

音声の発話速度と休止時間が話者の性格印象と 自然なわかりやすさに与える影響

内 田 照 久*

音声の中の実音声区間と休止区間の時間配分と、話者の性格印象、音声の自然性との関係を検討した。パラグラフ・レベルの音声材料を用いて、大学生581名を被験者として聴覚実験を行った。実音声時間長は、性格印象、自然性に特徴的な影響を与えていた。休止時間長の違いによる影響は比較的小さかった。実音声部の発話速度の変化に伴う性格印象の変化パターンは、Big Fiveの特性ごとに固有の2次回帰式で近似できた。勤勉性は速い発話で評価が高く、遅くなると急峻に低下した。協調性では逆に、速い発話で評価が低く、遅い発話で高かった。外向性と経験への開放性はやや速い発話で評価が高く、より速くても遅くても下降した。情緒不安定性には影響は見られなかった。自然性はオリジナルの時間長で評価が最も高く、より速くても遅くても左右対称の形状で低下した。これらのことから音声の変換に伴う自然性の低下の影響は皆無とは言えないまでも、性格特性ごとの独自性が確認された。この結果は、性格印象全体の複雑な変化を捉えるために、音声の時間構造の変化に応じて性格特性ごとに特性値を近似関数で推定し、その推定値を組合せて全体像を捉えるというモデルを支持するものであった。

キーワード： 音声言語教育、発話速度、ポーズ、印象形成、性格特性5因子モデル

問題と目的

言語教育における音声言語重視への転換

近年、わが国の言語教育においては、歴史的な方向転換が進んでいる。すなわち、文章を読んで理解することを中心とした従来型の教育から、対話を行うためのコミュニケーション能力重視への方針の転換である。具体的な動向としては、小中学校の国語科にも「音声言語」の項目が新規に導入され、音声コミュニケーション能力の育成を教育の中に位置づけた営みが行われるようになってきた(瀬川, 1995)。

音声にかかわる心理学の研究の役割とその困難性

しかし、この音声コミュニケーション教育の基盤を成すべき、広く音声に関わる知見の蓄積に関しては、これまでのところ必ずしも十全とは言えない。

従来、音声が担っている情報は、いわゆる文字に還元できる言語情報と、それ以外の韻律的特徴 (prosodic feature) とに分けて考えられてきた。音声学を含めた言語学は、言語規則の確立を大局的な目的としているため、韻律的な音響特徴を、純粋な言語情報を伴わない事象として扱い、これまで重要視してこなかった。

しかし、文字の形では捨象されてしまうものの、時々

刻々と変容する音声の韻律的な特徴の中には、話し手の人柄や、その人の今の気持ちや意思といった、話し手に関わる様々な情報が織り込まれ、音声を介して伝達されている。この話し手自身に関わる情報が、音声中にどのような形で格納されており、聞き手がその特徴をどのように認知しているのかを明らかにしていくことは、まさに心理学の範疇であり、その知見の蓄積が教育にも生かされていくべきものである。

さて、これまでのわが国の心理学においては、音声を対象とした研究は、視覚情報を対象としたものと比べると格段に少ない。その背景には、声の質は変えずに抑揚だけをコントロールするとか、音声の時間的構造のみを任意に操作するといった、音声の韻律的特徴の基本的な要素を操作する上での技術的な難しさがその根底にあったためと考えられる。

音声情報処理技術を利用した心理学研究の進展

一方、音声工学の領域では、計算機の飛躍的な性能の向上に伴って、音声情報処理技術もめざましい進歩を遂げてきている。最近、その技術的成果を応用して、音声に関わる心理学的な研究が進められるようになってきた。

たとえば、内田 (2000) や内田 (2002) は、音声の発話速度と話者の性格印象との関係について、聴覚実験の結果を基にモデル化を試みている。そこでは、発話速度の操作に、Malah, Crochiere, & Cox (1981) の時間

* 大学入試センター研究開発部
〒153-8501 東京都目黒区駒場2-19-23
uchida@rd.dnc.ac.jp

領域調波構造伸縮 (TDHS) を発展させた, 森田・板倉 (1986) のポイント移動量制御による重複加算法 (PICOLA) を用いている。

また, その話者の性格印象の評価にあたっては, パーソナリティ心理学の領域で1990年以降, 劇的な進展を遂げた性格特性5因子論 (Goldberg, 1990; John, 1990) を基盤にした測定を行っている。そこでは, 性格特性5因子論の展開の中で見出された Big Five を, 対人認知場面での他者認知の次元としてとらえ直し, 性格印象測定の評価軸として用いている。

この Big Five とは, 外向性 (extraversion, E), 情緒不安定性 (neuroticism, N), 経験への開放性 (openness to experience, O), 勤勉性 (conscientiousness, C), 協調性 (agreeableness, A), などとよばれるものである。一般に性格や人柄の印象といったものは, 抽象度が高く, ややもすると定義が曖昧になってしまいがちな概念であるが, この Big Five をフレームワークとして用いることで, これらの研究では信頼性の高い測定を実現している。また, その結果から, 音声から想起される話者の性格印象は, Big Five の性格特性ごとに独立に認知されることが見出されてきた。

さらに Uchida (2003) は, 音声の発話速度に加えて, 声の高さ, すなわち基本周波数と話者の性格印象との関係についても検討している。そこでは, 河原 (1998), 及び Kawahara, Masuda-Katsuse, and de Cheveigne (1999) による, 高品質音声分析変換合成アルゴリズム (STRAIGHT) を用いて, 基本周波数や発話速度の操作を行っている。この研究でも性格特性5因子論の観点から話者の性格印象の測定を行い, 声の高さ・発話速度と性格印象との関係のモデル化を試みている。

これまでの聴覚実験による研究の問題点

この内田 (2000) や内田 (2002), 及び Uchida (2003) の一連の研究には, 大きく2つの問題点がある。

まず第1点は, これらの聴覚実験で用いられている音声資料が, いずれも1-2秒程度の短い単文であることである。対人認知場面での印象形成を考えた場合, たとえ電話のような音声コミュニケーション場面を想定したとしても, 1-2秒の音声ではあまりにも情報が不足していると思われる。生態学的な妥当性を検討するためには, 最低でも一定の意味内容を構成し得るパラグラフ・レベルの音声資料を用いることが望まれる。

第2点目は, 変換音声の自然性の問題である。たとえ最新の優れた音声情報処理方式を採用したにせよ, 計算機上での変換に伴う音質の劣化や自然性の低下は

避けられない。したがって, 聴覚実験で得られる話者の性格印象評価に, それらの影響が重畳している可能性は払拭できない。そこで, 聴覚実験で変換音声を用いる場合には, その音声の自然性についても, 主観的な評価に基づいたデータを同時に収集しておく必要がある。

