

# 音と意味の対応関係に関する事象関連電位を用いた実験研究\*

中村絵里香      福盛貴弘

## 1 序

音声学は、発出・伝播・受容に関わる人間の言語音のしくみを扱う学問である。発出については、調音音声学あるいは舌運動や呼気流量などを計測する生理音声学が、伝播については音響音声学がめざましい進展をみせている。その一方で受容にあたる聴覚音声学については、自然言語音・合成音声を聴取させて判断させる知覚実験が行われているものの、インフォーマントが判断した結果に関する研究であり、その結果に至る過程を実証的に扱った分野については、全くといっていいほど進展がみられない。筆者らは、この点を非常に不満と感じており、この分野の開拓こそが、音声学の3分野を相互補完させるために必要であると確信した<sup>1</sup>。そこで、我々はERP(事象関連電位)という電気生理学的手法を方法論として採択し、言語音をどのように聴覚情報系から脳内処理しているか、すなわち受容における過程の一端を垣間見ることを考えた。本稿では、言語音をどのように音声知覚処理しているかを、ERPによる波形データから判断したいと考える。音声は、ともすれば、トップダウン処理のみで、換言

---

\*事象関連電位という方法論を採択するにあたって、城生佰太郎先生のご指導・恩恵を賜った。言語学・音声学にたずさわる者として、脳を扱うなどということを思いもしない頃があったが、非侵襲的方法、すなわち脳を切開によって露呈しなくてよい方法から、音声処理の過程を扱う可能性を見出せたことは、まさに一筋の光明であった。この場をかりて、心から感謝の意を申し上げる次第である。本稿に最も有益な示唆を与えて下さった論者は、城生佰太郎(1997)である。ERPを用いた実験方法論の詳細については、こちらを参照していただきたい。また、実験手順に関していろいろ助言をいただいた島田武氏にも改めてお礼を申し上げる次第である。なお、本稿の執筆分担は1～3が福盛、4～7が中村となっている。

<sup>1</sup>城生佰太郎(1997:245)参照。

すると周囲の状況や自身の知識などからあらかじめ予測して音声を知覚していると考えられがちである。しかし、人間は確かにトップダウン処理だけでない、いわば「ボトムアップ」的な音声知覚を行っている。これを簡単に確認する実験として、Pinker(1994)で次のことが提案されていた。訳書である椋田直子訳(1995:253)から、以下に引用する。

読者のなかで、気のいい友人を持っている人には、つぎの実験をお勧めする。辞書を開き、適当に単語を 10 語選んで、その友人に電話をかけ、10 語を明瞭に発音する。友人は、音波のデータと英語の語彙と音韻の知識しか手がかりがないにもかかわらず、完璧に 10 語を繰り返せるだろう。

電話であれば、文脈情報や相手の表情といった周辺情報から予測するということは困難である<sup>2</sup>。筆者はこれを日本語で確認してみた。電話を通じて行った実験概要を以下略述して説明する。実験は 1999 年 1 月に、福盛の自宅から電話をかける形で行った。調査票の語彙項目は、ランダムな条件を設定するべく、手元にあった国語辞典<sup>3</sup>の頁を適当にめくり、目についた語をあげた。あげた語項目は「命運・住民税・段だらぞめ・四次元・念入り・科挙・美しい・アクアラング・ぶらさがる・チョンガー・心ひそかに・場・かぶる・切磋琢磨・老女」の 15 例である。発音は福盛貴弘(男・1970 年生・大阪(=言語形成期(5,6 ~ 12,3 歳の時、過ごした場所))が行った。気のいい友人として、ご協力いただいたインフォーマントは 10 名<sup>4</sup>である。実験の提示は、「今から読みあげる語を繰り返して行って下さい。ただし、アクセントはまねしなくても結構です。」というようにした。結果は、完璧にとはいかなかったものの、全員がほぼ完璧に繰り返した。ききとりにくかった語項目に「場」があげられる。「場」については、6 人が 1 度目に「あ」ときき、課題終了後ききなおすと、「場」とききとることが

