

発話速度の認知に関する一考察*

基本周波数変動との関連性に注目して

丸島 歩

要 旨

従来、発話速度は発出側の立場から一定時間におけるモーラ数などによって表わされてきた。本研究では聴取側の視点から、発話速度がどのように認知されるのかを探った。

モーラ数と時間長がほとんど変わらない音声資料を複数組み合わせることで、モーラ数以外に発話速度の認知に影響を与える要因を見定めた。その結果、基本周波数変動が大きいものの方が、そうでないものより発話速度が速いと認知されることが示唆された。

キーワード

発話速度 基本周波数変動 実験音声学 聴覚印象 聴取実験

1 はじめに

発話速度は「モーラ毎秒」のように、しばしば一定時間内のモーラ数で表される。それ以前は文字数で表されることが多かったという(最上勝也 1999)が、文字数は漢字を用いるか仮名を用いるかで全く同じ音声でも数字上異なってしまい、発話速度の表現には不適當であることは明白である。

では、文字数ではなくモーラ数に基準を置けば発話速度を忠実に表せると言えるだろうか。我々が日常において他人の話し方に対して「あの人は早口だ」などと評価するとき、それは一定時間内のモーラ数が平均よりも多いということを指しているのだろうか。必ずしもそうとは限らないだろう。

例えば小林聡・北澤茂良(1996)では、発話速度を含めたプロソディー情報の聴覚印象が物理量¹に必ずしも反映されない、ということが明らかになっている。そこで、発話速度の実態に迫るために聴取側の視点から検証を行なおうと考えた。つまり、聴き手の発話速度

* 本稿は、2007年度に筑波大学に提出した中間評価論文(主査:城生佰太郎先生)の一部に加筆・修正を行なったものである。この論文では脳波実験と音響・聴取実験を併用したが、本稿では枚数規定の都合上音響・聴取実験の一部分をまとめることとした。脳波実験での成果については、丸島歩(近刊予定)に譲る。

¹ 小林聡・北澤茂良(*ibid.*)で言われている「物理量」が具体的に何を指しているかは明記されていない。しかし、現段階で発話速度を表す指標として一定時間内のモーラ数が広く用いられていることから考えると、それと同様に近い基準を用いていると思われる。ただし、果たしてそのような基準を「物理量」と称するに相応しいかには疑問が残る。

認知に何が影響をもたらすかを明らかにする。特に本稿では、基本周波数変動との関係について明らかになった部分を扱うこととする

なお、本稿はボトムアップの研究姿勢をとっている。そもそも実験音声学が具体的な発声発語を観察するところからスタートする学問であり、事象探査型の帰納的方法を基本的な方法論としているためである (城生 佰太郎 2005: 11)。

また、発話速度を聴取側の視点から検証する研究は、筆者の認識する限り城生 佰太郎 (1999) のほかに過去に行われたことがない。このような先行研究のほとんど存在しない研究テーマにおいては、仮説を立ててそれを検証するというトップダウン式の研究姿勢をとることができない。従ってこのような研究姿勢を採ったことは必然であると言えよう。

2 目的

本稿の目的は、従来モーラ数を基準として計測されてきた発話速度について、基本周波数変動が聴覚レベルにおいてどのように影響を及ぼしているのかを明らかにすることである。

3 課題 : 通常発話の聴取課題

まずは自然発話・朗読など、あまり (もしくは全く) 統制されていない音声を用いた実験を行なった。

3.1 方法・手順

3.1.1 分析資料

市販の CD や DVD から音源を得たものと、独自に録音したものを分析資料とした。この資料の選定は一貫性に欠けるといふ謗りを免れ得ない。しかし、聴取側の観点による発話速度についての先行研究が筆者の知る限り存在しないため、どのような資料を用いるべきか見当がつかない。そこで話すことに熟達しておりなおかつ幾分の編集が加えられているであろう資料と、ごく自然な発話に近い音声資料の両方を使用することにした。

3.1.1.1 市販のメディアからの資料

市販のものについては映画とインタビューを用いた。映画の音声を使用したのはできるだけ現実の談話に近い場面での音声を得るためであり、インタビューの音声を用いたのは映画や演劇のように恣意的に発話されたものではない音声を得るためである。

なお、使用した CD・DVD は以下の通りである。

- ・ DVD 『男はつらいよ 8 寅次郎恋歌』松竹ホームビデオ
(1971 年公開・2002 年発売)
- ・ DVD 『明日の記憶』東映 (2006 年公開・同年発売)

