

速さの異なる読み上げ音声の 時間的特性*

丸島 歩

要 旨

異なる速さで読み上げられた音声がどのような時間的特性を持つのかについて明らかにするために、6種類のテキストを3種類の速さで読み上げた音声进行分析した。

その結果、発話速度は意図した速さにしたがって変化するが、調音速度はあまり変化しない場合があった。また、句点位置にあるポーズは速さにしたがって時間長が変化するが速さに関わらずほぼ全ての位置にポーズが挿入された。それに対し、読点位置にあるポーズは速くなってもその時間長はあまり変化しない代わりに、速さが速くなるにしたがってポーズが挿入されにくくなる傾向が見られた。

キーワード

発話速度 調音速度 ポーズ 読み上げ音声 句読点

1. はじめに

従来、聞き手にとっての音声言語の速度感について、いくつかの研究がなされてきた。最近のものでは、籠宮ほか (2008) などがある。筆者自身も丸島 (2008) で聴覚的な発話速度が音声のピッチ動態に影響を受ける可能性を、丸島 (2009) ではポーズを含めた発話速度¹が調音部分のみを問題にした調音速度²よりもより速度感に強い影響がある可能性をそれぞれ指摘してきた。

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について、広実 (1994)、杉藤 (1999) が挙げられる。広実 (1994) では、ニュース原稿の一部を日本語母語話者に音読させた資料 (ポーズ 11 個) と、短いポーズを長いポーズと合体させてポーズの合計時間長は同じだがポーズ数を少なくした資料 2 つ (ポーズ 7 個・ポーズ 4 個) の、計 3 種の資料を作成し、ほかの日本語を母語とする被験者に 2 つを一对として提示して聴かせ、

* 本稿は、2010年日本音声学会研究大会 (於・國學院大学渋谷キャンパス) での口頭発表の内容の一部を加筆・修正したものである。コメント・アドバイス等を下さった方々に厚く御礼を申し上げたい。

¹ 発話全体における、1秒あたりのモーラ数で算出した。ポーズを含めた発話全体の時間長をもとにして算出される。Speech Rate。

² 発話速度がポーズを含めた発話全体の時間長をもとに算出されるのに対し、ポーズを除いた発話部分から算出された値。Articulation Rate。

どちらがより速く聴こえるかを強制2択式で判断させた。結果、ポーズの数が少ないものほどより速く聴こえるという結果になった。

杉藤(1999)では、一般に早口と言われているテレビタレントなどの談話から、発話部分の一秒毎の平均拍数と最高拍数を算出した。また、それらの音声を大学生10名の被験者に聴かせた。その結果、特に速い話者1名以外はそれほど大きな違いがなかった。また聴き取り実験では、その最も速く話していた話者の音声だけではどの被験者にも内容が掴めなかった。さらに、速さ感覚が内容の理解とも関わりがあることが示唆された。次に、ニュースの音声からポーズを全て切除したものとそのままの音声を比較して20名の被験者に聴かせた。ポーズを切除したものは速いと感じられ、内容の理解も難しいという結果になった。また筆者は、ポーズが長ければ発話速度が速くても速く感じないと主張している。

以上の研究では主に聴覚的な速度感にポーズがどのように影響するかについて触れられているが、速度感についての知見を深めるためには、実現としての音声の実態を明らかにする必要があるだろう。すなわち、上の研究で扱われているのは自然言語だけではなく、一部の要素に音響的な操作を施した半合成音声である。当然このような音声をを用いることは、実験条件の統制の必要性から考えて当然かつ重要なことである。しかし聴覚的な速度感が影響を受ける要因について、現実の音声言語に戻って観察し直してみる必要があるのではないだろうか。聞き手と話し手の両方を視野に入れることで、言い換えれば、速さが異なるものとして実現された音声と、速さが異なって聞こえる音声の特性を両輪としてとらえることで、音声の速さを総合的に捉えることができるのではないかと筆者は考える。

そこで本研究では、発話者が意図的に速さを変化させた音声の時間的特性、特にポーズの実態を明らかにし、先行研究で明らかになった速度感にポーズが及ぼす影響についてもあらためて考察したい。

2. 目的

本研究での目的は、発話者が意図的に話す速さを変えることで、同一内容で速さの異なる音声を得、それらがポーズを含めてどのような時間特性を持つのか、また、速さが変化することでそれらの特性がどのように変化するのかを明らかにすることである。

3. 方法

3.1. 音声資料・録音方法

音声資料として、100モーラ前後の文章6種類を用いた。具体的には、説明文(テキストa³)、エッセイ(テキストb)、哲学書(テキストc)、小説(テキストd)、映画脚本(テキストe)、演劇脚本(テキストf)である。これらの文章を3種類の異なる速さ⁴で読んでもらい、

³ 全ての読み上げテキストを本稿の末尾に付した。

⁴ 以下本稿では、特に断りがない場合「速さ」という用語を発話者の意図としての速さのことを指すこととする。

それぞれを“normal”、“fast”、“slow”とした。具体的な指示として、まず特に指示を与えずに読んでもらい (“normal”)、次に不自然でない範囲で速く読むように指示し (“fast”)、最後に不自然でない範囲で遅く読むように指示した (“slow”)。それぞれのテキスト、速度について2回ずつ読むように指示した。一部に流暢に発話されていないものがあったため、より滑らかに発話されている1回分のみを分析対象とした。