実音声区間と休止区間を操作した実験の必要性

上記の問題点をふまえて, 音声の韻律的特徴と話者の性格印象との関係を検討することを考える。実験で用いる音声試料を30-45秒程度のパラグラフ・レベルのものにすることを考えると, 従来の単文を用いた実験では問題にならなかった新たな検討課題が生ずる。それはすなわち, 句や節の区切れ, また息継ぎの箇所などで観測される無音休止区間の扱いである。

杉藤 (1997) によれば, この休止区間のポーズは, 音声の意味内容の理解や記憶に影響を与えるとされる。したがって, 音声から想起される話者の性格印象を検討する際にも, 実音声区間と休止区間の時間配分に配慮した実験のデザインがもとめられよう。

本報告の実験では, 音声の韻律的特徴が話者の性格印象に与える影響を検討する上で, 音声の時間的側面に焦点を絞る。そして, パラグラフ・レベルの音声を材料とし, 実音声区間と休止区間の時間配分の操作を行う。そして, 話者の性格印象の評価に加えて, 音声の自然性や, 発話速度感に関する主観的評価の測定も行う。その結果を基に, 音声コミュニケーション時の対人認知場面における印象形成の特徴の記述を目指すものとする。

方 法

被験者

3つの国立大学 (静岡大学・東京大学・名古屋大学) の文科系・理科系の1年生が参加した。男性444名, 女性137名の計581名 (18-23才) であった。文科系・理科系の被験者の割合はほぼ半数ずつである。実験は大学ごとに行なわれ, 実験実施1回あたりの被験者数はおおよそ100名ずつであった。

実験デザイン

原音声データ 実験に用いられた原音声は, 日本留学試験の中の日本語リスニング・テスト問題である聴解・聴読解問題の音声部分である (日本国際教育協会 2001)。発話者は, 発声の訓練を受けた日本語話者で, 男声1名 (M1), 女声1名 (F1) であった。発話内容は, 単独の発話による, 4つのスクリプト (S1-S4) である。男声と女声は, それぞれ2つのスクリプトを担当

している (M1-S1, M1-S2, F1-S3, F1-S4)。スクリプトは181-242文字相当で、全音声時間は30.7-44.6 Sec.であった。これらの音声について、量子化16bit, サンプリング周波数44.1kHzの音声ファイルとして保存されたものを原音声データとして用いた。

原音声の中の実音声区間と休止区間の区分について、具体例 (M1-S1) を FIGURE 1 に示す。音声中の200ms以上の無音区間を、句節区切りによって生ずる休止時間と定義した。その上で、4つの原音声について、音声波形上で実音声区間と休止区間を特定した。全音声時間の中で、休止区間の占める割合は、22.3-29.4%であった。

発話速度ブロックと実験条件の設定 音声全体についての発話速度の影響を検討するために、発話速度を3段階で操作する発話速度ブロックを設定した。具体的には、原音声の総時間長を Fast (75%), Original (100%), Slow (133.33%) の3段階で一様に変換する発話速度ブロックを配置した。

次に、休止時間の操作にあたり、「可変時間長条件」と「固定時間長条件」の2つの実験条件を設定した。

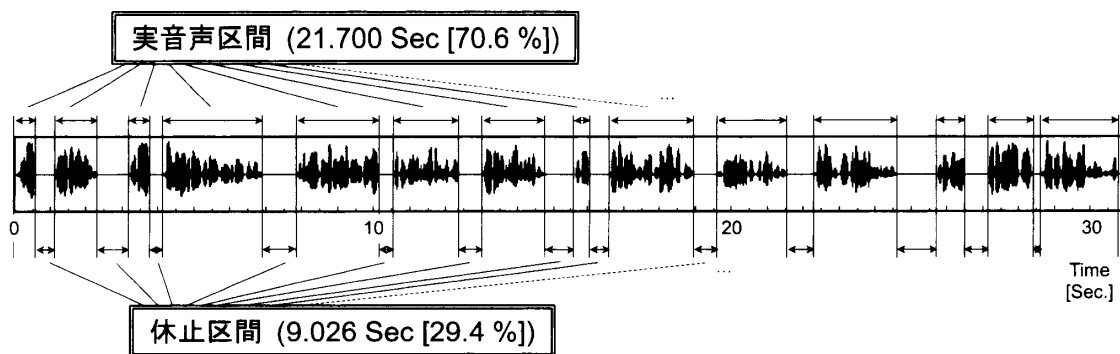
一般に学会発表や講演、各種会合でのスピーチでは一定の制限時間が存在する。その場合、設定時間内での実音声時間と休止時間の配分の調整が重要となる。このようなターゲット時間中の実音声時間と休止時間

の配分比を直接対象とする条件が「固定時間長条件」である。しかし、この条件では、総時間長が固定されているために、休止時間を伸長した場合には必然的に実音声時間が圧縮される。そのため、もしも何らかの影響が観測されたとしても、それが実音声時間の影響なのか、休止時間の影響なのかを特定することができない。

そこで実音声時間の影響と休止時間の影響を切り分けて検討するため、いずれかの時間長を固定し、もう一方の時間長だけを操作した条件で実験を行う。これまでの先行研究で実音声時間については検討がなされてきているので、今回は休止時間を操作して検討する。そのようにして設定された条件が、「可変時間長条件」である。

可変時間長条件 可変時間長条件は、休止時間の操作に伴って、実験刺激の総時間長が発話速度ブロック内で異なったものになる条件である。この条件では、各発話速度ブロックの一様変換後の休止時間を新たな基準とし、その休止区間の時間長を Short 2 (75%), Short 1 (86.60%), Equal (100%), Long 1 (115.47%), Long 2 (133.33%) の5段階で変換する。この条件では、発話ブロック内での実音声区間の時間長は、常に一定である。この可変時間長条件の実験操作の概念図を FIGURE 2 に示す。

【原音声データ (M1-S1) の音声波形】



【発話スクリプト [6.61 mora/Sec.] ▽: 休止区間(200ms 以上)】

それでは、▽₁今日の講義は以上です。▽₂それから、▽₃来週の試験についての注意事項を言っておきます。▽₄私の文化人類学の試験は、▽₅テキストの持ち込みは可としますが、▽₆ノートの持ち込みは認めません。▽₇また、▽₈携帯電話については、時計代わりにあっても▽₉机の上に置いておくことは出来ません。▽₁₀電源を切って、かばんの中に入れてください。▽₁₁試験中に▽₁₂携帯電話が鳴ったら、▽₁₃他の人に大変迷惑になりますからね。

FIGURE 1 原音声データの音声波形と発話スクリプトの例

【可変時間長条件】 発話速度の一樣変換後に休止時間を操作: 5段階
($q = 75\%, 86.6\%, 100\%, 115.47\%, 133.33\%$)

