

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	http://hdl.handle.net/10119/1571
Rights	
Description	Supervisor: 赤木 正人, 情報科学研究科, 修士

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

田高 礼子 (010066)

北陸先端科学技術大学院大学 情報科学研究科

2002 年 2 月 15 日

キーワード: 時間伸張、時間圧縮、雑音置換、SFTR (スペクトル変化率) .

1 はじめに

連続音声の中の音素は、調音器官の連続的な動きにより、単独発話時の音響的特徴に到達せずに次の音素へ推移する場合がある。実際、3 連続母音から切り出した中心母音に対して異聴が起こることも確認されている [1]。それでも聞き手は連続音声の中の曖昧な音素を、発話者の意図した音韻として正しく知覚できるが、それは何故なのか。

このとき正しい音韻の知覚に用いられる特徴は、その曖昧な音素の前後音素の音韻やわたりなどに含まれると考えられている。その中でもわたりは、調音結合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、特に有効に用いられていると考えられる。実際にわたりの重要性は、自然音声を用いた聴取実験で示されている [2][3]。

しかし、わたりの詳しい役割、そこに含まれる特徴はわかっていない。よって、連続音声の中の曖昧な音素を正しく知覚するためには、どのような音響的特徴がどのように用いられて (処理されて) いるのかを明らかにする必要があると考える。これらのことを明らかにすることは、音声知覚単位の解明や、上記の Text-to-Speech などの音声処理技術への応用などにも有用な知見を与えると考えられる。

これまで、わたりに着目した研究は多数行われている。合成音声を用いた実験もその刺激の操作が容易であるため、様々なものが行われている。Lindblom ら [4] は、放物状のフォルマント遷移を持つ合成音声 CVC 音節の中心母音知覚について調べ、フォルマント遷移を用いた補正機構と、その機構には変化速度と方向が影響することを提案している。また、鈴木 [5] の実験では、フォルマント遷移の変化量と変化速度を変化させた合成音声刺激の音素同定実験の結果から、これらのパラメータが相補的な関係をもって影響することが提案されている。これらのことから、わたりの変化速度は連続音声の中の音素の知覚に影響があると考えられる。しかし、わたりを模擬した合成音声では複雑な構成の自然音声

が含む情報を正しく与えられない可能性があり、これらの結果が自然音声の知覚にあてはまるかはわからない。

一方、自然音声を用いた Strange ら [2][3] の聴取実験では、CVC 音節中心の定常的な部分が無音置換をした音声刺激を主に用いており、わたりとわたりに与えられる音節全体の時間構造が母音の弁別を決定することを示している。しかし、音声操作の難しさが原因となり、わたりによってどのような時間構造が決定されているかなど、それ以上の詳しいことはわかっていない。

本研究では、3 連続母音の時間構造を変形したとき、中心母音が異聴ではなく元の正しい音韻に知覚できるわたりの時間構造とそれに関連する物理量を聴取実験により調べた。3 連続母音の時間伸縮を分析合成系 STRAIGHT により行い、中心母音部分を雑音置換したものを音声刺激として用い、その刺激の時間構造を様々に変化したときの影響を測定している。

2 実験 1

様々なわたりの時間構造が中心母音知覚に与える影響とその傾向を調べる。

2.1 刺激の構成

3 連続母音として ATR 音声データベース男性話者 mtm の単語データから、/wariai/ の /iai/、/zaiaku/ の /aia/、/nioi/ の /ioi/ を用いた。

刺激の全体長は固定し、その中でわたりの時間構造を変えている。図 1 の (1) に示される、○ は母音中心点、□ はわたり中心点の基準点とし、各基準点間を区間 1 ~ 4 と決定している。この基準点を固定し、各区間の中で時間伸縮を行っている。時間伸縮による時間構造の変形は、大きく図 1 の (1)、(2)、(3) に分けられる。

パターン (1) は時間伸縮しない normal である。ここで、○ は母音中心点、□ はわたり中心点の基準点である。パターン (2) は元の音節のわたり区間を伸張部とし、母音区間を圧縮部とした刺激であり、刺激のわたりの時間構造は中心母音方向へ { 圧縮部・伸張部・圧縮部 } と変化する。パターン (3) は元の音節のわたり区間を圧縮部、母音区間を伸張部とした刺激であり、刺激のわたりの時間構造は中心母音方向へ { 伸張部・圧縮部・伸張部 } と変化する。