3.2. 録音環境・機器

録音は静穏な室内で行った。パーソナルコンピュータ (NEC 製 VersaPro VJ17M/ED-19) にオーディオインターフェース (CREATIVE 社製 Sound Blaster Digital Music SX) を USB 接続で介し、マイク (オーディオテクニカ社製 AT-VD4。吹かれノイズを避けるため、スポンジ状のウインドスクリーンをかぶせるか、ポップスクリーンを用いた) を接続して行なった。録音ソフトは Cool Edit2000、OS は Windows XP Home Edition である。ファイル形式は Windows PCM、モノラル、サンプリング周波数は 44.1kHz、量子化 16bit である。

3.3. 被験者

読み上げ音声の録音に抵抗感が少ないと思われる、アナウンス・演劇などの発話訓練の経験のある男女4名を対象とした。

01⁵ (YK)、21歳男性。言語形成地は長崎県長崎市。

02 (HK)、20歳男性。言語形成地は東京都青梅市 (8~10歳時は広島県広島市)。

03 (MK)、25歳女性。言語形成地は東京都東久留米市。

04 (MA)、26歳女性。言語形成地は東京都東久留米市。

3.4. 解析方法

原波形とスペクトログラムをもとに、ポーズ部分と発話部分を計測した。その際、200ms以上の無音部分と句点直後の無音部分を暫定的にポーズと設定した。解析ソフトは Praat。Ver.4.6.22 と 5.1.43 を用いた。

さらに、計測したポーズ部分と発話部分の時間長から、発話速度⁶と調音速度⁷を算出した。

3.4.1. 発話速度と調音速度

以上で算出した発話速度と調音速度の統計分析を行なった。速度、テキストの種類をそれぞれ要因とした二元配置分散分析 (対応あり) を行なった。そのうち、有意な差の現れた要因について LSD の多重比較を行なうこととした。また、交互作用が見られた要因については単純主効果について分散分析を行い、有意差が出たものについて LSD の多重比較を行な

⁵ 01~04の数字は被験者番号である。

⁶ 脚注1参照。

⁷ 脚注2参照。

った。

3.4.2. ポーズ位置

テキストとポーズの関係を観察するために、計測したポーズがどの位置に現れたかを集計した。具体的には、句点位置・読点位置・それ以外、である。句読点の位置については、各被験者が速さによってポーズの出現のしかたが異なるかどうかを判断するために、コクランの Q 検定を行った。さらに「それ以外」の位置にあるポーズの数について、意図した速さとどのような関係にあるかを明らかにするために、速さを要因とした一元配置分散分析 (対応あり) を行なった。

3.4.3. ポーズの時間長

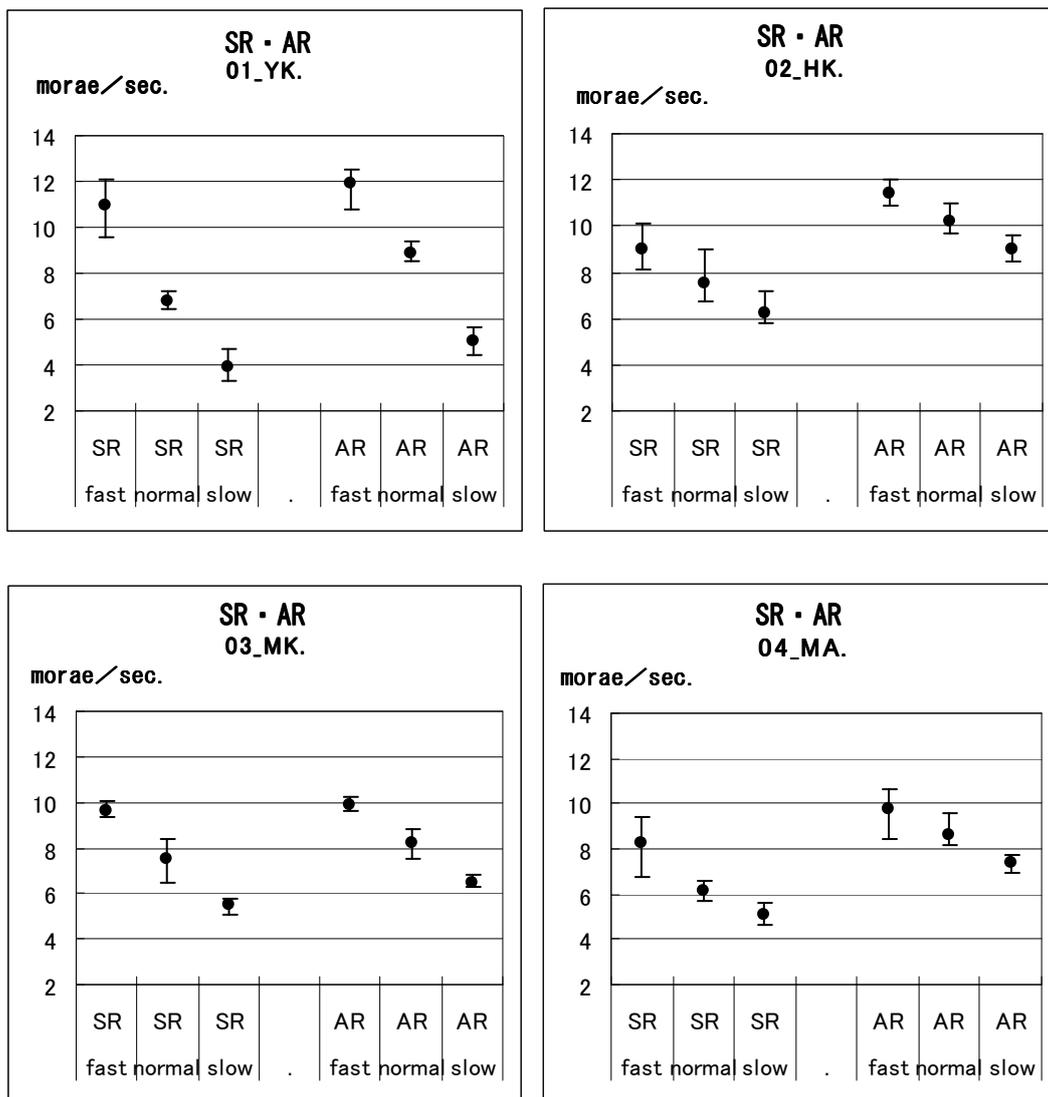
各ポーズの時間長が速さ・位置によって異なるか否かについて、シェッフェの多重比較を行った⁸。