・ CD 『サウンド文学館パルナス 1 高瀬舟』 学習研究社

(1995 年発売・江守徹朗読)

DVD についてはパーソナルコンピュータ (NEC 製 VersaPro VJ17M/ED-1) 上で DVD 再生ソフト Win DVD5.0 (Inter Video 社製) を用いて再生しながら、録音ソフトぼけっとれこーだー Ver2.17Final (P++ 氏作製・<http://chrono.pos.to/ppp/>にて無償配布) でサンプリングレート 44.1kHz・量子化 16bit・モノラルで wave ファイル化した。CD については、音楽ファイル変換ソフト Rip!AudiCO Ver4.03 フリーウェア版 (PINO 社製) でサンプリングレート 44.1kHz・量子化 16bit・モノラルで同様に wave ファイル化した。以上の作業での OS は Windows XP Home Edition である。

複数の音声を重ねておらず、BGM や大きな背景音が被っていない任意のまとまった一定時間分を切り出して用いた。その際、できるだけ切れても不自然でないところを切り出すよう努めた。

3.1.1.2 録音による資料 (自然言語音)

できるだけ自然談話に近いものを得ることを心がけた。実際の会話の音声を用いることが出来ればそれが最良の方法だが、複数の音声がかぶってしまうと音響解析が不可能になってしまうので、被験者が単独で話しながらもできるだけ自然に話すことができるように心がけた。具体的には、録音の前に被験者に「最近印象に残ったこと」というテーマを提示し、テーマに基づいて自由に話してもらうようにした。その際、話す話題についてあらかじめメモを用意する事を認めた。また、聞き手として筆者が第三者が頷くなどして、被験者の発話を促した。ただし、実際に過去に行なった発話を再現しようとした被験者もあり、その場合には特に働きかけはしなかった。また、「目上の人に話すつもりで」、「友達など、身近な人に話すつもりで」と 2 種類の状況を設定し、それぞれについて録音を行なった²。

実際に観察・実験を行なう際は、プロソディックパターンが崩れないように適当部分を切り出して用いた。

3.1.1.3 録音による資料 (読み上げ)

比較を容易にするために一部文章の読み上げも用いた。文章の読み上げは、ネットラジオやインタビューなど音声の文字起こし・脚本の台詞部分や書籍の文章を用いた。

² 録音はパーソナルコンピュータ (NEC 製 VersaPro VJ17M/ED-1) 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。サンプリングレート 44.1kHz・量子化 16bit でモノラル録音を行なった。この際、マイクと手の摩擦でノイズが発生しないように、マイクスタンドを用いた。なお、録音は静謐な和室で行なった。

文字起こしは以下の音声資料を、15000ms 分を目安に切り出した上で行なった³。

- ・ 時計 podcast (<http://gekidantokei.web.fc2.com/podcast.html>)
: 第 1 回 (2007/08/14 配信・2007/08/24 取得)
- ・ CD 『CD セレクション ラジオ深夜便 ラグビーの心～山口良治』

読み上げに用いた文章は以下から採ったが、録音の際にそれぞれを比較しやすくするためにあらかじめ文の適当な部分を選んでおき、15000ms を目安に読んでもらった。その際基本周波数などの動態にバリエーションを持たせるため、通常の音読と任意の 3 箇所を強調してプロミネンスをつけたもの、さらに一部は抑揚を抑えたもの⁴も併せて録音した。録音機材などは、自然言語音と同様である。

- ・ 『津軽三味線ひとり旅』より、解説「竹山さんとその芸について」(高橋竹山著。解説は佐藤貞樹。1991 年、中公文庫)
- ・ 『つっこみ力』(随想。パオロ・マツツァリーノ著。2007 年、筑摩書房)

読み上げについては、全ての被験者について同じ資料を用いたわけではないことを注記しておく。一定時間を目安に読まなくてはならない上、抑揚を操作する技術を要するため、被験者に重い負担をかけると同時に高い技術も要求されるからである。後述の通り今回の被験者は筆者を含めて 4 名だが、それぞれの技術の程度や向き不向きは異なる。そのためあらかじめ筆者が判断したり予備実験を行った結果をもとに、適当な資料を選んだ。

3.1.2 被験者⁵

録音を行なったのは、以下の被験者と筆者自身の音声である。以下に、録音時の被験者の情報を示す。

KR: 24 歳女性。言語形成期 (4、5~12、3 歳) は埼玉県所沢市で過ごした。

HS: 23 歳女性。生後から 8 歳まで鹿児島県鹿児島市、8 歳から 9 歳まで鹿児島県内之浦、9 歳から 10 歳まで千葉県鴨川市、10 歳から 18 歳まで鹿児島県鹿児島市で過ごした。