パターン (2)・(3) に与えられた伸縮率はそれぞれ 3 種類である。

また、置換雑音はパターン (2)・(3) の各刺激に 3 通りが与えられる。図 1 の long, half, short で表されるように、雑音置換により残る音節の長さが長い、半分、短いの 3 通りである。(1) normal には全 6 通りが与えられる。

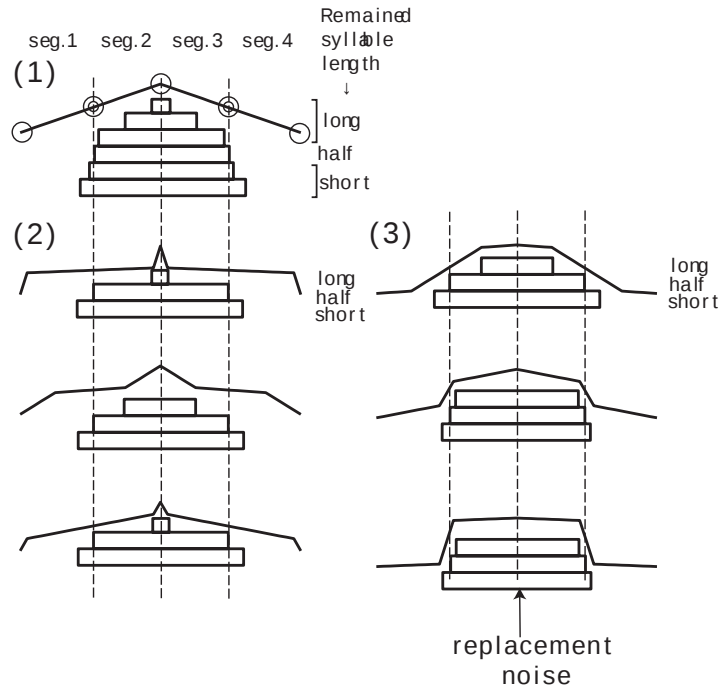


図 1: 実験 1 音声刺激の概略図
各時間伸縮刺激をフォルマント遷移の概略により表している。

2.2 実験方法

被験者は大学院生の男性 3 名。聴こえた音韻をノート PC のキーボード入力による回答とした。教示として、3 連続母音を変形したものに雑音が混在した刺激を呈示することを伝えた。また注意点として、雑音が音節の一部として聴こえる場合は、その部分を例えば「x」などでわかりやすく表記するよう伝えた。刺激の聴き直しを許した。1 試行では全刺激 72 個がランダム呈示され、それを防音室で個別に 5 回行った。また刺激はヘッドホンで両耳に、音圧レベルは最大が 55dB SPL 以下になるよう呈示した。全刺激 20kHz sampling、16bit 量子化である。

2.3 実験 1、結果のまとめと考察

/iai/の結果を図 2 と図 3 に示す。図 2 は時間伸縮していない normal と図 1 のパターン (2) の刺激に対する結果で、図 3 は図 1 のパターン (3) の刺激に対する結果である。

2.3.1 スペクトル変化率極大点の存在

結果は大きく中心母音が知覚できる、できないに分かれた。その刺激を再分析し、得られたスペクトルから線スペクトル (LSF)、LSF からスペクトル変化率 (SFTR) を求め