4. 結果

4.1. 発話速度と調音速度

以下の図1は被験者ごと、速さごとの発話速度 (SR) と調音速度 (AR) である。それぞれについて平均値と最大値、最小値を示してある。

⁸ 本来であれば、位置と速度をそれぞれ要因とした分散分析を行うべきであろうが、サンプル数が著しく異なるのと、欠損した値が多い (特に句読点のない位置でのポーズは、速度や被験者によっては全く現れない場合があった) ために、ここでは用いなかった。



(平均値・最大値・最小値を記した)

図1 各被験者の音声の発話速度・調音速度

4.1.1. 発話速度

発話速度と調音速度それぞれについて二元配置分散分析 (対応あり) を行った。速さとテキストの種類をそれぞれ要因として設定した。

まず発話速度だが、速さ { $F(2,6) = 18.99; p < .01$ } とテキスト { $F(5,15) = 6.48; p < .01$ } の要因それぞれの主効果に有意差が見られた。また、速さとテキストの交互作用は見られなかった { $F(10,30) = 1.74, n.s.$ }。

そこで、それぞれの要因について LSD の多重比較を行なった。有意水準は 5%、有意傾向の基準を 10% とした。速さの要因については、fast が normal より、fast が slow より、normal が slow よりそれぞれ速いという結果になった (いずれも $p < .05$)。テキストの要因について

は、テキスト a より f が、d より b が、c より f が、d より e が、d より f がそれぞれ有意に速かった (いずれも $p < .05$)。

4.1.2. 調音速度

調音速度については、速さの要因の主効果に有意差がみられた $\{F(2,6) = 12.67; p < .05\}$ 。また、テキストの要因の主効果に有意傾向が認められた $\{F(5,15) = 2.28; p < .10\}$ 。さらに、速さとテキストの交互作用に有意傾向が見られた $\{F(10,30) = 1.95; p < .10\}$ 。

二要因に交互作用が有意であったため、それぞれの単純効果の分散検定を行なった。まず速さの要因に対しては、全てのテキストにおいて有意差が認められた {テキスト a では $F(2,6) = 11.96; p < .01$ 、b では $F(2,6) = 14.92; p < .01$ 、c では $F(2,6) = 12.36; p < .01$ 、d では $F(2,6) = 13.47; p < .01$ 、e では $F(2,6) = 11.64; p < .01$ 、f では $F(2,6) = 6.04; p < .05$ }。

さらに単純主効果に有意差が表れたものについて LSD の多重比較を行った。有意水準は 5% とした。まずは速さの単純主効果について記述する。テキスト b、d、においては、すべての速度間 (fast>normal、fast>slow、normal>slow) において有意差が見られた (いずれも $p > .05$)。テキスト a、c、e、f においては、一部の速度間に有意差が見られた。具体的には、テキスト a、c は fast>normal、fast>slow で有意差が認められ (いずれも $p > .05$)、normal-slow 間には有意差が見られなかった。テキスト e は fast>slow、normal>slow で有意差が見られ (いずれも $p > .05$)、fast-normal 間には見られなかった。テキスト f では fast>slow にのみ有意差が見られたが ($p > .05$)、fast-normal 間、normal-slow 間には有意差が認められなかった。

テキストの要因については、速さが fast の場合に有意差が見られた $\{F(5,15) = 3.79; p < .05\}$ 。normal のときは有意差は見られなかった $\{F(5,15) = 0.61 \text{ n.s.}\}$ 。また、速さが slow の場合に有意傾向が見られた $\{F(5,15) = 2.54; p < .10\}$ 。

テキストの要因においても単純主効果に有意差が表れたものについて LSD の多重比較を行った。有意水準は 5% とした。テキストの単純主効果について述べる。fast においては、d>a、c>f、d>e、d>f で有意差が見られた ($p < .05$)。normal においては有意差が見られたものはなかった。slow においては、a>e、d>e、f>e で有意差が見られた ($p < .05$)。

