

速度変化をともなう音声の速度感と その規定要因

丸島 歩[†]

【要旨】自然言語音ではしばしば発話内で発話速度が変化する。しかし、それが聴覚的な速度感にどう影響するかは、あまり考慮されてこなかった。そこで、発話の速度が一定の音声と次第に加速する音声、減速する音声を3種類の速度 (fast, normal, slow) でそれぞれ3トークンずつ用意して聴取実験を行った。

実験1ではサーストンの一対比較法を用いた。その結果、速度変化をともなう fast, normal の音声は、速度が一定の音声とは速度感が異なるという結果が得られた。

実験2ではそれぞれのトークンの速さを評価させた。その結果、fast, normal の速度変化をともなう音声は、判断のゆれが生じやすいということが示唆された。

キーワード： 発話速度、速度感、速度変化、聴取実験

1. はじめに

日本語音声の速さはしばしば、単位時間あたりに話されたモーラ数で算出される発話速度と、発話全体からポーズ時間を除いた時間長で算出される調音速度で表される。この単位はどちらも日本語音声ではモーラ数などを基準に算出されるもので、その音声が早く聞こえるのか遅く聞こえるのかという知覚的な側面とは必ずしも一致しない。

音声言語の速さについて、音の高さの知覚的尺度であるメル、音の強さの聴覚的単位であるホンのような知覚量をあらかず単位は存在しない。しかしこれまで、日本語音声の速度感が何によって規定されるのか、もしくは影響されるのかについて、さまざまな視点から述べられてきた。

1.1 分節音に着目した研究

分節音のレベルで、日本語音声の時間構造を規定する要因を扱っているものとして、加藤ほか (2004) がある。加藤ほか (2004) では、人が音声に時間構造を感じる際の手がかりを検討している。V-onset (子音から母音への境界) と V-offset (母音から子音あるいはポーズへの境界) とで時間マーカ¹としての強さに違いがあるか、役割に違いがあるかを、実音声に時間伸縮を施した音声刺激を用いた聴取実験で調査した。その結果、時間マーカとしての強さに違いはないが、テンポ判断の手がかりとしては V-onset のみが有効に働いた。

[†] 筑波大学大学院人文社会科学部研究科一貫性博士課程

¹ 音声から時間構造を感じさせる手がかり。加藤ほか (2004) では V-onset や V-offset のみを移動させた音声は原音声と異なって聞こえるかどうかで、V-onset や V-offset の時間マーカとしての強さを検出している。これに対してテンポ判断を検出するためには、被験者の判断基準をテンポのみに絞っている。

1.2 プロソディーに着目した研究

プロソディー的な要因を述べたものとして内田 (2000) がある。内田 (2000) では、日本音響学会研究用連続音声データベースから抜粋した音声を用いて、ピッチや発話速度の物理量や実測値とその知覚の関連について明らかにしている。まずオリジナルの音声とその話速を80%、125%、156.25%にそれぞれ変換したものを、40名の日本語母語話者に聴かせた。その結果、 F_0 値に有意差はなかったものの遅く変換されたものほどピッチも低く感じられるという傾向が見られた。さらに、同じ音声刺激を用いて性格印象についても実験が行われており、協調性についてはやや遅いものが、勤勉性・経験への開放 (独創性など)・外向性についてはオリジナルの音声かやや速い発話が、それぞれ最もあてはまるという結果になった。また、同じ音声と基本周波数を90%に下げたものと111.11%に上げたものを被験者に聴かせたところ、高いものほど速く聴こえる傾向があった。丸島 (2008) でも発話速度が近似の自然発話音声を組み合わせて速度感を聞き比べさせる実験が行われ、そこから得られた結果から、ピッチ変動の大きい音声のほうが速く聞かれる可能性が指摘された。

1.3 ポーズに着目した研究

聴覚的な音声言語の速度感とポーズの関連について明らかにしたものとして、広実 (1994)、杉藤 (1999) が挙げられる。広実 (1994) は、ポーズ数が知覚上の発話速度に及ぼす影響を、ニュース原稿の一部を用いた資料を読ませた音声を加工することで検証した。録音された資料には300ms以上の比較的長いポーズが4個、300ms未満の比較的短いポーズが7個あった (刺激音 A)。この元の音声を用いて、短い7個のうちの任意のポーズ4個を、後続または先行する比較的長いポーズの位置に移動、合体させ、ポーズ7個の音声を作成した (刺激音 B)。さらに残りの比較的短い3個のポーズを、すべて後続または先行する比較的長いポーズの位置に移動、合体させてポーズを4個にした音声 (刺激音 C) も作成した。これら3種類の音声刺激を2つ1対 (AB、BC、CA) にしたものを、各対10試行ずつ連続でランダムに被験者に聞かせてその遅速を判断させた。その結果、刺激音はC-B-Aの順でより速いと判断された。つまり、発話速度や調音速度が同じでもポーズ数が少ないほど「遅い」と知覚されることが明らかになった。

杉藤 (1999) ではポーズを切除した音声を被験者に聞かせ、その印象を述べさせている。ポーズを切除した音声は速く聞かれ、内容の把握も難しいとしている。丸島 (2009) でも2トークン1ペアの音声を数種類被験者に聞かせて音声の速度比較を行わせる実験が行われており、その結果ポーズを除いた発話部分の調音速度²より、ポーズを含めて算出した発話速度³の方が聴覚的な速度感とより相関が高いことを示した。

1.4 複合的な要因に着目した研究

発話速度知覚に影響を与えている発話特徴を複合的にとらえた研究として、籠宮ほか (2008) がある。籠宮ほか (2008) では『日本語話し言葉コーパス』に収められている模擬講演音声⁴を用いている。講演の録音時に各講演1名の収録スタッフ (のべ人数9名) がその講演全体の速度印象を判定した。これらのデータを分析した結果、発話速度の知覚に影響を与える要因として「ポーズ比⁵」

² モーラ数を発話部分の時間長で割った数値である。

³ モーラ数をポーズを含めた全体の時間長で割った数値である。

⁴ 学生や派遣社員があらかじめ与えられたテーマで、10～15分程度のスピーチを聴衆3～5名の前で行ったものである。

⁵ 「(講演の時間長 - 講演の転記基本単位時間長の総計) / 講演の時間長」で計算される。なお、「転記基本単位時間長」とは200ms以上のポーズや明確な文末形式を切れ目とした単位である。

「ポーズ/秒⁶」「モーラ/秒⁷」を挙げ、発話中のポーズの数や長さも重要な要因であると述べた。言い淀みやフィラーなどは発話速度知覚に影響を及ぼす要因として抽出されなかった。

1.5 自然言語音声における速度変化

前節までで述べたように、これまで日本語音声の速度感に影響を与うるさまざまな要因が指摘されてきた。実験には単語レベルの短い音声を用いられているものもあるが、ある程度の長さをもつ音声を刺激として用いる場合も多い。

自然発話の音声はある程度の長さをひとまとまりとしてとらえると、しばしば発話の途中で速度が大きく変化する。以下の図1・2は丸島(2007)で収録された自然発話の音声を解析したものである。棒グラフの横幅は発話部分の時間長を、棒と棒の間の距離はポーズの時間長を示している。また、棒の長さは該当する発話部分の調音速度を示しており、その具体的な数値が上に赤字で付されている。これらはごく日常的な発話であり、発話者に対して速度に関わる実験であることを提示したわけではない。これらの音声は、発話内で非常に大きく速度が変化していることがわかる。