AT: 23 歳男性。言語形成期は石川県金沢市で過ごした。

筆者 (MA) は 24 歳女。言語形成期は東京都東久留米市で過ごした。なお、父は神奈川県出身・母は鹿児島県出身である。

³ 予備実験を行い、発話速度が認知されるに十分な時間長として 15000ms に設定した。

⁴ 当初はより「抑揚を大きく」という指示を出していたが、それがどの被験者にとっても難しかったため、逆に抑える方法に切り替えた。

⁵ 本来、実験音声学的研究を目的とする実験データは網羅的に公開されるべきであり、被験者についての情報についても例外ではない。父母の出身地・本人の言語生活についてなども記載すべきだが、個人情報の扱いに關しての法制限により被験者名はイニシャルのみとし、年齢・性別・言語形成期の居住地以外の情報は割愛する。また、敬称は省略する。以下同様。

3.1.3 解析装置

音響解析に用いたソフトは、Praat Ver.4.6.22 (<http://www.praat.org> にて無償配布) である。同ソフトには基本周波数の音圧の標準偏差や中央値を算出する機能がないことから、Excel2003 SP2 (Microsoft Office 社製) にデータを入力し、各統計データを算出した。解析に用いたのは、パーソナルコンピュータ (NEC 製 VersaPro VJ17M/ED-1) で、OS は Windows XP Home Edition である。

3.1.4 解析方法

時間長とモーラ数⁶が近似した複数の資料を組み合わせ、聞き取りによって聴覚印象上の遅速差を判断する。資料の一覧を、下の表 3.1 に提示する。聴取はヘッドフォンによった。その上で、それぞれの基本周波数の中央値・標準偏差・最大値・最小値を計測した。

また、遅速の判断は、筆者自身によった。本研究は前述したとおり、聴覚情報処理系の営みに着目している。発話速度の認知についてほとんど何も明らかにはなっていない現段階で聴取者として多くの被験者を用いることは、安定した結果を導けないおそれがあり危険である。本稿での実験において筆者が聴取者となったことで、安定した基準を得られたと確信する。また聴取者である筆者は、誓って資料を無心に聴いた上で自身の聴覚印象に忠実に遅速の判断を行なったことを、ここで注記しておきたい。

表 3.1 聴取実験解析資料 (市販メディア)

		A	B
比較 1		んでえ。たまに帰ってくらあ、どいつもこいつも膨れっ面ぱっかりしやがって。潰れた工場 (こうば) 行ってなあ、テメエの亭主の顔見てくるよ。おいちゃん、これあおいちゃんに買って来た土産の孫の手だ。ほれ。ったく。	ショーレン寺の和尚と二人であの寺の古文書を調べてるんだが、これがなかなか面白い仕事でね。ま、五年や十年はかかるだろう。や、身の回りのことは横山の婆さんがいてくれるから、何とか。
	時間長	約 15000ms	約 15000ms
	モーラ数	95 モーラ	96 モーラ
	話者	渥美清 (『男はつらいよ 8 寅次郎恋歌』より)	志村喬 (『男はつらいよ 8 寅次郎恋歌』より)

⁶ 1. において「一定時間内のモーラ数」だけが発話速度の認知を決定しているとは限らないと述べたが、ここではモーラ数を基準として解析を行なっている。先行研究の存在しない現時点においてはほかに妥当な基準を立てることができないこと、また、「一定時間内のモーラ数」をできるだけ揃えることによって発話速度の認知を決定する要因となるほかの要素を見出し得ると判断したことが、このような解析方法を選択した理由である。

比較 2		あたし達はね、あなた達が、子どもや旦那の面倒見てる何年もの間、ずーっと一人で働いてきたの。セクハラじじいに嫌味言われて、実家の親に孫は諦めるって泣かれながら、それでも一人でずーっと続けてきたの。だからこの年齢になっても続けられる仕事があんのよ。子育てがひと段落して、家にいてもつまないから外出で働きましょって思ったって、そんなうまく行くもんじゃないのよ。	去年の秋のことです。わたくしは弟といっしょに、西陣の織り場にはいりまして、空引き（そらびき）ということをしたことになりました。そのうち弟が病気で働けなくなったのでございます。そのころわたくしどもは北山の掘っ立て小屋同様のところに寝起きをいたして、紙屋川の橋を渡って織り場へ通っておりましたが、わたくしが暮れてから、食べ物などを買って帰ると、弟は待ち受けていて、
	時間長	約 30000ms	約 30000ms
	モーラ数	198 モーラ	196 モーラ
	話者	渡辺えり子（『明日の記憶』より）	江守徹（『サウンド文学館パルナス1 高瀬舟』より）