<sup>2</sup> 音声情報の中にも、言外の情報は数多く含まれている。例えば、相手が誰であるかという属性を特定する要素、相手に話す順番を交替させることを示す要素などである。Laver(1994:14-15) 参照。ここでの周辺情報というのは、電話という前提では音声から判断する可能性が最も高くなるため、音声以外の周辺情報を用いにくいという範囲で使っている。

<sup>3</sup> 旺文社編(1979)『標準国語辞典新版』旺文社

<sup>4</sup> 岡田正氏(男・1972 年生・島根)、上高真琴氏(女・1972 年生・広島)、桑原雅子氏(女・1973 年生・山口)、島田武氏(男・1970 年生・徳島)、菅井康祐氏(男・1973 年生・兵庫)、菅生美恵子氏(女・1973 年生・栃木)、辻伸之氏(男・1973 年生・岡山)、中村絵里香氏(女・1975 年生・愛知)、増田幸子氏(女・1977 年生・千葉)、三松国宏氏(男・1966 年生・東京)(アイウエオ順)この場をかりて、お礼を申し上げる。

できた<sup>5</sup>。また、調査項目の中には、インフォーマントによって、「段だらぞめ・アクアラング・チョンガー」など意味が分からない語も含まれていたが、意味が分からなくてもこちらから与えた音声情報を繰り返すことはできるのである。しかし、この実験は日本語の語彙知識 (例えば「染め」の部分) から予測して判断した、換言するとトップダウン処理が加わったという可能性をぬぐい去ることはできない。そこで、筆者らは、言語音と非言語音、意味を考える・考えないというパラメータを設定した音声知覚に関する基礎実験を行ってみたいというところにたどりついた次第である。

## 2 目的

本実験は、人間の脳内における情報処理の過程を受容・認知の観点から、ERP を用いた実験によって確認すること、その中でも特に音声を音声情報として処理する場合と、意味ある語として処理する場合との間に差異があるのかを検証することが目的である。しかし、ERP を用いた実験で、現在我々が行っている全ての処理をみるというのは不可能である。従って、情報処理の一端をみるものとして、言語音・非言語音を聞かせ、

1. 言語音と非言語音との間に、音声知覚の区別があるのか。
2. 言語音・非言語音を基に、意味を考えさせた場合、両者の違いが見られるか。

を検証することを目的とする。

## 3 方法

### 3.1 被験者

本実験に快く協力してくれたのは、以下の 4 名と筆者 2 名の計 6 名である。

---

<sup>5</sup>この点にいたっては、本実験の範囲から一拍語がききとりにくい、筆者の発音が明瞭でなかったなどの要因が考えられるが、特定の要因は不明である。

	氏名	言語形成期	性別	生誕年
(1)	上高真琴氏	広島県	女性	1972 年生
(2)	関 裕子氏	新潟県	女性	1973 年生
(3)	長嶋宏和氏	愛知県	男性	1976 年生
(4)	林 和宏氏	愛知県	男性	1975 年生
(5)	福盛貴弘	大阪府	男性	1970 年生
(6)	中村絵里香	愛知県	女性	1975 年生

全員右利きである。言語形成期の項目は、5,6 ~ 12,3 歳までを過ぎたところを示す。今回の実験では、特に方言圏の指定はせず、日本語母語話者という条件だけを整えた。なお、筆者らは主に本実験の方法を決定するために、被験者となっただけなので、結果からは除外している<sup>6</sup>。

### 3.2 分析資料 (刺激音)

言語音については、肉声による自然言語音<sup>7</sup>を、シェアソフト Cool-Edit を用いて、サンプリングレート 44.1kHz、モノラル、16bit で録音・編集した。/papa/・/sasa/を平板アクセントで発音したものを刺激音として用いている。非言語音については、シェアソフト Cool-Edit で、1kHz 純音・ホワイトノイズ・ブラウンノイズを編集した音を用いた。純音は/papa/と、ノイズ 2 種は/sasa/とおよそ同じ音節時間長の配分にしてある。

