

Title	連続母音中の中心母音知覚に寄与するわたり部の時間構造に関する研究
Author(s)	田高, 礼子
Citation	
Issue Date	2002-03
Type	Thesis or Dissertation
Text version	author
URL	<a href="http://hdl.handle.net/10119/1571">http://hdl.handle.net/10119/1571</a>
Rights	
Description	Supervisor:赤木 正人, 情報科学研究科, 修士

修 士 論 文

連続母音中の中心母音知覚に寄与する  
わたり部の時間構造に関する研究

北陸先端科学技術大学院大学  
情報科学研究科情報処理学専攻

田高 礼子

2002年3月

# 修士論文

## 連続母音中の中心母音知覚に寄与する わたり部の時間構造に関する研究

指導教官 赤木正人 教授

審査委員主査 赤木正人 教授

審査委員 党建武 助教授

審査委員 宮原誠 教授

北陸先端科学技術大学院大学  
情報科学研究科情報処理学専攻

010066 田高 礼子

提出年月: 2002年2月

## 概要

本研究では、わたり部の時間情報がどのように知覚に用いられているのかを詳しく調べるため、3連続音声の中のわたり部に対して時間構造の変形操作と、中心母音部分の雑音置換を行った分析合成音声による音声刺激を用いて聴取実験を行った。実験結果を被験者ごとの母音の音韻境界と比較することで、中心母音が回復されるわたり部の変化速度と中心母音方向へ向かう定常部の長さの規則的な関係と、それによる中心母音の予測機構が示唆された。

# 目次

<b>第1章</b>	<b>序論</b>	<b>1</b>
1.1	はじめに	1
1.2	背景	2
1.2.1	連続音声知覚に含まれる問題	2
1.2.2	フォルマント遷移を用いた補正機構の存在	4
1.2.3	フォルマント遷移の時間変化の追跡と予測機構の存在	5
1.2.4	合成音声を用いることの問題	6
1.2.5	自然音声によるわたり部の重要性の証明	6
1.3	本研究の目的	9
1.4	本論文の構成	10
<b>第2章</b>	<b>実験1、音節の時間構造変形の影響</b>	<b>11</b>
2.1	実験目的	11
2.2	刺激作成方法	11
2.2.1	時間伸縮による変化速度の操作	11
2.2.2	雑音置換	17
2.3	刺激の構成	18
2.4	実験方法	19
2.5	実験結果	21
2.5.1	/iai/の結果	21
2.5.2	/ioi/の結果	25
2.5.3	/aia/結果	28
2.6	実験1、結果のまとめと考察	31
<b>第3章</b>	<b>実験2、時間圧縮部の圧縮率と時間伸張部の時間長の影響</b>	<b>33</b>
3.1	実験目的	33

3.2	刺激の構成	33
3.2.1	時間伸縮の仕方	33
3.3	雑音置換方法	34
3.4	実験方法	34
3.5	実験結果	36
3.6	実験 2、結果のまとめと考察	46
<b>第 4 章</b>	<b>実験 3、単独母音の音韻境界の調査</b>	<b>47</b>
4.1	実験目的	47
4.2	刺激の作成方法	47
4.3	刺激の構成	48
4.4	実験方法	48
4.5	実験 3、結果と考察	48
<b>第 5 章</b>	<b>実験結果のまとめと関数近似による考察</b>	<b>56</b>
5.1	実験 1、2、3 のまとめ	56
5.2	関数近似の目的	57
5.3	音声分析方法	57
5.4	関数近似の結果	58
<b>第 6 章</b>	<b>考察</b>	<b>63</b>
<b>第 7 章</b>	<b>謝辞</b>	<b>64</b>

# 第1章 序論

## 1.1 はじめに

連続音声では、音素が単独発話時とは異なる音響的特徴を示す調音結合が起こる場合がある。それでも聞き手はその音素に関する情報を上手く利用し、発話者の意図した通りの音素が知覚できる。これは知覚の際に、音素の前後の特徴に含まれる情報を用いているためであると考えられる。その音素前後の特徴の中でも特にわたり部は調音の動きの表れであり、また連続音声では定常部が少ないことから、連続音声の知覚では重要な働きをすると考えられる。

実際、わたり部に関する研究は数多く行われているが、わたり部を模擬した合成音声では自然音声に含まれる情報を正しく与えられない可能性がある。事実それは、Text-to-Speechの品質低下の原因としても見られている。反対に、自然音声を用いた聴取実験は、自然音声の操作の難しさのため、わたり部のその詳しい特徴を調べることは難しくなっている。このようなことから、わたり部がどのように知覚に用いられているのかは未だ詳しい説明ができていない。

そこで、なるべく自然性を失わない音声刺激を用いて、連続母音中における中心母音知覚において、わたり部がどのように用いられているのかを詳しく調べる必要がある。そして、わたり部の役割を明らかにすることは音声知覚単位を明らかにする手がかりとなり、またより自然な音声合成への応用も期待される。

## 1.2 背景

### 1.2.1 連続音声知覚に含まれる問題

音素を連続的に発話しようとするとき、調音器官の連続的な動きのために、連続音声中の音素の音響的特徴も連続的に推移する。そのため、連続音声は定常部が少なく、時間的に変化し続ける部分が多く存在する。

その中には、音素から音素への過渡部分で「わたり」と呼ばれる部分が存在する。わたり部は各音素間の中間的な音響的特徴を通過するが、その中間的な音韻が知覚されることはない。

また連続的な発話の中では、前後音素の影響を受けて、単独発話時の調音位置まで十分に達せずに次の音に移る場合がある。この現象は「なまけ」と呼ばれている。

これらの「わたり」や「なまけ」の現象を引き起こす原因を調音結合と言うが、この調音結合により、連続音声中の音素は、単独発話時とは異なる音響的特徴を示す場合があり、同一話者の音声にも関わらず、連続音声中の母音の音響的特徴が単独発話母音のものと対応がとれなくなる原因となっている。調音結合の影響の例を、単独発話母音を弁別できる優勢な特徴の一つとして考えられている  $F_1$ 、 $F_2$ （フォルマントのピーク周波数の小さい方から、第1、第2、第3フォルマントと順に呼ぶ、その略式。）の値について見てみる。図1.1のような各母音を決定する  $F_1$ 、 $F_2$  周波数値を各母音の目標値とすると、図1.2の概略図で示すように、連続音声中の母音はこの値に到達せず、次の音素へと移ってしまう場合がある。

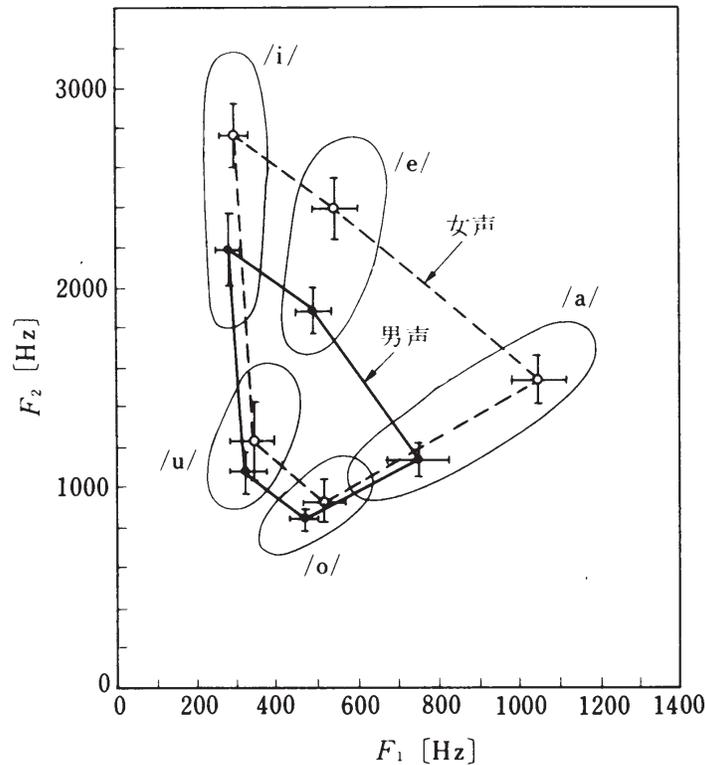


図 1.1: 母音の F1、F2

さらに、図 1.2 に見られるような音響的特徴の変形は、3 連続母音から切り出した中心母音に対する同定実験によっても示されている [4]。表 1.1 の 3 連続母音から切り出して呈示した中心母音に対する同定結果と、表 1.2 の 3 連続母音で呈示したときの中心母音の同定結果を比べると、切り出した中心母音に対して異聴が起こっていることがわかる。

調音結合のほかにも、発話様式、発話速度、単語中での位置の違いによって、同一話者中の同じ音素でも、その特徴が大きく異なってくる。それでも、私たちはそのような曖昧になった母音を、話者の意図する音素として知覚することができている。

以上に見られるように、単独発話母音知覚の説明のみでは、連続音声での母音の知覚を説明することはできない。単独発話母音の知覚とは異なり、連続音声での母音の知覚は母音部分だけでなく、母音に関する利用可能な情報をできるだけ用いていると考えられる。その利用可能な情報は、わたり部や前後音素などにも含まれると考えられているが、連続音声は時間変化部分が多いことを考えると、わたり部は大きな情報源となっていると考えられる。そして、わたり部の重要性は多くの聴取実験で示されている。しかし、どのよう

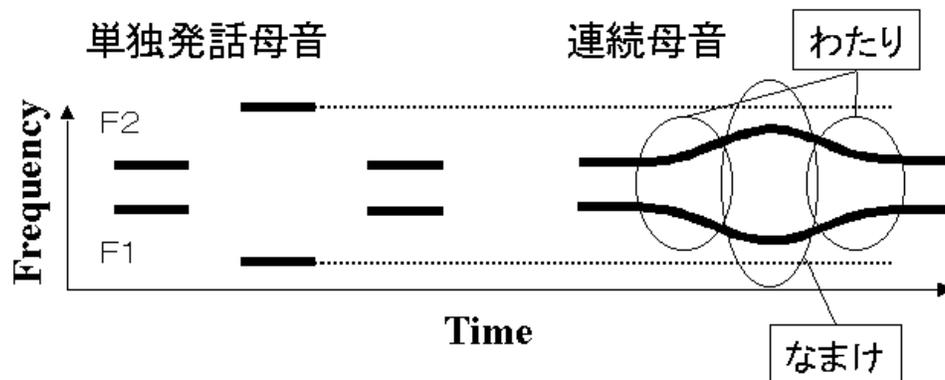


図 1.2: 調音結合を受けた連続音声の F1、F2 の概略図

刺激	回 答				
	i	e	a	o	u
i	218	22			
e	57	171			
a		88	148	3	1
o		1	32	125	82
u				9	231

表 1.1: 3 連続母音から切り出した中心母音の同定結果 [4]

に用いられているのかはよくわかっていない。

### 1.2.2 フォルマント遷移を用いた補正機構の存在

フォルマント遷移の影響によると考えられる補正機構の存在は、合成音声を用いた聴取実験によって多く示唆されている。

Lindblom ら [5] の聴取実験では、補正機構が初めて提案されている。

ここで用いられた音声刺激は、F1 を定常とし、F2、F3 のみが放物状に変化する CVC 音節 /wVw/ (周波数値が高い方へ凸型に変化) /jVj/ (凹型に変化) である。そして、凸型変化の極大値、凹型変化の極小値を /u/-/i/ の間で変化させ、中心母音 V の弁別実験を行い、母音の音韻境界を決定している。放物状変化の速度は、全体長が 200ms と 100ms になるものを 2 種類用意している。

またこの他に、F2、F3 も定常とした単独母音 /u/-/i/ の音韻境界を決定し、CVC 音節の V の音韻境界が、単独母音の音韻境界からずれて (shift) していることを確認した (図 1.3)。

刺激	回 答				
	i	e	a	o	u
i	225	15			
e	8	232			
a			234	2	4
o				228	12
u	5		2	5	228

表 1.2: 3 連続母音のままでの中心母音の同定結果 [4]

そしてこの音韻境界の shift が、図 1.4 で表すような補正を表すものとして考えられた。音韻境界の shift の大きさは図 1.3 の概略で表すように、/jVj/ よりも /wVw/ で、また変化速度が速い刺激（母音部 100ms 長の刺激）でより大きくなっている。よって補正機構には、フォルマント遷移の変化速度と方向が影響し、調音的に曖昧になった母音をその補正機構により知覚的に回復しているだろうと考えられた。

### 1.2.3 フォルマント遷移の時間変化の追跡と予測機構の存在

音声知覚ではなく、聴覚による情景解析に焦点をあてた研究では、Masuda ら [6] によって、連続的に変化しているフォルマント遷移に対する追跡と予測の機構が存在することが示されている。

この実験で用いた刺激は F 1 または F 2 を直線的に変化させ、その他の F4 までは一定とした 2 連続母音である。この第 2 母音について、その後続が無音ではなく雑音があるとき、フォルマントの変化が延長した方の母音に弁別する傾向が見られた。その他の証明実験からも、後続雑音の中に目標値を抽出していると考えられ、フォルマント遷移の追跡と予測の存在を示している。

この結果は、Lindblom らの言うような、フォルマント遷移を用いた補正機構による中心母音の回復についての説明を大きく助ける結果になると思われる。その結果の一部を図 fig:Masuda に示す。

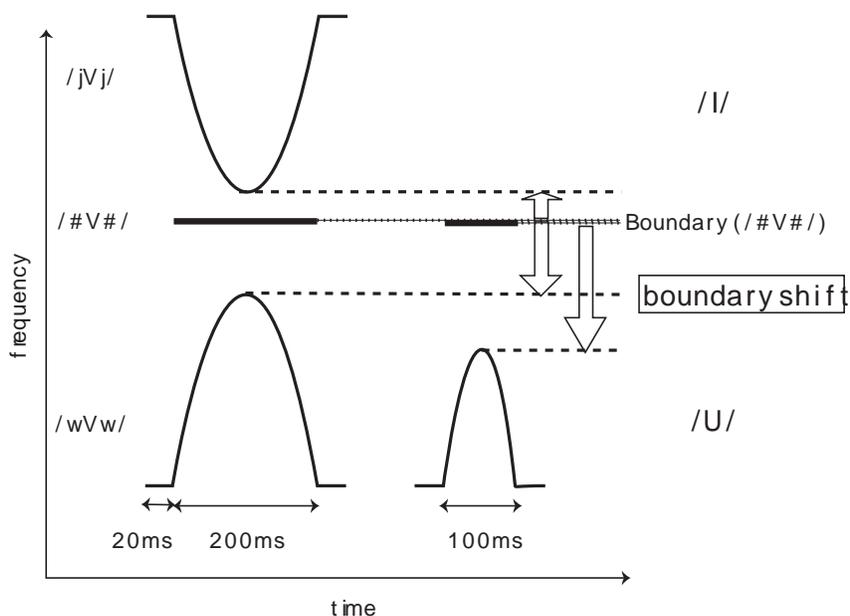


図 1.3: Lindblom らの実験結果の概略.

#### 1.2.4 合成音声を用いることの問題

合成音声を用いることで、自然音声のわたり部を簡単な構成として模擬し、様々に操作することができる。よって、注目する特徴のみを変化させ、その他の特徴は変化させていない合成音声を用いることで、その注目する特徴のみに対する知覚現象を測定できるという利点がある。

しかし反対に、簡単な構成の合成音声では、複雑な構造をもった自然音声を再現することはできない。実際には、直線に変化するフォルマント遷移や、単一の特徴のみが変化するような自然音声は存在しない。さらに、複雑な構成である自然音声のわたり部の中で、その注目した特徴が実際に用いられているのかもわからない。

自然性を失った合成音声を用いて得られた知覚現象は、自然音声の知覚にあてはまるかわからないため、上で述べられている補正機構や、フォルマント遷移の追跡と予測機構も自然音声の知覚で存在するのかわからない。

#### 1.2.5 自然音声によるわたり部の重要性の証明

1.1 で述べられている桑原の実験 [4] では、3 連続母音の中心母音だけでは異聴が起こることを示しているが、わたり部の重要性を証明してはいない。

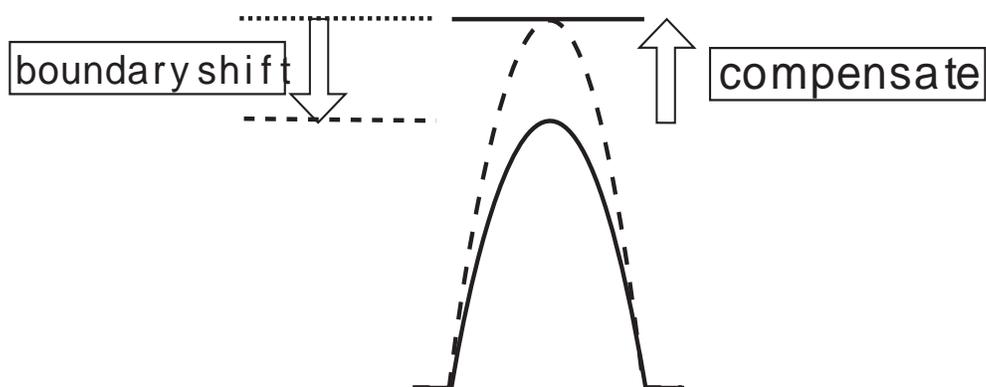


図 1.4: 音韻境界移動の解釈.

わたり部の重要性は、主に音節の中心母音部分を取り除いた刺激に対する中心母音知覚の結果により示されている。Strange[2] は、CVC 音節の中心母音 V の知覚に十分な情報を与える部分は、

- (a) 中心母音
- (b) 母音固有の時間長
- (c) わたり部

のいずれであるかを調べるため、その保存されている部分がそれぞれ異なる、以下の 7 種類の刺激の弁別実験結果の誤り率を比べている。

- (1)Control (unmodified syllables) : (a),(b),(c) を保存.
- (2)Ctr (Centers, 各母音の固有の長さで切り出し) : (a),(b) を保存.
- (3)NDC (neutral-duration center, 各母音の平均長で切り出し) : (a) を保存.
- (4)SC (silent-center syllables) : (b),(c) を保存.
- (5)NDSC (neutral-duration silent-center syllables, SC 無音区間を各母音平均長に操作) : (b) を保存.
- (6)I (initials only).
- (7)F (finals only).

図 1.7 は弁別実験結果の誤り率である。

- 母音固有時間長が保存されている刺激 (SC) と (Ctr) は、保存されていない刺激 (NDSC)、(NDC) より誤り率が小さい。よって、母音固有の時間長が情報源であると見ることができる。
- しかしその誤り率の差は、わたり部が保存されている (SC) と (NDSC) 間の差よりも、母音部が保存されている (Ctr) と (NDC) 間での差の方が大きいことから、わたり部が

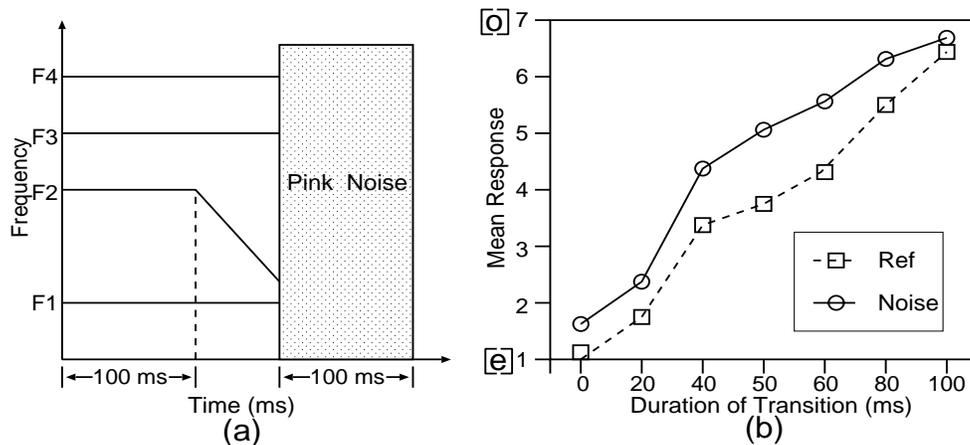


図 1.5: Masuda らの実験結果

(a) 刺激構成 (後続雑音ありの/e/→/o/について). (b) (a) の結果 (Ref は雑音なし, Noise は雑音あり). 横軸は遷移の長さ. 縦軸は回答率. 回答は Ref より Noise の方が第 2 母音へ偏っている.

大きな情報源であることがわかる。

- さらに、(I)、(F) では誤りがかなり大きくなることから、入わり部、出わり部の両方が必要であることがわかる。

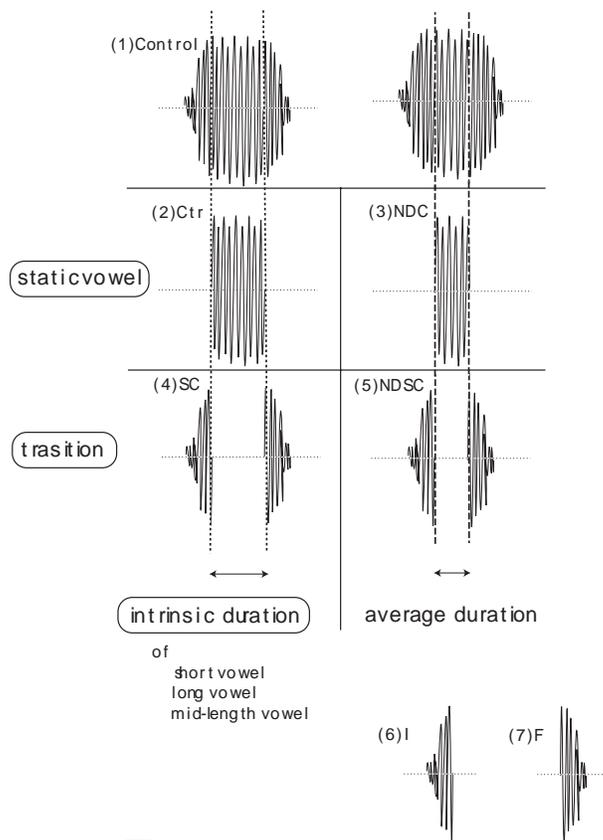


図 1.6: Strange の刺激構成.

