

音楽と音声に共通するリズム要素の単音列音声における検証

吉田友敬^{†1} 原史恵^{†1} 梅原綾花^{†1} 棚橋紀幸^{†1}
田添詩奈^{†1} 行村涼^{†1} 武田昌一^{†2}

概要：音楽演奏におけるフレージングで重要な役割を果たす、緊張-弛緩の基本構成はグレゴリオ聖歌など、詩の韻律におけるリズムの概念に帰着される。これは言語学におけるフット（詩脚・韻脚）の概念と共通するが、このリズム的要素が、音声の中でどのように現れ、検証可能であるかどうかを検討した。先行研究では音響パワーを見るなどしたが、その後モーラ長に着目し、日本語フットがモーラ長にどのような影響を与えるかを検討した。音素の違いがモーラ長に影響することから、今回は、単一音素の繰り返しによる音声における、2モーラフットのリズム構造を検証した。2モーラごとに抑揚をつけた場合、つけない場合でのモーラ長の変化について、3種の無声破裂音、3種の母音を組み合わせた9種類のモーラによる測定を行った。その結果、多くの測定で奇数モーラより偶数モーラの方が長くなり、日本語音声におけるフットの実在性を示唆する結果となった。

キーワード：リズム、フット（詩脚、韻脚）、モーラ長、破裂音

Verification of the Common Rhythmical Factor both in Music and in Speech

TOMOYOSHI YOSHIDA^{†1} FUMIE HARA^{†1} AYAKA UMEHARA^{†1}
NORIYUKI TANAHASHI^{†1} SHIINA TAZOE^{†1} RYO YUKIMURA^{†1}
SHOICHI TAKEDA^{†2}

Abstract Musical performance has its important characteristics of the typical formula : tense – relax, which is related to the prosody of poems like Gregorian chants. This notion is connected to the term a “foot” in linguistics. We consider how a foot influences the length of each mora, which is phenome-dependent, so, we tested the case that same phoneme is succeeded, especially the case of three voiceless plosive sounds and three vowels. As a result, many data show a consistent tendency that even-numbered mora are longer than odd-numbered mora.

Keywords: Rhythm, Foot, Length of Mora, Plosive Sound

1. はじめに

本研究は、音楽演奏などの感性的時系列データにおける構造的特徴を、同調（Synchronization, Entrainment）などの概念を手がかりに追究していたことに端を発する。

現象の背後には、音楽、音声などを含む様々の時系列全般における普遍的特徴の存在も見いだされるが、また、同時に音楽など感性的要素の強い時系列現象に特有の傾向も見いだされる。それは、一言で言えば、緊張-弛緩という図式で表現できる共通した構造であり、これについては、様々の研究や文献で言及されている。例えば、最近の研究で、中西らは、音楽に合わせた動きの中に同様の構造を見だしている[1]。

こうした構造は、言語における主語と述語のような文法的構造との関連もあり、その意味で音楽と言語、音楽と音声の特徴は多く共通しているといえるであろう。

本研究は、このような考えから、音楽自体の分析をいったん離れ、音声におけるリズム的特徴の一端を検証するも

のである。

2. 演奏における緊張-弛緩の構造と韻律

前述のように、音楽などにおける基本的なフレーズ構造として、緊張-弛緩という二項図式を挙げることができる。これは、音楽の演奏において、緊張と弛緩が繰り返され、それが演奏に独特の息づかいを与えることによって、音楽的な感動や感興を生み出すという考え方である。

このこと自体は、演奏者の立場からすれば当たり前のことであるが、これだけでは、現象の本質を明らかにしたことにはならない。そこで、これを明らかにするためのヒントとして、音楽家が演奏の本質としてよく学ぶことの多い、グレゴリオ聖歌におけるフレージングの考え方を参照する。

グレゴリオ聖歌、中でもソレム派の演奏哲学に、独特の考え方を見いだされる[2]。ここでは動律（リズム）という要素の中で、この緊張と弛緩の構造について Arsis（アルシス）と Thesis（テジス：教会ラテンの読み方による）という概念がそれぞれ緊張と弛緩に対応する。これは、Ictus（イクトゥス）と呼ばれる時点を契機として、そこから気持ちが高揚していき、そしてまた、Ictus に向かって着地すると