### 3.3 手続き

被験者には、それぞれの課題ごとに内容を説明した上で、シールドルーム内に座らせ、スピーカーから流れてくる刺激音を、フリーフィールドで聴取してもらった。刺激音の呈示感覚は 2 秒で、再生条件は 65dBSL に統一した。まず、課題 1 として、刺激音を/papa/・純音・/sasa/・ブラウンノイズ・ホワイトノイズの順で、各刺激音における加算平均回数が 30 回になるまで聞かせた。その際、普通に流れてくる音を聞いてもらうように指示した。次に 5 分程度の休憩・

<sup>6</sup>実験の提示を事前を知ってしまったため、他の被験者との等条件を維持することが困難であるという理由による。

<sup>7</sup>発音は福盛が行った。

閑談をはさみ、課題として、「刺激音は同じ順番で出てくるので、今度は何か意味のある語を考えて、頭の中でくり返して聞いてください」という指示を与えた。なお、意味のある語を考えてもらう方の取り込みにあたっては、5 回程考える時間を与え、それ以降は同じ語を想定して聞いてもらっているため、はじめの 5 回はデータから除外した。

### 3.4 実験・解析装置

取り込み装置は、筑波大学人文社会学系棟 B613 音声実験室に設置された、NEC 社製 BIO TOP 6R12 型生体アンプを、NEC 社製 PC9821 Xv20 型コンピューターに接続した装置を用いた。(配置は図 1 を参照。) 取り込み用ソフトは、キッセイコムテック社製誘発電位研究用ソフト EPLYZER、同コンピューターの DOS 上で動かした。取り込みの際、EPLYZER 上で、サンプリングレートは 500Hz、トリガーは 100msec.、取り込み時間は 100 ~ 2000msec. に設定した。解析は、キッセイコムテック社製 EEG マッピング研究用プログラム ATAMAP ver.3.0 を使用し、同コンピューターの DOS 上で作動させた。電極の配置は、国際 10-20 法に基づき、図 2 に示した 14 チャンネルのモデルを採用した。電極の装着は、Electro-Cap International 社製エレクトロキャップ ECI-2 を被験者の頭部にかぶせておこなった。音源装置は、IBM 社製 PS/V Model2408 型コンピューターで、Windows 上で作動する Winstim というソフトを用いた。コンピューターのヘッドホンジャックから、Technics 社製プリアンプ Stereo Flat Preamplifier Technics70A、同社製アンプ Stereo Power Amplifier Technics60A を介して、同社製 2-Way Speaker System SB-6000 に接続し、被験者に刺激音を呈示した。また、Winstim からトリガーを読み込むために、プリンターポートから生体アンプを介し、PC9821 機に接続した。

### 3.5 解析方法

解析は 170 ~ 260msec. 近傍にみられる大きな振幅を伴った陽性波 (P2 成分波) を中心におこなった<sup>8</sup>。波形の解析は、実験装置の項記載の ATAMAP によ

---

<sup>8</sup>当初、城生 (*ibid.*) に従い、一般的な思考、判断等の高次機能を反映するとされる P3 成分波のあらわれやすい 2 音節語を刺激音とすることで、P3 を扱う予定だった。しかし、計測を進めていく中で、被験者によっては P3 のたちあがりの低い結果がでた。その結果、P2 を中心に解析を

