

Title	遅延聴覚フィードバックが日本語文章読み上げ課題における発話リズムに与える影響
Author(s)	姫子松, 洋介
Citation	
Issue Date	2024-03
Type	Thesis or Dissertation
Text version	author
URL	<a href="http://hdl.handle.net/10119/18892">http://hdl.handle.net/10119/18892</a>
Rights	
Description	Supervisor: 鶴木 祐史, 先端科学技術研究科, 修士(情報科学)

修士論文

遅延聴覚フィードバックが日本語文章読み上げ課題に  
おける発話リズムに与える影響

姫子松 洋介

主指導教員 鶴木 祐史

北陸先端科学技術大学院大学  
先端科学技術研究科  
(情報科学系)

令和6年3月

## Abstract

In spoken communication, the speaker uses speech to express his intention, and the listener listens to the speech to understand the speaker's intention. However, as it is said that "the speaker is also the listener," the speaker's speech is heard by two routes: the listener and the speaker himself. In this study, we focus on the speaker's own listening route. We listen to speech with our own ears, and if there is an error in the speech, we can adjust speech production based on the error. This is called auditory feedback, and humans perform this behavior unconsciously. In addition to verbal information, speech also contains linguistic and paralinguistic information (such as the speaker's emotion), and auditory feedback of this information enables us to understand speech. In everyday life, when communicating with others, the speaker can listen to his or her own speech and adjust it to make sure others can better understand it. Smooth adjustment and communication are possible when auditory feedback works well. However, there are some people for whose auditory feedback may not function well. Examples include speech impediments such as stuttering and pronunciation errors that occur when learning another language. Stuttering is defined as a speech style that disrupts fluency through repetition of sounds, stuttering, and pauses in speech. The relationship between stuttering and auditory feedback has been studied extensively, and it is hoped that methods to improve speech production will be established soon. Regarding pronunciation errors that occur when learning other languages, there are many languages other than Japanese, such as English, German, French, etc., and humans are not able to speak all of them. Linguistic speech contains phonetic information in the form of phonemes, syllables, mora, and other phonetic units. All languages are known to have isochrony of speech units, and the phenomenon of isochrony speech units is defined as the rhythm of speech production. In this study, the rhythm of speech production is defined as the rhythm of speech production. It is known that the world's languages can be divided into three speech rhythms, and Japanese is classified as a mora rhythm. Thus, it is necessary to understand and perceive vocalization and the speech rhythm of a language in order to acquire the speech of another language, but it is thought that misperception of the speech rhythm can lead to pronunciation errors. To solve these problems, speech improvement methods using delayed auditory feedback (DAF) have been studied. has been studied. It is known that when normal speakers speak under DAF, the larger the amount of delay, the more unstable the speech motor control becomes, such as repetition of sounds similar to stuttering symptoms. Assuming that speech motor control based on auditory feedback information is not fully functional, this study aims to realize a speech improvement method using delayed auditory feedback by examining the relationship between the delay time of auditory feedback

information and speech motion on a speech unit basis.

There have been several studies on DAF and speech motor. Yamamoto and Kawabata found that the larger the feedback delay, the longer the average mora duration and the slower the speech rate in a reading task to adapt the DAF (adaptation-in-progress task). In particular, the mean mora duration was longest when the delay was 200 ms. In the post-adaptation task, speech rate increased with a delay of 200 ms, and the mean mora duration was shorter than in the adaptation task. This suggests that the speech motor itself adapts to the DAF and is capable of producing somewhat fluent speech. Ichinose and Noda also found that speech rate decreased the most (was most affected by DAF) under DAF with a delay of 200 ms when significant and meaningless sentences were read aloud. These results indicate that speech produced by reading Japanese sentences aloud under DAF is affected by the speech rate and the mean mora duration, which is the speech unit. However, it is not clear whether the mora duration of the entire utterance or a portion of the utterance is lengthened, and in either case, the mean mora duration is lengthened. Furthermore, they do not mention whether the isochrony of mora changes as the average mora unit lengthens. Dellwo and Wagner used DeltaC, the standard deviation of consonant intervals, and %V, the ratio of vowels to total utterance time, as rhythm indices to quantify speech rhythm, a phenomenon in which speech units are isochronous. An inter-linguistic analysis was conducted to classify rhythms based on differences in speech rate among English, German, and French. Intra-linguistic analysis was also conducted to classify rhythms based on differences in speech rate within the same language, including individual differences. The results showed that DeltaC increased with increasing speech rate in all languages for intra-linguistic analysis, indicating that rhythm classification is possible even with the inclusion of individual differences. On the other hand, no significant change in %V was observed with increasing speech rate. In the inter-lingual analysis, the data for each speech rate were plotted on a graph with %V on the horizontal axis and DeltaC on the vertical axis, and English and German were plotted at positions where %V was small and DeltaC was large, while French was plotted at positions where %V was large and DeltaC was small, indicating that rhythm classification is possible. In a similar experiment, Dellwo analyzed VarcoC, the coefficient of variation of consonant intervals, to look at the relative variability of speech. The results showed that in the intra-linguistic analysis, VarcoC values differed from language to language as speech rate increased, and that some languages could be classified as containing individual differences while others could not. The inter-linguistic analysis showed that the plots were in the same position as DeltaC, indicating that rhythm classification was possible. However, since the speech rhythms of these three languages are

classified into strong and syllabic beats, it has not been mentioned whether the rhythmic index of mora beat rhythm changes depending on individual differences such as speech rate.

In this study, we will conduct a speech experiment in a Japanese sentence reading task under DAF in order to observe how the rhythmic indexes of the above phonemes and mora beats change with the amount of feedback delay. also, Phonemes and mora are extracted from the speech data obtained in the experiment, and their respective rhythm index are obtained. In the speech experiment, 10 sentences of approximately 30 mora were read at four delay conditions (0 ms, 100 ms, 200 ms, and 300 ms) and the speech was recorded. The recorded speech was used for speech rhythm analysis. The first step in the analysis is annotation, which extracts phonemes and morae. Next, the annotated data is used to obtain a rhythmic index for each participant. Next, an intra-participant analysis is performed to see how each participant's rhythmic index varies with the amount of delay in the DAF. The rhythm index is then averaged for each delay amount and evaluated using an inter-participant analysis and statistical tests to see if there is a difference in the mean rhythm index for all delay amounts and if there is a difference between the mean value for delay amount 0 ms and the mean value for other delay amounts. The analysis yielded the following findings. First, the mean durations of consonants, vowels, and moras tended to increase with increasing delay time, and the two experimental participants tended to vary significantly. However, there were large differences in the mean duration of consonants and moras across all conditions, but no differences in vowels across conditions. Second, DeltaC, DeltaV, and DeltaM tended to increase with increasing latency, and two experimental participants tended to vary significantly. These three rhythm indices differed significantly among all conditions. Third, VarcoC, VarcoV, and VarcoM tended to increase with increasing latency, with the majority of VarcoC tending to vary significantly. Half of the VarcoV indicators also tended to change significantly. However, VarcoC and VarcoV differed significantly among all conditions, while VarcoM did not differ significantly among all conditions.

The results of the speech rhythm analysis yielded the following findings. First, the average duration between consonants and moras was found to increase with feedback delay. Second, the variability between consonants, vowels, and moras was found to increase with feedback delay time. Third, feedback delay time was found to increase the relative variability of consonant and vowel durations in speech compared to speech without feedback delay.

# 目次

<b>第1章 序論</b>	<b>1</b>
1.1 はじめに	1
1.2 研究背景	2
1.3 研究目的	7
1.4 論文構成	7
<b>第2章 関連研究</b>	<b>9</b>
2.1 遅延聴覚フィードバックと発話運動に関する研究	9
2.2 発話リズム指標に関する研究	10
2.3 問題点	12
<b>第3章 本研究の方略</b>	<b>13</b>
3.1 方法論	13
3.2 リズム指標の定義	16
3.3 発話リズムの評価	17
<b>第4章 DAF 下での日本語文章読み 上げ課題における発話実験</b>	<b>18</b>
4.1 実験目的	18
4.2 実験条件	18
4.3 実験参加者	20
4.4 実験装置	20
4.5 実験方法	22
4.6 実験結果	24
4.7 分析方法	27
4.7.1 アノテーション	27
4.7.2 発話リズム分析	29
4.8 分析結果	29
4.8.1 実験参加者内分析結果	29
4.8.2 実験参加者間分析結果	50
4.9 考察	61
<b>第5章 全体考察</b>	<b>62</b>

<b>第6章 結論</b>	<b>64</b>
6.1 本研究で明らかにしたこと . . . . .	64
6.2 残された課題 . . . . .	64
<b>謝辞</b>	<b>66</b>
<b>参考文献</b>	<b>67</b>

# 目次

1.1	言葉の鎖と聴覚フィードバックの関係 [1]	4
1.2	世界の言語の発話リズム [49]	5
1.3	遅延聴覚フィードバックの影響 [48]	6
1.4	論文構成	8
3.1	本研究の方略	14
3.2	撥音の有無においての子音区間のヒストグラム	15
4.1	実験環境の概略図	21
4.2	実験手順の概略図	23
4.3	遅延量なしのときと遅延量ありのときのモーラ区間の例	25
4.4	遅延量なしのときと遅延量ありのときの音素区間の例	26
4.5	自動ラベリングの手順の概略図	28
4.6	各実験参加者の遅延量ごとにおける平均子音区間	32
4.7	各実験参加者の遅延量ごとにおける平均母音区間	34
4.8	各実験参加者の遅延量ごとにおける平均モーラ区間	36
4.9	各実験参加者の遅延量ごとにおける DeltaC	38
4.10	各実験参加者の遅延量ごとにおける DeltaV	40
4.11	各実験参加者の遅延量ごとにおける DeltaM	42
4.12	各実験参加者の遅延量ごとにおける VarcoC	44
4.13	各実験参加者の遅延量ごとにおける VarcoV	46
4.14	各実験参加者の遅延量ごとにおける VarcoM	48
4.15	実験参加者間の遅延量ごとにおける平均子音区間	52
4.16	実験参加者間の遅延量ごとにおける平均母音区間	53
4.17	実験参加者間の遅延量ごとにおける平均モーラ区間	54
4.18	実験参加者間の遅延量ごとにおける DeltaC	55
4.19	実験参加者間の遅延量ごとにおける DeltaV	56
4.20	実験参加者間の遅延量ごとにおける DeltaM	57
4.21	実験参加者間の遅延量ごとにおける VarcoC	58
4.22	実験参加者間の遅延量ごとにおける VarcoV	59
4.23	実験参加者間の遅延量ごとにおける VarcoM	60



# 表 目 次

2.1	先行研究で用いられたリズム指標の一覧表 . . . . .	11
4.1	提示文章の一覧表 . . . . .	19
4.2	各実験参加者の遅延量ごとにおける平均子音区間の値 . . . . .	33
4.3	各実験参加者の遅延量ごとにおける平均母音区間の値 . . . . .	35
4.4	各実験参加者の遅延量ごとにおける平均モーラ区間の値 . . . . .	37
4.5	各実験参加者の遅延量ごとにおける DeltaC の値 . . . . .	39
4.6	各実験参加者の遅延量ごとにおける DeltaV の値 . . . . .	41
4.7	各実験参加者の遅延量ごとにおける DeltaM の値 . . . . .	43
4.8	各実験参加者の遅延量ごとにおける VarcoC の値 . . . . .	45
4.9	各実験参加者の遅延量ごとにおける VarcoV の値 . . . . .	47
4.10	各実験参加者の遅延量ごとにおける VarcoM の値 . . . . .	49

# 第1章 序論

## 1.1 はじめに

私たちは日常生活において様々な場面で音声を発している。カラオケで歌声を出したり、驚きや喜びなどを感じたときに反射的に声が出たりする。しかし、ほとんどの人が日常で多く音声を使用する場面は他者とのコミュニケーションである。コミュニケーションにはメールなどの文字ベースでやり取りするものもあれば、表情や身体の動きでやり取りする場面もある。しかし、日常的に良く使われているものは音声でのコミュニケーションである。音声コミュニケーションは音声以外に表情やジェスチャといった視覚的要素も含まれるが、音声は主体となって情報の伝達を行っている。音声コミュニケーションの仕組みについて、話し手が音声を使って発話し、聞き手がそれを聴取することにより成り立っている。また、音声には非言語音声や音声合成などの機械音声があり、近年様々な手段で伝達することができるが、現在もなお音声コミュニケーションでは言語音声は主体となっている。その言語音声には音素や音節、モーラといった分節音（音声セグメント）の他に、イントネーションやアクセントといった超分節的特徴などの言語情報が含有している。それらを声帯などの発声関連筋を巧みに操作することで言語音声を生成できる。この言語音声を無意識に生成して無意識に知覚し合うことでコミュニケーションができる。しかし、この言語情報は言語によって異なり、発声方法も様々である。ヒトは言語音声を意識せずに話すことができるが、生まれつき話すことができるわけではない [49]。これは他者の音声を聞いたり、自分自身が発した音声を聞いたりすることで言語音声を学習し、獲得することができる。そのため、他言語の学習のときも聞くだけでなく、話すことで音声をフィードバックし、学習の効率が上がると考えられる。このように話すことと聞くことは音声コミュニケーションにおいて重要な役立ちをしている。

## 1.2 研究背景

音声コミュニケーションにおいて、話し手と聞き手が存在するが、「話し手は同様に聞き手でもある」と言われるように話し手が発話した音声は聞き手と話し手自身に聴取されている [17]. このことは「言葉の鎖 (Speech chain)」として言語音声のコミュニケーションがどのように成り立っているかをモデル化されている. 本研究では話し手自身が聴取する経路に重点を置く. 私たちは発話した音声を自分の耳で聞いており (聴覚フィードバック), 聞いた音声に誤りがあれば, 聴覚などのフィードバック情報に基づいて発話運動を制御することができる. 本研究では聴覚フィードバック情報に基づく発話運動制御に焦点を当てる. 図 1.1 に言葉の鎖と聴覚フィードバックの関係を示す. 図 1.1 より, 右側の青い部分が話し手自身が聴取する経路である. また, 発話音声には言語情報や非言語情報 (発話者の感情など) が含まれており, それらも聴覚フィードバックされることによって発話音声を理解することができる. 日常では, 他者との会話で自身の発話音声を聴取し, 話す内容を理解することで相手に伝わったり, カラオケで自身の歌声を聞きながら音程を調整し, 音程にズレがあれば即座に修正したりしている. これらは聴覚フィードバックが効果的に機能することによって行うことができると考えられる. しかし, 聴覚フィードバックが効果的に機能していない可能性がある人々も存在する. 例えば, 吃音症などの発話障がい [42] や他言語学習時に生じる発音の誤りなどが挙げられる. 吃音症は音の反復 (連発), 引き伸ばし (伸発), 言葉が詰まって沈黙する (難発) という症状から発話の流暢性を乱す話し方として定義されている [24]. 吃音症と聴覚フィードバックの関係はこれまでに数多く研究されており, 音声発話の改善法の早期確立が期待されている [3]. 他言語学習時に生じる発音の誤りについて, 言語というものは日本語以外に英語, ドイツ語, フランス語など様々あり, ヒトはそれらすべての言語を話すことができるわけではない. 言語音声には音素・音節・モーラなどの音声セグメントといった音声単位の言語情報が含有されている. 全ての言語はその音声セグメントに等時性があることが知られており, この音声セグメントが等時性を持つ現象のことを音声生成のリズムと定義されている [14][29][22]. 本研究では音声生成のリズムのことを発話リズムとして定義している. 図 1.2 に世界の言語の発話リズムを示す. 世界の言語は3つの発話リズムに分けることができると知られており, 日本語はモーラ拍リズムとして分類される [33][34][35]. このように言語音声を獲得するために発声の仕方や言語の発話リズムなどを理解して知覚する必要があるが, それを誤って理解することで発音の誤りが生じる [18]. 聴覚-発話運動制御のメカニズムの解明はこれらの音声コミュニケーションにおける QoL 向上に寄与できる可能性がある. この課題に対して, 発話音声を遅らせて知覚させ発話への影響を調べる遅延聴覚フィードバック (Delayed Auditory Feedback: DAF) 手法を用いた音声発話の改善法が検討されている [44]. 図 1.3 に遅延聴覚フィードバックの影響を示す. 正常発話者は DAF 下での発話の際, 遅延量が大きいほど音の繰り返しなどの発話運動が不安定となる

現象が知られている [4][15]. 本研究では, この現象を聴覚フィードバック情報に基づく発話運動制御が効果的に機能していないと仮定して, 聴覚フィードバック情報の遅延時間と発話運動の関係について調査する.

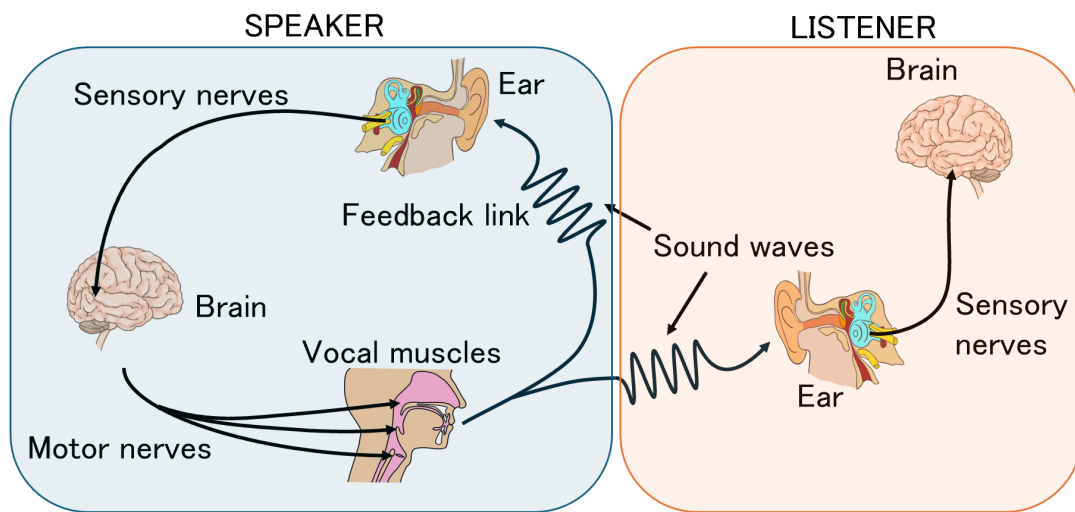


図 1.1: 言葉の鎖と聴覚フィードバックの関係 [1]

強勢拍リズム(英語, ドイツ語)

Humpty dumpty sat on a wall.

