JAIST Repository
https://dspace.jaist.ac.jp/

Title	発話目的の違いに注目した感情音声発話時の調音運動の比較に関する研究
Author(s)	浅井，拓也
Citation	
Issue Date	2017-06
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/14710
Rights	
Description	Supervisor:赤木　正人，情報科学研究科，修士

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

発話目的の違いに注目した
感情音声発話時の調音運動の比較に関する研究

北陸先端科学技術大学院大学
情報科学研究科

浅井 拓也

平成 29 年 6 月

修 士 論 文

発話目的の違いに注目した
感情音声発話時の調音運動の比較に関する研究

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　鵜木祐史 教授
審査委員　党建武 教授

北陸先端科学技術大学院大学

情報科学研究科

平成 29 年 4 月

Copyright © 2017 by Asai Takuya

概 要

ヒトが音声を通じて伝えたいものは音素や, 語といった言語的な情報だけでなく, 感情表現や個人性などといったパラ言語情報や非言語情報も含まれる. ヒトの音声発話にはこれら様々な情報を同時に音声に表出させることが可能であるという特徴がある. 一方でヒトの音声生成に注目すると, 音声の生成に使用できる調音的パラメータの数は限られている. それではヒトはどのように限られた調音パラメータを使用し, 複数の情報を同時に伝達しうる音声を生成しているのであろうか？ この問いに答えることは音声科学における重要な課題である.

近年の感情音声生成メカニズムに関する研究では, 特定の感情表現を行う際に, 特徴的な調音運動が存在することが明らかになってきた. しかし, 感情表現と音声発話時の言語的情報はどの程度関連をしており, 感情表現は, どのように言語的情報の生成を妨害または高めることができるのかという議論にまでは踏み込めていない. 感情表現が含まれる音声には, その感情の伝達を目的に発声された音声 (演技感情音声) と, その感情の伝達は目的とはしていないにも拘わらず感情が滲み出てしまっている音声 (自発感情音声) が存在する. 感情表現が, どのように言語的情報の生成を妨害または高めることができるのかを明らかにするためには, 言語情報の伝達と感情表現の伝達という二つの発話目的の違いに注目し, 感情音声発話時の調音運動を調査することが必要である.

本研究では発話目的が異なる音声発話時の調音運動を調査し, それぞれの音声発話時の特徴を明らかにすることを目的とする.

まずは発話目的に言語情報の伝達と感情表現の伝達が同時含まれる音声発話 (演技感情音声発話) 時の感情表現と言語情報の生成との関連を調査するために, 感情音声コーパスを作成し, 母音 /a, e, o/ を発話している際の調音運動を, 五種類の感情表現毎に比較検討を行った. ついで, 感情表現の伝達は目的とはしていないにも拘わらず感情が滲み出てしまっている音声発話 (自発感情音声発話) 時の調音運動の解析を行い, 感情表現が言語情報の伝達という発話目的の達成をより強く妨害する場合の調音運動の特徴を解析した. 最後に, 発話目的に差が存在するであろう演技感情音声発話時と自発感情音声発話時の調音運動を比較し, それぞれの発話目的の差が, どのように音声生成時の調音運動に影響を与えるのかを検討した.

解析結果から, 発話目的の差が感情音声発話時の言語的情報の生成に影響を与えることを確認した. すなわち, 演技感情音声発話時には, 二つの発話目的を同時に達成するために, 言語情報の生成に重要な調音運動に関しては, その言語的特徴を保持したまま感情表現を行い, 言語情報の生成に重要ではない調音運動に関しては, 感情表現に対し一定の値に変化させること, それに対し, 自発感情音声発話時には, 上記のような調音運動の使い分けは行われず, 生成される音声が言語的に異なる音であるという特徴を, 強調するような調音運動を行うことが示された.

目 次

第 1 章　序論 **2**
1.1　はじめに 2
1.2　背景 3
1.2.1　先行研究 3
1.2.2　問題点 3
1.3　目的 4
1.4　構成 4

第 2 章　音声の生成メカニズム **6**
2.1　調音点 6
2.2　調音運動 6

第 3 章　演技感情音声発話時の調音運動 **8**
3.1　目的 8
3.2　解析対象 8
3.3　感情表現の影響 10
3.3.1　開口度 10
3.3.2　円唇性 12
3.3.3　舌運動 12
3.3.4　考察 18
3.4　個人性の影響 19
3.4.1　開口度 19
3.4.2　円唇性 21
3.4.3　舌運動 21
3.4.4　考察 25
3.5　程度の影響 26
3.5.1　開口度 27
3.5.2　円唇性 29
3.5.3　舌運動 31
3.5.4　考察 33
3.6　母音の影響 33
3.6.1　開口度 33

3.6.2 円唇性 . . . 34
3.6.3 舌運動 . . . 35
3.6.4 考察 . . . 38
3.7 本章のまとめ . . . 39

第 4 章 自発感情音声発話時の調音運動 41
4.1 目的 . . . 41
4.2 解析対象 . . . 41
4.3 感情表現の影響 . . . 42
4.3.1 開口度 . . . 42
4.3.2 円唇性 . . . 43
4.3.3 舌運動 . . . 43
4.3.4 考察 . . . 43
4.4 母音の影響 . . . 45
4.4.1 開口度 . . . 46
4.4.2 円唇性 . . . 46
4.4.3 舌運動 . . . 48
4.4.4 考察 . . . 48
4.5 本章のまとめ . . . 49

第 5 章 演技/自発音声発話時の調音運動の比較 53
5.1 目的 . . . 53
5.2 感情表現が調音運動に対し与える影響の比較 . . . 53
5.2.1 開口度 . . . 54
5.2.2 円唇性 . . . 55
5.2.3 舌運動 . . . 57
5.2.4 考察 . . . 57
5.3 感情表現が言語情報の生成に与える影響の比較 . . . 58
5.3.1 開口度 . . . 58
5.3.2 円唇性 . . . 59
5.3.3 舌運動 . . . 61
5.3.4 考察 . . . 61
5.4 本章のまとめ . . . 64

第 6 章 全体考察 66

第 7 章 まとめと結論 69
7.1 本研究の成果 . . . 69
7.2 今後の課題 . . . 69

図 目 次

2.1 調音点の模式図 7
2.2 日本語母音とその生成に関する舌の位置 7

3.1 Valence-Activation 空間 の概念図 9
3.2 EMA による調音運動計測点 10
3.3 母音 /a/ における上下門歯間の距離の時間変化 11
3.4 上門歯-下門歯間の平均距離と感情表現との関係 14
3.5 上唇-下唇間の平均距離と感情表現との関係 14
3.6 上門歯-上唇間の平均距離と感情表現との関係 16
3.7 下門歯-下唇間の平均距離と感情表現との関係 16
3.8 舌背-上門歯間の平均距離と感情表現との関係 18
3.9 上門歯-下門歯間の平均距離と感情表現及び個人性との関係 20
3.10 上唇-下唇間の平均距離と感情表現及び個人性との関係 21
3.11 上門歯-上唇間の平均距離と感情表現及び個人性との関係 22
3.12 下門歯-下唇間の平均距離と感情表現及び個人性との関係 23
3.13 舌背-上門歯間の平均距離と感情表現及び個人性との関係 25
3.14 舌背-下門歯間の平均距離と感情表現及び個人性との関係 26
3.15 上門歯-下門歯間の平均距離と感情表現及び感情程度との関係 28
3.16 上唇-下唇間の平均距離と感情表現及び感情程度との関係 28
3.17 上唇-上門歯間の平均距離と感情表現及び感情程度との関係 30
3.18 下唇-下門歯間の平均距離と感情表現及び感情程度との関係 30
3.19 舌背-下門歯間の平均距離と感情表現及び感情程度との関係 32
3.20 舌背-上門歯間の平均距離と感情表現及び感情程度との関係 32
3.21 上唇-下唇間の平均距離と感情表現及び母音との関係 35
3.22 上唇-上門歯間の平均距離と感情表現及び母音との関係 36
3.23 下唇-下門歯間の平均距離と感情表現及び母音との関係 36
3.24 舌背-上門歯間の平均距離と感情表現及び母音との関係 38
3.25 舌背-下門歯間の平均距離と感情表現及び母音との関係 39

4.1 上唇-下唇間の平均距離と感情表現との関係 43
4.2 上門歯-上唇間の平均距離と感情表現との関係 44
4.3 下門歯-下唇間の平均距離と感情表現との関係 44

4.4 舌端-上門歯間の平均距離と感情表現との関係 45
4.5 上門歯-下門歯間の平均距離と感情表現及び母音との関係 47
4.6 上唇-下唇間の平均距離と感情表現及び母音との関係 47
4.7 上唇-上門歯間の平均距離と感情表現及び母音との関係 49
4.8 下唇-下門歯間の平均距離と感情表現及び母音との関係 50
4.9 舌端-上門歯間の平均距離と感情表現及び母音との関係 51
4.10 舌端-下門歯間の平均距離と感情表現及び母音との関係 52

5.1 上門歯-下門歯間の平均距離と発話条件との関係 54
5.2 上唇-下唇間の平均距離と発話条件との関係 55
5.3 上門歯-上唇間の平均距離と発話条件との関係 56
5.4 下門歯-下唇間の平均距離と発話条件との関係 56
5.5 舌端-下門歯間の平均距離と発話条件との関係 58
5.6 上門歯-下門歯間の平均距離と発話条件との関係 60
5.7 上唇-下唇間の平均距離と発話条件との関係 60
5.8 上門歯-上唇間の平均距離と発話条件との関係 62
5.9 下門歯-下唇間の平均距離と発話条件との関係 62
5.10 舌端-下門歯間の平均距離と発話条件との関係 63

表 目 次

3.1 開口度に対する 1 要因 (感情表現) 分散分析の結果 12
3.2 開口度に対するテューキーの範囲検定の結果 13
3.3 円唇性に対する 1 要因 (感情表現) 分散分析の結果 13
3.4 円唇性に対するテューキーの範囲検定の結果 15
3.5 舌運動に対する 1 要因 (感情表現) 分散分析の結果 17
3.6 舌運動に対するテューキーの範囲検定の結果 17
3.7 開口度に対する 2 要因 (感情表現, 発話者) 分散分析の結果 20
3.8 円唇性に対する 2 要因 (感情表現, 発話者) 分散分析の結果 22
3.9 舌運動に対する 2 要因 (感情表現, 発話者) 分散分析の結果 24
3.10 開口度に対する 2 要因 (感情表現, 感情程度) 分散分析の結果 27
3.11 円唇性に対する 2 要因 (感情表現, 感情程度) 分散分析の結果 29
3.12 舌運動に対する 2 要因 (感情表現, 感情程度) 分散分析の結果 31
3.13 開口度に対する 2 要因 (感情表現, 母音) 分散分析の結果 34
3.14 円唇性に対する 2 要因 (感情表現, 母音) 分散分析の結果 34
3.15 舌運動に対する 2 要因 (感情表現, 母音) 分散分析の結果 37

4.1 開口度に対する 2 要因 (感情表現, 母音) 分散分析の結果 46
4.2 円唇性に対する 2 要因 (感情表現, 母音) 分散分析の結果 48
4.3 舌運動に対する 2 要因 (感情表現, 母音) 分散分析の結果 51

5.1 開口度に対する 1 要因 (発話条件) 分散分析の結果 54
5.2 円唇性に対する 1 要因 (発話条件) 分散分析の結果 55
5.3 舌運動に対する 1 要因 (発話条件) 分散分析の結果 57
5.4 開口度に対する 2 要因 (発話条件, 母音) 分散分析の結果 59
5.5 円唇性に対する 2 要因 (発話条件, 母音) 分散分析の結果. 61
5.6 舌運動に対する 2 要因 (発話条件, 母音) 分散分析の結果. 63

第1章　序論

1.1　はじめに

ヒトが音声を通じて伝えたい情報には, 音素や語といった言語的情報だけではなく, 感情や個人性といった様々な情報が含まれる [1]. Fujisaki は音声が伝達する情報を, 言語的情報・パラ言語的情報・非言語的情報という 3 つの情報に整理した [2]. 言語的情報とは言語記号が伝達する情報であり, 話者による意識的な制御を受け, 離散的な表現が可能である. パラ言語情報は対話の制御に関わる情報の他, 話者の意図 (強調の有無や程度など) や心的態度 (疑い, 関心, 感心の有無等) に関する情報が含まれる. パラ言語的情報は, 話者によって意図的に制御される点は言語的情報と同一であるが, 単一の意味カテゴリの内部で, 量的かつ連続的な変化 (例えば, ちょっとした「疑い」から強烈な「疑い」まで) が生じうる点において, 離散的な言語情報から区別される. 非言語情報は発話の言語的内容とは独立に存在し, 話者の身体的ないし精神的状態に関するものをさす. 身体的な状態としては, 例えば年齢・性別・個人性・体調の良否など, また精神的状態としては, 感情・気分・性格などが挙げられる. これらの状態のうち, あるものは離散的な範疇 (例えば, 性別や感情) をなすが, その多くは連続的な特徴をもち, また通常は話者が意図的に制御出来ないものである. この中で, 感情に関しては他者への伝達を意図的に行う場合がある. この場合には, 感情自体は意図的に制御されないが, その表出・伝達は意図的に制御される. さらに, 社会的・文化的要因によって, その表出に誇張・抑制・隠蔽, あるいは虚構などの意図が働くこともある. ヒトの音声発話にはこれら様々な情報を同時に音声に表出させることが可能であるという特徴がある. 一方でヒトの音声生成に注目すると, 音声の生成に使用できる調音的パラメータの数は限られている. それではヒトはどのように限られた調音パラメータを使用し, 複数の情報を同時に伝達しうる音声を生成しているのであろうか？この問いに答えることは音声科学における重要な課題である [3].

この問いに答えるためには, 感情表現と言語情報を同時に伝達する音声発話時の調音運動を調査し, どのように感情表現が, 言語情報の生成を阻害または強調することができるのかを明らかにすることが重要である. 感情表現が, どのように言語的情報の生成を妨害または高めることができるのかを明らかにするためには, 言語情報の伝達と感情表現の伝達という二つの発話目的の違いに注目し, 感情音声発話時の調音運動を調査することが必要である.

本研究では発話目的が異なる音声発話時の調音運動を調査し, それぞれの音声発話時の特徴を明らかにすることを目的とする. 音声発話時の発話目的および, その達成のしやす

さが異なる音声発話時の調音運動を調査し, ここで得られた知見を整理, 比較することを通して, 感情表現が, 言語情報の生成にどのような影響を与えるのかに関する知見を得る.

感情音声発話時の感情表現と言語的情報との関係がどのようなものであるのかを知ることができれば, ヒトがどのように限られた調音パラメータを使用し, 複数の情報を同時に伝達しうる音声を生成しているのであろうかという問いの解明に貢献し, 延いては, ヒトの音声生成メカニズムの本質的な特徴の解明にとって重要な知見をもたらすことが期待される.

1.2 背景

1.2.1 先行研究

先行研究において, 感情音声の生成メカニズムの解明のために大きく分けて二通りのアプローチが存在する. 一つは生成された音声の音響特徴量を解析する研究であり, もう一つは, 音声発話時の調音運動を直接的に観察する研究である.

前者のアプローチにより, 感情発話時に生成される音声の音響特徴は感情表現により変化することが分かっている (例えば, [4], [5], [6], [7], [8]). 一方で, 後者のアプローチに関しては, x-ray マイクロ波 [9], 超音波 [10], 舌運動装置 (EMA) [11], [12], そして, MRI [13], [14] など, 様々な機材を使って, 発話時の調音運動を計測する方法を模索している段階であり, 感情音声表現の調査のためのデータ収集環境は充分であるとは言い難い. しかし, 特に EMA を使用し感情音声発話時の調音運動の計測を行ったいくつかの研究では特定の感情表現に対し特徴的な調音運動があらわれることが示されている [15], [16], [3].