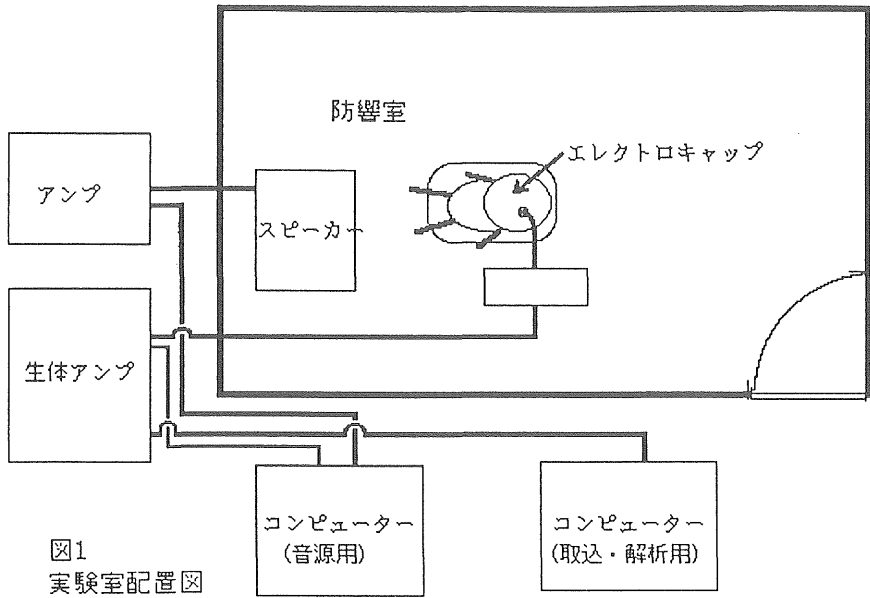


図1  
実験室配置図

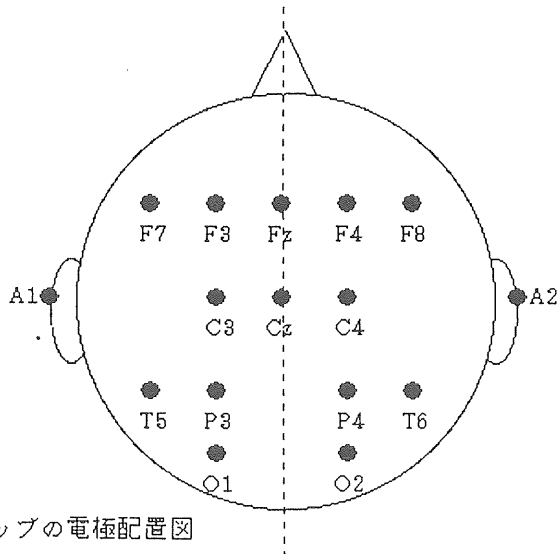


図2  
エレクトロキャップの電極配置図

すすめることとした。P2 は基本的には聴覚皮質の反応だが、高次機能の認知と関わるともいわれている。本実験では、意味を考えさせるというパラメーターを設定したため、高次機能との対応を考えることができる P2 を中心に扱っても支障はないと考えた。なお、神経生理学や心理学で主に扱われている P300 は、本実験方法ではあらわれにくいので、本稿では取り扱っていない。

る脳電位トポグラフィーをもとにして、刺激開始時から陽性波の示す頂点（ピーク）までの潜伏時間（ピーク潜時）とその電圧（ピーク電位）を計測した。

## 4 結果

結果を以下に示す。

表 1:全体での結果 (電圧と P2 の潜時)

KAMITAKA				NAGASIMA			
刺激音	潜時	最大電圧	電極	刺激音	潜時	最大電圧	電極
beep01	216	-5.3	Cz	beep01	222	-6.4	F4
beep02	208	-6.5	Cz	beep02	230	-6	F4
brown01	206	-6.2	Cz	brown01	228	-5.9	Cz
brown02	248	-6.9	F4	brown02	246	-5.3	F4
papa01	262	-3.2	F4	papa01	246	-5.3	F4
papa02	236	-8.8	F4	papa02	220	-6.9	F4
sasa01	308	-4.4	Cz	sasa01	314	-2.4	F4
sasa02	334	-6.8	F4	sasa02	316	-4.7	F3
white01	232	-2.5	F3	white01	228	-4.4	F4
white02	204	-8.1	F3	white02	240	-4.4	F4