ゆっくり発声

早く発声

モーラ拍リズム(日本語)

ほんじつはせいてんなり

カナ1文字を同じ長さの時間で発声

音節拍リズム(フランス語, スペイン語)

J e n e parle pas fran cais

音節を同じ長さの時間で発声

図 1.2: 世界の言語の発話リズム [49]

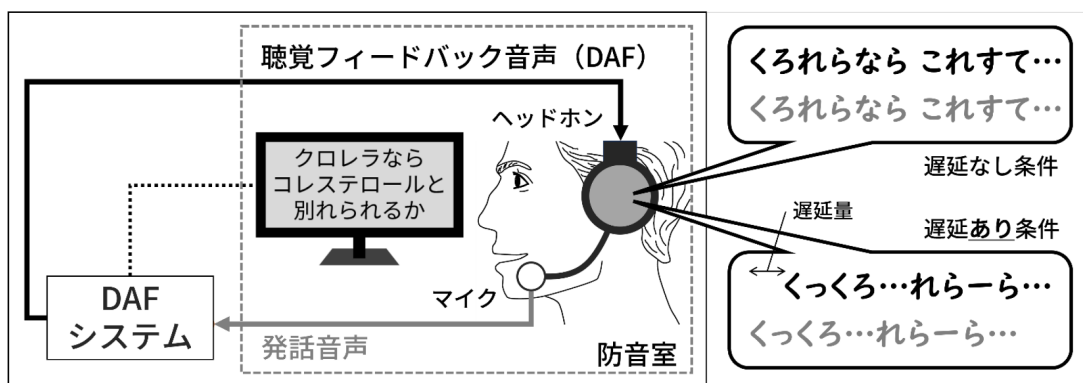


図 1.3: 遅延聴覚フィードバックの影響 [48]

## 1.3 研究目的

遅延聴覚フィードバックによる発話改善法の実現を目指す。そのために聴覚フィードバック情報の遅延時間が発話リズムにどのように影響するのかを明らかにすることを目的とする。本研究では、DAFによる発話運動の影響を見るために、DAF下での日本語文章読み上げ課題における発話実験を行う。実験によりDAF下での音声データを録音し、音声データに含まれる音声セグメントの区間を調べることで発話リズムにどのように影響するのかを明らかにする。詳しい発話リズムについての説明は3.1で行う。

## 1.4 論文構成

本論文は6つの章で構成される。図1.4に本論文の構成を示す。第1章では、音声生成と音声知覚の関係性について説明し、研究背景と研究目的について述べる。第2章では、遅延聴覚フィードバックと発話運動に関する研究と発話リズム指標に関する研究を示し、それらの問題点を述べる。第3章では、本研究の方法論について説明し、本研究で用いるリズム指標の定義と発話リズムの評価方法について述べる。第4章では、DAF下での日本語文章読み上げ課題における発話実験の内容について述べる。第5章では、第4章の結果から考えられる全体考察を述べる。第6章では、本研究で明らかにしたことと、残された課題を述べる。



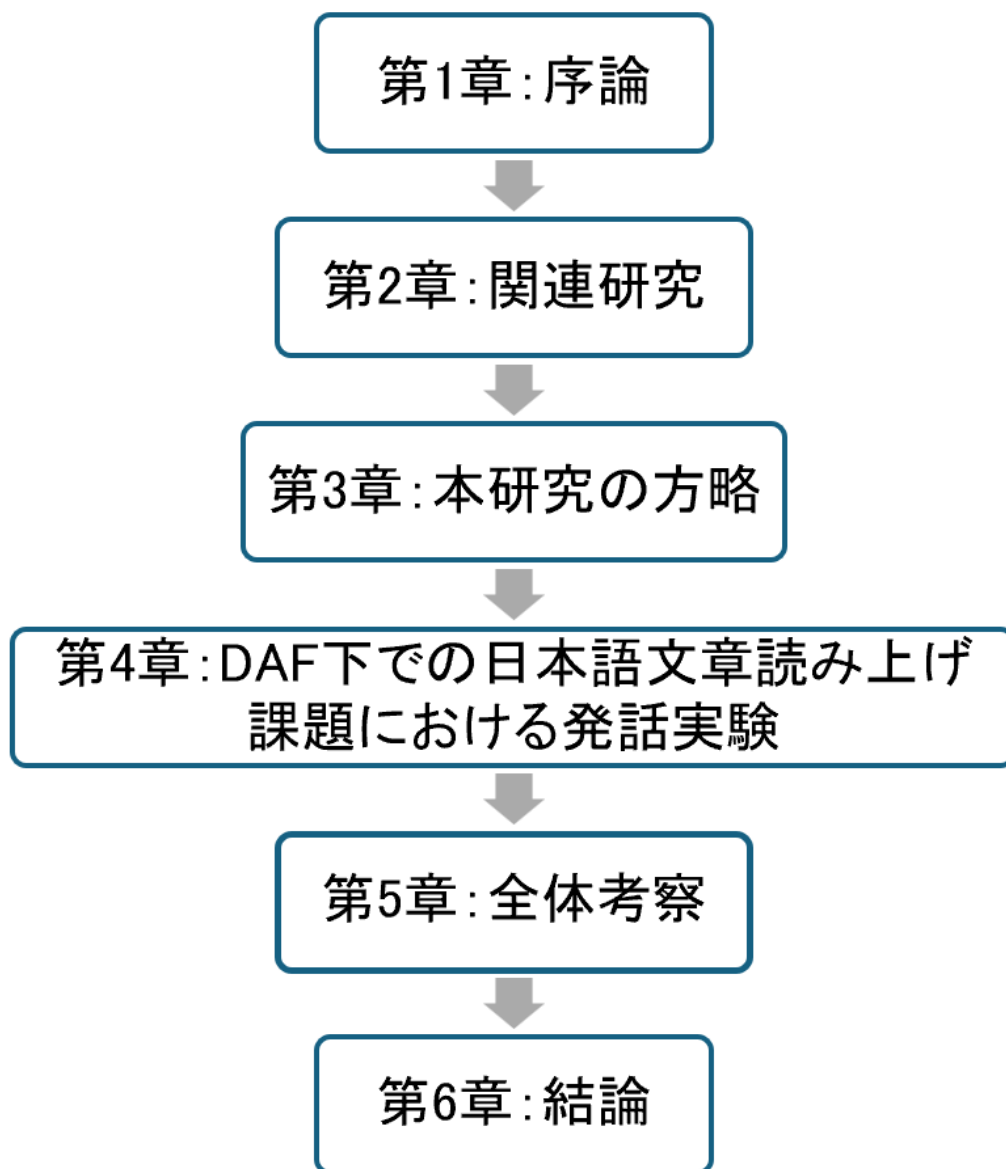


図 1.4: 論文構成

## 第2章 関連研究

### 2.1 遅延聴覚フィードバックと発話運動に関する研究

Yamamoto and Kawabata は DAF 下での発話運動の適応効果に関する研究を行った。実験参加者は発話と聴力に問題のない日本語母語話者 12 名を対象とした。また、提示文章は夏目漱石の「道草」から選定した。実験は合計で 3 回行われ、1 回目の実験では遅延量を 0 ms, 30 ms, 75 ms, 120 ms の 4 条件とした。2 回目の実験では遅延量を 0 ms, 100 ms, 200 ms, 300 ms の 4 条件とした。3 回目の実験では遅延量を 0 ms, 200 ms の 2 条件とした。実験内容について、まず DAF を適応させるための読み上げ課題（適応中課題）を 3 分間実施し、間髪入れずに DAF を適応させた後の課題（適応後課題）を 10 秒間実施した。次に 10 秒間のうちに読み上げができなかった場合は、間髪入れずに自身の声が発声時と同時に聞こえたかどうかを判定する課題（テスト課題）を実施した。適応後課題とテスト課題は合計で 35 回の試行を行った。実験結果より、適応課題中の発話音声はフィードバック遅延量が大きいほど平均モーラ区間が長くなり、発話速度が低下することがわかった。特に遅延量 200 ms の場合に平均モーラ区間が最も長くなることがわかった。また、適応後課題では適応中課題より遅延量 200 ms で発話速度が増加し、平均モーラ区間が短くなることがわかった。これによって発話運動自体が DAF に適応してある程度流暢な発話が可能であることを示唆した [2][5][20]。

一ノ瀬と能田は文章内容・発話速度・遅延時間が遅延聴覚フィードバック条件下での発話に与える影響を調べるために、DAF 下での日本語文章読み上げ課題を行った。実験参加者は発話と聴力に問題のない日本語母語話者 48 名を対象とした。また、提示文章は NTT データベースから親密度が高い有意味文と親密度が低い無意味文をそれぞれ 20 文選定した。発話速度は通常の方法で読み上げるように指示したものと読み上げる直前に聞かせるナレーションの方法で読み上げるように指示したものの 2 条件とした。遅延量は 0 ms, 100 ms, 200 ms, 300 ms の 4 条件とした。実験内容について、遅延量 1 条件ごとに 2 つの文を読み上げて、読み上げにかかった時間を測定した。その測定を提示文章 20 文に対して遅延量と発話速度のすべての条件で行った。また、実験参加者ごとに組み合わせを変えて行った。実験結果より、通常の方法で読み上げるように指示したときの発話音声において発話速度が遅延量 200 ms の DAF 下で最も低下する（最も DAF の影響を受ける）ことがわかった [11][25][32]。また、DAF と発話運動に関する研究は他にもある [37][43][45]。

## 2.2 発話リズム指標に関する研究

Dellwo and Wagner は、子音区間の標準偏差である DeltaC と全発話時間のうちの母音の割合である %V をリズム指標として用い、英語・ドイツ語・フランス語の異なる発話速度の音声でリズム分類を行い、また個人差を含む発話速度で言語内のリズム分類が行えるかを分析した [10]. 発話速度は最も遅い、遅い、普通の速さ、速い、最も速いの 5 段階で読み上げるように指示した. 提示文章は B.Schlink の小説から英語とドイツ語のテキスト 3 文 (76 音節) とフランス語のテキスト 4 文 (93 音節) を選定した. また、実験参加者は英語母語話者 5 名、ドイツ語母語話者 7 名、フランス語母語話者 4 名を対象とした. 結果は、言語内のリズム分類では発話速度の増加で全ての言語の DeltaC が大きくなることがわかり、個人差を含むものでもリズムを分類できることがわかった. また、%V については、発話速度の増加で大きな変化は見られないことがわかった. また言語間のリズム分類では、横軸に %V 縦軸に DeltaC をとったグラフに各言語の各発話速度のデータをプロットしたところ、英語とドイツ語は %V が小さく DeltaC が大きい位置にプロットされ、フランス語は %V が大きく DeltaC が小さい位置にプロットされ、リズム分類ができたことがわかった. Dellwo は、[10] と同様の実験を行い、再度 DeltaC と %V を分析し、言語音声を相対的に観測するために子音区間の変動係数である VarcoC の分析を行った [9][26]. 提示文章は Bonn Tempo Copus というデータベースから約 80 音節の文章を選定した. 実験参加者は英語母語話者 7 名、ドイツ語母語話者 12 名、フランス語母語話者 7 名を対象とした. 結果は、言語内のリズム分類では発話速度の増加に対して異なる言語で VarcoC の値が変わることがわかった. 英語とドイツ語に関しては、VarcoC は発話速度が増加することで小さくなる傾向が見られたが、英語での発話速度が最も速いときの VarcoC と発話速度が最も遅いときの VarcoC の値と近似した. さらに、ドイツ語での発話速度が最も遅いときの VarcoC と通常の水速のときの VarcoC の値が近似した. また、フランス語に関しては、発話速度の増加に対して VarcoC が大きくなる傾向が見られた. このことから発話速度などの個人差を含むもので分類が可能である言語もあれば分類できないものもあった. 言語間のリズム分類では、VarcoC のプロットは DeltaC と同様の位置となり、リズム分類ができることがわかった. また、言語間のリズム分類を目的としてリズム指標を分析した研究は他にもある [8][19][27][28][36][38][39][40][41]. 表 2.1 にリズム指標の一覧表を示す.

表 2.1: 先行研究で用いられたリズム指標の一覧表

リズム指標	意味	出典
PVI	母音・子音の持続時間の変動	Grabe and Low (2002)
DeltaC	子音区間の標準偏差	Ramus+ (1999)
DeltaV	母音区間の標準偏差	Ramus+ (1999)
%V	全発話時間のうち母音区間 が占める割合	Ramus+ (1999)
VarcoC	子音区間の変動係数 × 100	Ferragne and Pellegrino (2004)
VarcoV	母音区間の変動係数 × 100	Ferragne and Pellegrino (2004)

## 2.3 問題点

Yamamoto and Kawabata の研究と、一ノ瀬と能田の研究では日本語を読み上げたときに遅延量 200 ms の DAF 下で発話速度が最も低下することがわかった。また、Yamamoto and kawabata の研究では、遅延量 200 ms の DAF 下で平均モーラ区間が最も長くなることもわかった。しかし、発話全体のモーラ区間が長くなったのか、一部のモーラ区間が長くなったのか明らかにされていなく、どちらの場合でも平均モーラ区間は長くなる結果となる。さらに、平均モーラ区間が長くなることによって、モーラの等時性に変化があるかは言及されていなかった。

Dellwo and Wagner の研究と、Dellwo の研究ではリズム指標の DeltaC と %V を用いて、異なる発話速度で読み上げたときの英語・ドイツ語・フランス語の音声を言語内と言語間でリズム分類可能であるかを分析した。結果は、DeltaC は言語内と言語間で分類できたが、%V は言語内では分類できず、言語間で分類ができることがわかった。また、Dellwo の研究では、リズム指標の VarcoC も同様に分析したが、言語内では分類できる言語もあれば分類できない言語もあることがわかった。また、言語間では分類できることがわかった。これらから、3つの言語は発話速度の違いで DeltaC が変化することがわかり、VarcoC は発話速度の違いで変化する言語もあることがわかった。しかし、この3つの言語は強勢拍か音節拍に分類される [13][14][15] 発話リズムであるため、発話速度などの個人差があるものに対して日本語のモーラ拍リズムに変化があるのかが言及されていなかった。

## 第3章 本研究の方略

### 3.1 方法論

本研究では聴覚フィードバック情報の遅延時間が発話リズムにどのように影響するのかを解明するために図 3.1 に示した方略に沿って進める。まず、発話リズムの定義を行い、発話リズムを定量化する方法を調査する。本研究では発話リズムを音声セグメントの音素・音節・モーラが等時性を持つ現象として定義するため、音声セグメントによって発話リズムを定量化する方法を調査する。調査により、発話リズムを定量化する方法として音素区間を用いたリズム指標 [7][47] を採用する。次に、調査した方法が日本語音声においてどのように変化するかを分析する。分析する前に日本語特有の特殊拍が音素区間に及ぼす影響を調べる。例として図 3.2 に撥音の有無においての子音区間のヒストグラムの例を示す。上の図は撥音を含めたときの子音区間のヒストグラムで、下の図は撥音を含めないときの子音区間のヒストグラムである。横軸が子音区間の長さ、縦軸が頻度値を示している。また、青い部分が発話速度が速いときの発話音声に含まれる子音区間で、赤い部分が発話速度が遅いときの発話音声に含まれる子音区間である。図 3.2 より、発話速度が速い方を見ると、特殊拍である撥音は他の子音区間より 50 ms 以上長いことがわかった。また、発話速度が遅くなると、撥音以外の子音区間は変化が見られないが、撥音は約 200 ms ほど長くなっていることがわかった。撥音を含めた特殊拍は発話速度の低下で大幅に長くなることから、DAF 下で発話速度が低下する [5][11] と仮定すると、DAF 下での平均子音区間が特殊拍によって大幅に長くなりリズム指標に影響を及ぼす可能性があると考えられる。したがって、撥音などの特殊拍は別で分析する必要があることがわかった。次に、日本語音声はモーラ単位で表されるため、モーラ単位で発話リズムを定量化する方法を考案する。次に、調査した方法と考案した方法がフィードバック遅延量に応じてどのように変化するかを調べるために、発話実験を行い、実験で得られた音声データを分析する。分析結果から統計的検定を用いて評価する。

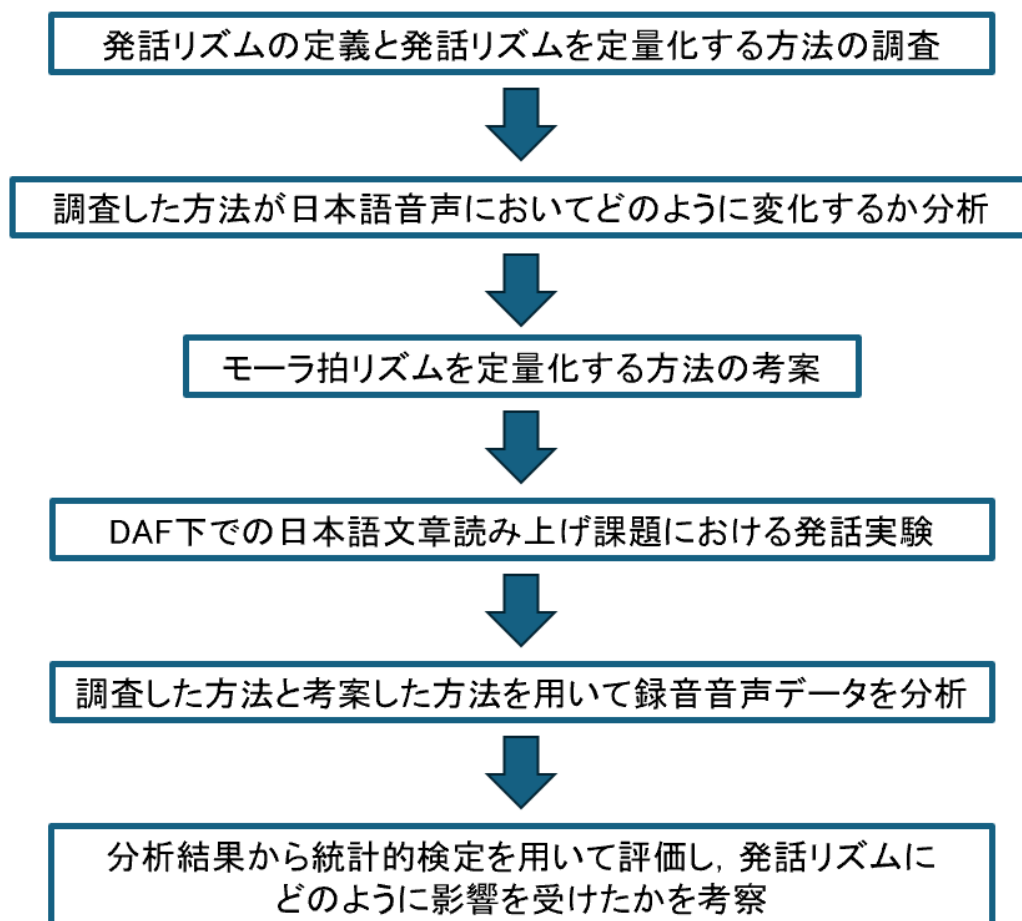
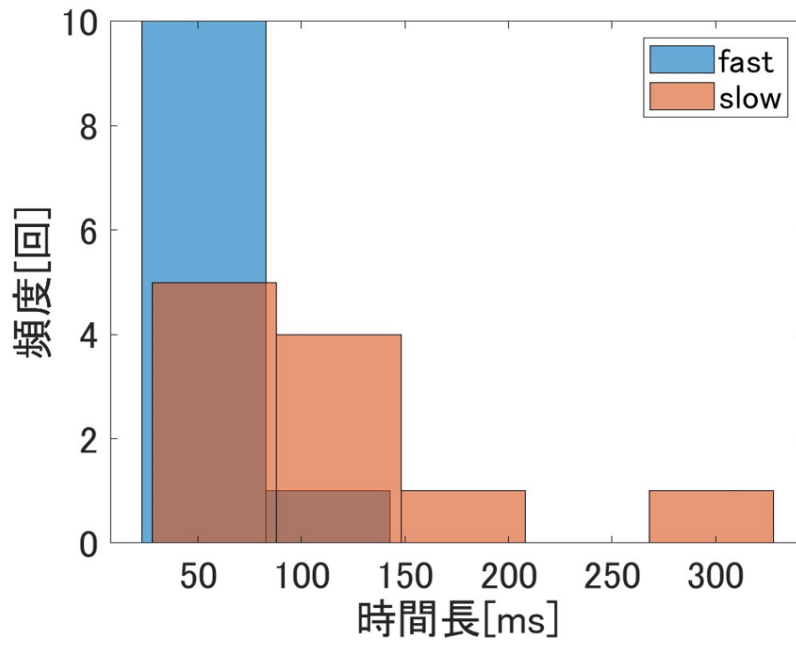


