事象関連電位を用いた発話速度の認知実験*

丸島　歩†

キーワード： 実験音声学、脳波実験、ERP、発話速度、ピッチ変動

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数値的に表現される。しかし、このような基準が果たして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のではないだろうか。

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。つまり発話速度の実態に迫るためには、その認知の様態に焦点を当てる必要があるだろう。しかし、発話速度を聴取側の視点から検証

---

*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、本稿では紙幅の都合上脳波実験で得られた知見をまとめることとした。音響・聴取実験での成果については、丸島歩 (近刊予定) に譲る。

†筑波大学大学院一貫制博士課程

[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、それと同様か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。

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ていない。

2 目的

本稿の目的は、従来モーラ数を基準として記述されてきた発話速度について、ERP を用いた実験音声学的手法を用いてその認知の様相の一端に迫ることである。

3 方法

3.1 実験機材

本実験の取り込みには、筑波大学文社会系棟 B613 音声実験室内に設置されている機材を用いた。以下の図 3-1 に機材のおおまかな配置図を示す。

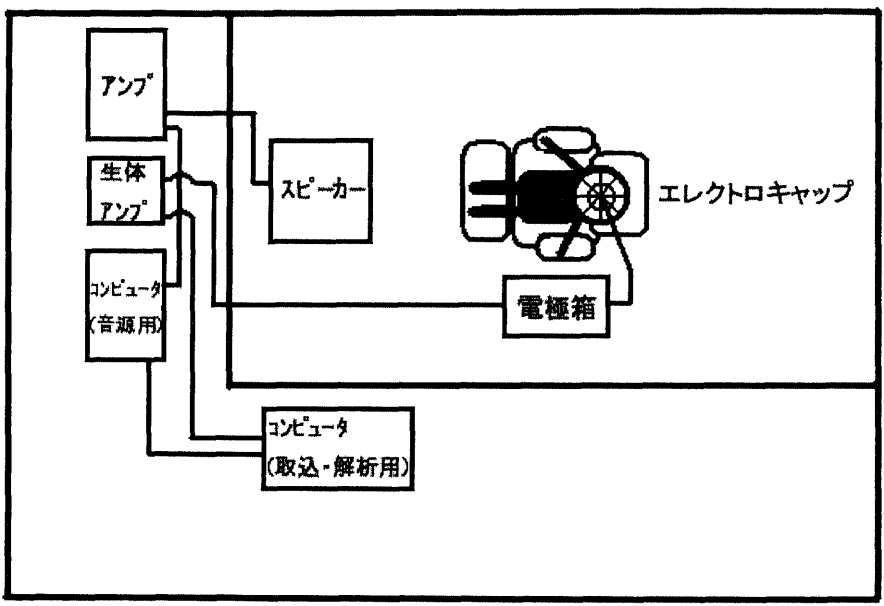

図 3-1：本実験における機材および被験者の配置

3.1.1 収録機材

被験者に装着したエレクトロ・キャップ[2]から導入した脳波を、電極箱を介して生体アンプで増幅し、コンピュータで記録した。使用した機器の詳細は以下の通り。

エレクトロ・キャップ： ECI-2 (Electro-Cap International 社製)。電極配置は国際 10-20 法[3]に基づいて、F3・F4・C3・C4・P3・P4・O1・O2・F7・F8・T5・T6・Fz・Cz の 14 チャンネルを選択した (図 3-2 参照)。これを被験者に被せた上で、同社製の electro-gel を電極と頭皮の間に注入した。

基準電極：耳朶、同側耳朶方。(Electro-Cap International 社製) を使用。

電極箱：NEC ELECTRODE BOX／TYPE 6R12-2。

生体アンプ：NEC BIOTOP 6R12-2。低域遮断フィルタ 0.5Hz、高域遮断フィルタ 60Hz、感度 50μV/fs に設定した。

収録・記録用コンピュータ：PC98xv20 (OS は NEC98 対応の MS-DOS6.1)。用いたソフトは EPLYZER2.1 (キッセイコムテック社)。標本化 500Hz、プレトリガ-100msec、取り込み時間-100～1948msec、加算回数は 35 回に設定した。