SEKI				HAYASI			
刺激音	潜時	最大電圧	電極	刺激音	潜時	最大電圧	電極
beep01	234	-5.4	F4	beep01	194	-6.8	Cz
beep02	230	-6	F4	beep02	202	-6.7	F4
brown01	224	-8.5	F4	brown01	236	-10.5	F4
brown02	216	-5.3	F4	brown02	224	-8.1	F4
papa01	240	-6.3	F3	papa01	236	-8.1	F3
papa02	206	-6.8	Cz	papa02	236	-8.1	F3
sasa01	236	-4.3	Cz	sasa01	284	-5.9	F3
sasa02	258	-4.4	Cz	sasa02	258	-5.5	Cz
white01	216	-5.4	Cz	white01	236	-10.6	F3
white02	210	-3.4	F4	white02	228	-8.5	F3

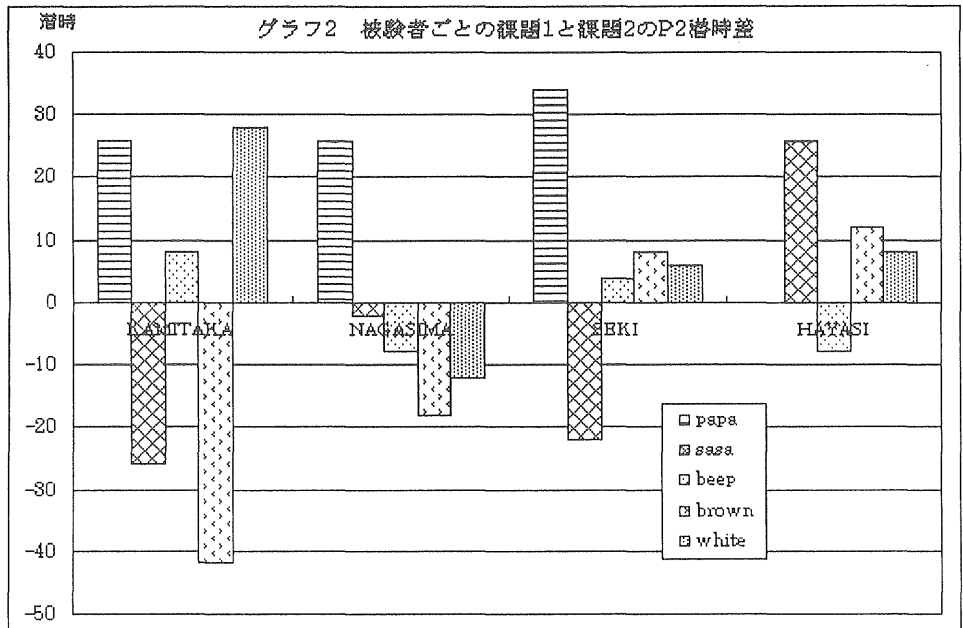
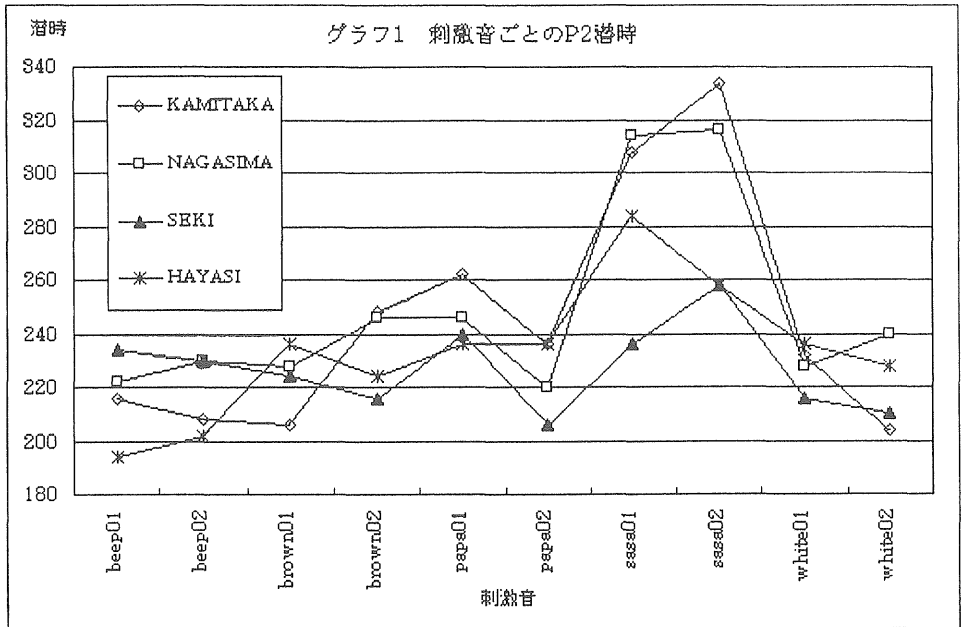
表 2: 課題 1 と課題 2 との潜時差