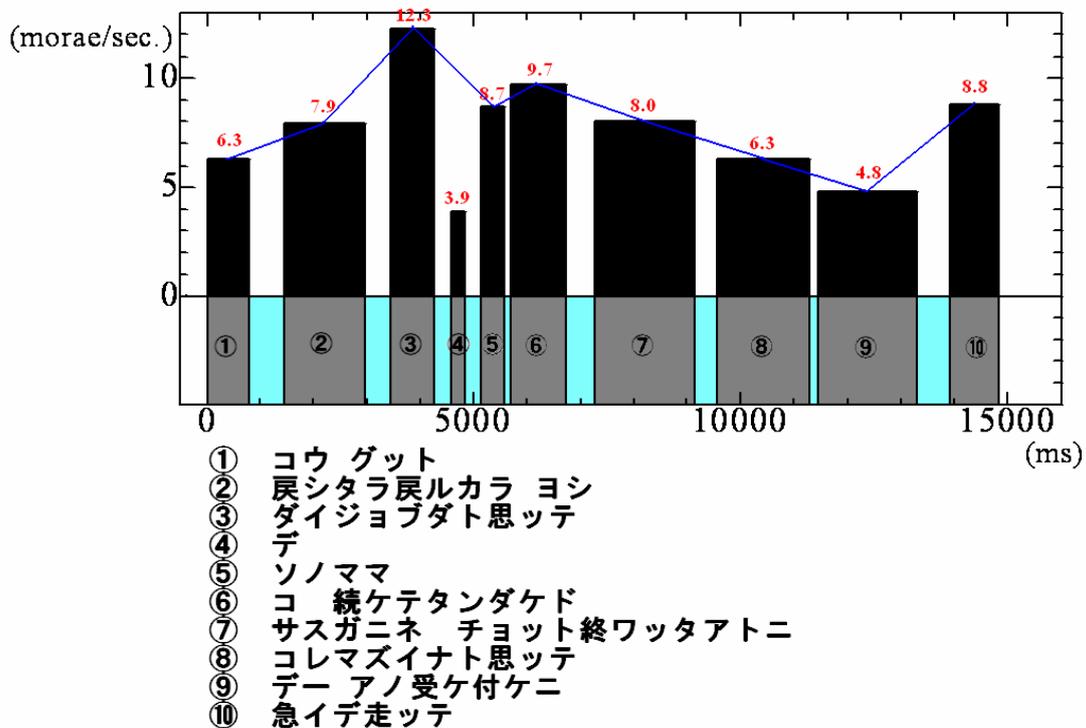


図1：自然発話の発話部分とポーズ部分

⁶ 「講演中のポーズ数の総計/講演の時間長」のことである。

⁷ 「講演中のモーラ数の総計/講演の転記基本単位時間長の総計」のことである。「転記基本単位時間長」については、注5を参照されたい。

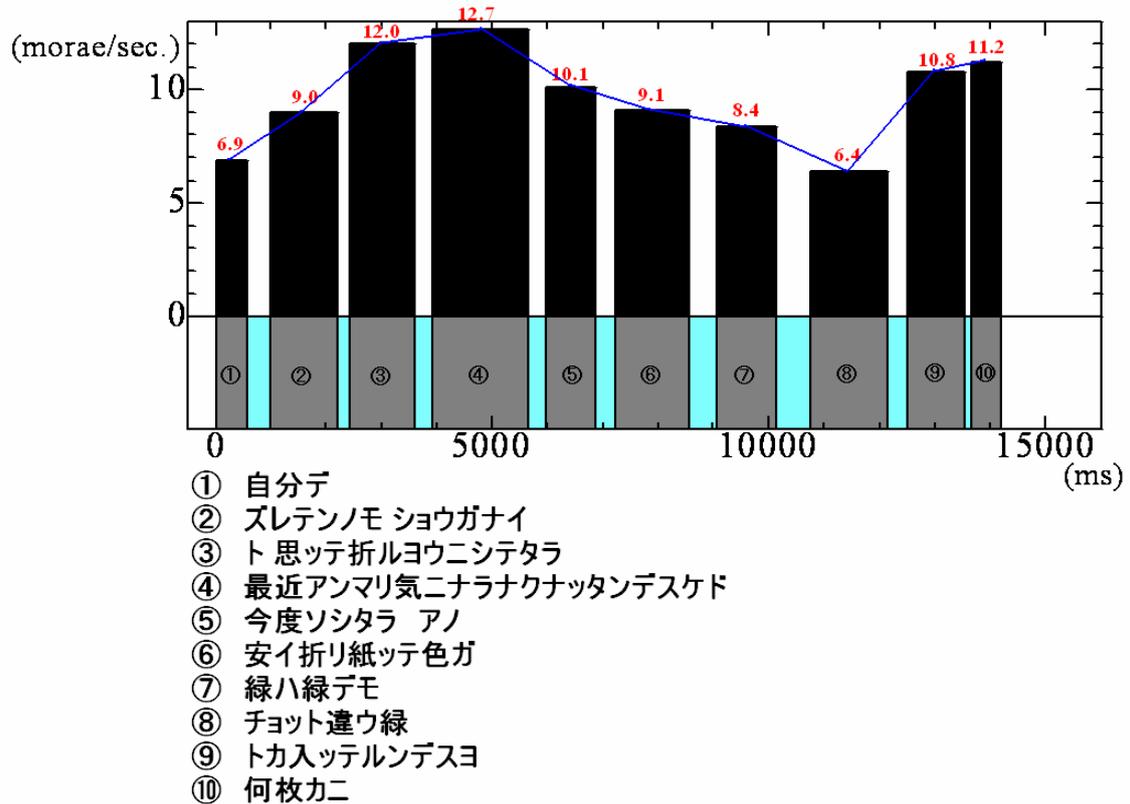
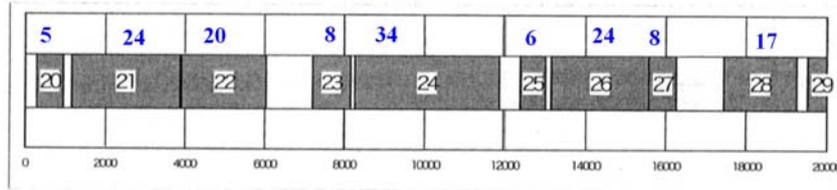


図 2：自然発話の発話部分とポーズ部分

このような特徴を持つ音声は自然発話の音声だけとは限らない。ある程度統制されていると思われるキャスターやアナウンサーの音声でも、日常的な発話ほど大きな変化ではないにしてもこのような現象が見られる。福盛 (2008) では 2006 年 5 月 3 日に放送されたニュース番組のアナウンサーやキャスターの音声の発話速度やポーズの割合を計測している。以下の図 3 は福盛 (2008) から転載したものである。グラフのグレーの部分が発話部分で、白い部分がポーズである。左端から右端までで 20 秒になっている。なお、各発話部分の上部に青字で数字が付してあるが、これは筆者がその発話部分のモーラ数を記したものである。例えば、23 と 27 はともに 8 モーラであるが、その時間長を見ると明らかに 23 のほうが長くなっている。また、28 は 27 のおよそ 2 倍のモーラ数であるが、28 の時間長は明らかに 27 の 2 倍以上の長さになっている。ここからも調音速度が必ずしも一定に発話されていないことが明らかである。



- 20：旅客機は
21：沿岸から6キロ離れた黒海の上空で
22：レーダーから機影が消えたということです。
23：この旅客機には
24：乗客乗員およそ100人が乗っているとみられているということで
25：今のところ
26：日本人の乗客が乗っているという情報は
27：入っていません。 28：ロシア非常事態省によりますと
29：機影が消えた現場、現場付近で

図1-3：NHK ニュースにおけるポーズの分布例（3）

図3：NHK ニュースにおけるポーズの分布例（福盛（2008）より転載）

では、速度変化をともなう音声を聞いた際、我々はその速さをどのようにとらえているのだろうか。発話全てを平均化して聴いているのだろうか。先行研究ではある程度の長さをもった音声を扱っているものも少なからず存在するが、発話途中での速度変化は考慮されていない。小林ほか（1996）は速度が11%変化すると卓立が感じられるとしているが、その変化が速度感にどのような影響を及ぼすかについては述べられていない。全ての発話の速度が平均化してとらえられていれば考慮の必要性はないが、発話の特定の部分に強く印象づけられ、そこに速度感が大きな影響を受ける可能性もあるだろう。例えば始端もしくは終端部分、特に速い部分や遅い部分が注目して聞かれ、その部分の調音速度が全体の印象を左右している可能性が考えられる。