感情表現と言語情報との関連を調査した先行研究としては Kim らの研究が存在する. この研究では, 言語情報の生成に重要な調音運動はそうでない調音運動と比較して, 感情表現の影響を受けにくいと仮定し, 特定のシラブル発話時の調音運動を, 特に子音に注目して調査した [17]. 調査の結果, 子音の生成に重要な調音運動は感情表現の影響を受けにくいことが示された. この結果から, Kim らは言語的情報とパラ言語的情報の生成を同時に行うために, 感情音声発話時には, 言語的情報の生成に重要な調音運動以外で感情表現を行うことを主張している.

1.2.2 問題点

感情表現と言語情報の同時生成を行うための調音運動上の特徴を明らかにするために, Kim らは特定の子音発話時における調音運動の観察を行った. その結果, 特定の子音を生成するために重要な調音運動は感情の影響を受けにくいことを示している [17]. 一方, Maekawa, Erickson らの研究では, 母音 /a/ の発話時の調音運動を調査し, 言語的情報の生成に重要な調音運動である舌の位置も, 感情表現の影響を受け, 変化することを示して

いる [15], [3]. これら二つの知見は, 音素の生成に重要な調音運動が感情表現の影響を受けるのかという視点でみると異なる見解を示すように見える.

だたし, Kim らの研究と Maekawa, Erickson らの研究では, 解析対象となる音声の発話条件間に様々な違いが見られる. まず, Kim らと Maekawa, Erickson らの研究では音声発話時に表現している感情の種類が異なる. そのため, 言語情報と感情表現との関係を明らかにするためには統一的な感情表現空間の中で音声発話時の調音運動を比較する必要がある.

また, Kim らは解析の対象として特定の子音に注目しているが, Maekawa, Erickson らの研究では母音発話時を解析の対象としている. 母音と子音とでは, 生成される音声の時間的性質が異なることから知覚に与える影響も異なるものと考えられる. そこで, 特に音響的には長い間定性的な状態を保つ母音を対象に, 言語情報と感情表現との関係を調査する必要がある.

最後に Kim らと Erickson らの研究では音声発話時の発話目的が異なることが想定される, Kim らの研究では, プロの声優が特定の台本を読みながら教示された感情表現をおこなった音声 (演技感情音声) を対象にしているが, Erickson の研究では素人の特に言語情報に対する統制を行っていない音声 (自発感情音声) を解析の対象としている. 前者の条件で収録された音声は, 言語的情報を伝達しつつ, 同時に感情表現を行う必要があるが, 後者の条件で発話された音声の場合, 発話者は言語情報の伝達のみを目的とし, 感情表現の伝達は発話目的に含まれない. このような発話目的の差が, 音素の生成に重要な調音運動が感情表現の影響を受けるか否かのを決定している可能性が存在する.

1.3 目的

前節で述べた過去の研究の問題点を踏まえて, 本研究では発話目的が異なる音声発話時の調音運動を調査し, それぞれの音声発話時の特徴を明らかにすることを目的とする. そして上記の知見を整理比較することで, 発話目的の違いが, 音声発話時の感情表現と言語的情報の生成との関係にどのような影響を与えるのかを考察する.

発話目的の違いが, 音声発話時の感情表現と言語的情報の生成との関係にどのような影響を与えるのかを知ることができれば, ヒトがどのように限られた調音パラメータを使用し, 複数の情報を同時に伝達しうる音声を生成しているのであろうかという問いの解明に貢献し, 延いては, ヒトの音声生成メカニズムの本質的な特徴の解明にとって重要な知見をもたらすことが期待される.

1.4 構成

本論文の構成を以下に示す. 第 1 章では, 本論文が対象としている研究分野の現状と問題点を指摘し, 本論文の位置付けと目的を明らかにする,

第 2 章では, 音声の生成メカニズムについて, 特に言語情報の生成に注目しながら過去の知見をまとめる.

第 3 章では, 演技感情音声の調音運動の解析を通して, 発話目的に言語情報の伝達と感情表現の伝達が同時に含まれる場合の音声発話において, 感情表現が言語的情報の生成にどのような影響を与えるのかを調査する.

第 4 章では, 自発感情音声の調音運動の解析を通して, 感情表現の伝達が発話目的に含まれていないにも拘わらず, 音声に感情が滲みでてしまっている音声発話時に, 感情表現が言語的情報の生成にどのような影響を与えるのかを調査する.

第 5 章では, 演技感情音声及び自発感情音声発話時の調音運動の比較を行い, それぞれの音声発話時の調音運動上の差異を整理する.

第 6 章では, 第 3, 第 4 5 章で得られた知見をまとめ, 音声発話時の発話目的の違いが感情表現を行う調音運動にどのような影響を与えるのかを考察する.

第 7 章では, 本論文で得られた結果を要約し, 今後の課題を示す.

第2章　音声の生成メカニズム

以下に, 本研究で扱う調音運動に関して特に言語情報の生成に注目しながら過去の知見をまとめる.

どのような調音運動がそれぞれの音素の知覚に対し重要であるのかを調査する学問体系の一つに調音音声学 [18] が存在する. 調音音声学では発話過程を調音行動により構成されたものであると考える. 言い換えれば, 音声の言語的な情報は, その言語的文脈に対応した, 声道の狭窄の生成や解放により生成されると考える.

2.1　調音点

上記枠組みにおいて, 声道の狭窄の生成や解放を行う器官は, 一種のパラメータとして考えることができる. 調音点とは肺から唇までの発声器官の中で, 音の区別に大きく係る声道位置のことを指す. 図 2.1 に代表的な調音点を示す. これらの調音点を操作すること (調音運動) で, 母音や子音などの音素の知覚に関連する音響的特徴量は決定される.

2.2　調音運動

本研究では, 日本語母音を前提に解析を行うため, ここでは日本語母音の生成に 2.1 節で紹介した調音点がどのような影響を及ぼすのかを概説する.

まず, 舌の高さと前後方向の位置とによって母音の音色は決定される. 日本語母音とその生成に関する舌の位置を図 2.2 に示す. 日本語 5 母音に関しては, 舌の高さに応じて, 狭母音 (/i, u/), 半狭母音 (/e, o/), 広母音 (/a/) に分類される. 次に, 舌の前後方向の位置により, 前舌母音 (/i, e/) と 後舌母音 (/u, o/), その中間である中舌母音 (/a/) に分類することができる. ただし, これは物理的に舌の位置を計測したものではなく, 聴覚印象上の音の距離によって決められたものである.

例えば, 舌は下顎に物理的に接続されているため, 舌の高さに関しては, 下顎の動きの影響も受ける. また, 唇の突き出し, 円唇性も母音の音色に影響する. 例えば, 日本語母音 /u/ に関しては音声生成時に唇を突き出すが, 母音 /i/ は他の母音と比較し円唇性が弱くなることが知られている.

本研究では上記の知見に従い調音運動の観察箇所を決定し, 感情表現が言語情報の生成にどのような影響を与えるのかを解析していく.

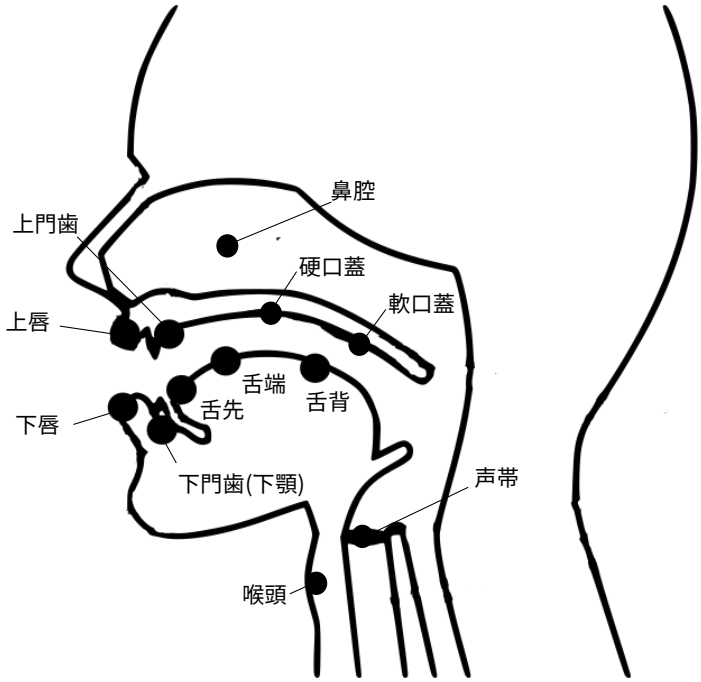


図 2.1: 調音点模式図. 田窪 1998 を参考に作成 [19]. 調音音声学においては, これらの調音点に基づき音素を分類する.

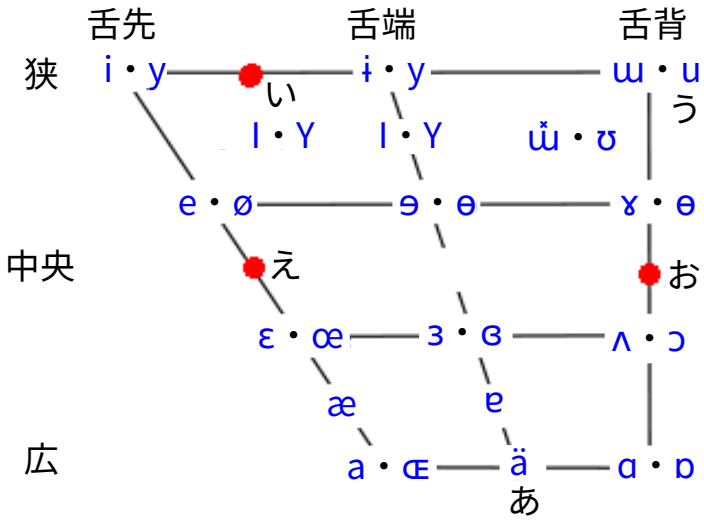


図 2.2: 日本語母音とその生成に関する舌の位置. ただし, これは物理的に舌の位置を計測したものではなく, 聴覚印象上の音の距離によって決められたものである. 田窪 1998 を参考に作成 [19].

第3章　演技感情音声発話時の調音運動

3.1　目的

本研究では, 感情音声発話時の感情表現と言語的情報との関係を解明することを目標にしている. 本章の目的は, 発話目的に言語情報の伝達と感情表現の伝達が同時に含まれる場合の音声発話において, 感情表現が言語的情報の生成にどのような影響を与えるのかを調査することである.

過去の研究では, 2.2 に示した通り, 調音運動と言語情報の生成に関する関係は体系化されつつある. 一方, 感情表現と調音運動との関係に関する研究は特定の感情を表現する際に, 調音点がどのように変化するのかを調査するに留まっており, 表現目標となる感情表現と, 調音運動との関係を体系的に説明するまでには至っていない. 感情音声発話時の感情表現が言語的情報の生成にどのように影響を与えるのかを調査するためには, 一般的な感情表現と調音運動との関係を整理する必要がある.

Vogt らは音声の感情空間が快不快 (Valence) の次元と活性度 (Arousal) の次元の 2 次元空間として整理しうることを示している [20] . Vogt らがまとめた Valence-Arousal 空間 (VA 空間) の概念図を図 3.1 に示す.

本章の解析ではこの知見を踏まえ, この二軸を前提に選択された 5 種類の感情音声を発話する際に調音運動がどのように変化するのかを調査する. まず, 3.3 節では VA 空間上に布置される感情表現が調音運動にどのような影響を与えるのかを整理する. 続く, 3.4 節では感情を表現するための調音運動に個人差が存在するのかを解析する. その他, 感情表現に関わる要素として, 赤木らは音声に表出される感情表現の知覚空間には”ちょっと” や ”かなり” と表現するのに相応しい度合いが存在するを指摘している [21]. そのため, 3.5 節では, 感情表現の程度が調音運動にどのような影響を与えるのかを調査する. 3.5 節では, 演技感情音声発話時の調音運動を言語情報の生成という観点で解析し, 演技音声発話時には母音発話においても Kim らが主張するように調音運動の使い分けを行っているのかを確認する.

3.2　解析対象

上記目的を達成するために, JAIST 舌運動感情音声データベース (JEESD) を作成した. 声優・演劇経験者は一般人に比べ, 音声によって感情状態の表現を行う手法を的確に心得ている [22] ことがわかっているため, 本章で扱う感情音声データはプロの声優を使用する.

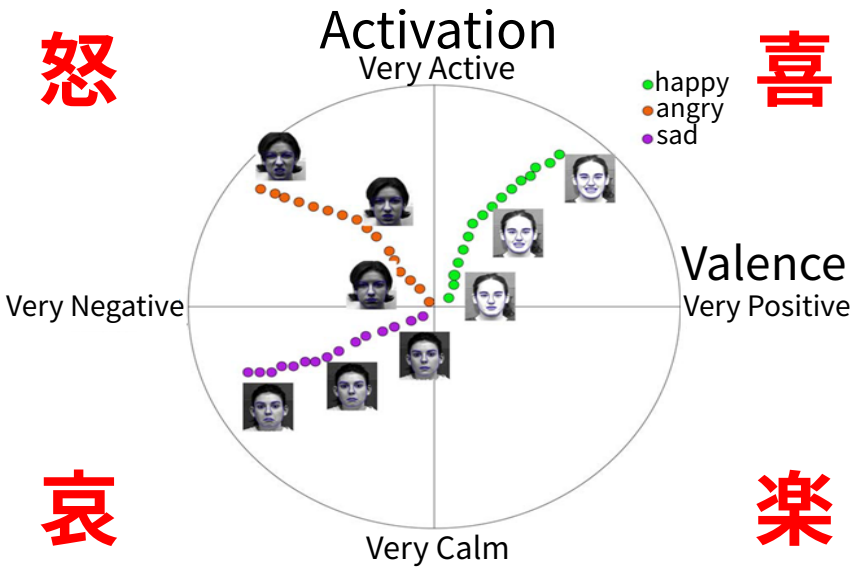


図 3.1: Valence-Activation 空間 の概念図

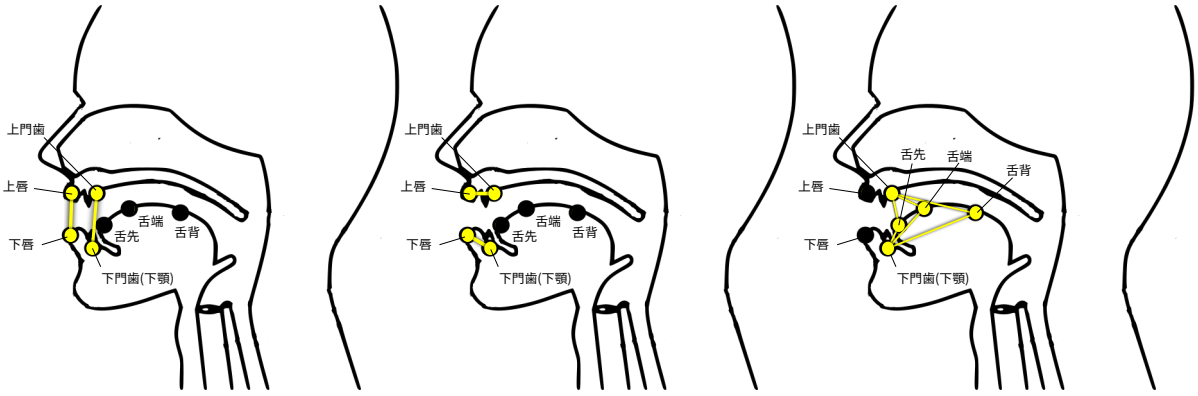
このデータベースには 2 人のプロの声優が, 5 種類の感情表現 (Anger, Joy, Neutral, Relax, Sad) を, 3 通りの程度 (Very, Nomal, Little) で, 11 通りの台本に基づき発話した音声データおよび, EMA により同時に計測された調音データが収録されている. このデータベースに収録した感情表現は, Valence-Arousal 空間 [20] に基づき選択された (図 3.1).

本研究では感情表現の生成と, 言語情報の生成を同時に考慮したいため, 収録された全ての発話に対し,『日本語話し言葉コーパス』に基づき, 分節音ラベル [23] を転記した. この際に, 特に音素境界が曖昧なものに関しては解析から除外している.

EMA で得られる調音データは 3 次元座標軸で表現される座標点としてデータ化されている. なお, これらの座標点はサンプリング周波数 200Hz で収録されており, 50Hz のローパスフィルタにかけられている. また 鼻根点および左右乳様突起の値を利用して, 発話時の顔全体の傾きを正規化した.