表 3.2 聴取実験解析資料（録音 - 読み上げ）

		A	B	C
比較 3		いつかは役に立つこともあるかと、できるだけ記録しておいただけで、発表目的の取材でもなく、もちろんこういう形で本にするなど全く考えてもいなかった。	いつかは役に立つことあるかと、 <u>できるだけ</u> 記録しておいただけで、 <u>発表目的</u> の取材でもなく、もちろんこういう形で本にするなど全く <u>考えても</u> いなかった。（下線部を強調）	/
	時間長	約 10000ms	約 10000ms	
	モーラ数	83 モーラ	83 モーラ	
	話者	MA（『津軽三味線ひとり旅』「竹山さんとその芸について」）	MA（『津軽三味線ひとり旅』「竹山さんとその芸について」）	

比較 4		今や関西以外の全国区でも当たり前のように使っている言葉なのですが、実は「ボケ」と「ツッコミ」という言葉の組み合わせがいつ頃から定着したのかは、定かではありません。言葉自体の歴史としては思いのほか浅いようなのです。	今や関西以外の <u>全国区</u> でも当たり前のように使っている言葉なのですが、実は「ボケ」と「 <u>ツッコミ</u> 」という言葉の組み合わせがいつ頃から定着したのかは、定かではありません。言葉自体の歴史としては思いのほか <u>浅い</u> ようなのです。(下線部を強調)	今や関西以外の全国区でも当たり前のように使っている言葉なのですが、実は「ボケ」と「ツッコミ」という言葉の組み合わせがいつ頃から定着したのかは、定かではありません。言葉自体の歴史としては思いのほか浅いようなのです。(抑揚を抑えて)
	時間長	約 15000ms	約 15000ms	約 15000ms
	モーラ数	118 モーラ	118 モーラ	118 モーラ
	話者	KR (『つっこみ力』)	KR (『つっこみ力』)	KR (『つっこみ力』)
比較 5		で、あのツインだったら、まだ良かったんだけど、あの、そんな時、北海道の函館に泊まったんですけども、シングルルームを予約してたの。で、フロントに行って、ま、「フクオカです、よろしく願いします」みたいなこと言ったら、また「シングルルームは埋まってます」って言われたんですよ。	で、あのツインだったら、まだ良かったんだけど、あの、そんな時、北海道の函館に泊まったんですけども、シングルルームを予約してたの。で、フロントに行って、ま、「フクオカです、よろしく願いします」みたいなこと言ったら、また「シングルルームは埋まってます」って言われたんですよ。(抑揚を抑えて)	
	時間長	約 15000ms	約 15000ms	
	モーラ数	133 モーラ	133 モーラ	
	話者	KR (時計 podcast)	KR (時計 podcast)	

比較 6		ほんつとにいい学校でしたけどねえ。そらもうあの学校行ってドキドキしながら。ラグビー部なかったんですけどね。その学校に赴任して、あの時はみんなそうだと思うんですけどね、教師になった時に、「おい、山口君。君の受け持ちのクラスだ」ってその名簿をもらうんですよ。	ほんつとにいい学校でしたけどねえ。そらもうあの学校行ってドキドキしながら。ラグビー部なかったんですけどね。その学校に赴任して、あの時はみんなそうだと思うんですけどね、教師になった時に、「おい、山口君。君の受け持ちのクラスだ」ってその名簿をもらうんですよ。 (下線部を強調)	ほんつとにいい学校でしたけどねえ。そらもうあの学校行ってドキドキしながら。ラグビー部なかったんですけどね。その学校に赴任して、あの時はみんなそうだと思うんですけどね、教師になった時に、「おい、山口君。君の受け持ちのクラスだ」ってその名簿をもらうんですよ。 (抑揚を抑えて)
	時間長	約 15000ms	約 15000ms	約 15000ms
	モー数	133 モー	133 モー	133 モー
	話者	AT(『CD セレクション ラジオ深夜便 ラグビーの心 ~ 山口良治』)	AT(『CD セレクション ラジオ深夜便 ラグビーの心 ~ 山口良治』)	AT(『CD セレクション ラジオ深夜便 ラグビーの心 ~ 山口良治』)

表 3.3 聴取実験解析資料 (録音 - 自然談話)

