている「u 2」の区間においては、「戸惑い」と「不安」の音声で「u」のローマ字に対応する音声が観測されたが、ほかの音声刺激では「u」の音声が観測されなかった。また、吐息やポーズの区間については次のようなことが観測された。「sp 2」の区間においては、「怒り」の音声ではポーズが観測されず、ほかの音声刺激ではポーズが観測され、「喜び」の音声ではポーズにくわえ息の音が観測された。「sp 1」の区間においては、「楽しさ」の音声で息の音が観測されただけで、ほかの音声刺激ではなにも観測されなかった。「sp 3」の区間においては、「怒り」と「積極的な」の音声で息の音が観測されただけで、ほかの音声ではなにも観測されなかった。「喜び」と「怒り」については、図2.4.1.参照し、「戸惑い」については、図2.2.1.と図2.2.2.を参照するとよい。

息の音については数値化できなかった。なぜなら、今回の方法では、音声の区間の始点と終点のあいだにある音圧を数値化しておらず、息の音については、始点と終点のあいだにある音圧は大きいが、始点と終点における音圧は小さいからである（図2.4.1.参照)。同様の理由で、すべての区間におけるピッチや音圧については数値であらわせなかった。

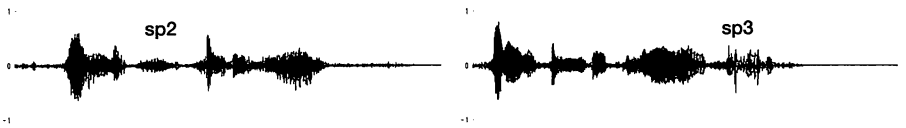

図2.4.1. 息の音の入っている音声波形
右図は「喜び」、左図は「怒り」の音声波形であり、横軸が時間で縦軸が振幅（音圧）を表わしている。また、図の上の「sp2」と「sp3」は実際の音声と対応させて筆者が書き込んだものである。

さて、各言語単位（ローマ字単位や仮名文字単位、単語単位、文節単位、文全体）において、「印象に関する因子」と「音声特性変数」との相関係数のなかでも0.6以上かつ5％水準で有意な相関係数から、本研究の目的である「物理的特性がどのように印象に関する因子にたいして影響を与えているのか」について考察をおこなう。ただし、今回のデータでは5％水準で有意な相関係数はすべて0.6以上であった。

考察するときの注意点として次の2つがある。相関係数からの考察では、相関係数は高いものの実際には相関がなかったり、相関係数は低いものの実際には相関があったりするので、そういったことを考慮にいれて考察するべきだろう。また、本研究は探索的であり、分析であつかった音声は11種類と少ないため、10%水準で有意な相関係数も考慮にいれて考察するとよいかもしれない。しかし、こういったことを考慮して、そのまま文章にする

と煩雑になるため、以降で考察するとき、これらの点を省いて述べた。

2.4.2. ローマ字単位の音声特性と4因子との相関

ローマ字単位において、音声特性変数と4因子とのピアソンの積率相関係数を求めた。5％水準で有意な相関係数を表2.4.3.に示す。全体的にみて因子と音声特性との相関は自信因子において多かった。ローマ字単位においては、表2.4.3.から次のようなことがいえるだろう。

表2.4.3. ローマ字単位の特性と4因子との相関

単位	音声特性	自信	友好	軽快さ	落ち着き
h1	⊿t	-0.69			
i1	⊿a	-0.65			
i1	|⊿a|	0.65			
i1	⊿a/⊿t	-0.65			
i1	|⊿a/⊿t|	0.65			
sp2	⊿t	-0.74			
s1	⊿a			0.71	
s1	|⊿a|			0.71	
o1	⊿a	-0.66			
o1	⊿a/⊿t	-0.72			
e2	⊿p		0.61		
e2	|⊿p|		-0.70		

** $p<.01$　* なし $p<.05$

a. 自信因子への影響

自信因子には、「h1」と「sp2」の部分での持続時間という特性以外は、音圧に関する特性が影響していると考えられる。

- ポーズの時間や音声のはじめの「h」の吐く息が短いほど、自信のある印象として受け取られ、逆にそれらが長いほど、自信のない印象として受け取られる。
- 「はい」の「i」の音圧レベルの下がり方が急なほど、自信のある印象として受け取られ、音圧レベルの下がり方が緩やかなほど、自信のない印象として受け取られる。
- 「そう」の「o」の音圧レベルが急に下がるほど、自信のある印象として受け取られ、逆に音圧レベルが急に上がるほど、自信のない印象として受け取られる。

b. 友好因子への影響

友好因子には、「e2」の部分でピッチに関する特性のみが影響していると考えられる。

- 音声の最後にあたる「e」のピッチの下降の幅が小さいほど、友好的な印象として受け取られるが、ピッチの下降の幅が大きいほど友好的ではない印象として受け取られる

c. 軽快さ因子への影響

軽快さ因子には、「s1」の部分で音圧に関する特性のみが影響していると考えられる。

- ポーズのあとに発声される「s」の音圧レベルの上がり方が大きなほど、軽快さのある印象として受け取られるが、音圧レベルの上がり方が小さなほど、軽快さのない印象として受け取られる。

d. 落ち着き因子への影響

落ち着き因子には、どの特性からも影響していないと考えられる。

e.　見かけの相関について

ローマ字単位だけでなく、すべての言語単位において、5％水準で有意な相関係数が示されるわけだが、このなかには、実際には相関のない見かけの相関の場合もある。本研究では、見かけの相関があるのは、共通の原因があるからだけではなく、そもそもピアソンの相関係数を求める際に、音声特性変数には互いに独立ではない変数もふくませているからである。

共通の原因については「2.4.7.g.　言語あるいは発話位置による見かけの相関」で述べることにするが、音声特性変数の独立ではない変数には、音声特性の派生関係にある変数（派生変数）や、言語単位の包含関係にある変数（包含変数）がある。変数の派生関係とは、変化量とそれが持続時間で除された変化率などの関係であり、ほかにその関係には、変化量とそれが絶対値であらわされた変動量、持続時間とそれが変化率で乗算された変化量、変化率とそれが変化量で除された持続時間などの関係がある。変数の包含関係とは、ある言語単位の変数とほかの言語単位の変数との関係である。たとえば、仮名文字単位でのピッチの変化量は、ローマ字単位でのピッチの変化量を包含するが、単語単位でのピッチの変化量や文節単位のピッチの変化量に包含される、といった関係のことである。