本稿では, EMA で得られた調音データを元に, 開口度, 円唇性, 舌運動の 3 類種の観点から調音運動を解析した. 開口度は母音発話時に口がどの位開いているのかを示す指標である. 本研究では, 上下門歯間の距離, および, 上下唇間の距離を算出した (図 3.2a). 円唇性とは唇の突き出しを示す指標である. 本研究では, 上唇-上門歯間の距離, および, 下唇-下門歯間の距離を算出することで円唇性を計測した (図 3.2b). 舌運動は, 母音発話時の舌の位置を示す指標である. 本研究では, 上下門歯と舌先, 舌端, 舌背間の距離を算出した (図 3.2c). この内, 上門歯と舌先, 舌端, 舌背間の距離は舌の高さを, 下門歯と舌先, 舌端, 舌背間の距離は舌の前後方向の位置を示している.

発話中の調音運動はある開始点から, 目的の音素を生成するための構えを行い, 次の音素に移行するものであると考えられる. 調音運動は感情表現のみでなく, 前後音素の影響を受けることから, 母音開始 1/4 の区間および, 母音終端 1/4 の区間に関しては解析から除外した. 参考として典型的な母音 /a/ を発話している際の開口度の時間変化を図 3.3 に



(a) 開口度　(b) 円唇性　(c) 舌運動

図 3.2: EMA による調音運動計測点. 図中黒点は磁気センサーの取り付け位置を, 黄点は解析に使用した計測点を示す.

示す.

3.3 感情表現の影響

本節では, 感情表現が音声発話時の調音運動に与える一般的な影響を整理することを目的とする. そのため JEESD に収録されている音声の内, /だからマナは頭がさらさらだ/ という文章中, 母音 /a/ を発話している際の調音データの解析を行った. 上記目的に従い, 5 種類の感情を, 通常の程度で表現しているデータのみを選択した.

調音データは発話者の口腔内の構造により, 大きく変化することが想定されたため, ここでは 1 名の発話者のみを解析の対象とした. なお解析に使用した母音数は 219 個であった.

感情表現と調音点間の距離との関係を統計的に解釈するために, 音声発話時の感情表現を独立変数とした 1 要因 5 水準の分散分析を行った. 従属変数には, 母音発話時の調音位置を示す指標として調音点間の距離の平均値 (mean) 及び, 調音運動の大きさを示す指標として, 母音始端/終端時の距離の差分 (diff) を使用した. なお, 統計的検定の有意水準はいずれも .01 と定めた.

3.3.1 開口度

開口度に関して, 分散分析の結果を表 3.1 に示す. 分散分析の結果, 上下門歯 (UI-LI) 間, 及び, 上下唇 (UL-LL) 間の距離平均値において, 感情表現間で有意差が確認された (UI-LI: $F(4, 219) = 5.66, p < .01$, UL-LL: $F(4, 219) = 11.37, p < .01$).

UI-LI, 及び, UL-LL の平均値に対するテューキーの範囲検定の結果を表 3.2 に示す. 下位検定の結果, UI-LI に関しては Anger-Relax 間および, Joy - Relax 間に, LL-LL に関

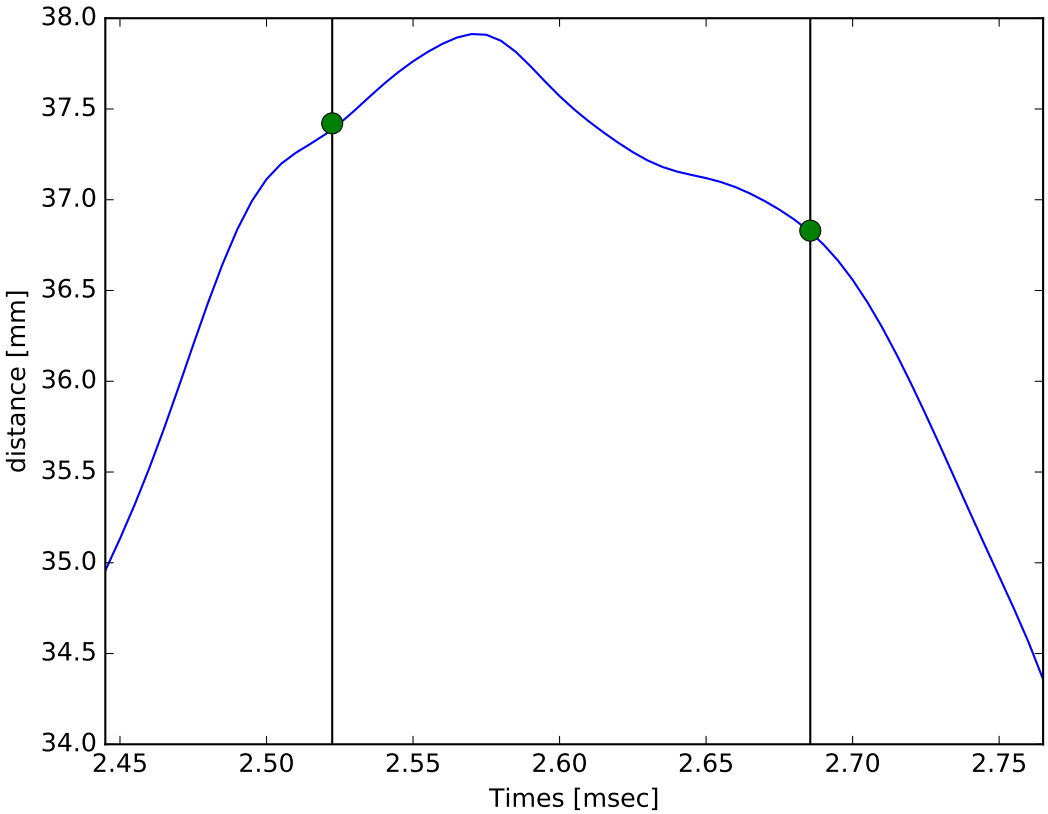


図 3.3: 母音 /a/ における上下門歯間の距離の時間変化例. 縦軸は距離を示し, 横軸は時間を示している. 図中黒線は本研究で使用した解析範囲の始端時刻と終端時刻を示す.

表 3.1: 開口度に関する 1 要因 5 水準分散分析の結果. 独立変数は音声発話時に指定された感情表現. 従属変数は表の通り. ただし, 表中 UL, LL はそれぞれ上下唇を, UI, LI は上下門歯を示す. また, 従属変数には, 調音点間の距離の平均値を使用した場合には mean, 母音始端/終端時の距離の差分を使用した場合には diff と表示した.

Dependent Variable	$d.f.$	F	p
UI-LI (mean)	4, 214	5.66	$p < .01$
UL-LL (mean)	4, 214	11.37	$p < .01$
UI-LI (diff)	4, 214	0.22	$p = .922$
UL-LL (diff)	4, 214	0.69	$p = .598$

しては Anger-Neutral 間, Neutral - Relax 間, そして Neutral - Sad 間に有意差が確認された.

UI-LI 及び, UL-LL の距離平均と感情表現との関係を図 3.4, 3.5 に示す. 図 3.4, 3.5 を確認すると, Relax や Sad のような感情表現を行う際には, Nreutral 時と比較して口の開きが小くなることが確認された. 特に上下門歯間の距離 (図 3.4) に関しては, Anger, Joy, Neutral を表現する場合には Relax, Sad を表現する場合と比べ, 顎を大きく開くことが判明した.

3.3.2 円唇性

円唇性に関して, 分散分析の結果を表 3.3 に示す. 分散分析の結果, 上唇上門歯 (UL-UI) 間, 及び, 下唇下門歯 (LL-LI) 間の距離平均値において, 感情表現間で有意差が確認された (UL-UI: $F(4, 21) = 83.43, p < .01$, LL-LI: $F(4, 21) = 112.29, p < .01$).

UL-UI, 及び, LL-LI の平均値に対するテューキーの範囲検定の結果を表 3.4 に示す. 下位検定の結果, UL-UI, LL-LI 共に Anger-Sad 間および, Neutral - Relax 間以外に対して有意差が確認された.

UL-UI 及び, LL-LI の距離平均と感情表現との関係をそれぞれ, 図 3.6, 3.7 に示す. 図 3.6, 3.7 を確認すると, Anger や Sad のような感情表現を行う際には, Nreutral 時と比較して唇を前に突き出すことが確認された. 一方で Joy を表現する際には唇の突き出しが弱くなることも確認された.

3.3.3 舌運動

舌運動に関して, 分散分析の結果を表 3.5 に示す. 分散分析の結果, 舌背上門歯 (TD-UI) 間の距離平均値においてのみ, 感情表現間で有意差が確認された ($F(4, 21) = 4.59, p < .01$).

TD-UI の平均値に対するテューキーの範囲検定の結果を表 3.6 に示す. 下位検定の結果, Neutral - Relax 間のみに対して有意さが確認された.

表 3.2: 開口度に関するテューキーの範囲検定の結果. group1, 2 は下位検定の水準間の組合せをしめす. meandiff, lower, upper はそれぞれ, 水準間における平均値の差分, 信頼区間の下限値, および上限値を示す. reject は水準間における有意さの有無を示し, 有意さがある場合 True としている. なお, 有意水準は .01 と定めた.

Dependent Variable	**group1**	**group2**	**meandiff**	**lower**	**upper**	**reject**
UI-LI (mean)	Anger	Joy	-0.0482	-1.5604	1.464	False
	Anger	Neutral	-0.2012	-1.7302	1.3277	False
	Anger	Relax	-1.5411	-3.0533	-0.0288	True
	Anger	Sad	-1.3846	-2.8968	0.1276	False
	Joy	Neutral	-0.153	-1.6468	1.3407	False
	Joy	Relax	-1.4929	-2.9695	-0.0162	True
	Joy	Sad	-1.3364	-2.813	0.1403	False
	Neutral	Relax	-1.3398	-2.8335	0.1539	False
	Neutral	Sad	-1.1833	-2.6771	0.3104	False
	Relax	Sad	0.1565	-1.3202	1.6331	False
LL-LI (mean)	Anger	Joy	1.1058	-1.2716	3.4832	False
	Anger	Neutral	3.2529	0.8493	5.6565	True
	Anger	Relax	-0.2425	-2.6198	2.1349	False
	Anger	Sad	-1.2093	-3.5866	1.1681	False
	Joy	Neutral	2.1471	-0.2011	4.4954	False
	Joy	Relax	-1.3483	-3.6697	0.9731	False
	Joy	Sad	-2.3151	-4.6365	0.0063	False
	Neutral	Relax	-3.4954	-5.8436	-1.1471	True
	Neutral	Sad	-4.4622	-6.8104	-2.1139	True
	Relax	Sad	-0.9668	-3.2882	1.3546	False

表 3.3: 円唇性に対する 1 要因 (感情表現) 分散分析の結果. 独立変数は音声発話時に指定された感情表現. 従属変数は表の通り. ただし, 表中 UL, LL はそれぞれ上下唇を, UI, LI は上下門歯を示す. また, 従属変数には, 調音点間の距離の平均値を使用した場合には mean, 母音始端/終端時の距離の差分を使用した場合には diff と表示した.

Dependent Variable	$d.f.$	F	p
UL-UI (mean)	4, 214	83.43	$p < .01$
LL-LI (mean)	4, 214	112.29	$p < .01$
UL-UI (diff)	4, 214	0.61	$p = .65$
LL-LI (diff)	4, 214	1.80	$p = .12$

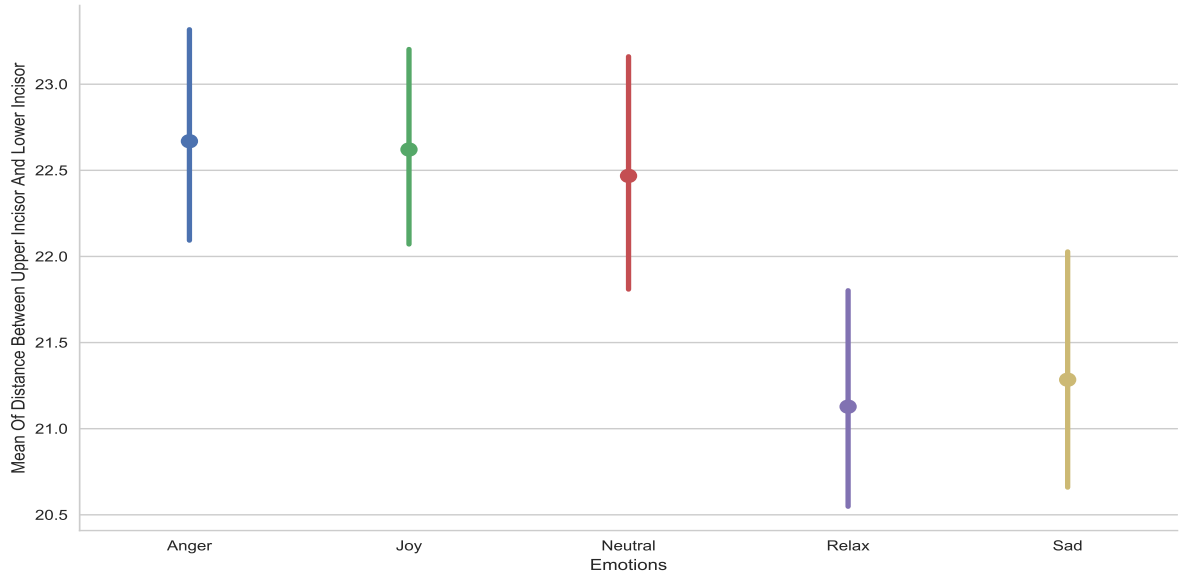


図 3.4: 上門歯-下門歯間の平均距離と感情表現との関係. 縦軸は上門歯-下門歯間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

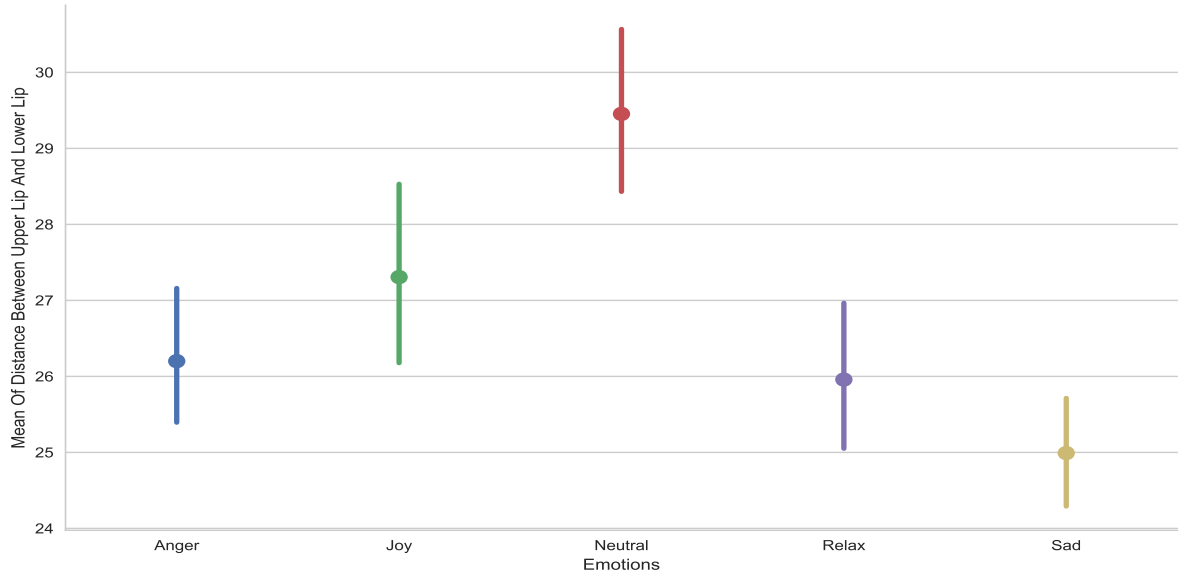


図 3.5: 上唇-下唇間の平均距離と感情表現との関係. 縦軸は上門歯-下門歯間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

表 3.4: 円唇性に対するテューキーの範囲検定の結果. group1, 2 は下位検定の水準間の組合せをしめす. meandiff, lower, upper はそれぞれ, 水準間における平均値の差分, 信頼区間の下限値, および上限値を示す. reject は水準間における有意さの有無を示し, 有意さがある場合 True としている. なお, 有意水準は .01 と定めた.