	A	B	
比較 7	大変、あたしにもできないくらい大変なことやってるんだけど、んで、わかんなかったらほら、子どもって、「わかんない」とか駄々こねたりとか、で、それが、こつ高じて、塾に来なくなっちゃったりとか、する子って、まあいるわけよ。	でも確かに表情が少ない人からやっぱり表情って学べないよなーと思って。やっぱりこつ、お母さんとか、になったときはやっぱり表情豊かに子どもを育てないとなーみたいなことはちょっと思ったね。	
	時間長	約 15000ms	約 15000ms
	モー数	100 モー	101 モー
	話者	HS (友達など、身近な人に話すつもりで)	HS (友達など、身近な人に話すつもりで)
比較 8	あさつての方向見るっていうか、手遊びとかしないのよ、ほんつとぼつとするのね。最初はすごくそれでびっくりして、あの一、なんかこつテキストを出させ、出させたりとか、いろいろ働きかけてたんだけど。	見れる芝居とかもない中で、あの一自分達なりに必死に頑張つて、まあ、あのお芝居作つてるんだなあと思つたら、ん一。あの一、私もまあ実習頑張ろうつていうふう思えたんですよ。	

	時間長	約 15000ms	約 15000ms
	モーラ数	94 モーラ	95 モーラ
	話者	HS (友達など、身近な人に話すつもりで)	HS (目上の人に話すつもりで)
比較 9		あれ、いやいいんだけど、全然いいんだけど、あのガクトが出てるじゃない。あれって今。で、あのガクトさあ、あの、一人だけ物凄い浮いてるって話は聞いてたの。ほんとに浮いてんのね。ねえ。だって絶対さ、日本人の顔つきじゃないじゃん。	台詞も気にしながら、周りの人がしゃべってる声も聞きながら合わせなきゃいけないって、それがすごい難しいのを再認識、しております。プロの人たちはやっぱりすごいんですね。現場とかに行くと、わかるんですけども。
	時間長	約 15000ms	約 15000ms
	モーラ数	108 モーラ	105 モーラ
比較 10	話者	KR (友達など、身近な人に話すつもりで)	KR (目上の人に話すつもりで)
		演じてた人が、ほんとに、な、だからだらだらだ、涙を流しながら読んで、「ああああ」って感動したのを、なー、も、私はすごいなと思って、も、その人みたいになろうと思った瞬間だったね。あれはまた再確認っていうか、再度思った。	最近やったのがドイツ語で、十歳の女の子の役を、やりまして。それがすごく難しかったんですけど、もう、声がもうあがっちゃってあがっちゃって。でも「もっと若くして」って言われたのがすごい大変だった、最近の出来事です。
	時間長	約 15000ms	約 15000ms
比較 11	モーラ数	112 モーラ	111 モーラ
	話者	KR (友達など、身近な人に話すつもりで)	KR (目上の人に話すつもりで)
		会話がかみ合わないのよ、これが。もうだからもう私、話すこと放棄しちゃって、で、ま、教材準備をするふりして、こ、ぼけーっと。も、い、きよ、教材準備をしながらこう、ほかの子、ほかの人たちとあんま会話を交わさなかった。	⁷ 君と、その空気読めない、あ、空気読めない×××君、空気読めない君、の悪ぐちを言うために、その実習終わった日には、3人で飲んで、もう、その二人の悪ぐちを言うためだけに、こう、集まって。
比較 11	時間長	約 15000ms	約 15000ms
	モーラ数	104 モーラ	102 モーラ
	話者	HS (友達など、身近な人に話すつもりで)	HS (友達など、身近な人に話すつもりで)
比較 11	時間長	約 15000ms	約 15000ms
	モーラ数	107 モーラ	109 モーラ

⁷ 資料中に人名が現れるものには などの記号を代わりにあてた。なお、記号一文字が1モーラに対応している。以下同様。

	話者	KR (目上の人に話すつもりで)	KR (目上の人に話すつもりで)
比較 12		なんか昔タッキーとかもやってたでしょ。あれも、ま、良かったっちゃ良かったんだけど。だってさ、なんか顔、怪我して、一回あの時。	その、その舞台の、誰だっけ。 × × × 先生っていう先生に、いま教わってるんだけど、その さんの舞台だったのね。
	時間長	約 7500ms	約 7500ms
	モーラ数	58 モーラ	58 モーラ
	話者	KR (友達など、身近な人に話すつもりで)	KR (友達など、身近な人に話すつもりで)