Strange は、2つのわたり部それぞれに母音に関する調音的な特徴が含まれており、さらにそのわたり部により与えられる音節全体の時間変化構造が母音の弁別を決定するだろうと考えている。この立場では、中心母音は補正機構により回復されることは言及せず、刺激そのものが持つ母音固有の構造により弁別されると考えている。

しかし、自然音声の操作が難しいためもあり、わたり部の詳しい特徴はまだわかっておらず、わたり部によって時間構造がどのように決定され、どのように用いられるのかもわかっていない。

ここでわかっていることは、無音置換で残されたわたり部により音節の時間構造が決定され、その時間構造が母音知覚に影響することである。

### 1.3 本研究の目的

合成音声を用いた実験と自然音声を用いた実験の両方で、わたり部の時間変化が母音知覚に影響を及ぼすことがわかっている。合成音声を用いた実験からは、補正機構の存在

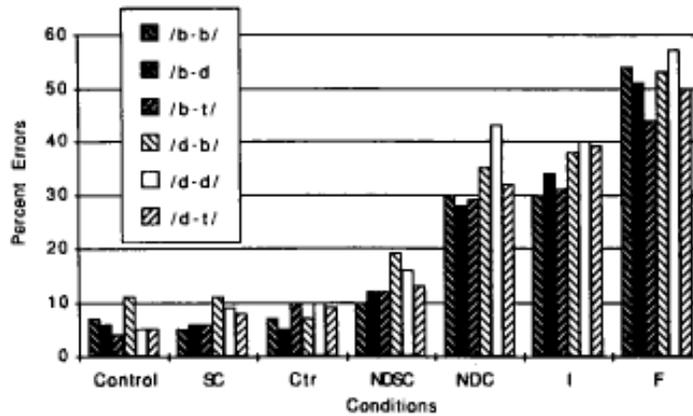


図 1.7: 各刺激の誤り率 [2]

やフォルマント遷移の追跡や予測機構が示されているが、自然性を失った合成音声による結果を、どれだけ自然音声にあてはめることができるかはわからない。また、自然音声を用いる実験では音声の操作が難しいため、わたり部の詳しい特徴を調べることも難しくなっている。そのようなことから、わたり部は実際どのように用いられているのかは未だわかっていない。よって、わたり部について詳しく調べるためには、合成音声に見られる自然性の劣化と、自然音声における操作性の制限を克服した音声刺激を用いて、有用な知覚結果を得ることが必要であると考え。そこで今回は、音声刺激に分析合成系 STRAIGHT[7] による分析合成音を用いる。そして本研究では、日本語 3 連続母音のわたり部について時間操作を行い、その中心母音知覚における影響を調べ、わたり部がどのように用いられているのかを調べることを目的とする。

## 1.4 本論文の構成

第 2 章では、実験 1 を示す。3 連続母音の全体長一定とし、その中での時間構造を変形したときの知覚の傾向を述べる。主な傾向の 1 つに、圧縮部の圧縮率と伸張部の時間長の影響が見られた。

第 3 章では、実験 2 として、圧縮部の圧縮率と伸張部の伸張率を変え、実験 1 の傾向のさらに詳しい知覚結果を述べる。

第 4 章では、実験 2 の結果を考察するために、単独母音の音韻境界を被験者ごとに調べた結果を示す。

第 5 章では、実験 2 の結果の分析を単独母音の音韻境界の結果と比較しながら行う。

第 6 章では、全体を通しての考察を述べる。

# 第2章 実験1、音節の時間構造変形の 影響

## 2.1 実験目的

自然音声を用いた Strange の結果では、無音置換で残されたわたり部により音節の時間構造が決定され、その時間構造が母音知覚に影響することがわかっていた。

ここでは、音節を一定長に固定し、その中での時間構造を様々に変形したときの、中心母音知覚への影響を調べ、その傾向を見つけることを目的とする。

時間構造の変形は、時間伸縮と中心母音部分の雑音置換によって行った。雑音置換の置換部分の長さを変えることで、残る音節部分の長さを変えている。そして残された音節部分の中心母音知覚への影響を見ることとした。

## 2.2 刺激作成方法

### 2.2.1 時間伸縮による変化速度の操作

音声の時間伸縮は、分析合成系 STRAIGHT[7] により音声の詳細な分析をし、その結果を用いた音声再合成の際に伸縮を行った。音節の全体長を固定して、その中で時間伸縮を行った。その一連の流れを以下に述べる。

例) 図 2.1 のソナグラムで表現される原音声の伸縮をする場合。

- STRAIGHT の音声分析により、高精度な時間周波数表現、基本周波数、駆動音源情報（非周期成分と周期成分）の分析を行う。
- 時間周波数表現の分析結果から、図 2.2 のような 30 次の線スペクトル（Line Spectral Frequency : LSF）を計算する。
- LSF からスペクトル変化率（Spectral Feature Transition Ratio : SFTR）図 2.3 を求める。ある時間  $n$  についての SFTR の計算は、区間  $[n-M, n+M]$  の LSF 一次ずつ

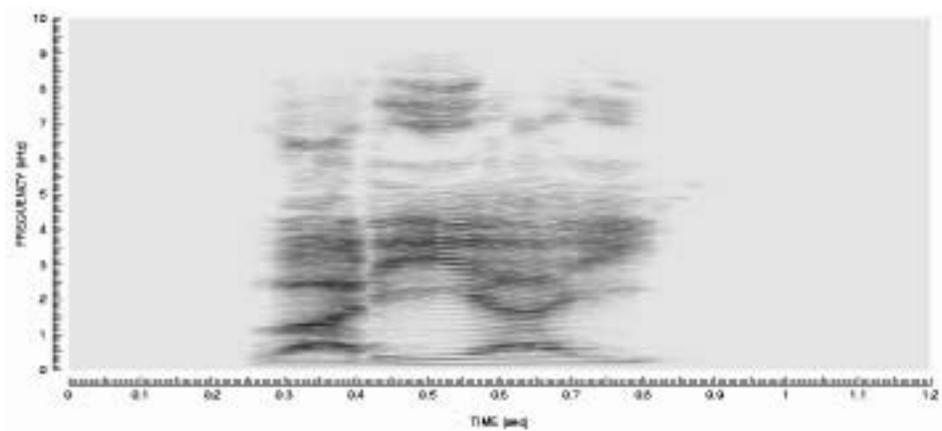


図 2.1: 原音声のソナグラム

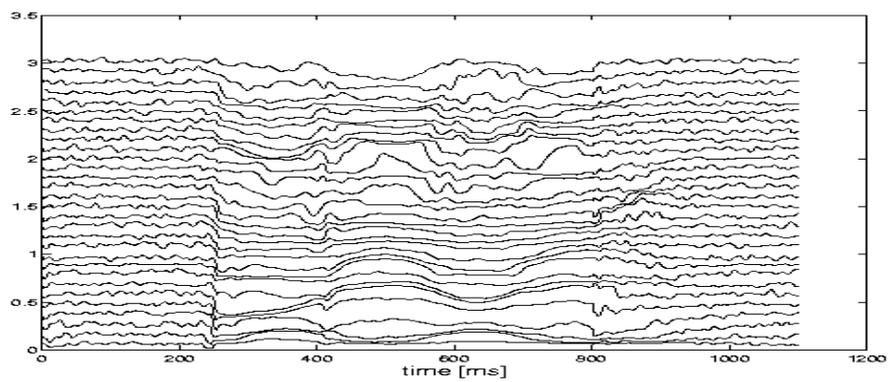


図 2.2: 線スペクトル

について、自己回帰直線の勾配の2乗を求め、それを30次分加算して得られる。以下表2.1にこれらの計算条件を示す。

LSF 次数	30 次
フレーム長 ( $2 * M + 1$ )	50 [ms]
フレームシフト	1 [ms]

表 2.1: 計算条件

- さらに図 2.3 のように、SFTR の極大点をわたり部の中心点、極小点を母音部の中心点として、これらを固定点とする。さらに、固定点間の区間を区間 1、2、3、4 と決定する。

区間 1	第 1 母音中心点から、第 1 と第 2 母音間のわたり部中心点まで。
区間 2	第 1、第 2 母音間のわたり部中心点から、第 2 母音中心点まで。
区間 3	第 2 母音中心点から、第 2 と第 3 母音間のわたり部中心点まで。
区間 4	第 2 と第 3 母音間のわたり部中心点から、第 3 母音中心点まで。
- 図 2.4 のように、区間 1 から区間 4 までのそれぞれの区間の中で、わたり区間と母音区間を決定し、それぞれの区間に時間伸縮を与える。図 2.4 はわたり区間を 10、50、90 % と決定した例。

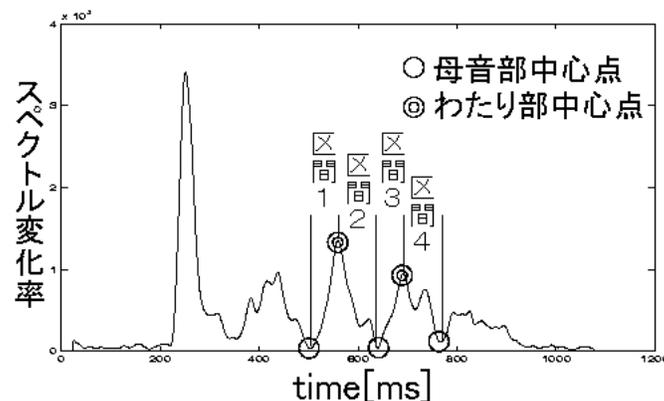


図 2.3: スペクトル変化率 (SFTR)

- わたり区間を圧縮 (母音区間は伸張) または、わたり区間を伸張 (母音区間を圧縮) するための時間伸縮関数  $imap$  を作成する。図 2.5 は 10 % わたり区間が 90 % わたり区間になるように伸張するための関数の考え方を表したものである。縦軸が原音声の時間、横軸が時間伸縮後音声の時間で表現される。

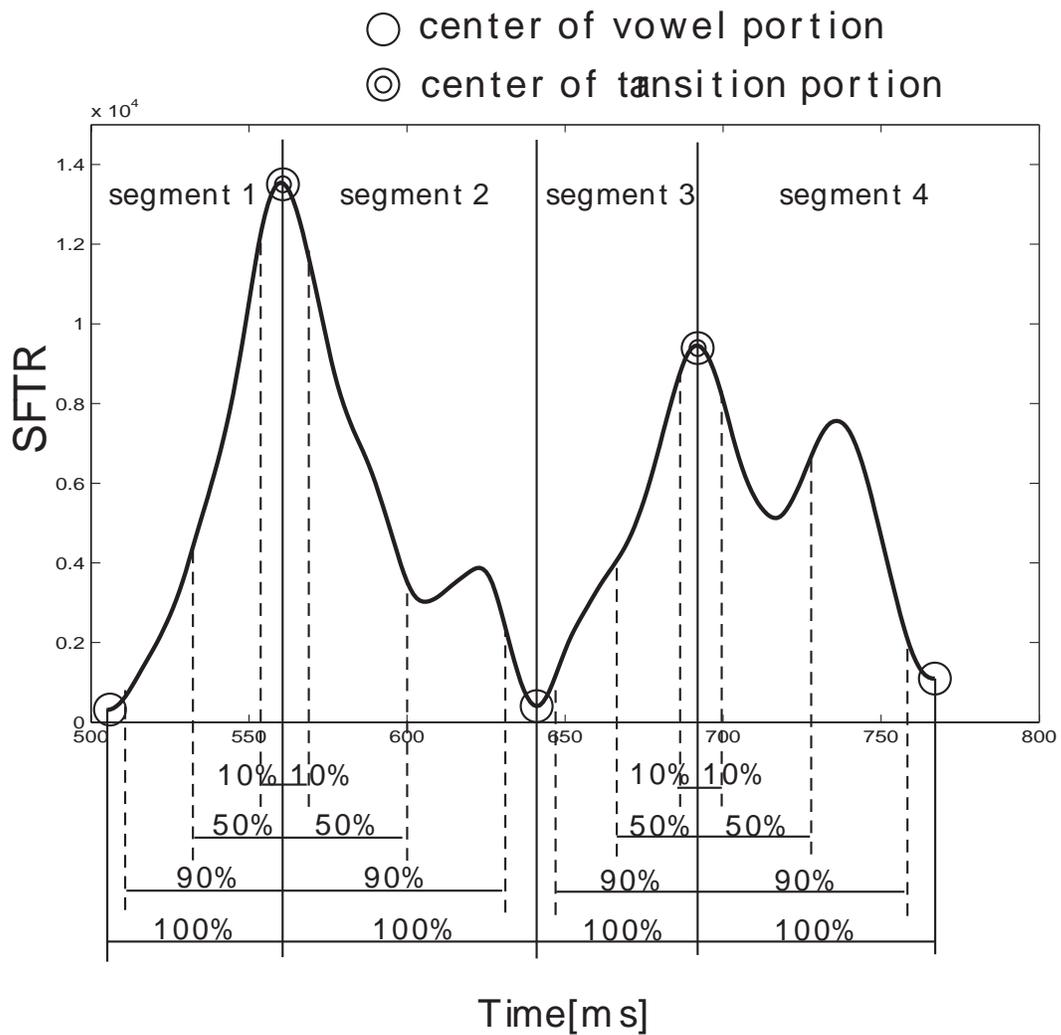


図 2.4: スペクトル変化率とわり、母音区間決定 (わり区間を固定点間区間の 10 %、50 %、90 %とした例)

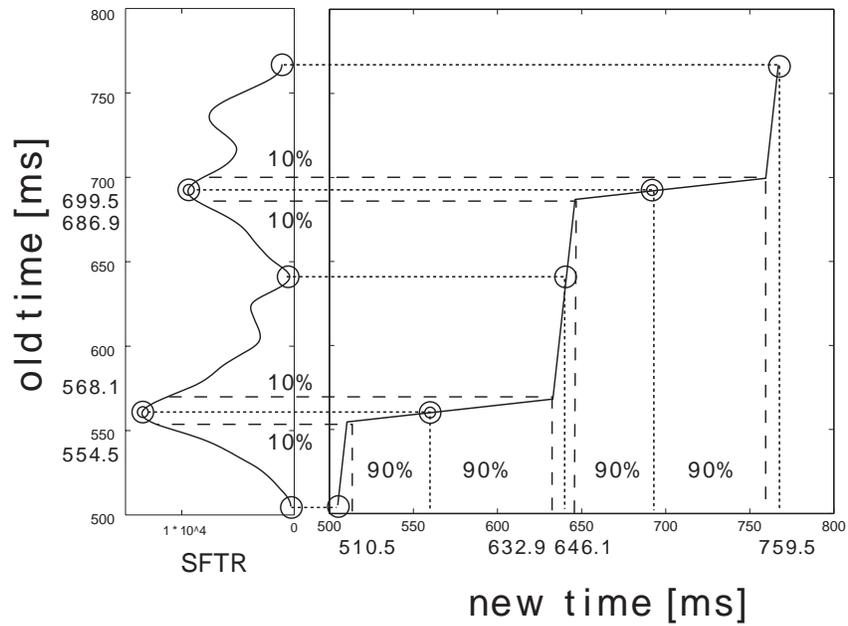


図 2.5: 時間伸縮関数の考え方 (10%あたり区間を90%区間まで伸張する場合)

- 図 2.6 に表されるのが実際の関数である。この時間伸縮関数  $imap$  と先に行った分析結果とを用いて、SRAIGHT により時間伸縮合成を行う。

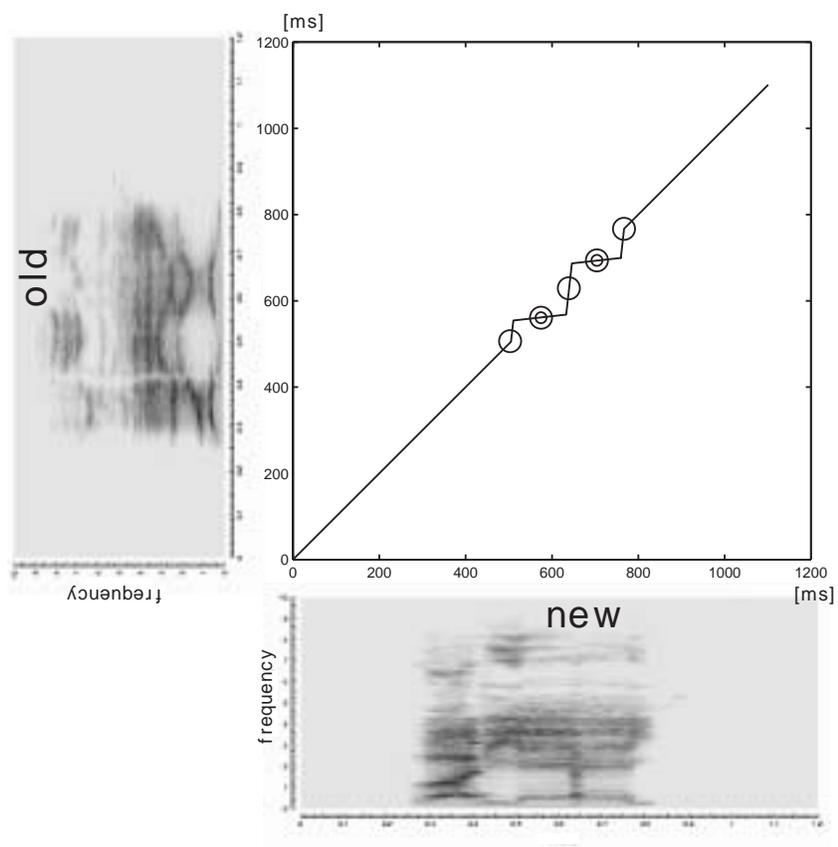


図 2.6: 時間伸縮関数  $imap$

## 2.2.2 雑音置換

今回は Strange らなどが用いている無音置換ではなく、雑音置換を行った。

柏野 [8][9] によると、雑音が音声の一部に重なることは、日常で起こる自然なことであり、無音置換に比べてアーティファクトは少ないと考えられている。また、自然音声 VCV 音節の一部を雑音置換または無音置換した実験でも、雑音置換の方が刺激の同定率が高く、特に有声の判断の劣化が小さいことが示されているためである。

雑音レベルは柏野の実験を参考としたが、今回は音節全体ではなく置換部分の音声の平均パワーを基準とし、それより 5 dB 大きい白色雑音を用いた。

雑音置換操作は時間伸縮操作後に以下の手順で続いて行われる。

- 雑音置換区間を決定し、その区間にある音声信号のパワーより 5 dB 大きい白色雑音により置換を行う。(図 2.7 に、例としてそのソナグラムを示した。)
- 第 1、第 3 母音が正しい母音として知覚され、その音韻が与える中心母音知覚への影響を除くために、第 1 母音中心点の 30ms 前、第 3 母音中心点から 30ms 後で単語から 3 連続母音を切り出し(図 2.7 では実線位置での切り出し)、3 連続音声の刺激とした。30ms という長さは、刺激を実際に聴いて、第 1、第 3 母音ができるだけ長くとれる範囲で決定した。

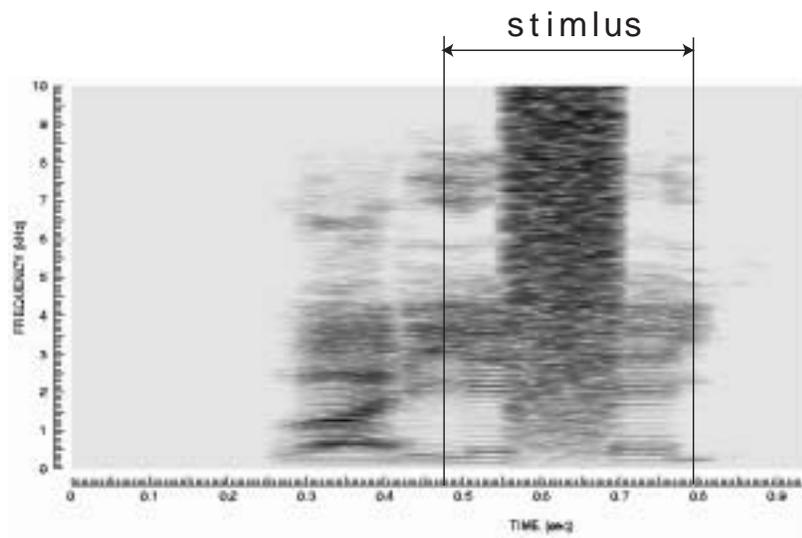


図 2.7: 雑音置換後のソナグラム

## 2.3 刺激の構成

- 用いた音声・ /iai/ in /waiiai/,
  - /aia/ in /zaiaku/,
  - /ioi/ in /nioi/.

(全て ATR 音声データベースの男性話者 mtm の単語データ.)

- 時間伸縮方法 (図 2.8 はその概形) SFTR から決定したわたり (母音) 区間 → 伸縮後わたり (母音) 区間長 (図 2.8) ・ 10 % (90 %) → 50 % (50 %) or 90 % (10 %).
  - 10 % (90 %) → 50 % (50 %) or 90 % (10 %).
  - 50 % (50 %) → 10 % (90 %) or 90 % (10 %).
  - 90 % (10 %) → 10 % (90 %) or 50 % (50 %).
  - 伸縮なしの normal.