この見せかけの相関については、ローマ字単位だけの問題ではなく、すべての言語単位の問題であるが、ここでは、ローマ字単位におけるa.からd.の考察に、以上のことをふまえて、新たな考察をつけくわえる。まず、派生関係について考慮して考察をつけくわえ、つぎに、包含変数について考慮して考察をつけくわえる。ほかの言語単位では派生変数について考慮して考察はつけくわえないが、同様の考察ができるので省くことにする。

f.　派生関係を考慮した考察

ローマ字単位におけるa.の考察のなかに、「o1」の音圧レベルが急に下がるほど（変化率が大きいほど）自信のある印象として受け取られるとした考察がある。実際には、この考察にくわえ、「また、音圧レベルの下降が大きいほど（変化量が大きいほど）自信のある印象として受け取られる可能性がある」と述べたほうがより適切であったと思われる。なぜなら、「o1」には、⊿aと自信因子、⊿a/⊿tと自信因子とのあいだに高い相関があるからである。しかし、そう述べなかったのは、⊿a/⊿tと自信因子とのあいだの高い相関にひきずられて、⊿aと自信因子とのあいだに高い相関が生じた可能性があると考えたからである。もちろん、実際には、⊿a/⊿tと自信因子とのあいだに高い相関がなくても、⊿aと自信因子とのあいだに高い相関が生じている可能性もあると考えられ、両方の変数から

自信因子に影響を与えている可能性も考えられる。⊿aと自信因子とのあいだに、⊿a/⊿tと自信因子とのあいだに高い相関があるとき、考察をわかりやすくするためにも、⊿aのことは考察の対象に入れなかった。ほかにも変化量と変化率の組み合わせで高い相関が生じた場合、より正確に考察をするなら、ここで述べたような考察がなされるべきだろう。

g. 包含関係を考慮した考察

ローマ字単位におけるa.からd.の考察から、ローマ字の部分、いいかえれば、母音や子音の部分だけからでも、全体の印象に影響を与えている可能性が考えられる。もし、相関のでたローマ字を含んだ仮名文字に同じ音声特性において相関がでなかった場合には、母音や子音の部分だけからでも、全体の印象に影響を与えていることが明らかになるが、その仮名文字に同じ音声特性において相関がでた場合には、母音や子音の部分だけから、全体の印象に影響を与えているかどうか分からなくなるだろう。全体の印象に影響を与えているかどうか分からないというのは、仮名文字の部分にある相関にひきずられて、母音や子音の部分に相関があるのか、仮名文字の部分にある相関にひきずられることなく母音や子音の部分に相関があるのか、どちらも考えられるからである。

2.4.3. 仮名文字単位の音声特性と4因子との相関

仮名文字単位において、音声特性変数と4因子とのピアソンの積率相関係数を求めた。5％水準で有意な相関係数を表2.4.4.に示す。全体的にみて因子と音声特性との相関は自信因子において多かった。仮名文字単位においては、表2.4.4.から次のようなことがいえるだろう。

表2.4.4. 仮名文字単位の特性と4因子との相関

単位	音声特性	自信	友好	軽快さ	落ち着き
は	|⊿p|		0.64		
い	⊿a	-0.65			
い	|⊿a|	0.65			
い	⊿a/⊿t	-0.65			
い	|⊿a/⊿t|	0.65			
sp2	⊿t	-0.74			
で	|⊿p|				-0.68
で	⊿p/⊿t2	-0.76**			
で	|⊿p/⊿t2|	0.69			-0.71
ね	⊿p		0.62		
ね	|⊿p|				-0.66
ね	|⊿p/⊿t2|				-0.88**

** p<.01　* なし p<.05

a. 自信因子への影響

自信因子には、「い」の部分で音圧に関する特性が影響していると考えられるが、「で」の部分ではピッチに関する特性が影響していると考えられる。

・「い」の部分に関しては、ローマ字単位と同様である。

・ポーズの時間に関しては、ローマ字単位と同様である。
・「で」のピッチの下がり方が急なほど、自信のある印象として受け取られるが、ピッチの下がり方が緩やかなほど、自信のない印象として受け取られる。

b. 友好因子への影響

友好因子には、ピッチに関する特性のみが影響していると考えられる。

・「は」のピッチの上昇下降の幅が大きいほど、友好的な印象として受け取られ、逆に、ピッチの上昇下降の幅が小さいほど、友好的ではない印象として受け取られる。
・「ね」のピッチが高くなっていくほど、友好的な印象として受け取られ、逆にピッチが低くなっていくほど、友好的ではない印象として受け取られる。

c. 軽快さ因子への影響

軽快さ因子には、どの特性からも影響していないと考えられる。

d. 落ち着き因子への影響

落ち着き因子には、ピッチに関する特性のみが影響していると考えられ、とくに「ね」の部分でピッチの変化率が影響していると考えられる。

・「で」や「ね」のピッチの上がり方が急なほど、落ち着きのない印象として受け取られるが、ピッチの上がり方が緩やかなほど、落ち着きのある印象として受け取られる。

e. 包含関係を考慮した考察

各因子にたいして、ローマ字の部分だけから影響があるかどうかを2.4.2.g.で述べたが、自信因子と⊿t[h 1]や⊿a[o 1]、⊿a/⊿t[o 1]とのあいだに、軽快さ因子と⊿a[s 1]や|⊿a|[s 1]とのあいだに相関が出ているが、その言語単位にふくまれた仮名文字の部分においてそれらと同じ音声特性との相関が出ていないことから考えると、母音や子音の部分だけからでも、全体の印象に影響を与えていることが考えられる。

また、ローマ字の部分の影響がなくても、各因子にたいして影響があるかどうかについては、2.4.2.で述べたローマ字単位だけから影響があるかどうかの考察と同じように考えることができる。落ち着き因子と|⊿p|[ね]、|⊿p/⊿t 2 |[ね]とのあいだに相関が出ているが、その言語単位にふくまれたローマ字や単語の部分の両者においてそれらと同じ音声特性との相関が出ていないことから考えると（「ね」の仮名文字を単語として見なさなければ)、仮名文字の部分からだけでも、全体の印象に影響を与えている可能性がある。