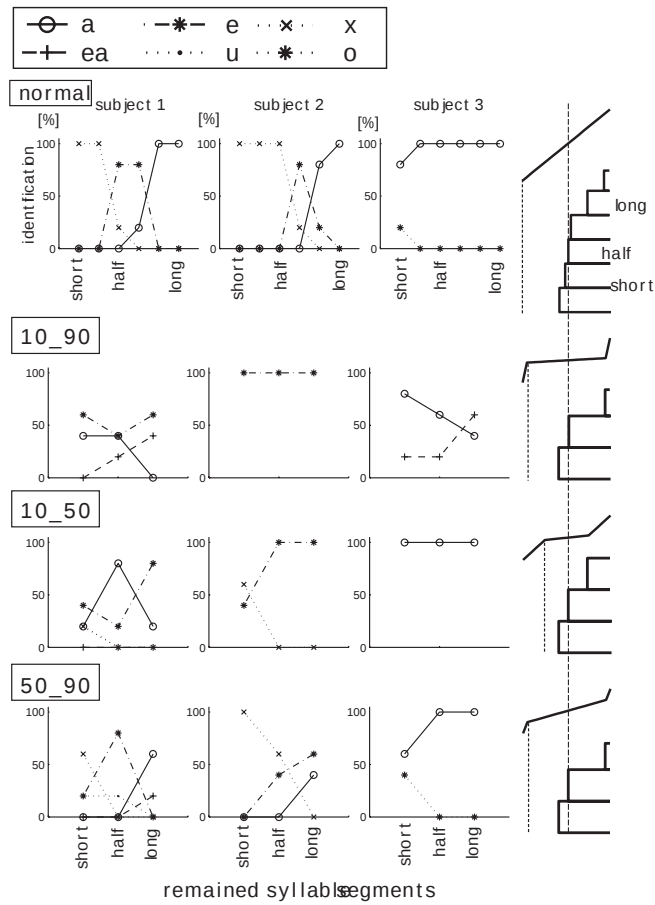


図 2: /iai/の結果その 1 (normal とパターン (2) の刺激の結果)

ると、前者の刺激のわりには SFTR 極大点が含まれることが特徴として見られた。

SFTR 極大点の重要性は Furui によって示されている [7]。その実験では、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知られている。

これらのことから、2つの母音の連鎖を知覚するためには、SFTR 極大点が必要であることがわかる。

2.3.2 パターン (2) の刺激について

刺激のわりには、圧縮部全部が含まれており SFTR 極大点を含むので、中心母音を知覚される結果となった。さらに、伸張部が長いときに異聴が起こり、伸張部が短くなると元の音韻が知覚されるような傾向 (図 2 の被験者 1・3 の結果) が見られた。特に図 1(2) の一番上の刺激のような時間伸縮率が大きい刺激で見られた。

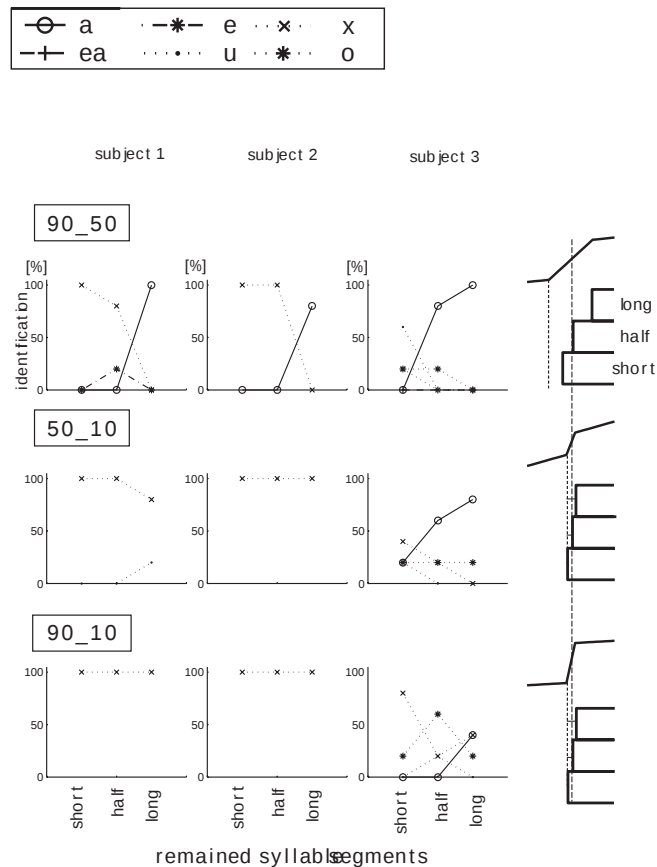


図 3: /iai/の結果その 2 (パターン (3) の刺激の結果)