4.2. ポーズ位置

以下の図 2 にテキストの句点位置にポーズが見られた割合を、図 3 に読点位置にポーズが見られた割合をそれぞれ示す。

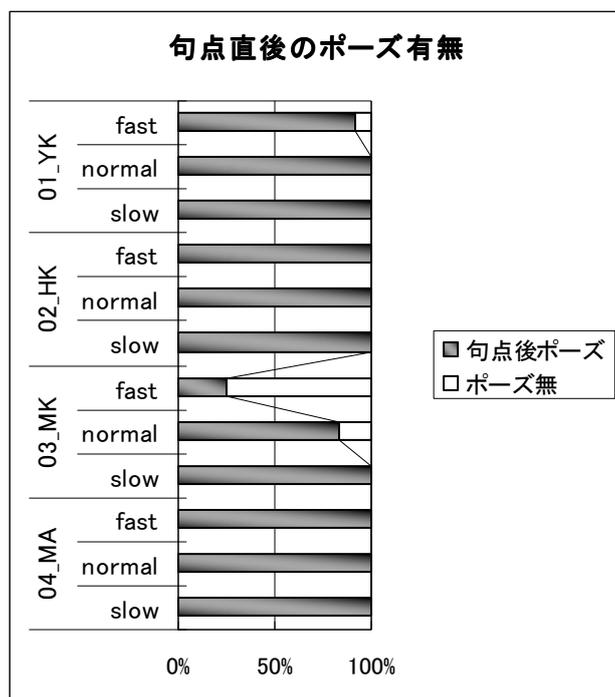


図2 句点位置のポーズの出現割合

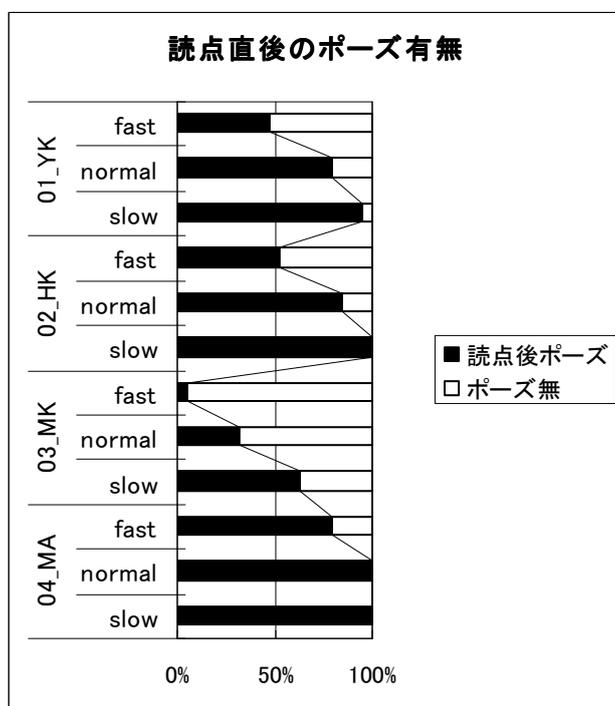


図3 読点位置のポーズの出現割合

句点位置は、速さを問わずほぼ全てにポーズが置かれているが、被験者 03 (MK) につい

でのみ、速さが速くなるほどポーズが置かれる割合が減少していることがわかる。句点位置にポーズがおかれた割合について、被験者ごとにコ克兰の Q 検定を行なったところ、被験者 03 (MK) のみ有意差が見られた $\{\chi^2(2)=14.88889; p<.01\}$ 。被験者 01 (YK) $\{\chi^2(2)=2.00000; n.s.\}$ では有意差が見られず、被験者 02 (HK)、04 (MA) では速さを問わず全ての句点位置にポーズが置かれていたため、検定を行なうことができなかった。

読点位置のポーズは、速度が速くなるにしたがって減少していく傾向がおおむね見られるが、被験者 04 (MA) に関しては fast 以外で 100%の値になっている。読点位置についても被験者ごとにコ克兰の Q 検定を行なったところ、全ての被験者において有意差が見られた (被験者 01 では $\{\chi^2(2)=24.57143; p<.01\}$ 、被験者 02 では $\{\chi^2(2)=14.00000; p<.01\}$ 、被験者 03 では $\{\chi^2(2)=15.16667; p<.01\}$ 、被験者 04 では $\{\chi^2(2)=8.00000; p<.05\}$)。

次に、図 4 にポーズが置かれた位置とそれぞれの回数を示した。ここでは、句読点位置以外に置かれたポーズに特に注目する。どの被験者も fast にはポーズが置かれておらず、全体としては速さが遅くなるほどポーズの数が多くなっている。このデータに速さを要因とした一元配置分散分析 (対応あり) を施したところ、速さの主効果に有意差が認められた $\{F(2,6)=5.92; p<.05\}$ 。これに LSD の多重比較 (5%有意水準) を行なったところ、fast<slow に有意差が認められた ($p<.05$)。fast-normal 間 (n.s.)、normal-slow 間 (n.s.) に有意差は認められなかった。