2.4.4. 単語単位の音声特性と4因子との相関

単語単位において、音声特性変数と4因子とのピアソンの積率相関係数を求めた。5％

水準で有意な相関係数を表２.４.５.に示す。全体的にみて因子と音声特性との相関は自信因子と友好因子において多かった。単語単位においては、表２.４.５.から次のようなことがいえるだろう。

表２.４.５. 単語単位の特性と４因子との相関

単位	音声特性	自信	友好	軽快さ	落ち着き
はい	⊿p	0.73			
はい	|⊿p|		0.71		
はい	⊿p/⊿t2	0.65			
はい	|⊿p/⊿t2|		0.72		
はい	⊿a	-0.69			
はい	|⊿a|		0.73		
はい	⊿a/⊿t	-0.66			
はい	|⊿a/⊿t|		0.72		
sp2	⊿t	-0.74			
です	⊿p/⊿t2	-0.78**			
です	|⊿p/⊿t2|	0.71			-0.72
ね	⊿p		0.62		
ね	|⊿p|				-0.66
ね	|⊿p/⊿t2|				-0.88**

** p<.01　* なし p<.05

a. 自信因子への影響

自信因子には、「はい」の部分でピッチに関する特性と音圧に関する特性が影響していると考えられるが、「です」の部分ではピッチに関する特性のみが影響していると考えられる。

- 「はい」のピッチが急に高くなるほど、自信のある印象として受け取られ、逆にピッチが急に低くなるほど、自信のない印象として受け取られる。
- 「はい」の音圧レベルが急に大きくなるほど、自信のない印象として受け取られ、逆に音圧レベルが急に小さくなるほど、自信のある印象として受け取られる。
- ポーズの時間に関してはローマ字単位と同様である。
- 「です」のピッチの下がり方が急なほど、自信のある印象として受け取られるが、ピッチの下がり方が緩やかなほど、自信のない印象として受け取られる。

b. 友好因子への影響

友好因子には、「はい」の部分でピッチに関する特性と音圧に関する特性が影響していると考えられるが、「ね」の部分ではピッチの変化に特性のみが影響していると考えられる。

- 「はい」のピッチの上昇下降が急なほど、友好的な印象として受け取られ、逆にピッチの上昇下降が緩やかなほど、友好的ではない印象として受け取られる。
- 「はい」の音圧の上昇下降が急なほど、友好的な印象として受け取られ、逆に音圧の上昇下降が緩やかなほど、友好的ではない印象として受け取られる。
- 「ね」のピッチに関しては仮名文字単位と同様である。

c. 軽快さ因子への影響

軽快さ因子には、どの特性からも影響していないと考えられる。

d. 落ち着き因子への影響

落ち着き因子には、「です」と「ね」の部分でピッチに関する特性のみが影響していると考えられる。

・「です」のピッチの上昇下降が急なほど、落ち着きのない印象として受け取られ、逆にピッチの上昇下降が緩やかなほど、落ち着きのある印象として受け取られる。
・「ね」のピッチに関しては仮名文字単位と同様である。

e. 包含関係を考慮した考察

ここでは、以上の結果と考察から、ローマ字の部分の影響がなくても（仮名文字の部分の影響がある可能性があるが)、各因子にたいして影響があるかどうかについて考える。このことは2.4.2.で述べたローマ字単位だけから影響があるかどうかの考察と同じように考えることができる。自信因子と⊿p/⊿t2[で]や|⊿p/⊿t2|[で]、⊿p/⊿t2[はい]、|⊿p/⊿t2|[はい]とのあいだに、友好因子と|⊿p|[は]や|⊿a|[はい]、|⊿a/⊿t|[はい]とのあいだに、落ち着き因子と|⊿p|[で]や|⊿p/⊿t2|[で]、|⊿p|[ね]、|⊿p/⊿t2|[ね]、|⊿p/⊿t2|[です]とのあいだに相関が出ているが、それらの言語単位にふくまれたローマ字の部分においてそれらと同じ音声特性との相関が出ていないことから考えると、ローマ字の部分の影響がなくても、全体の印象に影響を与えていることが考えられる。

ローマ字や仮名文字の部分の影響がなくても、いいかえれば、単語といった中程度（文節の単位と比べて）の単位になってはじめて、各因子にたいして影響があるかどうかについては、2.4.2.で述べたローマ字単位だけから影響があるかどうかの考察と同じように考えることができる。自信因子と⊿p[はい]や⊿p/⊿t2[はい]とのあいだに、友好因子と|⊿a|[はい]、|⊿a/⊿t|[はい]とのあいだに相関が出ているが、その言語単位にふくまれたローマ字や仮名文字の部分の両者においてそれらと同じ特性との相関が出ていないことから考えると、ローマ字や仮名文字の部分から全体の印象に影響を与えなくても、単語の部分だけからでも、全体の印象に影響を与えることが考えられる。

2.4.5. 文節単位の音声特性と4因子との相関

文節単位においても、音声特性変数と4因子とのピアソンの積率相関係数を求めたが、5％水準で有意な相関係数の表は示さなかった。「はい」の部分は単語単位の場合と同じで、「そうですね」の部分では、どの因子にたいしても、すべての音声特性変数との相関係数が5％水準で有意な相関がなかったからである。「はい」に関しては、「2.3.4. 単語単位の音声特性と4因子との相関」を参照されたい。

2.4.6. 文全体の音声特性と4因子との相関

文全体において、始点終点とローマ字単位、仮名文字単位、単語単位のそれぞれで得られた音声特性変数の平均と4因子とのピアソンの積率相関係数を求めた。5％水準で有意な相関係数を表2.4.6.に示す。全体的にみて因子と音声特性の相関は自信因子において多かった。