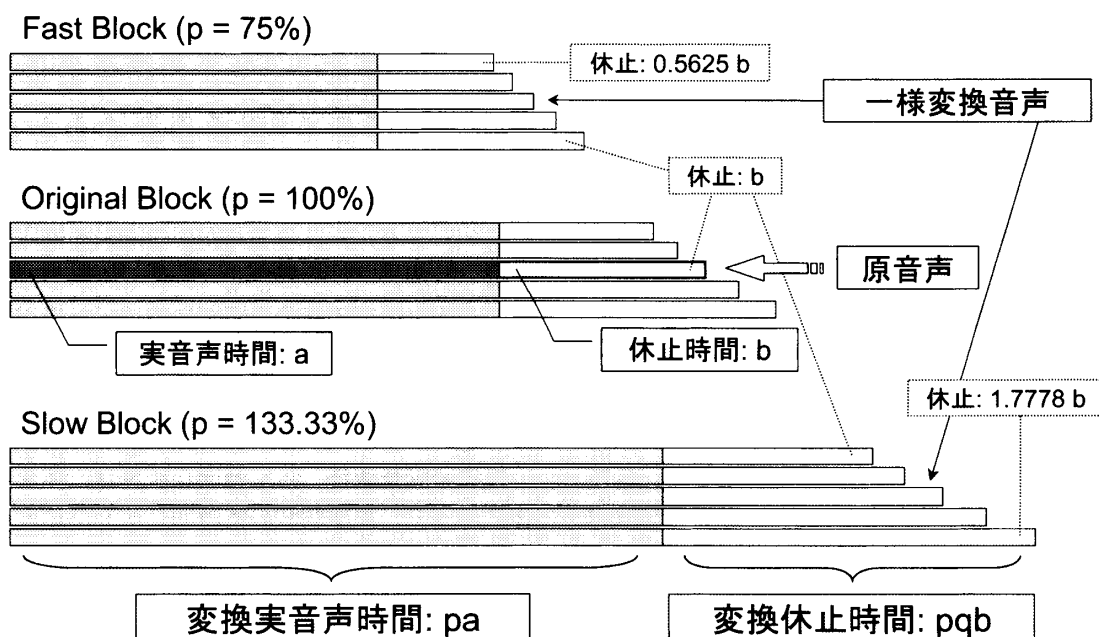


FIGURE 2 可変時間長条件における実験音声刺激の作成

(a : 原音声内の実音声時間 [Sec.], b : 原音声内の休止時間 [Sec.],

p : 発話速度ブロック操作のための一樣変換比率 [%], q : 休止時間操作のための発話速度ブロック内での変換比率 [%])

【固定時間長条件】 1. 可変時間長条件と同様に休止時間を操作
2. 一樣変換時間長の保持のため実音声時間を伸縮

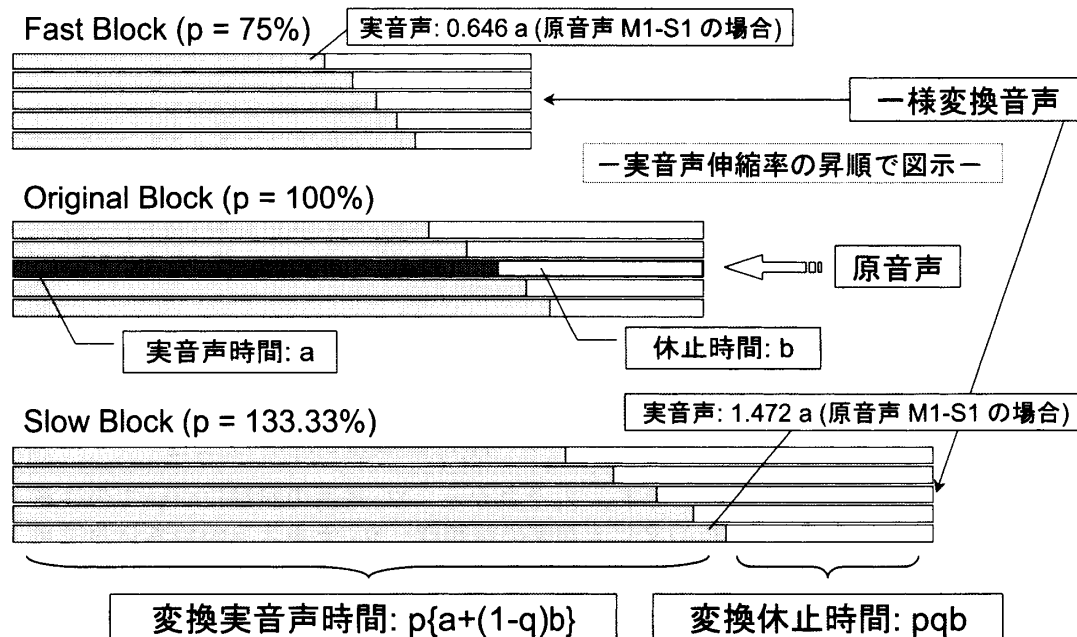


FIGURE 3 固定時間長条件における実験音声刺激の作成

固定時間長条件 固定時間長条件は、一様変換後の発話ブロックごとでの総時間長は固定したまま、先般と同様に休止時間を操作する条件である。操作に際しては、総時間長の保持のため、実音声区間は適応的に時間伸縮率を調整して変換を行う。したがって、休止区間の時間長操作に伴い、各発話速度ブロック内では、実音声区間の時間長は系統的に異なったものになる。この固定時間長条件の実験操作の概念図を FIGURE 3 に示す。

実験音声刺激の作成 原音声データを基に、各条件ごとに算出される実音声区間と休止区間の時間伸縮率に基づいて、音声の時間長の変換処理を行い、実験音声刺激を作成した。具体的には、4 [原音声]×3 [発話速度ブロック]×2 [可変・固定時間長条件]×5 [休止伸縮段階]=計120刺激を作成した。実音声区間、休止区間の音声時間長の変換にあたっては、内田(1998)の PICOLA plus 2を用いた。また、これらの音声刺激については予備実験を行い、人間の発話音声として自然なものであることを確認した。

実験実施ブロックの割当て 作成された実験刺激を15刺激ずつ、8つの実験実施ブロックに割当てた。割り当て時の単位は、可変・固定時間長条件の中の同一発話速度ブロック内に含まれる5つの刺激を組にしたものを最小単位とした。各実験実施ブロックには、男声・女声、可変時間長条件・固定時間長条件、また3段階の発話速度ブロックが必ず含まれるように配置した。その上で、各実施ブロックに割り当てられた実験刺激について、提示順序をランダムにした実験用のコンパクト・ディスク・レコーダブル (Compact Disk Recordable, CD-R) を作成した。CD-R の作成にあたっては、音楽 CD 用の CD-DA フォーマットで記録した。

次に、被験者も8つの実験実施ブロックにランダムに割当てた。実施ブロックあたり、72名程度ずつが割振られた。そして、各実施ブロックの被験者を、さらに2等分し、第1群は話者の性格印象の評定を、もう一方の第2群は音声の自然性などを評定するように配置した。

評定項目 話者の性格印象の評価には、和田(1996)の Big Five Scale を基に、萩生田・繁樹(1995)が項目を精選した20項目の短縮版を使用した (TABLE 1)。