---

[2]電極を頭皮上の正しい位置に固定できるよう、あらかじめ定位置に電極が取り付けられた伸縮性のある帽子。

[3]頭の前後径、左右径、周径をそれぞれ 10 等分した上で、さらに電極間の距離をそれぞれの 10%か 20%となるようにする電極の配置方法のこと(城生佰太郎 1997:248)。

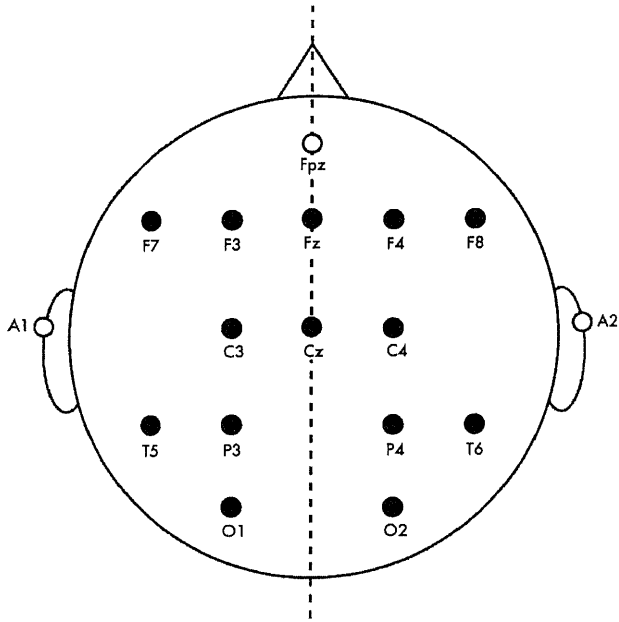

図 3-2：電極配置図

3.1.2 刺激発生装置

コンピュータから再生させた刺激音を、ライン・アウト端子を介してアンプ類で増幅、シールド・ルーム内のスピーカから発生させ、被験者にフリーフィールド[4]の状態で聴取させた。また、同コンピュータから音声発出と同時にトリガを発し、収録用コンピュータに送ることで刺激音発出と脳波収録のタイミングを同期させた。使用した機器の詳細は以下の通り。

刺激発生用コンピュータ：IBM Vision PS/V Model 2408 (OS は Windows 3.1)。用いたソフトは Winstim (榊氏作製、私家版ソフト)。単一の刺激音が 3000msec の提示間隔で繰り返し発出されるように設定した。

アンプ類： Technics Stereo Cassette Desk ／RS-678U、Technics Stereo Flat Preamplifier／SU-9070、Technics Stereo Universal Frequency Equalizer／SH-9010E、Technics Peak/Average Meter Unit／SH-9020M、Technics Stereo

[4]ヘッドフォンを用いず、スピーカから流れる音声を両耳で聴取した、ということ。

Power Amplifier／SE-9060 (全て松下電器産業社製)。被験者の耳元の位置で再生音圧が 65dBSL[5]になるように設定した (音圧計はリオン社製・型式 NL-14 を使用)。

スピーカ：Technics Linear Phase Speaker System (松下電器産業社製)。

3.2 被験者

以下の 2 名にご協力いただいた。

KM、女性、22歳、右利き。言語形成地は埼玉県草加市。

NN、女性、23歳、右利き。言語形成地は東京都江東区。

3.3 刺激音

刺激音は、自然言語音の [papapapapa][6]を三種類の速度 (fast・normal・slow) で読み上げたものと、加工した半合成音を用いた。半合成音は二種類で、normal のものを音響編集ソフト Cool Edit 2000 (Syntrillium 社製) を用いてピッチを変更した。20%ピッチを上げたところから次第に下がっていくもの (down) と、35%ピッチを下げたところから次第に上がっていくもの (up) である。したがって、刺激音は全部で 5 種類となる。

調音を行なったのは KR 氏、女性、24 歳。言語形成期は埼玉県所沢市で過ごしている。

刺激音の時間長は表 3-1 の通り (単位は msec。down・up の時間長は normal と同じ)。

表 3-1：刺激音の時間長

fast : 599	normal : 723	slow : 968

[5]デシベル表示には SPL (sound pressure level) と SL (sensation level) の 2 種類のレベルが設けられている。成人男性の聞き取ることのできる最も小さい音と言われている 0.00002Pa を 0Hz として、物理的に規定した SPL に対し、SL はヒトの感覚を基準としたレベルで、被験者の最小可聴域をもとにしている (城生佰太郎 2005:458)、(中村健太郎 2001:34)。