図 3.1: 本研究の方略

## 撥音あり



## 撥音なし

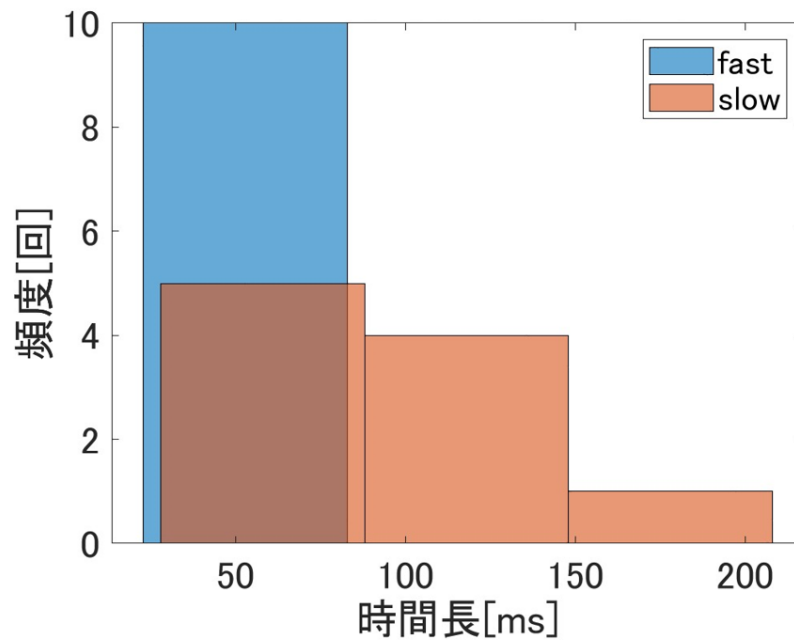


図 3.2: 撥音の有無における子音区間のヒストグラム



## 3.2 リズム指標の定義

先行研究より，リズムは等時性周期の単位継続時間に基づくとし [12][23][30]，音声生成のリズムを「音声セグメントの音素・音節・モーラが等時性を持つ現象」として定義する [6]．また，本研究では音声生成のリズムを発話リズムとして定義する．本研究では日本語を用いるため，モーラが等時性を持つ発話リズムであるモーラ拍リズムを分析し，フィードバック遅延量によってモーラ拍リズムがどのように影響を受けるのかをモーラの等時性で評価する．遅延量なしのときのモーラ区間を基準として，遅延量の増加でモーラ区間を分析することでモーラが等時性を持っているかを評価する．本研究では，モーラ区間の長さがどのように変化するかを音素単位で調べるために音素区間を用いたリズム指標を用いる．本研究で用いる音素区間のリズム指標について，子音区間の標準偏差を  $\Delta C$ ，子音区間の変動係数を  $\text{VarcoC}$ ，母音区間の標準偏差を  $\Delta V$ ，母音区間の変動係数を  $\text{VarcoV}$  とする． $\Delta C$  を  $D_C$  として式 (3.1) に示す [7]．

$$D_C = \sqrt{\frac{\sum_{i=1}^{N_C} (x_{C_i} - \bar{x}_C)^2}{N_C - 1}} \quad (3.1)$$

$N_C$  は子音区間の数， $x_{C_i}$  は  $i$  番目の子音区間の値， $\bar{x}_C$  は平均子音区間を示す．次に  $\text{VarcoC}$  を  $V_C$  として式 (3.2) に示す [48]．

$$V_C = \frac{D_C}{\bar{x}_C} \times 100 \quad (3.2)$$

$\text{VarcoC}$  は各遅延量の子音区間を相対的に観測するために分析する．次に  $\Delta V$  を  $D_V$  として式 (3.3) に示す [7]．

$$D_V = \sqrt{\frac{\sum_{i=1}^{N_V} (x_{V_i} - \bar{x}_V)^2}{N_V - 1}} \quad (3.3)$$

$N_V$  は母音区間の数， $x_{V_i}$  は  $i$  番目の母音区間の値， $\bar{x}_V$  は平均母音区間を示す．次に  $\text{VarcoC}$  を  $V_V$  として式 (3.4) に示す [47]．

$$V_V = \frac{D_V}{\bar{x}_V} \times 100 \quad (3.4)$$

$\text{VarcoV}$  も同様に各遅延量の母音区間を相対的に観測するために分析する．音素区間を用いたリズム指標において，日本語特有の特殊拍が平均子音区間を大幅に長くし，音素区間のリズム指標に影響を及ぼす可能性があることがわかった．それらを考慮しないようにするためにモーラ単位でも調べる必要がある．そのためにモーラ区間のリズム指標を考案する．本研究で用いるモーラ区間のリズム指標について，モーラ区間の標準偏差を  $\Delta M$ ，モーラ区間の変動係数を  $\text{VarcoM}$  とする． $\Delta M$  を式 (3.5) のように定義する．

$$D_M = \sqrt{\frac{\sum_{i=1}^{N_M} (x_{Mi} - \bar{x}_M)^2}{N_M - 1}} \quad (3.5)$$

$D_M$  は DeltaM,  $N_M$  はモーラ区間の数,  $x_{Mi}$  は  $i$  番目のモーラ区間の値,  $\bar{x}_M$  は平均モーラ区間を示す. 次に VarcoM を式 (3.6) のように定義する.

$$V_M = \frac{D_M}{\bar{x}_M} \times 100 \quad (3.6)$$

$V_M$  は VarcoM を示す. また, VarcoM も同様に各遅延量のモーラ区間を相対的に観測するために分析する. 以上の式を本研究ではリズム指標として用いる.

### 3.3 発話リズムの評価

本研究では, 3.2 で説明したリズム指標がフィードバック遅延量に応じてどのような変化が出るかを観測するために, DAF 下での日本語文章読み上げ課題における発話実験を行う. 発話実験で得られた録音音声データから音素・モーラ区間を抽出し, 各区間のリズム指標で発話リズム分析を行う. 発話リズム分析では, まず各実験参加者の発話リズムが DAF の遅延量によってどのような変化があるのかを確認するため, 実験参加者内分析を行う. 次に各遅延量でのリズム指標の平均をとり, 各遅延量の平均値を比較するために実験参加者間分析を行う. 分析結果から発話リズムを評価するために統計的検定を用いる. 評価方法として, はじめに実験参加者間分析結果により, 全ての遅延量においてリズム指標の平均値に差があるかを検定する. 次に, 遅延量なしのときと他の遅延量のときのリズム指標の平均値に差があるかを多重比較で検定する. 実験参加者内分析結果と検定結果からリズム指標がフィードバック遅延量に応じてどのように変化したかを調べる. これにより, DAF 下での日本語文章読み上げ課題における発話リズムを評価する.

# 第4章 DAF下での日本語文章読み上げ課題における発話実験

## 4.1 実験目的

音素・モーラ区間を用いたリズム指標がフィードバック遅延量に応じてどのような変化が出るかを観測するために、DAF下での日本文章読み上げ課題における発話実験を実施する。実験結果として、先行研究 [5] より遅延量の増加で平均モーラ区間が長くなることが予想される。また、モーラは子音と母音のセグメント、母音のみ、子音のみの3種類で構成されるため、平均モーラ区間が長くなることで平均子音区間と平均母音区間も長くなることが予想される。先行研究 [9][10] より、英語・ドイツ語・フランス語の発話音声は発話速度の増加に対して  $\Delta C$  が小さくなる傾向が見られたことが知られている。これらからすべての言語において  $\Delta C$  が小さくなるとすると、日本語でも発話速度の増加に対して  $\Delta C$  が小さくなると予想される。また、遅延量の増加で発話速度が低下すると仮定し、遅延量の増加で  $\Delta C$  が大きくなると予想される。さらに、 $\Delta C$  が大きくなることで子音と母音のセグメント、子音のみで構成される2種類のモーラ区間もばらつきが大きくなり、 $\Delta M$  が大きくなると予想される。[10] より、英語とドイツ語では発話速度の増加に対して、 $\Delta C$  が小さくなることがわかった。また、フランス語では発話速度の増加に対して、 $\Delta C$  が大きくなることがわかった。よって、日本語は発話速度に対して  $\Delta C$  がどちらかの挙動を見せると考えられる。このことから、異なる遅延量の DAF によってモーラの等時性を持った発話リズムがどのような影響を受けるのかを音素・モーラ区間を用いたリズム指標で分析し、リズム指標の変化量で発話リズムの評価を行う。

## 4.2 実験条件

先行研究 [5][11] では、遅延量 200 ms で発話への影響が最も大きくなる傾向を示すことが知られている。本研究では、[5][11] と同様に遅延量 200 ms 以外の時の発話への影響も調査するために遅延量 0 ms, 100 ms, 200 ms, 300 ms の4条件を設定した。また、提示文章は ATR 音素バランス文から1文約30モーラの文章を10文選定した。提示文章の一覧表を表 4.1 に示す。

表 4.1: 提示文章の一覧表

文章番号	文章内容
1	ちょっと遅い昼食をとるため、ファミリーレストランに入ったのです。
2	言論の自由は、一步譲れば百歩も千歩も攻めこまれる。
3	老人ホームの場合は、健康器具やひざ掛けだ。
4	パンまたは肉入りのおかゆ、スープ、ソーセージと野菜、コーヒーというメニューだ。
5	午前八時、健康な捕虜(ほりょ)は作業所へトラックで出発する。
6	叔父さんは岬(みさき)の一軒家に、独りぼっちで住んでいた。
7	立春がすぎても厳しい寒さの日々が続く。
8	最初早いテンポで巻きこんでおいて、中盤じっくり見せる。
9	広い構内にレストランや列車食堂などを配(はい)し、札幌ビール園に似た雰囲気だ。
10	あんな物は新聞や雑誌の白いスペースを埋めるための砂利だ。

### 4.3 実験参加者

実験には発話と聴力に問題のない日本語母語話者 10 名（男性 8 名，女性 2 名，平均  $24.7 \pm 1.85$ ，23-29 歳）が参加した．また，実験前に発話の練習をいずれの話者も数回実施し，問題なく発話可能なことを確認した．

### 4.4 実験装置

実験は暗騒音  $< 30$  dB の北陸先端大の防音室内で実施した．フィードバック音声を聴取するためのヘッドホン（Sennheiser HD 280）と発話音声を集音するためのマイク（Sennheiser MKH 416 P48U3）を使用した．また，オーディオインターフェース（Roland Quad Capture）を使用した．遅延量と文章の提示について，LaptopPC（Lenovo Windows11）上のソフトウェア（Mathworks MATLAB 2022a）および DAF システム（Audapter）[46] を用いて提示する．発話音声の録音は PC 上の録音ソフトウェア（Audacity）を使用．また，マイクとヘッドホンの音量について，人工耳と騒音計を使用してマイクの入力とヘッドホンの出力を調整した．本研究では，マイクの入力を 80 dB，ヘッドホンの出力を 85 dB として設定した．マイクと発話者の距離について，拳一つ分空けるように指示した．音圧レベルの設定について，実際にフィードバック音声を聞き，matlab 上で音圧を 1.3dB に設定した．図 4.1 に実験環境の概略図を示す．

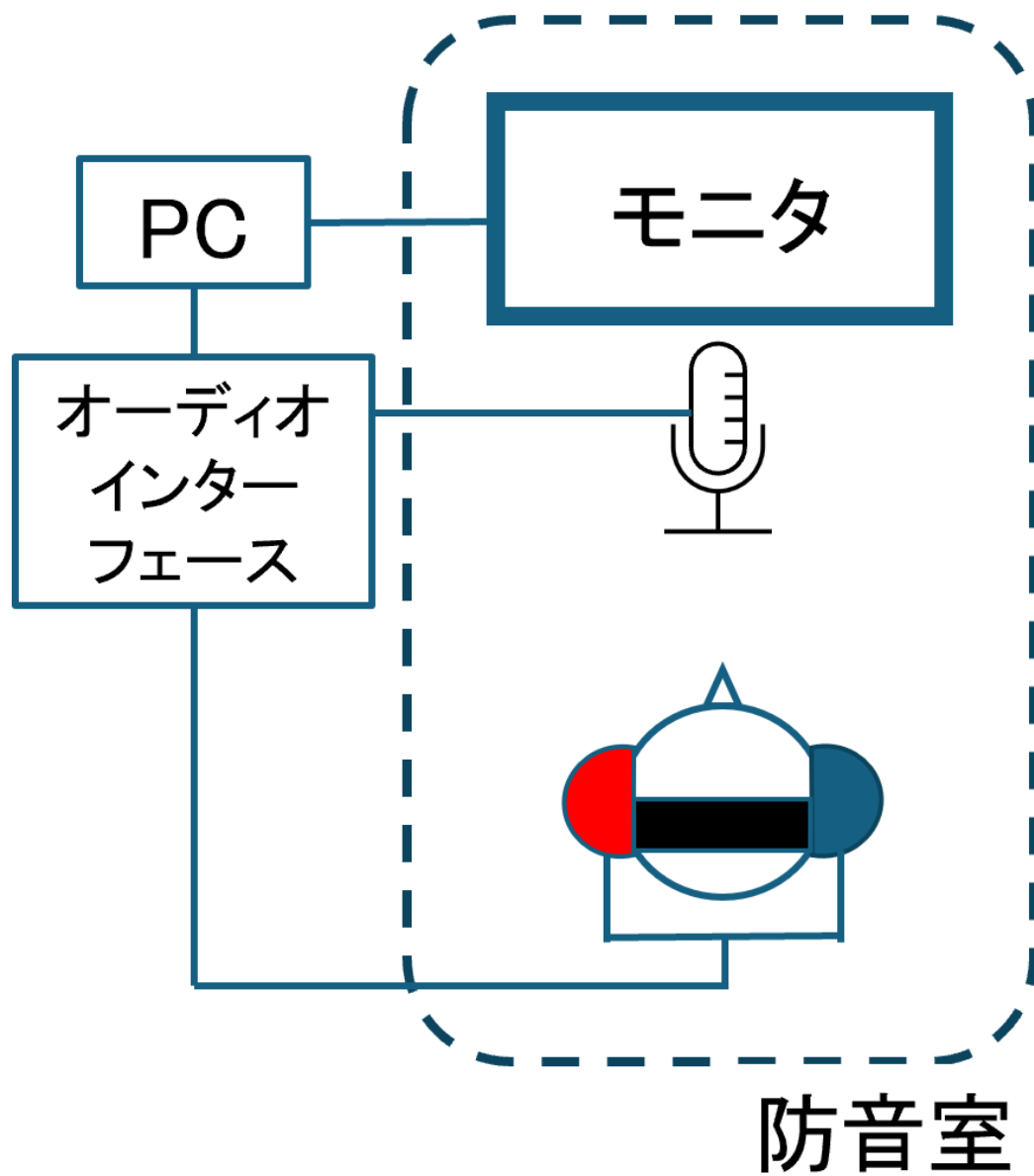


図 4.1: 実験環境の概略図

## 4.5 実験方法

図 4.2 に実施した実験手順の概略を示す。本実験では日本語文章を読み上げる課題を実施した。提示文章 1 文に対して 1 条件の遅延量を提示した。課題開始前に練習を 2 回実施し、練習終了ごとに 2 分の休憩を入れた。1 回目の練習ではヘッドホンを着用せずに、提示文章 10 文をランダムで 1 文ずつ読み上げるように指示し、2 回目の練習はヘッドホンを着用して読み上げるように指示した。練習終了後に課題を実施した。課題には 1 文 1 条件を 1 Trial として合計 8 Trial あり、これを 1 ブロックとしてまとめて実施した。実験では 5 ブロック実施し、ブロック間で 10 分間の休憩を入れた。1 Trial 20 秒間あり、その 20 秒間でモニタに表示された 1 文を読み上げるように指示した。Trial 終了ごとに 10 秒の待機時間を入れ、待機時間中はモニタには何も表示せず、実験参加者には音を出さずに待機するように指示した。

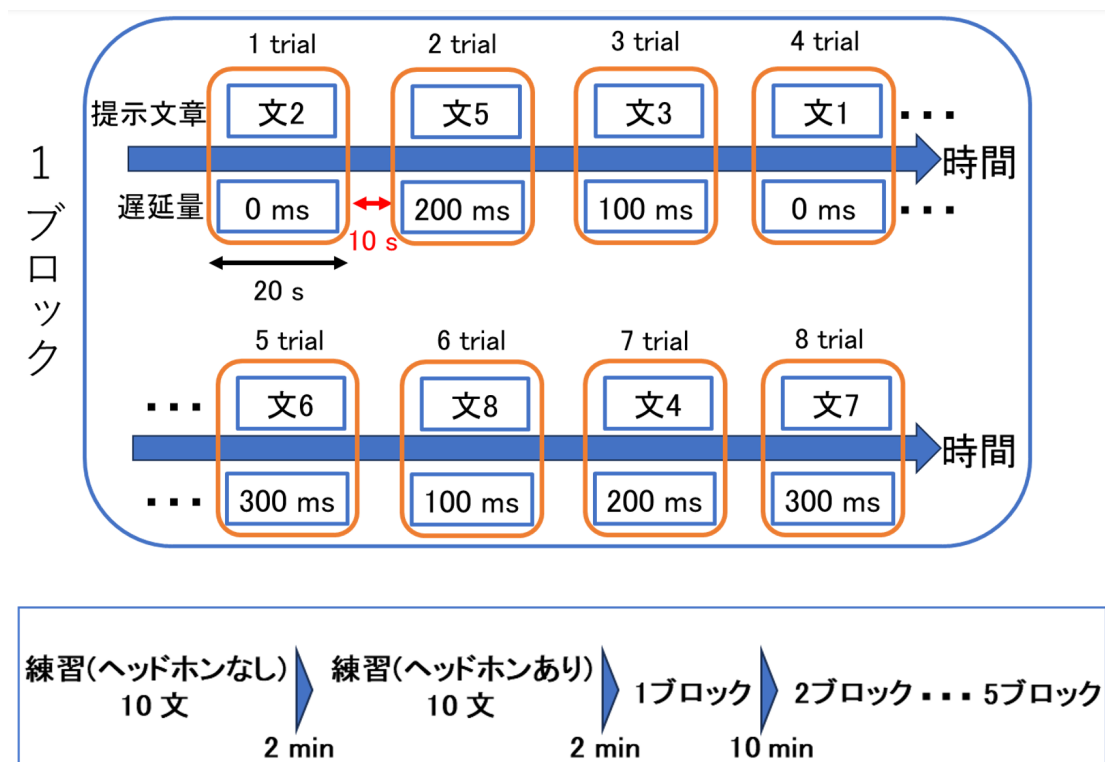


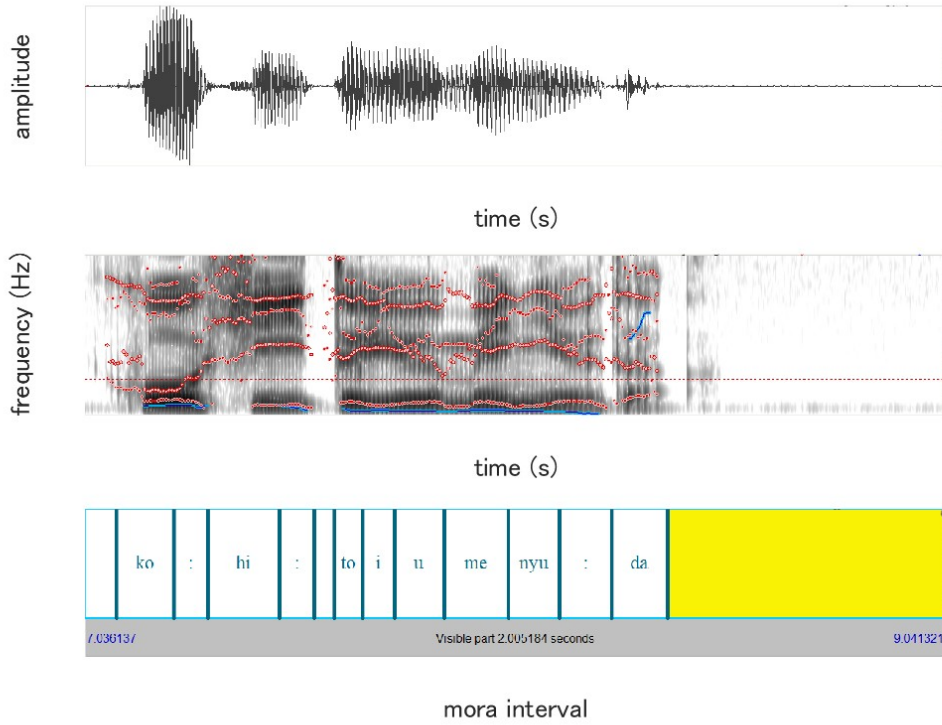
図 4.2: 実験手順の概略図



## 4.6 実験結果

図 4.3 に遅延量なしのときと遅延量ありのときのモーラ区間の例を示す。図より、1 番目のグラフでは横軸が発話時間、縦軸が振幅を示す。2 番目のグラフでは横軸が発話時間、縦軸が周波数を示す。3 番目の図では 1 番目、2 番目のグラフと同じ時間軸のモーラ区間を示す。図 4.4 に遅延量なしのときと遅延量ありのときの音素区間の例を示す。図より、1 番目のグラフでは横軸が発話時間、縦軸が振幅を示す。2 番目のグラフでは横軸が発話時間、縦軸が周波数を示す。3 番目の図では 1 番目、2 番目のグラフと同じ時間軸の音素区間を示す。また、区間の長さの比較を視覚的に捉えやすくするために、発話開始から約 2 秒間のところで切り取った結果を示す。