^{†1} 名古屋文理大学
Nagoya Bunri University

^{†2} 高野山大学
Koyasan University

というのが、フレーズを歌うときの考え方である。この考え方は、詩の韻律に由来するものであるため、ラテン語の歌詞を読み上げることに際しても、同様に Ictus と Arsis や Thesis が定められる。こうして、音楽のフレーズングの源は、詩の韻律に結びつけられることとなる。

3. 音声リズムの単位としての Foot

一方、言語学、とりわけ音韻論において、音声のリズムに関する知見が深められてきている。ここでは、前述の Ictus と起源を同じくすると思われる Foot (フット) という概念が提唱されている。

たとえば、英語の詩において、シェイクスピアを引用するならば、

> > > > >

|To Be | or not | To Be, | That is | the Question. |

のように、音声はアクセントを含む数音節ごとに区切られ、韻を踏んでリズムを構成することになり、このリズム単位を Foot と呼ぶ。

4. 日本語における Foot

こうした特徴は、英語のように強弱アクセントを伴い、音節の長さがかなり変化する(柔らかい)いわゆる音節言語に特有のものであると考えられてきた。

しかし、窪菌らは、日本語においても同様のリズム単位があることを提唱している[3]。それによれば、日本語は、基本的に2モーラごとにリズム単位があるとする、2モーラ Foot の存在を見出している。

このリズム的特徴は、たとえば、俳句や短歌に見出される五七調や七五調を支えるものであると共に外来語などにも影響している。たとえば、アイスクリームという語は、語源を考えれば

アイス・クリーム

というように分割されるが、これを日常言語の中で発話するとき、

|アイ|スク|リー|ム

のように感じていると考えられる。また、省略語などにも、2モーラずつを組み合わせたものが少なくなく、就職活動を|シュウ|カツ|、パーソナルコンピューターを|パソ|コン|のように略す例は、枚挙に暇がない。

このようなことが見出された1980年頃には日本語にお

けるフット研究が盛んに行われたようであるが、その後、下火となり、2000年を過ぎてからは、この概念に対する見直しや再検討、あるいは、再評価が行われている[4][5]。

5. Foot の音響的検証

このように議論の分かれる Foot であるが、その要因の一つが、こうした現象、とりわけ言語学・音韻論で言われる現象が、基本的には心理的なものであり、必ずしも物理的に正しいわけではないということである。

たとえば、モーラの等時性と言われる日本語の特徴をとっても、確かに音声イメージすれば、各モーラは同じ時間ずつ区切られて日本語らしいリズムを構成しているように思われる。しかし、これを物理的に測定してみると、各モーラの長さは全く等時ではない。このことは、本研究にも影響を与えている。

このような事情の下で、Foot というものの存在を果たして物理的・音響的に見いだすことができるのか、問題となる。

まず、基本的な音響物理量として、ピッチと音響パワーによる分析を考えることができる。窪菌らも、可能性の段階としながら、2モーラごとにピッチが上行、下行を繰り返す傾向を指摘している[3]。筆者らの先行研究においても、その傾向は多少見いだされ、このことから、2モーラ Foot の上位に4モーラのリズム単位が存在する可能性も考えられる[6]。

また、音響パワーによる分析では、2モーラごとにパワーの途切れる傾向も一部に見いだされた。しかし、音響パワーは当然のことながら、音素の種類によって大きく変化するもので、その傾向は、Foot によると言うよりは、音素によるパワーの途切れといわざるを得ない。もっとも、2モーラずつにパワーが切れるように、音素が並ぶ傾向があると言うこともいえるかもしれないが、これはまた、別の研究課題であろう。

いずれにしても、ピッチの傾向も明示的なものではなく、音響パワーも Foot の検出には適さない。そこで、今回は、音素境界を同定することによってモーラ長に Foot の特徴がみられるかどうかを検証することにした。