[6]最も脳が鋭敏に反応する子音であることがわかっている /p/(林実・篭一彦 1989) と、最も安定した母音 /a/(城生佰太郎 1997) を用いた。また、速度を知覚させるためにはできるだけ音節数を増やした方がよいと考えられるが、解析時に最大で約 1000ms を目安に観察することから、slow の刺激音が 1000ms 以下になるように 5 音節とした。

3.4 実験手順

まずは音声実験室に設置されたシールド・ルーム内にある安楽椅子に被験者を座らせ、エレクトロ・キャップの装着などを行なう。電極の状態を確認後、被験者に「今から流れてくる音声を頭の中で繰り返してください」と指示し、実験中は目を半眼にしてリラックスした状態を保つように伝えた。

被験者それぞれに、刺激音一種類ごとに一回、合計5回の試行を行なった。各試行では同一の刺激を3000msecの間隔で35回提示[7]し、35回分の脳波を収録した。

3.5 解析装置

解析に用いたコンピュータは、PC98-xv20・PC9821-Nr166(両機ともOSはNEC98対応のMS-DOS6.1)。解析ソフトはATAMAP(キッセイコムテック社)を用いた。

3.6 解析方法

被験者の瞬目などによるアーチファクト[8]を除去するため、取り込みに使用したソフト(EPLYZER)を用いてRAWデータ再加算[9]を行なった。

次に、解析ソフトATAMAPの波形とトポグラフィー[10]を摺り合わせ、ソフト搭載のマーキング機能で陰性波・陽性波[11]で最も色濃くトポグラフィーが反応したところでカーソルを立てた。

3.7 データ処理方法

ピーク潜時(PL[12])・ピーク間潜時(IPL[13])をもとに情報処理を行った。

---

[7]各試行の冒頭は脳波の状態が安定しないため、収録コンピュータのモニターで脳波の安定が確認されてから収録を開始した。従って、実際に被験者に刺激音が提示される回数は40回弱である。

[8]収録中、特に電極に近い部分を動かすと、脳波に重畳してしまう筋電によるノイズのこと。

[9]収録した脳波の中からアーチファクトなどの見られないものだけを選び、再加算すること。

[10]脳のどの箇所がどの程度反応しているかを視覚的に見ることの出来る色分け分布図。陰性の電位を帯びた部分は赤く、陽性の電位を帯びた部分は青く表示される。電圧が強ければ強いほど色濃く表れる。

[11]陰性波は上向き、陽性波は下向きに表示される。

4 結果

4.1 実験 1 回目 (KM)

target wave[14]を見定めるために、一回目の実験では約 1000ms までの波形を解析・観察した。ここでは速さを変えた自然言語音のみを用いた。被験者は KM である。

以下の表 4-1 に PL、表 4-2 に IPL を示す。空欄になっているものは、陽性波もしくは陰性波がほとんどあらわれなかったものである。アスタリスク (*) つきのものは、電圧が低いことを表す (以下同様) 。

表 4-1：実験 1 回目 PL (KM)

[KM]	N1[15]	P2	N2a[16]	N2b	P3	N3	P4	N4a	N4b	P5
FAST	122	186	272	370	446			**670**	758	824
NORMAL	122	220	284	440	482	538	608	**664**		
SLOW	110	180	282		450	502	554	**630**	700	828

表 4-2：実験 1 回目 IPL (KM)

[KM]	P2-N1	N2a-P2	P3-N2a	N3-P3	P4-N3	N4a-P4	P5-N4a
FAST	64	86	174				154
NORMAL	98	64	198	56	70	56	
SLOW	70	102	168	52	52	76	198

[12]Peak Latency の略。

[13]各ピーク潜時間の時間長のこと。Inter Peak Latency の略。

[14]まずノイズを含んだ波形を除去して加算することで特徴的な波形を見出し、もう一度ある項目に特化した波形が明瞭に表れているものを加算し直すこと。

[15]陰性波を N、陽性波を P で表す。後続する数字は潜時の早い順に番号を振ったもの。このような表記の仕方は医学の分野で用いられる方法であり、心理学の分野においては N (もしくは P) のあとにおおよその潜時をミリ秒単位で示す。筑波大学大学院人文社会学研究科文芸・言語専攻一般言語学研究室においては、基本的には医学系の名称を用い、心理学で発見された P300、N400 などの著名で特徴的な波形についてはその名称を用いている。