### 遅延量なし



### 遅延量あり

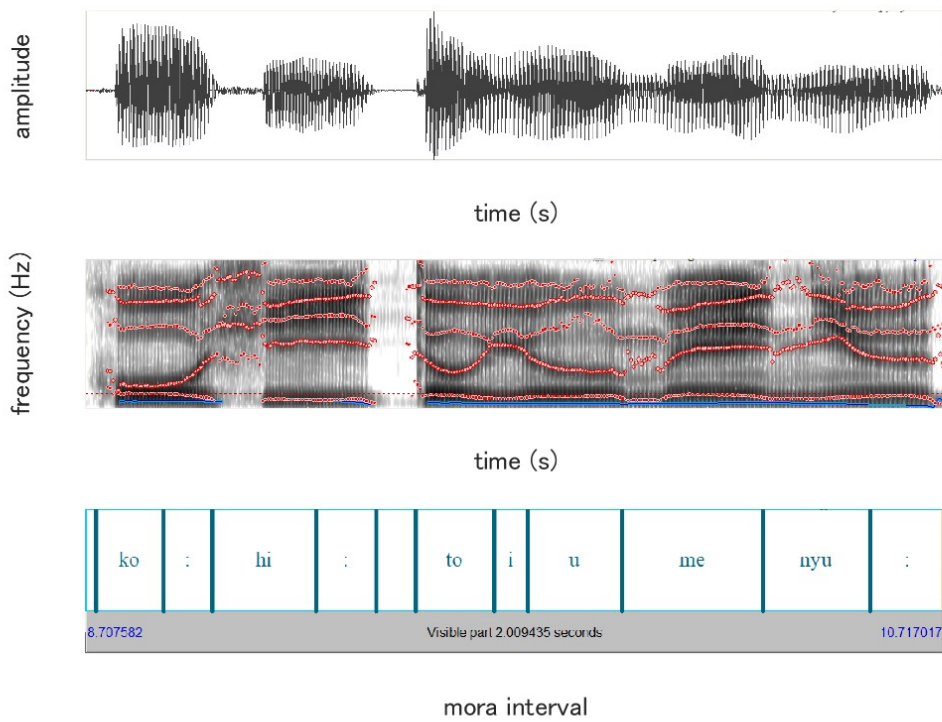
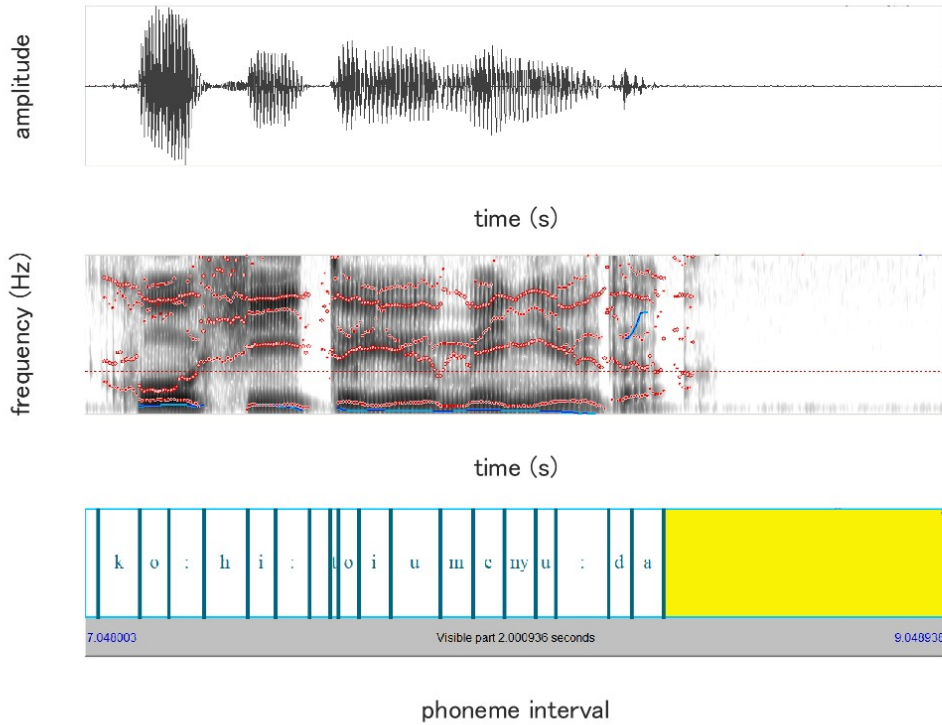


図 4.3: 遅延量なしのときと遅延量ありのときのモーラ区間の例

### 遅延量なし



### 遅延量あり

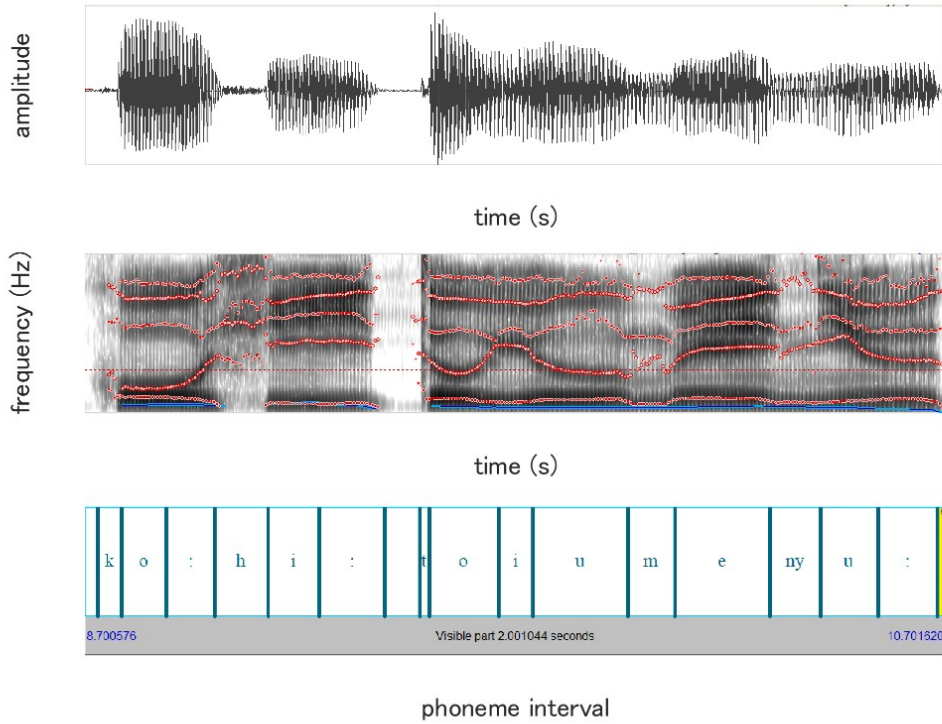


図 4.4: 遅延量なしのときと遅延量ありのときの音素区間の例

## 4.7 分析方法

発話音声から音素・モーラの区間を分析するために、音素・モーラ単位で時間情報のアノテーションを行う。本研究における時間情報のアノテーションの方法は以下の通りである。

### 4.7.1 アノテーション

アノテーションは音声分析ソフトの Praat と音声認識ツールの Julius を用いた。まず、Julius を使ってラベリングを自動的に付与した。自動ラベリングは無音区間も文字として判断されラベリングされるため、Praat 上で手動調整する [31]。図 4.5 に自動ラベリングの手順の概略図を示し、内容を下記に示す。

1. 発話実験で録音した音声データを 1 文ごとに切り取り、wav ファイルに保存する。次にその 1 文の文章内容を text ファイルにひらがなで入力し、保存する。
2. Julius を用いて wav ファイルと text ファイルから lab ファイルと log ファイルを生成する。
3. Python で lab ファイルから TextGrid ファイルに変換し、wav ファイルと TextGrid ファイルを Praat で同時表示させることで自動ラベリングを行う。

次に、手動でのモーラ単位のラベリングの調整方法について説明する。本研究では、カナ 1 文字をモーラとしてカウントしているため、日本語特有の特殊拍（促音、撥音、拗音、長音）は 1 モーラとする。また、無声化された音声も 1 モーラとする。モーラ間の区別は音声を聞きながら基本周波数とフォルマント周波数の位置で行った。次に手動での音素単位のラベリングの調整方法について説明する。自動ラベリングではモーラ単位でラベリングされるため、音素単位で分析する場合は手動でモーラを子音と母音に分ける必要がある。子音と母音の区別は、音声を聞きながら基本周波数とフォルマント周波数の位置を確認して行った。特殊拍の促音・撥音と無声化された音声は Praat 上で子音と母音の区別ができないため、全て子音として扱うと音素区間を用いたリズム指標に影響を及ぼす可能性があることがわかった。本研究ではそれらを省いて分析を行う。また、特殊拍の拗音は子音として扱い、長音は母音として扱う。

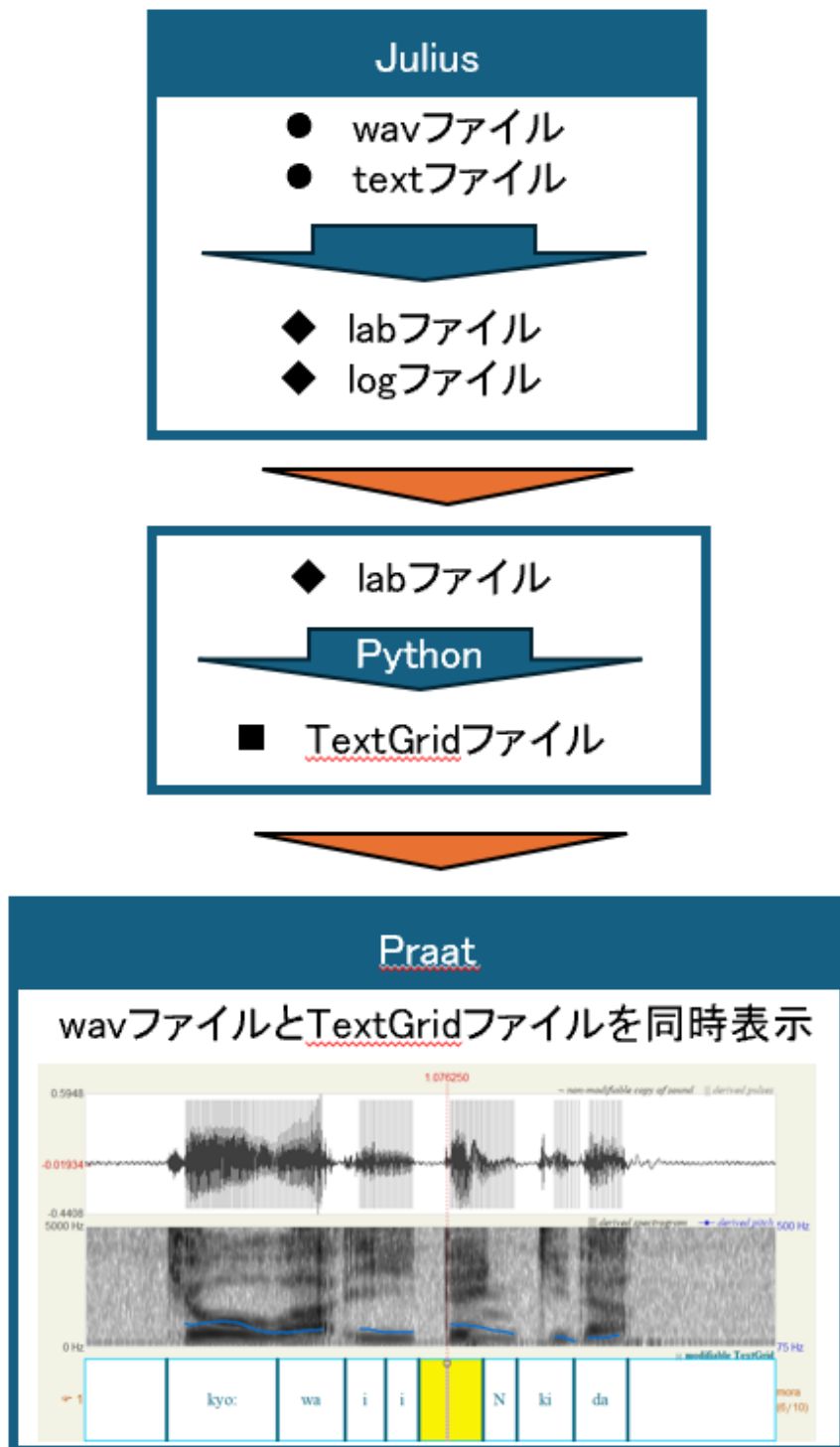


図 4.5: 自動ラベリングの手順の概略図

## 4.7.2 発話リズム分析

次にリズム指標を用いた発話リズム分析の方法について説明する。はじめに、各実験参加者の発話リズムがDAFの遅延量によってどのような変化があるのかを確認するために実験参加者内分析を行う。実験参加者内分析の方法は、各実験参加者の発話音声ごとに音素・モーラの平均区間とリズム指標を求め、その値を実験参加者ごとで平均した。その平均値を各遅延量でプロットした。次に、各遅延量でリズム指標の平均をとり、全ての遅延量においてリズム指標の平均に差があるかの確認と、遅延量0 msと他の遅延量の平均に差があるかを確認するために実験参加者間分析と統計的検定により評価する。実験参加者間分析の方法は、実験参加者ごとの平均値を遅延量ごとで平均し、その平均値を各遅延量でプロットした。反復測定分散分析(Repeated Measures Analysis of Variance:RM-ANOVA)で各遅延量の条件間に差が存在するかを検定した。また、Dunnett検定により遅延量0 msを対象群としたときの各条件との差を検定した [21]。

## 4.8 分析結果

### 4.8.1 実験参加者内分析結果

図 4.6 に各実験参加者の遅延量ごとにおける平均子音区間の結果を示す。横軸は各遅延量、縦軸は各遅延量の全ての平均子音区間を平均した値を示す。また、各プロットは各実験参加者のデータを指す。表 4.2 に各実験参加者の平均子音区間の値を示す。図 4.6 より、遅延量の増加で平均子音区間は長くなることがわかった。また、Subject H, I は遅延量の増加で平均子音区間が大きく変化し、遅延量 300 ms で平均子音区間が最も長くなる傾向があった。他の実験参加者は遅延量 100 ms で平均子音区間が最も長くなる傾向があったが、遅延量 200 ms で平均子音区間が短くなる傾向が見られた。また、遅延量 300 ms で平均子音区間が長くなる傾向も見られ、短くなる傾向も見られた。

図 4.7 に各実験参加者の遅延量ごとにおける平均母音区間の結果を示す。横軸は各遅延量、縦軸は各遅延量の全ての平均母音区間を平均した値を示す。また、各プロットは各実験参加者のデータを指す。表 4.3 に各実験参加者の平均母音区間の値を示す。遅延量の増加で平均母音区間は長くなることがわかった。また、Subject H, I は遅延量の増加で平均母音区間が大きく変化する傾向が見られた。他の実験参加者は遅延量の増加で大きな変化は見られなかったものの、表 4.3 を見ると遅延量 100 ms で平均母音区間が長くなる傾向が見られ、遅延量 200 ms で平均母音区間が短くなる傾向が見られた。また、遅延量 300 ms で平均母音区間が長くなる傾向も見られ、短くなる傾向も見られた。

図 4.8 に各実験参加者の遅延量ごとにおける平均モーラ区間の結果を示す。横軸は各遅延量、縦軸は各遅延量の全ての平均モーラ区間を平均した値を示す。ま

た、各プロットは各実験参加者のデータを指す。表 4.4 に各実験参加者の平均モーラ区間の値を示す。図 4.8 より、遅延量の増加で平均モーラ区間が長くなることがわかった。また、Subject H, I は遅延量の増加で平均モーラ区間が大きく変化し、遅延量 300 ms で最も長くなる傾向があった。他の実験参加者は遅延量の増加で大きな変化は見られないが、表 4.4 を見ると遅延量 100 ms で平均モーラ区間が最も長くなる傾向が見られ、遅延量 200 ms で平均母音区間が短くなる傾向が見られた。また、遅延量 300 ms で平均モーラ区間が長くなる傾向も見られ、短くなる傾向も見られた。

図 4.9 に各実験参加者の遅延量ごとにおける DeltaC の結果を示す。横軸は各遅延量、縦軸は各遅延量の全ての DeltaC を平均した値を示す。また、各プロットは各実験参加者のデータを指す。表 4.5 に各実験参加者の DeltaC の値を示す。図 4.9 より、遅延量の増加で DeltaC が大きくなることがわかった。また、Subject H, I は遅延量の増加で DeltaC が大きく変化し、遅延量 300 ms で最も大きくなる傾向が見られた。他の実験参加者は大きな変化は見られなかったものの、表 4.5 を見ると遅延量 100 ms で DeltaC が最も大きくなる傾向も見られ、遅延量 200 ms でも最も大きくなる傾向も見られた。また、遅延量 300 ms で DeltaC が大きくなる傾向も見られたが、半数以上は小さくなる傾向が見られた。

図 4.10 に各実験参加者の遅延量ごとにおける DeltaV の結果を示す。横軸は各遅延量、縦軸は各遅延量の全ての DeltaV を平均した値を示す。また、各プロットは各実験参加者のデータを指す。表 4.6 に各実験参加者の DeltaV の値を示す。図 4.10 より、遅延量の増加で DeltaV は大きくなることがわかった。また、半数の実験参加者は遅延量の増加で DeltaV が大きく変化し、遅延量 200 ms で最も大きくなる傾向が見られた。他の実験参加者は遅延量の増加で大きな変化は見られなかったものの、表 4.6 を見ると遅延量 100 ms で DeltaV が大きくなる傾向が見られた。さらに過半数の実験参加者は遅延量 300 ms で DeltaV が小さくなる傾向が見られた。

図 4.11 に各実験参加者の遅延量ごとにおける DeltaM の結果を示す。横軸は各遅延量、縦軸は各遅延量の全ての DeltaM を平均した値を示す。表 4.7 に各実験参加者の DeltaM の値を示す。図 4.11 より、遅延量の増加で DeltaM は大きくなる傾向があった。また、Subject H, I, G は遅延量の増加で DeltaM が大きく変化し、遅延量 300 ms で最も大きくなる傾向があった。他の実験参加者は、遅延量の増加で大きな変化は見られなかったものの、表 4.7 を見ると遅延量 100 ms で最も大きくなる傾向も見られ、遅延量 300 ms で最も大きくなる傾向も見られた。

図 4.12 に各実験参加者の遅延量ごとにおける VarcoC の結果を示す。横軸は各遅延量、縦軸は各遅延量の全ての VarcoC を平均した値を示す。表 4.21 に各実験参加者の VarcoC の値を示す。図 4.12 より、遅延量の増加で VarcoC は大きくなることがわかった。また、過半数の実験参加者は遅延量の増加で VarcoC が大きく変化した。その中で遅延量 300 ms や遅延量 100 ms で最も大きくなる傾向も見られたが、遅延量 200 ms で最も大きくなる傾向が多かった。