ただし、予備実験によって、音素の種類によってモーラ長が変動する割合の方が、Foot による変化よりも影響が大きいことがわかった[7]ため、まず、手がかりとして、同じ音素を何回か繰り返した場合のモーラ長の変化を見た。また、モーラ長の測定が容易であることから、まずは、破裂音に限定して、これを3種類の母音と組み合わせたモーラの長さを測定することにした。

6. 単音の繰り返しにおけるモーラ長の測定実験

6.1 実験の目的

同じ音を繰り返して音読した場合のモーラ長に Foot の影響が見られるかどうかを調べる。すなわち 2 モーラ Foot が想定されるため、音声の中の奇数番目のモーラ長と偶数番目のモーラ長を比較する。また、音素による相違を検証する。

6.2 実験の方法

次の 3 つの無声破裂音と 3 つの母音を組み合わせた 1 モーラの音節を 9 回繰り返す文字列を被験者が読み上げ、それを録音した。9 回繰り返すことによって、その間の 8 モーラ分のモーラ長を測定する。無声破裂音は、

/p/, /t/, /k/

の 3 種であり、母音は、各フォルマントが特徴的になる

/a/, /i/, /u/

の 3 種を組み合わせる。その結果、読み上げるパターンは以下の 9 種類となる

1. ぱぱぱぱぱぱぱぱ
2. びびびびびびびび
3. ぶぶぶぶぶぶぶぶ
4. たたたたたたたた
5. ててててててててて
6. ととととととととと
7. かかかかかかかか
8. きききききききき
9. くくくくくくくく

これらを、順に 3 回ずつ繰り返して音読し、録音した。また、この際に、取り立てて抑揚なしで音読した場合と、2 モーラごとにアクセント（多くは、ピッチ+ストレス）をつけたパターンの 2 種類を録音した。パターンの学習効果を避けるため、被験者ごとに読み上げる番号をずらした。

録音に用いたマイクは Shure 社の SM58、Windows パソコンに M-Audio 社のオーディオインターフェース Mobile Pre を用いた。また、録音はフリーソフト Audacity を用いて行った。

6.3 被験者

被験者は、名古屋文理大学情報メディア学部にも所属する学生 10 人（女子 2 名、男子 8 名）である。いずれも東海地

方に在住している。

7. 実験の結果

7.1 波形データからのモーラ長の測定

まず、それぞれの破裂音の音素境界の定義について検討した。

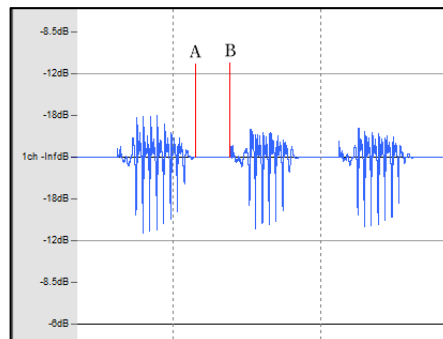


図 1 /papapa/の波形と音素境界

Figure 1 The wave form of /papapa/ and the boundary of phonemes.

図 1 において、破裂音の音素境界は、一般的には破裂音の始まり、すなわち閉鎖区間の始まる図の A によって定義される。しかし、本実験では、破裂の瞬間である B の時点をもって、モーラ長を測定することにする。

予備実験で、A と B の双方で測定を行って比較したところ、A を用いた測定では、音声リズムの特徴がより不明確になった[8]。その理由としては、A は前の音の終了時点であるが、それがすなわち次の音の閉鎖区間の開始時点とは限らず、そもそも閉鎖区間の開始時点を特定するのは容易ではないということ、それに対して、B は破裂の瞬間であるためその特定も容易であること、そして、もう一点は、音声のリズム的要素を考えたとき、破裂の瞬間によって、発話者がリズムを感じていることを推測できるからである。

7.2 Foot とモーラ長の傾向

測定は、被験者一人当たり、

音素 9 パターン × 読み方 2 パターン × 各 3 回 = 54 セット

で実施した。1 セット中にモーラ長のわかるモーラが 8 モーラ (=4Foot) 含まれており、それぞれ奇数モーラと偶数モーラの組みが 4 組含まれる。この 1 セット中のモーラ長の変化の例を図 2 に示す。ここで、奇数モーラと偶数モーラでモーラ長が異なる様子が見て取れる。