1.6 本研究の目的

1.5 節までで、自然発話の音声では発話内でしばしば速度が変化することを述べた。また、音声の速度の知覚についてさまざまな角度から研究が行われているにも関わらず、発話内の速度の変化はあまり考慮されていないことにも触れた。

そこで本研究では、一定速度で話された音声と発話内で速度が変化する音声それぞれどのように聞かれているかを観察することを目的とする。発話途中でその速度が変化する音声の速度感がどのようにになっているのかを明らかにするために、同一文の読み上げ音声を用いた聴取実験を行うこととした。

2. 実験

発話速度、調音速度といった速度に関する定量的単位と異なり、知覚量としての発話の速度感には明確な基準が存在しない。そこで本稿では、2種類の聴取実験を行った。まず実験Aでは、速度が一定の音声と変化する音声の速度感の相対差を算出することにより、尺度化を試みる。次に実験Bでは、被験者の聴覚印象のみを基準としてこれらの音声の速度感を記述する。

2.1 実験A

実験 A では、一対比較法を用いることで速度感の尺度化を試みる。具体的には後述するが、2 トークン 1 組で音声を聴き比べて判定させる方法である。

2.1.1 音声資料

文「まわりのさんごを見渡すと、それこそぼきぼきに折られ、白い折れ口をさらしていた」を、各読点部分にポーズを置くようにして読んだ。速度は“fast”、“normal”、“slow”の 3 種を設定し、読み手自身が普通に読んだものを“normal”とした。また読み方は、「ほぼ一定の速度で」(EV=“even”)「だんだん速く」(AC=“accelerating”)、「だんだん遅く」(DE=“decelerating”)の 3 種の読み方で読んだ。EV では普通に読み、AC ではだんだん読みの速度が速くなるように読み、DE ではだんだん読みの速度が遅くなるように読んだ。これらの 3 速度と 3 種の読み方を組み合わせて、全部で 9 のパターンを作成した。これら計 9 パターンを約 10 回ずつ発話した。3 速度内でそれぞれの読み方の平均の発話速度ができるだけ統一されており、全体のピッチの中央値にできるだけ差がないものを各パターン 3 トークンずつ選んだ。10 回ほど発話しても 3 トークン得られないパターンについては、ある程度統一されたパターンが得られるまで何度も追加録音をして補った。なお、録音者は筆者自身⁸である。

各速度群の時間構造を以下の図 4~6 に示した。なお、頭文字が a のものは発話部分、p のものはポーズ部分である。a1 は「まわりのさんごを見渡すと」の部分、a2 は「それこそぼきぼきに折られ」の部分、a3 は「白い折れ口をさらしていた」の部分にあたる。p1 は a1 と a2 の間のポーズ、p2 は a2 と a3 の間のポーズである。グラフのそれぞれの棒が各トークンの情報で、横軸は時間長である。図 4 の fast が全体としてもっとも時間長が短く、図 6 の slow がもっとも時間長が長い。また、各速度の AC (加速) 音声は EV (一定) に比べて a1 が長く a3 が短い傾向にあり、DE は逆に a1 が短く a3 が長い傾向にある。

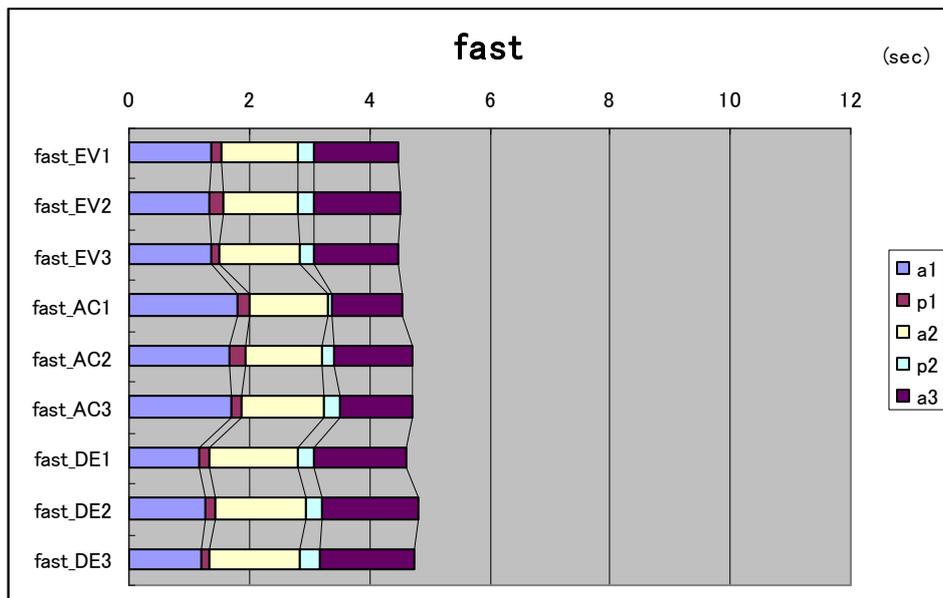


図 4 : fast 群音声の時間構造 (a1~a3 が発話部分、p1~p2 がポーズ)

⁸ 録音時 27 歳、女。言語形成地は東京都東久留米市である。

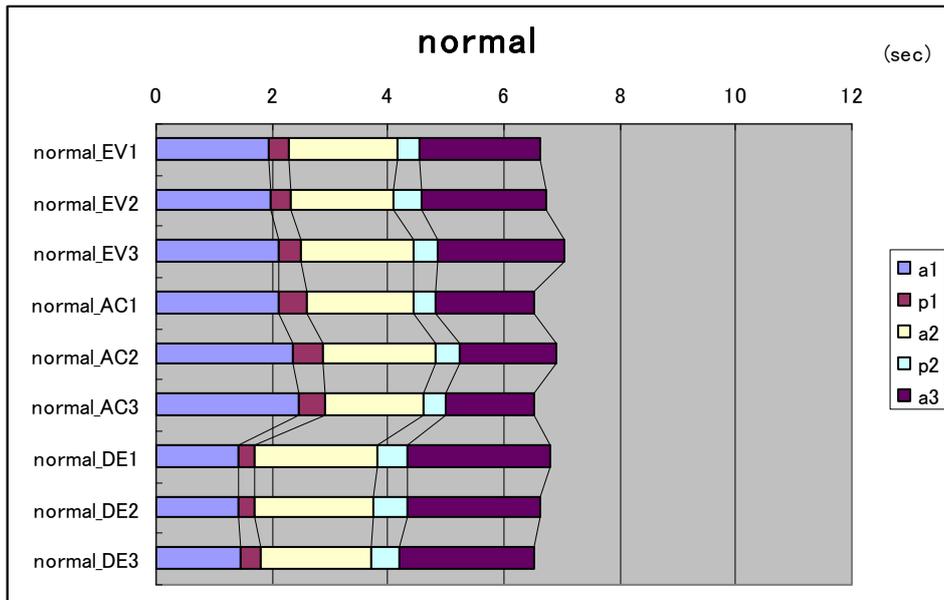


図 5 : normal 群音声の時間構造 (a1~a3 が発話部分、p1~p2 がポーズ)