図 4.13 に各実験参加者の遅延量ごとにおける VarcoV の結果を示す。横軸は各

遅延量，縦軸は各遅延量の全ての VarcoV を平均した値を示す．表 4.9 に各実験参加者の VarcoV の値を示す．図 4.13 より，遅延量の増加で VarcoV が大きくなることがわかった．また，半数の実験参加者は遅延量の増加で VarcoV が大きく変化し，遅延量 200 ms で最も大きくなる傾向が見られた．他の実験参加者は，遅延量の増加で大きな変化は見られなかったものの，表 4.9 を見ると，遅延量 100 ms で最も大きくなる傾向も見られ，遅延量 100 ms で最も小さくなり，遅延量 300 ms で最も大きくなる傾向も見られた．

図 4.14 に各実験参加者の遅延量ごとにおける VarcoM の結果を示す．横軸は各遅延量，縦軸は各遅延量の全ての VarcoM を平均した値を示す．表 4.10 に各実験参加者の VarcoM の値を示す．図 4.14 より，遅延量の増加で VarcoM が大きくなることがわかった．また，一部の実験参加者は遅延量の増加で VarcoM が大きく変化し，遅延量 300 ms で最も大きくなった．また，半数の実験参加者は遅延量 100 ms で VarcoM が小さくなり，遅延量 200 ms で大きくなる傾向が見られた．



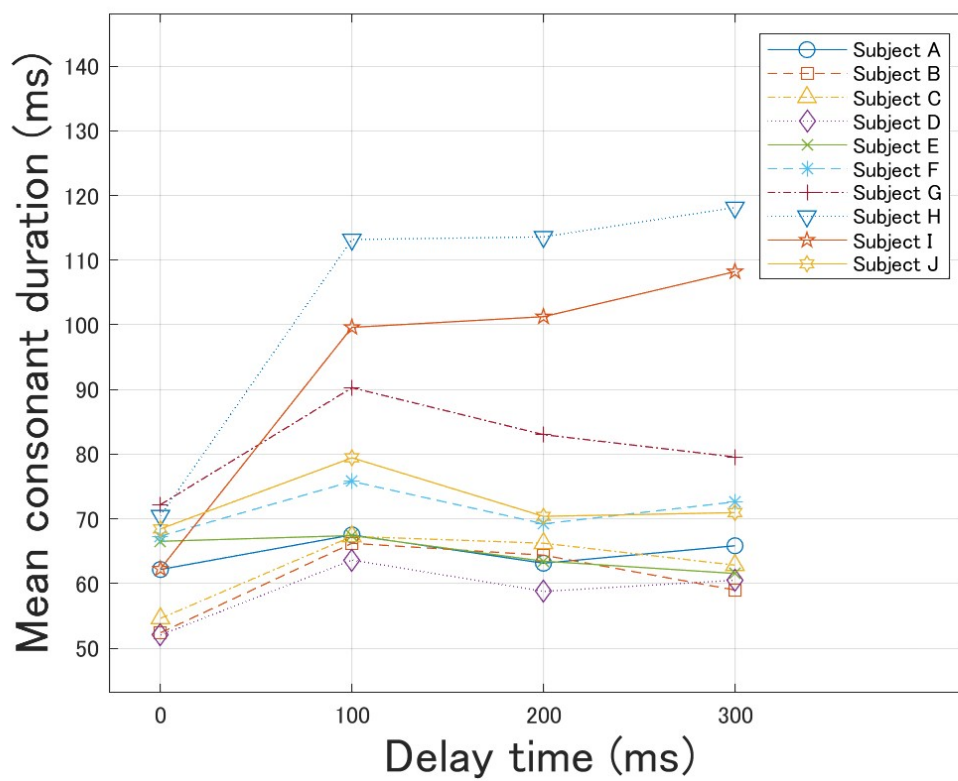


図 4.6: 各実験参加者の遅延量ごとにおける平均子音区間

表 4.2: 各実験参加者の遅延量ごとにおける平均子音区間の値

実験参加者	遅延量			
	0 (ms)	100 (ms)	200 (ms)	300 (ms)
A	62.185	67.531	63.153	65.848
B	52.388	66.230	64.389	59.000
C	54.624	67.297	66.243	62.851
D	52.051	63.646	58.799	60.516
E	66.538	67.429	63.467	61.567
F	67.334	75.795	69.242	72.637
G	72.198	90.281	83.024	79.571
H	70.487	113.185	113.595	118.129
I	62.286	99.609	101.259	108.227
J	68.459	79.444	70.391	70.972

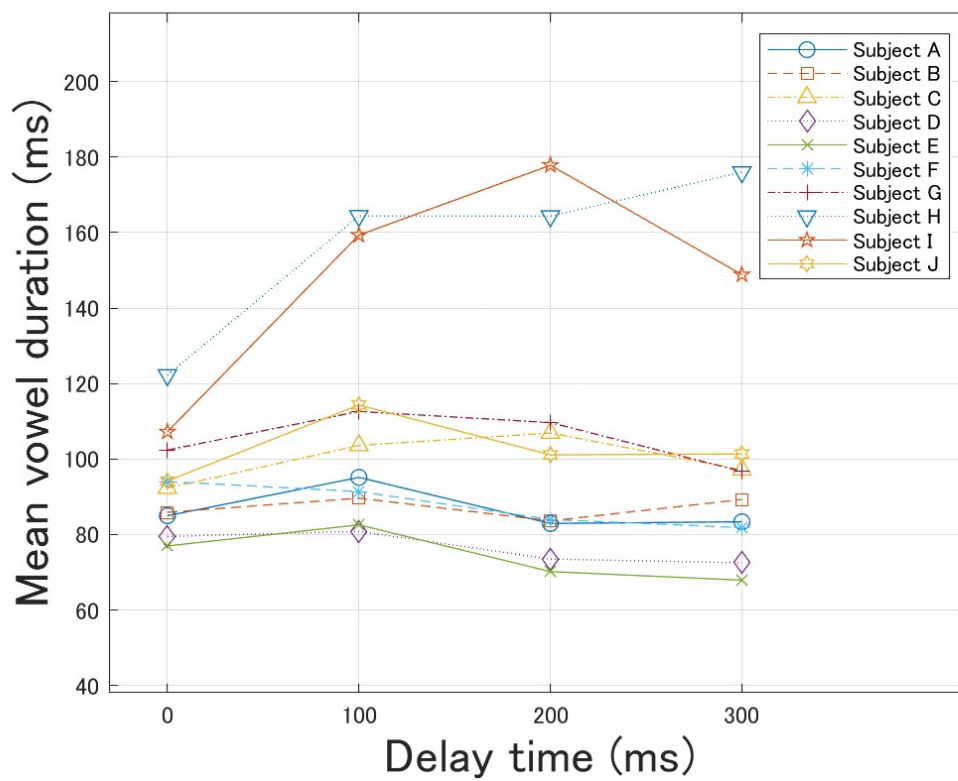


図 4.7: 各実験参加者の遅延量ごとにおける平均母音区間

表 4.3: 各実験参加者の遅延量ごとにおける平均母音区間の値

実験参加者	遅延量			
	0 (ms)	100 (ms)	200 (ms)	300 (ms)
A	84.972	95.132	82.916	83.367
B	85.879	89.619	83.607	89.263
C	92.300	103.603	106.888	97.144
D	79.459	80.824	73.493	72.500
E	76.964	82.550	70.162	67.892
F	94.020	91.374	83.968	81.824
G	102.367	112.601	109.641	96.692
H	122.239	164.414	164.348	176.012
I	107.203	159.339	177.833	148.848
J	94.158	114.303	101.104	101.328

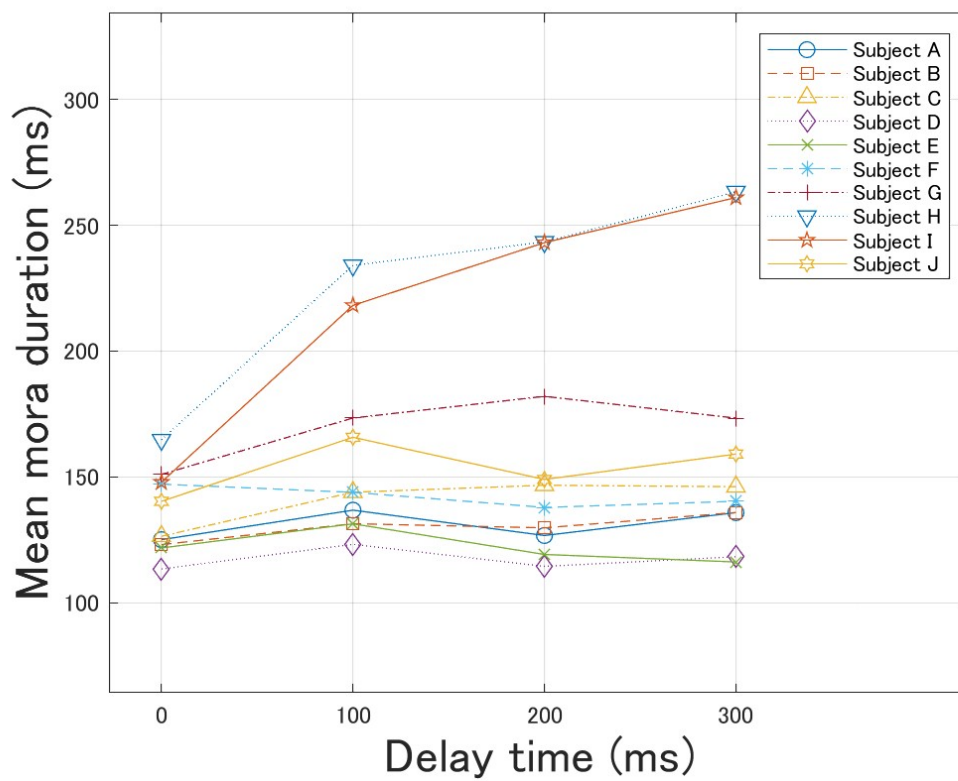


図 4.8: 各実験参加者の遅延量ごとにおける平均モーラ区間

表 4.4: 各実験参加者の遅延量ごとにおける平均モーラ区間の値

実験参加者	遅延量			
	0 (ms)	100 (ms)	200 (ms)	300 (ms)
A	125.047	136.769	126.704	135.796
B	123.091	131.395	129.746	135.852
C	126.409	143.903	146.683	146.094
D	113.409	123.192	114.440	118.261
E	121.757	131.342	119.172	116.181
F	147.121	143.849	137.864	140.306
G	150.987	173.398	182.019	173.359
H	164.739	234.071	243.384	263.330
I	147.885	218.173	243.128	261.030
J	140.287	165.733	148.995	159.068

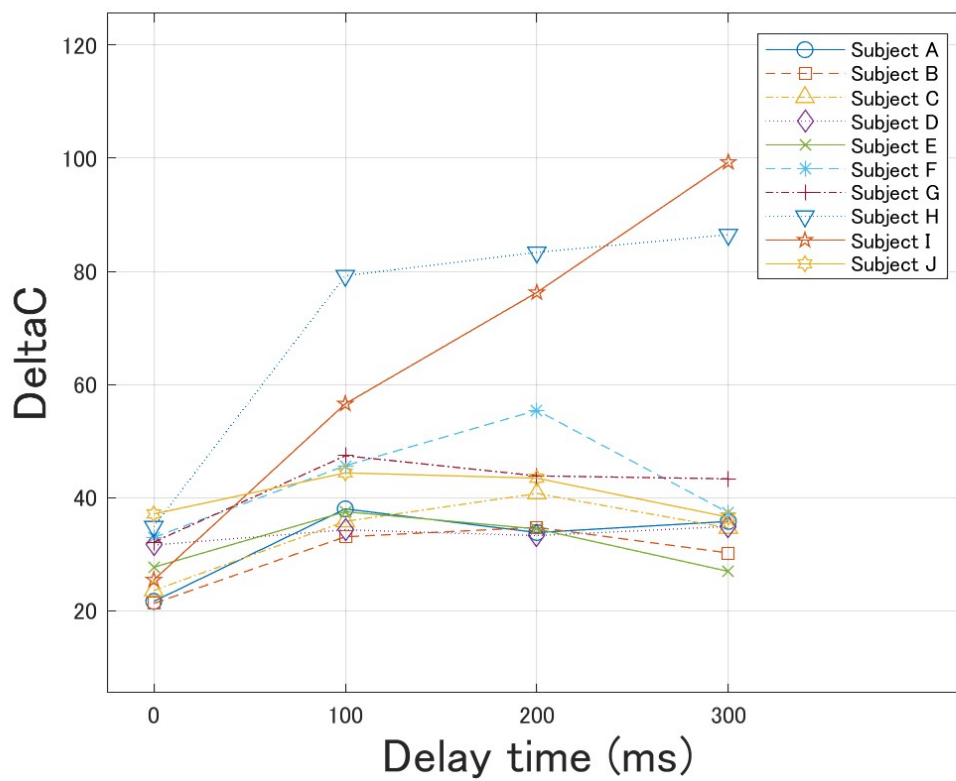


図 4.9: 各実験参加者の遅延量ごとにおける DeltaC

表 4.5: 各実験参加者の遅延量ごとにおける DeltaC の値

実験参加者	遅延量			
	0 (ms)	100 (ms)	200 (ms)	300 (ms)
A	21.696	38.028	33.856	35.816
B	21.323	33.099	34.775	30.238
C	23.619	35.820	40.747	34.651
D	31.650	34.330	33.309	34.925
E	27.733	37.531	34.491	26.993
F	32.984	45.661	55.443	37.351
G	32.168	47.421	43.859	43.340
H	34.884	79.219	83.349	86.466
I	25.478	56.587	76.263	99.269
J	37.128	44.397	43.456	36.564



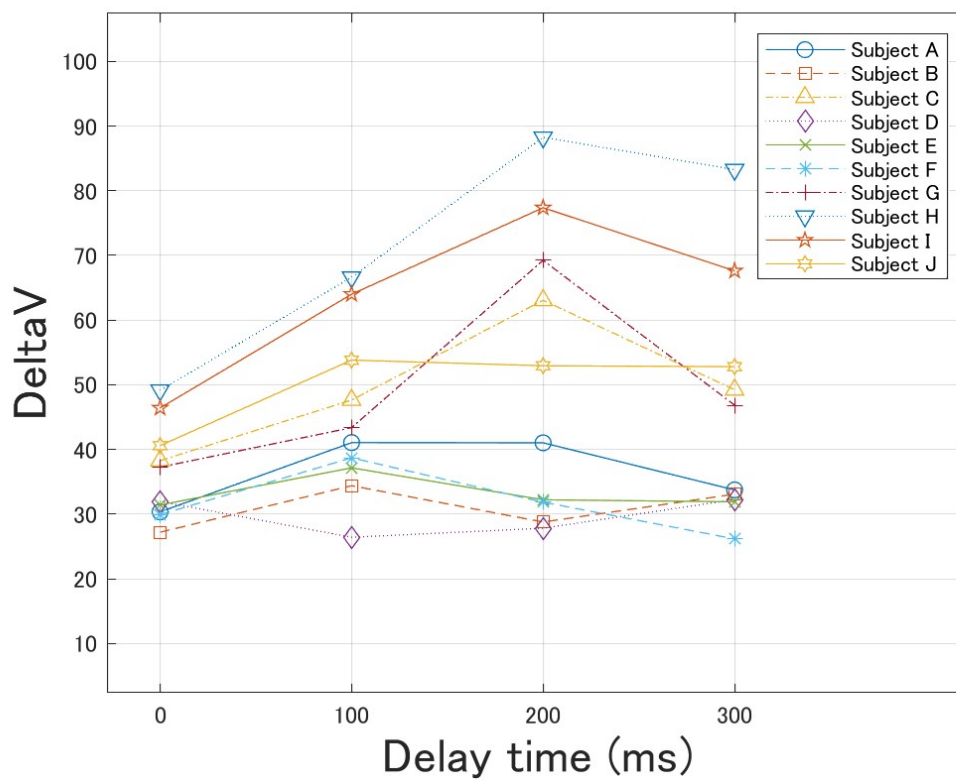


図 4.10: 各実験参加者の遅延量ごとにおける DeltaV

表 4.6: 各実験参加者の遅延量ごとにおける DeltaV の値

実験参加者	遅延量			
	0 (ms)	100 (ms)	200 (ms)	300 (ms)
A	30.355	41.065	41.034	33.751
B	27.211	34.401	28.790	33.096
C	38.274	47.648	63.068	49.230
D	31.970	26.436	27.831	32.267
E	31.428	37.166	32.229	31.915
F	29.935	38.712	31.854	26.180
G	37.272	43.377	69.240	46.850
H	49.169	66.621	88.257	83.256
I	46.419	64.004	77.368	67.612
J	40.597	53.819	52.945	52.803

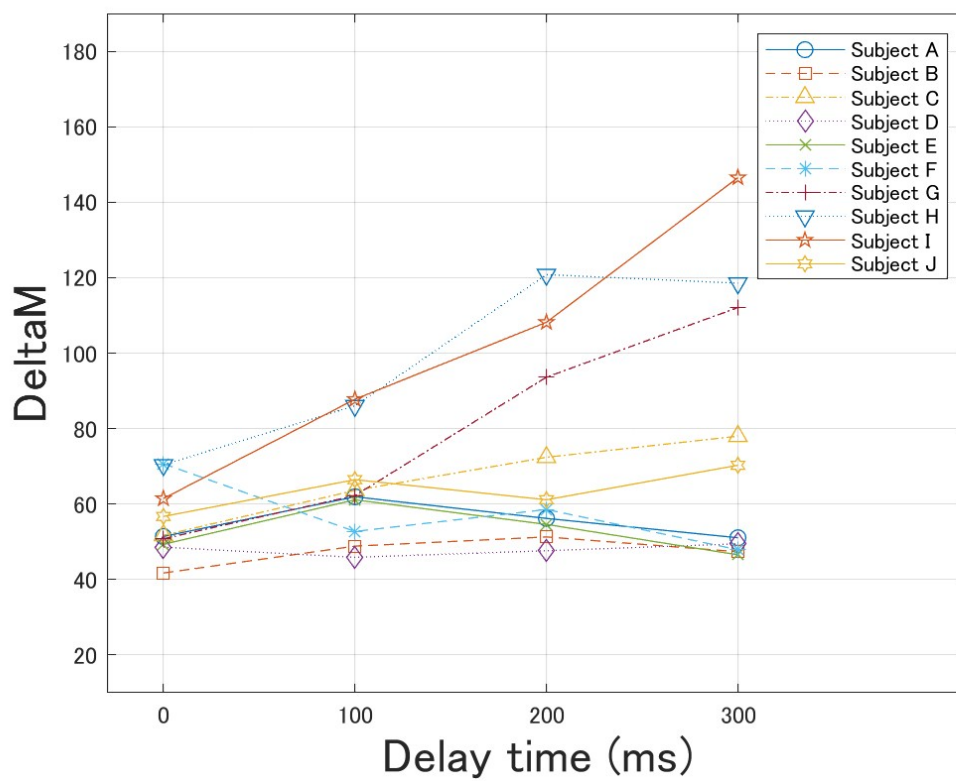


図 4.11: 各実験参加者の遅延量ごとにおける DeltaM

表 4.7: 各実験参加者の遅延量ごとにおける DeltaM の値

実験参加者	遅延量			
	0 (ms)	100 (ms)	200 (ms)	300 (ms)
A	51.409	61.928	56.200	51.067
B	41.738	48.852	51.291	47.374
C	51.727	63.629	72.393	77.985
D	48.581	45.892	47.639	49.472
E	49.340	61.128	54.610	46.546
F	70.680	52.746	58.568	47.902
G	50.813	62.246	93.632	112.074
H	70.351	86.037	120.794	118.526
I	61.455	87.777	108.167	146.500
J	56.714	66.468	61.140	70.290