また、それぞれ 3 セット分の測定データの中に奇数モーラと偶数モーラの組みが 12 組あるので、この 12 組のデータについて、奇数番目と偶数番目のモーラ長に差異があるかどうか検定を行った。その結果を表 1、表 2 に示す。

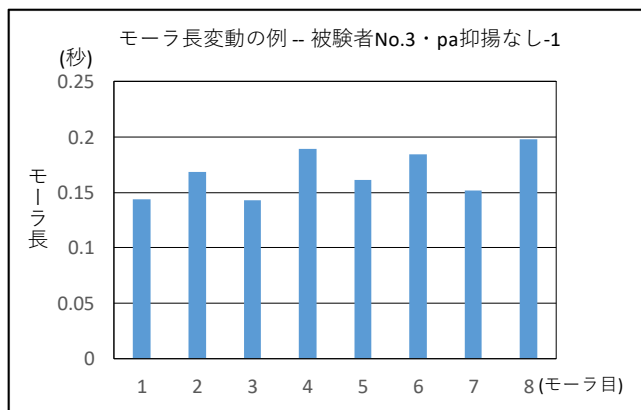


図 2 モーラ長の変化の様子为例

Figure 2 An example of the fluctuation of the length of each mora.

表 1 各パターンの奇数モーラと偶数モーラのモーラ長の差の平均値 (抑揚あり)

Table 1 The Averages of the difference of the length between odd and even mora at each pattern (with accent)

被験者	pa		pi		pu	
	差の平均値	判定	差の平均値	判定	差の平均値	判定
No.1	0.001917		-0.00625		0.00875	
No.2	0.02975	**	0.027833	**	0.030583	**
No.3	0.06325	**	0.031583	**	0.0305	**
No.4	0.008		0.0085		0.02475	*
No.5	0.002583		0.010083	*	-0.00092	
No.6	0.06325	**	0.031583	**	0.0305	**
No.7	0.01725	**	-0.21925		0.00325	
No.8	0.014333	**	0.009667		0.00125	
No.9	0.002583		0.010083	*	-0.00092	
No.10	0.008		0.0085		0.02475	*

被験者	ta		ti		tu	
	差の平均値	判定	差の平均値	判定	差の平均値	判定
No.1	-0.01883	**	-0.0015		-0.00908	
No.2	0.052833	**	0.027917	**	0.045833	**
No.3	0.0235	**	0.015	*	0.016667	*
No.4	0.022083	**	0.0125		0.013583	*
No.5	0.004417		-0.00358		0.00825	
No.6	0.0235	**	0.015	*	0.016667	*
No.7	0.011417	*	0.00925	*	0.010833	*
No.8	0.0225	**	0.02825	**	0.02025	**
No.9	0.004417		-0.00358		0.00825	
No.10	0.022083	**	0.0125		0.013583	*

被験者	ka		ki		ku	
	差の平均値	判定	差の平均値	判定	差の平均値	判定
No.1	0.008833		-0.00658		0.00025	
No.2	0.046667	**	0.039417	**	0.024333	**
No.3	0.020417	*	-0.01467		0.0265	**
No.4	0.004333		0.025833	*	0.020667	*
No.5	0.045917		0.006833		0.0165	*
No.6	0.020417	*	-0.01467		0.0265	**
No.7	0.095	*	-0.00067		-0.06375	
No.8	0.023167	**	0.035917	**	0.0395	**
No.9	0.045917		0.006833		0.0165	*
No.10	0.004333		0.025833	*	0.020667	*

表 2 各パターンの奇数モーラと偶数モーラのモーラ長の差の平均値 (抑揚なし)

Table 2 The Averages of the difference of the length between odd and even mora at each pattern (without accent)