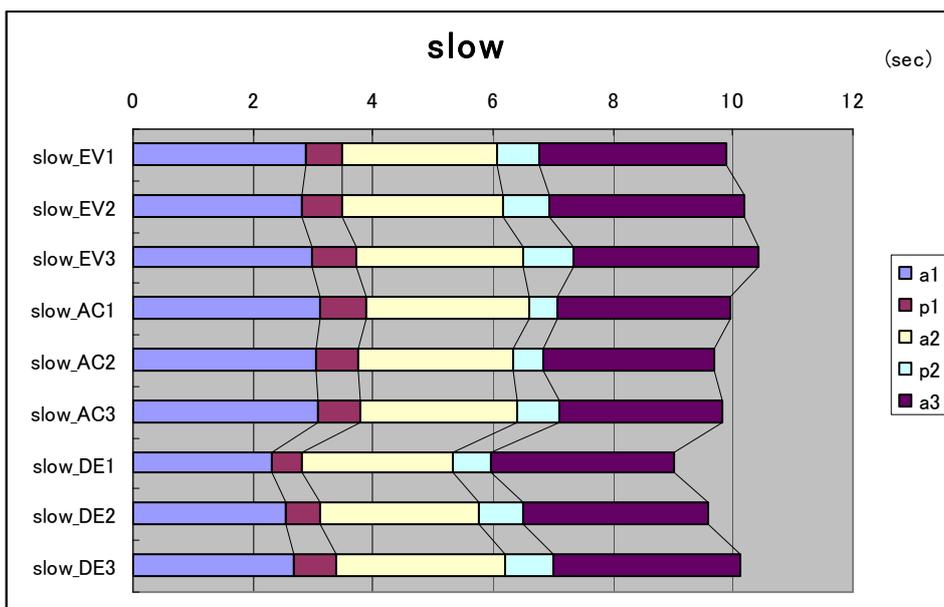


図 6 : slow 群音声の時間構造 (a1~a3 が発話部分、p1~p2 がポーズ)

さらに、各トークンのピッチ情報を以下の図 7~9 で箱ひげ図にあらわした。箱の上下が四分位範囲、ひげの部分が高値と最低値である。また、箱の中の横線が中央値をあらわしている。それぞれのピッチの中央値にはあまり差がないことがわかる。

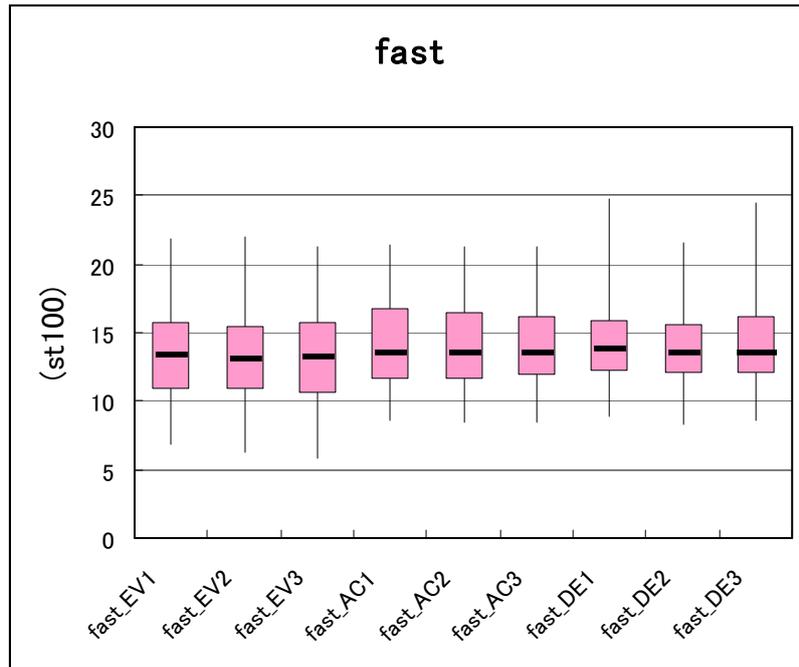


図 7 : fast 群のピッチ情報 (最高値、最低値、中央値、四分位範囲)

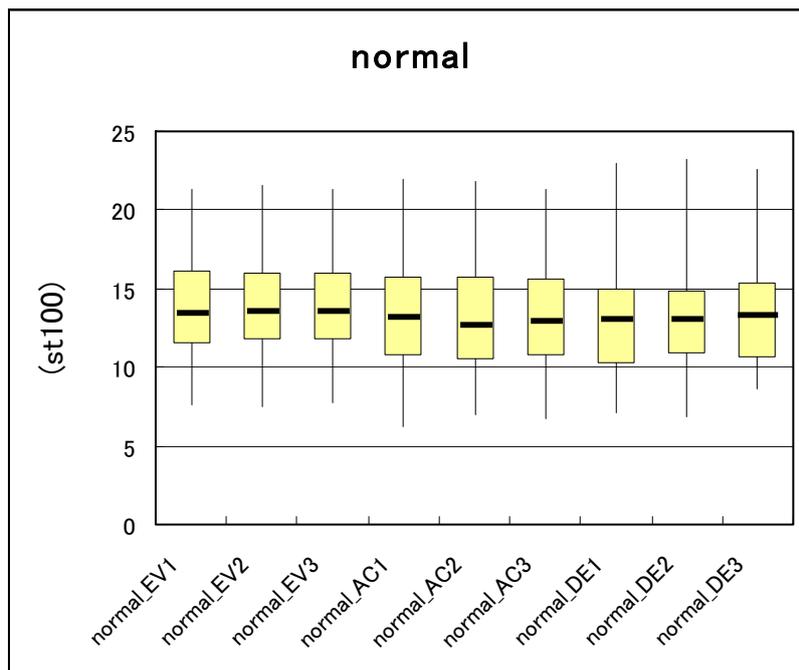


図 8 : normal 群のピッチ情報 (最高値、最低値、中央値、四分位範囲)

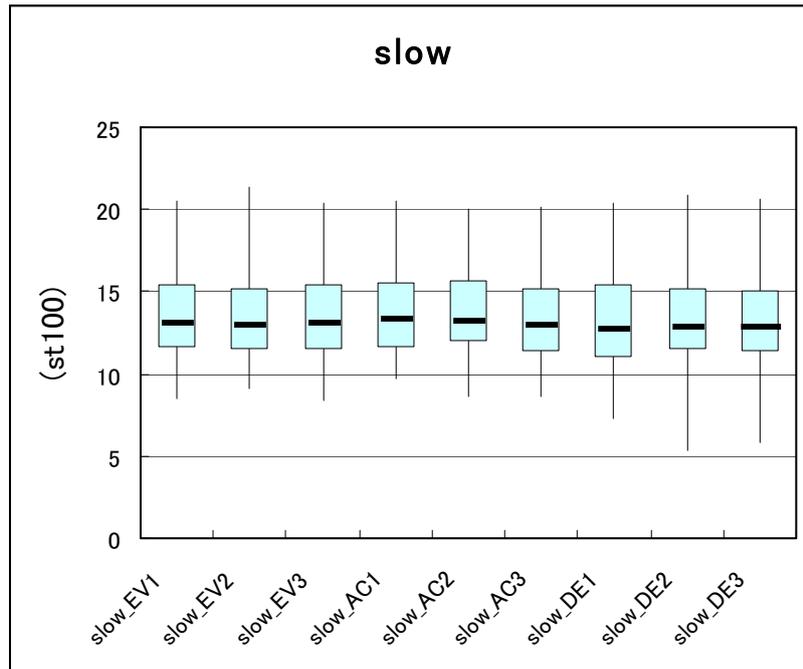


図 9：slow 群のピッチ情報（最高値、最低値、中央値、四分位範囲）

2.1.2 実験手順

上記で得られた 27 トークンを、サーストンの一対比較法⁹にもとづいた組み合わせで計 351 組のペアを作成し、どちらが速く聞こえるかを被験者に回答してもらった¹⁰。

実験は Praat ver.5 の ExperimentMFC の機能で作成し、ペアの始めと各トークン同士の間には 500ms のポーズを入れた。聴き逃した場合を考慮し、各ペアを 1 回まで聞きなおせるようにした。