また、音声の自然性の評価については、音声の自然性やわかりやすさ、発話速度感や早口感の評定項目をはじめとして、話者の信頼感や親しみやすさに関する評定項目を新たに作成して用いた (TABLE 2)。

この2系列の評定項目について、それぞれ20項目の

TABLE 1 Big Five Scale短縮版の評定項目

性格特性	評定項目(各4項目)
外向性：E	話し好き、無口な*、陽気な、外向的
情緒不安定性：N	悩みがち、不安になりやすい、心配性、気苦労の多い
経験への開放性：O	独創的な、進歩的、洞察力のある、想像力に富んだ
勤勉性：C	いい加減な*、ルーズな*、怠惰な*、計画性のある
協調性：A	温和な、寛大な、親切的な、協力的な

*逆転項目

TABLE 2 音声の自然性・発話速度感・話者の信頼感・親しみやすさの評定項目

評価観点	評定項目
自然性・わかりやすさ	自然な、わかりやすい、はっきりした、間の取り方のよい、ごこちない*、聞き取りにくい*、話のリズムの悪い、歯切れのよい
発話速度感・早口感	早口な、せわしげな、間の取り方の長い*、のんきな*
話者の信頼感	頼りがいのある、威圧的な、従順な*、依存的な*
話者の親しみやすさ	親しみやすい、好意的な、無愛想な*、敵意のある*

*逆転項目

順序をランダムにした評定用紙を準備した。

手続き

被験者に携帯型 CD プレーヤ (AIWA: XP-V30S) を配布して実験を実施した。実験では、1 kHz 純音の信号音の後、イヤークラスパーから音声刺激を単独で提示した。被験者は音声を聞き終わってから、各評定項目の内容に即して、“全くあてはまらない”から“非常によくあてはまる”まで、10段階 (1-10) で評定した。1試行では、1つの音声刺激につき20項目の評定を行い、制限時間は60秒であった。被験者1人あたりの試行数は15試行で、実験全体の所要時間は約20分であった。

分析と結果

可変時間長条件の分析と結果

話者の性格印象の特性値 5つの性格特性について、特性ごとの合成得点を各4項目の評定値の加算によってもとめた。逆転項目では、評定値を反転させて加算した。さらに、音声刺激ごとに各合成得点の平均を算出し、それらをその音声刺激から想起される話者の性格特性印象の特性値とした。この特性値を性格特性ごとに、発話速度ブロック別、休止区間の操作段階別に平均した結果を TABLE 3 に示す。

話者の性格印象に影響を与える要因 性格印象に影響

TABLE 3 可変時間長条件での発話速度ブロック・休止区間伸縮段階別の性格印象特性値の平均 (1-10 点換算)

性格特性	発話速度ブロック	休止区間の時間伸縮段階と特性値の平均										発話速度ブロック ごとの平均
		Fast(75%)	Short2	Short1	Equal	Long1	Long2					
		Original(100%)			Short2	Short1	Equal	Long1	Long2			
		Slow(133.3%)					Short2	Short1	Equal	Long1	Long2	
		休止時間伸縮率	56.3%	65.0%	75.0%	86.6%	100.0%	115.5%	133.3%	154.0%	177.8%	
外向性：E	Fast	5.43	5.65	5.92	5.63	5.59					5.65	
	Original			6.04	5.90	5.49	5.89	5.92			5.85	
	Slow					5.09	4.81	4.70	5.00	4.89	4.90	
情緒不安定性：N	Fast	5.25	5.20	4.89	5.02	5.21					5.11	
	Original			4.80	4.92	4.81	4.90	4.71			4.83	
	Slow					5.26	5.10	5.24	5.20	5.27	5.21	
経験への開放性：O	Fast	5.19	5.16	5.22	5.20	5.24					5.20	
	Original			5.48	5.38	4.92	5.35	5.52			5.33	
	Slow					4.50	4.22	4.21	4.43	4.42	4.36	
勤勉性：C	Fast	7.46	7.73	7.80	7.69	7.62					7.66	
	Original			7.21	7.20	7.30	7.16	7.45			7.26	
	Slow					5.36	5.00	5.02	5.32	5.10	5.16	
強調性：A	Fast	4.33	4.16	4.34	4.38	4.48					4.34	
	Original			6.34	6.45	5.87	6.56	6.15			6.27	
	Slow					6.11	5.97	5.93	6.16	6.15	6.06	

響を与えている要因を整理するため、5つの特性ごとに2要因の分散分析を行った。分析にあたっては、発話速度ブロック要因 (3水準: Fast, Original, Slow), 休止区間要因 (5水準: Short 2, Short 1, Equal, Long 1, Long 2), 及び発話速度ブロック要因と休止区間要因の交互作用を分析モデルに設定して行った。

結果として、情緒不安定性を除いた性格特性においては、いずれも発話速度ブロック要因の主効果が有意であり、音声全体の発話速度が変わると話者の性格印象が変化することが示された (E: $F_{(2,45)}=25.71, p<.0001$, N: $F_{(2,45)}=1.96, ns$, O: $F_{(2,45)}=31.61, p<.0001$, C: $F_{(2,45)}=123.91, p<.0001$, A: $F_{(2,45)}=59.74, p<.0001$)。

一方、休止区間要因の主効果については、どの性格特性においても有意差は見出されなかった (E: $F_{(4,45)}=0.22, ns$, N: $F_{(4,45)}=0.06, ns$, O: $F_{(4,45)}=0.88, ns$, C: $F_{(4,45)}=0.05, ns$, A: $F_{(4,45)}=0.45, ns$)。また、発話速度ブロック要因と休止区間要因の間にも有意な交互作用は見られなかった (E: $F_{(8,45)}=0.89, ns$, N: $F_{(8,45)}=0.14, ns$, O: $F_{(8,45)}=0.39, ns$, C: $F_{(8,45)}=0.37, ns$, A: $F_{(8,45)}=0.30, ns$)。

したがって、本実験での休止時間の伸縮範囲内では、休止時間長の違いは、話者の性格印象に必ずしも顕著な影響は与えていないことが示された。

音声の自然性・発話速度感の特性値 音声の自然性や発話速度感に関する評価項目が、それらの評価観点をよく反映しているかについて検討した。ここでは主成分分析を行った。それぞれの音声刺激についての、

各被験者による20項目分の一連の評定結果を1つのオブザベーションとみなして分析した。固有値プロットから、4つの主成分を抽出することとした。中でも、第1固有値と第2固有値が特に高かったので、第1主成分と第2主成分に寄与率の高い項目をまとめ、ここでの尺度として扱うこととした。

第1主成分を「音声の自然性」を示す成分ととらえた。その成分の尺度を構成する項目は、“自然な”、“わかりやすい”、“はっきりした”、“間の取り方のよい”、“ぎこちない*”、“聞き取りにくい*”、“話のリズムの悪い*”の7項目となった。*印は逆転項目である。第2主成分は「発話速度感」としてとらえた。尺度項目は、“早口な”、“せわしげな”、“間の取り方の長い*”、“のんきな*”の4項目となった。