- 雑音置換区間により、残される音節の長さ (図 2.9 はその概形) ・ 各時間伸縮刺激には 3 通りずつ与える。

long	時間伸張 (または圧縮) 後のわたり区間が半分以上残るもの。
half	時間伸張 (または圧縮) 後のわたり区間が丁度半分残るもの。
short	時間伸張 (または圧縮) 後のわたり区間が半分以下残るもの。

- 上記 3 種類のうち half は全刺激に共通で与えられている。
  - normal については図 2.8 にみられるように全 6 通りが与えられる。
- 全刺激 20kHz sampling, 16bit 量子化。

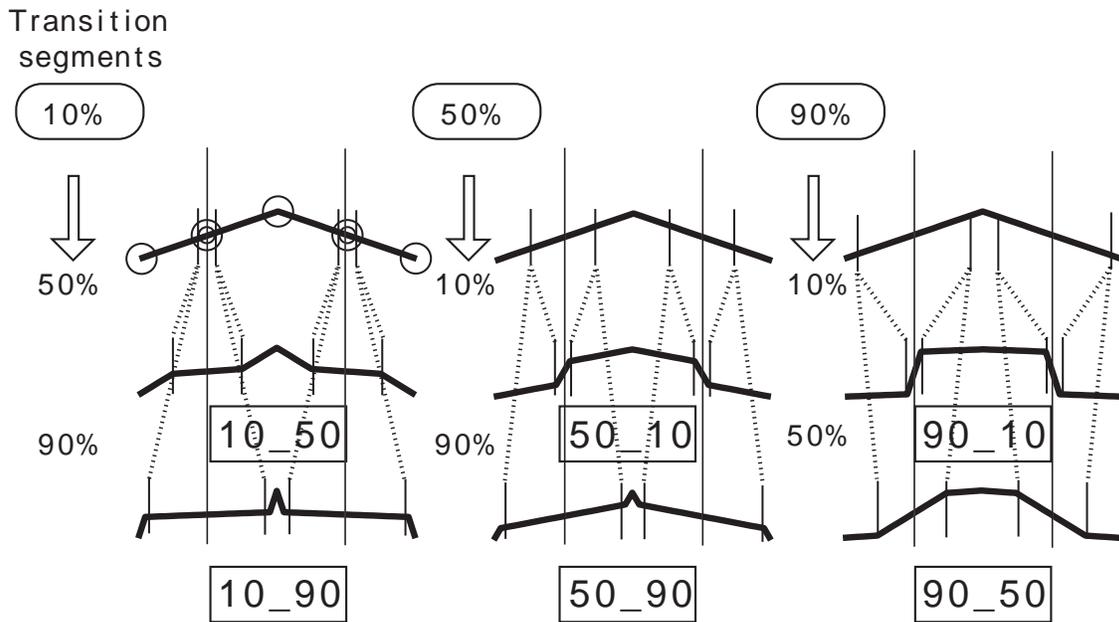


図 2.8: 時間伸縮区間とその伸縮

(図はわり区間の値を表示している。わり区間が伸張されるとき、母音区間は反対に圧縮されている。)

## 2.4 実験方法

- 被験者は大学院生の男性 3 名。
- 実験は 1 試行ずつ、防音室で個別に行った。
- 1 試行は全 72 個の刺激がランダムに呈示される。それを 5 回ずつ行った。
- 回答は、聴こえた音韻をノート PC のキーボードにより入力してもらった。
- 教示は、3 連続母音を変形したものに雑音混在した刺激が呈示されることを伝えた。注意点として、雑音が音節の背景に聴こえるときは音節の回答を、雑音が音節の一部として聴こえる場合は雑音をわかりやすい表記で回答するように伝えた。また、刺激の聴き直しを許した。
- 刺激はヘッドホンで両耳に、最大で 55dB SPL を越えない範囲で呈示した。

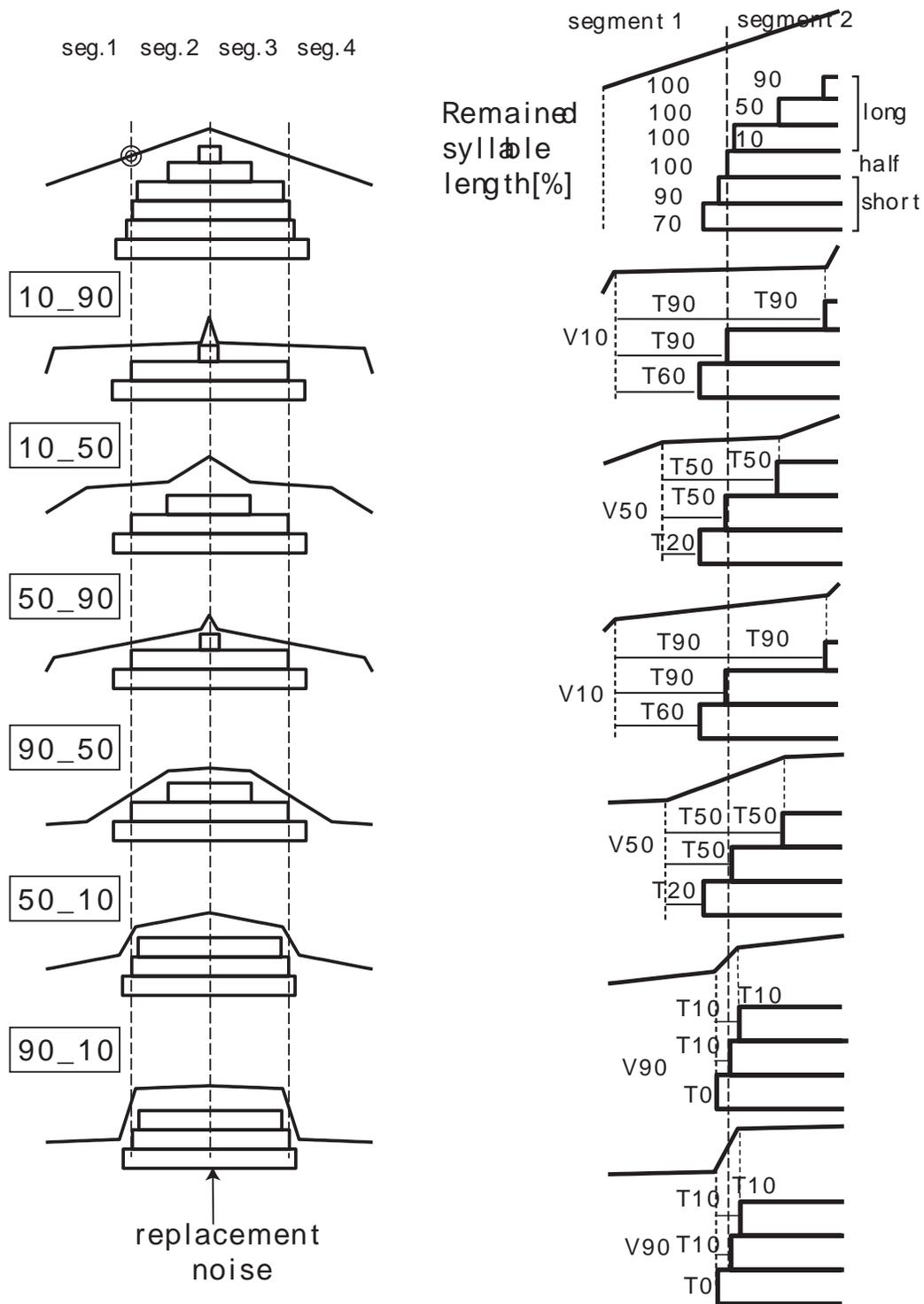


図 2.9: 雑音置換区間

## 2.5 実験結果

- 集計方法

雑音が音節の一部として知覚された場合についての回答は、雑音の表記をして、聞こえた通りの音節を回答してもらっていた。例えば、/iai/が雑音により途切れて聞こえる場合は、/iaxai/ というような回答である。

今回の集計では雑音の有無は区別せず、回答された中心母音で集計した。例に挙げた /iaxai/ については、中心母音に /a/ が聞こえているので、/iai/ と同じ回答結果「a」として扱っている。また、中心母音が全く知覚されていない（雑音にしか聞こえない）回答は「x」として扱っている。

しかし、回答された音節の中に対称 3 連続母音が含まれない場合は中心母音を決定できないため、そのままの音節を回答結果として扱った。例えば /aia/ に対する回答に見られる、被験者 3 の (ao) などがそうである。

また、回答率がかなり低い回答結果はグラフに示していない。

- 図の見方

3 連続母音ごとの結果は、2 つの図に分かれて表示してある。各図では、縦に被験者ごと、横に刺激の種類ごとに結果が並んでいる。刺激の並び順は、1 番目が normal で、2 番目からはわたり区間の伸張率が大きいものから圧縮率が大きいものへと続いている。その図の中の各グラフは、縦軸に回答された中心母音の回答率を、横軸に残された非雑音置換区間の長さを短い方から short, half, long で表している。長さの具体的な値は図の右端に概略図で示した。

### 2.5.1 /iai/ の結果

図 2.10 は時間伸縮していない normal と、わたり区間を伸張（母音区間を圧縮）した /iai/ の刺激に対する回答結果であり、図 2.11 はわたり区間を圧縮（母音区間を伸張）した /iai/ の刺激に対する回答結果である。

- normal とわたり区間伸張刺激との差。

normal の結果を見ると、被験者 1、2 はどちらもちょうど half の付近で中心母音の異聴が起こっており、short では中心母音が知覚されない「x」となっている。

しかし時間伸縮した刺激に対する被験者 1、2 の回答をみると、50 \_\_ 90 に対しては被験者 1 は、short における「x」の回答率が小さくなり、母音が知覚されている。被験者 2 は、normal の half では完全に「x」だったのが、この刺激の half ではその半分の率になり、中心母音が知覚されている。10 \_\_ 50 になるとさらに「x」の回答は少なくなっている。10 \_\_ 90 になると被験者 1 は「x」が見られなくなり母音の回答のみとなる。被験者 2 も同様、母音の回答のみとなっている。

よって、被験者 1、2 の結果からは、わたり区間の伸張率（母音区間の圧縮率）が大きくなるにつれ、回答「x」の割合は少なくなり、母音が知覚されるようになっている。

- わたり区間伸張刺激の伸張率と残った音節長の関係。

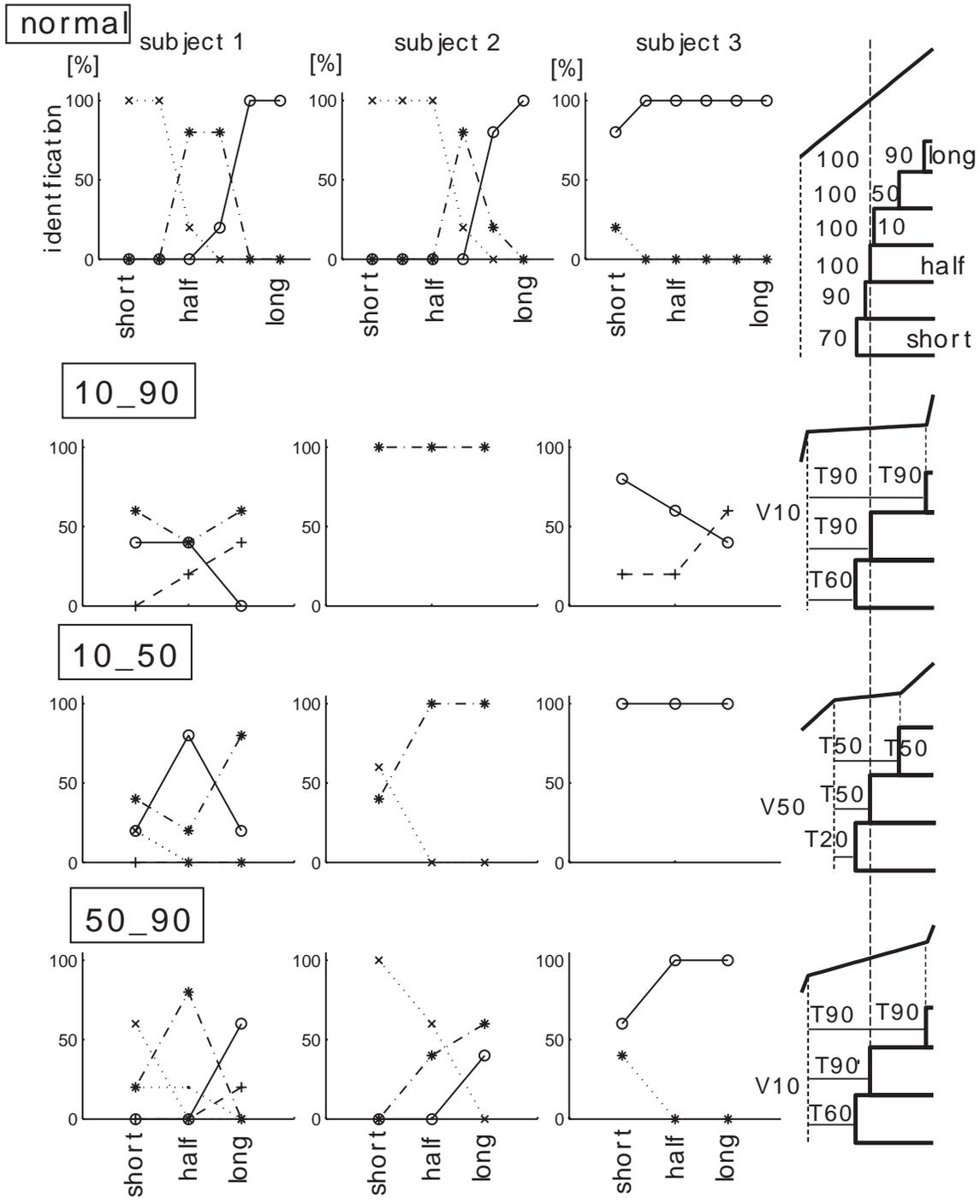
上記のように被験者 1、2 について normal と比べると、わたり区間伸張刺激では、その伸張率が大きくなるにつれて short の位置であっても中心母音が回答されるようになっている。

ただし、このとき被験者 1 では/e/から/a/の率が、被験者 2 では/e/の率が大きくなっているという違いがある。被験者 2 は結局 10-90 で/e/ のみの回答になっているが、被験者 1 は、long のときは異聴/e/が多く見られ、short になるにつれて元の母音と同じ音韻の母音が知覚されるようになっている。被験者 1 の 10-90 の結果は、被験者 3 の 10-90 にも類似して見られる。

よって、時間伸縮の影響と残る音節の長さに関係があるかもしれない。

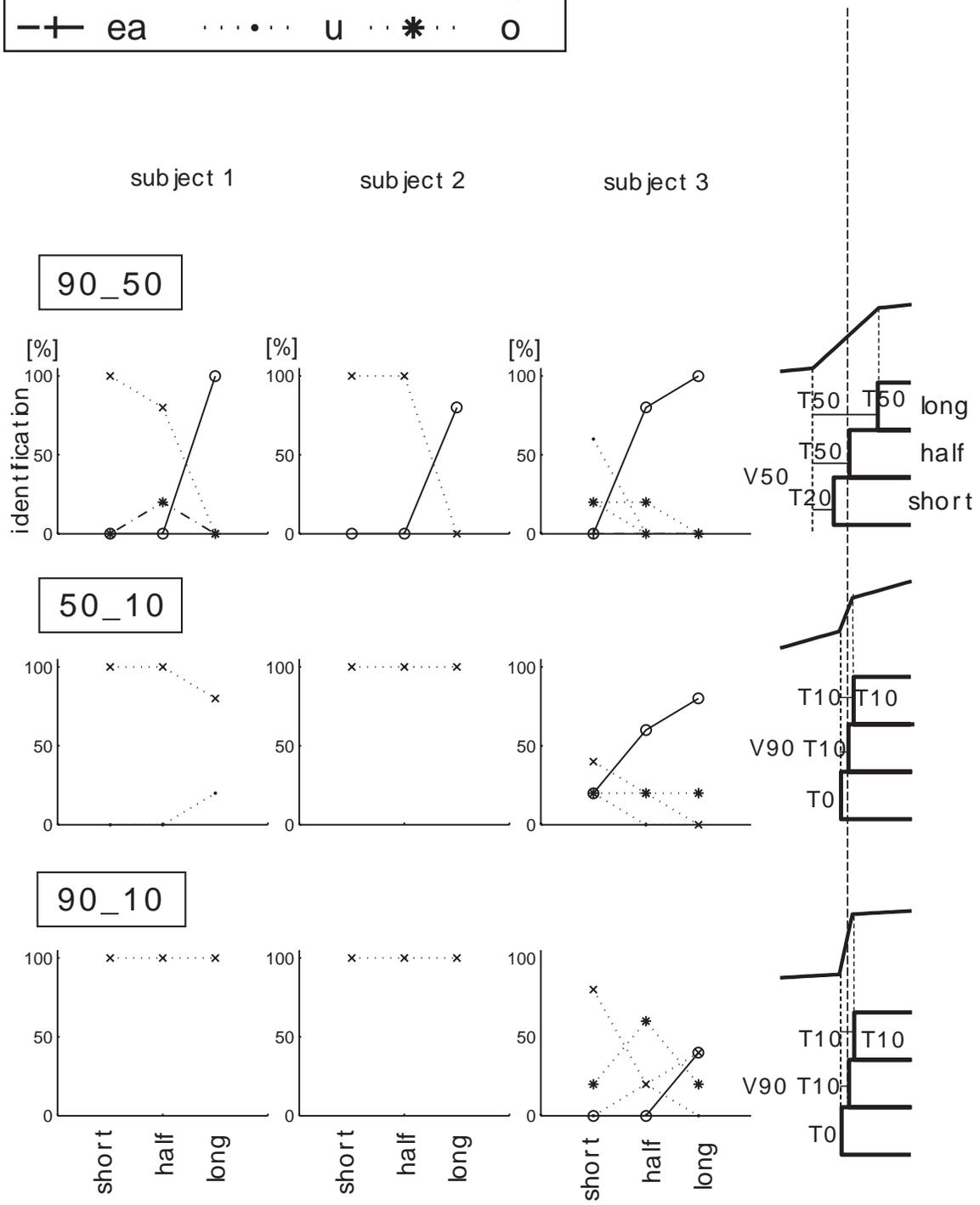
- わたり区間圧縮刺激と normal の比較。

被験者 1 の結果について、刺激 90-50 の half と normal の half に対する回答を比べると、90-50 に対して「x」が多くなっており、90-50 で母音が知覚されるためには、long でなければならない。さらにわたり区間の圧縮率が大きい刺激になると、全体的に「x」の回答になってしまう。被験者 2 も全体的に「x」になり、被験者 3 でも「x」が増している。



remained syllable segments

図 2.10: /iai/の結果その 1 (normal とわたり区間を伸張した場合の結果)



remained syllable segments

図 2.11: /iai/の結果その2 (わたり区間を圧縮した場合の結果)

## 2.5.2 /ioi/の結果

図 2.12 は時間伸縮していない normal と、わたり区間を伸張（母音区間を圧縮）した /ioi/ の刺激に対する回答結果であり、図 2.13 はわたり区間を圧縮（母音区間を伸張）した /ioi/ の刺激に対する回答結果である。

- わたり区間伸張刺激の結果について。

normal の回答と比べると、被験者 1、2 では /iai/ の結果と同様に、normal の half 付近から short の範囲で「x」の回答が増えているが、わたり区間伸張刺激では「x」は増えず、中心母音が知覚されている。被験者 2 の回答について normal の short と刺激 10-90 の short を比べると違いが明確にでている。

今回は、伸張率が一番大きい 10-90 になるにつれて、被験者全員で異聴が増えている。しかし被験者 2 の 10-90 の結果をみると、/iai/ の被験者 1、3 に見られた結果と同様の、わたり区間の伸張率が大きいとき、残る音節が短くなると元の母音の音韻が知覚されてくるという結果が見られる。

- 図 2.13 の結果について

被験者 1、2 に関しては、/iai/ の結果と同様、わたり区間の圧縮率が大きい刺激 90-10 に向かうほど、中心母音は知覚されなくなっている。

被験者 3 は被験者 1、2 とは異なる知覚結果となっており、かなりばらつきがある。刺激 90-10 では大きく伸張された母音区間しか残らないような刺激になっているが、それでも被験者 3 では中心母音の知覚が起こっている。被験者 3 は実験方法の解釈が多少異なっている可能性が考えられる。