被験者	pa		pi		pu	
	差の平均値	判定	差の平均値	判定	差の平均値	判定
No.1	-0.00392		-0.00692	*	0.00575	
No.2	0.015583	**	0.009417	**	0.019833	**
No.3	0.030417	**	0.032917	**	0.0215	**
No.4	0.024583	**	0.020083	**	0.022083	
No.5	0.00425		0.01425	*	0.031	**
No.6	0.030417	**	0.032917	**	0.0215	**
No.7	0.00475		0.00025		0.007	
No.8	0.008083		0.01425		0.01075	
No.9	0.00425		0.01425	*	0.031	**
No.10	0.024583	**	0.020083	**	0.022083	

被験者	ta		ti		tu	
	差の平均値	判定	差の平均値	判定	差の平均値	判定
No.1	-0.00692		-0.01575	*	0.001667	
No.2	0.018333	**	0.025833	**	0.025583	**
No.3	0.017667	**	0.012667		0.013333	**
No.4	0.016333	*	0.013917		-0.004	
No.5	0.0065		0.004667		0.008417	*
No.6	0.017667	**	0.012667		0.013333	**
No.7	0.01075	*	0.015583	*	0.008	*
No.8	0.004667		0.017083	**	0.0055	
No.9	0.0065		0.004667		0.008417	*
No.10	0.016333	*	0.013917		-0.004	

被験者	ka		ki		ku	
	差の平均値	判定	差の平均値	判定	差の平均値	判定
No.1	0.02425	**	-0.01942	*	0.015167	
No.2	0.032833	**	0.03325	**	0.030333	**
No.3	0.020667	**	-0.002		0.03125	*
No.4	0.031	*	0.00975		0.01925	
No.5	0.006583		0.013	*	0.013	*
No.6	0.020667	**	-0.00342		0.025583	
No.7	0.000417		0.003083		-0.00825	
No.8	0.02075	**	0.005167		0.016083	**
No.9	0.006583		0.013	*	0.013	*
No.10	0.032083	*	0.00975		0.01925	

表の中で、差の平均値は、

(偶数モーラ長 - 奇数モーラ長) の平均値

である。単位は秒である。予備実験から奇数モーラより偶数モーラの方が長くなる可能性が示唆されていたため、後者-前者の値を表示した。実際、ほとんどのパターンでこの値はプラスになっており、偶数モーラの方が長くなる傾向が示されている。また、*は危険率 5%未満、**は危険率 1%未満で、差異が有意であることを示す。一見してわかるように、非常に多くのパターンで有意な差が見いだされた。

7.3 被験者間の差異

それぞれのパターンのデータを被験者ごとに集計し、モーラ長の差の平均値を示したのが、図3、図4である。

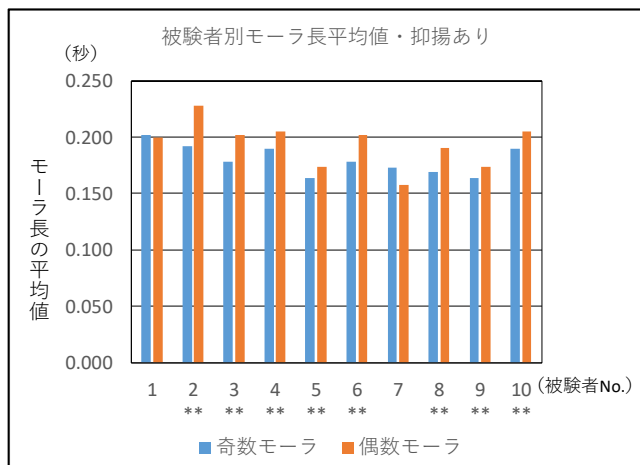


図3 被験者別のモーラ長の平均値 (抑揚あり)

Figure 3 The averages of the length of each subject's mora (with accent).