表2.4.6. 文全体における平均と4因子との相関

単位	音声特性	自信	友好	軽快さ	落ち着き
始点終点	ps,pe				-0.77**
ローマ字	⊿t	-0.71			
ローマ字	|⊿p/⊿t2|	0.65			
ローマ字	⊿a/⊿t		-0.69		
仮名文字	⊿t	-0.74**			
仮名文字	⊿p	0.64			
仮名文字	|⊿p/⊿t2|	0.61			
単語	⊿t	-0.75**			
単語	⊿p	0.81**			
単語	⊿p/⊿t2	0.61			
単語	|⊿p/⊿t2|		0.67		
単語	|⊿a/⊿t|		0.68		

** p<.01　* なし p<.05

さらに、文全体において、始点終点とローマ字単位、仮名文字単位、単語単位のそれぞれで得られた音声特性変数の標準偏差と4因子とのピアソンの積率相関係数を求めた。5％水準で有意な相関係数を表2.4.7.に示す。全体的にみて因子と音声特性の相関は友好因子において多かった。

表2.4.7. 文全体における標準偏差と4因子との相関

単位	音声特性	自信	友好	軽快さ	落ち着き
ローマ字	⊿t	-0.73**			
ローマ字	|⊿p/⊿t2|		0.61		
ローマ字	|⊿a|			0.70	
仮名文字	⊿t	-0.74**			
仮名文字	|⊿p|		0.65	0.69	
仮名文字	⊿p/⊿t2	0.70	0.62		
仮名文字	|⊿p/⊿t2|		0.78**	0.64	
単語	⊿t	-0.73			
単語	⊿p		0.67	0.65	
単語	|⊿p|		0.65		
単語	⊿p/⊿t2		0.65		
単語	⊿a/⊿t		0.69		
単語	|⊿a/⊿t|		0.64		

** p<.01　* なし p<.05

表2.4.6.と表2.4.7.の両方にある単位の列について述べる。たとえば、ローマ字単位で得られた音声特性変数と因子とのあいだに5％水準で有意な相関があるときには、単位の列に「ローマ字」という語を書き込んだ。文字単位や単語単位もまた同様である。文節単位で得られた音声特性変数については平均や標準偏差を求めていないので、単位の列には「文節」という語はあらわれない。

文全体においては、表2.4.6.と表2.4.7.から次のようなことがいえるだろう。

a. 自信因子への影響

自信因子には、持続時間とピッチに関する特性の平均が影響していると考えられる。

・どの単位においても、持続時間が長く、ばらつきが大きいほど、自信のない印象として

受け取られ、逆に、持続時間が短く、ばらつきが小さいほど、自信のある印象として受け取られる。
・ピッチについては、単語単位では、上がり方が急なほど、単語より小さい単位では上昇下降にばらつきがあるほど、その上昇下降が急なほど、自信のある印象として受け取られる。音圧レベルについては、文全体からは自信のある印象には影響していない。

b. 友好因子への影響

友好因子には、ピッチに関する特性や音圧に関する特性のばらつきが大きく影響していると考えられる。

・ピッチについては、単語単位では、上昇下降にばらつきがあるほど、その上昇下降が急なほど、また、単語より小さい単位では、上昇下降にばらつきがあるほど、緩急にばらつきがあるほど、友好的な印象として受け取られる。
・音圧レベルについては、単語単位では、上昇下降が急なほど、上昇下降にばらつきがあるほど、緩急にばらつきがあるほど、また、ローマ字単位では、下がり方が急なほど、友好的な印象として受け取られる。

c. 軽快さ因子への影響

軽快さ因子には、ピッチに関する特性や音圧に関する特性のばらつきのみが影響していると考えられる。

・ピッチについては、単語単位では、上昇下降の幅にばらつきがあるほど、また、単語より小さい単位では、緩急にばらつきがあるほど、友好的な印象として受け取られる。
・音圧レベルについては、ローマ字単位で、上昇下降の幅にばらつきがあるほど、友好的な印象として受け取られる。

d. 落ち着き因子への影響

落ち着き因子には、ピッチの程度のみが影響していると考えられる。

・文全体をとおして、ピッチが低いほど、落ち着いている印象として受け取られ、ピッチが高いほど、落ち着いていない印象として受け取られる。

2.4.7. 全体的な考察

ローマ字や仮名文字単位、単語単位、文節単位、文全体をとおして、5%水準で有意な相関がでたが、さまざまな観点から全体的に考察をおこなうことにする。以降、有意といえば、5%水準で有意ということである。

考察するうえで注意すべきことの1つに、本研究でつかわれる変数は、文全体をとおしての始点と終点におけるピッチ（ps,pe）と音圧（as,ae）以外は、すべて変化量に関する変数だということがある。こういった音声特性だけを変数として選んだ理由は、SUGI

Speech Analyzerから得られたデータをつかって、簡単に求められる変数が、そういった変数だったからである。本来なら、各単位におけるピッチや音圧の平均や標準偏差を求めるべきであろう。これに関しては今後やっていかねばならないことである。

ほかには、見かけの相関については、a.からe.にかけては考慮せずに考察をおこなう。ただし、g.では見かけの相関について考察をおこなう。見かけの相関については、「２.４.２. ローマ字単位の音声特性と４因子との相関」のe.からg.にかけてと「２.４.７.g. 言語あるいは発話位置による見かけの相関」で詳しく述べられている。

a. 全体を通した考察

４つの因子はどんな音声特性によって影響されているのかを全体をとおして考えることにする。軽快さ因子は、区分された単位においては、ローマ字単位の「ｓ１」の部分での「音圧の変化に関する特性」だけから影響されており、文全体においては、おもに「ピッチの変化に関するばらつき」から影響されている。友好因子は、区分された単位においては、単語単位の「ピッチの変化に関する特性」や「音圧の変化に関する特性」からよく影響されており、文全体においては、「ピッチの変化に関するばらつき」や「音圧の変化に関するばらつき」からよく影響されている。落ち着き因子は、区分された単位においては、仮名文字単位の「ピッチの変化に関する特性」からよく影響されており、文全体においては、「全体的なピッチの程度」だけから影響されている。自信因子は、区分された単位においては、どの単位の特性からも影響されており、ローマ字単位では「音圧の変化に関する特性」だけから影響されているが、仮名文字単位や単語単位では「ピッチの変化に関する特性」や「音圧の変化に関する特性」から影響されている。また、自信因子は、ほかの因子とは異なり、「持続時間」からも影響されている。文全体においては、自信因子は「ピッチの変化に関する特性の程度」からよく影響されている。また、ほかの因子とは異なり、「持続時間の程度」や「持続時間のばらつき」からも影響されている。

b. 音声特性が４因子に影響しやすい単位

まず、ローマ字単位、仮名文字単位、単語単位、文節単位における相関行列の結果から、どの単位で有意な相関が出やすいのかを確認する。単語単位においては、その単位で求められたすべての相関係数のうち有意に相関のある割合がほかの単位にくらべて高かった。仮名文字単位においては、その単位で求められたすべての相関係数のうち有意に相関のある割合がつぎに高く、ローマ字単位においては、その単位で求められたすべての相関係数