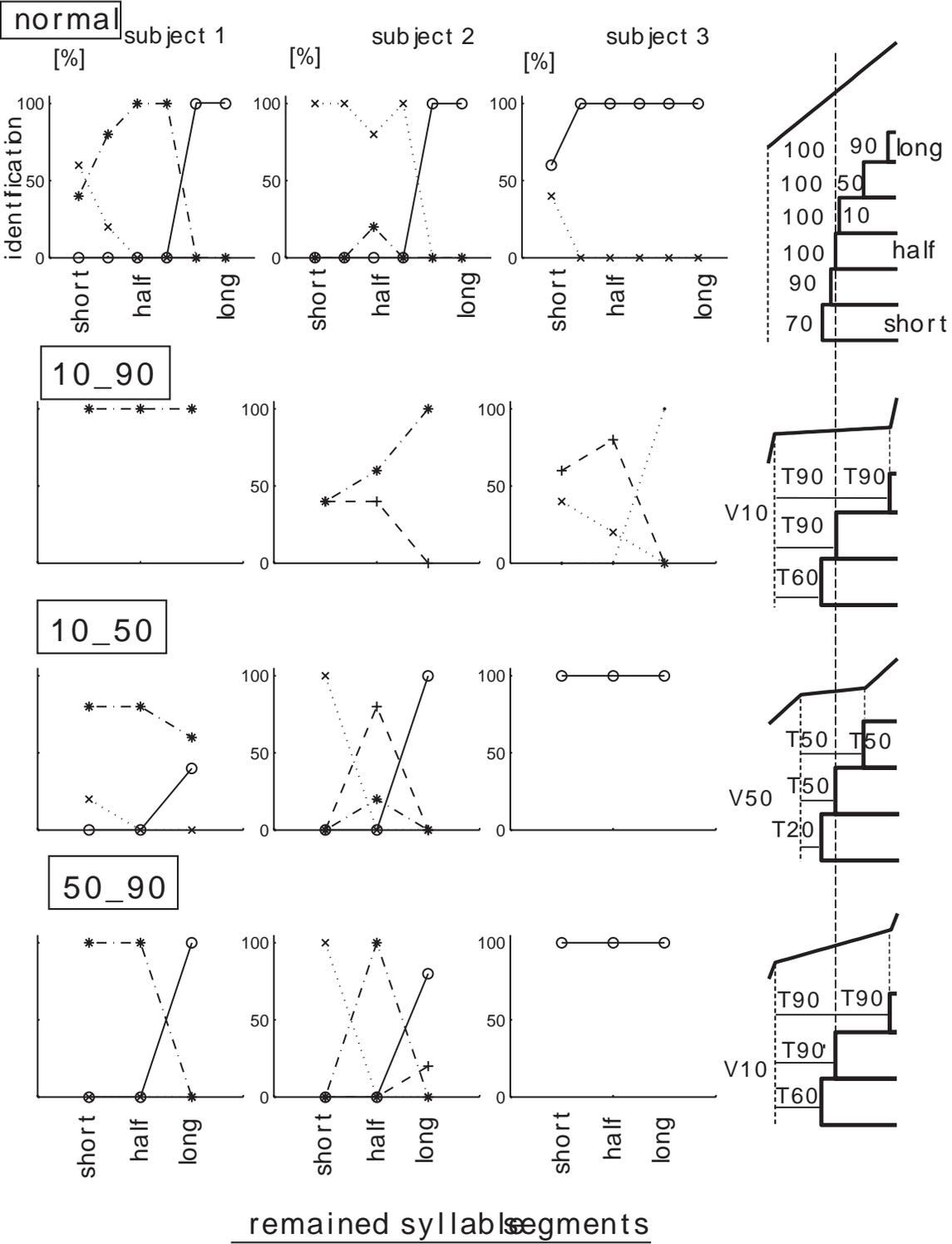


図 2.12: /ioi/の結果その 1 (normal とわたり区間を伸張した場合の結果)

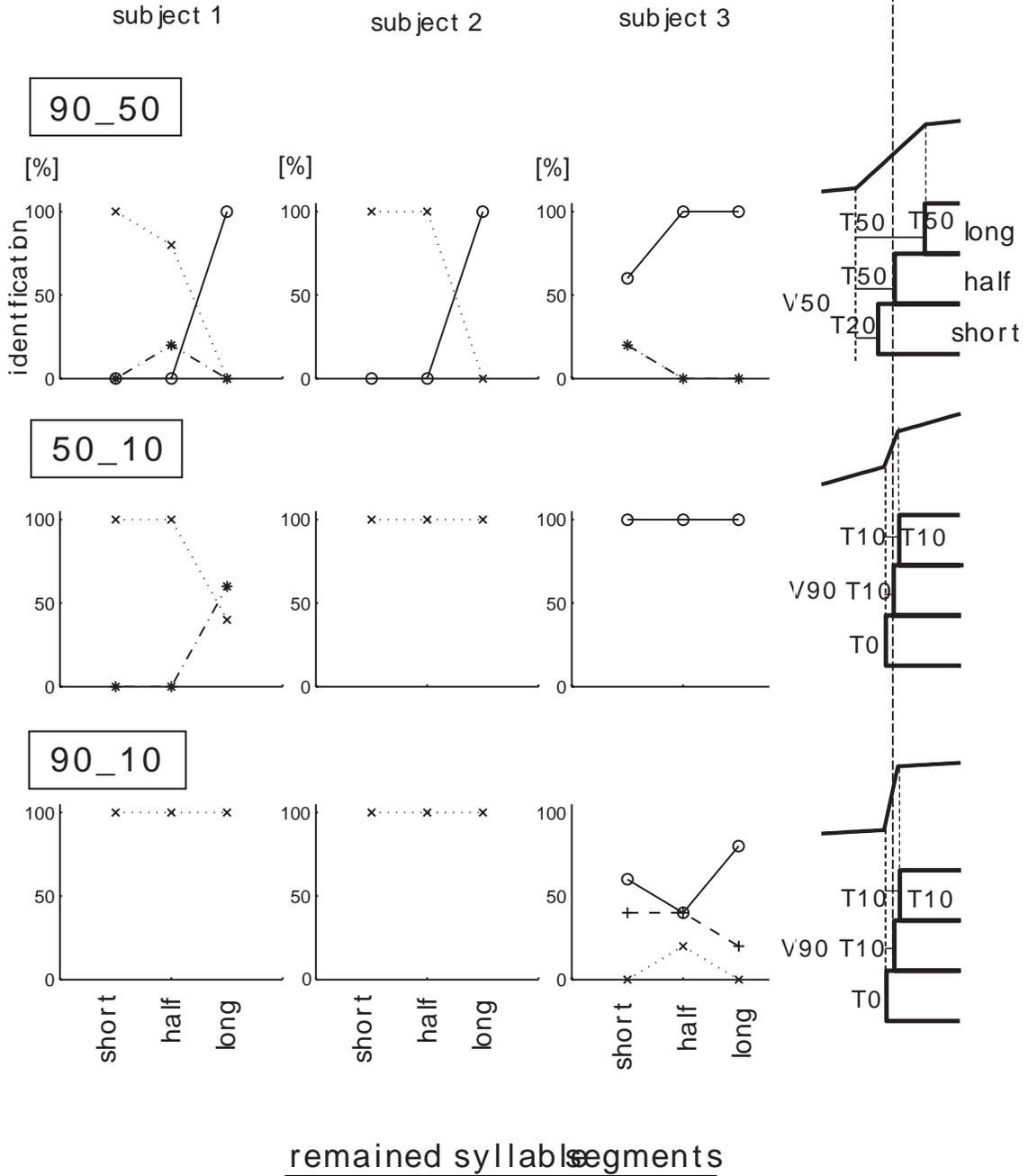


図 2.13: /ioi/の結果その2 (わたり区間を圧縮した場合の結果)

### 2.5.3 /aia/結果

図 2.14 は時間伸縮していない normal と、わたり区間を伸張（母音区間を圧縮）した /aia/ の刺激に対する回答結果であり、図 2.15 はわたり区間を圧縮（母音区間を伸張）した /aia/ の刺激に対する回答結果である。

- 図 2.14 の被験者 1、2 について

normal と 時間伸縮刺激と「x」の回答範囲の差は、今回も同じように見られる。

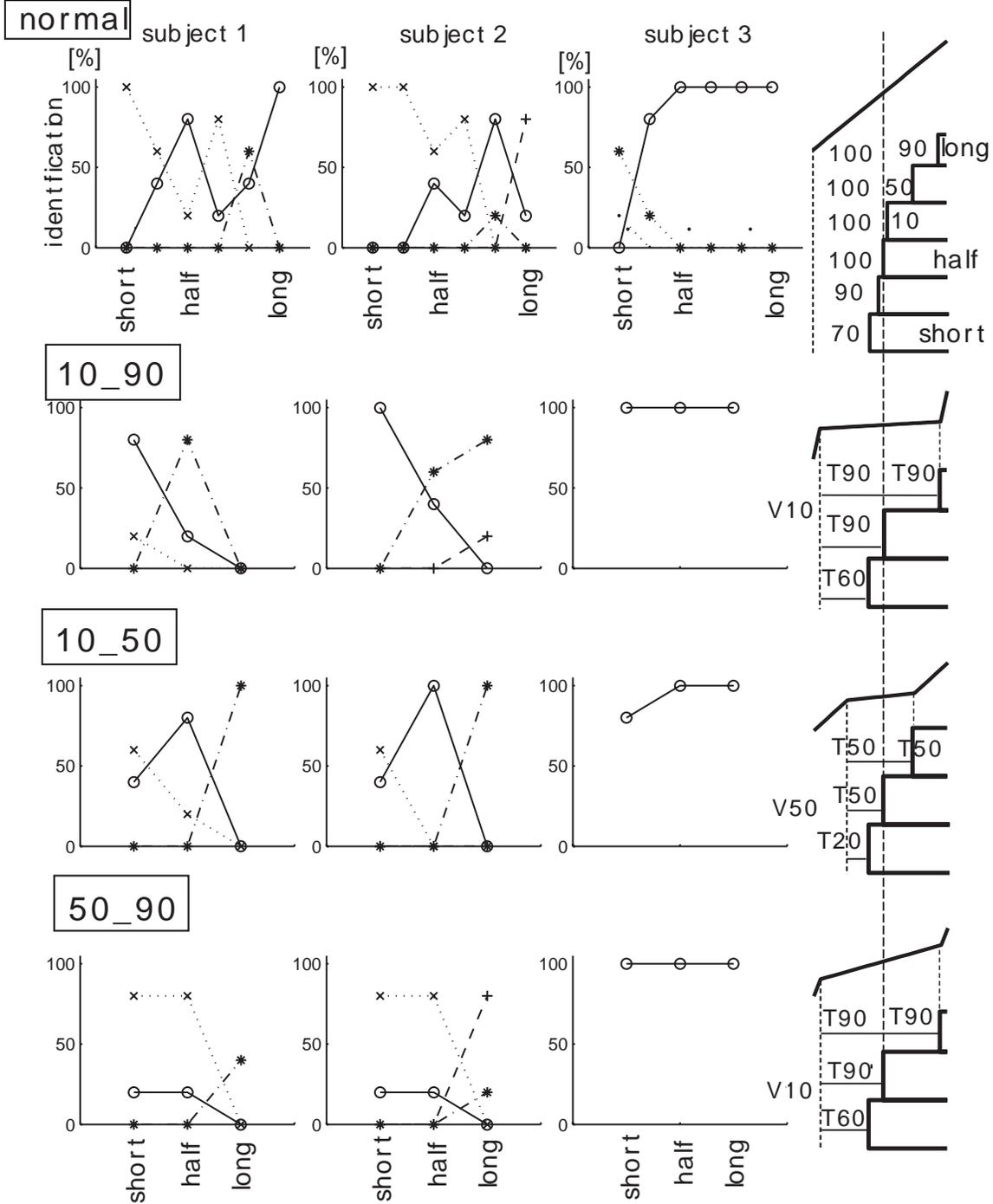
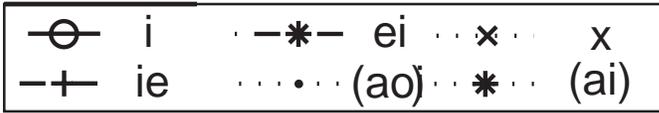
そして、わたり区間伸張率が大きく、さらに残る音節が短いほど、異聴が少なくなっている結果も、被験者 1、2 で明確に表れているようである。しかし、ここでは /ei/ または /ie/ から /i/ への中心母音の変化が起きている。これは完全な異聴 /e/ からの変化ではないため、/ioi/ における /u/ から /uo/ または /ou/ への変化や、/iai/ における /e/ から /a/ への変化の結果と、同じように見ることができないかはわからない。

雑音置換により /ei/ または /ie/ の /e/ が無くなり、/i/ が残ったとも考えられるので、伸縮率と音節長の関係に対する傾向を明確に示すためには、/e/ から /i/ への変化の確認が必要である。

- 図 2.15 の被験者 1、2 について

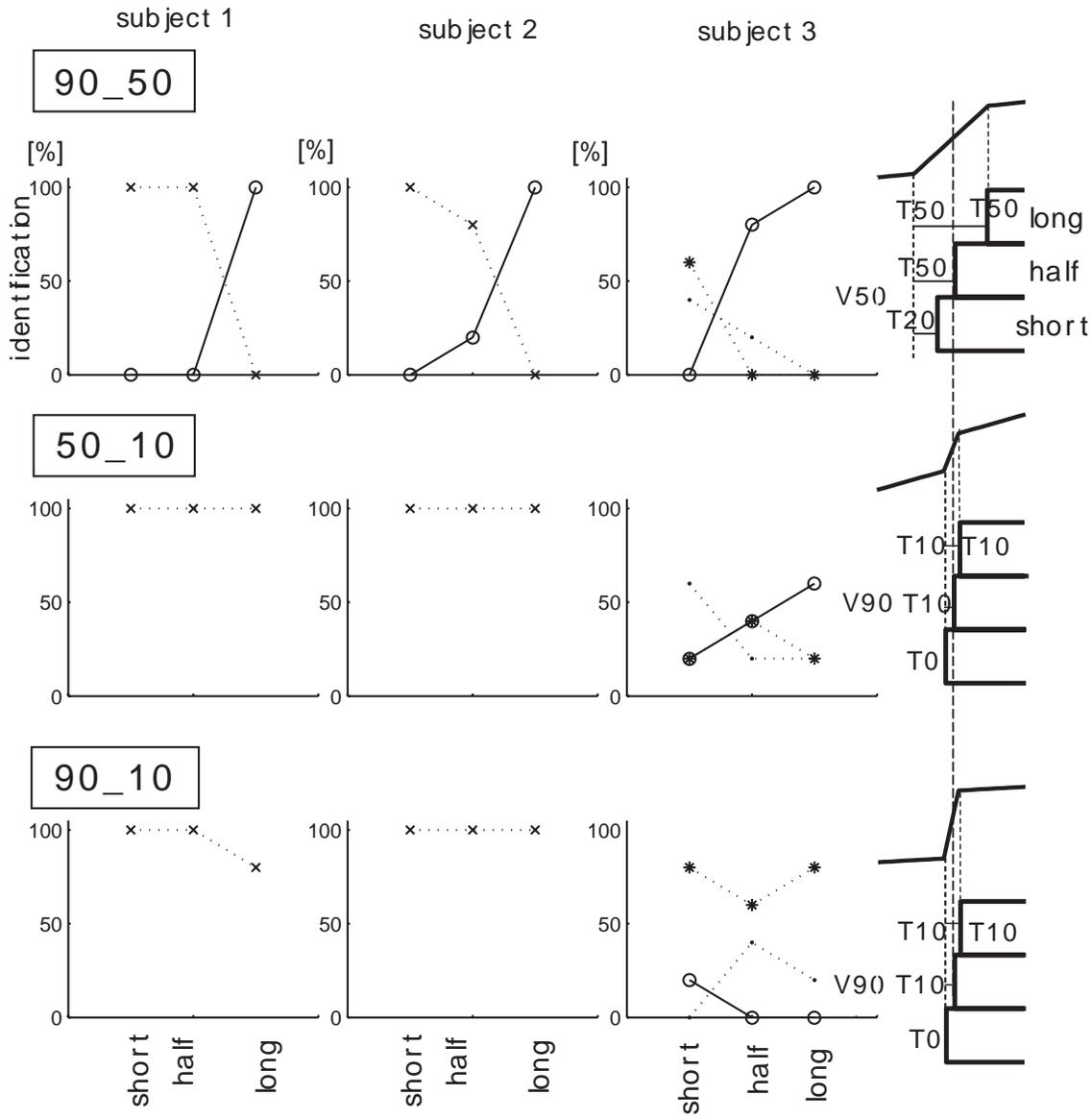
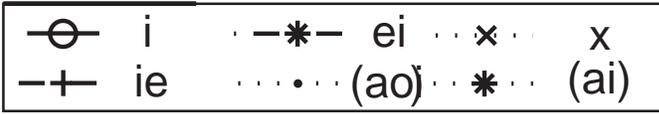
他の結果と同様に、わたり区間の圧縮率が大きくなる 90-10 に向かうほど、残る音節を長く残し（このとき同時にスペクトルの位置も中心母音に近づいている）ても、中心母音の知覚はされなくなっている。

- この 3 連続母音は第 1 母音にアクセントがあり、第 3 母音が第 1 母音に比べて弱いために、被験者全員にとって少し判断が難しいようであった。特に被験者 3 では /ao/、/ai/ の回答があるなど、対象 3 連続母音として聴こえないことが多く起こった。これでは、回答された中心母音を決定できないため、対象な 3 連続母音として聴くことができる刺激への改善が必要である。



remained syllable segments

図 2.14: /aia/の結果その1 (normal とわたり区間を伸張した場合の結果)



remained syllable segments

図 2.15: /aia/の結果その2 (わたり区間を圧縮した場合の結果)

## 2.6 実験 1、結果のまとめと考察

(1) half 位置で雑音置換した normal と同じく half 位置で雑音置換した時間伸縮刺激では、中心母音の回答の違いがみられた。half は各刺激の残る音節の時間長が等しくなり、置換雑音に隣接するまでにその音節が到達するスペクトルの位置が等しくなる。よって、この音節をわたり部として考えると、入わり部終点と出わり部始点の到達スペクトル位置が同じであっても、変化速度の違いで中心母音が知覚される・されないの違いが起こると見ることができる。

そこで、中心母音が知覚される刺激を再分析し、LSF、SFTR を求めた。すると、中心母音が知覚できる刺激の残された音節部分には、SFTR の極大点が含まれていることが特徴として見られた。

SFTR 極大点の重要性は Furui[10] によって示されている。Furui の実験では、自然音声 CV 音節を前後から徐々に切り出し、同定実験をした結果、スペクトル変化率の極大点が含まれる 10ms 区間が重要な情報を与えており、さらに約 50ms 区間があれば十分であることがわかっている。

これらのことから、2つの音素の連続を聴くためには、SFTR の極大点が必要であることがわかる。しかし、SFTR の極大点を含む場合でも、次の(2)、(3)のような知覚結果の違いがあった。

(2) わたり区間を圧縮して変化速度が急激になった (SFTR 極大値が大きくなった) 場合、中心母音の知覚には、さらに長い音節部分が残される必要があった。

- わたり区間の変化が急激であるため、その追従ができなかった影響と思われる。
- わたり区間の変化速度の違いによって、その用いられ方が異なると考えられる。

(3) わたり区間伸張刺激の場合、伸張後のわたり区間を伸張部、圧縮後の母音区間を圧縮部と呼ぶとする。SFTR の極大点は圧縮部に含まれている。

このとき、伸張部は母音間の特徴を持っているので、伸張部が長いときは異聴が起こる。しかし、伸張部が短くなる (終端のスペクトル位置が中心母音から離れていく) と、中心母音の異聴が少なくなり、母音が元の音韻で知覚されるようになった。特にそれは圧縮率が大きいときに見られた。

よって、その圧縮率の違いの影響に加えて次のことが考えられる。

- 伸張部が長いとき、その伸張部の影響が大きく、伸張部が短いとき、圧縮部の影響が大きくなっているのではないか。
- 伸張部の長さの違いにより、圧縮部の影響が異なるだろう。

(2),(3) の以上のような知覚傾向が見られたが、この傾向についてさらに詳しく調べるべきポイントに次の2点が挙げられる。

- わたり区間を圧縮したときの、必要な非置換部の長さの違い。
- わたり区間を伸張したときの、その非置換部に含まれる伸張部の長さ、圧縮部の圧縮率の違いによる影響。

ここで、第2点目は知覚される母音の音韻に影響するものであるなので、異聴が起こらないのはどのような時間構造のときであるのかを詳しく調べることができる。よって、次の実験では第2点目について詳しく調べていく。

# 第3章 実験2、時間圧縮部の圧縮率と時間伸張部の時間長の影響

## 3.1 実験目的

実験1の考察では、わたり区間伸張刺激で見られた傾向である、わたり区間伸張率と残る音節の時間長の関係を、母音区間圧縮率と伸張後のわたり区間の時間長の関係として置き換えた。またこのとき、圧縮された母音区間は圧縮部、伸張されたわたり区間を伸張部とすることにする。

実験2では、この圧縮部の圧縮率と伸張部の長さの2つにどのような関係があるとき、中心母音の知覚に変化があるのかを調べるのが目的である。

## 3.2 刺激の構成

- 圧縮部の圧縮率と伸張部の時間長を変える刺激を作成した。
- 分析合成音の時間伸縮を行うための一連の作業は実験1と同様である。
- 時間伸縮、雑音置換の条件以外も実験1と同じである。

### 3.2.1 時間伸縮の仕方

今回は、時間伸縮区間を2種類の Type A と Type B に決定し、その伸縮を行っている。その時間伸縮のしかたは、図3.1で説明していく。

実験 1 の刺激作成法と同様に、母音中心点、わたり中心点を決定し、区間 1、2、3、4 を決定している。

Type A の刺激伸縮は図 3.1 Type A の

- (a) をそのまま。
  - (b) を (1)1 倍長、(2)0.75 倍長、(3)0.5 倍長、(4)0.25 倍長。
  - (c) を 5 倍長。
  - (d) を 5 倍長。
- に伸縮する。

Type B の刺激伸縮は図 3.1 Type B の

- (a) をそのまま。
  - (b) を (1)1 倍長、(2)0.75 倍長、(3)0.5 倍長、(4)0.25 倍長。
  - (c) を 10 倍長。
- に伸縮する。

### 3.3 雑音置換方法

置換雑音は一定長として固定した。その前後に伸張部の長さを切り詰めて変化させた刺激を与えて、3 連続母音の音声刺激としている。

図 3.2 は、Type A と Type B の圧縮率 (4)0.25 倍の例を示す。刺激の全体長が長いものから、その伸張部の長さは 100 %、70 %、30 %となっている。

- 今回の刺激では、伸縮部の 100 % から 0 %まで 10 %刻で取っている。
- 雑音長は (区間 2 + 区間 3) の長さである。実験 1 の傾向で、異聴から元の母音と同じ音韻が知覚されるようになったときの雑音長に等しい値 (half を与えていた雑音長) を用いた。