KAMITAKA	潜時	最大電圧	電極		潜時	最大電圧	電極
papa01	262	-3.2	F4	beep01	216	-5.3	Cz
papa02	236	-8.8	F4	beep02	208	-6.5	Cz
潜時差	26				8		
sasa01	308	-4.4	Cz	brown01	206	-6.2	Cz
sasa02	334	-6.8	F4	brown02	248	-6.9	F4
潜時差	-26				-42		
				white01	232	-2.5	F3
				white02	204	-8.1	F3
				潜時差	28		
NAGASIMA							
papa01	246	-5.3	F4	beep01	222	-6.4	F4
papa02	220	-6.9	F4	beep02	230	-6	F4
潜時差	26				-8		
sasa01	314	-2.4	F4	brown01	228	-5.9	Cz
sasa02	316	-4.7	F3	brown02	246	-5.3	F4
潜時差	-2				-18		
				white01	228	-4.4	F4
				white02	240	-4.4	F4
				潜時差	-12		
SEKI							
papa01	240	-6.3	F3	beep01	234	-5.4	F4
papa02	206	-6.8	Cz	beep02	230	-6	F4
潜時差	34				4		
sasa01	236	-4.3	Cz	brown01	224	-8.5	F4
sasa02	258	-4.4	Cz	brown02	216	-5.3	F4
潜時差	-22				8		
				white01	216	-5.4	Cz
				white02	210	-3.4	F4
				潜時差	6		
HAYASI							
papa01	236	-8.1	F3	beep01	194	-6.8	Cz
papa02	236	-8.1	F3	beep02	202	-6.7	F4
潜時差	0				-8		
sasa01	284	-5.9	F3	brown01	236	-10.5	F4
sasa02	258	-5.5	Cz	brown02	224	-8.1	F4
潜時差	26				12		
				white01	236	-10.6	F3
				white02	228	-8.5	F3
				潜時差	8		

表 3: 実験後インタビューの結果 (後ろの数字は人数)

papa	パパ 2 ・ 馬場 ・ パバイヤ
sasa	笹 3 ・ ささ (よびかけ)
beep	道 ・ ほら (よびかけ) ・ 危機 ・ アラーム
brown	怒涛 ・ 土像 ・ 膝 ・ ((ゲームの) 攻撃する効果音)
white	中国 ・ ズボン ・ 支持 ・ スプレーの音





## 5 考察

### 5.1 グラフからの考察

全体での結果をまとめた表 1 およびグラフ 1 から考察する。言語音 papa と sasa を比較した場合に、papa よりも sasa の方が遅くなる傾向が看取される。/p/ より /s/ の方が反応時間が遅いというのは、林・笈 (1989) の結果と照らし合わせても妥当な結果といえる。しかし、課題 2 が課題 1 よりもより顕著にその傾向があらわれており、課題 1 の場合には逆転している被験者もある。ここから、脳波実験において、単発課題の教唆は意味を考えるよう示唆したほうがより有意な結果が得られるものと考えられる。60msec 程度の幅を仮定してデータを眺めた場合に、言語音 sasa は各被験者でのばらつきがはげしい。最もばらつきが少なかったのは人工音の white ノイズの課題 1 の場合であった。非言語音は 3 つ全てあわせてみても 190 ~ 250msec 程度のばらつきにおさまるが、言語音はそれでは収まりきらない。筆者ら自身のデータ<sup>9</sup>も併せて考えた場合ではあるが、多人数でデータを集めていった場合には、非言語音は 60msec (±30msec) 程度の範囲にデータの分散が収まるのではないかと推測できる。