のうち有意に相関のある割合がほかの単位とくらべて低かった。文節単位においては、その単位で求められたすべての相関係数のうち有意になっている相関が、単語単位における「はい」にしかなく、文節を単位にすることによってあらたに出た相関はないことが確認できた。

したがって、本研究で分析した変化に関する音声特性についていえば、単語単位で、印象に関する因子にたいして、もっとも影響しやすいことがわかる。このことから、変化に関する特性が印象の因子に影響を与えるものとして、単語が重要な役割をになうことが示唆された。

c. 音声特性が各因子に影響する単位の相違

因子ごとに考察すると、軽快さ因子については、ローマ字単位のみで音声特性変数と有意な相関が出ているが、ほかの単位では有意な相関は出なかった。落ち着き因子では、ローマ字単位だけは変数と有意な相関が出なかったが、仮名文字単位と単語単位では音声特性変数と有意な相関が出ており、同じくらいの割合で有意な相関が出ている。自信因子については、すべての単位で有意な相関が出ており、単語単位でももっとも大きい割合で有意な相関が出ている。友好因子については、すべての単位で有意な相関が出ており、単語単位でもっとも有意な相関が出ている。このことから、因子のちがいによって音声特性の影響する単位が異なることが示された。

また、因子ごとに考察しても、変化に関する特性が印象の因子にもっとも影響しやすいのは、軽快さ因子は例外であるが、b.と同じく単語単位であることも示唆された。

d. 各因子に影響する音声特性の種類

落ち着き因子については、4つの各単位と文全体から得られたすべての音声特性のうち、ピッチに関する変数に有意な相関が出ている。ほかの因子については、ピッチに関する変数と音圧に関する変数に有意な相関が出ている。このことから、落ち着き因子はピッチに関する特性のみから影響されると考えられるが、自信因子や友好因子、軽快さ因子はピッチに関する特性や音圧に関する特性のどちらからでも影響されると考えられる。

e. どこかの単位のみで各因子に影響する音声特性

「2.4.3. 仮名文字単位」と「2.4.4. 単語単位」の「e. 包含関係を考慮した考察」で述べてきたことをまとめると、どこかの単位においてのみ、因子と相関のでる音声特性

があることが確認された。ローマ字単位だけで、ある因子に影響を及ぼす音声特性には、自信因子にたいして⊿t[h 1]や⊿a[o 1]、⊿a/⊿t[o 1]があり、軽快さ因子にたいしては⊿a[s 1]や|⊿a|[s 1]がある。仮名文字単位だけでは、落ち着き因子にたいして|⊿p|[ね]や|⊿p/⊿t 2 |[ね]がある。単語単位だけでは、自信因子にたいして⊿p[はい]や⊿p/⊿t 2 [はい]があり、友好因子にたいして|⊿a|[はい]や|⊿a/⊿t|[はい]がある。

以上のことから、全体の印象に関する因子にたいして、ローマ字単位だけで影響を及ぼす音声特性もあれば、仮名文字単位だけで影響を及ぼす音声特性もあり、単語単位だけで影響を及ぼす音声特性もあることが示唆された。

f.　各因子に影響する言語あるいは発話位置

すべての因子にたいして、「そう」の部分では音声特性からの影響はほとんどない。自信因子にたいして「o 1 」の部分が影響しているだけである。言語要素との関わりが問題なのか、発話の位置との関わりが問題なのか、あるいは、今回の音声データで、たまたま、「そう」の部分で印象に影響するような音声特性の多様性がなかっただけなのかは、明確に判断できない。

因子ごとに考察すると、落ち着き因子は「ですね」の部分の音声特性だけから影響されている。「ですね」という単語から影響しているのか、「ですね」の発話全体のなかの位置、つまり、さいごの位置から影響しているのかのどちらかだと考えられる。軽快さ因子は「s 1 」の音声特性だけから影響されている。この場合は、おそらく発話の位置からの影響だと考えるより、「s」の音素の音声学的特徴との関わりだと考えるほうがよいと思われる。自信因子は「そ」と「ね」の部分からは影響されていないが、ほかの部分からはかなり影響されている。「はい」と「です」という単語が影響しているのか、「はい」と「です」の発話全体のなかの位置が影響しているのかのどちらかだと考えられる。友好因子は「そうです」の部分からは影響されておらず、「はい」と「ね」の部分からはかなり影響されている。これについても、単語の種類の影響と、発話の始めと終わりという位置の影響とが考えられる。

音声のある部分のみに印象因子に対する有意な相関が出たとしても、このことから、言語要素が関係しているのか、発話位置が関係しているのかは明確に判断できない。これを明確に判断するためには、音声刺激を統制して実験しなければならない。

g. 共通の原因による見かけの相関

友好因子にたいして、「はい」と「ね」の部分からはかなり影響しているとしたが、実際には、友好因子に「はい」の部分だけからしか影響していないことも考えられる。なぜなら、「はい」と「ね」にあるなんらかの共通の原因により、「ね」の部分で相関が出ている可能性があるからである。これを明確にするためには、「はい」と「ね」のあいだに相関があるかどうか調べるべきであろう。ただし、相関がない場合には、両者の部分から影響しているということが明確になるが、相関がある場合には、どちらの部分から影響しているのか、両者から影響しているのかは明確にはできない。これを明確にするためには、さきと同じで音声刺激を統制して実験しなければならない。

h. 知見の一般化について

2.4.7.では、さまざまな観点で考察してきたが、今回は1名の話者による「はいそうですね」という音声しか用いていないので、そこで得られたいくつかの知見を一般化することは不可能であり、それらを一般化するための手がかりにしかならない。それでも、今回の研究によって、どのような影響があるのかを知るための有益な手がかりを得たと思われるし、ありえないだろうという考えを否定する知見を得たと思われる。