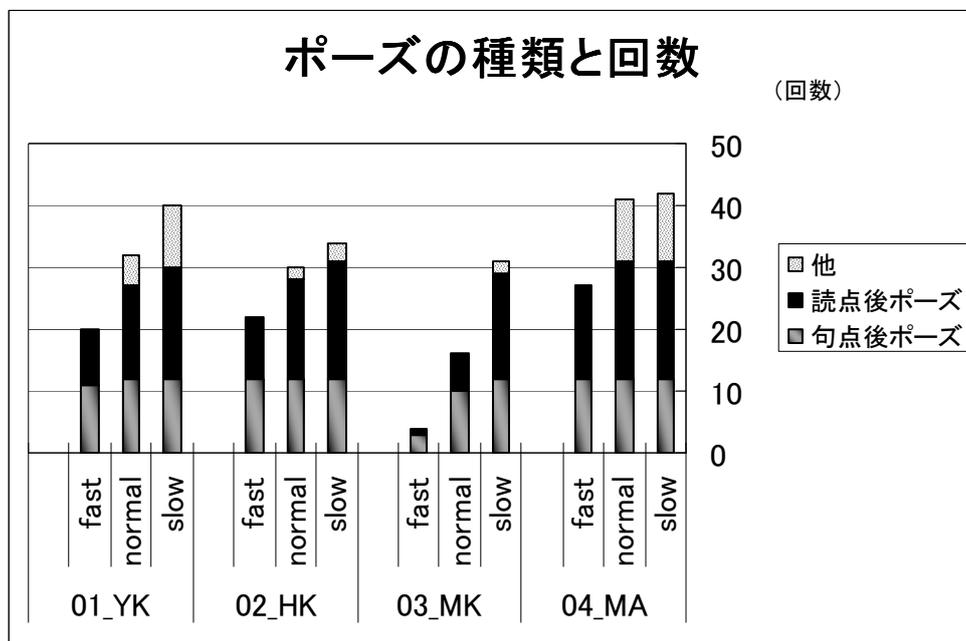


図 4 ポーズが置かれた位置と回数

4.3. ポーズの時間長

以下の図 5 は、ポーズの位置と速さごとのポーズの時間長の平均を、被験者ごとに示した

ものである。グラフには、それぞれの平均値と標準偏差を示した。

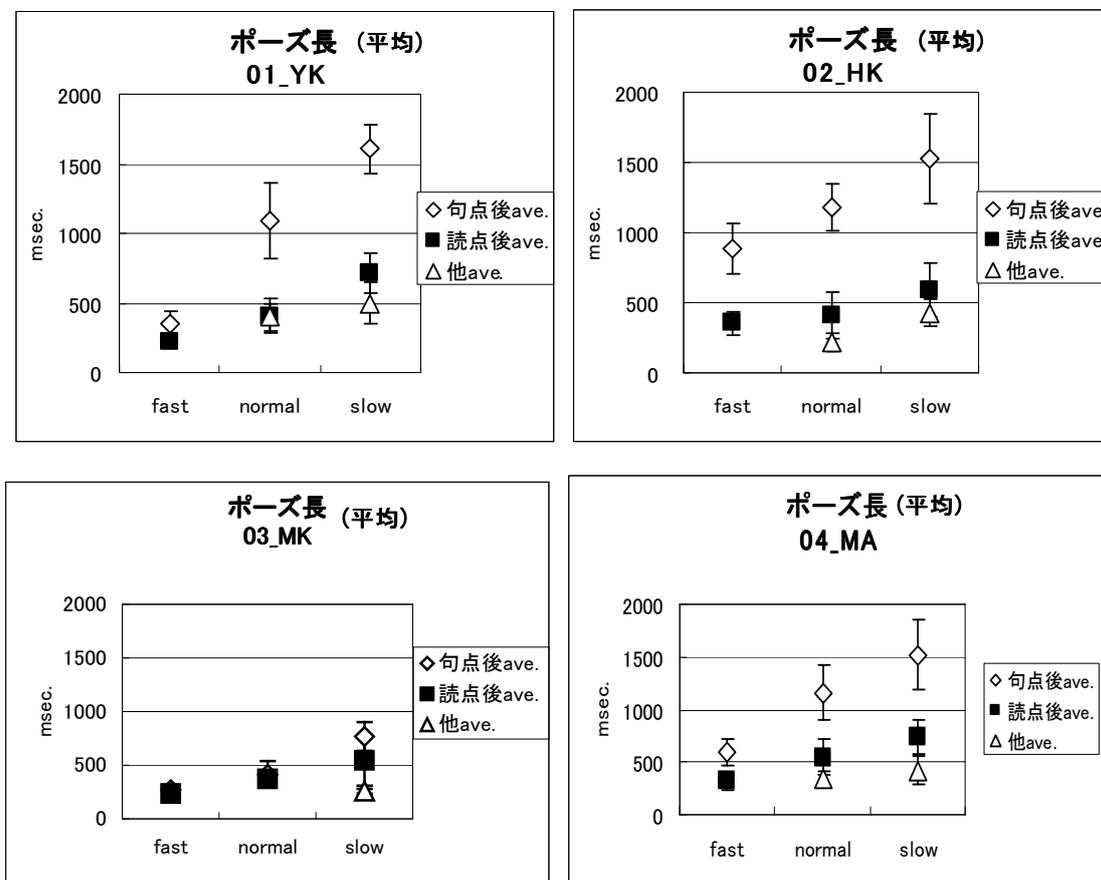


図5 速さ・ポーズ位置ごとのポーズの時間長

被験者ごとに、ポーズ位置を要因としたシェッフェの多重比較を行なったところ、以下の表1のようになった。なお、「**」は $p < .01$ であることを示している。

表1 ポーズ位置を要因とした多重比較の検定結果

		01YK	02HK	03MK	04MA
句点後	読点後	**	**	n.s.	**
句点後	その他	**	**	n.s.	**
読点後	その他	n.s.	n.s.	n.s.	n.s.

さらにそれぞれの被験者について、ポーズの位置ごとに速さを要因としたシェッフェの多重比較を行なったところ⁹、以下の表2のようになった。なお、「*」は $p < .05$ であることを

⁹ 句読点位置以外である「その他」にポーズが現れず、三者の比較ができなかったものに関しては、Welchのt検定を行なった。具体的には、被験者01(YK)と02(HK)の「その他」である。

示し、「**」は $p < .01$ であることを示している。値が欠損しており比較が行なえなかったものについては「-」と記した。

表2 速さを要因とした多重比較の検定結果

	水準1	水準2	句点	読点	その他
01YK	fast	normal	**	n.s.	-
	fast	slow	**	**	-
	normal	slow	**	**	n.s.
02HK	fast	normal	*	n.s.	-
	fast	slow	**	**	-
	normal	slow	**	**	*
03MK	fast	normal	**	n.s.	-
	fast	slow	**	n.s.	-
	normal	slow	**	n.s.	-
04MA	fast	normal	**	**	n.s.
	fast	slow	**	**	n.s.
	normal	slow	**	**	n.s.