伸張部は母音間の中間の音響的特徴を持ち、その部分が長く存在するために異聴が起こるのではないかと考えられる。そのことから、伸張部が長いときはその影響が大きく、伸張部が短いときは圧縮部の影響が大きいのではないかと考えられる。

しかし、伸張部が母音間の中間の音響的特徴を持っていること、また圧縮部の圧縮率と伸張部の長さの影響の関係は詳しく調べる必要がある。

2.3.3 パターン (3) の刺激について

刺激のわりに SFTR 極大点が含まれていても、中心母音の知覚はできない場合が多かった。おそらく圧縮部における変化があまりにも急激なために、わりには中心母音が知覚できるための働きをしていないと考えられる。

3 実験 2

実験 1 で用いたパターン (2) の刺激について、そのわたりの時間構造に関連する物理量と中心母音知覚との関係を調べる。

3.1 予備実験

主に伸張部の持つ音響的特徴はどのような音韻を表すのかを調べるために、実験 1 で用いた 3 連続母音の周波数スペクトル断面から作成した単独母音の音韻境界を実験 2 の被験者ごとに調べた。結果の音韻境界は被験者 3 名で大差はなく、実験 2 の刺激で伸張部とする部分は、ほぼ母音と母音の中間の音韻を示している。

3.2 刺激の構成

用いた 3 連続母音は実験 1 ・ 予備実験と同じである。

刺激の全体長は可変にし、雑音置換によって残されるわたりの時間構造が、中心母音方向に { 圧縮部・伸張部 } と変化する刺激を作成し、その圧縮率、伸張部の長さを変化させている。

時間伸縮による時間構造の変形は、図 4 のように伸張部の到達点が異なる Type A・B に分けられる。Type A は伸張部の中心に図 1(1) の基準点 が位置する。Type B は圧縮部と伸張部の境に基準点 が位置する。

さらに Type A・B に、伸張率は 1 種類 (5 倍)、圧縮率は 4 種類 (図 4 の (1)1 倍、(2)0.75 倍、(3)0.5 倍、(4)0.25 倍) を与えた。

置換雑音は一定長 (区間 2 + 区間 3 を置換した雑音) に固定している。その雑音の前後に伸張部を切り詰めた刺激を与えて雑音置換とした。伸張部全長を 100 % とすると、100 ~ 0 % まで 10 % 刻みで切り詰めている。(図 5 は例)

3.3 実験方法

被験者は実験 1 の被験者 1・3 と大学院生の男性もう 1 名の 3 名である。1 試行は 2 つに分けられ、各々で全刺激 264 個の半分がランダム呈示され、試行回数は 8 回とした。他は実験 1 と同様である。