なお、実験を始める前に本実験とは別の音声で作成した 8 組のトークンペアを用いて練習を行った。聞き取りはヘッドフォンを用いて行った。

2.1.3 被験者

被験者は日本語を母語とする 19 歳～29 歳（平均 21.4 歳、標準偏差 3.5）の男女 10 名（男女 5 名ずつ）である。

2.1.4 解析手順

上記で得られた結果を Praat ver.5、Microsoft Excel で集計・統計処理した。具体的には、サーストンの一対比較法¹¹を用いた。その後の分散分析、多重比較などの統計処理には、R 言語 (ver.2.13.0) を用いた。

2.1.5 結果

⁹ 主観評価法のひとつ。複数の刺激を 2 つずつ組み合わせ、それらを被験者の主観で比較させる。その結果から各刺激を数直線上にプロットし、刺激の順位と心理的距離を示すことができる。

¹⁰ 351 組を a・b・c ブロック (各 90 組) と d ブロック (81 組) に分けて提示した。各ブロック内はランダムで提示されるよう設定した。各組内での刺激の提示順は、全被験者が同じにならないようにした。具体的には、各ブロックについて各組内の刺激提示順が逆順の実験ファイルを用意し、正順のファイルと逆順のファイルを組み合わせることでカウンターバランスを取った。

¹¹ ただし、サーストンの一対比較法は各ペアの選択率が 0.0、もしくは 1.0 の場合は計算が不可能である。その場合は 0.0 を 0.0001、1.0 を 0.9999 の近似値に置き換えて算出した。

実験 A の結果を以下の図 10 に示す。ポイントの色が各速度を示しており、赤が fast、黄が normal、青が slow である。また、形が読み方を表しており、○が EV (一定)、△が AC (加速)、□が DE (減速) である。数値が少ないほど速いと判断されたということを意味する。

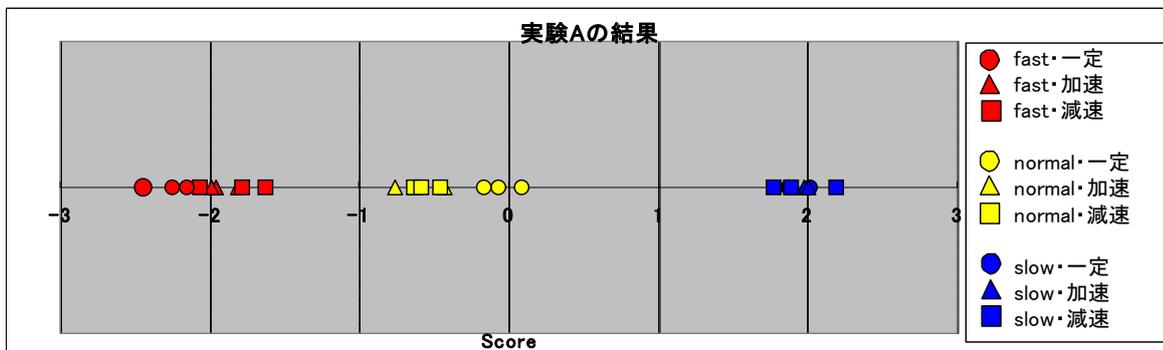


図 10：実験 A の結果

上の図 10 の結果から、読み方が一定のものはほぼ速度ごとに均等に配されているが、加速したり減速したりしているものは同じ速度の一定のトークンからずれているものが多い。

この結果に対して、速度と読み方を要因とした二元配置分散分析 (対応なし) を行った¹²。速度の要因については有意であった ($\{F(2,18)=1695.2225, p<.001\}$) が、読み方の要因については有意ではなかった ($\{F(2,18)=0.3059, p=0.740\}$)。また、速さと読み方の交互作用が有意であった ($\{F(4,18)=10.8089, p<.001\}$) ので、単純主効果分析として、速さ・読み方それぞれの要因の各水準で、一元配置分散分析を行った¹³。結果は、以下の表 1・2 のとおりである。速さの各水準における読み方を要因とした分析では、fast と normal では有意であり、slow では有意ではなかった。読み方の各水準における速さを要因とした分析では、EV、AC、DE 全ての水準において有意であるという結果になった。

表 1：単純主効果の検定結果 (速さの各水準における読み方を要因とした分析)

	第一自由度	第二自由度	F 値	p 値	有意差 ¹⁴
fast	2	6	6.622	0.030	*
normal	2	6	16.201	0.004	**
slow	2	6	0.0753	0.928	n.s

¹² $p<.05$ を有意、 $p<.10$ を有意傾向とする。

¹³ $p<.05$ を有意、 $p<.10$ を有意傾向とする。

¹⁴ $p<.05$ を*、 $p<.01$ を**、 $p<.001$ を***としてある。

表 2：単純主効果の検定結果（読み方の各水準における速さと要因とした分析）

	第一自由度	第二自由度	F 値	p 値	有意差 ¹⁵
EV	2	6	895.97	p>.001	***
AC	2	6	894.79	p>.001	***
DE	2	6	315.24	p>.001	***

さらに、有意であった各分析について Tukey の多重比較を行った¹⁶。

速度が fast のものについては、AC > EV が有意傾向 (p = .076)、DE > EV が有意 (p = .032) となった。normal については、EV > AC (p = .005)、EV > DE (p < .001) が有意であった。

読み方が EV (一定) のものは、normal > fast、slow > fast、normal > fast いずれも有意であった (p 値はいずれも p < .001)。読み方が AC (加速) のものも、normal > fast、slow > fast、normal > fast いずれも有意であった (p 値はいずれも p < .001)。読み方が DE (減速) のものも、normal > fast、slow > fast、normal > fast いずれも有意であった (p 値はいずれも p < .001)。

以上の結果から、読み方が一定でも刺激内で速度が変化していても、fast がもっとも速いと判断され、slow がもっとも遅く、その中間が normal であるという結果になった。また、fast と normal では読み方によって速さの知覚が異なるという結果になった。具体的に述べると、normal は加速・減速ともに一定のものより速いと判断されているが、fast では全く逆の結果になっており、速度がトークン内で変化するとより遅く聞こえるという傾向が見られる。slow に関しては読み方による速度判断の違いはみられない。以上の結果をまとめると表 3 のようになる。

表 3：実験 A 「加速」「減速」音声の「一定」音声との相対差

速度/読み方	加速	減速
fast	(より遅い) ¹⁷	より遅い
normal	より速い	より速い
slow	ほぼ同じ	ほぼ同じ

2.1.6 考察

上の結果では、速度が加速していても減速していても、速度知覚への影響は類似していた。fast 音声では加速・減速問わず一定の音声より遅いと判断され、normal 音声では加速・減速どちらもより速く聞かれるという結果になった。slow 音声では加速・減速ともに速度感に影響を及ぼさないという結果になった。

このことから、音声の始端もしくは終端のどちらかが固定的に速度感に影響を与えるわけではないということが言える。もし加速音声が一貫してより速く聞かれ、減速音声が一貫してより遅く聞かれるという結果になれば、音声の終端が速度感に重要な影響を及ぼす可能性が考えられる。もし逆の結果になれば、音声の始端が速度感に強く影響するとも考えることができる。しかし、本実験では音声が増速しても減速しても一定の音声との相対的な関係は共通していた。

また、速度変化が速度感にもたらす影響は、全体としての速度によって異なることが言えるだろう。具体的には、normal 群においては速度変化があるものが一定のものより速く聞こえるという傾

¹⁵ 注 14 と同様である。

¹⁶ p<.05 を有意、p<.10 を有意傾向とする。

¹⁷ 統計結果が有意傾向であったため、括弧を付した。

向が見られた。それに対し、fast 群には速度変化があるもののほうがより遅く聞こえる傾向が見られ、normal 群とはまったく逆の結果になった。slow 群では、加速・減速に関わらず、一定のものとの有意な差が見られなかった。