[16]二峰性の波形 (頂点が二股になっている波形。一度陽性 (もしくは陰性) に振れてから、完全に陰性 (陽性) に振れずに再び陽性 (陰性) に振れる波形のこと。) は潜時の速いピークを a、遅いピークを b とする。

4.2 実験 2 回目 (NN)

次に、ピッチを操作した半合成音も加えた 5 種類の刺激を用いた実験を実施した。一回目の再現性を確認することもここでの目的としている。被験者は NN である。

表 4-3 には PL を、表 4-4 には IPL をそれぞれ示した。

表 4-3：実験 2 回目 PL (NN)

[NN]	N1	P2	N2a	N2b	P3	N4	P4	N5a	N5b	P5
FAST	102	204	270	382	476	*494	576	**704**	784	824
NORMAL	80	212	292		334	392	608	**660**	730	780
SLOW	132	204	264	330	338	406	472	**536**	712	774
135%UP	128	206	282	390	458	528	608	**642**	714	778
80%DOWN	156	204	288	404	466	510	550	**602**	756	818

表 4-4：実験 2 回目 IPL (NN)

[NN]	P2-N1	N2a-P2	P3-N2a	N4-P3	P4-N4	N5a-P4	P5-N5a
FAST	102	66	210	(18)	(82)	128	120
NORMAL	132	80	224	58	216	52	120
SLOW	72	60	208	68	66	64	238
135%UP	78	76	176	70	80	34	136
80%DOWN	98	84	178	44	40	52	216

4.3 実験 3 回目 (KM)

二回の実験を終え、target wave が N4a にあることがおぼろげながら見えてきた。ここで、さらに再現性を保障するための実験を行なった。二人の被験者のうち瞬目などによるアーチファクトが少なく、脳波そのものも安定していた被験者 KM の脳波再度収録した。

また target wave の考えから、より特徴的な部分に注目できるよう、N3 以降にしぼって解析を行なった。

表 4-5 には PL を、表 4-6 には IPL をそれぞれ示した。

表 4-5：実験 3 回目 PL (KM)

[KM]	N3	P4	N4a	N4b	P5
FAST	620	*674	**744**	802	874
NORMAL	556	616	**692**	814	860
SLOW	440	616	**680**	740	776
135%UP	528	596	**672**	734	810
80%DOWN	534	588	**652**	724	810

表 4-6：実験 3 回目 IPL (KM)

[KM]	P4-N4	N5a-P4	P5-N5a
FAST	54	70	130
NORMAL	60	76	168
SLOW	176	64	96
135%UP	68	76	138
80%DOWN	54	64	158

4.4 実験 4 回目 (KM)

最後にもう一度、被験者 KM の脳波を収録した。

表 4-7：実験 4 回目 PL (KM)

[KM]	N3	P4a	P4b	N4a	N4b	P5
FAST		526	636	**704**	804	852
NORMAL	506	620		**694**	820	900
SLOW	468	574		**638**		746
135%UP	442	546		**680**		760
80%DOWN	368	464		**664**	738	776

表 4-8：実験 4 回目 IPL (KM)

[KM]	N4-P3	P4-N4	N5-P4	P5-N5
FAST			178	148
NORMAL	92	114	74	206
SLOW	136	106	64	108
135%UP	88	104	134	80
80%DOWN	40	96	200	112

5 考察

本研究では 4 度にわたる脳波実験を行なったが、特に注目しておきたいのは N4a の潜時である。以下の図 5-1 は各実験の各刺激音の N4a 潜時についてグラフ化したものである。