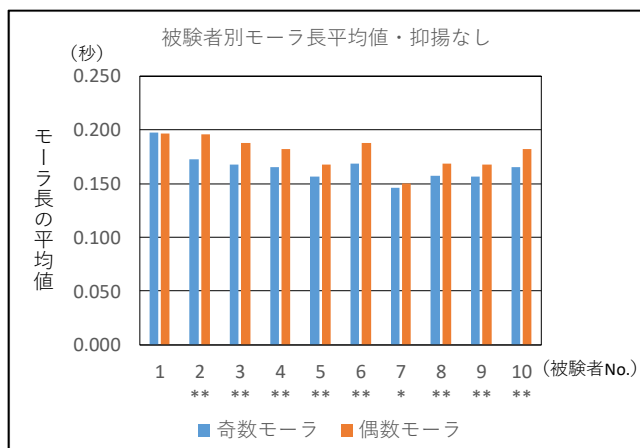


図4 被験者別のモーラ長の平均値 (抑揚なし)

Figure 4 The averages of the length of each subject's mora (without accent).

図中、*は5%、**は1%水準で、その被験者の奇数モーラと偶数モーラのモーラ長で有意な差があることを示している。被験者 No.1 と No.7 だけが、他の被験者と異なる傾向を示している。発話速度によるモーラ長の違いを相殺するために、次のような正規化を行った上で、各被験者のデータの間で有意差があるかどうかを多重比較した。

$$(\text{偶数モーラ長} - \text{奇数モーラ長}) / \text{奇数モーラ長}$$

その結果、やはり、被験者 No.1 と No.7 の関わる組み合わせにおいて、有意差が見出された。特に、No.1 といくつかの被験者間では1%水準の有意差があった。

7.4 音素間の差異

被験者の場合と同様に、それぞれのパターンのデータを

音素ごとに集計し、モーラ長の差の平均値を示したのが、図5、図6である。

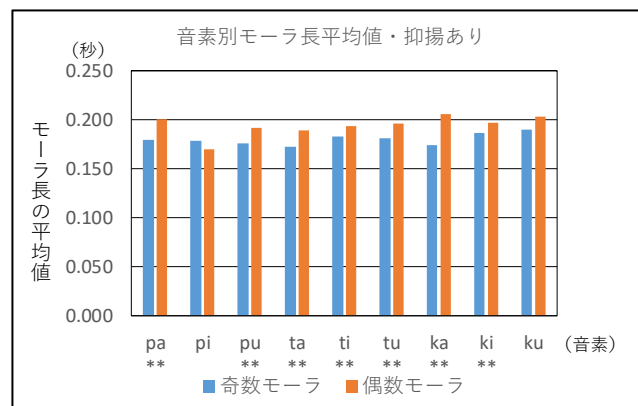


図5 音素別のモーラ長の平均値 (抑揚あり)

Figure 5 The averages of the length of each phoneme's mora (with accent).

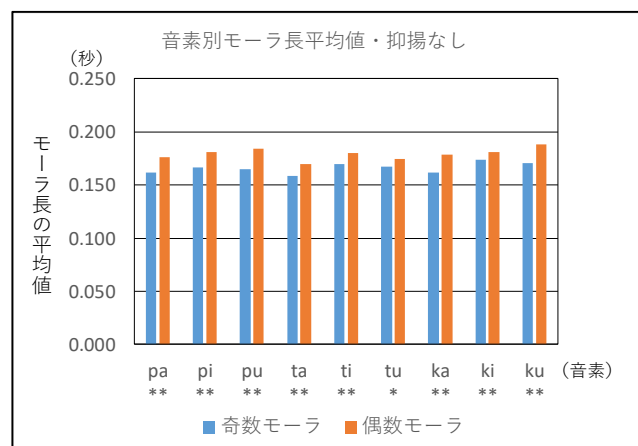


図6 音素別のモーラ長の平均値 (抑揚なし)

Figure 6 The averages of the length of each phoneme's mora (without accent).