このことから、速度が変化する音声は、速い部分が注目して聞かれるとも、遅い部分が注目して聞かれるとも限らないと考えられる。速度変化が音声の速度感にどう影響するかは、全体としての発話速度が速いか遅いかによって異なることが示唆された。

では、fast 群と normal 群ではなぜ全く逆の傾向が見られたのであろうか。また、slow 群では速度変化が速度知覚にあまり影響を及ぼさなかったのはなぜだろうか。

ここでは速度変化を伴う音声が、fast 群においては normal 群の方向に寄り、normal 群においては fast 群の方向に寄っていることに注目したい。つまり、両群の速度変化を伴うトークンの速度感はある一定のものに比べてそれぞれが互いに近い位置にあることを意味する。したがって一つの可能性として、ある特定の速度が速度感に影響を与えやすいということが考えられる。その速度に比較的強い影響を受けて、fast 群・normal 群の加速・減速音声のスコアが一定音声と離れた値になったかも知れない。ただし本実験の結果だけで、特定の速度が速度判断に影響を及ぼしやすいと断言することはできない。これを明らかにするためには、より細かなパラダイムによる実験が必要だろう。

2.2 実験 B

前述したとおり、実験 B では被験者の速度感のみを基準とした速度知覚を観察する。具体的には、音声を聞いてそれが速く感じられるかもしくは遅く感じられるかを判断してもらうという方法をとる。

被験者の直感的な判断を引き出すため、選択肢はできるだけ単純なものにした。もっとも単純なのは「速い」「遅い」の 2 つの選択肢を設定することであるが、速いとも遅いとも感じられない場合があり得ることを考え、「どちらでもない」という選択肢¹⁸を加えた三択で回答してもらうことにした。

2.2.1 音声資料

音声資料は実験 A と同じものを用いた。さらに、A 実験で使われなかったトークンを加工して 6 種類のダミー音声を作成した。

2.2.2 実験手順

上記のトークンを 3 回ずつランダムで聞かせ、それぞれ「速い」「遅い」「どちらでもない」から選択させた。27 トークンにダミーの 6 トークンを加えて 33 トークンとなり、これを 3 回ずつ聞かせるので計 99 試行を行った。

実験は Praat ver.5 の ExperimentMFC の機能で作成した。聴き逃した場合を考慮し、各ペアを 1 回まで聞きなおせるようにした。

なお、実験を始める前に本実験とは別の音声で作成した 8 つのトークンを用いて練習を行った。聞き取りはヘッドフォンを用いて行った。

¹⁸ 「速い」と「遅い」の中間的な選択肢であるため、「ふつう」などの表現を用いることも考えられる。しかし、「ふつう」の速度という表現は直感的にわかりづらく、被験者によってその解釈や判断がまちまちになってしまうおそれがあると考えた。音声言語の「ふつう」の速度とは何であるかという問いは、本稿で扱う問題とは別に扱わなければならない問題である。本稿の実験で意図したのは単に「速い」と「遅い」の中間という意味でしかないので、「どちらでもない」という選択肢を用いることにした。

2.2.3 被験者

被験者は実験 A と同じ被験者で、日本語を母語とする男女 10 名である。

2.2.4 解析手順

上記で得られた結果を Praat ver.5、Microsoft Excel で集計した。集計したのち、速度群ごとにカイ二乗検定を行った。カイ二乗検定は R (ver.2.13.0) を用いて計算を行った。

2.2.5 結果

回答の割合を、速度群ごとにまとめた。以下の図 11～13 にグラフで示す。

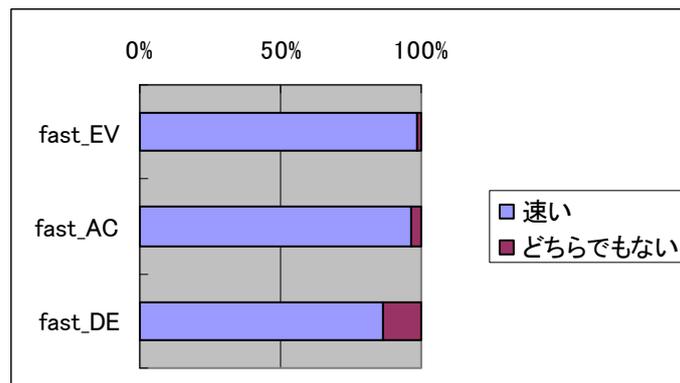


図 11：fast 群の回答

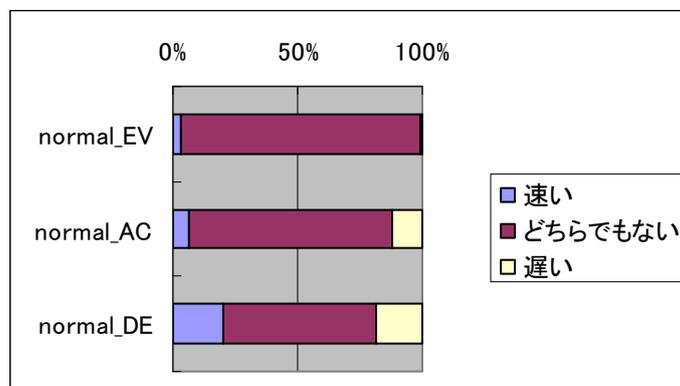


図 12：normal 群の回答

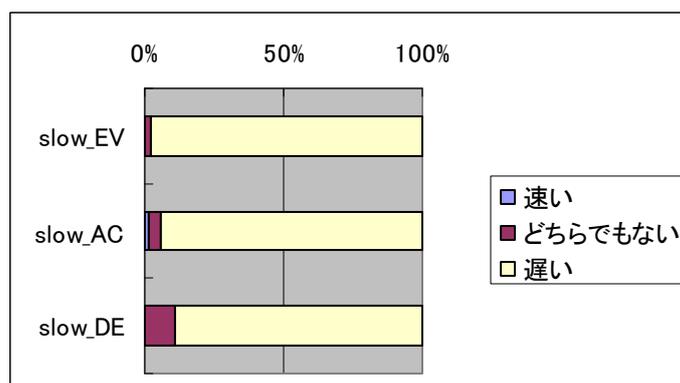


図 13：slow 群の回答

fast 群の回答割合を見ると、一定 (EV)、加速 (AC) では 100%に近い割合で「速い」と回答されているが、減速 (DE) ではやや「どちらでもない」の割合が増えている。normal 群ではどの読み方でも「どちらでもない」が最も多い。ただし、「一定」に比べて「加速」では若干その割合が低くなっている。「減速」ではさらにその割合が低い。slow 群では「一定」では 100%が「遅い」という回答になっているが、「加速」・「減速」ではややその割合が下がっている。

各速度群で回答割合についてカイ二乗検定¹⁹を行った。その結果、fast 群では 1%水準で有意 ($\chi^2=13.686$, $df=2$, $p=0.001$)、normal 群でも 1%水準で有意で ($\chi^2=27.663$, $df=4$, $p<0.001$) あり、ともに回答の割合は等確率ではないという結果になった。slow 群では有意差はあらわれなかった ($\chi^2=8.8874$, $df=4$, $p=0.064$)。

さらに、これらの結果を被験者ごとに細かく算出した。その結果を、以下の表 4~6 に示す。なお、各表の左上に被験者番号を付してある。各被験者の結果の下に被験者ごとに行ったカイ二乗検定の p 値を示し、有意差の有無を記した²⁰。

¹⁹ $p<.05$ で有意であるとしている。

²⁰ 被験者ごとの検定も $p<.05$ で有意であるとした。また、 $p<.05$ には*、 $p<.01$ には**、 $p<.001$ には***を付してある。

表 4：被験者別結果 (fast)