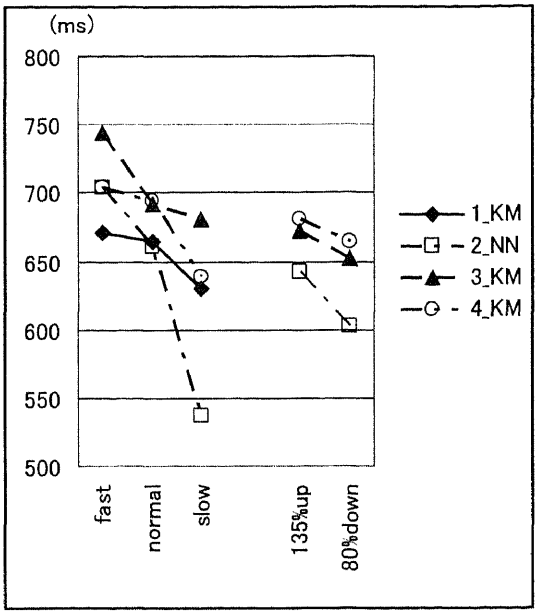

図 5-1：各実験の N4a 潜時

これによれば刺激音 fast・normal・slow について、刺激音の速度が速くなるほど N4a の潜時が遅くなっているのが見て取れる。すなわち 4 回行なわれた実験全てにおいて N4a の潜時は、刺激音 fast が最も遅く、それに次いで刺激音 normal、最も速いのが刺激音 slow となっている。つまり、刺激音の発話速度が速くなれば速くなるほど、N4a の潜時が遅くなっている。したがって、発話速度の認知の反映として現われるのは N4a であるという蓋然性が大きい。さらに、個体差を超えて 2 名の被験者に同様の特徴が現われた上、被験者 KM については 3 度収録を行なっての結果であることから、再現性が非常に高いことが認められるだろう。

さらに、2・3 回目の実験で用いた刺激音 135%up・80%down を N4a の潜時について比較した場合、135%up が 80%down よりも潜時が遅くなっている。この結果は 3 回の実験で 2 人の被験者について同様の結果となった。この刺激についても個体差を跨いで共通した結果が得られたことから、かなりの再現性が認められる。また、これらの N4a の潜時が刺激音 normal

で得られた脳波のN4aよりも早く、slowの潜時寄りになっていることと、前述したピッチ変動の激しいものが速く聴こえるという音響実験の結果とは一見矛盾するようにも思われる。しかし、一般的におよそ600msを超えて現われる陰・陽性波は複数の要因が重畳することによって現われると言われていることから、この結果が音響実験の結果と矛盾していると断定することはできない。ピッチ変動のあまりない自然言語音であるfast・normal・slowの解析結果と、ピッチ変動を機械的に施した半合成音である135%up・80%downを単純に比較するのは短絡的である。むしろこの結果から、ピッチ変動と発話速度は認知の上においてはかなり深い繋がりがあるであろうということが想定できる。

なお、ピッチが上昇している135%upのN4a潜時の方が、ピッチが下降している80%downのN4a潜時よりも遅くなっていることから、この2名の被験者はピッチが上昇する方が下降する方に比べて、発話速度を速く知覚しているということが想定できる。

また、本研究での脳波実験では、城生佰太郎 (1999) に挙げられたようなPzのP300の極性の逆転のような現象は起きなかったが、刺激音や実験パラダイムの違いからこのような結果が現われたと考えられる。とは言え、本研究での実験パラダイムや刺激音が特殊で偏りのあるものとは言えない。なぜなら、発話速度の知覚にはほかのさまざまな要因が影響しているからである。このことは、丸島 (2007) において明らかになっている。

6 展望

本研究における被験者は2名であった。また、ピッチ変動を与えた刺激音は半合成音を用いた。今回発見されたN4aの再現性を確認するとともに、ピッチ変動を与えた刺激音に自然言語音を用いた場合でも同様の結果が得られるのかを検証する必要があると考える。

また前述した通り、丸島 (2007) では聴取実験により発話速度の知覚にはさまざまな要因が関わっていることが明らかになっている。これらと発話速度の認知の相関についても新たな実験パラダイムを構築し、脳波実験

によって確認することで、発話速度の認知の様相をより詳細に探索していきたいと考えている。

【参考文献】

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』、平成 9 年度科学研究費補助金による助成出版、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 (2005) 『日本音声学研究』、平成 16 年度科学研究費補助金による助成出版、勉誠出版

中村健太郎 (2001) 『図解雑学 音のしくみ』ナツメ社

林実・筧一彦 (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」『第 5 回日本生体磁気学会論文集』 3-1: 135-136. 日本生体磁気学会

丸島歩 (2007)「発話速度の実験音声学的研究ー聴取側の視点からー」筑波大学人文社会科学研究科修士論文

丸島歩 (近刊予定) 「発話速度の認知に関する一考察 基本周波数との関連性に注目して」『言語学論叢』オンライン版創刊号 (通巻 27 号)