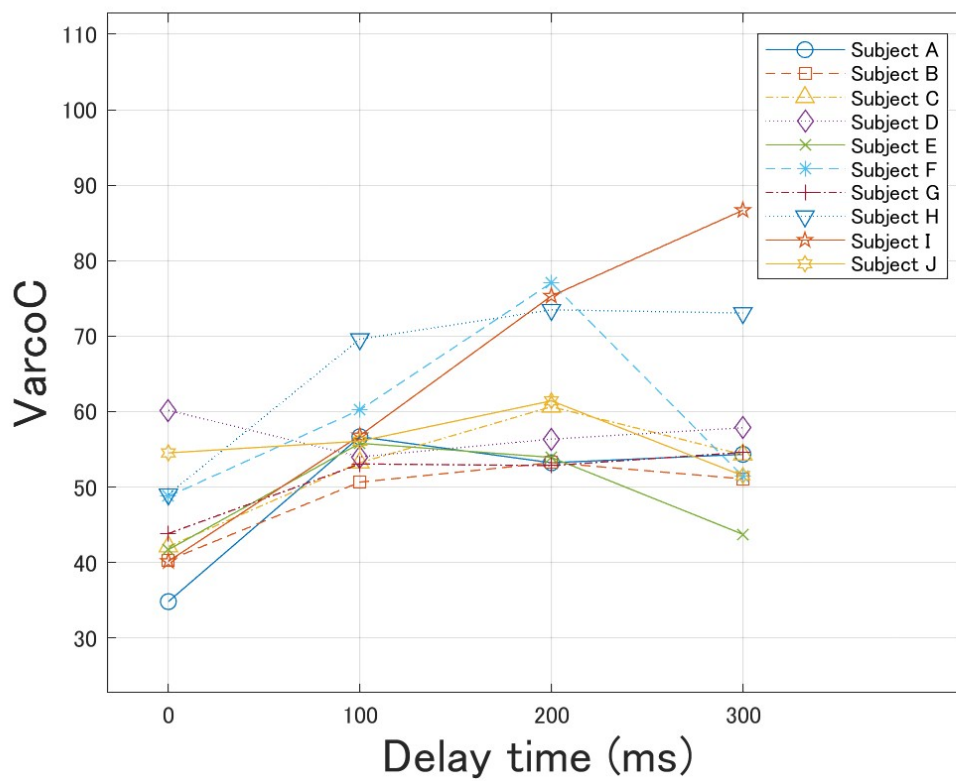


図 4.12: 各実験参加者の遅延量ごとにおける VarcoC

表 4.8: 各実験参加者の遅延量ごとにおける VarcoC の値

実験参加者	遅延量			
	0 (ms)	100 (ms)	200 (ms)	300 (ms)
A	34.827	56.673	53.224	54.342
B	40.313	50.655	53.237	51.090
C	42.123	53.234	60.647	54.294
D	60.182	54.002	56.358	57.860
E	41.710	55.802	53.935	43.745
F	48.771	60.230	77.055	51.422
G	43.852	53.066	52.851	54.605
H	49.049	69.604	73.489	73.065
I	40.168	56.827	75.322	86.680
J	54.531	56.080	61.431	51.511

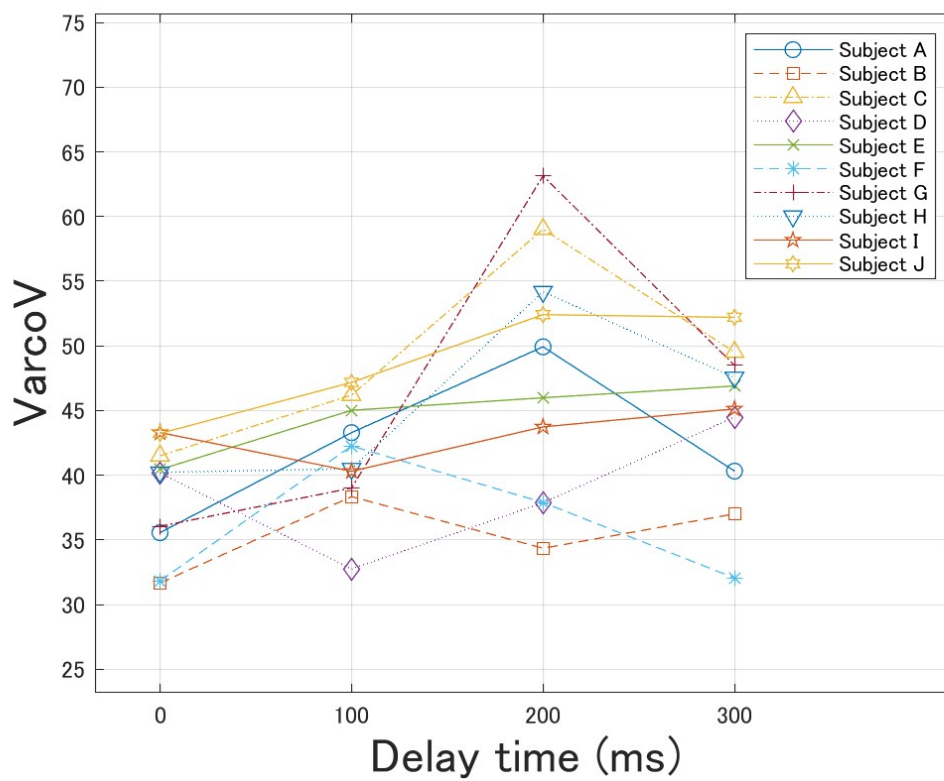


図 4.13: 各実験参加者の遅延量ごとにおける VarcoV

表 4.9: 各実験参加者の遅延量ごとにおける VarcoV の値

実験参加者	遅延量			
	0 (ms)	100 (ms)	200 (ms)	300 (ms)
A	35.564	43.293	49.929	40.322
B	31.679	38.351	34.360	37.016
C	41.509	46.194	59.017	49.527
D	40.173	32.763	37.872	44.494
E	40.447	45.022	45.991	46.902
F	31.836	42.263	37.898	32.061
G	36.062	39.011	63.131	48.494
H	40.225	40.505	54.181	47.559
I	43.301	40.309	43.741	45.130
J	43.232	47.180	52.414	52.203



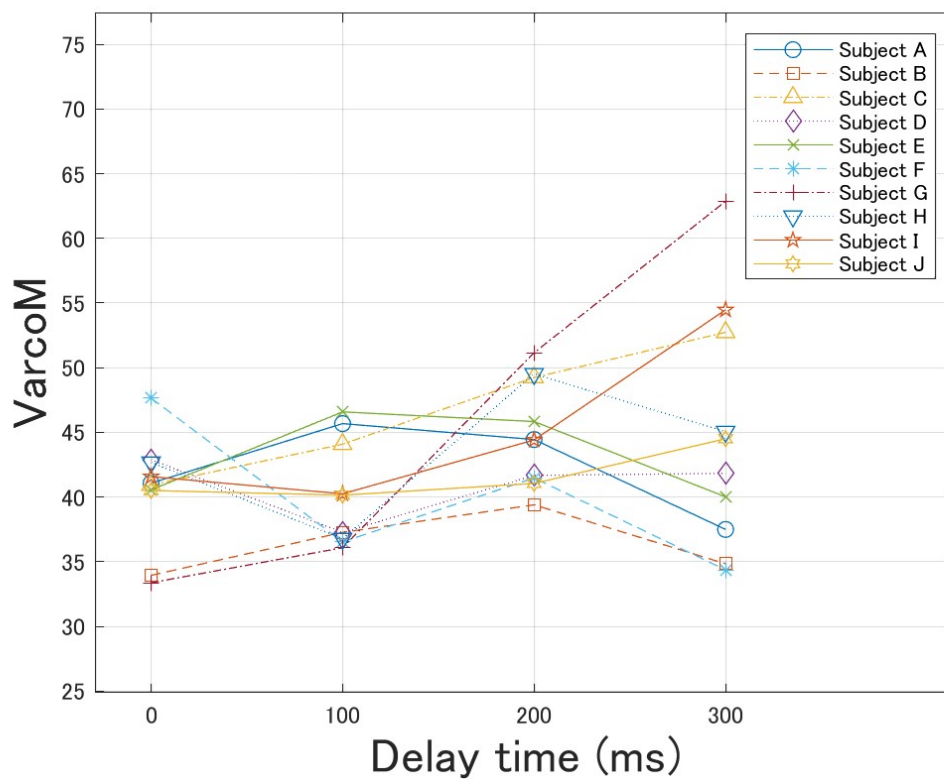


図 4.14: 各実験参加者の遅延量ごとにおける VarcoM

表 4.10: 各実験参加者の遅延量ごとにおける VarcoM の値

実験参加者	遅延量			
	0 (ms)	100 (ms)	200 (ms)	300 (ms)
A	41.079	45.679	44.447	37.499
B	33.954	37.275	39.401	34.848
C	40.932	44.091	49.239	52.737
D	42.887	37.273	41.681	41.833
E	40.556	46.586	45.838	40.017
F	47.649	36.541	41.477	34.368
G	33.381	36.087	51.161	62.862
H	42.686	36.781	49.521	45.056
I	41.581	40.254	44.417	54.477
J	40.510	40.149	41.057	44.510

## 4.8.2 実験参加者間分析結果

図 4.15 に実験参加者間の平均子音区間の結果を示す。RM-ANOVA で検定したところ、遅延量で平均子音区間に有意差が認められた ( $F(1, 8) = 7.2920, p < 0.05$ )。また、遅延量 0 ms の平均子音区間と他の遅延量の平均子音区間の差を Dunnett で検定したところ、遅延量 0 ms と 100 ms 間で平均子音区間に有意差が認められた ( $t(0, 100) = 4.358, p < 0.001$ )。遅延量 0 ms と 200 ms 間で平均子音区間に有意差が認められた ( $t(0, 200) = 3.365, p < 0.01$ )。遅延量 0 ms と 300 ms 間で平均子音区間に有意差が認められた ( $t(0, 300) = 3.520, p < 0.01$ )。

図 4.16 に実験参加者間の平均母音区間の結果を示す。RM-ANOVA で検定したところ、遅延量で平均母音区間に有意差が認められなかった ( $F(1.28, 10.21) = 2.7077$ )。また、遅延量 0 ms の平均母音区間と他の遅延量の平均母音区間の差を Dunnett で検定したところ、遅延量 0 ms と 100 ms 間で平均母音区間に有意差が認められた ( $t(0, 100) = 2.843, p < 0.05$ )。また、他の遅延量で平均母音区間に有意差が認められなかった。

図 4.17 に実験参加者間の平均モーラ区間の結果を示す。RM-ANOVA で検定したところ、遅延量で平均モーラ区間に有意差が認められた ( $F(1, 8) = 4.51, p < 0.1$ )。また、遅延量 0 ms の平均モーラ区間と他の遅延量の平均モーラ区間の差を Dunnett で検定したところ、遅延量 0 ms と 100 ms 間で平均モーラ区間に有意差が認められた ( $t(0, 100) = 2.872, p < 0.05$ )。遅延量 0 ms と 200 ms 間で平均モーラ区間に有意差が認められた ( $t(0, 200) = 2.757, p < 0.05$ )。遅延量 0 ms と 300 ms で平均モーラ区間に有意差が認められた ( $t(0, 300) = 3.437, p < 0.01$ )。

図 4.18 に実験参加者間の DeltaC の結果を示す。RM-ANOVA で検定したところ、遅延量で DeltaC に有意差が認められた ( $F(1.15, 9.19) = 5.5696, p < 0.05$ )。また、遅延量 0 ms の DeltaC と他の遅延量の DeltaC の平均値の差を Dunnett で検定したところ、遅延量 0 ms と 100 ms 間で DeltaC に有意差が認められた ( $t(0, 100) = 3.309, p < 0.01$ )。遅延量 0 ms と 200 ms 間で DeltaC に有意差が認められた ( $t(0, 200) = 3.865, p < 0.01$ )。遅延量 0 ms と 300 ms で DeltaC に有意差が認められた ( $t(0, 300) = 3.583, p < 0.01$ )。

図 4.19 に実験参加者間の DeltaV の結果を示す。RM-ANOVA で検定したところ、遅延量で DeltaV に有意差が認められた ( $F(1.53, 12.24) = 5.8903, p < 0.05$ )。また、遅延量 0 ms の DeltaV と他の遅延量の DeltaV の差を Dunnett で検定したところ、遅延量 0 ms と 100 ms 間で DeltaV に有意差が認められた ( $t(0, 100) = 2.671, p < 0.05$ )。遅延量 0 ms と 200 ms 間で DeltaV に有意差が認められた ( $t(0, 200) = 4.421, p < 0.001$ )。遅延量 0 ms と 300 ms 間で DeltaV に有意差が認められた ( $t(0, 300) = 2.78, p < 0.05$ )。

図 4.20 に実験参加者間の DeltaM の結果を示す。RM-ANOVA で検定したところ、遅延量で DeltaM に有意差が認められた ( $F(1.12, 8.95) = 4.1174, p < 0.1$ )。また、遅延量 0 ms の DeltaM と他の遅延量の DeltaM の差を Dunnett で検定し

たところ、遅延量 0 ms と 200 ms 間で DeltaM に有意差が認められた ( $t(0, 200) = 2.493, p < 0.05$ )。遅延量 0 ms と 300 ms 間で DeltaM に有意差が認められた ( $t(0, 300) = 3.122, p < 0.05$ )。遅延量 0 ms と 100 ms で DeltaM に有意差認められなかった。

図 4.21 に実験参加者間の VarcoC の結果を示す。RM-ANOVA で検定したところ、遅延量で VarcoC に有意差が認められた ( $F(1.99, 15.95) = 6.5127, p < 0.01$ )。また、遅延量 0 ms の VarcoC と他の遅延量の VarcoC の差を Dunnett で検定したところ、遅延量 0 ms と 100 ms 間で VarcoC に有意差が認められた ( $t(0, 100) = 3.231, p < 0.01$ )。遅延量 0 ms と 200 ms 間で VarcoC に有意差が認められた ( $t(0, 200) = 4.732, p < 0.001$ )。遅延量 0 ms と 300 ms 間で VarcoC に有意差が認められた ( $t(0, 300) = 3.595, p < 0.01$ )。

図 4.22 に実験参加者間の VarcoV の結果を示す。RM-ANOVA で検定したところ、遅延量で VarcoV に有意差が認められた ( $F(1.9, 15.19) = 5.5247, p < 0.05$ )。また、遅延量 0 ms の VarcoV と他の遅延量の VarcoV の差を Dunnett で検定したところ、遅延量 0 ms と 200 ms 間で VarcoV に有意差が認められた ( $t(0, 200) = 4.371, p < 0.001$ )。遅延量 0 ms と 300 ms 間で VarcoV に有意差が認められた ( $t(0, 300) = 2.760, p < 0.05$ )。遅延量 0 ms と 100 ms 間で VarcoV に有意差が認められなかった。

図 4.23 に実験参加者間の VarcoM の結果を示す。RM-ANOVA で検定したところ、遅延量で VarcoM に有意差が認められなかった ( $F(1.5, 12) = 2.6327$ )。また、遅延量 0 ms の VarcoM と他の遅延量の VarcoM の差を Dunnett で検定したところ、すべての遅延量で VarcoM に有意差が認められなかった。