検定結果の記号の意味は同様である。抑揚ありのパターンの/pi/と/ku/でのみ、奇数モーラと偶数モーラの間で有意差が生じなかった。また、被験者間と同様に正規化して、多重比較したところ、ほとんどの組み合わせで有意差は見出されなかった。これについては、Footによる奇数モーラと偶数モーラの長さの違いの傾向が、音素によらず普遍的に存在することを示唆していると考えられる。ただし、抑揚なしの場合の/pu/と/tu/および/pu/と/ki/の間でのみ、5%水準で有意差があった。

8. 考察

本実験で被験者に1モーラおきにアクセントをつけるという抑揚を求めたのは、予備実験からFootがモーラ長に反映する度合いは、きわめて微妙であり、まとまった傾向を容易に見出すのは難しいという観測があったためである。

元々Footは詩の韻律にルーツを持つことから、韻を踏むようにアクセントをつけて読み上げれば、多少なりとも傾向が見出されるかもしれないという予想があったのである。また、対照実験として、特にアクセントをつけない抑揚なしのパターンを測定したが、こちらの方は、当初の予想では一定の傾向が見出される可能性は低いと思われた。

しかし、実際に測定してみると、意外なことに、むしろ抑揚なしの方が奇数モーラと偶数モーラの差が大きいくらいに見える結果となった。ただし、両者間の有意な差は見出されていない。いずれにしても、Footで考えられるような2モーラごとのリズムの存在が、特段の韻を踏まなくても読み上げ音声の中に見出されたことは、Footの実在性に対して、積極的な結果となった。

この傾向は、測定した中のほとんどの音素において安定的に見出され、また、ほとんどの被験者においても見出された。一部の被験者で異なる傾向が見出された要因は不明であるが、これが、今回実験を行った東海地方に特有なものなのか、他の方言地域においても共通に見出されるものなのかは、今後の検証に俟つことになろう。

一方、今回の実験において、/pi/の場合だけ抑揚ありの場合で異なる傾向となったことについては、要因は不明である。もっとも、抑揚なしでは、すべての音素で同様の傾向を示している。抑揚ありの方が、恣意的なリクエストをしたことによって、自然な発話からかえって外れることになったのかもしれない。

9. おわりに

今回の実験で、Footと呼ばれるリズム要素がモーラ長という音響的パラメータで検出できる可能性が示された。しかし、今回の測定は単音のくり返しパターンの読み上げであるので、これを異なった音素を含む音声でどのように検出するかが次の課題である。ここで、前述のように、音素の種類によってモーラ長が大きく異なるため、その効果をあらかじめ相殺した上でデータを処理しないと、Footの検出は困難である。音素の種類によってどの程度モーラ長が変わるかについては、先行研究もあるので、それらを参照して今後の検証を進めていく予定である。

また、今回の被験者は全員東海三県に在住しているので、他の地方の被験者の場合との比較も必要であろう。

謝辞

本研究の被験者諸氏、協力いただいた学生諸氏に感謝する。本研究の一部はJSPS科研費JP26330317の助成を受けて行ったものである。

参考文献

[1] 中西里果, 山本佐代子, 桐生昭吾, 武田昌一. 音楽と感情の相

関における動き情報の役割 一緊張と弛緩の交代に着目して一. 日本感性工学会論文誌, 2017, Vol.16 No.5, p.1-5.

- [2] 水嶋良雄. グレゴリオ聖歌. 音楽之友社, 1970, p. 102-201.
- [3] 窪菌晴夫, 太田聡. 音韻構造とアクセント. 研究社, 1998, p. 185-202.
- [4] 藤川直也. 日本語東京方言におけるモーラ・音節・フット. 熊本大学言語学論集, 2013, vol. 12, p. 79-112.
- [5] 五十嵐陽介. 南琉球宮古語池間方言・多良間方言の韻律構造. 言語研究, 2016, vol. 150, p. 33-57.
- [6] 吉田友敬, 武田昌一. 音楽リズムと音声リズムの共通性についての基礎検討. 名古屋文理大学紀要, 2012, vol. 12, p. 113-119.
- [7] 吉田友敬, 武田昌一. 音声リズムの音響的検出方法の検討. 科研費中間成果報告書 (JP26330317), 2015, p. 13-15.
- [8] 吉田友敬, 原史恵, 加藤瞳, 山岸舞子, 武田昌一. モーラ長推定によるフットの特徴抽出の検討. 科研費中間成果報告書 (JP26330317), 2017, p. 17-20.