3.2 結果

聴取実験の結果と、基本周波数の情報を併せて以下の表 4.1 に挙げる。聴覚印象の欄の不等号は、大きいものがより早いと認知されたことを意味する。なお、速度の差がそれほど感じられなかったものについては“ ”で、速度の違いを認知できなかったものについては“ = ”で結果を示した。数値で示されているものは全て基本周波数に関する情報で、単位は Hz である。全体的な基本周波数として中央値を、変化の大きさを表す指標として標準偏差をそれぞれ計測した。また、参考として最大値・最小値も挙げておく。

表 4.1 課題 実験結果

	聴取実験の結果		標準偏差	中央値	最大値	最小値
比較 1	A > B	A	49	161	266	50
		B	30	100	199	51
比較 2	A > B	A	61	149	326	143
		B	17	96	66	51
比較 3	B > A	A	39	226	330	174
		B	64	223	393	173
比較 4	B A = C	A	51	207	328	121
		B	54	210	339	123
		C	49	212	321	122
比較 5	A > B	A	81	173	461	104
		B	37	170	265	121
比較 6	C A = B	A	28	113	264	53
		B	25	115	227	85
		C	34	120	236	83

比較 7	A > B	A	52	202	403	128
		B	36	179	365	103
比較 8	A > B	A	58	188	427	138
		B	22	174	322	143
比較 9	A > B	A	114	248	599	104
		B	56	175	386	127
比較 10	A > B	A	89	180	486	117
		B	33	179	325	119
比較 11	A > B	A	98	206	611	128
		B	47	188	370	129
比較 12	A B	A	69	212	461	135
		B	60	189	379	133

3.3 考察

本稿で行なった比較では全て、速く認知されたものの方が基本周波数の標準偏差が高くなっていった。【比較 2】は話者の性別の異なる音声を組み合わせているために、単純に比較することがふさわしいかは疑問であるが、他の組み合わせを含めても基本周波数の標準偏差が最も聴覚上の発話速度に反映されていることがわかる。

また、聴取者である筆者がはっきりと発話速度の差を認知できたものの方が、より基本周波数の標準偏差の差も大きくなっていることから、基本周波数変動が聴覚的な発話速度に影響していることがより明確になった。発話速度の違いがはっきりと聞き分けられた【比較 1】・【比較 2】・【比較 3】・【比較 5】・【比較 7】・【比較 8】・【比較 9】・【比較 10】・【比較 11】では基本周波数の標準偏差に大きな差が見られる。それに比べて、【比較 4】・【比較 6】・【比較 12】は聴覚的にも大きな速度の差は認知されず、基本周波数の標準偏差にも大きな開きは見られなかった。

本研究では基本周波数の標準偏差を、基本周波数変動の指標として用いている。したがって上の結果に鑑みると、ピッチ変化の大きなものほど速く聞こえると考えるのが妥当であろう。内田照久 (2000) では、基本周波数全体を上げたものは発話速度も速く感じ、逆に下げたものは発話速度が遅く感じるという結果が出ている。しかし、【比較 3】【比較 4】においては、全体の基本周波数が (わずかではあるが) 低いものの方が速く聞こえている。したがって、基本周波数の標準偏差の方がより聴覚上の発話速度を表す一つの指標として妥当であると考えられる。また、内田照久 (*ibid.*) では基本周波数を機械的に上げたり下げたりした合成音を用いている。しかし、合成音は聴取時に不自然さを感じさせてしまう。このことが遅速の聴覚印象にも何らかの影響を及ぼした可能性が考えられる。心理学や音響学などの隣接科学においてはしばしば合成音が用いられるが、筆者の知る限りにおいて合成音

の「不自然さ」がもたらす悪影響を考慮に入れているものは存在しない。合成音を用いることで録音や音声編集の手間は大幅に省けるが、最も安全で最良の実験方法は自然言語音を用いることである。自然言語音を用いる実験音声学的方法がより妥当であることが、本稿の実験結果からも明白であると言えよう。