3.4 実験 2、結果のまとめと考察

図 6 に被験者 2 の /iai/ の結果を示す (グラフの R は圧縮率)。図 7 に被験者 1 の /iai/ の結果を示す (グラフの R は圧縮率)。

被験者 2・3 では同様の傾向を示し、それは主に図 6 に見られるような圧縮部の圧縮率

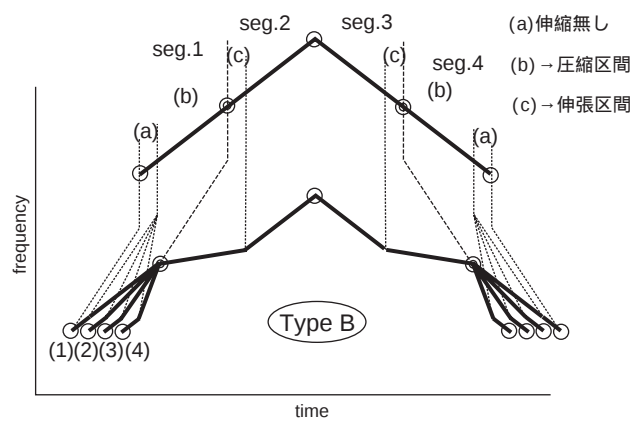
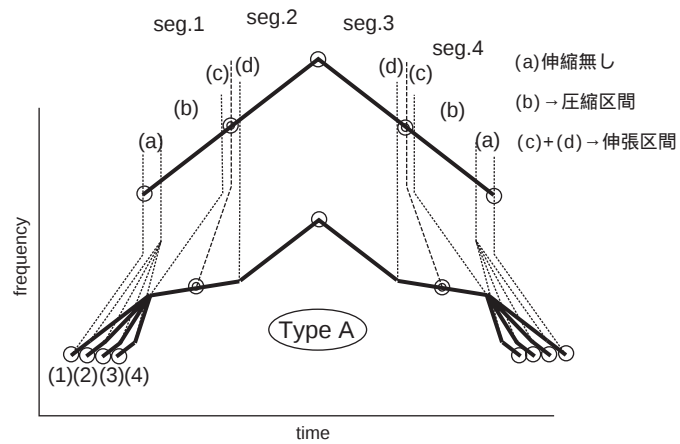
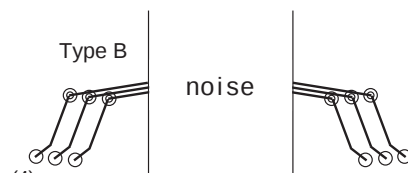
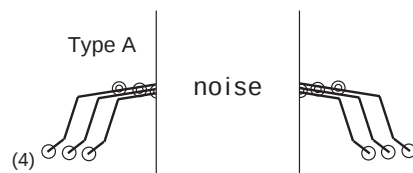


図 4: 時間伸縮方法



同じ伸張率で異なる伸張部の割合のものを3種類ずつ示した。

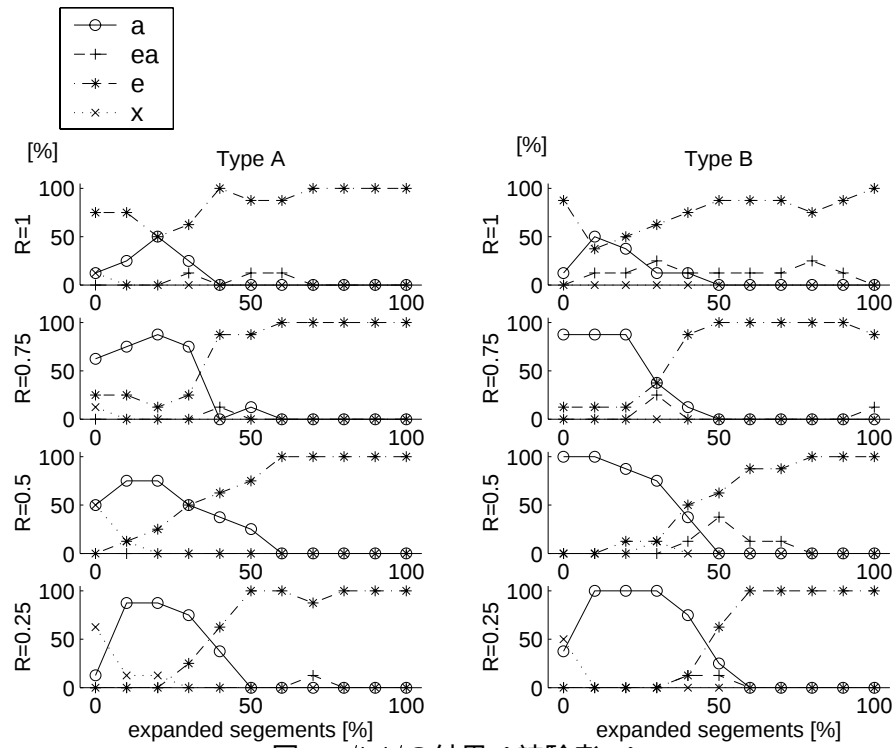


図 6: /iai/の結果 (被験者 2)