この音声の自然性と発話速度感について、それぞれの合成得点を各項目の評定値の加算によってもとめた。そして、音声刺激ごとに各合成得点の平均を算出し、その音声刺激の自然性と発話速度感の特性値とした。これらを発話速度ブロック別、休止区間の段階別に平均した結果をTABLE 4に示す。

音声の自然性・発話速度感に影響を与える要因 音声の自然性、発話速度感に影響を与えている要因を整理するため、分散分析を行った。発話速度ブロック要因、休止区間要因、及び発話速度ブロック要因と休止区間要因の交互作用について分析した。

結果として、音声の自然性、発話速度感ともに、発

TABLE 4 可変時間長条件での発話速度ブロック・休止区間伸縮段階別の音声の自然性・発話速度感の特性値の平均 (1-10点換算)

評価観点	発話速度ブロック	休止区間の時間伸縮段階と特性値の平均										発話速度ブロック ごとの平均
		Fast(75%)	Short2	Short1	Equal	Long1	Long2					
		Original(100%)			Short2	Short1	Equal	Long1	Long2			
		Slow(133.3%)			Short2	Short1	Equal	Long1	Long2			
音声の自然性	Fast	5.26	5.46	5.68	5.70	5.31						5.48
	Original			6.95	6.96	6.36	6.59	6.64				6.70
	Slow					5.11	4.73	4.55	4.75	4.58		4.75
発話速度感	Fast	8.63	8.81	8.33	8.26	8.27						8.46
	Original			5.76	5.35	5.01	4.80	5.41				5.26
	Slow					3.27	2.92	2.77	2.83	2.67		2.89

話速度ブロック要因の主効果が有意であり、全体の発話速度が変わるといづれもその評価が変わることが示された（自然性： $F_{(2,45)}=57.78, p<.0001$, 発話速度感： $F_{(2,45)}=603.89, p<.0001$ ）。

休止区間要因については、音声の自然性には有意な主効果は見られなかったが、発話速度感に有意水準5%で主効果が見られた（自然性： $F_{(4,45)}=0.48, ns$, 発話速度感： $F_{(4,45)}=1.74, p<.05$ ）。この結果から、本実験の操作範囲内での休止時間長の違いは、音声の自然性には、必ずしも明確な影響を及ぼさないことが示された。一方、発話速度感については、休止区間の時間長の影響が観測された。全般的に休止時間の短い方が、発話速度が速いと評価されているが、逆に最も長い休止時間のものが必ずしも発話速度が遅いと評価されている訳ではなかった。この休止区間要因について、Tukeyの多重比較を行ったところ、休止区間の5段階はいずれも同一のグループ内のものとされ、休止時間の段階別の明確な傾向は検出されなかった。

また、発話速度ブロック要因と休止区間要因の交互作用については、いずれも有意ではなかった（自然性： $F_{(8,45)}=0.69, ns$, 発話速度感： $F_{(8,45)}=0.56, ns$ ）。

固定時間長条件の分析と結果

実音声区間の時間長と話者の性格印象の関係 可変時間長条件の分析の結果、発話ブロック間では話者の性格印象が変化するものの、休止区間の時間長は性格印象には影響を与えていなかった。ここでFIGURE 2と照らし合わせて考えると、実質的には音声の中の実音声区間の時間長が、性格印象に影響を与えていることがわかる。そこで、固定時間長条件では、この実音声区間の時間長に着目し、話者の性格印象との関係を分析する。この条件では、FIGURE 3に見られるように、1つの原音声から作成された15の実験刺激の実音声時間長は、それぞれ連続的に変化している。したがって、

実音声時間長と性格印象との関係については、可変時間長条件よりも細やかにとらえることができると考えてよい。

まず、先般の可変時間長条件と同様にして、5つの性格特性について、音声刺激ごとに特性値をもとめた。そして、さらに各実音声時間長ごとに特性値を平均した結果をFIGURE 4に示す。

次に、実音声時間長と各特性値との関係性を明らかにするため、重回帰分析を行った。内田(2000)、内田(2002)、及びUchida(2003)は、発話速度と性格特性印象の関係については、2次の回帰予測式がよい近似を与えるとしている。そこで、1次の回帰式に加えて、2次の回帰式でも分析を行い、どちらが両者の関係をより適切に記述できるかを比較する。

性格特性印象ごとに、1次予測式、2次予測式で重回帰分析を行った際の情報量基準(AIC)をもとめ、TABLE 5にまとめた。その結果、情緒不安定性以外の性格特性では、いずれも2次の回帰予測式の方がよくフィットしていた。そこで、2次の予測式での重相関係数(R^2)、及びそのモデルの検定結果もTABLE 5に示

TABLE 5 固定時間長条件での実音声区間時間長と性格特性印象の重回帰分析における各次数の回帰予測式でのAIC情報量基準と2次の予測式で重回帰分析を行った際の R^2

	1次式	2次式	2次回帰式 での R^2
外向性：E	84.21	37.99	0.578****
情緒不安定性：N	36.52	37.92	0.104*
経験への開放性：O	74.50	30.46	0.619****
勤勉性：C	74.50	30.46	0.877****
強調性：A	109.90	67.14	0.748****

太字：モデル・フィットのよかったもの

*: $p<.05$, **: $p<.01$, ***: $p<.001$, ****: $p<.0001$

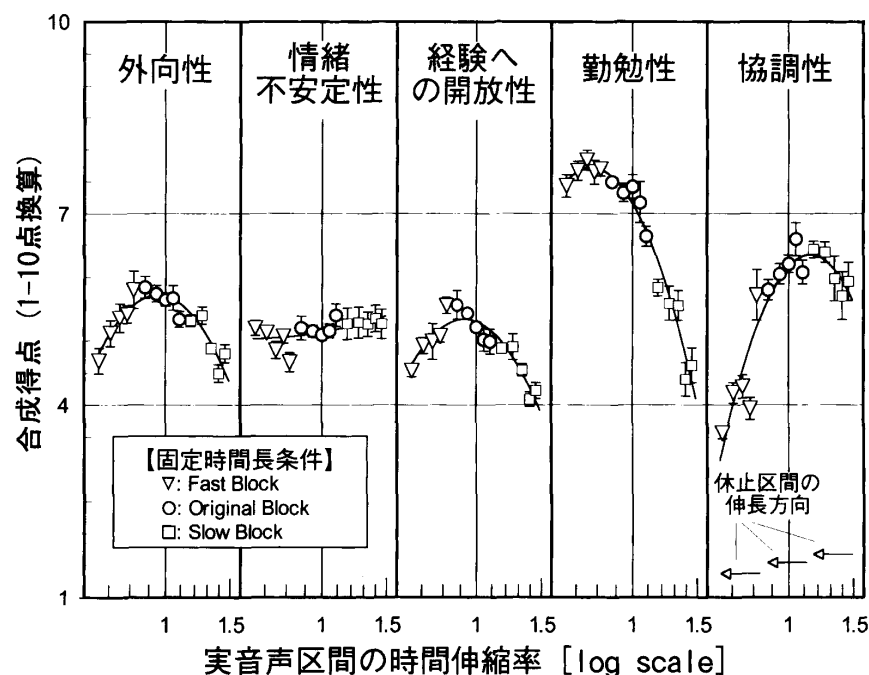


FIGURE 4 実音声区間の時間伸縮率と話者の性格特性印象の平均と S.E.