被験者 1				被験者 6			
	速い	どちらでもない	遅い		速い	どちらでもない	遅い
AC	9	0	0	AC	9	0	0
DE	6	3	0	DE	9	0	0
EV	8	1	0	EV	9	0	0
p=.034			有意差	n.s.	-算出不可 ²¹ -		
被験者 2				被験者 7			
	速い	どちらでもない	遅い		速い	どちらでもない	遅い
AC	8	1	0	AC	9	0	0
DE	8	1	0	DE	9	0	0
EV	9	0	0	EV	9	0	0
p=.583			有意差	n.s.	-算出不可-		
被験者 3				被験者 8			
	速い	どちらでもない	遅い		速い	どちらでもない	遅い
AC	9	0	0	AC	8	1	0
DE	9	0	0	DE	8	1	0
EV	9	0	0	EV	9	0	0
-算出不可-				p=.583			有意差 n.s.
被験者 4				被験者 9			
	速い	どちらでもない	遅い		速い	どちらでもない	遅い
AC	9	0	0	AC	9	0	0
DE	9	0	0	DE	9	0	0
EV	9	0	0	EV	9	0	0
-算出不可-				-算出不可-			
被験者 5				被験者 10			
	速い	どちらでもない	遅い		速い	どちらでもない	遅い
AC	8	1	0	AC	9	0	0
DE	2	7	0	DE	9	0	0
EV	9	0	0	EV	9	0	0
p<.001			有意差	***	-算出不可-		

²¹ 回答が全て同じである場合は、カイ二乗検定を行うことができない。

表 5 : 被験者別結果 (normal)

被験者 1				被験者 6					
	速い	どちらでもない	遅い		速い	どちらでもない	遅い		
AC	0	6	3	AC	1	7	1		
DE	0	9	0	DE	4	5	0		
EV	0	9	0	EV	1	8	0		
p=.034			有意差.	*	p=.223			有意差	n.s.

被験者 2				被験者 7					
	速い	どちらでもない	遅い		速い	どちらでもない	遅い		
AC	1	7	1	AC	1	6	2		
DE	2	6	1	DE	1	2	6		
EV	1	7	1	EV	1	8	0		
p=.963			有意差	n.s.	p=.032			有意差.	*

被験者 3				被験者 8					
	速い	どちらでもない	遅い		速い	どちらでもない	遅い		
AC	0	8	1	AC	0	8	1		
DE	1	6	2	DE	0	7	2		
EV	0	9	0	EV	0	8	1		
p=.330			有意差	n.s.	p=.746			有意差	n.s.

被験者 4				被験者 9					
	速い	どちらでもない	遅い		速い	どちらでもない	遅い		
AC	0	9	0	AC	3	5	1		
DE	3	4	2	DE	2	7	0		
EV	0	9	0	EV	0	9	0		
p=.015			有意差.	*	p=.204			有意差	n.s.

被験者 5				被験者 10					
	速い	どちらでもない	遅い		速い	どちらでもない	遅い		
AC	0	8	1	AC	0	9	0		
DE	0	5	4	DE	5	4	0		
EV	0	8	1	EV	0	8	1		
p=.145			有意差	n.s.	p=.007			有意差.	**

表 6：被験者別結果 (slow)

被験者 1				被験者 6					
	速い	どちらでもない	遅い		速い	どちらでもない	遅い		
AC	1	0	8	AC	0	2	7		
DE	0	1	8	DE	0	2	7		
EV	0	0	9	EV	0	1	8		
p=.395			有意差	n.s.	p=.782			有意差	n.s.

被験者 2				被験者 7					
	速い	どちらでもない	遅い		速い	どちらでもない	遅い		
AC	0	1	8	AC	0	0	9		
DE	0	3	6	DE	0	1	8		
EV	0	0	9	EV	0	0	9		
p=.128			有意差	n.s.	p=.354			有意差	n.s.

被験者 3				被験者 8			
	速い	どちらでもない	遅い		速い	どちらでもない	遅い
AC	0	1	8	AC	0	0	9
DE	0	1	8	DE	0	0	9
EV	0	0	9	EV	0	0	9
p=.583			有意差	n.s.	-算出不可-		

被験者 4				被験者 9				
	速い	どちらでもない	遅い		速い	どちらでもない	遅い	
AC	0	0	9	AC	0	0	9	
DE	0	0	9	DE	0	2	7	
EV	0	0	9	EV	0	1	8	
-算出不可-				p=.324			有意差	n.s.

被験者 5				被験者 10			
	速い	どちらでもない	遅い		速い	どちらでもない	遅い
AC	0	0	9	AC	0	0	9
DE	0	0	9	DE	0	0	9
EV	0	0	9	EV	0	0	9
-算出不可-				-算出不可-			

以上の表 4～6 に被験者ごとの結果を示した。fast 群 (表 4) はほとんどの被験者が AC・DE・EV 問わず「速い」と回答しているものが多いが、被験者 5 の DE のみ「速い」の回答がわずか 2 になっており、カイ二乗検定でも 1 名のみ有意という結果になっている。normal 群 (表 5) では、全体的には「どちらでもない」を多く回答している被験者が多いが、細かく見るとかなり個人差の大き

い結果になっている。被験者 1,4,7,10 は、カイ二乗検定で有意という結果になった。特に被験者 4,7,10 は DE での「どちらでもない」の回答が半数を下回っている。slow 群 (表 6) は一部の被験者の一部の回答で「どちらでもない」などが選択されているが、多くが「遅い」と判定されていることがわかる。個別のカイ二乗検定で有意となった被験者はいなかった。

2.2.6 考察

各速度群の回答の割合についてそれぞれカイ二乗検定を行ったところ、fast 群と normal 群では有意差が見られ、3 種の読み方 (AC・DE・EV) の回答割合が有意に等確率ではないという結果が得られた。

しかし、各被験者の結果を比較すると fast 群において 1 名の被験者 (被験者 5) のみが他の被験者と大きく異なる回答をしている。よって、この被験者 5 の回答が fast 群の統計結果に大きく影響している可能性が考えられる。もしそうであれば被験者全体の特徴をとらえた結果とは断言することはやや難しくなる。そこで、被験者 5 を除いた 9 名の結果でカイ二乗検定を行った。その結果、($\chi^2 = 3.361, df = 2, p = 0.186$) となり、有意差はみとめられなかった。このことから、fast 群全体の結果から見られた読み方ごとの回答の偏りは、被験者 5 の回答が特殊であったためとも考えられる。しかし、本稿の実験のみでそれを断言することも難しい。

normal 群についても、本節の冒頭で述べたように全体についての統計結果は有意であった。加速・減速音声で「どちらでもない」の回答割合が大きく下がっていたことから、normal 群では、速度が徐々に変化する加速・減速音声では速度判断にゆれが生じやすい傾向があると言える。

slow 群において、統計結果に有意差はあらわれなかった。被験者ごとの統計結果でも有意差があらわれなかったことから、slow 群では読み方が速度判断に影響を与えなかったと考えてよいだろう。以上、3 群の結果とその傾向を以下の表 7 にまとめた。

表 7：実験 B の回答

速度群	統計的有意差
fast	22
normal	
slow	×

では、速度群によってなぜ統計的な有意差の強さに違いが出たのであろうか。具体的には、normal 群においてのみ問題なく有意差を認めることができたのに対し、slow 群では有意差はあらわれず、fast 群では全員の結果では有意差が認められたものの、特徴的な 1 名の被験者の結果を除くと有意差が見られなくなった。

ひとつには、ゆれが観察しやすいか否かが理由のひとつとしてあるだろう。normal 群は判断にゆれが生じた際、速い方にずれても遅い方にずれてもそのゆれを観察することができる。それに対してたとえば fast 群は「速い」に回答が集中しているため、より遅い方向に判断のずれが生じた場合「どちらでもない」「遅い」が選択されるので観察することができるが、もしより速い方向に判断のずれが生じた場合にそれに相当する選択肢がないために観察することができない。slow 群も「遅い」に回答が集中しているのも同様である。それに比べ、normal 群は中間の選択肢である「どちらでも