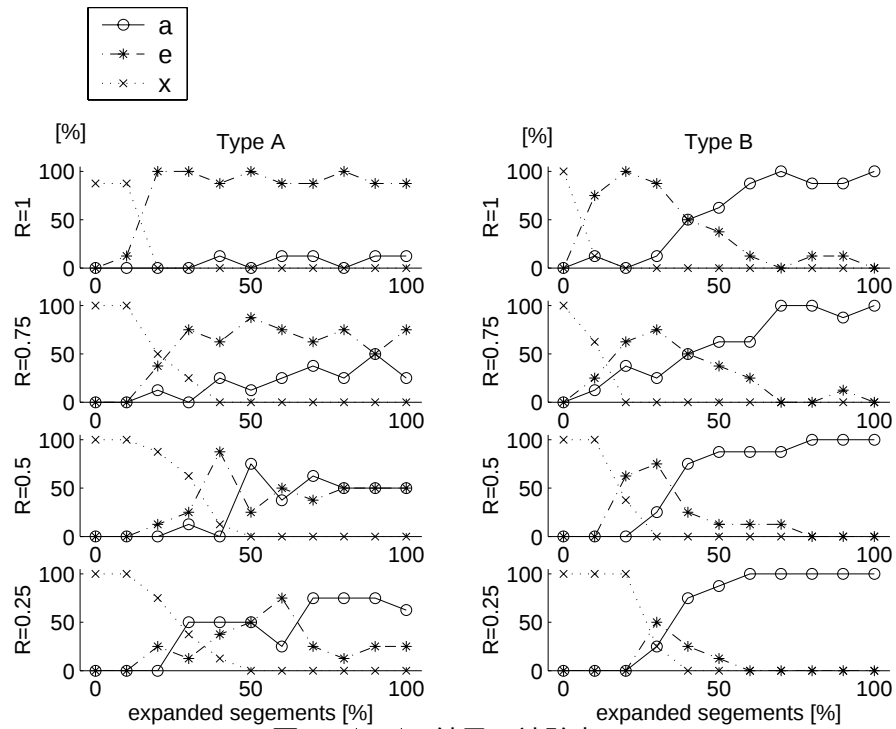


図 7: /iai/の結果 (被験者 1)

グラフの R は圧縮率を示す。

が小さいほど、元の中心母音の音韻が知覚され始める伸張部の長さは短くなるという傾向である。

一方被験者 1 の /iai/、 /ioi/ では、図 7 に見られるような結果となり、圧縮部の圧縮率が小さいほど、元の音韻の母音が知覚され始める伸張部の長さは長くなるという傾向を示した。

ここで、予備実験で調べられた、3 連続母音の周波数スペクトル断面から作成した単独母音に対する被験者ごとの音韻境界と、伸張部到達点を比べると、伸張部は元の音韻には到達していない。(/iai/ であれば /e/ の領域にある。)

よって、2 つの結果では伸張部の影響は異なっているが、知覚される中心母音が音韻が異聴から元の音韻へと変化するような、圧縮部の圧縮率と伸張部の長さとの規則的な関係が存在することがわかる。

4 全体考察

4.1 母音回復のための時間構造とその関連物理量

実験 1・2 から、中心母音が元の音韻として知覚される現象は、わたりの時間構造が { 圧縮部・伸張部・圧縮部 } である刺激において見られた。

その時間構造の中でも、異聴から元の音韻へと変化するような、圧縮部の圧縮率と伸張部の長さとの規則的な関係が存在した。

さらに、圧縮部は変化速度が大きい部分であるので、圧縮率の影響を変化速度の影響として見ると、合成音声を用いた Lindblom ら [4] や鈴木 [5] の結果をサポートすることがわかる。

ここで、Strange ら [2][3] の提案する、わたりによって全体の時間構造が決定される、ということを見ると、元の音韻が知覚されるとき、圧縮部の変化速度と伸張部の長さの 2 つの物理量によって中心の雑音区間の時間構造が決定され、元の音韻が回復されているのではないかと考えられる。

4.2 関数近似による全体の時間構造決定に関する考察

本研究では 2 つの物理量によって、全体の時間構造が規則的に決定されると仮定し、わたりのフォルマント遷移を 2 次関数で近似し、雑音部分を補間することによって、{ わたり・雑音置換部分・わたり } における時間構造が知覚的に決定されると見立てる。

そして、実験 2 で用いた刺激のわたりのフォルマント遷移に 2 次関数をあてはめ、実験結果と単独母音の音韻境界との対応を調べた。伸張部の長さの影響の違いについてはまだよくわからないため、その違いを表す近似は設けなかった。