した。

全体として、実音声時間長と性格特性印象の関係は、2次の予測回帰式での近似が良かったので、FIGURE 4に性格特性ごとの回帰曲線を記した。ここで情緒不安定性は、実音声時間長の変化にほとんど影響を受けておらず、2次式による近似は情報記述の観点では若干冗長となるが、他の特性と同様に扱うこととした。

実音声時間長と自然性・発話速度感の関係 可変時間長条件と同様に、自然性と発話速度感についての特性値をもとめた。そして、各実音声時間長ごとに特性値を平均した結果をFIGURE 5に示す。

次に、実音声時間長と自然性・発話速度感との関係を明らかにするため、重回帰分析を行った。性格特性印象と同様、1次と2次の回帰式で分析を行った。評価観点ごとに重回帰分析でのAICをもとめ、TABLE 6にまとめた。その結果、自然性も発話速度感のいずれも2次回帰式でのフィットがよかった。そこで、2次式での R^2 、モデルの検定結果をTABLE 6に示した。また、FIGURE 5に評価観点ごとに回帰曲線を記した。

総合考察

実音声時間・休止時間と話者の性格印象の関係

本実験の結果から、実音声区間の時間長は話者の性格印象に影響を与えるが、休止区間の時間長の影響は、本実験で扱った範囲では比較的小さいことが示された。

この実音声区間は、実際に声を出して話をしている部分であり、その時間長の伸長はゆっくり発声することに、短縮は早口で話すことに対応する。この実音声部の発話速度と話し手の性格印象には、性格特性ごとに特徴的な関係が見られた。

FIGURE 4を見ると、実音声部の発話速度が変わると各性格特性の印象は、情緒不安定性を除いて、それぞれに逆U字型の変化パターンを示しているのがわかる。そして、2次の回帰予測式による近似曲線は、その変化の概観をよく表現している。性格特性印象ごとに变化パターンの形状は異なり、性格特性によって最もその評価が高くなる発話速度に違いが見られ、またその上昇や下降の幅にも特徴がある。

勤勉性(C)の印象は、速い発話で評価が高く、発話がゆっくりになるとその評価が急激に低下する。一方、協調性(A)の印象は逆に、速い発話では評価が低く、ゆっくりになるにつれて評価が上がり、さらに遅くなっても評価はそれほど低下しない。外向性(E)と経験への開放性(O)の印象は、やや速い発話で評価が最も高く、発話がより速くなっても、遅くなってもなだらかに評価が下降する。情緒不安定性(N)には、発話速度の影響はほとんど見られない。

これらの結果から、話者の性格印象変化のパターンは、特性ごとに2次の関数形で近似でき、各特性の特徴は係数の値の取り方で表現が可能であると考えられ

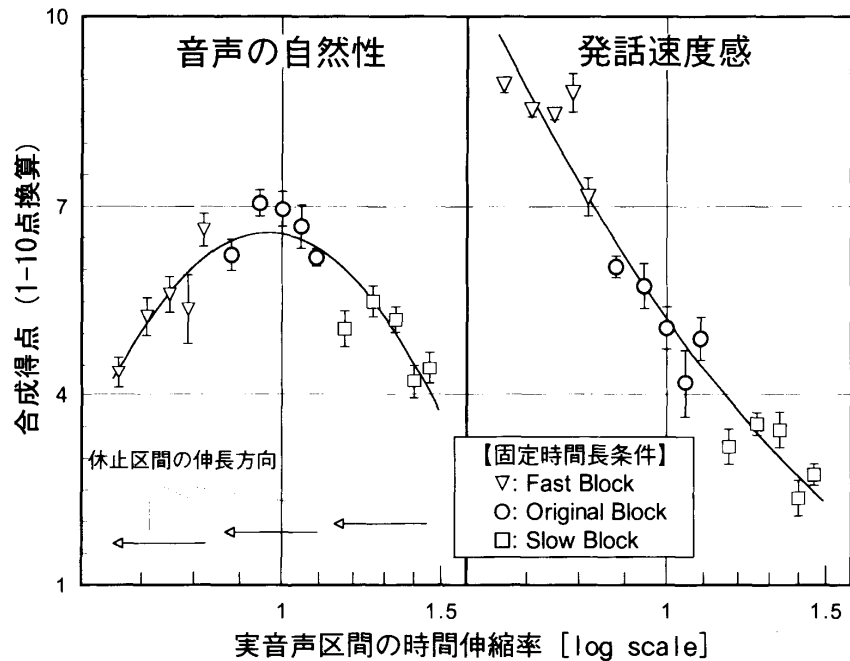


FIGURE 5 実音声区間の時間伸縮率と自然性・発話速度感の平均と S.E.

TABLE 6 固定時間長条件での実音声区間時間長と音声の自然性・発話速度感の重回帰分析における各次数の回帰予測式でのAIC情報量基準と2次の予測式で重回帰分析を行った際の R^2

	1 次式	2 次式	2 次回帰式 での R^2
音声の自然性	239.22	185.36	0.618****
発話速度感	132.78	129.76	0.908****

太字：モデル・フィットがよかったもの

*: $p < .05$, **: $p < .01$, ***: $p < .001$, ****: $p < .0001$

る。したがって、この結果は、性格特性ごとにその近似式を用いて印象評定値を推定し、それらの推定値を組合せることで、性格印象全体の複雑な変化に関するモデル化が可能であるとする、内田 (2002) のまとめを支持し得るものといえよう。なお、本実験の結果は、従来の結果と同様の傾向を示しており、話者の性別や発話内容の偏りによるバイアスのようなものは特に観測されなかった。

さて、内田 (2000) や内田 (2002) では、オリジナルの発話速度の実験刺激において、どの性格特性も最も評価が高かった。そのため、所詮、実験操作を行わない自然な音声は、単に一番良く評価されているに過ぎないのではないかという批判があった。しかし、FIGURE 4 を見ると、性格特性印象ごとに特徴的で、かつ、オリジナルとは異なる時間長の箇所でも最も高い評価がなされているのがわかる。このことは、本実験で従来の実

験刺激よりも長いパラグラフ・レベルの音声資料を用いて、より現実場面に近い状況で実験を行ったこと、また、従来よりも多くの観測点で測定したことによって、その検出が可能になったものと考えられる。

実音声時間・休止時間と音声の自然性・発話速度感の関係

音声の自然性は、全体の発話速度が変わると、その評価が変わることが示された。一方、休止時間長の違いは、音声の自然性にはあまり影響を与えなかった。

そこで実音声時間長に着目して検討する。FIGURE 5 を見ると、音声の自然性は、実音声部の発話速度に応じて、放物線状に評価が変化していた。その評価は、オリジナルの時間長の近傍で最も高く、それより発話が速くなっても遅くなっても、いずれの場合にもほぼ同等の傾斜、左右対称の形状で評価が低下することがわかった。