Dependent Variable	**group1**	**group2**	**meandiff**	**lower**	**upper**	**reject**
UL-UI (mean)	Anger	Joy	-3.2985	-4.0836	-2.5133	True
	Anger	Neutral	-1.1492	-1.9429	-0.3554	True
	Anger	Relax	-0.871	-1.6562	-0.0859	True
	Anger	Sad	0.6765	-0.1086	1.4617	False
	Joy	Neutral	2.1493	1.3738	2.9248	True
	Joy	Relax	2.4274	1.6608	3.1941	True
	Joy	Sad	3.975	3.2083	4.7416	True
	Neutral	Relax	0.2781	-0.4974	1.0537	False
	Neutral	Sad	1.8257	1.0502	2.6012	True
	Relax	Sad	1.5475	0.7809	2.3142	True
LL-LI (mean)	Anger	Joy	-2.6217	-3.1593	-2.0841	True
	Anger	Neutral	-1.9002	-2.4437	-1.3567	True
	Anger	Relax	-1.4878	-2.0254	-0.9502	True
	Anger	Sad	0.1461	-0.3915	0.6836	False
	Joy	Neutral	0.7215	0.1905	1.2525	True
	Joy	Relax	1.1339	0.609	1.6588	True
	Joy	Sad	2.7678	2.2429	3.2927	True
	Neutral	Relax	0.4124	-0.1186	0.9434	False
	Neutral	Sad	2.0463	1.5153	2.5773	True
	Relax	Sad	1.6339	1.109	2.1588	True

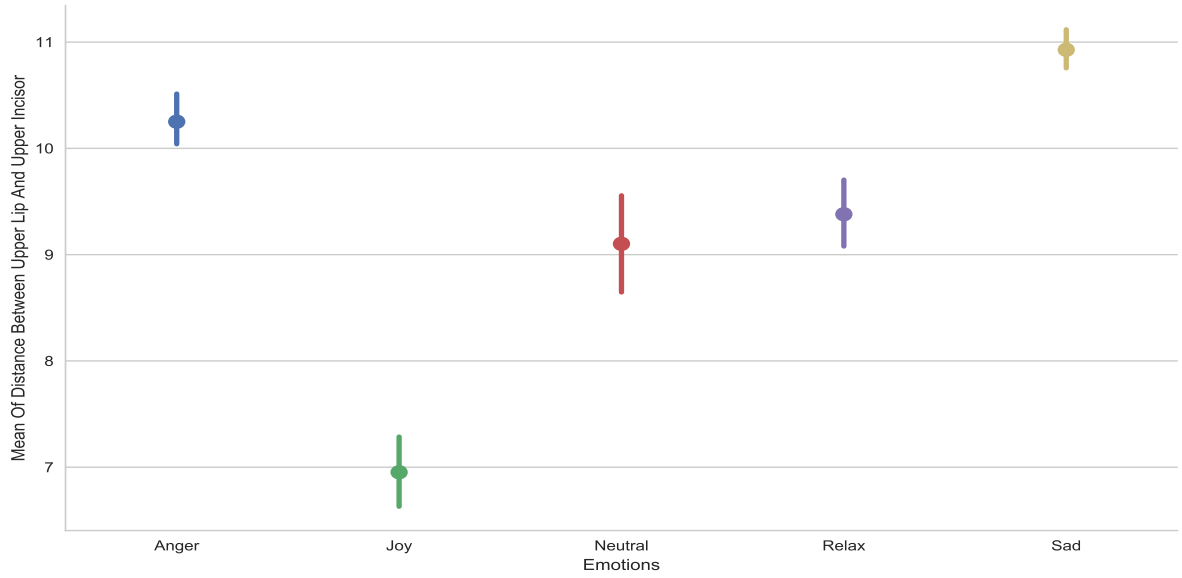


図 3.6: 上門歯-上唇間の平均距離と感情表現との関係. 縦軸は上門歯-上唇間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

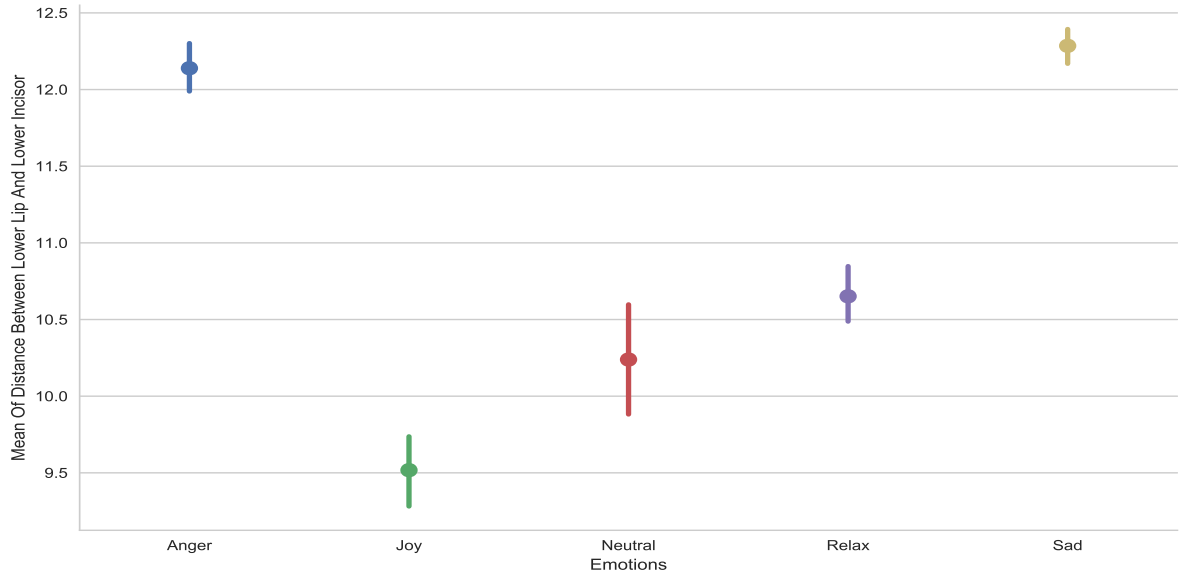


図 3.7: 下門歯-下唇間の平均距離と感情表現との関係. 縦軸は下門歯-下唇間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

表 3.5: 舌運動に対する 1 要因 (感情表現) 分散分析の結果. 独立変数は音声発話時に指定された感情表現. 従属変数は表の通り. ただし, 表中 TT は 舌先, TB は舌端, TD は 舌背 を示し, UI, LI はそれぞれ上下門歯を示す. また, 従属変数には, 調音点間の距離の平均値を使用した場合には mean, 母音始端/終端時の距離の差分を使用した場合には diff と表示した.

Dependent Variable	$d.f.$	F	p
TT-UI (mean)	4, 214	1.06	$p = .37$
TT-LI (mean)	4, 214	1.52	$p = .19$
TB-UI (mean)	4, 214	1.63	$p = .16$
TB-LI (mean)	4, 214	1.34	$p = .25$
TD-UI (mean)	4, 214	4.59	$p < .01$
TD-LI (mean)	4, 214	1.25	$p = .28$
TT-UI (diff)	4, 214	0.00	$p = .99$
TT-LI (diff)	4, 214	0.10	$p = .98$
TB-UI (diff)	4, 214	0.01	$p = .99$
TB-LI (diff)	4, 214	0.09	$p = .98$
TD-UI (diff)	4, 214	0.01	$p = .99$
TD-LI (diff)	4, 214	0.02	$p = .99$

表 3.6: 舌運動に対するテューキーの範囲検定の結果. group1, 2 は下位検定の水準間の組合せをしめす. meandiff, lower, upper はそれぞれ, 水準間における平均値の差分, 信頼区間の下限値, および上限値を示す. reject は水準間における有意さの有無を示し, 有意さがある場合 True としている. なお, 有意水準は .01 と定めた.

Dependent Variable	**group1**	**group2**	**meandiff**	**lower**	**upper**	**reject**
TD-UI (mean)	Anger	Joy	-0.6595	-1.6981	0.3791	False
	Anger	Neutral	-0.9746	-2.0246	0.0754	False
	Anger	Relax	0.0581	-0.9804	1.0967	False
	Anger	Sad	-0.8039	-1.8424	0.2347	False
	Joy	Neutral	-0.3151	-1.341	0.7108	False
	Joy	Relax	0.7176	-0.2965	1.7318	False
	Joy	Sad	-0.1444	-1.1585	0.8698	False
	Neutral	Relax	1.0327	0.0069	2.0586	True
	Neutral	Sad	0.1707	-0.8551	1.1966	False
	Relax	Sad	-0.862	-1.8761	0.1521	False

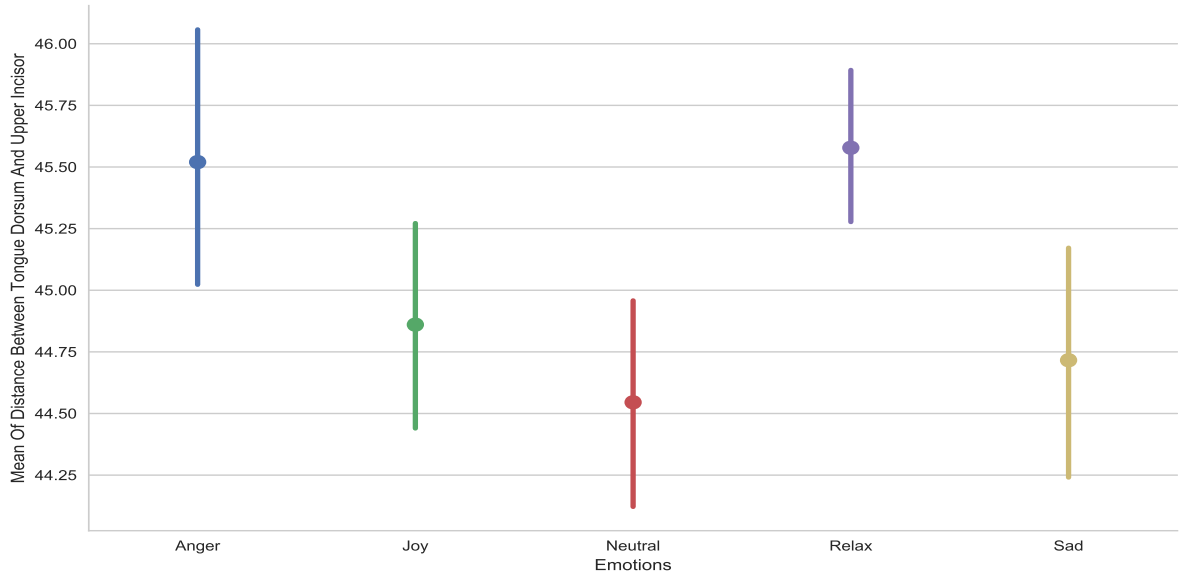


図 3.8: 舌背-上門歯間の平均距離と感情表現との関係. 縦軸は舌背-上門歯間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

TD-UI の距離平均と感情表現との関係を図 3.8 に示す. 図 3.8 を確認すると, Anger や Relax のような感情表現を行う際には, Nreutral 時と比較して舌背の位置が高くなっていることが確認された.

3.3.4 考察

本節の目的は感情表現が音声発話時の調音運動に与える一般的な影響を整理することであった. そのため, 本節では, 開口度, 円唇性, 舌運動のそれぞれの観点から, 感情表現が調音運動にどのような影響を与えるのかを調査した.

まず分散分析の結果, いずれの観点においても感情表現により, 調音点間の距離の平均値変化することが示された. 一方で, 感情表現による母音始端/終端時の距離の差分値の変化は確認されなかった. これは, 母音発話時においては, 感情表現は調音運動の大きさよりも母音発話時の調音位置に強く影響することを示唆する結果である. そのため, 以後の解析では, 調音点間の距離の平均値を中心に解析をしていくこととする.

つづいて, 感情表現が開口度, 円唇性, 舌運動のそれぞれにどのような影響を与えるのかを考察する. 開口度に関する解析では, Relax や Sad のような感情表現を行う場合には口の開きが小くなり, Anger や Joy のような感情表現を行う場合には口の開きが大きくなることが分かった. この対比は 感情音声の V-A 空間においては活性度の軸に相当する. 円唇性に関する解析では, Anger や Sad のような感情表現を行う場合には唇を突き出し, Joy を表現する際には唇の突き出しが弱くなることが確認された. この対比は 感情音声

の V-A 空間においては快不快の軸に相当する. 舌運動に関する解析では舌背の縦方向の位置が感情表現により影響されることが確認された. Anger や Relax のような感情表現を行う際には, 舌背の位置が高くなっていることが確認された. だたし, 舌の縦方向の位置は下顎の位置に依存する. そのため, 開口度に関する解析と合わせて考えると, 特に Relax を表現するために舌背の高さを変化させるものだと考察できる.

3.4 個人性の影響

本節では, 感情表現と調音運動との関係に関する個人差の影響を解析するために, JEESD に収録されている音声の内, /だからマナは頭がさらさらだ/ という文章中, 母音 /a/ を発話している際の調音データ二名分の解析を行った.

感情表現と調音点間の距離との関係を統計的に解釈するために, 5 種類の感情表現と 2 名の発話者を独立変数とした 2 要因分散分析を行った. 統計的検定の有意水準はいずれも .01 と定めた. 従属変数には, 母音発話時の調音位置を示す指標として調音点間の距離の平均値を使用した. ここで, データ収録時の欠損値等の影響から, 両名のデータセット数は異なる. このため, 分散分析の解析にはタイプ II 平方和を利用した. なお解析に使用した母音数はそれぞれ 219 個 および 252 個であった.

3.4.1 開口度

開口度に関して, 二要因分散分析の結果を表 3.7 に示す. 分散分析の結果, 上下門歯 (UI-LI) 間, 及び, 上下唇 (UL-LL) 間の距離平均値に対する感情表現の主効果が確認された (UI-LI: $F(4, 461) = 21.35, p < .01$, UL-LL: $F(4, 461) = 16.47, p < .01$). また両観測点ともに, 発話者の違いによる主効果が確認された (UI-LI: $F(1, 461) = 140.20, p < .01$, UL-LL: $F(1, 461) = 0.34, p < .01$). くわえて, 感情表現と発話者間には交互作用が確認された (UI-LI: $F(4, 461) = 4.84, p < .01$, UL-LL: $F(4, 461) = 12.64, p < .01$).

それぞれの個人差がどのように開口度に影響を与えるのかを解析するために, UI-LI 及び, UL-LL の距離平均と感情表現との関係を可視化した (図 3.9, 図 3.10). 結果を確認すると, 図 3.9, 3.10 ともに Neutral 表現時の開口度が異なることが確認された. また, Relax, Sad 表現時には両発話者ともに Neutral 表現時と比較し開口度が小くなることが確認された. 開口度において, 発話者の個人差が確認される感情表現は Anger 及び, Joy であった. 図 3.9 においては, Speaker 1 では Neutral に対する Anger, Joy 条件の開口度は小くなっているに対し, Speaker 2 では 大きくなっている. 一方, 図 3.10 を確認すると Speaker 2 では Neutral に対する Anger, Joy 条件の開口度は小くなっているに対し, Speaker 1 では大きくなっている.

表 3.7: 開口度にに対する 2 要因 (感情表現, 発話者) 分散分析の結果. 独立変数は音声発話時に指定された感情表現, および, 発話者. 従属変数は表の通り. ただし, 表中 UI, LI はそれぞれ上下門歯を, UL, LL はそれぞれ上下唇を示す.