5. 考察

5.1. 発話速度と調音速度

発話速度は、速さが速くなるほどその値が有意に高くなっている (図1、4.1.1 参照)。このことから、発話速度は意図した速さにほぼ対応しているとみて良いだろう。それに比して、調音速度は必ずしも速さに対応していない (図1、4.1.2 参照)。すなわち、速さを変化させて発話したにも関わらず調音速度に有意な差が見られない場合があったということである。意図した速さと調音速度が対応しているのは、テキスト b、d のみである。ちなみにテキスト b はエッセイ、テキスト d は小説である。説明文であるテキスト a と哲学書のテキスト c では normal と slow の間に有意差が見られなかった。映画脚本であるテキスト e は fast と normal の間に有意差が見られなかった。演劇脚本であるテキスト f では fast と slow の間にのみ有意差が見られた。以上のことから、意図した速さによる調音速度が受ける影響はある程度文体による可能性がある。すなわち、本研究の結果においては硬い文体では normal-slow 間に差があまり見られず、自然発話を模した文では fast-normal 間に差が見られないという傾向が見られた。本実験においては、被験者が fast・slow の音声を産出する際に、「不自然にならない程度に速く (遅く)」という指示をしている。したがって、これは硬い文では調音が冗長になると不自然になると発話者には感じられ、slow での調音があまり遅くならないようにする意識が働いたという可能性が考えられる。それとは逆に発話を模し

た文では発話者がある程度統一された軽快さを意識し、fast と normal との間に調音の違いが生まれにくかったという可能性も考えられる。

また、テキストによる発話速度、調音速度の違いも見られた。発話速度においては発話者の意図した速さを問わず、小説であるテキスト d が b (エッセイ)、e (映画脚本)、f (演劇脚本) より有意に速く、演劇脚本であるテキスト f が a (説明文)、c (哲学書)、d (小説) に比して有意に遅いという結果になった。

調音速度においては、normal の速さではテキストごとの有意な差が見られなかった。速さを特に気にしない状況 (自然な速さで読まれた場合) では、テキストの種類を問わず、調音速度はほぼ一定であることは注目に値する。おそらく、テキストの読み上げという状況では、発話者それぞれにとっての「適当な」調音速度が存在するのであろう。発話速度ではテキストごとに句読点の数や一文の長さ、文の構造が異なるためにポーズの取り方がテキストによって変わり、結果として上のような結果になったと考えられる。

一方、速度を意図的に変化させた時の調音速度は、テキストごとの差が現れた。fast では小説であるテキスト d が a (説明文)、e (映画脚本)、f (演劇脚本) より有意に速く、演劇脚本であるテキスト f が c (哲学書)、d (小説) より有意に遅いという傾向が見られた。これは発話速度の結果と類似しているが、このような結果になったのが偶然によるものであるのか、何らかの理由によるものなのかは本実験でのデータからはわからない。また、slow では映画脚本であるテキスト e が a (説明文)、d (小説)、f (演劇脚本) に比して有意に遅いという結果になっている。

つまり、本研究で測定した発話速度は小説で速く実現され、演劇脚本で遅く実現されたということである。注目したいのは、この傾向が fast の調音速度とも normal の調音速度とも異なる点である。このような結果が現れた原因は定かではないが、速さを意図的に調整した音声で調音速度にこのようなバリエーションが見られたことは、見逃せない点であろう。

5.2. ポーズの出現とその位置

句点位置のポーズ出現割合についてコクランの Q 検定を行なったところ、被験者 03 (MK) でのみ有意差が現れた。ほかの被験者では速さを問わずほぼ 100% の箇所ポーズが現れていたことを考えると、被験者 03 (MK) のポーズの置き方がほかの被験者に比べ異なる特徴を持っていると言えるだろう。被験者 03 (MK) の事例が特殊な例であるのか否かについては、別の被験者群で検証するしかないが、個人差が見られた事例として無視できないだろう。

読点位置のポーズ出現割合についてもコクランの Q 検定を行なったところ、全ての被験者で有意差が認められた。読点位置にポーズをどの程度の割合で置くかは、意図された速さに大きな影響を受けることが示唆された。この傾向は句点位置のポーズとは異なっている。

以上の結果から、読み上げ音声においては句点と読点は異なる特徴をもって実現されると

考えて良いだろう。繰り返しになるが、句点についてはほぼポーズとして現れるが、読点にポーズが現れるかどうかは、発話者が意図した速さにより影響を受けるということである。

句読点位置以外のポーズの出現数についても、意図された速さによって有意に異なるという結果になった。ただし、本実験では fast と slow の間に有意差が見られたのみであった。

5.3. ポーズの時間長

図5・表1より、被験者03(MK)以外の被験者では句点位置のポーズがそれ以外の位置のポーズに比べて有意に長いという結果になった。読点位置と句読点がない位置におかれたポーズについては、どの被験者においても有意差が見られなかった。したがって読み上げ音声においては、句点がそれ以外とは明らかに区別されると考えて良いだろう。

さらに、速さを要因とした多重比較の結果から、句点位置のポーズ長に意図した速さが反映されやすいという傾向が見える(表2)。具体的には、全ての被験者において意図された速さが速いほど、句点位置のポーズの時間長が短くなっているということである。それに比べて読点位置のポーズでは、被験者03(MK)以外で slow の場合が fast・normal に比べて有意に時間長が長い、被験者04(MA)を除いて fast-normal 間の有意差は見られなかった。つまり、速さを操作する際に読点位置のポーズの時間長は、発話を遅くする際にしか変化が見られないという傾向が見られた。

5.4. 意図された速さと発話速度、ポーズ

5.1 から 5.3 での考察から、発話を意図的に速くしようとする際と遅くしようとする際とで、どのような時間的特徴に変化が現れるかの傾向が明らかになった。それを以下の表3にまとめた。おおよその傾向をみるために3名以上の被験者で見られた変化を○、1名のみもしくはどの被験者にも有意な変化が見られなかった特性については×を付した。それ以外の記号については、脚注6~10に詳細を記した。