被験者ごとの課題 1 と課題 2 の潜時差を表 2・グラフ 2 にまとめた。課題 1 と課題 2 の間には、一応の差異はあるようだが、個人差がはげしいことがグラフ 2 から読み取れる。非言語音よりも言語音の時により大きな差異のた被験者が多かったが、逆転している被験者もあり、はっきりしたことはいえない。ただし、全体で見た場合には、言語音に関する差は有意であるようにみうけられる。P2 反応の遅速には一定の傾向は見られなかった。

### 5.2 意味付けの方法に関する考察

また、課題 2 は各被験者で意味付けの方法に傾向の違いが見られた。実験中にとったインタビューから、被験者の「意味付け」の方法は大きく 2 通りにわかれた。

主に音の形・音感からの連想

sasa=筈

---

<sup>9</sup>ここには公開していない

papa=パパ、パイパイ

brown=土蔵、怒濤、膝

主に音の役割からの連想

sasa=ささ (よびかけ)

beep=アラーム、危機、ほら (呼びかける声)

brown= (ゲームの) 攻撃する音

しかし、データから意味付けの方法と潜時との対応関係は読み取ることができなかった。

### 5.3 音と意味との対応関係

大辞典 6(1996:53-54) の意味の項をみると、「意味」を「意味 (Sinn)」と「意義 (Bedeutung)」に大別した上で、大辞典 6(*ibid.*:54) で次のような解説がある。

意義を内容とする考え方では、言語が容れ物でその中に意義を容れているという考え方である。言語の場合、その容れ物は音形式である。いうまでもなく、この表現は比喩的な表現であるが、音形式と意義とが密着していて分かちがたいことを、形式と内容の関係で表わしているのである。事実、1つの言語集団の中では、一定の音形式は一定の意義と不可分な状態にある。そこでは、一定の音形式を聞けば、一定の概念が自動的に喚び起こされる。この自動性が、音形式即概念という考えに導く。それが高じると、いわゆる「音義節」に陥る。確かに、意義の裏付けのない音形式は、単なる音連続であって言語音 (speech sound) ではない。言語音である以上、そこには意義がある。意義があってはじめて音連続は音形式となるともいえる。

「意義の裏付けのない音形式は、単なる音連続」であるという点には同意できるものの、「意義があってはじめて音連続は音形式となる」という点には、必ずしも同意できない。本実験では、言語音および非言語音の両者に語彙的意味を考えさせた。結果としてあらわれたのは、たとえ意義を考えたにしても、その出発点が言語音であるか非言語音であるかによって、ERP におけるピーク潜時に差異が生じるということである。本実験結果から考える限り、意義を優先的に

捉えた「意義があってはじめて」という言い方は、疑わざるを得ない。むしろ、言語の意義というものは、言語音があってはじめて成立するという方が妥当であると考えられる。

一言に「意味」といった場合でも、その解釈は人によって異なり、人工音にたいする反応と言語音（人間の「声」）に対する反応はあきらかに違う。これは、今後言語学分野での脳波実験課題を考えるうえでの一つの指針となるであろう。提示に対する同条件性を確保するために提示音を合成する場合にも、完全な人工合成音を使用しては本来の傾向を見失いかねない。また言語の「意味」に関する課題を考える上で、被験者に対してどのような教唆を与えるのが適切かは、慎重に検討する必要があるといえる。先行研究において確認されていた事象と同様の傾向を確認するとともに、より適切な実験結果を得るためには目的に沿った提示音と適切な教唆が不可欠であることを示した。適切な教唆を与え、目的意識をもってもらうことで被験者はより課題に集中して取り組むことができるといえる。