Dependent Variable	Independent Variable	$d.f.$	F	p
UI-LI (mean)	emotion	4, 461	21.35	$p < .01$
	speaker	1, 461	140.20	$p < .01$
	emotion:speaker	4, 461	4.84	$p < .01$
UL-LL (mean)	emotion	4, 461	16.47	$p < .01$
	speaker	1, 461	0.34	$p < .01$
	emotion:speaker	4, 461	12.60	$p < .01$

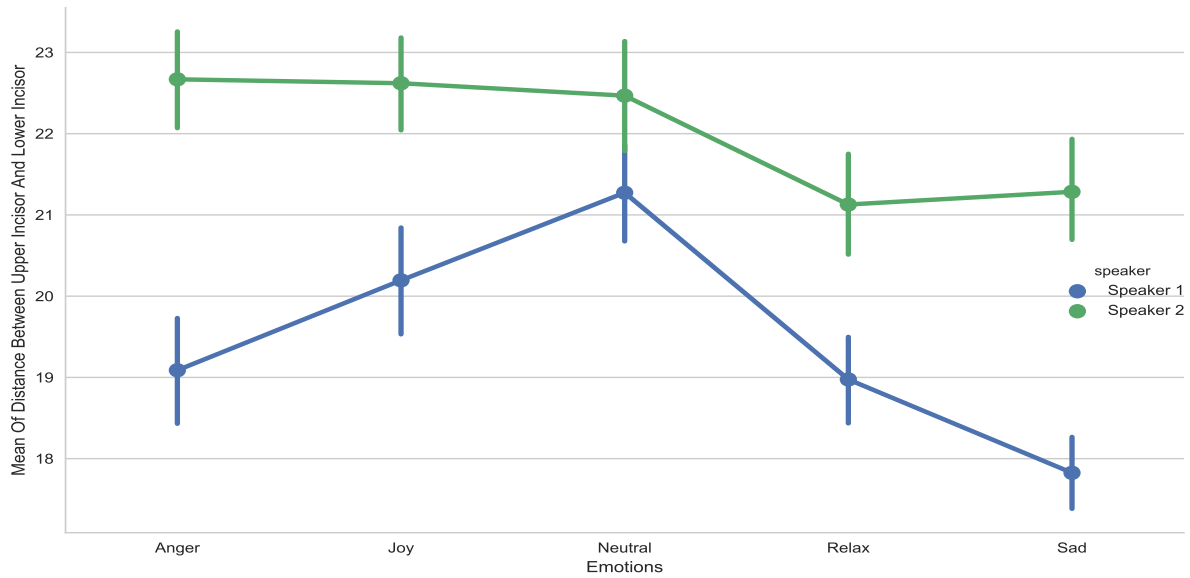


図 3.9: 上門歯-下門歯間の平均距離と感情表現及び個人性との関係. 縦軸は上門歯-下門歯間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中色分けはそれぞれの発話者を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

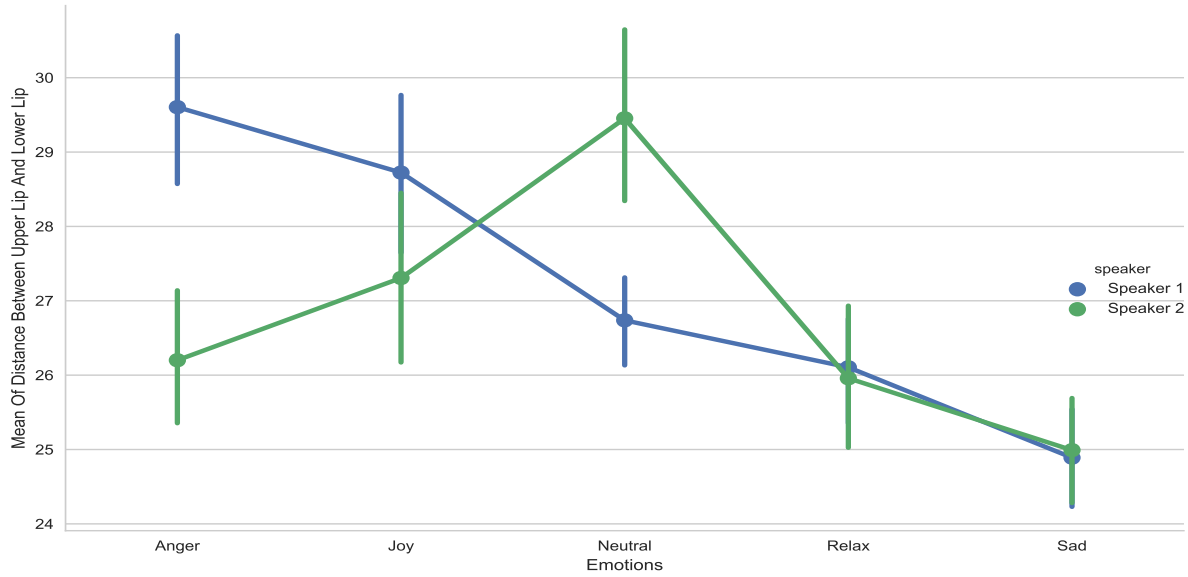


図 3.10: 上唇-下唇間の平均距離と感情表現及び個人性との関係. 縦軸は上門歯-下門歯間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中色分けはそれぞれの発話者を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

3.4.2 円唇性

円唇性に関して, 二要因分散分析の結果を表 3.8 に示す. 分散分析の結果, 上唇上門歯 (UL-UI) 間, 及び, 下唇下門歯 (LL-LI) 間の距離平均値において, 感情表現の主効果が確認された (UL-UI: $F(4, 461) = 114.56, p < .01$, LL-LI: $F(4, 461) = 133.41, p < .01$). また, 両観測点ともに発話者の違いによる主効果が確認された (UL-UL: $F(1, 461) = 121.56, p < .01$, LL-LI: $F(1, 461) = 53.58, p < .01$). くわえて, 感情表現と発話者の間には交互作用が確認された (UL-UI: $F(4, 461) = 35.68, p < .01$, LL-LI: $F(4, 461) = 45.27, p < .01$).

それぞれの個人差がどのように円唇性に影響を与えるのかを解析するために, UL-UI 及び, LL-LI の距離平均と感情表現との関係を可視化した (図 3.11, 図 3.12). 結果を確認すると, 図 3.11, 3.12 ともに Neutral 表現時の円唇性が異なることが確認された. また, Joy 表現時には両発話者ともに Neutral 表現時と比較し円唇性が小くなることが確認された. くわえて, Anger 表現時には両発話者ともに Neutral 表現時と比較し円唇性が大きくなることが確認された. 円唇性において, 発話者の個人差が確認される感情表現は Relax 及び, Sad であった. 図 3.11, 3.12 ともに Speaker 1 は Relax, Sad 表現時に唇の突き出しを弱めるのに対し, Speaker 2 は唇を突き出す動きをしていることが確認された.

3.4.3 舌運動

舌運動に関して, 二要因分散分析の結果を表 3.9 に示す. 分散分析の結果, 舌先上門歯 (TT-UI), 舌先下門歯 (TT-LI), 舌端上門歯 (TB-UI), 舌端下門歯 (TB-LI), 舌背上門歯 (TD-

表 3.8: 円唇性に対する 2 要因 (感情表現, 発話者) 分散分析の結果独立変数は音声発話時に指定された感情表現, および発話者. 従属変数は表の通り. ただし, 表中 UL, LL はそれぞれ上下唇を, UI, LI は上下門歯を示す.

Dependent Variable	Independent Variable	$d.f.$	F	p
UL-UI (mean)	emotion	4, 461	114.56	$p < .01$
	speaker	1, 461	121.56	$p < .01$
	emotion:speaker	4, 461	35.68	$p < .01$
LL-LI (mean)	emotion	4, 461	133.41	$p < .01$
	speaker	1, 461	53.58	$p < .01$
	emotion:speaker	4, 461	45.27	$p < .01$

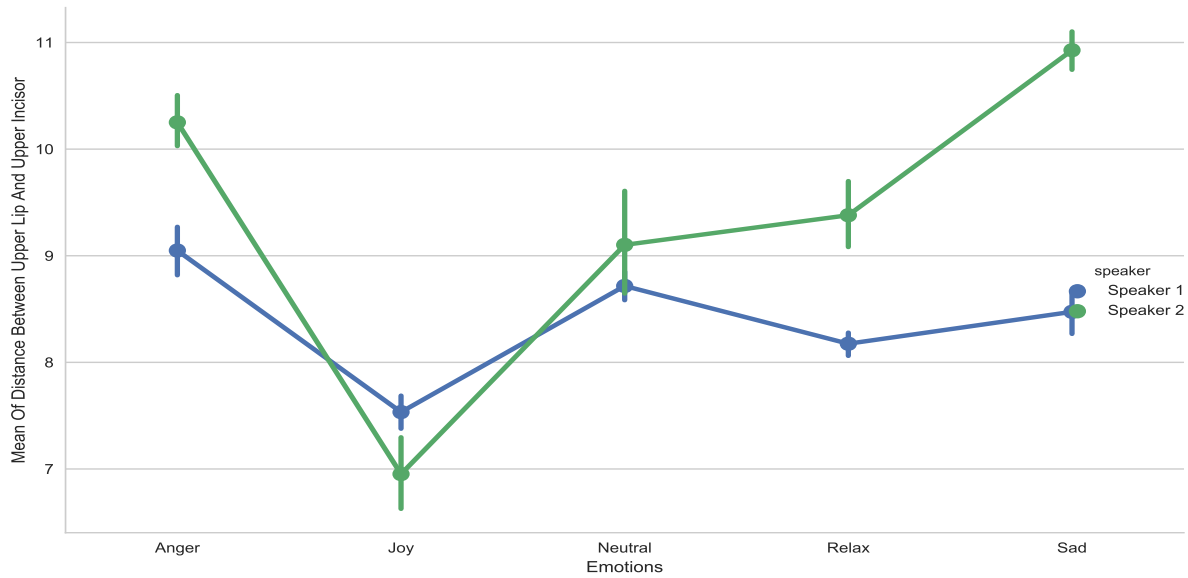


図 3.11: 上門歯-上唇間の平均距離と感情表現及び個人性との関係. 縦軸は上門歯-上唇間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中色分けはそれぞれの発話者を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

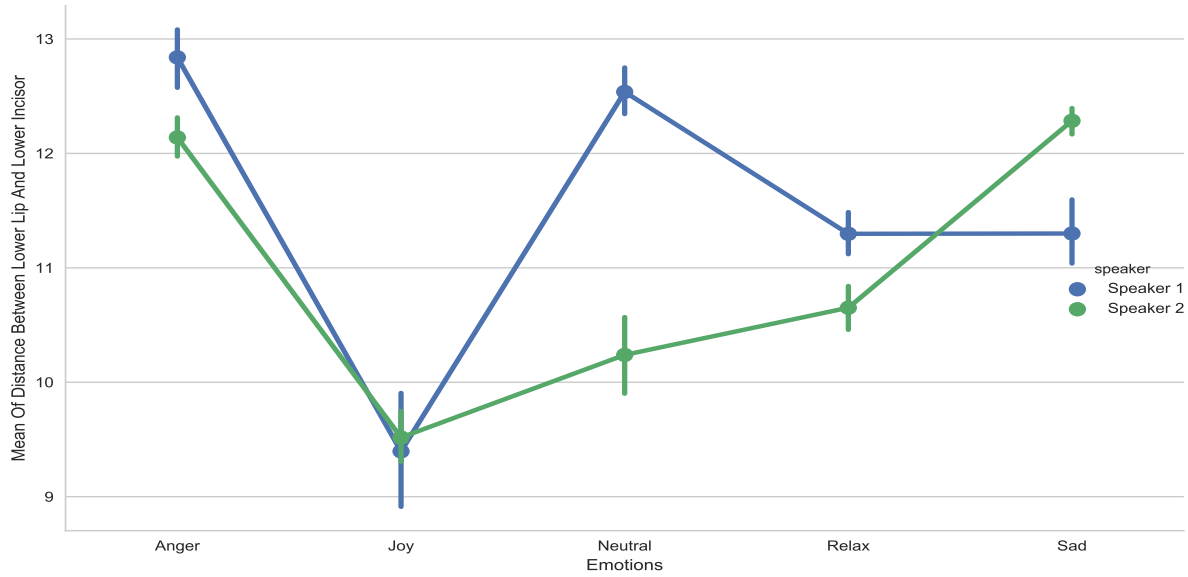


図 3.12: 下門歯-下唇間の平均距離と感情表現及び個人性との関係. 縦軸は下門歯-下唇間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中色分けはそれぞれの発話者を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

UI) そして 舌背下門歯 (TD-LI) 間の距離平均値において, 感情表現の主効果が確認された (TT-UI: $F(4, 461) = 2.64, p < .01$, TT-LI: $F(4, 461) = 1.19, p < .01$, TB-UI: $F(4, 461) = 3.05, p < .01$, TB-LI: $F(4, 461) = 1.20, p < .01$, TD-UI: $F(4, 461) = 3.69, p < .01$, TD-LI: $F(4, 461) = 2.58, p < .01$). また, これら全ての観測点に対し発話者の違いによる主効果が確認された. (TT-UI: $F(1, 461) = 5.33, p < .01$, TT-LI: $F(1, 461) = 99.37, p < .01$, TB-UI: $F(1, 461) = 456.47, p < .01$, TB-LI: $F(1, 461) = 126.26, p < .01$, TD-UI: $F(1, 461) = 2000.13, p < .01$, TD-LI: $F(1, 461) = 357.143, p < .01$). くわえて, 感情表現と発話者条件の間には交互作用が確認された (TT-UI: $F(4, 461) = 2.00, p < .01$, TT-LI: $F(4, 461) = 2.08, p < .01$, TB-UI: $F(4, 461) = 3.25, p < .01$, TB-LI: $F(4, 461) = 0.60, p < .01$, TD-UI: $F(4, 461) = 6.96, p < .01$, TD-LI: $F(4, 461) = 1.27, p < .01$).

それぞれの個人差がどのように舌運動に影響を与えるのかを解析するために, TD-UI 及び, TD-LI の距離平均と感情表現との関係を可視化した (図 3.13, 図 3.14). 結果を確認すると, 図 3.13, 3.14 ともに Neutral 表現時の舌の位置が異なることが確認された. また, 縦方向の舌の位置 (図 3.13) に関しては, Anger, Joy を表現している際には Neutral と比較して低い位置に存在することが確認された. 縦方向の舌の位置に関して, 発話者の個人差が確認される感情表現は Relax 及び, Sad であった. Speaker 1 は Relax, Sad を表現する際には, Neutral 表現時と比べ, 舌を位置を低くするのに対し, Speaker 2 は高くすることが確認された. 一方, 前後方向の舌の位置 (図 3.14) に関しては, 両発話者ともに, Anger, Sad を表現している際に, Neutral と比較して舌を後ろに位置に移動させることが確認された. また, Speaker 1 は Joy 表現を行う際に Neutral と比較して, 舌を後退させるのに

表 3.9: 舌運動に対する 2 要因 (感情表現, 発話者) 分散分析の結果. 独立変数は音声発話時に指定された感情表現, および 発話者. 従属変数は表の通り. ただし, 表中 TT は 舌先, TB は舌端, TD は 舌背 を示し, UI, LI はそれぞれ上下門歯を示す.