### 3.4 実験方法

- 実験 1 との違いは、被験者である。今回の被験者 1、被験者 3 は実験 1 と同じであるが、被験者 2 が実験 1 とは異なる。全員、大学院生の男性である。
- 2 Type \* 4 圧縮率 \* 11 伸張長 \* 3 音節 = 264 刺激を用いた。

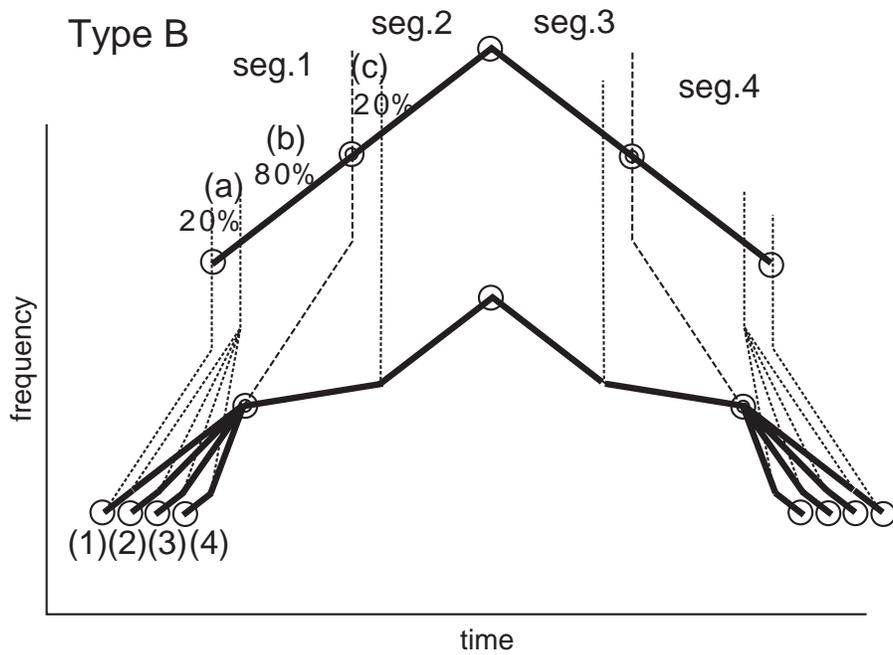
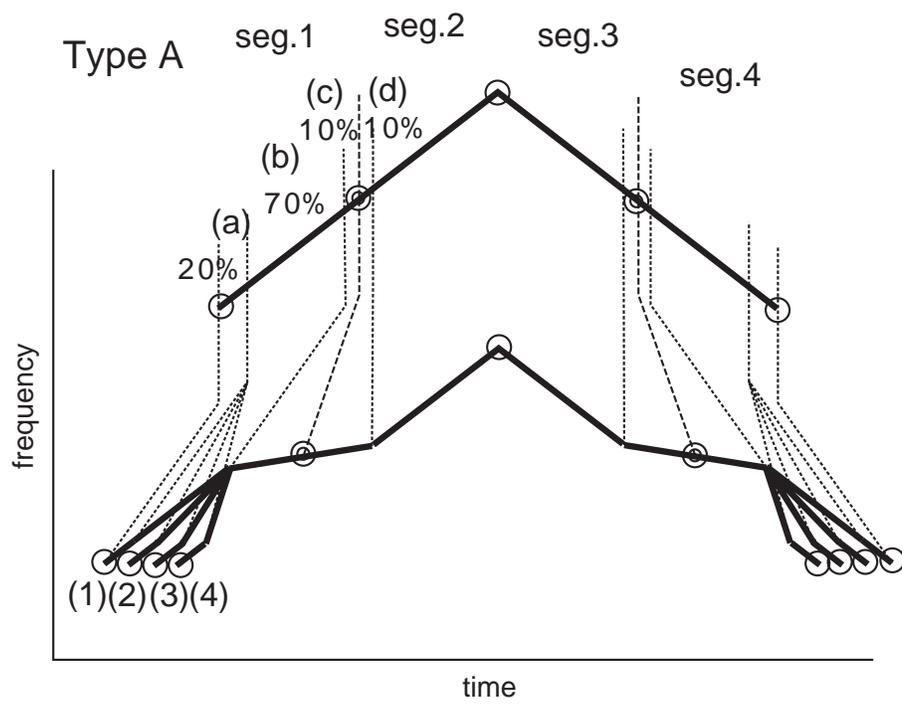


図 3.1: 時間伸縮方法

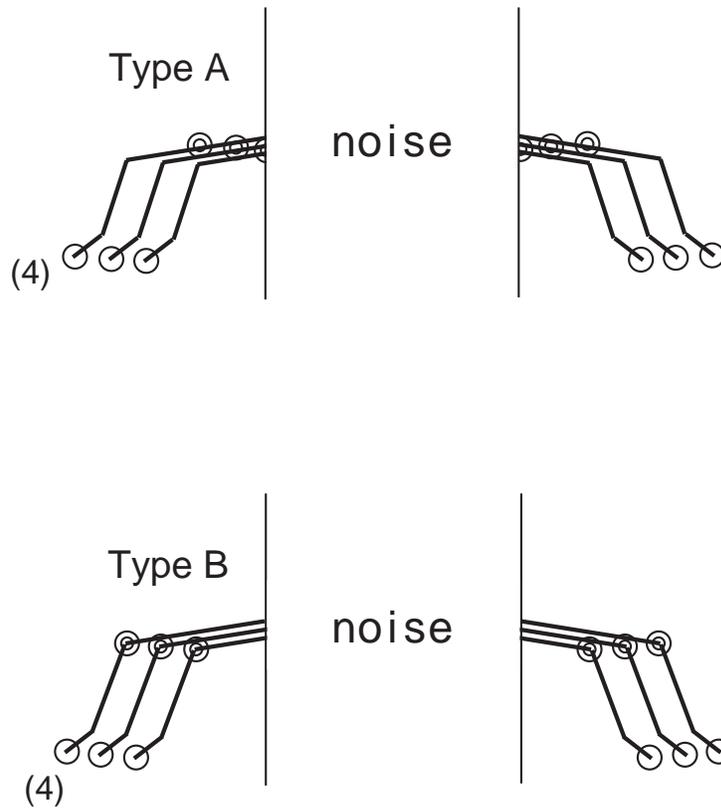


図 3.2: 雑音置換方法

- 1 試行分を 2 回に分けて、それぞれ 8 回行った。

その他は、実験 1 と同様。

### 3.5 実験結果

- 結果の集計方法は実験 1 と同じである。
- 図の見方。

各図は被験者 1 名分の、音節 1 種類に対する回答結果を示す。図の中の各グラフの左にある R は圧縮率を示す。各グラフの横軸は伸張部の長さを示し、縦軸に母音の回答率を示す。左列のグラフが Type A、右列が Type B である。Type A、の伸張部の長さ 50 % から 100 % までと、Type B の伸張部の長さ 0 % から 50 % までは伸張部の雑音に隣接した終端でのスペクトルが一致している。

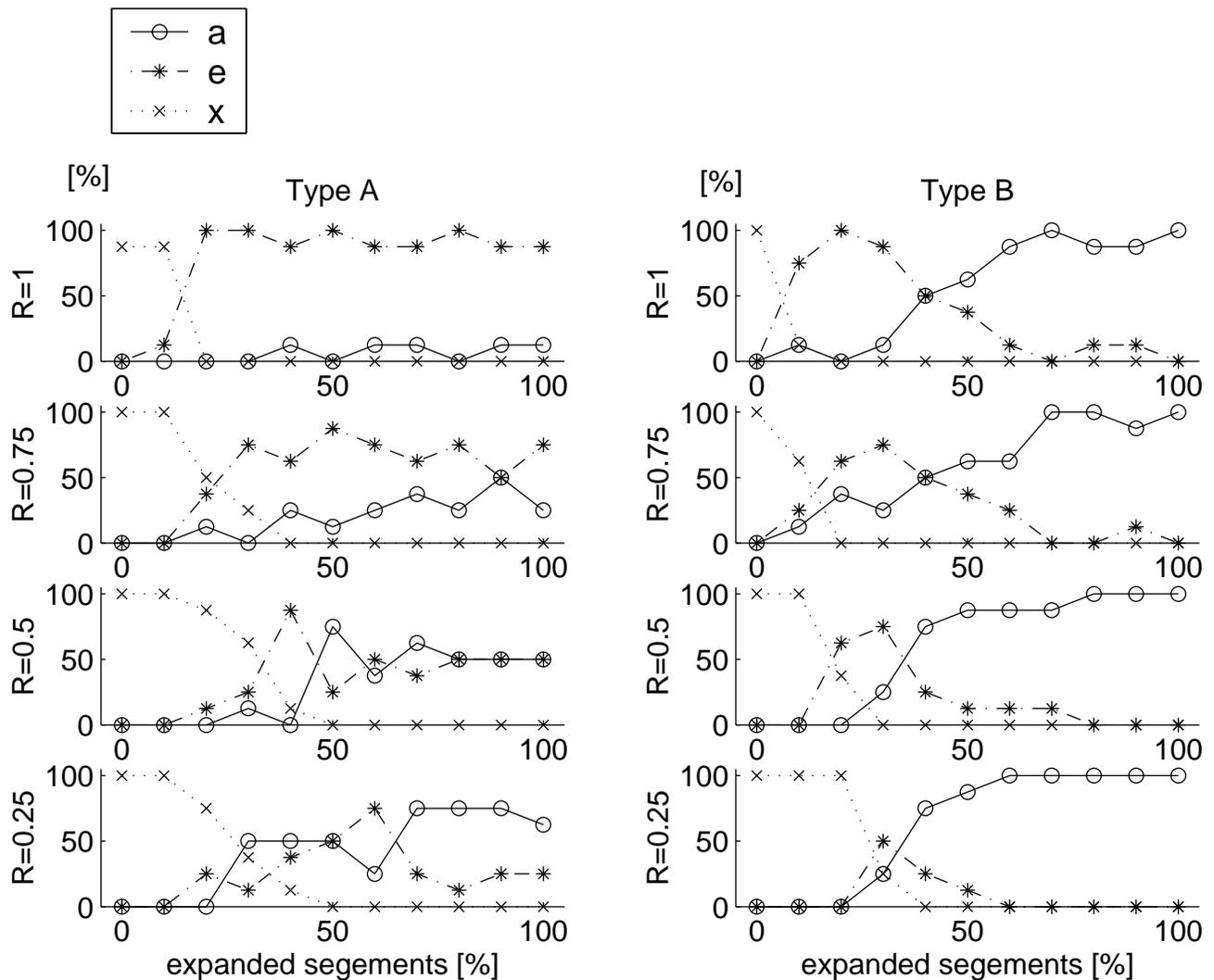


図 3.3: /iai/の結果 (被験者 1)

圧縮率が大きくなると、/a/の回答率が大きくなっている。/a/の回答は特に伸張部が長いほうの刺激でみられており、圧縮率が大きくなると、伸張部がすこし短くなっても/a/が知覚できるようになっている。

さらに Type A と Type B の伸張部の到達点が一致する点で、回答を比べてみると、R = 0.25 の Type A の 70 % では/a/であるが、R = 0.25 の Type B の 20 % では「x」となっている。

よって、到達点が一致していても、そこまでの時間変化の仕方により回答が異なっていることが再び確認される。

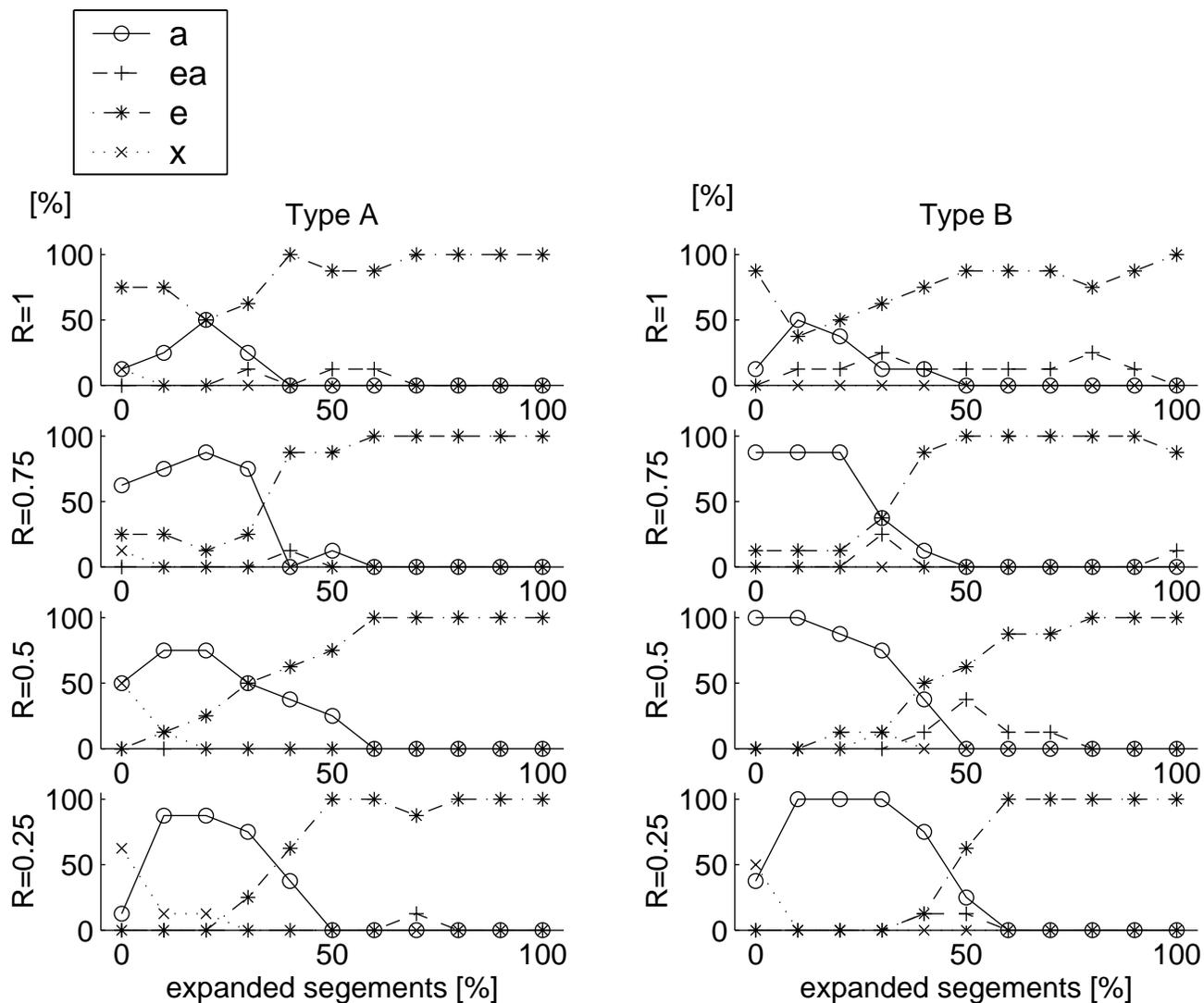


図 3.4: /iai/の結果。(被験者 2)

被験者 1 では、伸張部が長くなると元の母音の音韻が知覚されていたが、被験者 2 では、伸張部が短くなると元の母音の音韻が知覚されており、逆の結果に見える。しかし、圧縮率が大きくなるにしたがって、元の母音の音韻が知覚できる刺激の範囲が広がっていることは共通している。

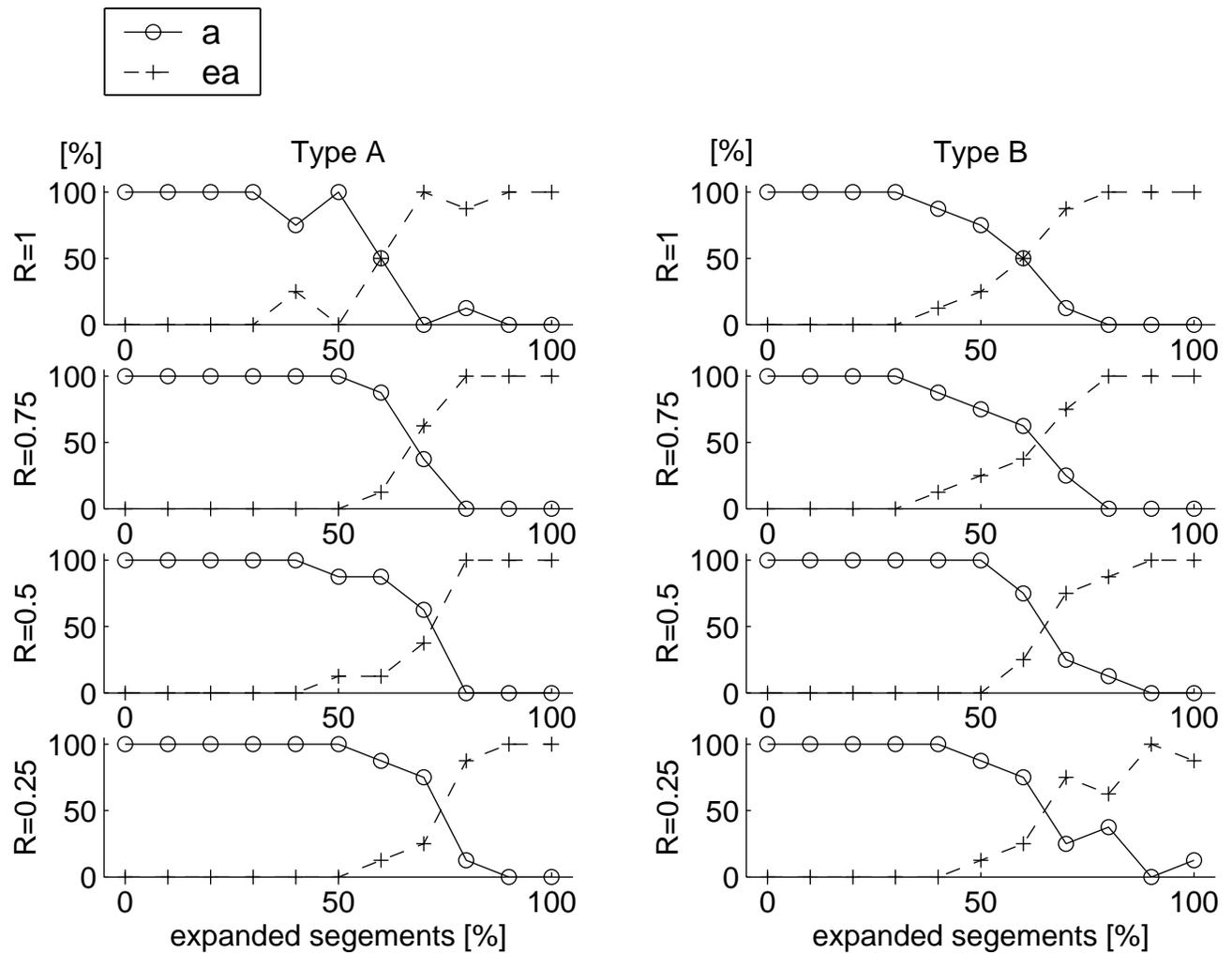


図 3.5: /iai/の結果 (被験者 3)

被験者 3 の結果では、圧縮率が大きくなっても、Type B では特に/a/と/ea/ の境界があまり動かないように見える。しかし、Type B の 80 % で各回答率を見ると、圧縮率が大きくなるにつれて、元の母音の音韻と一致する回答が増えており、Type A, Type B 両方に対して、被験者 2 と同じ結果の傾向が見られる。

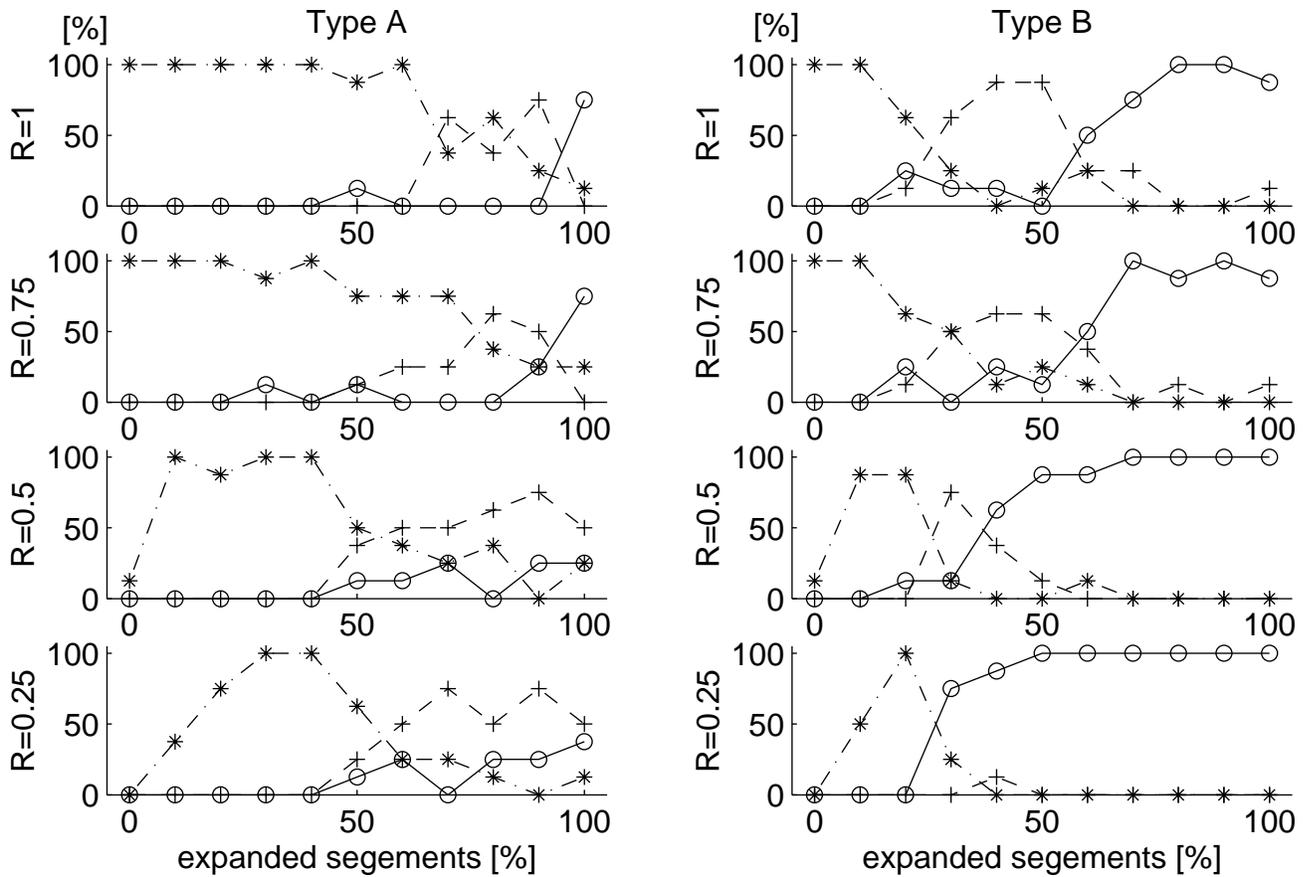
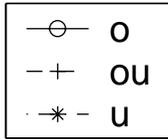