4 課題 : 基本周波数変動の向きに関する聴取課題

次に、基本周波数変動の向きが発話速度の認知にどのように影響するかを検証するために、あらかじめ実験的に統制した音声を用いた課題を行なった。

4.1 方法・手順

4.1.1 分析資料

基本周波数変動の方向が聴覚上の発話速度にどのような影響を及ぼすかを明らかにするため、以下のような資料を録音した⁸。

A : 「高木さんと山田さんが計画した⁹」の各モーラを、[ta]¹⁰に置き換えて話す。

各文節は上昇調で。

B : 「田辺さんと三上さんが迷惑した¹¹」の各モーラを、[ta]に置き換えて話す。

各文節は下降調で。

4.1.2 被験者

被験者は以下の 1 名である¹²。

KR : 24 歳女性。言語形成期(4、5 ~ 12、3 歳)は埼玉県所沢市で過ごした。

4.1.3 解析装置

課題 と同様。3.1.3 参照。

4.1.4 解析方法

⁸ 録音機材などは課題 と同様。3.1.1.2 参照。

⁹ アクセントの影響を配慮して、A と B の課題文を異なるものにした。つまり、A はピッチが繰り返し上昇する形にするために平板型の単語を使用し、B は繰り返し下降する形にするために頭高型の単語を使用した。

¹⁰ 林実・寛一彦(1989)では、子音の種類によって反応時間が異なることを示している。その序列は、反応時間の速い順から、[p],[t],[k],[h],[b],[d],[w],[g],[j],[m],[z],[r],[n],[s]となっている。ここでは、子音の中でも比較的脳が鋭敏に反応する [t] と、最も安定した母音 /a/(城生佰太郎 1997:253) を用いた。また、非常にこの課題の難易度が高く、自然談話と同じテンポで話すことが非常に困難であることから、ほかの脳が鋭敏に反応する子音 ([p],[k]) なども試した上で、被験者が最も発音しやすい音を選んだ。

¹¹ 注 7 と同様、B はピッチが次第に下降する形にするために、頭高型の単語を使用した。

¹² この課題の発話には高い技術が要求されたため、非常に活舌が良く時間長の制御にも長けていた KR 氏のみを被験者とした。それでもこの課題の録音には相当な時間と労力を要した。このような無理難題に根気良くご協力下さった KR 氏には、感謝の言葉が尽きない。

聞き取りによって、2種の資料の聴覚印象上の遅速差を判断する。この課題では、筆者のほかにもう一名の日本語母語話者にも聴取を行なわせた。その聴取者の情報は以下の通りである。

MK：22歳女性。言語形成期は東京都東久留米市で過ごした。

その他音響解析等は課題と同様の方法を採用している。3.1.4を参照されたい。

4.2 結果

まずは、A・B両資料の音声波形と基本周波数曲線を下の図4.1・4.2に挙げる。なお、基本周波数の表示レンジは0～700Hzである。これらの基本周波数曲線を見ると、Aは概ね繰り返し上昇、Bも繰り返し下降しているのが見て取れる。

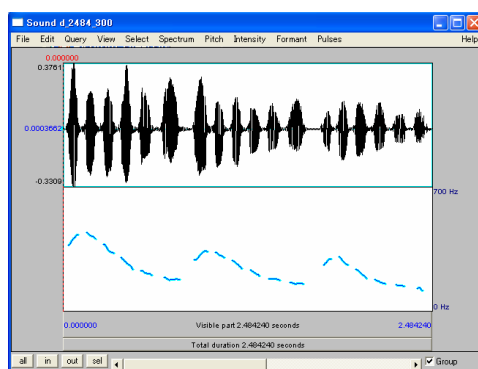
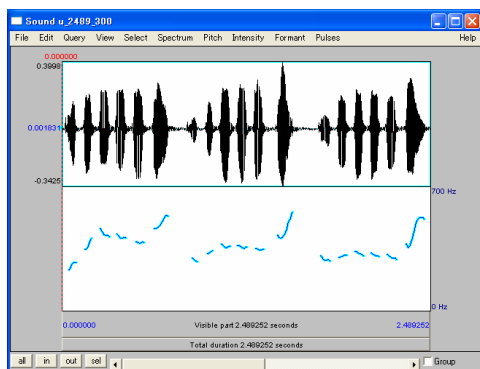


図 4.1 資料Aの音声波形・基本周波数曲線 図 4.2 資料Bの音声波形・基本周波数曲線

次に、下の表4.2に聴取実験の結果と音響解析の結果を示す。聴取実験の項目の不等号は大きい方が発話速度の早く感知されたことを示すものである。音響データの数字の単位は全て Hz である。