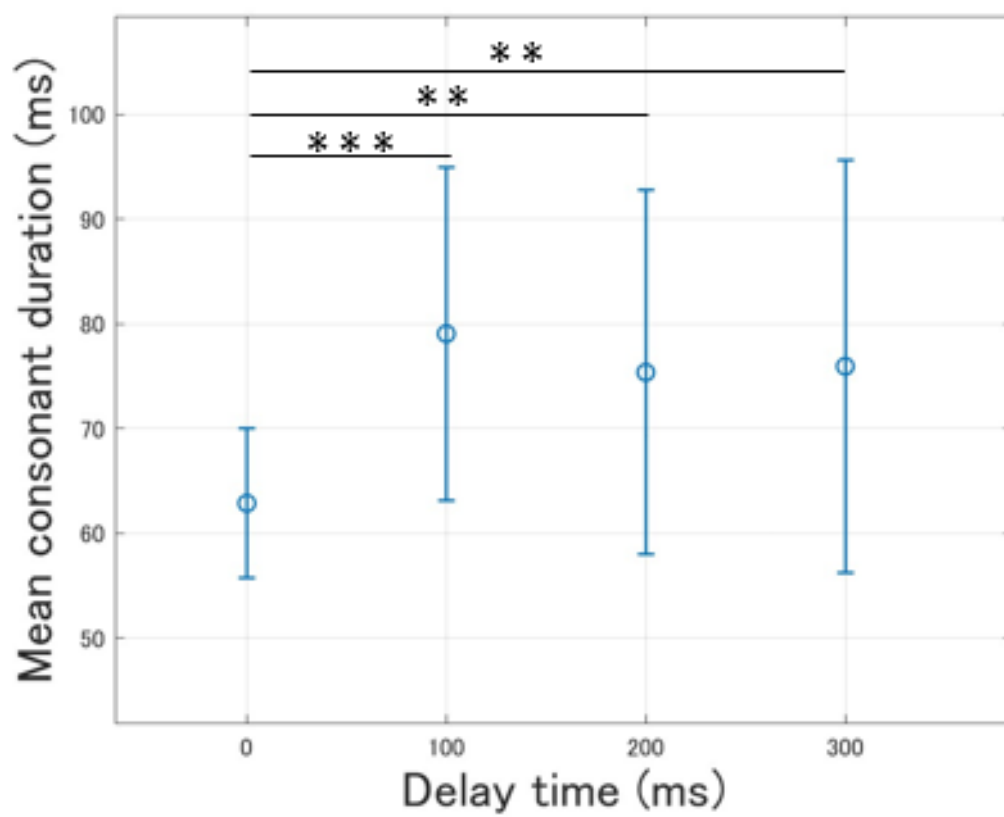


図 4.15: 実験参加者間の遅延量ごとにおける平均子音区間

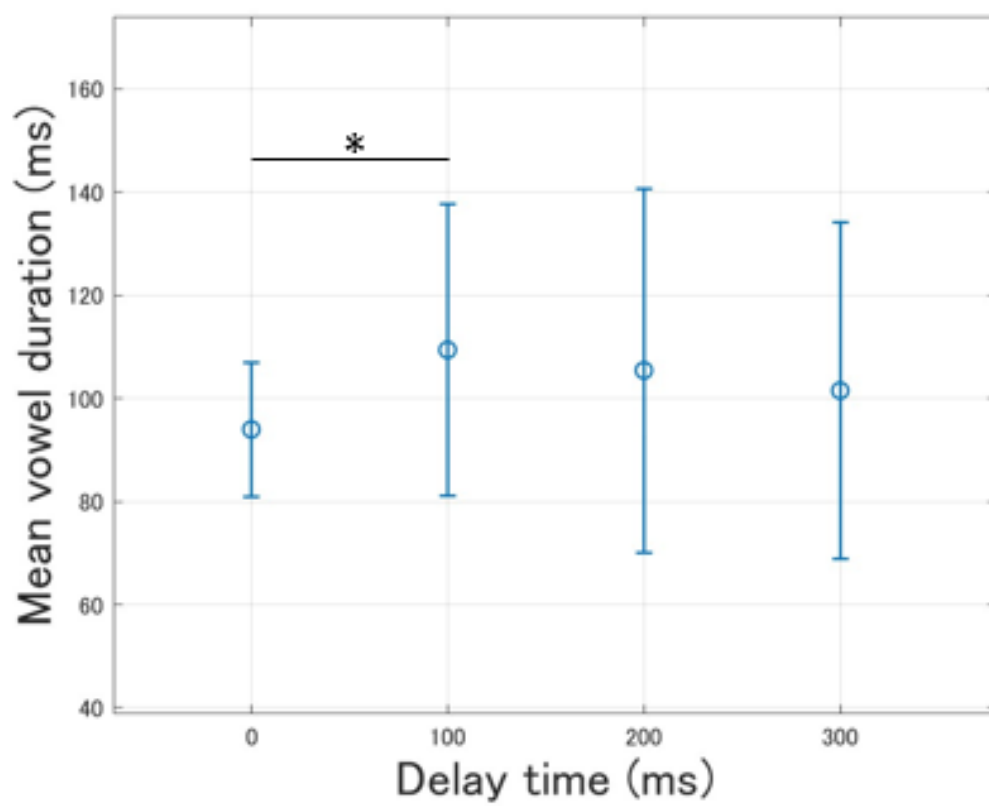


図 4.16: 実験参加者間の遅延量ごとにおける平均母音区間

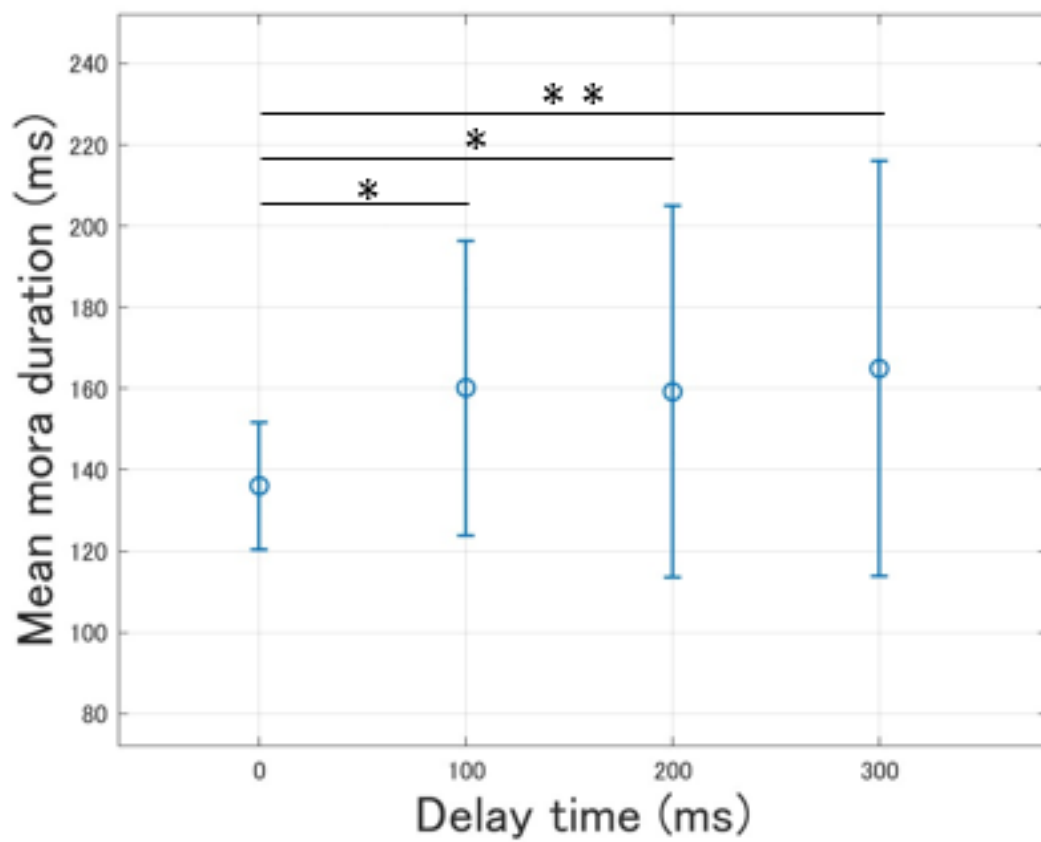


図 4.17: 実験参加者間の遅延量ごとにおける平均モーラ区間

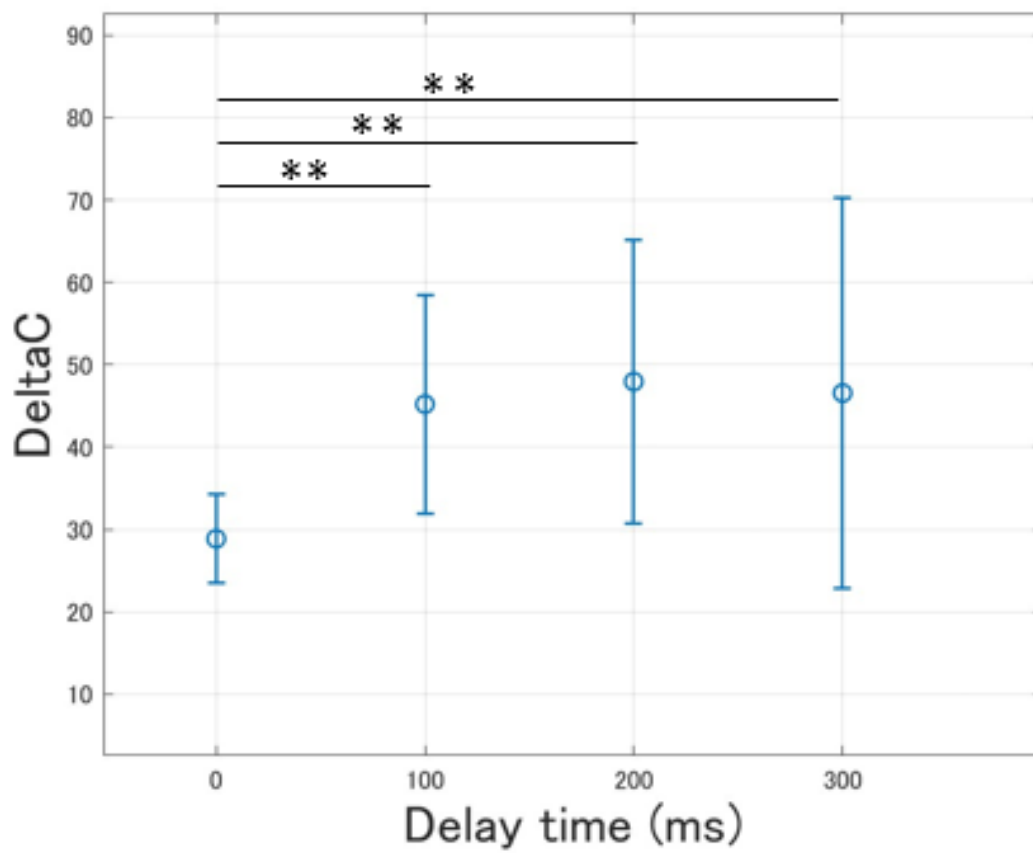


図 4.18: 実験参加者間の遅延量ごとにおける DeltaC



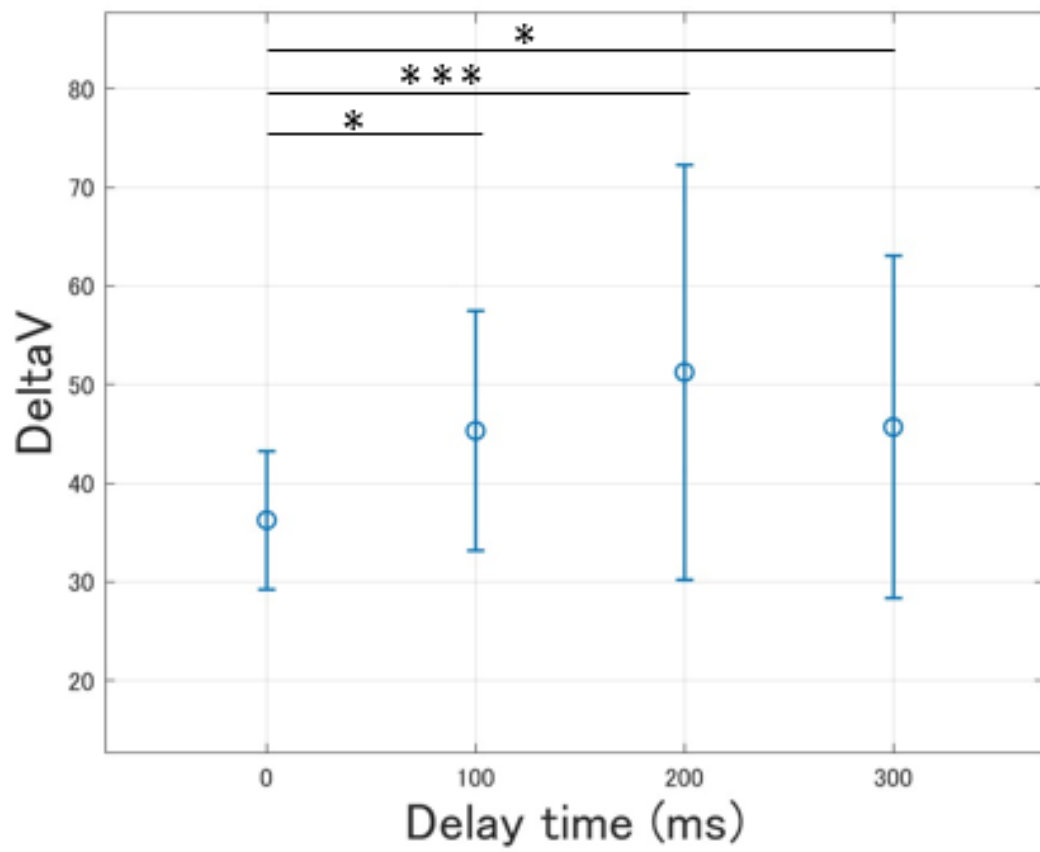


図 4.19: 実験参加者間の遅延量ごとにおける DeltaV

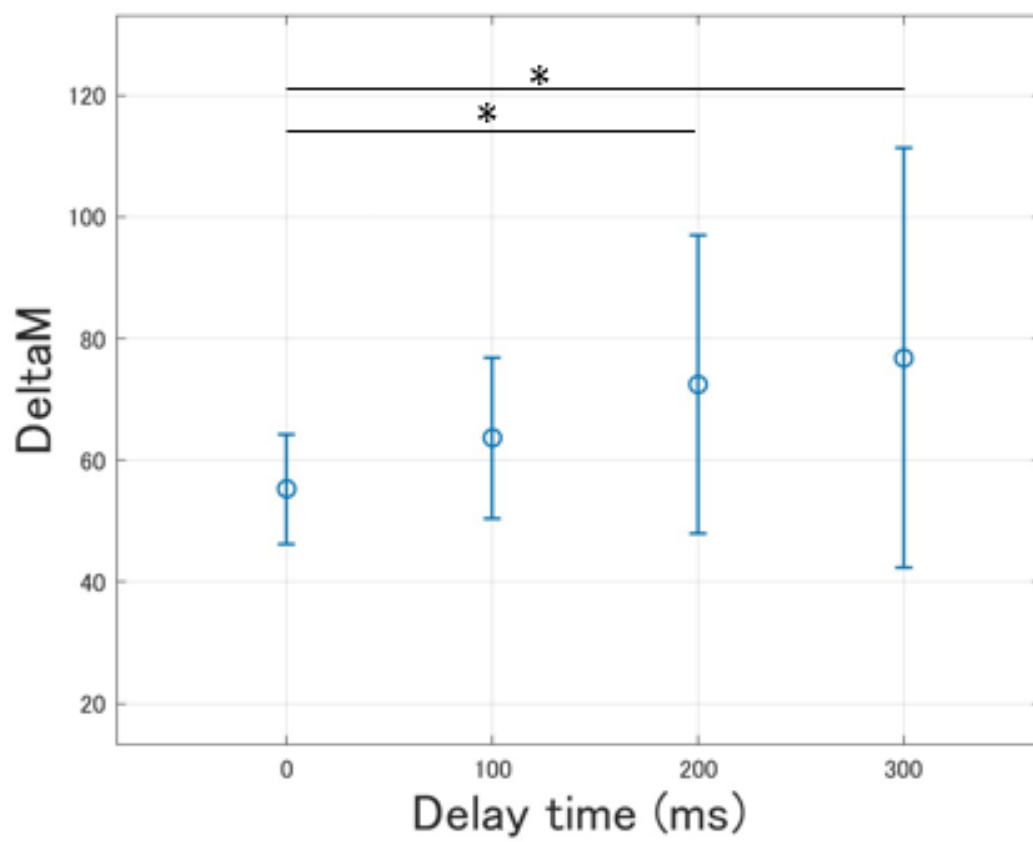


図 4.20: 実験参加者間の遅延量ごとにおける DeltaM

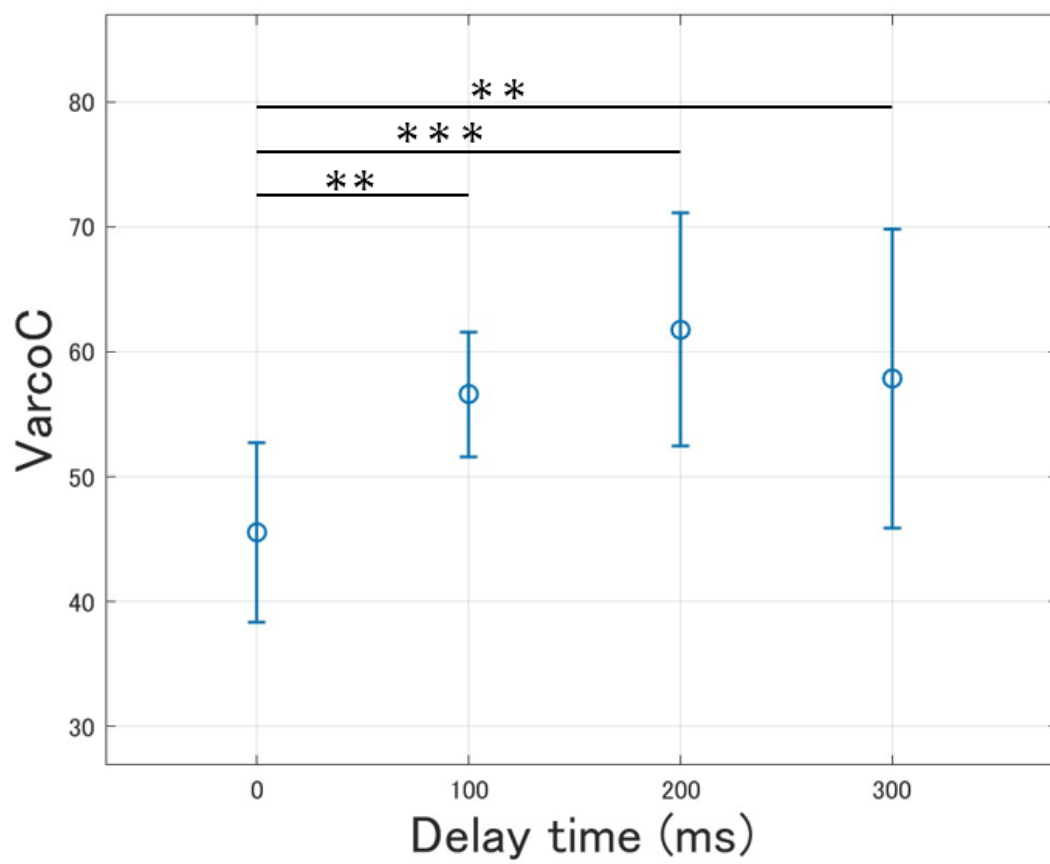


図 4.21: 実験参加者間の遅延量ごとにおける VarcoC

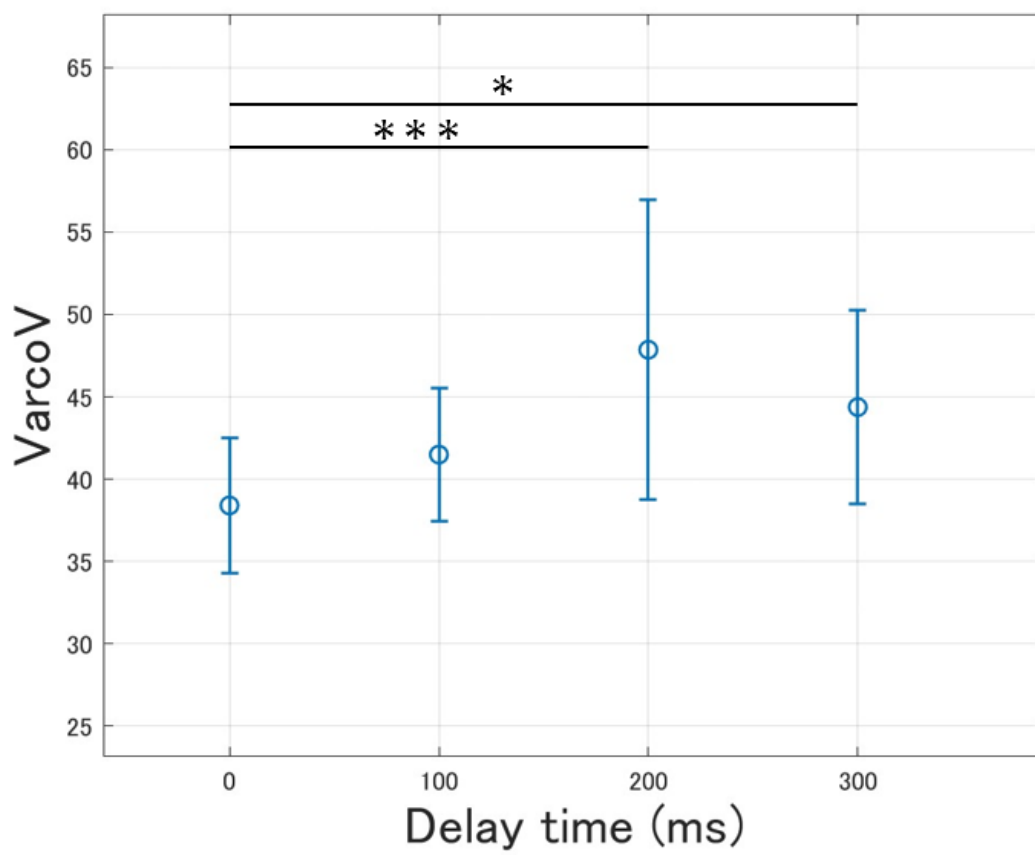


図 4.22: 実験参加者間の遅延量ごとにおける VarcoV

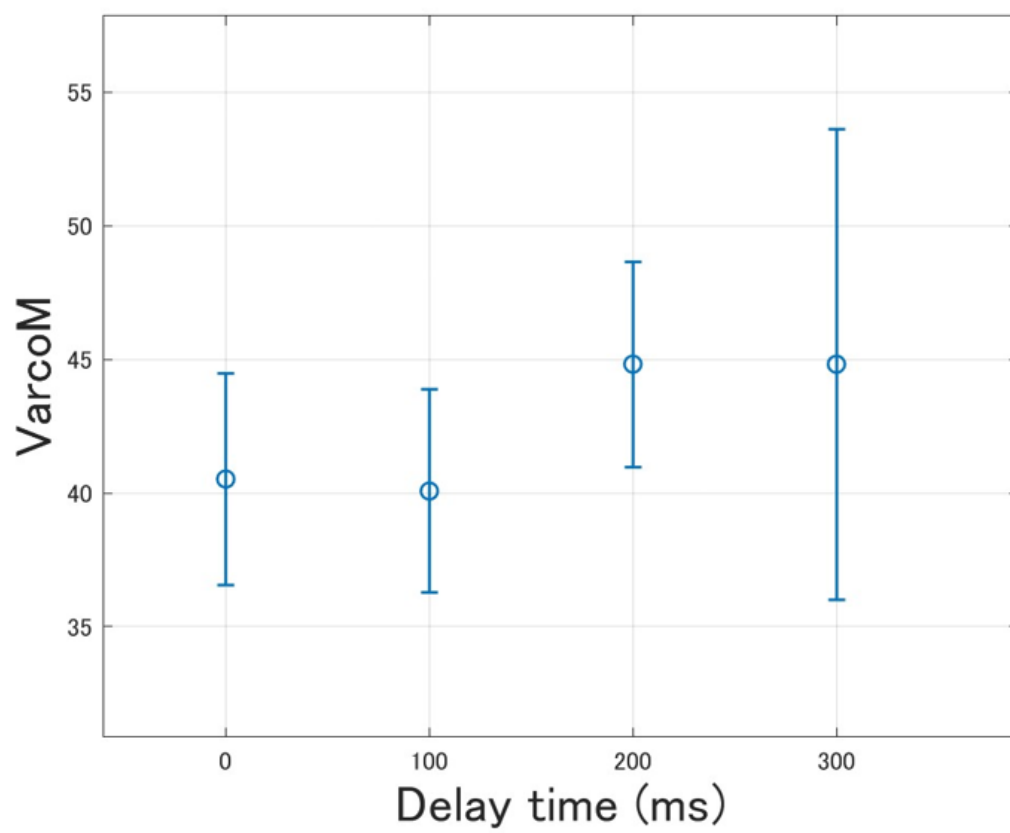


図 4.23: 実験参加者間の遅延量ごとにおける VarcoM

## 4.9 考察

実験で予想される結果として、先行研究 [5][11] より遅延量の増加で平均モーラ区間が長くなることを予想した。また、平均モーラ区間が長くなることで平均子音区間と平均母音区間も長くなることを予想した。この予想に対して、個人の差はあったが先行研究と同様に遅延量の増加で平均モーラ区間が長くなることがわかった。また、平均子音区間と平均母音区間も遅延量の増加で長くなる傾向が見られることがわかった。実験参加者間での分析結果を見ると、遅延量で平均子音区間と平均モーラ区間は有意差が認められたが、平均母音区間は有意差が認められなかった。このことから、平均子音区間と平均モーラ区間は実験参加者全体が遅延量増加に対して長くなり、平均母音区間は一部の実験参加者が遅延量増加に対して長くなったと考えられる。