表3 発話の速さを変化させた際の音声の時間的特性の変化

時間的 特性			句点位置の ポーズ		読点位置の ポーズ		句読点のない 位置のポーズ	
	発話 速度	調音 速度	回数	時間長	回数	時間長	回数	時間長
速くする	○	△ ¹⁰	×	○	○	×	△ ¹¹	_ ¹²
遅くする	○	△ ¹³	×	○	○	○	△ ¹⁴	×

上の表から、読み上げ音声においてはあらゆる音声の時間的特徴が、速さの操作に均等に用いられているわけではないことがわかる。具体的には、調音部分の速度についてはテキストによって、fast で有意に速いわけではなかったり slow で有意に遅いわけではなかったりする場合があった。句点部分のポーズについては、時間長の変化は見られたが、速さを問わずほぼ全ての箇所でもポーズが現れたため、ポーズの回数に変化はなかった。読点部分のポーズについては、句点とむしろ逆の傾向が見られた。すなわち、読点部分のポーズ数は速度によって変化する傾向があったが、ポーズの時間長は発話者が遅く発話しようとした際は長くなる傾向があったが、速く発話しようとした際にはあまり変化がなかった。句読点以外の位置のポーズについては、意図した速さによってある程度回数が変化する傾向が見られた。特に、全ての被験者で fast 発話において一度もポーズが置かれなかった点は注目して良いだろう。ただし、句読点のない位置でのポーズの時間長には、速さの違いによる有意な変化はほとんど見られなかった。

さらに、発話を速くしようと意図した際と、遅くしようと意図した際には、そのストラテジーに若干の違いが見て取れる。すなわち、読点位置にあるポーズの時間長を変化させるか否かという点である。速く発話した際は時間長にあまり変化は見られなかったが、遅く発話した際は時間長が長くなる傾向にあった。速く発話した場合と遅く発話した場合とでは、必ずしも同じ時間的特性が同じように操作されているわけではないことが示唆される。

5.5. 聴覚的な速度感と先行研究との比較

広実 (1994) ではポーズの数が音声の速度感に影響を及ぼすと述べられていた。本実験における速さの異なる音声においても、速さの違いによってポーズの数が異なっていた。ただし、それは句点位置に置かれたポーズ以外においてのことであり、句点位置では速さを問わずポーズが置かれる傾向があった。広実 (1994) では一文を刺激音として用いているた

¹⁰ テキストによって結果が異なったため、△とした。

¹¹ fast での回数について slow との間には有意差が見られたが、fast-normal 間には見られなかったため、△とした。

¹² fast ではいずれの被験者にも句読点以外の位置にポーズは現れなかった。

¹³ 脚注6に同じ。

¹⁴ slow での回数について fast との間には有意差が見られたが、normal-slow との間には見られなかったため、△とした。

め、句点位置でのポーズについては考慮されていないことから、句点位置のポーズの有無が読み上げ音声の聴覚的な速度感にどのように作用するかはここからは明らかにできない。しかし、句点以外の位置にどれだけの回数のポーズを置くかということは、聴覚的な発話の速度感においても発話の速さを発話者が操作する際にも重要な特性であると考えられる。

杉藤 (1999) ではテレビのニュースでの音声と、そこから全てのポーズを切除した音声で聴取実験を行なっている。本実験での用語に言い換えると、調音速度は等しいが発話速度の異なる音声ということができるだろう。ゆえに、一部の調音速度で意図した速さに対応しない箇所があった本研究の結果とは単純な比較ができないが、ポーズの有無が聴覚的な速度感にも、発話者が発話の速さを操作する際にも重要な要素となっていると行うことができるだろう。

また、杉藤 (1999) では「実際の発話速度¹⁵は速くても、各ポーズの時間が長いと、聞き手は速いとは感じない。同様の時間内での話でも、発話速度が遅く、ポーズが短い場合には、聞き手は速いと聞き取る」と述べている。発話者が速さを意図的に変化させた本稿での音声では、句点位置のポーズは読みあげの速さにしたがって時間長が変化するが、読点位置のポーズは発話を速くした際には時間長に有意な差がないという傾向が見られた。また、句読点のない位置のポーズは発話を遅くしても有意に時間長が長くなることはなく、発話を速くした際は時間長が短くなるどころか、全くポーズがなくなってしまった。杉藤 (1999) には上記の記述の根拠となったデータの詳細は示されていないため、置かれたポーズの位置と聴取者の速度感との関連性についてはここからはわからない。しかし、少なくとも発話者にとっての速さを扱った本稿では、速さが変化した際、ポーズの回数や時間長の変化のしかたがポーズが置かれた位置によって異なることが明らかになった。この特徴が聴取者にとっての速度の聞こえにも同様のことが言えるのかどうか、観察する必要があるだろう。

さらに杉藤 (1999) では、発話者の視点に立った際の発話の速さについても「誰でも早口でしゃべるとポーズの時間は減少する」と述べている。しかし本実験では繰り返しになるが、読点位置にあるポーズは発話を速くしても有意に短くはならなかった。つまり、発話を早くしても全てのポーズの時間長が必ずしも変化するわけではない。速さによって時間長が変化しやすいポーズと、そうでないポーズが存在するということができるだろう。

6. まとめと展望

本稿では、発話の速さを意図的に変化させた読み上げ音声を用いて、発話の速さが変化するのにもなってその時間的特性がどのように変化するかを観察した。その結果、速さを変化させる際に全てのポーズの数・時間長が同等に変化させられるわけではなく、ポーズの置かれる位置によって違いが見られた。また、発話を速くするか遅くするかによってもそのストラテジーに若干の違いが見られた。