本研究では、「そう」にふくまれる「s」の音素について相関があったように、ローマ字単位による音声特性からでも、印象の因子に影響するといった知見が得られたと思われる。ほかには、印象に関する因子への影響は、音声特性以外から完全に独立しているのだということを否定できたと思われ、音声特性だけではなく、音声特性以外の要素、たとえば、言語あるいは発話位置の要素からも印象の因子に影響しているだろう。

今後の研究の方向性として、「はいそうですね」の音声以外をつかって、本研究と同じ手続きで研究することで一般化を試みてよいし、この研究で得られた手がかりをつかって仮説をたて、音声刺激を統制した実験を試みてもよいだろう。

2.5. 印象の音声分析に関する問題点

2.5.1. ローマ字や仮名文字よる音声区分の難しさ

今回、「はいそうですね」という音声を、「h」と「a」、「i」、「s」、「o」、「u」、「d」、「e」、「s」、「u」、「n」、「e」というローマ字に対応する音声に区分することは容易ではなかった。また、「は」と「い」、「そ」、「う」、「で」、「す」、「ね」という仮名文字に対応する音声に区分することにも難しい点があった。とくに、「ai」の「a」と「i」、「ou」の「o」と「u」

を区切ることはかなり困難だった。「ne」の「n」と「e」を区切るのはさほど困難ではなかったが、「n」の影響がほとんどない「e」はどこからかは明確にしにくかった。

ここでは、「ai」を例にあげて区切りの難しさを述べたいと思う。11種類の「ai」という音声をとおして、分析者１名（出身地も育ちも京都である男性）は、音韻論的には「a」であろう部分では、「a」として聞こえるのではなく、「i」が弱く混じったような「ai」として聞こえ、音韻論的には「i」であろう部分では、「a」が混じって聞こえ、先の部分よりも「i」が強くなったような「ai」として聞こえた。さらには、「i」のほぼ最後まで「a」が混じっているように聞こえ、「ai」のどの部分をとっても「ai」として聞こえた。また、「ou」に関しても、「ai」と同様な現象で生じた。

以上のことから、音声学的には（実際の音声では)、「はい」の「a」と「i」や、「そう」の「o」と「u」を区切ることはできず、それぞれを「a」と「i」が順に混じった１つの母音、「o」と「u」が順に混じった１つの母音だと考えたほうがよいかもしれない。

そういったこともあり、日本語のどの単語、あるいは、同じ単語でもどのような音声に、「a」と「i」として、「o」と「u」として母音が使われているのか、「ai」や「ou」という母音が使われているのか分からないため、今のところ（いずれは自動的に判別できるようなアルゴリズムが公開されるかもしれないが)、音声を聞きながら、すべての音声にたいして「ai」で１つに区分するのか「a」と「i」で２つに区分するのかを吟味していかねばならない。しかしながら、「はいそうですね」という音声を、「は（ha)」や「い（i)」、「そ(so)」、「う（u)」という仮名文字単位で区切って分析していくほうが、日本人にとって直観的に分かりやすいかもしれず、また、筆者は日本語の文字表記につかわれるローマ字や仮名文字などの単位で区切って分析を進めていきたいと考えた。そのため、今回の分析では筆者は「ai」を１つの母音として見なして分析するのは避けることにした。

今回の分析では、暫定的に筆者は抽出されたピッチを参考にしつつ、実際に音声を聞くことによって、「ai」と「ou」を区切ることにした。具体的には、ピッチは図でドットとして表現されているが、そのドットの集合は連続的で、山や谷のような曲線に近似できるので、この山の頂点を目安にして、かつ、音声の「a」と「i」のどちらがよく聞こえるか、「o」と「u」のどちらがよく聞こえるかなどを目安にして、暫定的に「a」と「i」、「o」と「u」を区切ることにした。城生（2001）は、こういった音声を音圧の変化を目安にして区切っている。

また、「ね」に関しては、うえであげた「ai」や「ou」とは異なり、分析者１名は音韻論的には「n」であろう部分では、「n」しか聞こえなかったが、音韻論的に「e」であろう

部分では、「e」が聞こえるだけではなく、はじめのほうでは、「n」がかなり混じって聞こえ、時間が経つにつれて「n」の混入が少なくなって聞こえた。今回の分析では、「n」と「e」で区切ることが妥当かどうかはべつとして、暫定的に区切ることは困難ではなかった。なぜなら、「n」しか聞こえない部分があったからであり、「e」が聞こえたところで区切れるからである。

2.5.2. 音声特性変数算出の難しさ

音声特性変数の算出は、発話におけるローマ字単位の境界での、時刻とピッチ、音圧の計測値をもとにおこなった。時間については、境界の始点の時刻と終点の時刻の差が⊿tである。ピッチの変化（⊿p）と音圧の変化（⊿a）も同様にしてもとめた。この方法で、⊿tの算出にかんしては、ローマ字単位だけではなく、仮名文字単位、単語単位、分節単位、文単位についても、まったく問題がない。しかし、⊿p、⊿aについては、問題がしょうずる場合がある。

ピッチと音圧については、発話におけるローマ字単位の境界での始点と終点のあいだの区間で、単調に増加したり、減少したりするとはかぎらない。区間内にピッチや音圧の山や谷が存在する場合がある。この場合にも、区間の始点と終点のピッチや音圧の差によって、⊿p、⊿aをもとめ、これを、ある区間のピッチや音圧の変化の指標とすることができる。今回報告した分析では、この方法で、⊿p、⊿aをもとめた。べつの方法としては、始点と終点の値だけではなく、区間内でのピッチや音圧の値の推移を反映するように⊿pと⊿aをもとめることもできる。区間内での、ピッチや音圧の値の推移を10ミリ秒単位程度でこまかく計測し、この推移への直線での近似をもとめ、この傾きを⊿pと⊿aとする方法である。こうすれば、区間内の変化そのものの集約量としての⊿pや⊿aをえることができる。区間内の変化の集約量と、始点から終点への変化のどちらのほうが、心理学的・音声学的に適切な指標となるかはわからない。今回は、始点から終点への変化を、ローマ字単位、仮名文字単位、単語単位、分節単位におけるピッチ、音圧の変化の指標としてもちいた。