Dependent Variable	Independent Variable	$d.f.$	F	p
TT-UI (mean)	emotion	4, 461	2.64	$p = .03$
	speaker	1, 461	5.33	$p = .02$
	emotion:speaker	4, 461	2.00	$p = .09$
TT-LI (mean)	emotion	4, 461	1.19	$p < .01$
	speaker	1, 461	99.37	$p < .01$
	emotion:speaker	4, 461	2.08	$p < .01$
TB-UI (mean)	emotion	4, 461	3.05	$p < .01$
	speaker	1, 461	456.47	$p < .01$
	emotion:speaker	4, 461	3.25	$p < .01$
TB-LI (mean)	emotion	4, 461	1.20	$p < .01$
	speaker	1, 461	126.26	$p < .01$
	emotion:speaker	4, 461	0.60	$p < .01$
TD-UI (mean)	emotion	4, 461	3.69	$p < .01$
	speaker	1, 461	2000.13	$p < .01$
	emotion:speaker	4, 461	6.96	$p < .01$
TD-LI (mean)	emotion	4, 461	2.58	$p < .01$
	speaker	1, 461	357.14	$p < .01$
	emotion:speaker	4, 461	1.27	$p < .01$

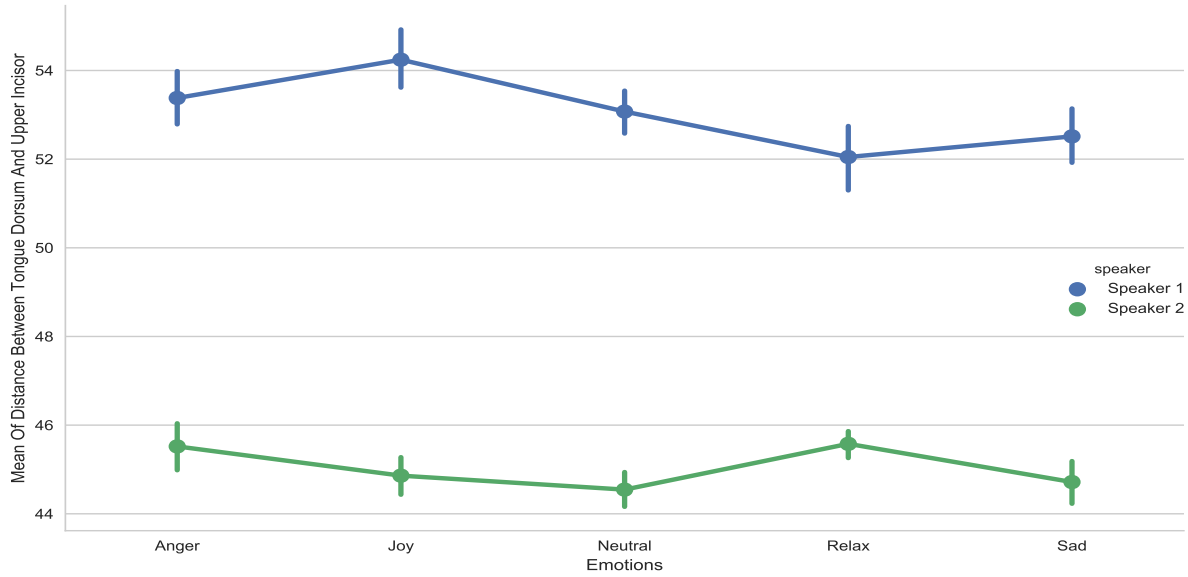


図 3.13: 舌背-上門歯間の平均距離と感情表現及び個人性との関係. 縦軸は舌背-上門歯間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中色分けはそれぞれの発話者を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

対し, Speaker 2 では, Relax 表現を行う際に 舌を後退させることが確認された.

3.4.4 考察

本節の目的は調音運動の観点から感情表現の個人差を明らかにすることであった.

開口度に関しては, 特に活性度が高い感情表現を行っている際には, 個人差が生じることが確認された. ただし, 活性度が低い感情表現時と比べ, 活性度の高い感情表現を行う際には, 開口度は大きくなる傾向が存在した. 円唇性に関して, 主に活性度の低い感情表現を行う際には, 個人差が生じることが確認された. 一方, 快不快の大小のみを考慮して観察すると, 不快な感情表現を行う際には, 唇を突き出す傾向がみられた. 舌の縦方向の位置は活性度の低い感情を表現する場合に個人差が生じることが示唆された. 一方, 舌の前後方向の位置は, 快不快軸上で快にあたる感情表現する場合に個人差が生じることが示唆された.

これらの結果を考察すると, 感情表現に対する調音運動上の個人差には以下のような特徴が存在すると考えられる. ある感情を表現するために開口度, 円唇性, 舌運動をどのように使用するのかは表現者による違いが存在する. この違いは, 特に活性度表現と関連して選択がなされている可能性がある. 一方, いずれの調音運動に対しても, 3.3 節で発見したような調音運動上の対比関係は保持されていた. この結果は, 感情表現を離散的なカテゴリととらえるのではなく, ある種の空間であると仮定した場合, その感情表現空間上の軸を表現するための, 個人差によらない一般的な調音運動が存在するのだという考えを支

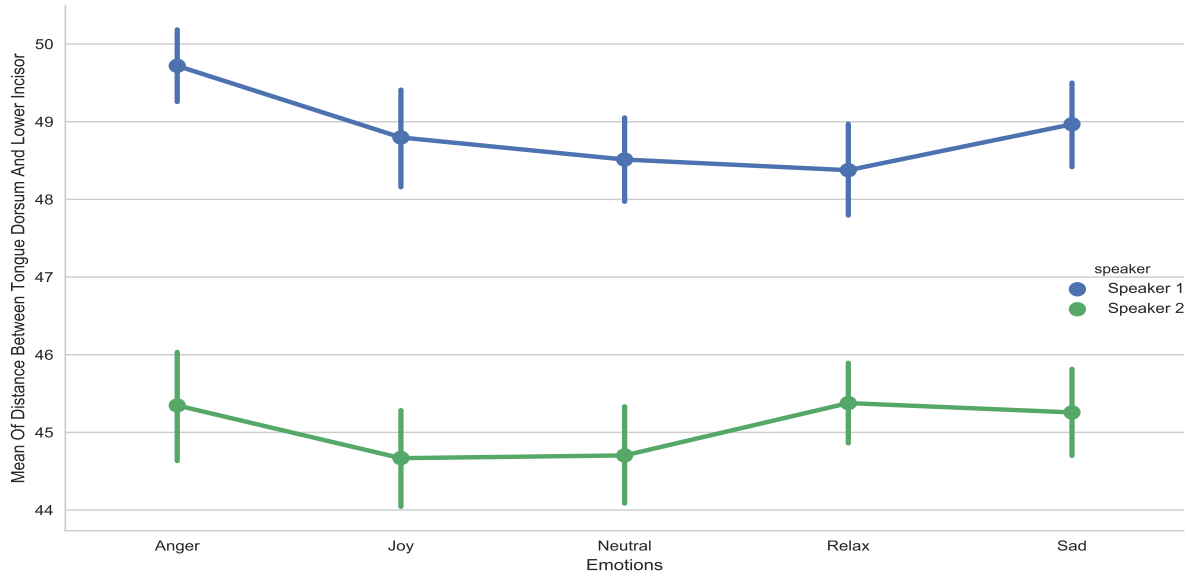


図 3.14: 舌背-下門歯間の平均距離と感情表現及び個人性との関係. 縦軸は舌背-下門歯間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中色分けはそれぞれの発話者を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

持する結果である.

3.5 程度の影響

本節では, 感情表現と調音運動との関係に関する程度表現の影響を解析するために, JEESD に収録されている音声の内, /だからマナは頭がさらさらだ/ という文章中, 母音 /a/ を発話している際の調音データの解析を行った. ここでは, 五種類の感情表現を Little, Normal, Very の三種類の程度で表現した音声発話時の調音運動を比較する. なお解析に使用した母音数はそれぞれ Little 174 個, Normal 219 個, および Very 169 個であった.

感情表現と調音点間の距離との関係を統計的に解釈するために, 感情表現要因と程度表現要因を独立変数とした 2 要因分散分析を行った. ここで, JEEAD に収録されている感情表現の内, Neutral 表現に関しては, 程度表現が存在しないため, 統計的解析からは除外した. 従属変数には, 母音発話時の調音位置を示す指標として調音点間の距離の平均値を使用した. また, 各程度表現数は, それぞれの感情表現により異なるため, データセット数は異なる. このため, 分散分析の解析にはタイプ II 平方和を利用した. なお, 統計的検定の有意水準はいずれも .01 と定めた.

表 3.10: 開口度に対する 2 要因 (感情表現, 感情程度) 分散分析の結果. 独立変数は音声発話時に指定された感情表現, および, 発話者. 従属変数は表の通り. ただし, 表中 UI, LI はそれぞれ上下門歯を, UL, LL はそれぞれ上下唇を示す.

Dependent Variable	Independent Variable	$d.f.$	F	p
UI-LI (mean)	emotion	3, 507	29.05	$p < .01$
	degree	2, 507	5.39	$p < .01$
	emotion:degree	6, 507	2.00	$p < .01$
UL-LL (mean)	emotion	3, 507	15.24	$p < .01$
	degree	2, 507	4.48	$p < .01$
	emotion:degree	6, 507	1.19	$p < .01$

3.5.1 開口度

開口度に関して, 二要因分散分析の結果を表 3.10 に示す. 分散分析の結果, 上下門歯 (UI-LI) 間, 及び, 上下唇 (UL-LL) 間の距離平均値において, 感情表現の主効果が確認された (UI-LI: $F(3, 507) = 29.05, p < .01$, UL-LL: $F(3, 507) = 15.24, p < .01$). また, 両観察点ともに程度表現の主効果が確認された (UI-LI: $F(2, 507) = 5.39, p < .01$, UL-LL: $F(2, 507) = 0.88, p < .01$). くわえて, 感情表現と程度表現の間には交互作用が確認された (UI-LI: $F(6, 507) = 2.00, p < .01$, UL-LL: $F(6, 507) = 1.19, p < .01$).

それぞれの程度表現がどのように開口度に影響を与えるのかを解析するために, UI-LI 及び, UL-LL の距離平均と感情表現との関係を可視化した (図 3.15, 図 3.16). なお, 比較のため, 上記二つの図においては Neutral 表現時の距離平均値を図にくわえている.

結果を確認すると, 図 3.15 においては, 全ての程度表現において, Joy 表現を行う際に, Neutral 表現時と比較して, 開口度が高くなり, Relax, Sad 表現を行う際には, 開口度が低くなることが確認された. 一方で, Normal, Very 程度で Anger 表現を行う場合には, Neutral より開口度が大きくなるのに対し, Little 程度で表現を行う場合には, 開口度が小くなることが確認された. また, 図 3.16 においては, 全ての感情表現を行う場合にも, Neutral 表現時と比較して開口度が小くなることが確認された. 特に, Anger, Sad 表現を Little 程度で表現する場合には, 開口度が減少が激しくなることが確認された.

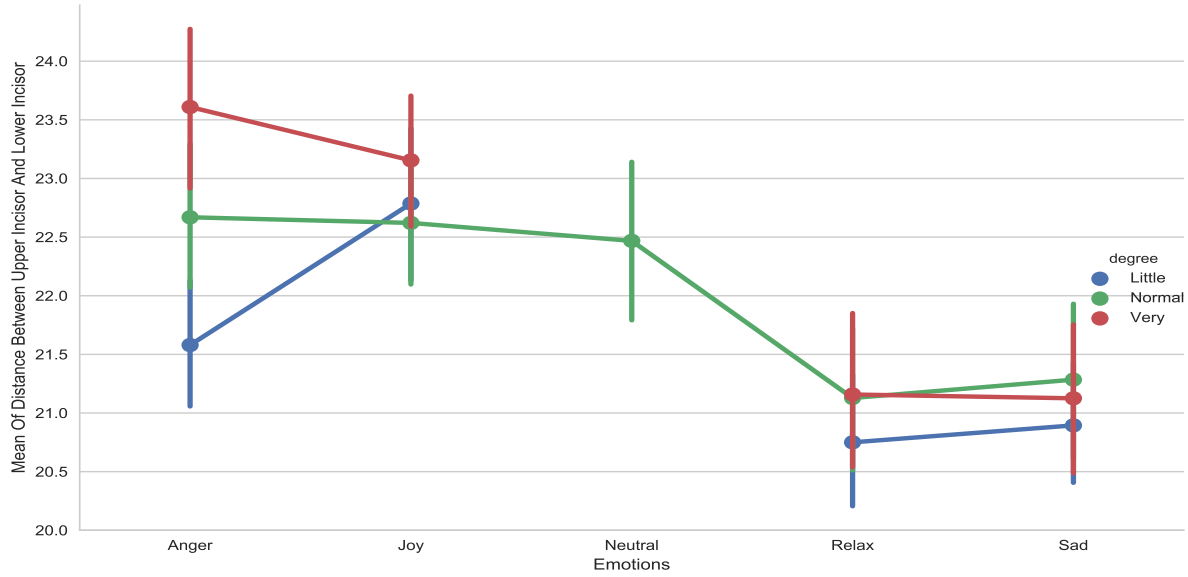


図 3.15: 上門歯-下門歯間の平均距離と感情表現及び感情程度との関係. 縦軸は上門歯-下門歯間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中色分けはそれぞれの程度表現を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

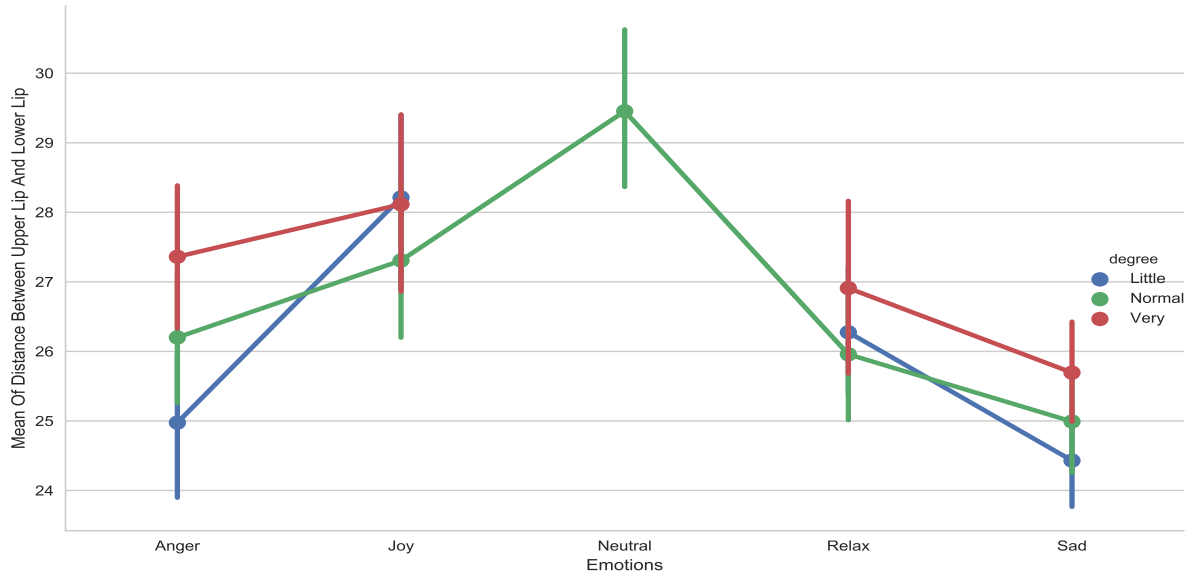


図 3.16: 上唇-下唇間の平均距離と感情表現及び感情程度との関係. 縦軸は上唇-下唇間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中色分けはそれぞれの程度表現を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

表 3.11: 円唇性に対する 2 要因 (感情表現, 感情程度) 分散分析の結果独立変数は音声発話時に指定された感情表現, および程度表現. 従属変数は表の通り. ただし, 表中, UI, LI は上下門歯を, UL, LL はそれぞれ上下唇を示す.

Dependent Variable	Independent Variable	$d.f.$	F	p
UL-UI (mean)	emotion	3, 507	316.29	$p < .01$
	degree	2, 507	1.69	$p < .01$
	emotion:degree	6, 507	3.79	$p < .01$
LL-LI (mean)	emotion	3, 507	507.34	$p < .01$
	degree	2, 507	1.47	$p < .01$
	emotion:degree	6, 507	3.81	$p < .01$

3.5.2 円唇性

開口度に関して, 二要因分散分析の結果を表 3.11 に示す. 分散分析の結果, 上唇上門歯 (UL-UI) 間, 及び, 下唇下門歯 (LL-LI) 間の距離平均値において, 感情表現の主効果が確認された (UI-LI: $F(3, 507) = 316.29, p < .01$, UL-LL: $F(3, 507) = 507.34, p < .01$). また, 両観測点ともに発話者の違いによる主効果が確認された (UI-LI: $F(2, 507) = 1.69, p < .01$, UL-LL: $F(2, 507) = 1.47, p < .01$). くわえて, 感情表現と発話者の間には交互作用が確認された (UI-LI: $F(6, 507) = 3.79, p < .01$, UL-LL: $F(6, 507) = 3.81, p < .01$).