¹⁵ ここでの発話速度とは本稿での発話速度を指すのか調音速度を指すのかは定かでない。

本稿の実験では、読み上げ音声を扱った。自然発話の音声では発話者の意図した速さによってどのように時間的特性に変化が起こるかについては、今後の課題として残されているであろう。特に本稿では句読点の位置を問題にできたが、自然発話においては速さの違いによってポーズが一様に変化するのか、置かれたポーズの位置環境によって変化の仕方が異なるのか、異なるならどのような環境の違いが変化の仕方に影響を及ぼすのかということが問題になるだろう。さらに自然発話を扱う場合、実験条件をどう統制するかを熟慮しなければならないであろう。

参考文献

籠宮孝之・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.

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.

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.

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注目して—」 『言語学論叢』 オンライン創刊号.

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.

(丸島 歩 筑波大学大学院生 ayumi_marushima@yahoo.co.jp)

読み上げテキスト

a. <説明文> ^{ひとくち}一口にサンゴ、^いと言ってもさまざま^{しゅるい}な種類があります。浅瀬^{あさせ}のサンゴはもろくて折れやすく、^お宝石^{ほうせき}になるような硬くて^{かた}美しいサンゴは、^{うつく}深い海^{ふか}に^{うみ}わずかしか^{せいそく}棲息しない、特殊^{とくしゆ}な種類^{しゅるい}なのです。

(出典：中村庸夫 (1994) 『サンゴ礁の秘密 彼らは僕らの肺である』 祥伝社)

b. <エッセイ> コドモ^{ころ}の頃、^{しゅうじ}習字の時間^{じかん}によくやってしまった^{しつぱい}失敗。一字一字^{いちじいちじ}を書くのに必死^{ひっし}なあまり半紙^{はんし}に入りきらなくて、最後^{さいご}は異様^{いよう}にちっちゃい字^じになってしまうんです。出来上^{できあ}がり^がり、^{わる}なんとも^じバランスの悪い字。

(出典：上大岡トメ (2004) 『キッパリ！ たった5分間で自分を変える方法』 幻冬舎)

c. <哲学書> また我々^{われわれ}は普通^{ふつう}に意志^{いし}は自由^{じゆう}であるといっている。しかし^いわゆる自由^{じゆう}とは如何なること^{いか}をいうのであろうか。元来^{がんらい}我々の欲求^{よくきゅう}は我々に与えられた者^{われわれ}であって、自由^{じゆう}にこれを^{しやう}生ずることはできない。

(出典：西田幾多郎 (1979) 『善の研究』 岩波書店)

d. <小説> 若荷谷^{みよがだに}さんは稽古場^{けいこば}で演技^{えんぎ}をつける時^{とき}みたい、テーブル^{えんりよ}を遠慮^{えんりよ}なくばんばん叩^{たた}きながら言^いった。こうい^いう時^{とき}はいくら理不^{りふじん}尽^{そくざ}でも即座^{ほう}にあやまった方^{ほう}がいいので、ぼくは小^こ声^{ごえ}で「すみません」と言^いった。

(出典：原田宗典 (1995)「何者でもない」『何者でもない』99-212 講談社)

e. <映画脚本> 今^{いま}インドってさあ、すんごい面白^{おもしろ}いのよ。都市部^{としぶ}の最新^{さいしん}型^{がた}建^{けん}築^{ちく}から、タージマハル^{たてん}みたい^な古典^{こてん}から、スラム^なのモスク^なから、何^{なに}から何^{なに}までい^いろいろあ^あって、結構^{けつこう}仲間^なも留^{りゅうがく}学^{がく}して^るんだ。で、留^{りゅうがく}学^{がく}して、力^{ちから}つ^つけて、イー^いうち^{うち}作^{つく}りたい^いじゃん!

(出典：大九明子 (2007)「恋するマドリ」『シナリオ』63-9 :66-98)

f. <脚本□(舞台演劇)> 最近^{さいきん}のビデオ^{かていよう}は、家庭^{ふつう}用^{あか}でも普通^との明^あるさ^さで撮^とれる。かえ^かって^りアル^るだ。ライテ^{そと}ィン^{つか}グ^{つか}も外^{はじ}づ^{はち}け^てマイク^ても使^しわ^しない。初^{しよ}め^{しん}て^か八^かミ^えリ^えを^みら^いに^らい^いに^いる。未^み来^{らい}の^みら^いの^みら^いの^みら^い人^{ひと}た^ちは^はな^なん^んて^てひ^ひど^どい^い画^えだ^だと^と思^{おも}う^うだ^だら^らう。

(出典：坂手洋二 (2004)「心と意志」『せりふの時代』9-1 :6-76)

Features of Different Reading Rates

MARUSHIMA Ayumi

The purpose of this study is to investigate what characterize(s) speech that has different rates. Four subjects read six kinds of text using three kinds of speech rate and I measured the poses duration and calculated the speech rates and articulation rates.

The results were as follows:

- (1) The speech rates decreased when the texts were read slowly, and increased when they were read quickly.
- (2) Some articulation rates did not decrease or increase, no matter whether the subjects read the texts quickly or slowly.
- (3) The duration of the poses at the periods (“。”) was shorter when the texts were read quickly, and longer when they were read slowly.
- (4) As the duration of poses at the commas (“、”) was longer when the texts were read slowly, it did not change when they were read quickly.
- (5) The subjects almost always inserted poses at the periods, whichever rates were intended, but at the commas they inserted more poses when they read slowly and fewer poses when they read quickly.