Delta に関しては、先行研究 [9][10] より遅延量の増加で DeltaC と DeltaM が大きくなると予想した。この予想に対して、遅延量が増加すると DeltaC は個人の差はあったが大きくなる傾向が見られるがわかった。また、遅延量が増加すると DeltaM も個人の差はあるが大きくなることがわかり、半数は遅延量 300 ms で最も大きくなる傾向が見られた。遅延量が増加すると DeltaV も個人の差はあるが、大きくなる傾向が見られ、遅延量 200 ms で最も大きくなる傾向が見られることがわかった。実験参加者間の分析結果を見ると、遅延量で DeltaC、DeltaV、DeltaM が有意差が認められたため、実験参加者全体が遅延量の増加に対して DeltaC、DeltaV、DeltaM が大きくなることが考えられる。

Varco に関しては、先行研究 [10] より遅延量の増加で VarcoC が大きくなるか小さくなるのどちらかが見られると予想した。この予想に対して、遅延量が増加すると VarcoC は大きくなることがわかり、過半数は遅延量の増加で VarcoC が大きく変化する傾向があった。遅延量が増加すると VarcoV も大きくなることがわかり、半数は遅延量の増加で VarcoV が大きく変化する傾向があった。遅延量が増加すると VarcoM も大きくなることがわかり、半数は遅延量 100 ms で VarcoM が小さくなり、遅延量 200 ms で大きくなる傾向があった。実験参加者間の分析結果を見ると、遅延量で VarcoC と VarcoV が有意差が認められたが、VarcoM は有意差が認められなかった。そのため、VarcoC と VarcoV は実験参加者全体で遅延量の増加に対して大きくなり、VarcoM は一部の実験参加者で遅延量に対して大きくなると考えられる。

これらの考察から、音素区間のリズム指標では、フィードバック遅延量によって区間の平均とばらつきが大きくなり、遅延量なしの時より音素区間のばらついている割合が多いと評価される。また、モーラ区間の指標では、フィードバック遅延量によって区間の平均とばらつきが大きくなるが、遅延量なしの時と比べるとモーラ区間のばらついている割合は変わらないと評価される。

## 第5章 全体考察

本研究では、調査したリズム指標がフィードバック遅延量に応じてどのような変化が出るかを観測するために、DAF 下での日本語文章読み上げ課題における発話実験を行い、発話実験で得られた録音音声データを音素・モーラ区間のリズム指標で発話リズム分析を行った。分析結果により以下のことが考えられる。

- 各平均区間について、遅延量の増加で平均モーラ区間が長くなる傾向があったことから、子音区間より比較的長い区間を占めている母音区間の方がモーラに依存して長くなる傾向があると予想していた。しかし、実験参加者間分析の結果を見ると平均子音区間の方は全体の条件間で有意差があり、各条件間でも有意差があると評価された。また、平均母音区間の方は全体の条件間で有意差が認められず、各条件間では遅延量なしのときと遅延量 100 ms のときの条件間のみ有意差が認められた。このことから、本研究では平均母音区間より平均子音区間の方が DAF の影響が大きくなった。平均母音区間の方は全体の条件間で有意差がない理由として、1つの条件間では有意差があったが他の条件間では有意差が認められなかったため、全体で有意差が認められなかったと判断された可能性がある。また、遅延量なしのときと遅延量 200 ms と 300 ms の平均母音区間に有意差が認められなかった理由として、遅延量 100 ms で最も DAF の影響が大きくなる実験参加者が多かったと考えられる。
- 各区間の標準偏差を表す Delta について、実験参加者間分析より DeltaC と DeltaV は全体の条件間で有意差が認められ、各条件間でも有意差が認められた。このことから DeltaM でも同様の検定結果が出ると予想した。しかし、全体の条件間で有意差は認められたが、各条件間では遅延量なしのときと遅延量 100 ms のときで DeltaM の有意差が認められなかった。これについて、各条件間での DeltaM に有意差が認められた割合が多いことから、全体の条件間でも有意差が認められたと予想する。また、遅延量なしのときと遅延量 100 ms の条件間で有意差が認められなかった理由として、条件間での DeltaM の差が小さいことと、特殊拍による影響が考えられる。
- 各区間の変動係数を表す Varco について、実験参加者間分析より VarcoC と VarcoV は全体の条件間で有意差が認められ、各条件間では遅延量なしのときと遅延量 100 ms のときの VarcoV 以外で有意差が認められた。このことか

ら VarcoM も同様に全体の条件間で有意差が認められ、遅延量なしと遅延量 200 ms または 300 ms の条件間で有意差が認められると予想した。しかし、VarcoM は全体の条件間で有意差が認められなく、各条件間でも有意差が認められなかった。これらはモーラの音節構造に関係していると考えられる。音素の場合、子音または母音のどちらかで評価されるが、モーラの場合は子音または母音のどちらかか子音と母音のセグメントで評価される。また、日本語文章は子音と母音のセグメントの割合が多いため、遅延量による影響が音素と異なる可能性がある。

- 各指標での個人差について、実験参加者内分析で実験参加者 2 名が他の実験参加者と比べて指標の値が大きくなる傾向があった。また、音素・モーラ区間は遅延量の増加であまり変化が見られないが、無音区間の頻度が多くなった実験参加者もいたため、DAF の影響を受けにくい可能性がある実験参加者もいれば、無音区間を増やして DAF の影響を受けにくくするような発話運動を行っている可能性がある実験参加者もいると考えられる。



## 第6章 結論

### 6.1 本研究で明らかにしたこと

本研究では、遅延聴覚フィードバックによる発話改善法の実現を目指し、聴覚フィードバック情報の遅延時間が日本語読み上げにおいて発話リズムにどのように影響するのかを解明することを目的とした。そのために、DAF下での日本語文章読み上げ課題における発話実験とリズム指標を用いた発話リズム分析を行った。リズム指標を用いた発話リズム分析では以下のことが明らかになった。

- フィードバック遅延量により発話音声の子音とモーラの平均区間が長くなることがわかった。
- フィードバック遅延量により発話音声の子音・母音・モーラ区間のばらつきが大きくなることがわかった。
- フィードバック遅延量により発話音声の子音・母音区間のばらついている割合が遅延量なしの時の発話音声より多いことがわかった。また、モーラ区間のばらついている割合は遅延量なしの時の発話音声と比べるとあまり変化がないことがわかった。

### 6.2 残された課題

遅延聴覚フィードバックによる発話改善法の実現を目指す上で以下のような残された課題が存在する。

- 特殊拍の分析方法の検討

日本語には特殊拍（促音、撥音、長音、拗音）と無声化という母音が発音されない音声がある。音素区間のリズム指標を用いた分析では、日本語特有の特殊拍（促音、撥音）と無声化された音声は子音と母音に区別ができなく、全て子音として扱うと分析結果に偏りが生じる可能性があることが考えられる。そのため、本研究では、音素区間の分析において促音と撥音、無声化された音声は省いて分析を行った。これらから、音素区間を用いた分析で特殊拍や無声化された音声をどのように分析するかを検討する必要がある。

- 無音区間の分析方法の検討

無音区間の分析において、本研究では文章の読み間違いを音素・モーラとしてカウントせず、無音区間として扱った。さらに、発話実験で遅延量なしの時より無音区間が増加した実験参加者が見られた。したがって、無音区間が発話リズムに及ぼす影響を調査し、無音区間の分析方法の検討を行う必要がある。

- PVIなどの他のリズム指標を用いた分析

本研究では、発話リズムを定量化するために音素・モーラ区間のリズム指標 ( $\Delta C \cdot \Delta V \cdot \Delta M \cdot \text{VarcoC} \cdot \text{VarcoV} \cdot \text{VarcoM}$ ) を用いた分析を行い、DAF 下でのモーラの等時性を調べることで発話リズムを評価した。これらの他に区間を用いたリズム指標として PVI や %V などが挙げられる。そのほかに、経験的モード分解 (empirical mode decomposition: EMD) などの別の観点からリズムを定義して発話リズムを分析する方法も検討されている [50]。これらを用いた分析ではどのような結果になるのかを調査し、発話リズムの分析に最適な方法を確立する必要がある。

# 謝辞

本研究を遂行するにあたり，研究に対する姿勢やご助言，研究室会議と打ち合わせでの的確なご指摘を賜りました主指導教員である鶴木祐史教授に深く感謝いたします。また，研究の相談や実験のサポート，資料作成時のアドバイス，RPのときの精神的なサポートなど自身の研究活動において全てのことを支えていただいた上江洲安史特任助教に心から感謝いたします。研究室会議において研究のご助言やご指摘を賜りました木谷俊介講師，赤木正人名誉教授，太田さん，磯山さんに心から感謝いたします。研究において気軽に相談に乗っていただいた Guo さんをはじめとした先輩方や，日常での雑談や実験への協力に応じてくれた同期のメンバーと後輩方に心から感謝いたします。最後に，長い学生生活を支えていただいた家族の皆様に心から感謝いたします。

## 参考文献

- [1] Danes, P. D. and Pinson, E. N., “The speech chain: The physics and biology of spoken language (Second Edition),” W. H. Freeman, 1993.
- [2] 廣谷定男 (編著), 聞くと話すの脳科学, コロナ社, 東京, 2017.
- [3] 岡崎俊太郎, 武田湖太郎, “吃音と聴覚フィードバック,” 脳科学とリハビリテーション研究会, Vol. 14, pp. 35–40, 2014.
- [4] Lee, B. S., “Effect of delayed speech feedback,” The Journal of the Acoustical Society of America 22, Vol. 22, pp.824-826, 1950.
- [5] Yamamoto, K. and Kawabata, H., “Adaptation to delayed auditory feedback induces the temporal recalibration effect in both speech perception and production,” Exp. Brain Res, 232, pp. 3707–3718, 2014.
- [6] Grabe, E. and Low, E. L., “Durational variability in speech and the rhythm class hypothesis,” In N. Warner, and C. Gussenhoven (Eds.), Papers in laboratory phonology 7, pp. 515–546, Mouton de Gruyter, 2002.
- [7] Ramus, F., Nespore, M., Mehler, J., “Correlates of linguistic rhythm in the speech signal,” Cognition, 73, pp.265-292, 1999.
- [8] White, L. and Mattys, S.L., “Calibrating rhythm: First language and second language studies,” Journal of Phonetics, Vol.35,No.4, pp.501-522, 2007.
- [9] Dellwo, V., “Rhythm and Speech Rate: A Variation Coefficient for deltaC,” Language and Language Processing: Proc. 38th Linguistic Colloquium Piliscaba, pp. 231–241, 2006.
- [10] Dellwo, V. and Wagner, P., “Relations between language rhythm and speech rate,” Proceedings of the 15th international congress of phonetics sciences, pp. 471–474, 2003.
- [11] 一ノ瀬 裕, 能田 由紀子, “音読内容・音読速度・遅延時間が遅延聴覚フィードバック条件下での発話に与える影響,” 日本音響学会誌, Vol. 73, pp. 404–406, 2017.

- [12] Merker, B. H., Madison, G. S. and Eckerdal, P., “On the role and origin of isochrony in human rhythmic Entrainment,” *Cortex*, 45(1), pp. 4-17, 2009.
- [13] Abercrombie, D., “*Studies in Phonetics and Linguistics*,” London, Oxford University Press, 1965.
- [14] Abercrombie, D., “*Elements of General Phonetics*,” Edinburgh, Edinburgh University Press, 1967.
- [15] Ramus, F., Nespors, M., and Mehler, J., “Correlates of linguistic rhythm in the speech signal,” *Cognition*, 72, pp. 1–28, 1999.
- [16] 河原英紀, “聴覚フィードバックの発声への影響,” *日本音響学会誌*, Vol. 59, No. 11, pp. 670–675, 2003.
- [17] 河原英紀, “声を使って聴覚を探る,” *日本音響学会誌*, vol. 53, No. 9, pp. 731–737, 1997.
- [18] Trofimovich, P. and Baker, W., “Learning second-language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech,” *Studies in Second Language Acquisition*, 28, pp. 1–30, 2006.
- [19] Dellwo, V., Leemann, A., and Kolly, M. J., “Rhythmic variability between speakers: Articulatory, prosodic, and linguistic factors,” *The Journal of the Acoustical Society of America*, Val. 137, No. 3, pp. 1513–1528, 2015.
- [20] Yamamoto, K. and Kawabata, H., “Temporal recalibration in vocalization induced by adaptation of delayed auditory feedback,” *PLoS ONE*, Vol. 6, No. 12, 2011.
- [21] Uezu, Y., Hiroya, S., and Mochida, T., “Articulatory compensation for low-pass filtered formant-altered auditory feedback,” *J Acoust Soc Am*, Vol. 150, No. 1, pp.64–73, 2021.
- [22] Haegens, S. and Golumbic, E. Z., “Rhythmic facilitation of sensory processing: A critical review,” *Neurosci, Biobehav, Rev.* 86, pp. 150–165, 2018.
- [23] F. Cummins, “Rhythm and speech,” Redford, M.A. (Ed.), *The Handbook of Speech Production*, first edition, John Wiley and Sons, Inc. p. 158, 2015.
- [24] 国立障害者リハビリテーションセンター研究所, 吃音について, 国立障害者リハビリテーションセンター, <http://www.rehab.go.jp/ri/departj/kankaku/466/2/>

- [25] Toyomura, A. and Omori, T., “Auditory feedback control during a sentence-reading task: Effect of other’s voice,” *Acoust. Sci. and Tech.*, 26, pp. 358–361, 2005.
- [26] Dellwo, V., Leeman, A. and Kolly, M. J., “Speaker idiosyncratic rhythmic features in the speech signal,” in *Proceedings of Interspeech 2012*, September 9–13, Portland, OR, pp. 1584–1587. 2012
- [27] Arvaniti, A., “The usefulness of metrics in the quantification of speech rhythm,” *Journal of Phonetics*, vol. 40, pp. 351–373, 2012.
- [28] Liss, J. M., White, L., Mattys, S. L., Lansford, K., Lotto, A. J., Spitzer, S. M., and Caviness, J. N., “Quantifying Speech Rhythm Abnormalities in the Dysarthrias,” *Journal of Speech Language and Hearing Research*, vol. 52, No. 5, pp. 1334–1352, 2009.
- [29] Pike, K., *The Intonation of American English*. 2nd edition. Ann Arbor: University of Michigan, Press. 1946.
- [30] 難波精一郎 (編著), *音と時間*, コロナ社, 東京, 2015.
- [31] 北原真冬, 田嶋圭一, 田中邦佳 (著), *音声学を学ぶ人のための Praat 入門*, ひつじ書房, 東京, 2017.
- [32] Kuhn, M., Schwanenflugel, P., Meisinger, E., “Aligning theory and assessment of reading fluency: automaticity, prosody, and definitions of fluency,” *Reading Research Quarterly*, Vol. 45, No. 2, p. 230–251, 2010.
- [33] Bloch, B., “Studies in colloquial Japanese IV: Phonemics,” *Language* Vol. 26, No. 1, pp. 86–125, 1950.
- [34] Han, M. S., “The feature of duration in Japanese,” *Onsei no kenkyuu*, 10, pp. 65–80, 1962.
- [35] Ladefoged, P., *A Course in Phonetics*, New York, Harcourt Brace Jovanovich, 1975.
- [36] Chesters, J., Ravary, L. B., Möttönen, R., “The effects of delayed auditory and visual feedback on speech production,” *Journal of the Acoustical Society of America*, Vol. 137, No. 2, pp. 873–883, 2015.
- [37] Howell, P. and Sackin, S., “Timing interference to speech in altered listening conditions,” *Journal of the Acoustical Society of America*, Vol. 111, No. 6, pp. 2842–2852, 2002.

- [38] Dankovicova, J. and Dellwo, V., “Czech speech rhythm and the rhythm class hypothesis,” In Proceedings of XVIth ICPHS, Saarbrücken, Germany, pp. 1241–1244), 2007.
- [39] Dihingia, L. and Priyankoo, S., “Rhythm and speaking rate in Assamese varieties,” 10th International Conference on Speech Prosody, Tokyo, Japan, pp. 561–565, 2020.
- [40] Dellwo, V., “Influences of speech rate on the acoustic correlates of speech rhythm: An experimental phonetic study based on acoustic and perceptual evidence,” Ph.D. dissertation, University Bonn, Bonn, Germany, pp. 1–185, 2010.
- [41] Yazawa, K. and Kondo, M., “A comparison of rhythm metrics for L2 speech,” In Proceedings of the 11th International Conference on Speech Prosody, International Speech Communication Association, pp. 332–336, 2022.
- [42] Bishop, D. V. and Adams, C., “A prospective study of the relationship between specific language impairment, phonological disorders and reading retardation. *Journal of Child Psychology and Psychiatry*, Vol. 31, No. 7, pp. 1027–1050, 1990.
- [43] Patel, R., Niziolek, C., Reilly, K., and Guenther, F. H., “Prosodic adaptations to pitch perturbation in running speech,” *Journal of Speech Language, and Hearing Research*, Vol. 54, No. 4, pp. 1051–1059, 2011.
- [44] Yates, A. J., “Delayed auditory feedback,” *Psychological Bulletin*, Vol. 60, No. 3, pp. 213–232, 1963.
- [45] Skodda, S. , Visser, W.,and Schlegel, U., “Vowel articulation in Parkinson ’ s disease,” *J Voice*, Vol. 25, No. 4, pp. 467–472, 2011.
- [46] Cai, S., Boucek, M., Ghosh, S.S., Guenther, F.H., and Perkell, J.S., A system for online dynamic perturbation of formant frequencies and results from perturbation of the Mandarin triphthong /iau/, In Proceedings of the 8th Intl, Seminar on Speech Production, Strasbourg, France, pp. 65–68, 2008.
- [47] Ferragne, E. and Pellegrino, F., “A comparative account of the suprasegmental and rhythmic features of British English dialects”, In Proceedings of Modelisations pour l ’ Identification des Langues, Paris, 2004.
- [48] 上江洲安史, “ “話す” と “聞く” の “フシギ” な関係,” *電子情報通信学会誌*, 2023.

- [49] 日本電信電話, 息づかい, ふるえ: NTT 研究所発触感コンテンツ専門誌, 2016, <https://furue.ilab.ntt.co.jp/book/201606/contents1.html>
- [50] Tilsen, S. and Arvaniti, A., “Speech rhythm analysis with decomposition of the amplitude envelope: characterizing rhythmic patterns within and across languages,” *J. Acoust. Soc. Am.* 134, pp. 628–639, 2013.