わたりのフォルマント遷移は Speech Tools の分析プログラムにより分析し、F1,F2,F3 について周波数軸を ERBrate に直し、その圧縮部と伸張部に対し最小自乗法により 2 次

関数の近似を行った。

図 8 は実験 2 の刺激 (/iai/、Type B、圧縮率 0.25) に対する被験者 2 の結果である。この刺激に対し関数近似を行った結果と被験者 2 の単独母音 /a/ と /e/ の境界を図 9 に示す。

F 1 に近似した 2 次関数の頂点に注目すると、関数の頂点が単独母音の音韻境界を越えて /a/ の領域に十分到達したところで、実験 2 の回答も /a/ に変化しているように (図 8 の 40 のときに対応)、回答結果と近似結果の対応がとれる。

また、図 10 は実験 2 の刺激 (/ioi/ の Type B、圧縮率 0.5) に対する被験者 1 の結果である。この刺激に対し関数近似を行った結果と被験者 1 の単独母音 /o/ と /u/ の境界を図 11 に示す。F2 に近似した 2 次関数の頂点に注目すると、同じように回答結果と近似結果の対応 (図 10 の 40 ~ 20 付近に対応) がとれる。

全体の時間構造の決定は、2 次関数によって一部の結果で対応がとれた。よって、このような規則的な方法によって、全体の時間構造が決定される可能性があることが示される。

さらに、2 次関数の近似方法で中心母音を予測するような機構が知覚に存在することが示唆される。

5 まとめ

3 連続音声の中の中心母音が正しい音韻に知覚できるわたりの時間構造とそれに関連する物理量を聴取実験により調べ、以下のことがわかった。

- 中心母音が元の音韻として知覚される現象は、わたりの時間構造が { 圧縮部・伸張部・圧縮部 } である刺激において見られた。
- 知覚される中心母音の音韻が異聴から元の音韻へと回復されるような、速い変化区間の变化速度と中心母音部分に近づく定常的な区間の長さとの規則的な関係があった。
- 全体の時間構造を 2 次関数の近似により決定した考察では、中心母音を予測するような機構が知覚に存在することが示唆された。

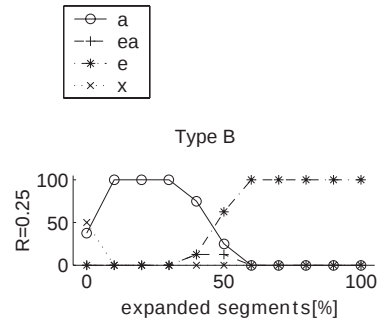


図 8: /iai/の圧縮率 0.25 の回答 (被験者 2)