図 3.6: /ioi/の結果 (被験者 1)

/iai/のときの結果と同様に、伸張部が長く残ったときにもとの母音の音韻が知覚されている。特に TypeB については、圧縮率が大きくなるにしたがって、/o/が知覚できる刺激の回答範囲が広がっている。すなわち、伸張部の長さが長くなくても、/o/が知覚されるようになる。

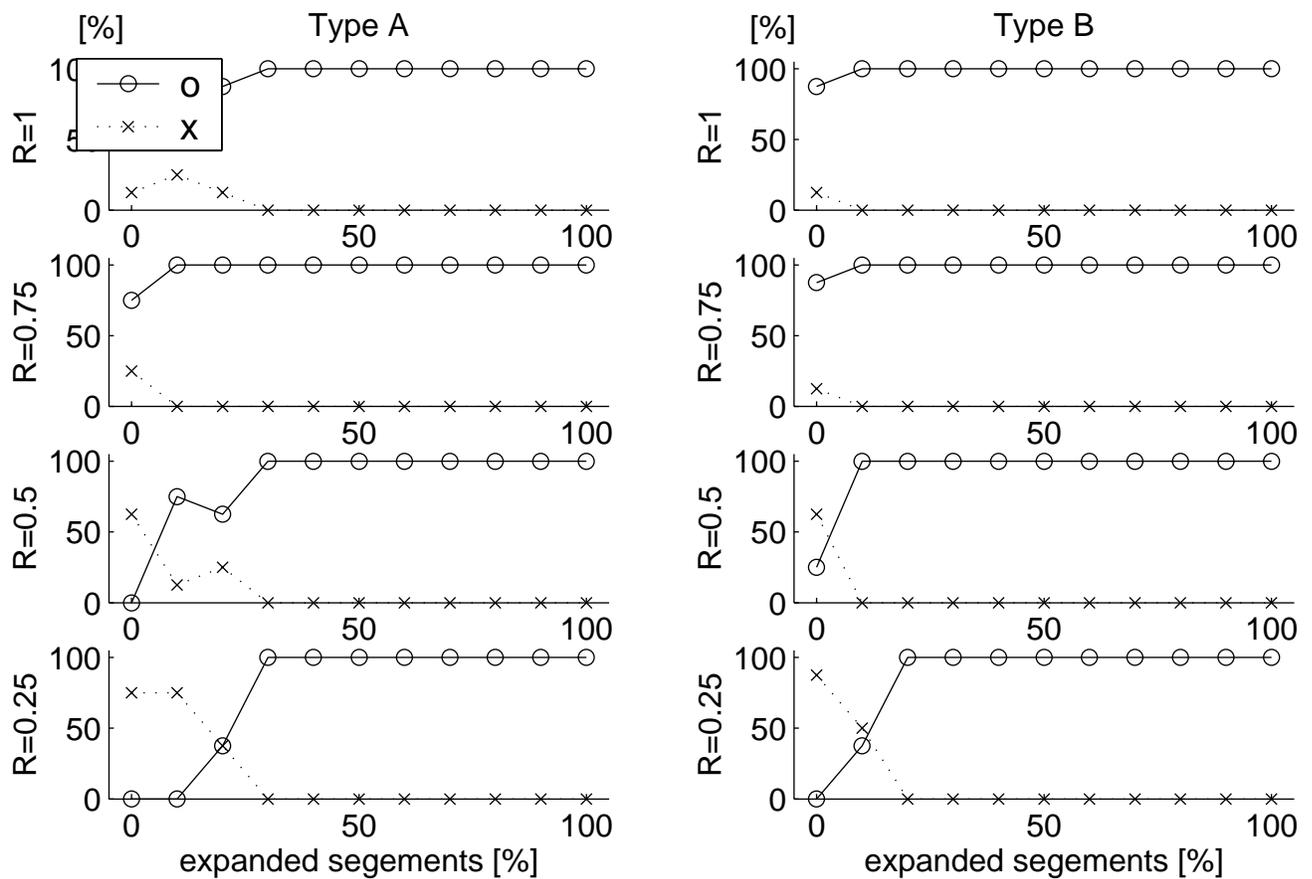


図 3.7: /ioi/の結果 (被験者 2)

この結果では、異聴が全く見られていない。圧縮率の影響もあまり見られない。この被験者の母音判別の音韻境界の差が影響しているかもしれない。

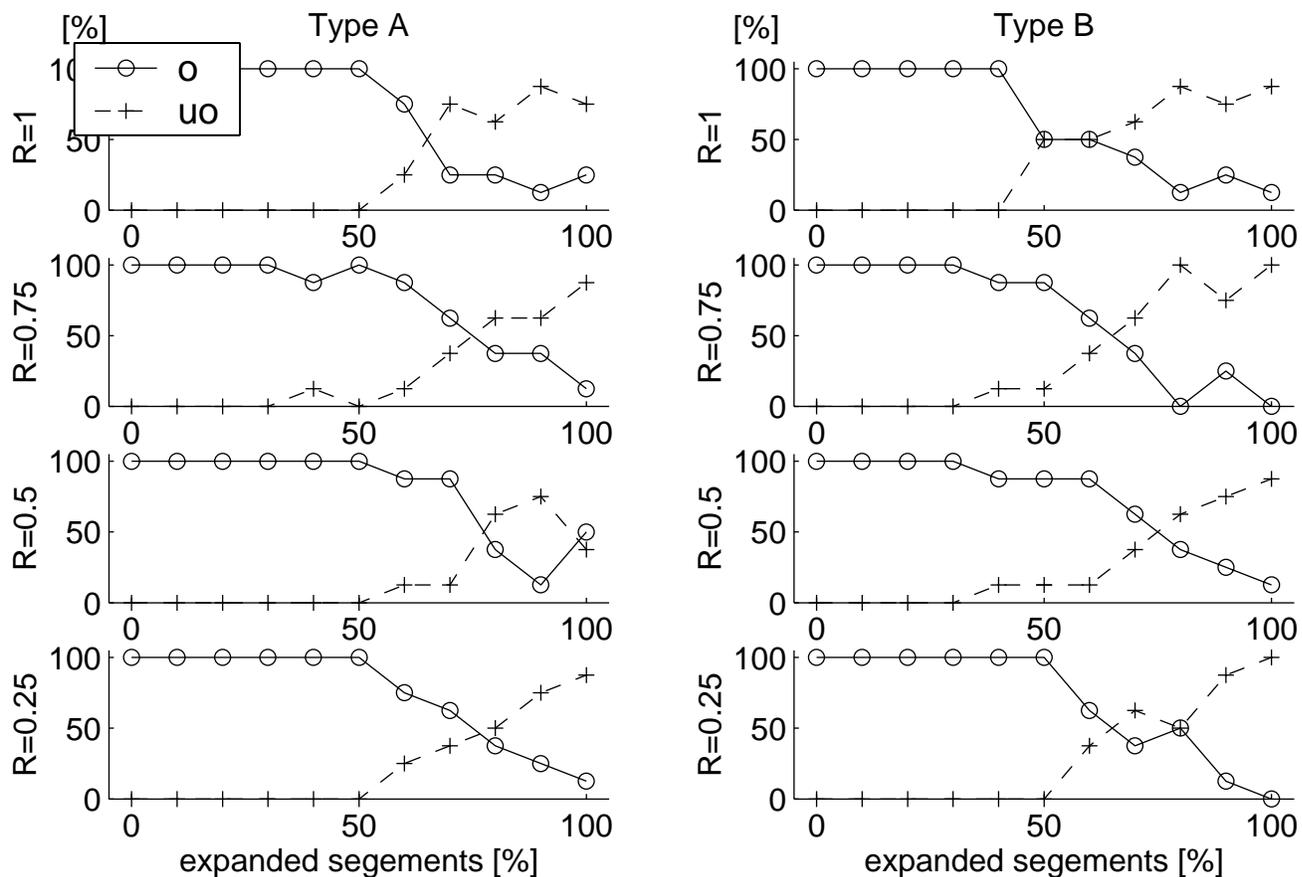


図 3.8: /ioi/の結果 (被験者 3)

伸張部が短いとき、元の母音の音韻が知覚されている。圧縮率が大きくなるとその範囲は広がっている。

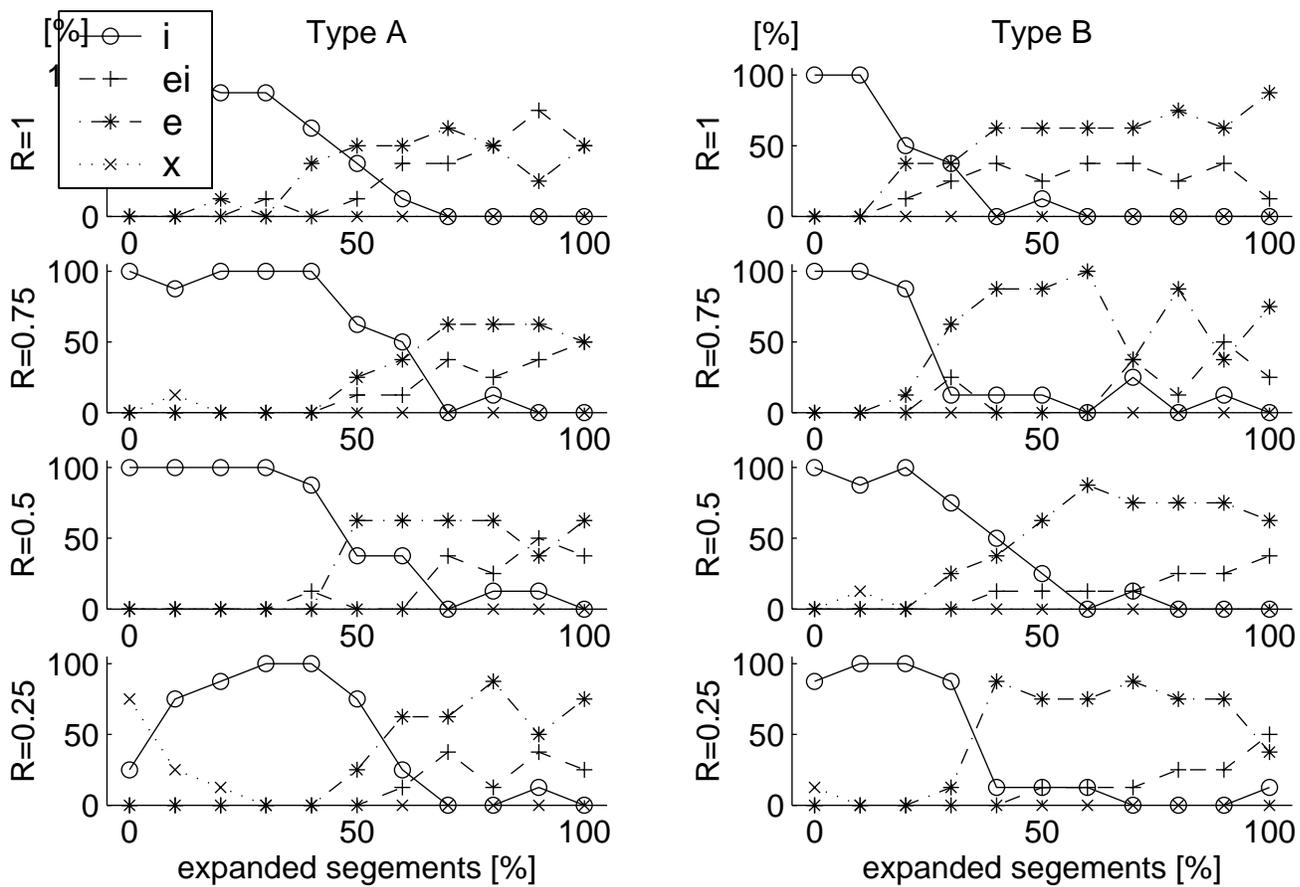


図 3.9: /aia/の結果 (被験者 1)

実験 1 では見られなかった /e/ という回答が含まれている。被験者 1 のこれまでの結果と同じ傾向が見られる。

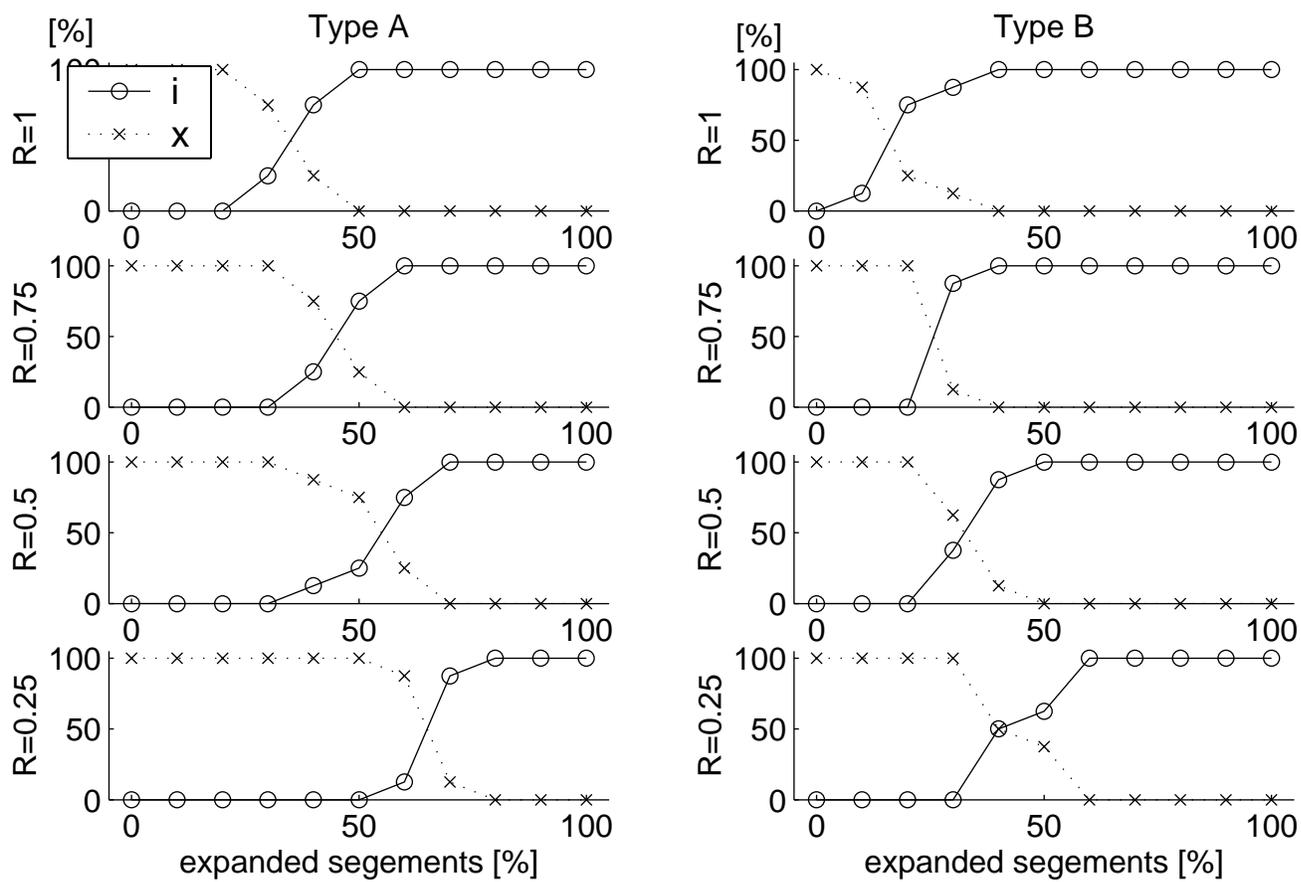


図 3.10: /aia/の結果 (被験者 2)

/ioi/の結果と同様、異聴は起こっていない。しかし、圧縮率が大きくなるにしたがって、伸張部の短いほうで知覚されていた「x」の範囲が広がってくる。音韻は違うが、圧縮率と伸張部の長さの影響が見られる。被験者 2 は母音の音韻境界が他の 2 人と異なるかもしれない。

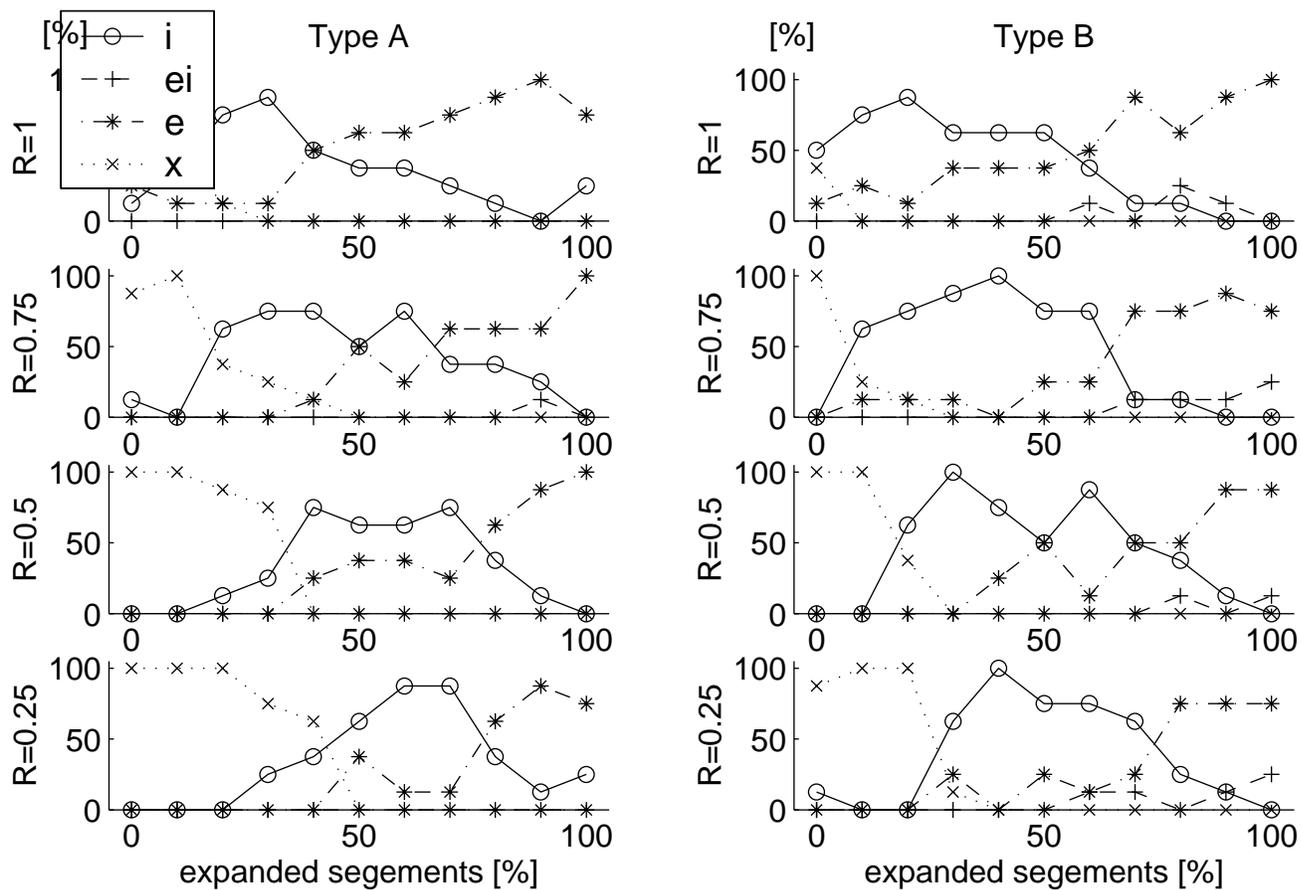


図 3.11: /aia/の結果 (被験者 3)

被験者 3 でも、実験 1 では見られなかった /e/ の回答が見られる。圧縮率が大きくなるにしたがって、/i/ が知覚される刺激の範囲はひろがっている。すなわち、ある程度伸張部の長さののこっても、/i/ が知覚されるようになる。

### 3.6 実験 2、結果のまとめと考察

被験者全員で見られたことは、

- 圧縮部の圧縮率を大きくすると、元の音韻の母音が知覚される伸張部の長さの範囲は広がった。

しかし、元の音韻の母音が知覚されるような、伸張部のその範囲は被験者によって異なった。

また、全く異聴が起こらない場合も見られた。この原因の 1 つとして、母音の音韻境界の個人差の影響が考えられる。

実験 2 の考察には、被験者ごとの母音の音韻境界を調べ、また、3 連続母音の音節に含まれるスペクトルはどのような音韻を表すのかを調べ、その結果と比較する必要がある。