表 4.2 課題の実験結果

聴取実験の結果			標準偏差	中央値	最大値	最小値
筆者	B > A	A	76	362	560	233
MK	A > B	B	85	234	446	117

上の表からもわかる通り、筆者とMK氏とは遅速の判断が逆になっている。

4.3 考察

課題 では、基本周波数の変化の向きによって発話速度の認知が変わるのかについての実験として、ピッチが繰り返し上昇するもの (A) と下降するもの (B) を比較した。その結果聴取者である筆者は B の方が速く聴こえたが、被験者 MK 氏の場合は A の方が速いという結果になった。このことから考えると、どちらの方向にピッチが変化するとより速く聞こえるかということについては、個体差があるということが示唆された。ただ、筆者と MK 氏は「この 2 つの資料から発話速度の差を認知した」という点では共通している。「どちらが速いと認知されたか」ということ以上に、「両者の間に速度の違いが聞き取れた」ということが重要ではないだろうか。すなわち今回の課題 の結果は、基本周波数の変化の向きが発話速度の認知に影響を及ぼすことを示唆するものである。ただしその影響がどのように現れるか (すなわち、基本周波数の変化がどちらを向くのかによって、遅く聞こえるのか、それとも速く聞こえるのか) には、個人差があるものと考えられる。

5 まとめ

本稿では 2 種類の課題によって、発話速度の認知が基本周波数変動にどのように影響を受けるのかを検証した。

課題 では様々な音声データを用いて発話速度の聴覚印象と基本周波数の統計データを比較した。その結果、基本周波数の変動が大きいものの方がそうでないものよりも相対的に速く聞こえることが明らかになった。城生百太郎 (2001:429-432) では、物理量として等量の周波数変動が生じてもその変化が遂行された時間長によって知覚レベルでは異なってくるとして、周波数変動に時間長のパラメータを取り込んだ「傾き」という概念を提示している。本稿の実験においても、基本周波数の「傾き」が大きいものほど発話速度が速く認知された。

また、課題 では実験的に統制された自然言語音の資料を用いて、基本周波数の変化の方向がどのように発話速度の認知に影響を及ぼすのかに迫った。その結果、変化の方向の違いが発話速度の認知にも影響を及ぼすことが検証された。

6 展望

本稿では発話速度の認知が受ける影響を基本周波数との関連のみで述べたが、筆者はこれ以外に発話速度の聴覚印象に変化をもたらす要因が存在しないとは考えていない。本稿では規定枚数の都合上触れることが出来なかったが、丸島歩 (2007) ではポーズのとり方など他のプロソディー情報、文体や方言差・発話内容などの言語レベルの問題が遅速の判断に影響を与えることが明らかになっている。

また丸島歩 (*ibid.*) では音響・聴取実験だけではなく、より客観的な裏打ちを得るために基本周波数変動と発話速度認知の関わりを検証するための脳波実験も行なっている。そこでの成果は丸島歩 (近刊予定) に譲るが、発話速度の認知に関してより深い洞察を得るには

基本周波数情報だけでなく、様々なプロソディー情報や言語情報との相関をより詳細に探っていく必要があるだろう。プロソディーそのものの全様の解明とともに、今後の課題としたい。

参考文献

- 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
- 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.
- 城生佰太郎 (1997) 『実験音声学研究』, 勉誠社.
- 城生佰太郎 (1999) 「現代日本語の自然音声談話のスピード」『言語』 28- 9: 44-50
- 城生佰太郎 (2001) 『アルタイ語対照研究 - なぞなぞに見られる韻律節の構造 - 』, 勉誠出版.
- 城生佰太郎 (2005) 『日本音声学研究』, 勉誠出版.
- 林実・箕一彦 (1989) 「音素・音節検出実験に基づく音声知覚の基本単位の検討」『日本音響学会講演論文集 - I - 』 355-356.
- 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 - 」筑波大学人文社会科学研究所修士論文.
- 丸島歩 (近刊予定) 「事象関連電位を用いた発話速度の認知実験」『言語学論叢』 特別号 城生佰太郎教授退職記念論文集.
- (丸島歩 筑波大学大学院生 ayumi_marushima@yahoo.co.jp)

Speech Rate Recognition: The Effect of Fundamental Frequency Movement

MARUSHIMA Ayumi

The purpose of this paper is to inspect correlation of speech rate recognition with movement of fundamental frequency (F_0).

To achieve the purpose, hearing various sets of natural speech sounds involving duration and morae approximate to one another / each other, I judged the sounds which are heard faster. And I measured the medians and the standard deviations of their fundamental frequency. The sounds are got from movies on DVDs or interview on CDs, or recorded some Japanese native speakers' natural speaking or reading.

The results were as follow:

- (1)The speech sounds with greater standard deviations of F_0 can be heard faster than the others. Therefore I conceive speech sounds involving greater movement of fundamental frequency to be faster cognitively.
- (2)The speech sound with rising pitch and with falling can be distinguished on cognition of rate. However, there seems to be individual difference which can be heard faster, involving rising pitch or falling.