²² 被験者全員での統計結果は有意であったが、回答傾向が特徴的だった被験者 5 の結果を除くと有意差が出なかったため、²²とした。

ない」の回答が最も多くなっているため、速い方向に判断がずれて「速い」が選択されても、遅い方向に判断がずれて「遅い」が選択されても、そのゆれを観察することが可能である。そのため、fast 群と slow 群で統計的な有意差が検出されなかった可能性も考えられる。

しかし、normal 群の音声に限れば、音声が一定の速度で読まれているか、加速していか減速しているかで、速度感の判断に違いがあらわれるという結果が得られた。

3. 実験 A と実験 B の比較

本稿では、速度変化をともなうと音声の聴覚的な速度感がどのように変化するかを、2 つの実験を行うことで検証した。実験 A では聞き比べを行ってそれぞれのトークンの速度感の相対的な差を算出することで尺度化を試み、実験 B では各トークンを単独で聞いた際の速度の印象を選択肢から選んでもらうことでそのトークンが聴覚的にどのような速さで聞かれているのかを観察した。その結果、2 つの実験の結果から、共通する特徴と異なる特徴が観察された。

まず共通点としては、速度変化しない音声の速度感と速度変化する音声の速度感が、必ずしも一致しない点が挙げられる。具体的には、実験 A における fast 群と normal 群では速度変化がないものとあるものの相対的な位置が統計的に有意に離れていた。なお、fast 群では速度変化があるものが遅いほうに偏り、normal 群ではそれとは逆に速いほうに偏っていた。実験 B では normal 群で速度変化の有無と変化の方向によって、速度判断の違いが見られた。速度判断のない音声では判断が比較的安定していたが、変化するものでは判断にゆれが見られた。

このことから、速度変化のある音声は速度が一定の音声と聴覚的な速度を比較した際、たとえ全体としての発話速度が近似であっても異なる速度感が得られやすいということが、一般的な傾向として言えるだろう。

また、slow 群で読み方による差があまり大きく現われなかった点も挙げられる。実験 A では slow 群で読み方による速度判断の有意差は見られなかった。実験 B においても読み方による統計的な有意差はあらわれなかった。全体的な速度が遅い音声では、速度変化の有無が速度間に影響しにくいと考えられる。

2 実験の相違点としては、まず一点として実験 A の slow 群では速度変化の有無が聴覚的な速度間の相対差として表れなかった点である。実験 B の slow 群では統計的には有意差はあらわれなかったが、速度変化のない音声の速度判断がもっとも安定しており、速度変化のある音声はそれに比して回答にゆれがあった。

もうひとつの相違点としては、速度の変化がない音声と比べた際に速度変化のある音声は速いほうに回答がずれやすいのか遅いほうに判断がずれやすいのかという点である。実験 A では速度変化を伴うと fast 群は遅いほうに、normal 群は速いほうに相対的な位置が移動した。実験 B では normal 群だけが実験パラダイムの両方向の速度感のずれを観察できたが、一概に速いほうに判断がずれやすいということはなく、速いほうにも遅いほうにもほぼ均等にずれていた。

実験方法によってこのような相違点があらわれたのは、一つには統計的な処理の違いにも原因があるだろう。相対的な位置関係を算出するのと、三択での回答を分析するのとでは同じトークンを扱っていても異なる結果が出てくる可能性は十分にあり得る。しかし、それ以外にも原因が考えられる。すなわち、実験 A は速度の聞き比べだったのに対し、実験 B では各トークンについての速度判断を求められた点である。同じように聴覚的な速度を扱っていても、速度比較の際に聞き手が用いるストラテジーと、単純な速度判断で用いるストラテジーが同じであるという保障はない。2.1.6 で、実験 A ではある特定の速度帯に聞き手がより注目して聞く可能性を述べた。実験 A では被験者の負担を考えて一対比較法の中でも比較的作業量の少ないサーストンの方法を用いたため、被験者

が必ずどちらかの音声が速いかを判断しなければならなかった。ゆえに、判定がシビアになった際はある特定の部分により注目してそこを判断のよりどころとして重く扱ったという可能性も考えられる。速度判断と一口に言っても、状況によってその判断基準は異なるものと思われる。

4. まとめと展望

本研究では速度変化をともなう音声が、速度が一定の音声と比べて速度感にどのような違いがあるのかを観察するため、2種類の知覚実験を行った。これらの実験の結果、速度変化をともなう音声の速度知覚のされかたは速度変化をともなわないものと同様ではないことが観察された。具体的には、速度変化をともなうと速度がより速く感じられたり遅く感じられたり、速度判断にゆれが生じやすいという傾向が確認できた。

しかしその判断基準は、聞き手や判断の状況、トークン全体の速度などによって異なると考えられる。本実験では速度判断に個人差が見られたり、全体の速度が遅くなると発話中に速度が変化しているにも関わらず、速度判断にあまり影響が見られないという結果が得られた。なぜこのような結果が得られたのかについて本実験の結果だけでは十分に議論できない。これについては、新たな実験パラダイムをもって検証していくべきだろう。

音声言語の速度感が定量的な単位である発話速度とある程度相関することが丸島 (2009) で述べられているが、現実の音声ではしばしば発話の途中でその速度が変化する。そして発話内で速度が変化するだけでその速度知覚は大きく変化する。しかもそれは一貫した変化ではなく、聞き手や速度判定の状況など、さまざまな要因に影響を受けると考えられる。

【参考文献】

- 福盛貴弘 (2008) 「ニュース番組におけるアナウンサー・キャスターの発話速度 —2006年5月3日のニュース番組を資料として—」大東文化大学外国語学部創設三十五周年記念論文集: 191-209
- 広実義人 (1994) 「知覚上の発話速度に及ぼすポーズ数の影響」『日本音声学会会報』205: 63-65
- 籠宮隆之・山住賢司・榎洋一・前川喜久雄 (2008) 「自発音声における大局的な発話速度の知覚に影響を与える要因」『音声研究』12-1: 54-62
- 加藤宏明・津崎実・匂坂芳典 (2004) 「音声のリズム・テンポのきこえとそのしくみ 持続時間長とタイミング処理の違い」音声文法研究会編『文法と音声<4>』207-229. ころしお出版
- 小林聡・北澤茂良 (1996) 「音声の高さ、大きさ、速さ感覚と物理関連量」『電子情報通信学会技術研究報告』NLC96-38, SP96-69: 1-8.
- 丸島歩 (2007) 「発話速度の実験音声学的研究 聴取側の視点から」筑波大学大学院人文社会科学研究科 中間評価論文(修士論文)
- 丸島歩 (2008) 「発話速度の知覚に関する一考察 - 基本周波数変動との関連性に着目して - 」『言語学論叢』オンライン創刊号: 70-85
- 丸島歩 (2009) 「音声言語のテンポに関する一考察 時間構造とピッチ構造に着目して」『言語学論叢』オンライン版第2号: 48-56
- 杉藤美代子 (1999) 「ことばのスピード感とは何か」『言語』28-9: 30-34
- 内田照久 (2000) 「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: 396-405

Perception of Gradually Changing Speech Rate

Ayumi MARUSHIMA †

The purpose of this paper is to observe how speech with gradual changes in rate is perceived. I created 27 stimuli in Japanese to solve this question. They consisted of three kinds of speech rates (fast, normal, and slow) and three kinds of ways of reading (even, acceleration, and deceleration). Utilizing these stimuli, I conducted two experiments on 10 Japanese native speakers.

In the first experiment, the styles of reading were found to affect judgments of speech rate on the stimuli.

In the other, the judgments were found not to be fixed, when the subjects listened to the stimuli with changes in rate.

†*Doctoral Program in Literature and Linguistics*
University of Tsukuba
1-1-1 Tennodai, Tsukuba, Ibaraki 305-8571, Japan
E-mail: ayumi_marushima@yahoo.co.jp