このことは、実音声区間の変換の程度が甚だしい周縁部にいくほど、自然性の評価が低下すると解釈することもできる。さて、先の性格特性印象の変化でも、周縁部では評価の低下傾向がみられる。その意味では、変換に伴う自然性の低下による性格印象評価への影響の可能性は皆無とは言えない。しかし、自然性の評価は、オリジナルの時間長を中心に左右対称形で低下しているのに対し、性格特性印象は必ずしも同じ変化形状をとらず、むしろそれに抗するような形の特徴的な変化パターンを示している。

例えば、勤勉性 (C) では、実音声部が早口になるほど評価が上昇しており、かなり速い発話でもあまり評価の低下は見られない。また、協調性 (A) は、遅い発話であっても比較的高い評価で安定している。したがって、これらの性格特性の印象変化パターンには、特性固有の特徴が反映されており、特性ごとの独自性が観測されたものと考えられる。

さて、発話速度感とは、全体の発話速度に応じて評価が変化した。また、休止時間長にも影響を受けていた。

FIGURE 5 を見ると、発話速度感とは、実音声区間の時間長の変化に鋭敏に反応し、実音声部の発話速度が速くなるにつれて、ほぼ単調にその評価が上昇している。それと比較すると、各発話ブロック内での休止時間長の影響は相対的に小さいことがわかる。したがって、発話速度感の評価の多くの部分は、実音声区間の時間長の影響によるところが多いことがわかる。

当然のことながら、休止区間の時間長が長くなれば、音声の総時間長も長くなるので、発話速度感に影響することは容易に推察できる。しかし、広実 (1994) は、休止区間のポーズは、単にその時間長だけでなく、挿入される回数も知覚的な発話速度に影響を与えている。すると、本実験で休止時間の最も長いものが必ずしも最も遅く評価されなかったことも、そのような観点から再検討することで解釈できる可能性もある。

また、本実験の休止時間長の操作範囲内では、話者の性格印象や音声の自然性はあまり影響を受けなかった。しかし、小森・長岡・河瀬・Draguna・中村 (2001) は、音声の中の内容面での区切りとして重要な文境界においては、ポーズ時間長は相対的に長めのものが最適と判断されることを示している。したがって、杉藤 (1997) の指摘と合わせて考えてみると、休止区間の役割としては、むしろ音声の中の言語的な意味や内容の理解に関連した機能の方に、より特化していると解釈することもできよう。

音声の時間構造と話者の性格印象の関係の頑健性と音声言語教育への寄与

従来、音声言語教育の場面で扱われてきた話し方やそれに対して聞き手が感じる印象などに関する知見は、アナウンサーや演劇関係者など、音声に関する経験の豊富な先達の主観的な印象や信念に基づいたものがほとんどであった (竹内, 1998 ; 広瀬, 2000 ; 米山, 2003)。また、落語や講談などの話芸に至っては、還元的な説明がなされることはなく、いわゆる徒弟制度の中で技の伝承がなされている。

さて、本実験で得られた知見の特色としては、FFM

といった心理学的な理論を基礎としたものであること、そして聴覚実験に基づいた実証的なデータに立脚した知見であることをあげることができる。さらに本研究では、従来の研究と比して、生態学的な妥当性の検証を目的の一つに据え、その実証を試みている点が特徴的である。

本実験は、厳密な実験操作を伴う音声刺激としては、かなり長い刺激を使用し、従来、扱われて来なかった休止時間長の検討を行っている。さらに音声の自然性の評価と性格特性印象との比較対照は新機軸であり、Big Five の特性ごとの独自性について、コントラスト鮮やかに示すことができたと言えよう。このような新しい取り組みは、いずれも生態学的な検証を行う上で、重要な役割を果たしている。

なお最近になって、籠宮・山住・楨・前川 (2004) は、日本語話し言葉コーパスの音声データベースを材料として、本実験と同様の Big Five Scale 短縮版を用いて話者の性格印象の評定を行っている。その結果、音声の速さ感と性格印象との間に、本実験での結果と類似した曲線関係を見出している。従来の研究は、ややもすると量質共に限られた音声刺激と実験条件の中での検討に過ぎなかったが、このような研究の蓄積と共に、得られた知見の頑健性が次第に明らかになりつつある。

このような実証的アプローチから発信されつつある心理学的な知見は、音声言語教育の内容面への提案として意義あるものと考えてよい。そしてそれは、従来の熟達者による主観的な信念に基づく知見を補うものとして、音声コミュニケーション教育一般に寄与するものと考えられる。

さて本研究では、音声の時間構造と話し手に関わる情報の関係性に着目したが、音声の時間構造は言語的な内容理解についても密接な関係がある。内田・菊地・中畝・前川・石塚 (2002) は、英語リスニング・テストの試験問題音声/material にし、本実験と同様の方法で実音声と休止時間を操作し、それらを用いて試験を行った。そして、その成績から音声の時間構造と内容理解との関係について定量的な検討を行っている。

今後さらに、音声とより様々な認知的な営みの関係を明らかにする研究が進み、音声コミュニケーションの実体の理解がいつそう深まっていくことを期待するものである。

ま と め

(1) 音声の中の実音声区間と休止区間の時間配分を実験

- 的に操作し、話者の性格印象、音声の自然性との関係を明らかにするための聴覚実験を行った。
- (2) 実音声時間長は、話者の性格印象、音声の自然性のそれぞれに、特徴的な影響を与えていた。休止時間長の違いによる影響は比較的小さかった。
- (3) パラグラフ・レベルの音声資料においても、実音声部の発話速度の操作に伴う、話者の性格印象の変化パターンは、Big Five の特性ごとに異なる固有の2次の関数形で近似することができた。
- (4) 勤勉性 (C) は速い発話で評価が高く、遅くなると評価が急激に低下する。協調性 (A) では速い発話で評価が低く、遅くなると評価が上がり、さらに遅くなってもそれほど低下しない。外向性 (E) と経験への開放性 (O) は、やや速い発話で評価が最も高く、それより速くなっても遅くなっても下降する。情緒不安定性 (N) には影響は見られない。
- (5) 音声の自然性は、オリジナルの時間長で評価が最も高く、速くなっても遅くなっても、左右対称の形状で評価が低下する。
- (6) 性格特性印象の変化パターンは、特性ごとにそれぞれが自然性のものとは異なっていた。音声の変換に伴う自然性の低下の影響は皆無とは言えないまでも、性格特性ごとに一定の独自性が確認された。
- (7) 話し手の性格印象の全体像を、Big Five の5つの特性印象値で表現する。すると、発話速度の違いに応じて、特性ごとに2次の近似式を介して特性印象値を推定し、それら5つの特性印象の推定値を組合せることで話者全体の印象が再構成できる。この特性印象の個別推定とその統合というモデルの下で、性格印象の全体像に迫る還元的なアプローチが可能である。
- (8) FFM といった心理学的理論を基盤とし、聴覚実験に基づいた実証的なデータに立脚した教育心理学的なアプローチから得られた本研究の知見は、音声言語教育の内容面での提案として意義深い。従来の音声の専門家による主観的な信念に基づく知見を補うものとして、音声コミュニケーション教育一般に寄与すると考えられる。