今回の分析では、ローマ字単位、仮名文字単位、単語単位、分節単位におけるピッチや音圧については、各単位における変化量のみを指標とし、値の水準そのものは指標として算出しなかった。この場合も、区間内での、ピッチや音圧の値の推移を10ミリ秒単位程度でこまかく計測すれば、その平均によって、ある区間のピッチや音圧の値の集約量とすることができる。しかし、今回は、こうしたこまかい計測をしていないので、区間ごとのピ

ッチや音圧の値の平均は、もとめられなかった。区間の始点と終点の値の平均をとる方法もかんがえられるが、区間内に山や谷がある場合には、実際の水準とずれた値がでてしまう。したがって、ローマ字単位から分節単位まで、区間の両端の値の平均でピッチや音圧の水準をあらわすことはしなかった。一方、文全体については、文にふくまれるローマ字単位の両端の値の平均としてピッチや音圧の水準をもとめた。文全体についてなら、ローマ字単位の両端の値の計測が、文全体の平均をもとめるのに十分なこまかさになるからである。今回の分析ではおこなわなかったが、単語単位や分節単位については、同様にして、ピッチや音圧の水準を算出してもよかったかもしれない。

おわりに

本論文のはじめでのべたように、音声の印象と音声特性の関連の研究がむつかしいのは、音声の印象そのものの基準も、関連する音声特性の定式化も、まだかなりの程度未確立な状態にあるからである。本論文の探索的な分析の結果から、どんな方向性をよみとることができるだろうか。

まず、１.の因子分析の結果として出てきた、音声の印象の四因子だが、これはSD法（Semantic Differential）で一般的に見出されるとされる印象の三因子と対応している。

（１）自信因子　　力量（Potency）
（２）友好因子　　評価（Evaluation）
（３）軽快さ因子　　活動性１（Activity１）
（４）落ち着き因子　　活動性２（Activity２）

因子間相関は、自信因子が独立で、活動性にかんするふたつの小さな因子は評価と正の相関をもっている。活動性因子は、評価がくわわったため、活動性にかんするふたつの因子にわかれたとかんがえられる。

SD法における、評価、力量、活動性の三因子は、経験的なものである。なぜこれらの因子が一般的か、理論的な説明で定説といえるものはない。SD法における印象には、より一般的な感覚的な印象や態度もふくまれているが、感情的な印象がコアとなっている。感情の次元説でいうと、活動性はArousal Levelとの、評価はValenceとの対応が基本である。感情の構成要素説からいうと、活動性はNovelty Checkとの、評価はValence Checkとの対応が、力量はCoping Potentialとの対応が基本である。感情の構成要素説における、既存の目標や社会的な規範との整合性のCheckはValence Checkよりもより複雑な認知的な過程をふ

くみ、評価と関連している。本論文では、悲しみ、怒り、喜びなどの基本感情が、音声の印象の四因子によって区別されることがしめされた。(感情の次元説、構成要素説、基本感情説については、つぎの論文でくわしくのべる。)

Caffi,C. and Janny,R,W. (1994) は、言語における感情的な側面の伝達についての諸理論をSD法における、評価、力量、活動性の三因子にてらして整理している。ここには、感情理論であつかわれるような狭義の感情だけでなく、態度などの複雑な要因もふくまれている。

三因子のなかでもっとも単純なのは、活動性である。本研究では、活動性因子は、評価がくわわったため、ふたつにわかれたが、中心となるのはArousalである。音声では、高いArousalは、一般的に、高いピッチ、速い発話、強い音圧とおおきな変動で表現される。これは、興奮するときのはなしかたで常識的だが、これまでの研究で一致して確認されている (Scherer,K,R.1986, Murray, I, R.and Arnott,J,L.1993)。今回のデータでも、表2.4.1.などで、Arousal のとくに高い楽しさとArousalの低い不安 (今回の音声には、図1.6.の第2象限にあたるArousalのとくにひくい音声はふくまれていなかった。) などを比較してみると、従来の研究と一致した結果がしめされていることがわかる。

評価と力量はより複雑である。評価は、環境刺激にたいする生得的なValence Checkだけでなく、内的な目標や社会的な基準との合致などもふくまれる。また評定対象と評定者に応じて、活動性や力量などの因子との相関をもつなどして、さまざまなタイプの評価的因子に分かれることもある。力量は、今回、独立のおおきな因子として出てきたが、言語的コミュニケーションと関連してもっとも複雑である。Caffi,C. and Janny,R,W. (1994) は、確信の有無、主張のつよさ、明確さ、関与度、など言語に関連した種々の要因が力量と関連するとしている。また、活動性の表現は、おもに内的状態の表出だが、評価と力量については、相手へのはたらきかけとしての表現の側面も重要になってくる。言語的コミュニケーションにおける感情表現についての、詳細な検討はつぎの論文でおこなうことにする。

2.でおこなった、音声の印象と音声特性との関連の分析で、しめされたもっとも重要な結果は、自信、友好、軽快さ、落ち着きなどの印象と音声特性の関連が、特定の言語内容にかたよってみられたことである。詳細は2.でのべたが、ごく、おおざっぱにいうと、たとえば、自信因子は最初の「はい」や「です」のピッチや音圧の変化やその間の長さ (sp2) などと関連しているが、友好因子は最後の「ね」の母音部分のピッチ変化と関連しており、軽快さは「そう」の「そ」の部分の子音の音圧変化と関連しており、落ち着き

は「ですね」のピッチの変化と関連しているなどである。今回の研究は、ひとりの話者による「はいそうですね」だけの分析で、各レベルの言語要素のピッチや音圧の水準自体もパラメーターとしてあつかっていないので、具体的な結果について一般化するのはむつかしい。しかし、音声の印象が発話全体の音声特性ではなく、特定の言語内容とむすびついた音声特性によって、になわれていることは、一般的な結果としていえそうである。これは、各レベルの言語要素ごとに音声特性を計測するという、かなりの手間を要する作業をつうじて、はじめてしめされたことである。