それぞれの個人差がどのように円唇性に影響を与えるのかを解析するために, UL-UI 及び, LL-LI の距離平均と感情表現との関係を可視化した (図 3.17, 図 3.18). なお, 比較のため, 上記二つの図においては Neutral 表現時の距離平均値を図にくわえている.

結果を確認すると, 図 3.17, 図 3.18 ともに全ての感情表現, どの程度表現においても Neutral 表現時との相対的な位置関係に変化は見られなかった. ただし, Anger-Joy, Relax-Sad 表現間においては, Little 程度の表現を行う場合に, 相対的な位置関係の逆転が起きることが確認された.

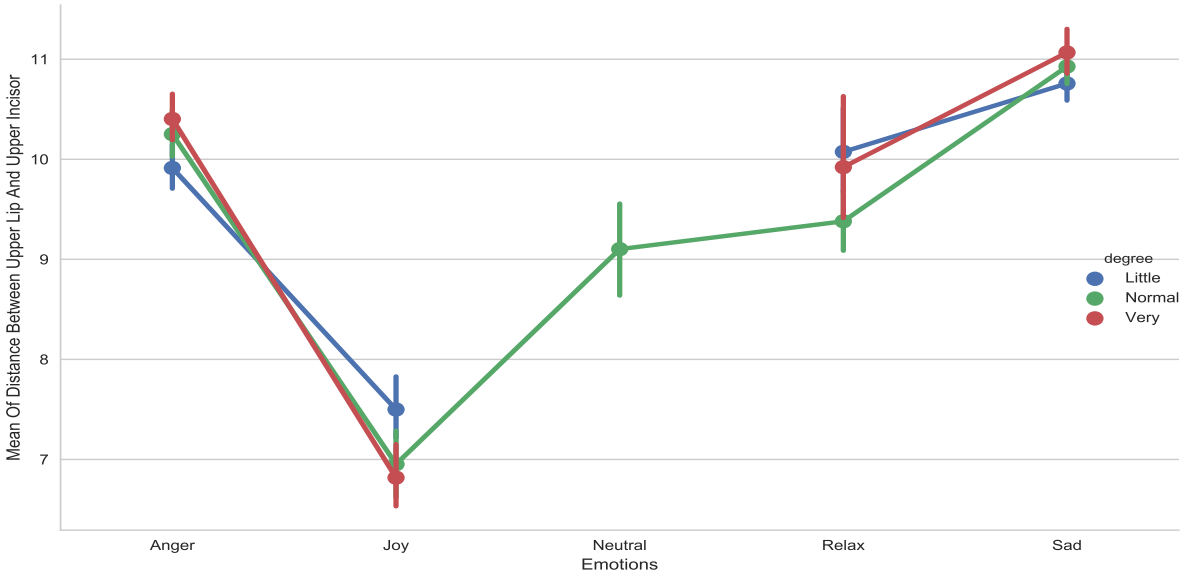


図 3.17: 上唇-上門歯間の平均距離と感情表現及び感情程度との関係. 縦軸は上唇-上門歯間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中色分けはそれぞれの程度表現を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

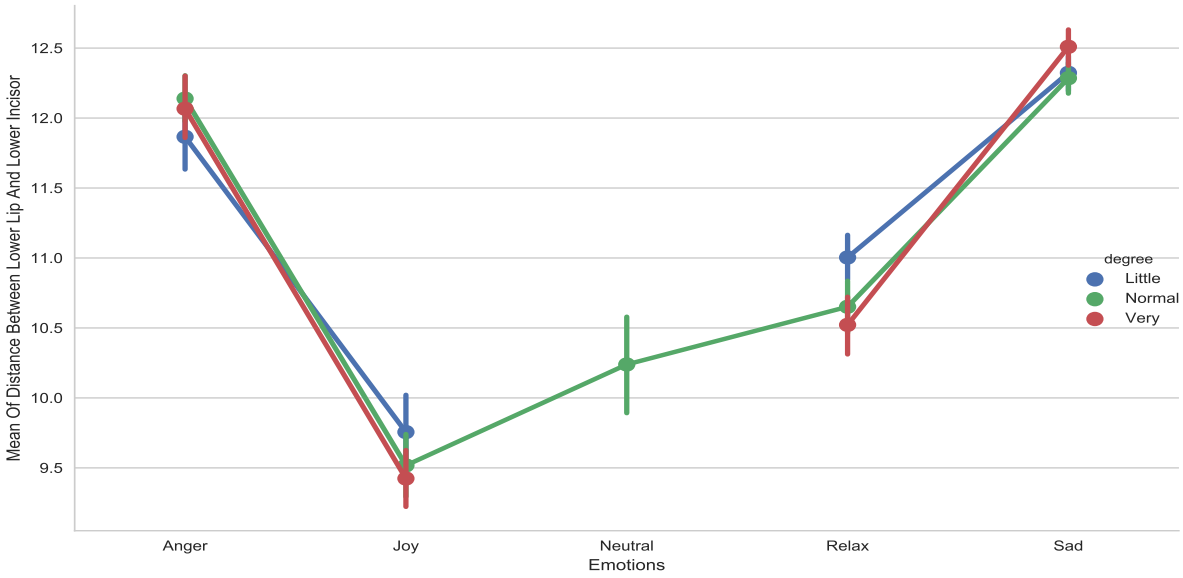


図 3.18: 下唇-下門歯間の平均距離と感情表現及び感情程度との関係. 縦軸は下唇-下門歯間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中色分けはそれぞれの程度表現を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

表 3.12: 舌運動に対する 2 要因 (感情表現, 感情程度) 分散分析の結果. 独立変数は音声発話時に指定された感情表現, および 発話者. 従属変数は表の通り. ただし, 表中 TT は 舌先, TB は舌端, TD は 舌背 を示し, UI, LI はそれぞれ上下門歯を示す.

Dependent Variable	Independent Variable	$d.f.$	F	p
TT-UI(mean)	emotion	3, 507	3.37	$p = .01$
	degree	2, 507	0.03	$p = .96$
	emotion:degree	6, 507	0.12	$p = .99$
TT-LI(mean)	emotion	3, 507	3.20	$p = .02$
	degree	2, 507	0.49	$p = .61$
	emotion:degree	6, 507	0.25	$p = .95$
TB-UI(mean)	emotion	3, 507	5.74	$p < .01$
	degree	2, 507	0.06	$p = .93$
	emotion:degree	6, 507	0.43	$p = .85$
TB-LI(mean)	emotion	3, 507	3.45	$p = .01$
	degree	2, 507	0.19	$p = .81$
	emotion:degree	6, 507	0.15	$p = .98$
TD-UI(mean)	emotion	3, 507	6.08	$p < .01$
	degree	2, 507	0.08	$p = .91$
	emotion:degree	6, 507	1.33	$p = .23$
TD-LI(mean)	emotion	3, 507	3.37	$p = .01$
	degree	2, 507	0.10	$p = .89$
	emotion:degree	6, 507	0.19	$p = .97$

3.5.3 舌運動

舌運動に関して, 二要因分散分析の結果を表 3.12 に示す. 分散分析の結果, 舌端上門歯 (TB-UI), 舌背上門歯 (TD-UI) 間の距離平均値において, 感情表現の主効果が確認された (TB-UI: $F(3, 507) = 3.37, p < .01$, TD-UI: $F(3, 507) = 5.74, p < .01$) 一方で, これらの観測点に対し, 程度表現による主効果は確認されなかった.

それぞれの程度表現の差がどのように舌運動に影響を与えるのかを解析するために, TB-UI 及び, TD-UI の距離平均と感情表現との関係を可視化した (図 3.19, 図 3.20). 結果を確認すると, Anger-Joy 間において程度差の影響は確認されなかった, 一方, Relax, Sad 間においては, Little, Very 程度で表現をしている場合には, Sad 表現寺の方が舌が低い位置に移動しているのに対し, Normal 表現時には関係が逆転することが観察された.

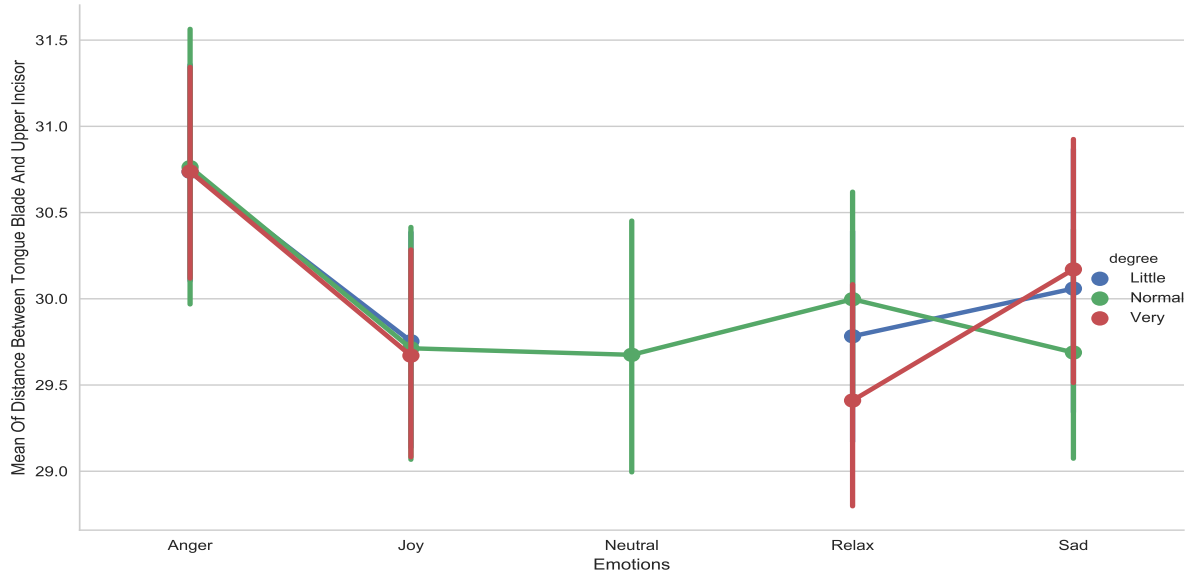


図 3.19: 舌背-下門歯間の平均距離と感情表現及び感情程度との関係. 縦軸は舌背-下門歯間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中色分けはそれぞれの程度表現を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

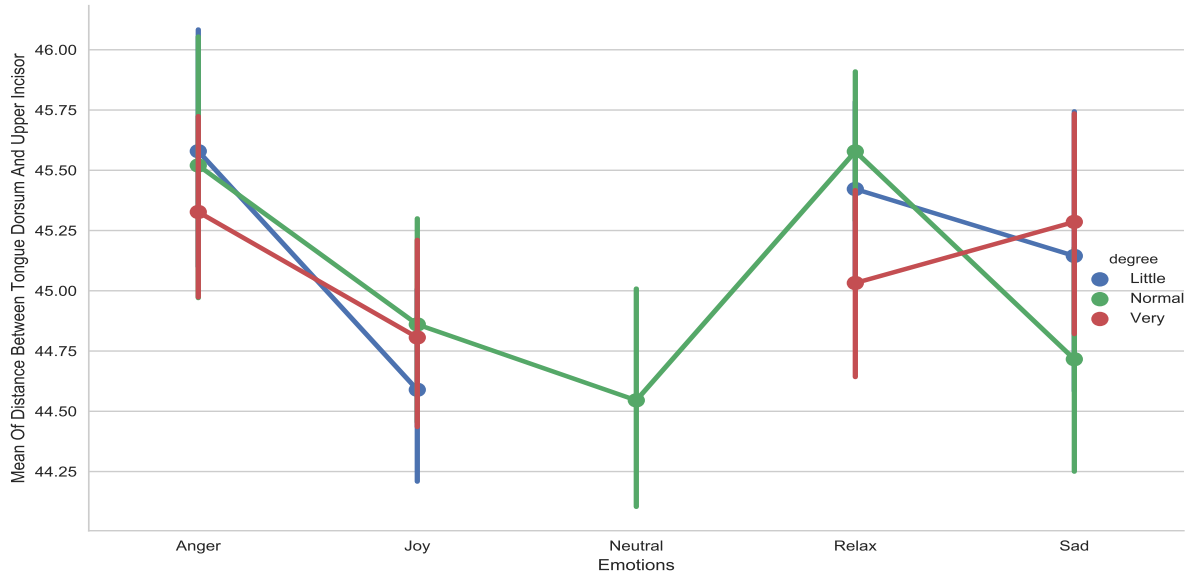


図 3.20: 舌背-上門歯間の平均距離と感情表現及び感情程度との関係. 縦軸は舌背-上門歯間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中色分けはそれぞれの程度表現を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

3.5.4 考察

本節の目的は感情表現を行う際に, その感情表現の程度表現がどのように調音運動に反映されているのかを明らかにすることであった.

開口度に関しては, 特に Anger 表現を行う際に, 程度表現による影響を大きく受けることが示された. 一方で, 円唇性および舌運動に関しては, 感情の程度表現を明確に反映した変化を観察することはできなかった. この結果は, 感情表現の内, 程度表現に関しては, 例えば, 音高やパワーの制御, あるいは持続時間の制御や声質の制御など, 調音運動を変化させる以外の方法を使用して行う可能性を示唆する.

3.6 母音の影響

本節では, 感情表現と言語情報との関連を調音運動から整理するために JEESD に収録されている音声の内, /これは感情音声ですね/ という文章, および, /新しい手紙が届いています/ という文章に含まれる母音 /a, e, o/ を発話している際の調音データの解析を行った. 上記目的に従い, 5 種類の感情を通常の程度で表現しているデータのみを選択した.

母音の音響的特性を決定するためには舌の位置が重要な役割を果す. 一方で, 3.4 節の解析により, 舌運動には個人差が存在することが確認されたため, ここでは特に舌運動を感情表現に使用していると思われる発話者 1 名のみを解析の対象とした. なお解析に使用したデータは, /a/ が 82 個, /e/ が 27 個, /o/ が 30 個である.

感情表現と各観測点間の距離との関係を統計的に解釈するために, 5 種類の感情表現と 3 種類の母音を独立変数とした 2 要因分散分析を行った. 従属変数には, 母音発話時の調音位置を示す指標として調音点間の距離の平均値を使用した. ここで, 解析に使用した文章内に含まれる母音数はそれぞれ異なる. そのため, 分散分析の解析にはタイプ II 平方和を利用した. なお, 統計的検定の有意水準はいずれも .01 と定めた.

3.6.1 開口度

開口度に関して, 二要因分散分析の結果を表 3.13 に示す. 分散分析の結果, 上下門歯 (UI-LI) 間, 及び, 上下唇 (UL-LL) 間の距離平均値において感情表現の主効果が確認された (UI-LI: $F(4, 81) = 7.84, p < .01$, UL-LL: $F(4, 81) = 25.60, p < .01$). その内, UL-LL に関しては母音の種類に対しても主効果が確認された ($F(2, 81) = 23.61, p < .01$). くわえて, 感情表現と母音の種類との間に交互作用が確認された ($F(8, 81) = 3.58, p < .01$).

それぞれの母音の影響がどのように開口度にどのように現れるのかを観察するため, UL-LL の距離平均と感情表現との関係を可視化した (図 3.21). 結果を確認すると, 特に母音 /o/ を発話している際には感情表現を行っている際の上下唇間の距離は Neutral 表現時時と比較し, 狭くなることが確認された. 一方で, 母音 /a, e/ 発話時関しては 3.3 節の解析と同様 Relax, Sad 表現時には Neutral 表現時と比較して開口度が小くなる傾向が確認さ

表 3.13: 開口度に対する 2 要因 (感情表現, 母音) 分散分析の結果. 独立変数は音声発話時に指定された感情表現, および, 母音/a, e, o/. 従属変数は表の通り. ただし, 表中 UI, LI はそれぞれ上下門歯を UL, LL はそれぞれ上下唇を示す.

Dependent Variable	Independent Variable	$d.f.$	F	p
UI-LI (mean)	emotion	4, 81	7.84	$p < .01$
	segment	2, 81	1.40	$p = .25$
	emotion:segment	8, 81	0.42	$p = .90$
UL-LL (mean)	emotion	4, 81	25.60	$p < .01$
	segment	2, 81	23.61	$p < .01$
	emotion:segment	8, 81	3.58	$p < .01$

表 3.14: 円唇性に対する 2 要因 (感情表現, 母音) 分散分析の結果. 独立変数は音声発話時に指定された感情表現, および母音 /a, e, o/. 従属変数は表の通り. ただし, 表中 UL, LL はそれぞれ上下唇を, UI, LI は上下門歯を示す.