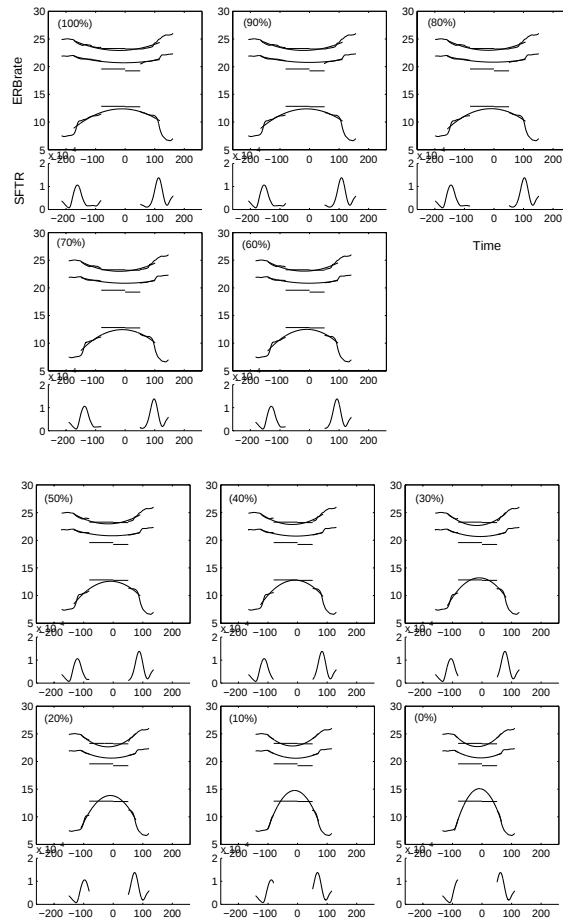


図 9: /iai/の R=0.25 (被験者 2) のフィッティングの結果

左上から右下へ順に、図 8 の横軸 100 から 0(伸張部の長さ [%]) に対応する。横軸 0ms のポイントは第 2 母音中心点である。単独母音の音韻境界は直線で表され、2 次関数の頂点が越えない範囲が/e/、越えた範囲が/a/である。

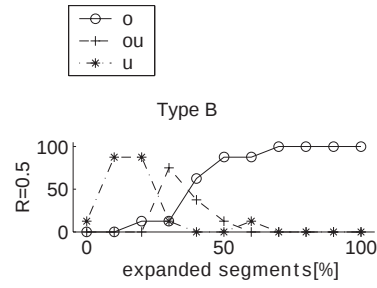


図 10: /ioi/の圧縮率 0.5 (被験者 1) の回答

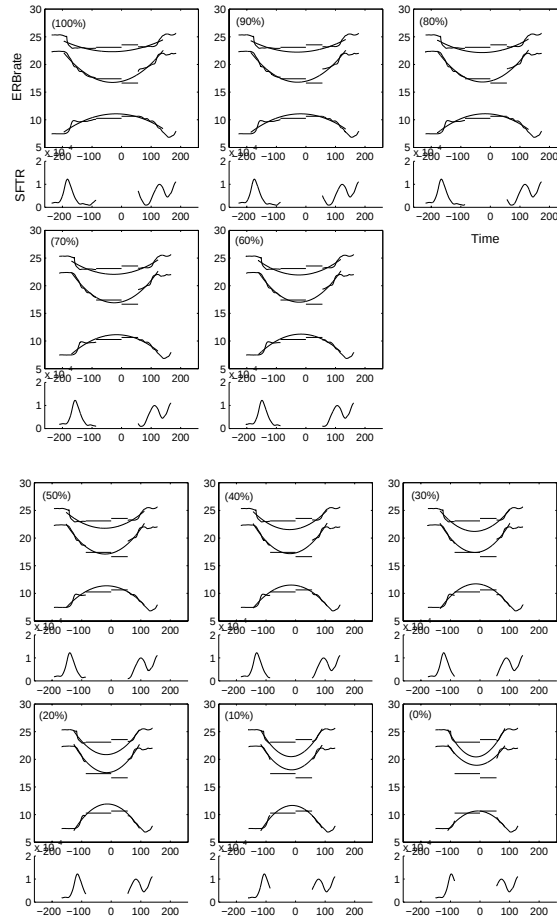


図 11: /ioi/の圧縮率 0.5 (被験者 1) のフィッティングの結果

左上から右下へ順に、図 10 の横軸 100 から 0(伸張部の長さ [%]) に対応する。横軸 0ms のポイントは第 2 母音中心点である。単独母音の音韻境界は直線で表され、2 次関数の頂点が越えない範囲が/u/、越えた範囲が/o/である。

参考文献

- [1] 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.
- [2] Strange, W., Jenkins, J. J., and Johnson, T. L. (1983). “Dynamic specification of
coarticulated vowels,” J. Acoust. Soc. Am. 74, 695-705.
- [3] Strange, W. (1989). “Dynamic specification of coarticulated vowels spoken in sen-
tence context,” J. Acoust. Soc. Am. 85, 2135-2153.
- [4] 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.
- [5] 鈴木 久喜. (1974). “母音・半母音・有声破裂音の知覚におけるホルマント遷移の変
化量と変化速度との間の相補性およびその識別機構,” 日本音響学会誌. 30 巻, 3 号,
169-180.
- [6] 河原 英紀. (2001). “高品質音声分析変換合成法 STRAIGHT,”.
- [7] Sadaoki Frui. (1986). “On the role of spectral transition for speech perception,” J.
Acoust. Soc. Am. 80, 1016-1025.