文字におきかえられるような言語の分節化された要素と、発話全体にわたるピッチや音圧や音質などの超分節的な音声特性との関連をどうとらえるかが問題である。言語学では、ピッチや強さなどのアクセントをあつかうが、文末などのアクセントを上昇か下降かなどのタイプに分類し、主張、疑問など、これもカテゴリー化された意味とどう対応しているかを定式化してきた。一方、心理学的研究では、超分節的な音声特性を連続量としてあつかい、分節化される言語要素の内容との関連は、あまり問題にしてこなかった。

Scherer,K,R.,Ladd,D,R. and Silverman,E,A.（1984）は、超分節的な音声特性を言語内容から独立なチャンネルとしてあつかう心理学的アプローチの妥当性を検討している。方法は、音声を文字化したり、音声にフィルターをかけたり、逆転再生したり、ばらしてつなげたりと、種々の加工をくわえて種々の印象を評定するものである。全体としては、音質、ピッチや音圧の平均、標準偏差などの超分節的な音声特性からだけでも、印象はかなりの程度つたわるとして、心理学的アプローチには一定の妥当性があるとしている。例外は、ピッチの輪郭の上昇・下降の別と、Wh疑問かYes/No疑問かといった疑問文のタイプが相互作用して、挑戦的か、愛想がよいか、丁寧かなどといった印象に影響することである。ふつう言語学では測定した印象との相関は問題にしない。しかし、ピッチの輪郭の分類は言語学で一般的なので、Scherer,K,R.たちの結論は、音声についてのある種の印象では音質、ピッチや音圧の平均、標準偏差などといった心理学的変数だけではなく、ピッチの輪郭と言語学的な文のタイプの分類の交互作用も考慮すべしとなる。（ピッチの輪郭は、心理学的研究がピッチ変化などの連続変量としてあつかい、言語学的研究がピッチの輪郭のタイプとしてあつかうなどといったちがいはあるが、ともにあつかってきた音声特性である）。

本論文でえられた結果は、超分節的な音声特性と言語内容の関連を、疑問文のタイプなどの言語学的な分類とピッチの輪郭の交互作用などといった、部分的な限定としてではなく、もっと一般的なものとしてあつかうべき必要があることを示唆している。今回の分析結果からだけでは、音声特性と交互作用するのが、発話の位置なのか、間投詞や終助詞な

どの品詞の種類なのか、摩擦音か鼻音かといった言語要素の音声学的なちがいなのか、言語要素の文中での機能的な役割なのか、などなど、これらの要因の検討が必要なことはいえても、具体的にこうだろうというのはむつかしい。分節的な言語内容と超分節的な音声特性がどう関連して印象をつたえているかの解明は、かなり複雑な課題になりそうである。心理学や言語学などでのこれまでの研究をレビューし、言葉の音声特性と印象の関係を研究するための枠組みをどう定式化できるかについての議論は、つぎの論文でおこないたい。

引用文献

Banse, R. and Scherer, K, R. 1996 Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*. 70(3),614-636.

Caffi,C. and Janny,R,W. 1994 Toward a pragmatics of emotive communication. *Journal of Pragmatics*. 22, 325-373.

Ekman, P. 1985 *Telling lies*. Norton & Company. (工藤他訳「暴かれる嘘」誠信書房)

遠藤利彦 2000 表情を解体する——構成要素説から見る表情の本性——. 心理学評論, 43 (2), 177-198.

Fernald, A. 1992 Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. in Barkow,J,H. and Cosmides, L, Ed. *The adapted mind: Evolutionary psychology and the generation of culture*. Oxford University Press.

井村恒郎 1984「分裂病家族の研究」みすず書房.

城生佰太郎 2001 音声研究の方法 城生佰太郎編「コンピュータ音声学」おうふう 所収.

神田真澄・五味聖子 1999 音声の印象について. 関西大学社会学部卒業論文.

Marks, L, E. 1975 On colored-hearing synesthesia: Cross-modal translations of sensory dimensions. *Psychological Bulletin*. 82(3), 303-331.

Moore, B. C. J 1989 *An introduction to the psychology of hearing*, 3rd Ed. Academic press. (大串健吾監訳「聴覚心理学概論」誠信書房)

向井秀一・山本雅規 2000 人間の感性について——音声による人間の印象認知に関する研究——. 関西大学社会学部卒業論文.

向井美幸・目加田慶人・春日正男・松本修一・小池淳 2000 感情を含む音声信号の特徴解析について. 信学技法IEICE, 17-22.

森山剛・斎藤英雄・小沢慎治 1999 音声における感情表現語と感情表現パラメータの対応付け. 電子情報通信学会論文誌. D-II Vol.J82-D-II No.4, 703-711.

Murray, I, R and Arnott, J, L. 1993 Toward the simulation of emotion in synthetic speech : A review of the literature on human vocal emotion. *Journal of the Acoustical Society of America*. 93(2), 1097-1108.

Reeves, B. and Nass, C. 1996. *The Media Equation:how people treat computers, television, and new media like real people and places*. CSLI Publications. (細馬訳「人はなぜコンピュータを人間として扱うか——メディア等式の心理学——」翔泳社)

斎藤収三・中田和男 1981 「音声情報処理の基礎」オーム社.

佐藤綾子 1995 「自分をどう表現するか：パフォーマンス学入門」講談社現代新書.

佐藤和之 2001 日本語音声の生成と音響特徴 城生佰太郎編 「コンピュータ音声学」 おうふう 所収.

Scherer, K, R. 1986 Vocal affect expression : A review and a model for future research. *Psychological Bulletin*. 99(2), 143-165.

Scherer,K,R., Ladd,D,R. and Silverman,E,A. 1984 Vocal cues to speaker affect: Testing two models. *Journal of the Acoustical Society of America*. 76(5), 1346-1356.

Picard, R, W. 1997 *Affective computing*. MIT Press.

杉藤美代子（監修・著）1999「杉スピーチアナライザー ガイドブック」アニモ.

Sullivan, H, S. 1954 *The psychiatric interview. Norton*.（中井久夫他訳「精神医学的面接」みすず書房）

―― 2001.12.21 受稿 ――