引用文献

- Goldberg, L. R. 1990 An alternative "Description of Personality": The Big-Five factor structure. *Journal of Personality and Social Psychology*, **59**, 1216-1229.
- 萩生田伸子・繁樹算男 1995 因子数選択のための新基準の提案 日本心理学会第59回発表論文集, 439.
- 広瀬久美子 2000 心を伝える話し方のコツ 永岡書店.
- 広実義人 1994 知覚上の発話速度に及ぼすポーズ数の影響 音声学会会報, **No.205**, 63-65. (Hirozane, Y. 1994 Effects of the number of pauses on perceived rate of speech. *Bulletin of the Phonetic Society of Japan*, **No.205**, 63-65.)
- John, O. P. 1990 The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research*. New York: Guilford Press. Pp. 66-100.
- 籠宮隆之・山住賢司・楨 洋一・前川喜久雄 2004 講演および講演者に対する評価の形成に関わる要因 日本音響学会2004年春季研究発表会, 343-344.
- 河原英紀 1998 聴覚の情景分析が生み出した高品質 VOCODER: STRAIGHT. 日本音響学会誌, **54**, 521-526. (Kawahara, H. 1998 STRAIGHT, a very high quality VOCODER: Insights from auditory scene analysis. *Journal of the Acoustical Society of Japan*, **54**, 521-526.)
- Kawahara, H., Masuda-Katsuse, I., & de Cheveigne, A. 1999 Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction: Possible role of a repetitive structure in sounds. *Speech Communication*, **27**, 187-207.
- 小森正嗣・長岡千賀・河瀬 諭・Draguna, M.・中村敏枝 2001 発話速度がポーズの時間長に及ぼす影響 ヒューマンインターフェイスシンポジウム2001論文集, 217-220.
- Malah, D., Crochiere, R. E., & Cox, R. V. 1981 Performance of transform and sub-band coding systems combined with harmonic scaling of speech. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **ASSP-29**(2), 273-283.
- 森田直孝・板倉文忠 1986 自己相関法による音声の時間軸での伸縮方式とその評価 電子通信学会技術報告, **EA86-5**, 9-16.
- 日本国際教育協会 2001 平成13年度日本留学試験

- (試行試験) 試験問題 日本国際教育協会.
- 瀬川榮志 1995 音声言語指導の重要性について 週間教育資料, No.445, 16-17.
- 杉藤美代子 1997 話しことばのアクセント, イントネーション, リズムとポーズ 杉藤美代子 (監修) 国広哲弥・廣瀬 肇・河野守夫 (編) 日本語音声 [2]アクセント・イントネーション・リズムとポーズ 三省堂 Pp.3-20.
- 竹内敏晴 1998 日本語のレッスン 講談社.
- 内田照久 1998 言語音声のための研究用話速変換プログラム PICOLA plus 2 の開発とその応用 日本教育心理学会第40回総会発表論文集, 311.
- 内田照久 2000 音声の発話速度の制御がピッチ感および話者の性格印象に与える影響 日本音響学会誌, 56, 396-405. (Uchida, T. 2000 Effects of the speech rate conversion on the impressions of pitch and the images of speakers' personality. *Journal of the Acoustical Society of Japan*, 56, 396-405.)
- 内田照久 2002 音声の発話速度が話者の性格印象に与える影響 心理学研究, 73, 131-139. (Uchida, T. 2002 Effects of the speech rate on speakers' personality-trait impressions. *Japanese Journal of Psychology*, 73, 131-139.)
- Uchida, T. 2003 The effects of fundamental frequency and speech-rate conversion upon impressions of talkers' personality. In C. Don (Ed.), *Acoustics on the move : Proceedings of the Eighth Western Pacific Acoustics Conference : WESPAC VIII*, Castlemaine : Australian Acoustical Society, Speech-TB23.
- 内田照久・菊地賢一・中畝菜穂子・前川眞一・石塚智一 2002 英語リスニング・テストにおける音声の時間構造と提示情報の様式が項目特性に与える影響 教育心理学研究, 50, 1-11. (Uchida, T., Kikuchi, K., Nakaune, N., Mayekawa, S., & Ishizuka, T. 2002 Effects of temporal structure and mode of presentation on item characteristics on an English listening comprehension test. *Japanese Journal of Educational Psychology*, 50, 1-11.)
- 和田さゆり 1996 性格特性語を用いた Big Five 尺度の作成 心理学研究, 67, 61-67. (Wada, S. 1996 Construction of the Big Five Scales of personality trait terms and concurrent validity with NPI. *Japanese Journal of Psychology*, 67, 61-67.)
- 米山文明 2003 声の呼吸法 美しい響きを作る 平凡社.

付 記

本研究は、日本学術振興会の科学研究費補助金『若手研究(B)：課題番号14710087』の援助を受けて行われました。また研究の一部は、日本音響学会2003年春季研究発表会、及び日本心理学会第67回大会で発表いたしました。

(2004.3.3 受稿, '04.9.18 受理)

Impression of Speaker's Personality and the Naturalistic Qualities of Speech : Speech Rate and Pause Duration

TERUHISA UCHIDA (NATIONAL CENTER FOR UNIVERSITY ENTRANCE EXAMINATIONS)

JAPANESE JOURNAL OF EDUCATIONAL PSYCHOLOGY, 2005, 53, 1–13

The present study investigated the relation between temporal structures of speech and listeners' impressions of the speaker's personality. The following speech stimuli were manipulated in paragraph-length speeches: speech durations of sound region (65%-147%) and duration of pause (56%-178%). University students (444 men, 137 women, 18-23 years of age) listened to the experimental stimuli and rated their impression of each speaker's personality and qualities of the speech. Pause duration did not distinctly affect impressions of the speakers' personality, whereas speech rate did. Characteristic quadratic functions provided good approximations of each "Big Five" trait from the Five-Factor Model of personality. Evaluations of Conscientiousness were high when speech was fast, and declined sharply for slower speech. In contrast, Agreeableness was rated high with slow speech, and dropped as speech became faster. Ratings of Extraversion and Openness to Experience were high for relatively rapid speech, and decreased somewhat when speech was faster or slower. Naturalness of speech was rated highest at the original speech rate, and declined symmetrically for faster or slower speech. Listeners' impression of each personality trait appears to have a unique changing pattern. Overall, the results supported the hypothesis that a formulation combining five approximations of personality traits can estimate the entire personality image.

Key Words : spoken language education, speech rate, pause duration, impression formation, Five-Factor Model of Personality (FFM)