# 第4章 実験3、単独母音の音韻境界の調査

## 4.1 実験目的

実験2で確認された被験者ごとの結果はどのように異なるのかを調べ、実験2の考察を与えるため、実験2に用いた3連続母音中のスペクトル断面の表す音韻(単独母音)について、その音韻境界を調べる。

## 4.2 刺激の作成方法

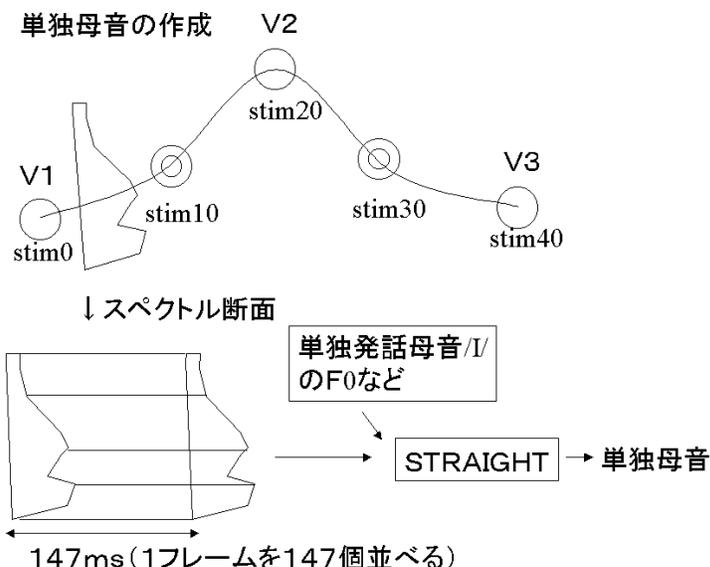
実験1、2で用いた3種類の3連続母音のスペクトルから単独母音を作成する。その作成方法を以下に示す。

- ATR 音声データベースから用いた3連続母音の話者 *mtm* の単独母音を STRAIGHT により分析する。その中で、得られた時間スペクトルからパワーの時間包絡を計算したものと、得られた基本周波数を用いる。
- 3連続母音の1フレーム分のスペクトルを、先に分析した母音の時間長分ならべて、単母音の時間スペクトルを作成する。そのとき、先に計算してあるパワーの時間包絡の平均が、等しくなるようにする。
- 3連続母音から取り出して作成した時間スペクトルと単独母音の基本周波数と、非周期成分のパラメータを用いて、再び STRAIGHT で合成する。最後に全刺激で平均パワーを等しくして、単独母音の音声刺激としている。

### 4.3 刺激の構成

今回、基本周波数などのパラメータを用いた単独母音は/i/を用いた。5母音で刺激を構成し、聴いた結果、一番良く聴こえたものであったことと、3連続母音3種類に共通して含まれている母音であるためである。

3連続母音から用いたスペクトルは、図4.3のように区間1、2、3、4の各区間を10%刻みで取り、全部で41個分を用いている。



### 4.4 実験方法

実験環境は、実験1、2と同じである。被験者も実験2と同じである。

教示としては、母音の回答を求めた。ただし、曖昧な音素についての回答の自由度を与えた。例えば、/a//e/が混じっているものに聴こえた場合は「a e」、または/a//e/が混じっているがどちらかというとき/a/が強く聴こえるという回答は「a > e」、全く弁別できない場合は「x」「？」などで回答してもらった。

全刺激 = 41母音 \* 3音節 = 123個を8回聴いてもらった。

### 4.5 実験3、結果と考察

- 結果の集計

「a > e」というような少しでも「a」が優位な回答であるときは、その回答を「a」として集計した。

- 図の見方

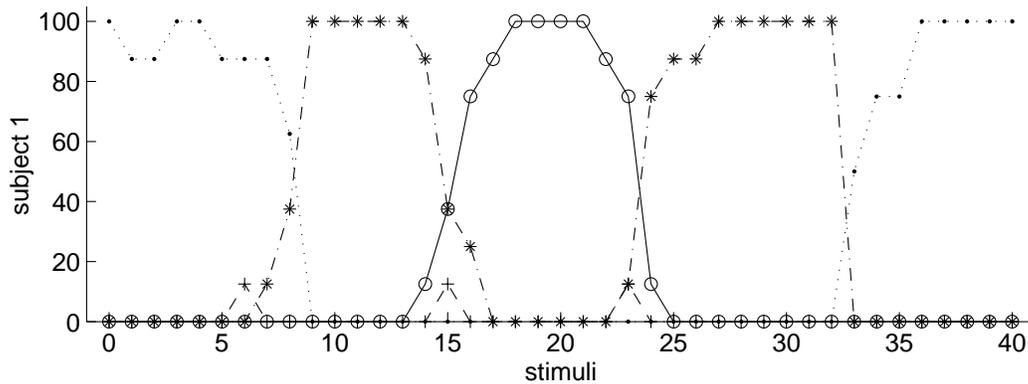
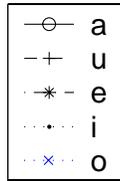
図の各グラフの縦軸は回答率であり、横軸は刺激の番号で、0 から 40 までである。各刺激の番号と、3 連続母音での基準点との対応を以下に示す。

刺激 0	第 1 母音の中心点
刺激 10	第 1、2 母音間のわたり部の中心点
刺激 20	第 2 母音の中心点
刺激 30	第 2、3 母音間のわたり部の中心点
刺激 40	第 3 母音の中心点

表 4.1: 単独母音音声刺激の番号と、そのスペクトルの 3 連続母音での位置

実験 2 で用いた伸張部の、雑音に隣接した到達点は Type A で単独母音の刺激 9 から刺激 110 間、または刺激 29 から刺激 31 の間にあり、Type B では単独母音の刺激 10 から刺激 12 の間、または刺激 28 から刺激 30 の間に位置する。

その区間（単独母音の刺激 9 から 12 の間と、刺激 28 から 31 の間）における単独母音の結果をみると以下のようなになる。



- 実験2の被験者2の/i oi//a ia/の結果で全く異聴が起こらなかったことは、音韻境界のずれのためであることがわかった。また、/a ia/は異聴は起こってはいないが、圧縮率と伸張部の長さの影響が見られる結果であることがわかった。
- もとの母音の音韻を知覚できるための伸張部の長さは、被験者によって、長い、短いの違いが見られたが、その被験者間での音韻境界の差は大きくないことから、その原因は音韻境界のずれではないことがわかる。
- 以上のことから、圧縮率が大きいほど、異なる領域の音韻の知覚が多くなっていくことは、全員に共通していることがわかり、圧縮部の圧縮率と伸張部の長さの違いにより、伸張部の到達するスペクトルとは離れた領域の音韻が知覚できていることがわかった。

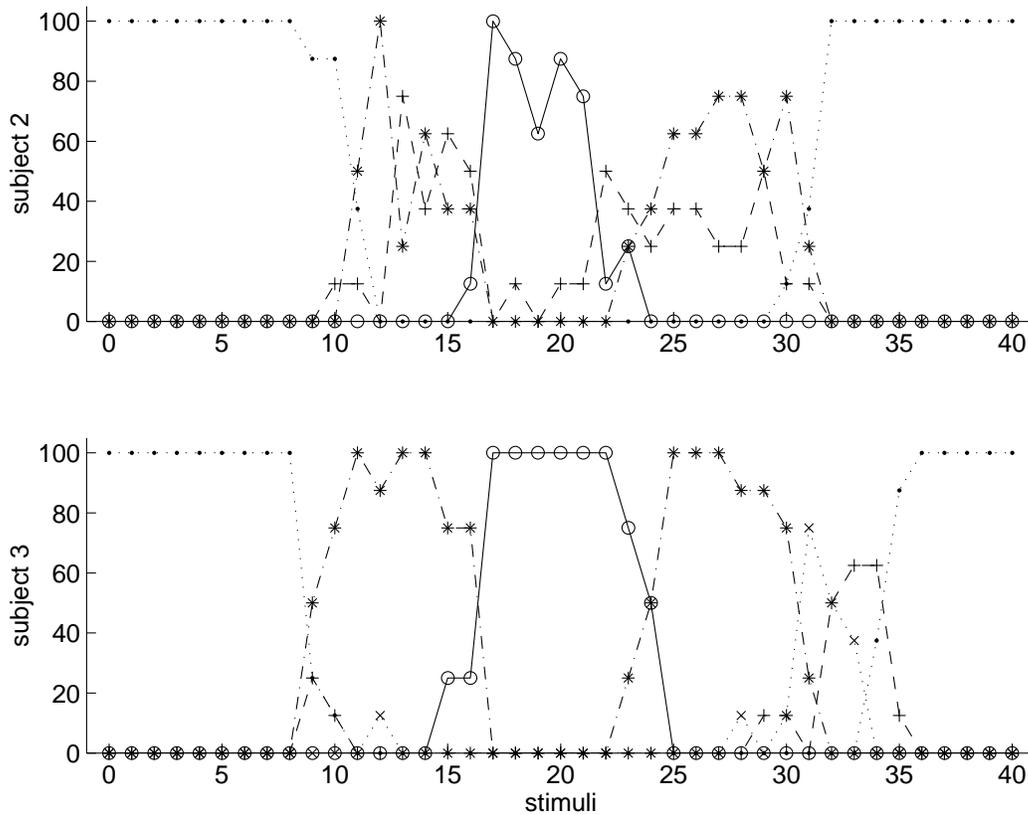
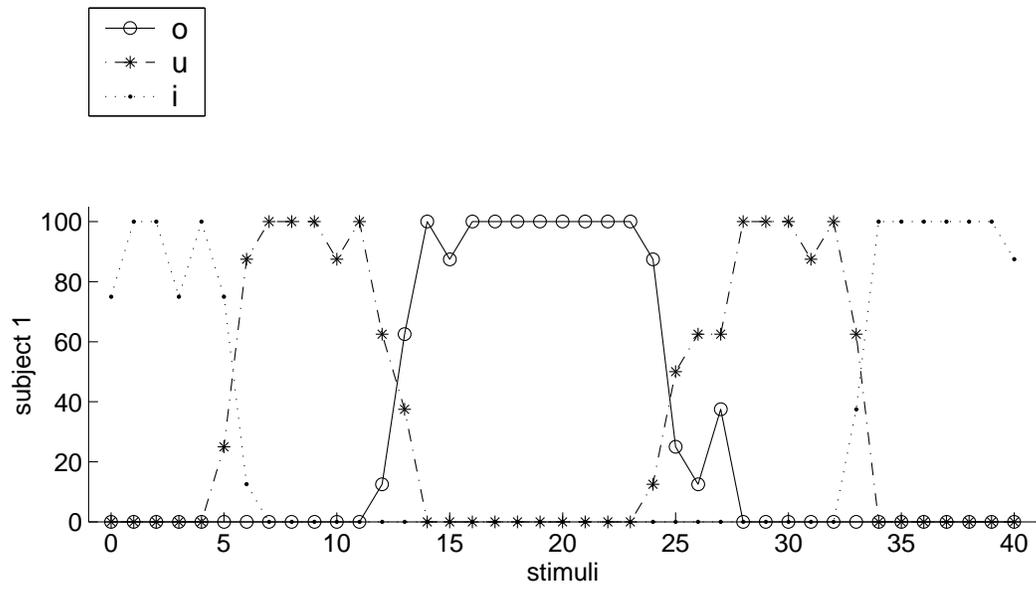


図 4.1: /iai/のスペクトルを用いた単独母音の結果

刺激 9 から 12、刺激 28 から 31 までの結果をみると、被験者 1、3 は /e/ の範囲にあり、被験者 2 は /e/ からその境界の範囲にある。この区間における被験者間での音韻境界のずれはこのデータの見方だと、さほどないと思われる。よって、実験 2 /iai/ の被験者 2 の回答には、母音の音韻境界の差の影響は含まれていないと見ることができる。

その他に、被験者 3 の /a/ 安定な範囲は他の被験者よりも広いことは、/iai/ の結果で /a/ が多く知覚されていることと、関係があるかもしれない。



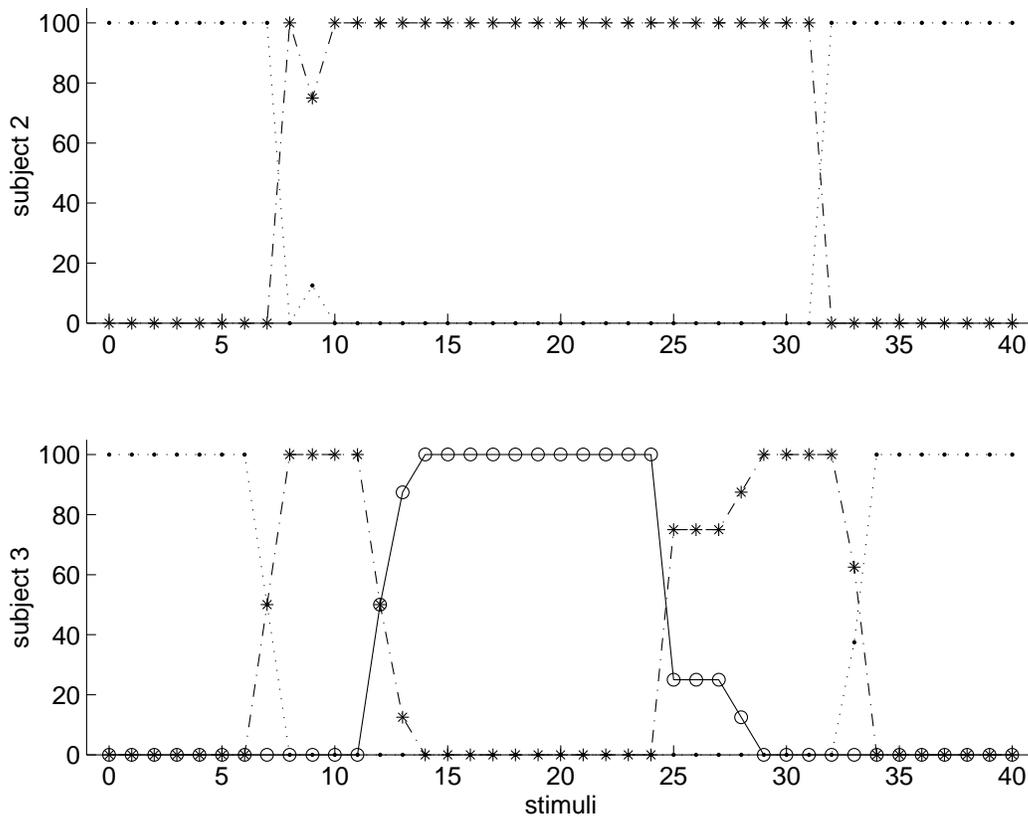
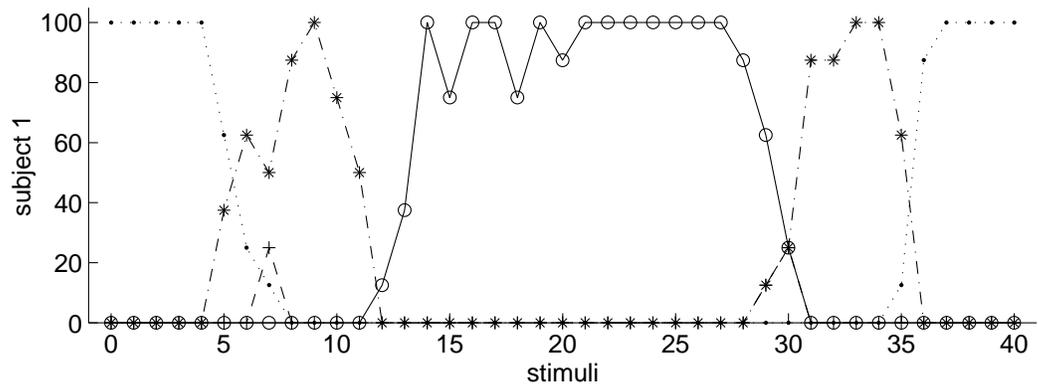
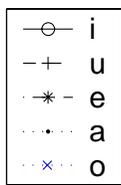


図 4.2: /ioi/のスペクトルを用いた単独母音の結果

刺激9から12、刺激28から31までの結果は、被験者1、2、3とも/u/の範囲にある。ただし、被験者2はそのほかの範囲で/o/が全く知覚されず、他の2名の被験者とは音韻境界が多少ずれていると思われる。/ioi/の結果で被験者2の回答母音数が少なかったことはこの音韻境界の差の影響かもしれない。しかし、被験者2の/ioi/の結果では/u/が全く知覚されず、ほとんど/o/のみである。圧縮部、伸張部により明らかに別の範囲の母音を知覚していることがよくわかる。

他には、被験者1、3は音韻境界がこのデータの見方では良く似ている。にもかかわらず、被験者1、3の/o/が知覚される範囲は伸張部が長い方と短い方と、逆の傾向にみられる。しかし、圧縮率の影響があることは共通している。



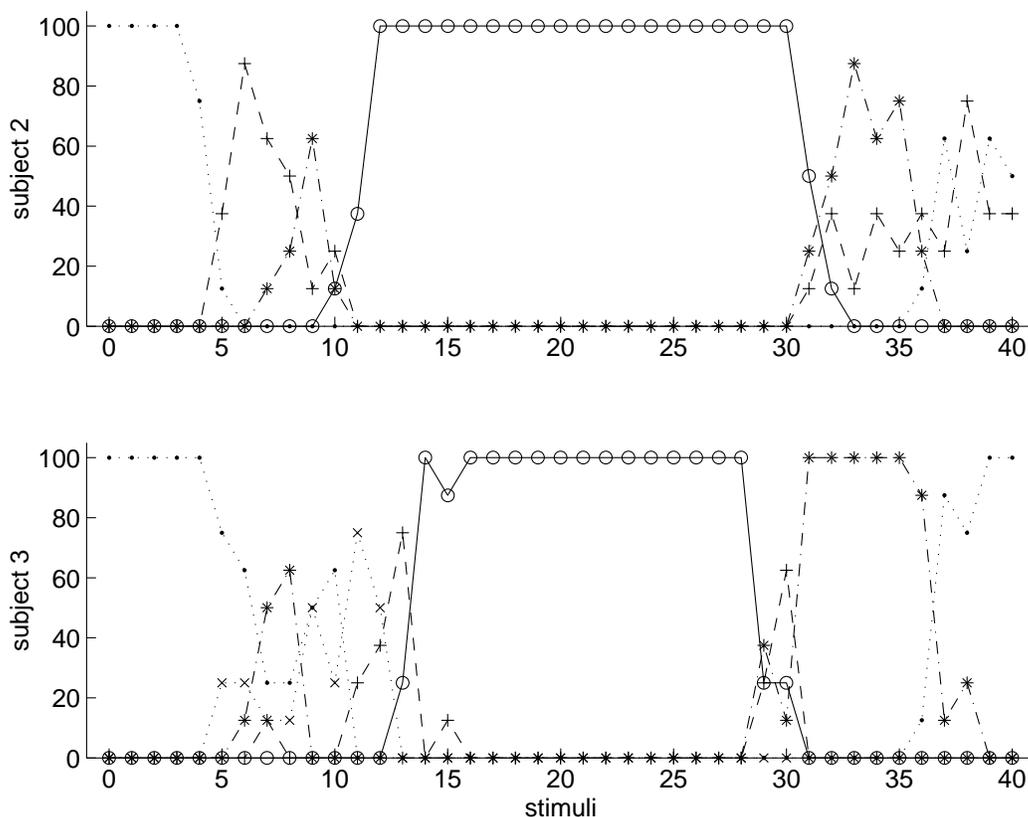


図 4.3: /aia/のスペクトルを用いた単独母音の結果

刺激9から12、の結果は被験者1、2、3とも/e/の範囲である。刺激28から31までの結果は、被験者1は/i/から/e/にあり、被験者2ではほとんど/i/の範囲、被験者3は境界にある。

/aia/で、被験者2が回答した母音は/i/のみで他は「x」だったことが、すでに、/i/の領域に入っていたからであることがわかる。よって、「x」は/i/の範囲を越えてしまっているための結果だとみることができる。このとき「x」の知覚できる刺激の範囲の傾向は被験者3と同じで、圧縮率が大きくなるにつれて、その音韻を知覚できる刺激の範囲が広がっている。すなわち、ある程度長くなっても、その音韻が知覚されるようになる。

# 第5章 実験結果のまとめと関数近似による考察

## 5.1 実験 1、2、3 のまとめ

実験 1 では、母音の中心点とわたりの中心点を決定し、その点を固定して全体長を一定として、その中で刺激の時間構造を時間伸縮と雑音置換により様々に変化させた。その刺激に対する実験結果から、以下の傾向が見られた。

(1) 図 5.1 の normal は、スペクトル変化率の極大点をちょうど含むか含まないので中心母音が知覚されない場合が増える。3 連続母音の中心母音を知覚できるためには、非雑音置換部にスペクトル変化率の極大点の存在が必要である。

(2) 図 5.1 のわたり区間を圧縮した刺激のように、スペクトル変化率の極大点が存在する非雑音置換部が与えられても、その変化速度があまりにも急激であると、中心母音は知覚できない。

(3) 図 5.1 のわたり区間を伸張した刺激のように、スペクトル変化率極大点を含み中心母音は知覚できる。長い伸張部があるときは、その部分のもつ音響的特徴が大きく影響し、異聴となるが、伸張部を置換雑音により短くしていくと、またもとの母音の音韻が知覚されるようになった。