## 6 結論

以下のような結論が得られた。

1. 言語音について、P2 潜時は非言語音より相対的に遅くなる傾向があった。
  2. 言語音 papa は sasa と比較して P2 潜時が早い。また、この傾向は課題 2 の場合により強くあらわれる。
  3. 非言語音 3 つについて、P2 潜時反応の個人差は比較的小さかった。一方、言語音 sasa では被験者間でのばらつきが激しかった
  4. 非言語音よりも言語音の P2 潜時反応で、課題 1 と 2 の結果に差異のある傾向が強かった。
3. 4. から、言語音そのものに意義はないにせよ、言語音（人の声）であることが意義を解釈する上で極めて重要な役割を果たしていると結論づけられる。言語の音と意味とは恣意的関係にあるとはいえ、どんな音でもいいわけではなく、言語の意味は言語音から成り立つことを傍証できたと考える次第である。

## 7 展望

### 7.1 多人数の被験者による実験と少人数の被験者による実験

今回の実験では教唆の有無による結果の差異についてははっきりとした結果を出すことができなかった。このテーマを実証するためにはより多人数の被験者での実験が必要になるだろう。また、多人数の被験者を用いてデータを集めていった場合には、非言語音は 60msec. ( $\pm 30$ msec.) 程度の範囲にデータの分散が収まるのではないかと推測できる。

言語音を聞いたときの絶対的な潜時には個人差があったが、papa より sasa が遅いという相対差には共通性がみられた。このことから単一被験者による実験計画も十分に意味あるものであるといえるだろう。

### 7.2 意味に関する実験について

個人の意味の解釈には大きな差がでた。城生百太郎 (*ibid.*:303) には、

ここから理論言語学者たちにとっての脳への関心は、専ら入力された言語形態に即した言語内容を出力する、あるいはその逆に、入力された言語内容に即した言語形態を出力するといった一種の計算プログラムを想定しこれを追究するといった、計算モデル論的レベルに置かれてきたということは否めない。

と述べられている。意味解釈の多様性は計算モデル論的レベルでは包括しきれない。意味に関する実験を続けるならば、この差をどのように分析、定量化するのか、その方法を模索していくことが今後の課題といえよう。

### 7.3 まとめ

ことばに関する脳波実験では、複雑に絡み合った要因をどこまで単純化するかが問題となることが多い。今回の実験は、今後の脳波実験課題での適切な刺激音と課題提示という方法論を考える上で、有効であったと考える。

## 参照文献

- 林実・笈一彦 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』2-5-4. 355-356. 日本音響学会
- 亀井孝・河野六郎・千野栄一編著 1996. 『言語学大辞典』6. 三省堂 (=大辞  
典 6)
- 城生佰太郎 1997. 『実験音声学研究』勉誠社
- Laver, J. 1994. *Principles of phonetics*. Cambridge Univ. Press.
- Pinker, S. 1994. *The Language Instinct*. William Morrow and Company. (棕  
田直子訳『言語を生みだす本能』(上)(下) NHK ブックス. 1995)

## Experimental ERP Study about Sound Influences on Supposition of Meanings

Erika NAKAMURA & Takahiro FUKUMORI

Purpose: This experience shows that what difference between "voice" and "sound" on cognition by ERP.

In this experience, use five stimuli (two kinds of voice, and two kinds of noise and beep). Two stimuli that kind of human-voice are Japanese monotonously voices, "papa" and "sasa". Three kinds of stimuli are a simple beep, and a brown noise, a white noise.

At session1, We ask a subject that "just listen to stimuli". At session2, We ask a subject that "listen to stimuli, and pay attention to meanings of that."

In this case, we focused on P2, measured and analyzed that peek latency.

Result: human-voices are late relatively human-voice "sasa" P2 peek latency late than "papa" Kinds of "Non human-voice" distribution fewer dispersals than kinds of "human-voice." Mostly subjects, human-voice has more difference between session1 and session2.

Therefore, one distinguishes human-voice from non voice, also if one is not an aphasic, and it is important factor that a sound is a kind of human-voice when one considers meanings of it.