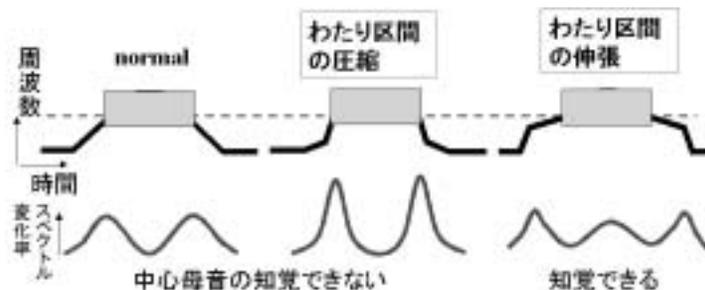


図 5.1: 実験 1 まとめ

実験 2 では、実験 1 で見られた傾向の (3) を確かめるためにも、圧縮部の圧縮率と伸張

部の長さが中心母音へ与える影響を調べた。

(4) 被験者によって、伸張部の長さが長いとき、または短いときに元の母音の音韻が知覚されるという違いがあった。しかし、圧縮率が大きくなると、元の母音の音韻が知覚される率とその刺激の範囲が大きくなった。

しかし実験2には、異聴が全く起こらないという結果もあり、このような知覚結果の差は、個人での音韻境界の差が影響しているのではないかと考え、実験3に単独母音の音韻境界を求め、実験2との比較を行った。

(6) 知覚結果の違いは、全く異聴が起こらない結果については、音韻境界のずれの影響があった。しかし、元の母音の音韻が知覚される伸張部の長さの違いは、音韻境界のずれによるものではなかった。

(7) 元の母音の音韻の知覚に用いられる伸張部の長さの範囲は、被験者によって異なったが、圧縮率の影響は共通して見られ、圧縮部の圧縮率と伸張部の長さの違いにより、伸張部の到達するスペクトルとは離れた領域の音韻が知覚できていることがわかった。

## 5.2 関数近似の目的

実験2で得られた結果から、圧縮率と伸張部の長さにより、伸張部の到達点より変化の方向に離れた領域の音韻を知覚できていた。このとき、圧縮部と伸張部の追跡機構と予測機構が働いていることが確認できたならば、補正機構の存在が明らかになるだろう。

今回は、簡単な2次関数近似を行ったもので、圧縮部と伸張部のフォルマント遷移の追跡機構と予測機構を表し、それを実験結果とを比較することで、圧縮部と伸張部の追跡機構と予測機構による補正機構の存在の可能性を調べる。

## 5.3 音声分析方法

- フォルマント遷移の分析は、Speech Toolsのフォルマント遷移の分析プログラムを用いて行った。これはLPC解析の方法を基礎として、フォルマント追跡をおこなう。以下表5.1に分析条件を示す。
- F1からF3までのフォルマント遷移について、その周波数軸をERBrateに直してから、2次関数の近似を行った。

LPC 次数	20 次
ハニング窓	30 [ms]
フレームシフト	1 [ms]
音声刺激	20 kHz sampling

表 5.1: 分析条件

## 5.4 関数近似の結果

今回の関数近似は、Type B の刺激について、その圧縮部と伸張部に対し、2 次関数を最小二乗法を用いて近似した。

図 5.2 は /iai/ の Type B、 $R = 0.5$  に対する被験者 2 の結果である。図 5.3 は /iai/ の Type B、 $R = 0.5$  の刺激の関数近似の結果と、被験者 2 の単独母音 /a/ と /e/ の境界を示す。F 1 に近似した 2 次関数の頂点について注目すると、関数の頂点が単独母音の音韻境界を大きく越え、/a/ の領域に十分到達したところで、実験 2 の回答も /a/ に変化している。

また、図 5.4 は /ioi/ の Type B、 $R = 0.5$  に対する被験者 1 の結果である。図 5.5 は /ioi/ の Type B、 $R = 0.5$  の刺激に関数近似をしたものと、被験者 1 の /o/, /u/ の音韻境界を示す。

F 2 に近似した 2 次関数の頂点に注目すると、F 2 に近似した関数の頂点が /o/ の領域にあるときに実験 2 での回答も /o/ になっている。

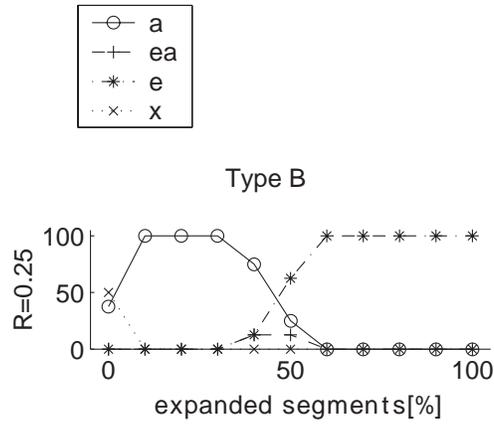
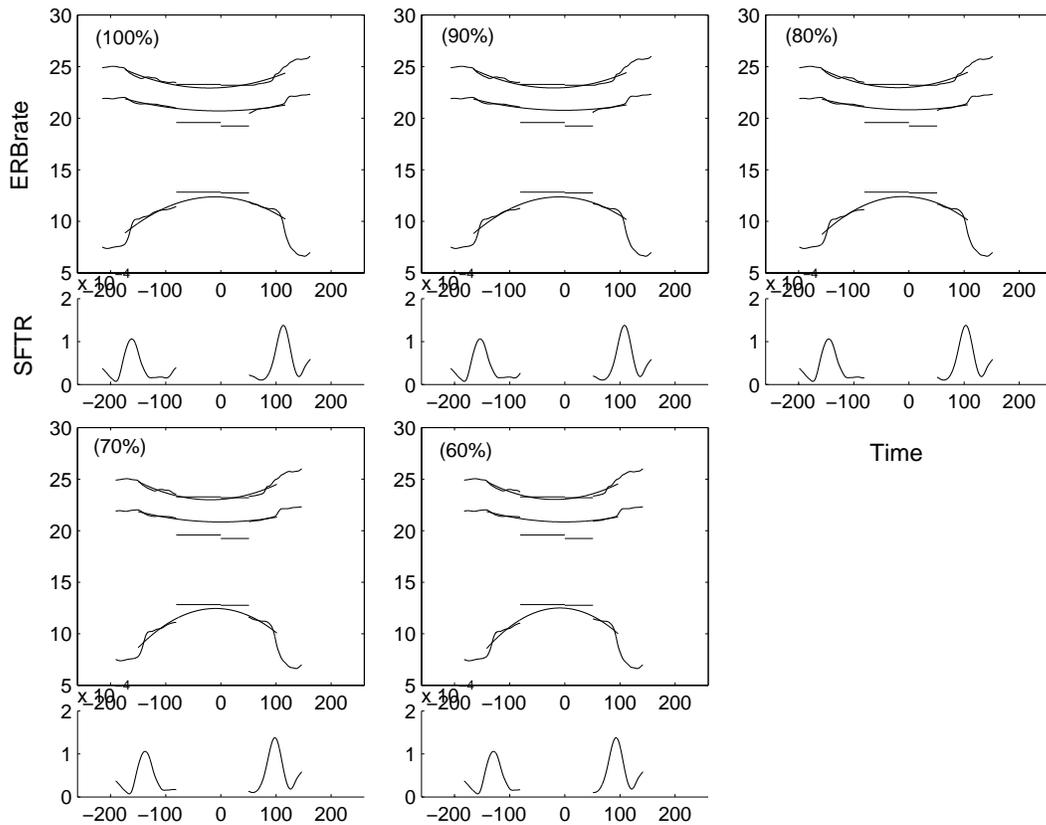


図 5.2: /iai/の R=0.25 の回答 (被験者 2)



左上から右へ順に、図 5.2 の横軸 100 から 60 までに対応する。横軸の 0ms のポイントは第 2 母音中心点である。母音の音韻境界は 0ms のポイントよりマイナス方向の時間に実験 4 /iai/ の刺激 1 7、プラス方向に刺激 2 1。

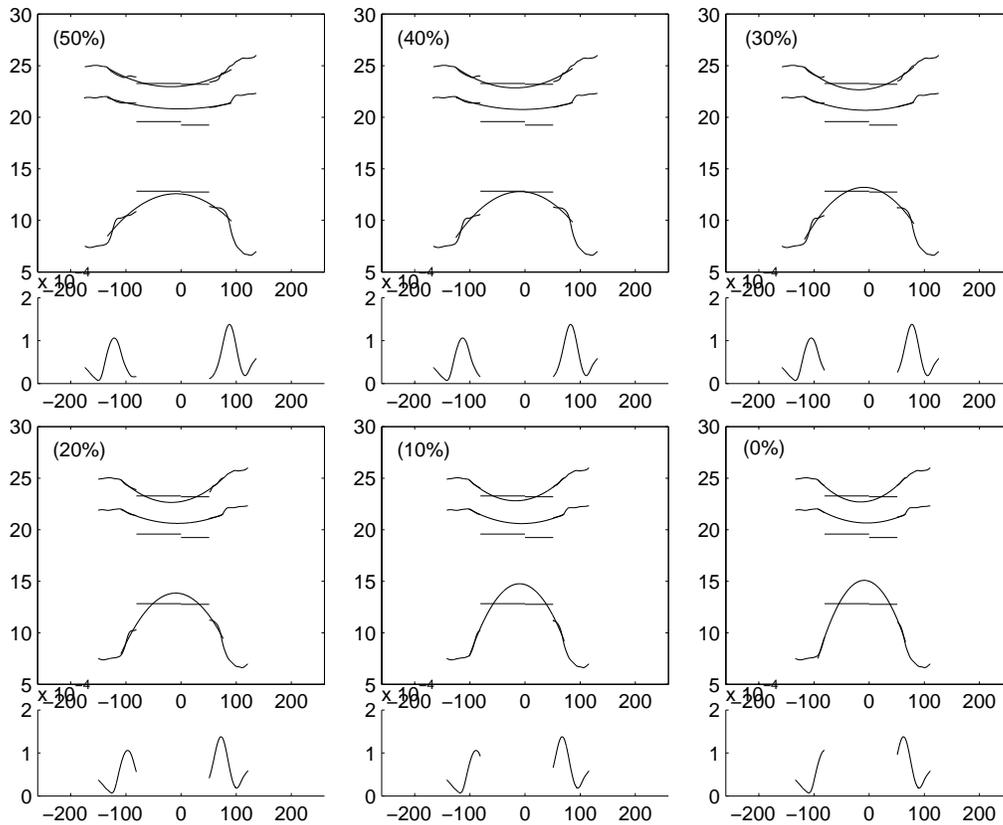


図 5.3: /iai/の R=0.25 (被験者 2) のフィッティングの結果

左上から右へ順に、図 5.2 の横軸 50 から 0 までに対応する。横軸の 0ms のポイントは第 2 母音中心点である。母音の音韻境界は 0ms のポイントよりマイナス方向の時間に実験 4 /iai/ の刺激 1 7、プラス方向に刺激 2 1。

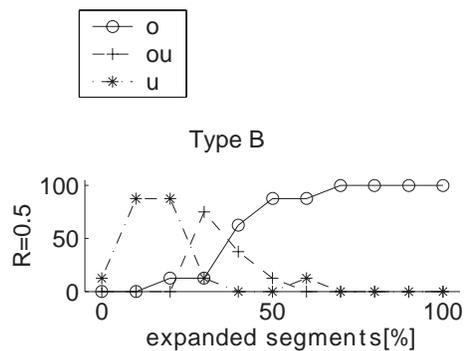
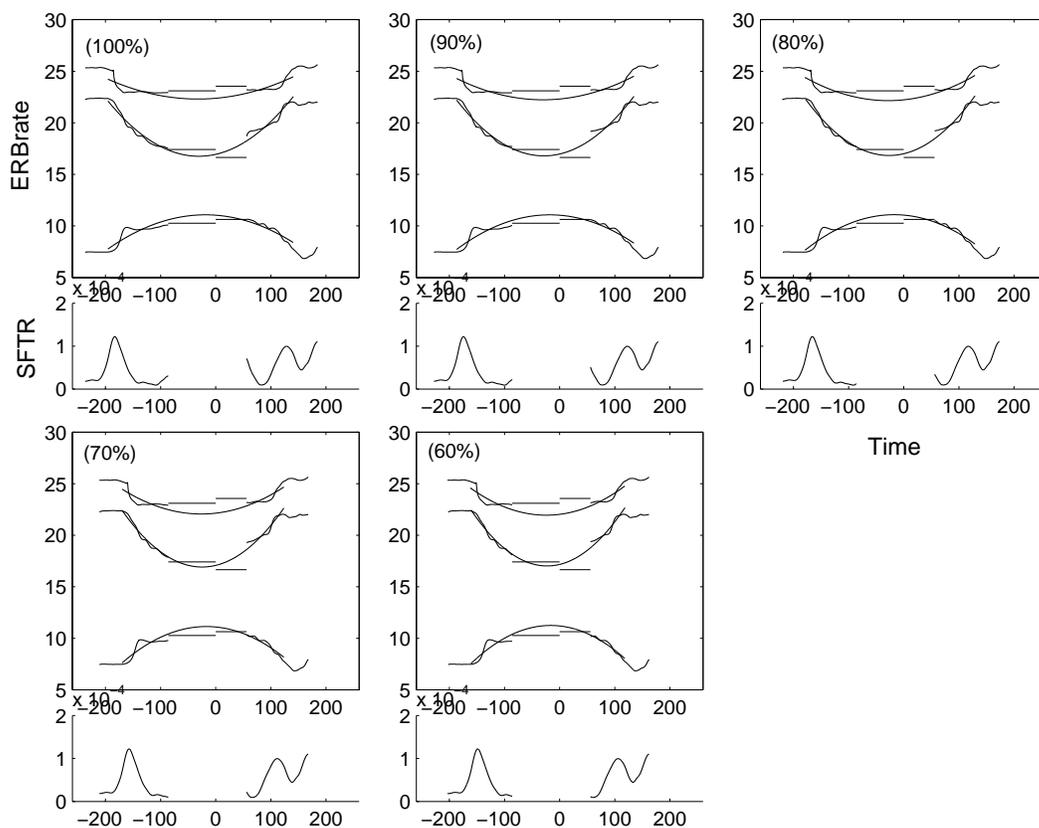


図 5.4: /ioi/の R=0.5 (被験者 1) の回答



左上から右へ順に、図 5.4 の横軸 100 から 60 までに対応する。横軸の 0ms のポイントは第 2 母音中心点である。母音の音韻境界は 0ms のポイントよりマイナス方向の時間に実験 4 /ioi/ の刺激 1 3、プラス方向に刺激 2 4。

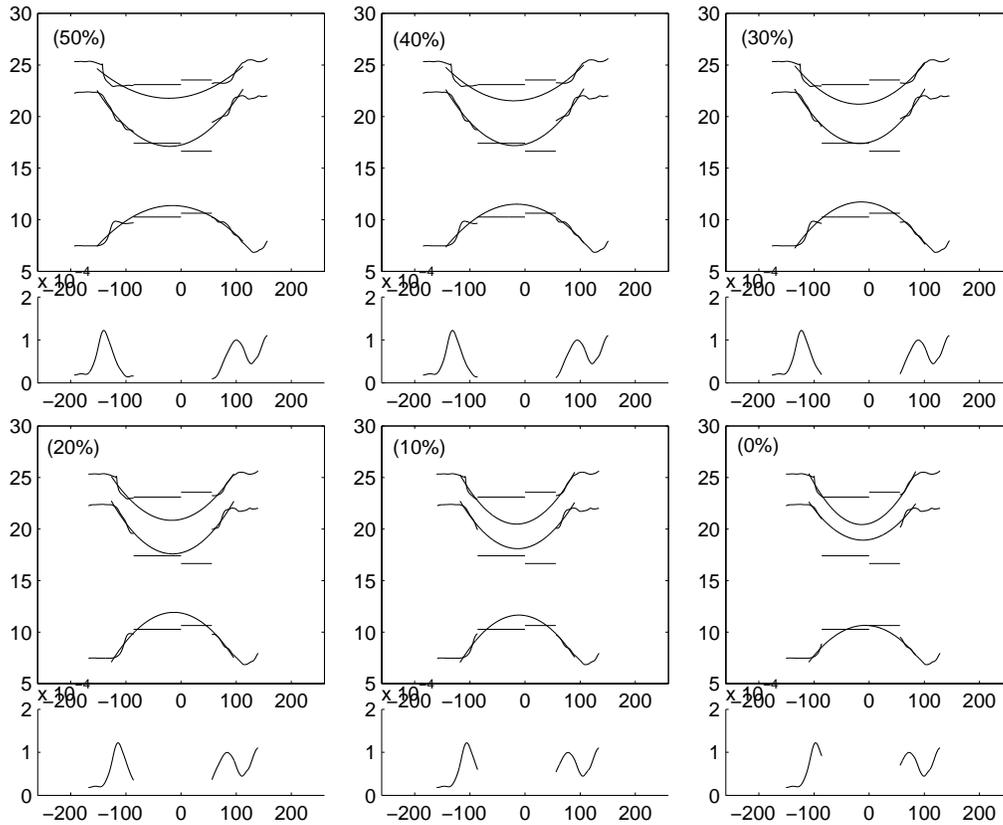


図 5.5: /ioi/のR=0.5 (被験者1) のフィッティングの結果

左上から右へ順に、図 5.4 の横軸 50 から 0 までに対応する。横軸の 0ms のポイントは第 2 母音中心点である。母音の音韻境界は 0ms のポイントよりマイナス方向の時間に実験 4 /ioi/ の刺激 1 3、プラス方向に刺激 2 4。

## 第6章 考察

実験1、2、3を通して、圧縮部の圧縮率と伸張部の長さによって、知覚される母音が異なることがわかった。伸張部の長さの用い方は、被験者によって異なったが、圧縮部の圧縮率の影響は被験者に共通に見られた。よって、わたりの変化速度は、母音知覚に大きく影響することがわかる。

実験1で見られたように、母音の連続を知覚するためには、スペクトル変化率の極大点が含まれたわたりがあれば十分であるが、わたりの変化速度があまりにも急激な場合にはあたりだけでは知覚できなかったことから、中心母音を知覚するためには、わたりの変化速度は有効に用いられていることがわかる。

さらに、Lindblom らの結果でも、変化速度の影響が言われており、その変化速度は補正機構に影響することも言われている。よって、ここで示される変化速度の影響は、Lindblom らの実験結果をサポートするものとなるだろう。

さらに、今回行った関数近似では、実験2の結果の一部と対応が取れるものとなり、圧縮部の変化の速度と、伸張部の長さの追跡による予測機構の存在の可能性が少なからず示唆されている。

/iai/のF1の近似関数と被験者2の結果の対応と、/ioi/のF2の近似関数と被験者1の結果とは、両者とも同じ2次関数の近似で良く対応がとれていた。

しかし、被験者2と被験者1では、元の母音の音韻を知覚できる伸張部の長さが、短い、長いの違いがあったが、この違いは母音の音韻境界の差によるものではないことが実験3からも分かっている。

よって、この違いが追跡または予測の方法の違いを表しているとする、関数近似の方法も変える必要があると思われる。今回の近似は1通りしか行っていないため、近似区間などを変えるとさらに対応が取れるかもしれない。

または、被験者1はF2、被験者2はF1の近似関数に対応していたので、その違いと、伸張部の長さの違いには何か関係があるのかもしれない。F1は凸型の変化であり、F2は凹型の変化であったので、この違いも考えられる。

## 第7章 謝辞

本研究を進めるにあたり、多大なる御指導ならびに御鞭撻を賜りました赤木正人教授に深く感謝の意を表します。

また、日頃から熱心に御討論頂き、有益な御助言を賜りました党建武助教授、ならびに心強いサポートで本研究の遂行を支えてくださった鷓木祐史助手に心より感謝致します。

また、日頃から多大なる議論と激励を頂きました赤木研究室の諸先輩方に厚くお礼申し上げますと共に、本研究の遂行の多面に渡りご協力頂いた赤木研究室の皆様へ感謝致します。

最後に、大学院での貴重な研究生活を与えて頂き、暖かく見守ってくれた両親、祖母、兄弟、そして友人に心から感謝しお礼を申し上げます。

## 関連図書

- [1] Strange, W., Jenkins, J. J., and Johnson, T. L. (1983). “Dynamic specification of coarticulated vowels,” J. Acoust. Soc. Am. 74, 695-705.
- [2] Strange, W. (1989). “Dynamic specification of coarticulated vowels spoken in sentence context,” J. Acoust. Soc. Am. 85, 2135-2153.
- [3] Strange, W. (1989). “Evolving theories of vowel perception,” J. Acoust. Soc. Am. 85, 2081-2087.
- [4] 桑原 尚夫, 境 久雄. (1973). “連続音声中の母音連鎖における調音結合効果の正規化,” 日本音響学会誌, 29 巻, 2 号, 91-99.
- [5] Lindblom, B. E. F., and Studdert-Kennedy, M. (1967). “On the role of formant transitions in vowel recognition,” J. Acoust. Soc. Am. 42, 830-843.
- [6] Ikuyo Masuda-Katsuse, Hideki Kawahara. (1999). “Dynamic sound stream formation based on continuity of spectral change,” Speech Commun. 27, 235-259.
- [7] 河原 英紀. (2001). “高品質音声分析変換合成法 STRAIGHT,”.
- [8] 柏野 牧夫. (1990. 3). “音韻知覚における部位的情報と全体的構造 - 新しい実験法による音韻情報の時間的分散の検討 -, ” 音講論集, 313-314.
- [9] 柏野 牧夫. (1992). “閉鎖区間の前後に分散する手がかりに基づく日本語語中閉鎖子音の知覚,” 日本音響学会誌 48 巻, 2 号, 76-86.
- [10] Sadaoki Frui. (1986). “On the role of spectral transition for speech perception,” J. Acoust. Soc. Am. 80, 1016-1025.