Dependent Variable	Independent Variable	$d.f.$	F	p
UL-UI (mean)	emotion	4, 81	49.26	$p < .01$
	segment	2, 81	8.45	$p < .01$
	emotion:segment	8, 81	4.80	$p < .01$
LL-LI (mean)	emotion	4, 81	48.66	$p < .01$
	segment	2, 81	19.30	$p < .01$
	emotion:segment	8, 81	1.52	$p < .01$

れた. また Anger 表現を行う際には, 母音 /a/ 発話時と母音/e/ 発話時の開口度が極めて近い値になることが確認された.

3.6.2 円唇性

開口度に関して, 二要因分散分析の結果を表 3.14 に示す. 分散分析の結果, 上唇上門歯 (UL-UI) 間, 及び, 下唇下門歯 (LL-LI) 間の距離平均値において, 感情表現の主効果が確認された (UI-LI: $F(4, 81) = 49.26, p < .01$, UL-LL: $F(4, 81) = 48.66, p < .01$). また, 両観測点ともに音声発話時の母音の種類の違いによる主効果が確認された (UI-LI: $F(2, 81) = 8.45, p < .01$, UL-LL: $F(2, 81) = 19.30, p < .01$). くわえて, 感情表現と発話者の間には交互作用が確認された (UI-LI: $F(8, 81) = 4.80, p < .01$, UL-LL: $F(8, 81) = 1.52, p < .01$).

音声発話時の母音の差がどのように円唇性に影響を与えるのかを解析するために, UL-UI 及び, LL-LI の距離平均と感情表現との関係を可視化した (図 3.22, 図 3.23).

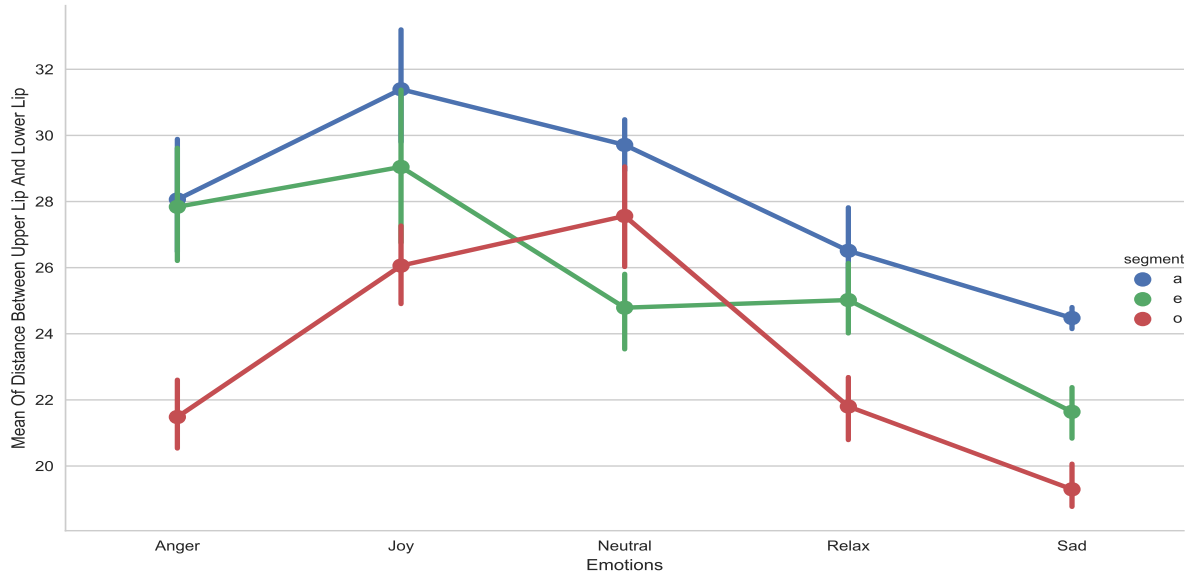


図 3.21: 上唇-下唇間の平均距離と感情表現及び母音との関係. 縦軸は上唇-下唇間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中色分けはそれぞれの音声発話時の母音の種類を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

結果を確認すると, とくに図 3.22 においては, Neutral 表現時には存在していた調音点間の距離の差が, 感情表現時には一定の値に収まることが確認された. 同様の傾向は図 3.23 の Relax 表現時や Anger 表現時にも観察された.

3.6.3 舌運動

舌運動に関して, 二要因分散分析の結果を表 3.15 に示す. 分散分析の結果, 舌先上門歯 (TT-UI), 舌先下門歯 (TT-LI), 舌端上門歯 (TB-UI), 舌端下門歯 (TB-LI), 舌背上門歯 (TD-UI) そして 舌背下門歯 (TD-LI) 間の距離平均値において, 感情表現の違いによる主効果が確認された (TT-UI: $F(4, 81) = 3.55, p < .01$, TT-LI: $F(4, 81) = 3.90, p < .01$, TB-UI: $F(4, 81) = 6.87, p < .01$, TB-LI: $F(4, 81) = 3.65, p < .01$, TD-UI: $F(4, 81) = 8.90, p < .01$, TD-LI: $F(4, 81) = 5.27, p < .01$). また, これら全ての観測点に対し発話時の母音の種類の違いによる主効果が確認された. (TT-UI: $F(2, 81) = 34.21, p < .01$, TT-LI: $F(2, 81) = 86.24, p < .01$, TB-UI: $F(2, 81) = 102.94, p < .01$, TB-LI: $F(2, 81) = 99.37, p < .01$, TD-UI: $F(2, 81) = 101.50, p < .01$, TD-LI: $F(2, 81) = 13.97, p < .01$). くわえて, 感情表現と発話者条件の間には交互作用が確認された (TT-UI: $F(8, 81) = 0.73, p < .01$, TT-LI: $F(8, 81) = 2.29, p < .01$, TB-UI: $F(8, 81) = 1.21, p < .01$, TB-LI: $F(8, 81) = 1.82, p < .01$, TD-UI: $F(8, 81) = 0.81, p < .01$, TD-LI: $F(8, 81) = 0.74, p < .01$).

それぞれの個人差がどのように舌運動に影響を与えるのかを解析するために, TD-UI 及び, TD-LI の距離平均と感情表現との関係を可視化した (図 3.24, 図 3.25). 結果を確認す

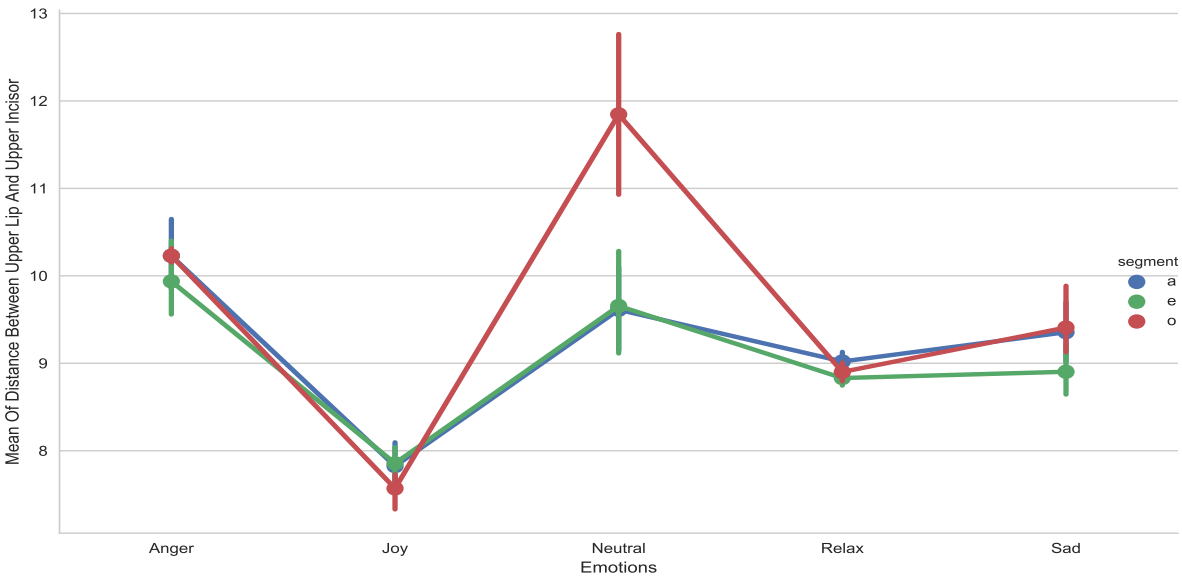


図 3.22: 上唇-上門歯間の平均距離と感情表現及び母音との関係縦軸は上唇-上門歯間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中色分けは音声発話時の母音の種類を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

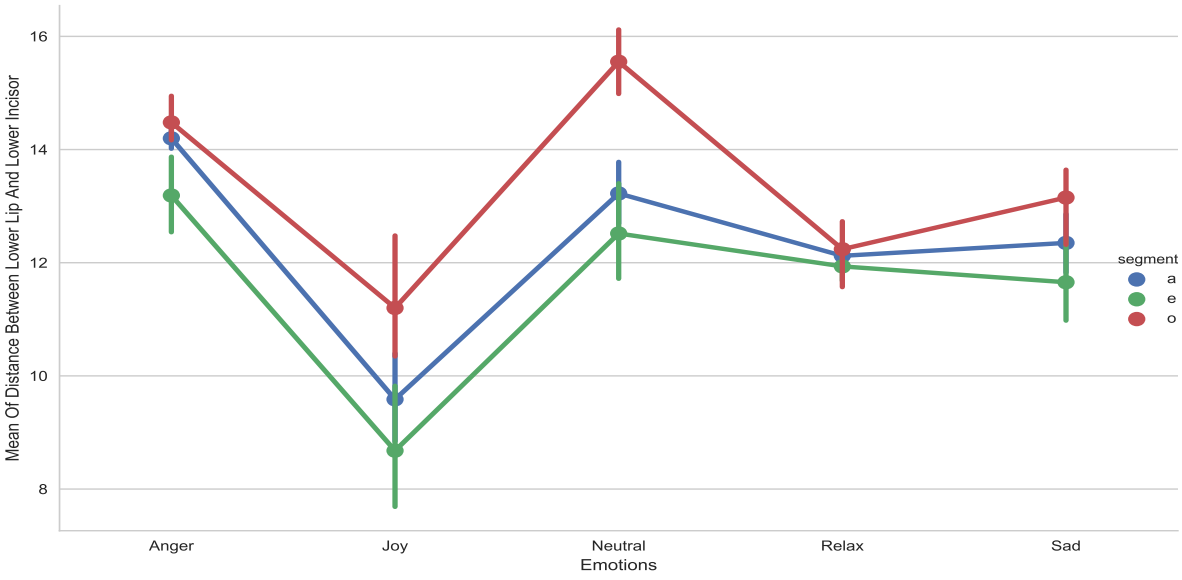


図 3.23: 下唇-下門歯間の平均距離と感情表現及び母音との関係. 縦軸は下唇-下門歯間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中色分けは音声発話時の母音の種類を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

表 3.15: 舌運動に対する 2 要因 (感情表現, 母音) 分散分析の結果. 独立変数は音声発話時に指定された感情表現, および 母音/a, e, o/. 従属変数は表の通り. ただし, 表中 TT は 舌先, TB は舌端, TD は 舌背 を示し, UI, LI はそれぞれ上下門歯を示す.

Dependent Variable	Independent Variable	$d.f.$	F	p
TT-UI (mean)	emotion	4, 81	3.55	$p < .01$
	segment	2, 81	34.21	$p < .01$
	emotion:segment	8, 81	0.73	$p < .01$
TT-LI (mean)	emotion	4, 81	3.90	$p < .01$
	segment	2, 81	86.24	$p < .01$
	emotion:segment	8, 81	2.29	$p < .01$
TB-UI (mean)	emotion	4, 81	6.87	$p < .01$
	segment	2, 81	102.94	$p < .01$
	emotion:segment	8, 81	1.21	$p < .01$
TB-LI (mean)	emotion	4, 81	3.65	$p < .01$
	segment	2, 81	99.37	$p < .01$
	emotion:segment	8, 81	1.82	$p < .01$
TD-UI (mean)	emotion	4, 81	8.90	$p < .01$
	segment	2, 81	101.50	$p < .01$
	emotion:segment	8, 81	0.81	$p < .01$
TD-LI (mean)	emotion	4, 81	5.27	$p < .01$
	segment	2, 81	13.97	$p < .01$
	emotion:segment	8, 81	0.74	$p < .01$

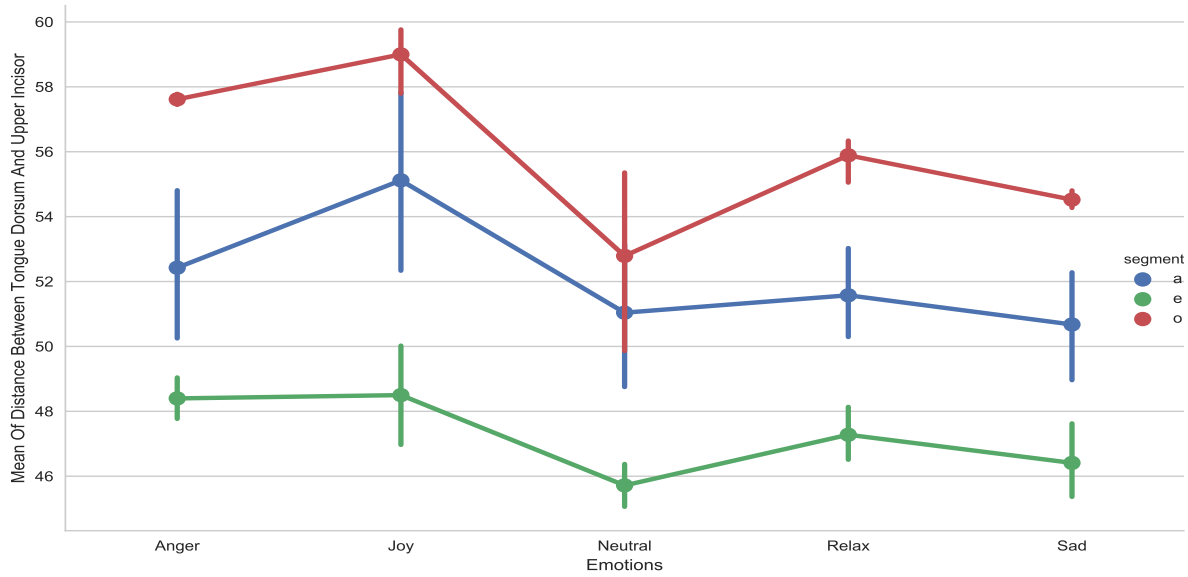


図 3.24: 舌背-上門歯間の平均距離と感情表現及び母音との関係. 縦軸は舌背-上門歯間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中色分けは音声発話時の母音の種類を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

ると図 3.24, 3.25 ともに, 特に母音 /a, o/ を比較すると, 感情表現時の方が, Neutral 表現時と比較して母音間の距離が大きくなっていることが確認された. 一方で感情表現間に対しては母音に対し相対的には同様の変化をする傾向がみられた.

3.6.4 考察

本節の目的は, 感情表現が調音運動に対し与える影響を母音毎に解析することであった. その結果, 円唇性に関しては, 特定の感情発話時に調音点間の距離が一定の範囲に収まることが確認された. この結果は, 円唇性に対しては言語的な情報の生成よりも優先して感情表現を行っている可能性を示唆する結果である. 一方舌運動に関してはすべての母音において感情表現の影響をうけることが確認されたが, 円唇性で確認されたような調音点の収束や逆転は観察されなかった.

上記のような感情表現から受ける影響の受け方の違いに関しては, その調音点が言語的な情報（本節の解析においては母音の音響的な特徴）の生成にどの程度重要であるのかが関わってくる可能性がある. 例えば, 円唇性は母音 /a, e, o/ の対比に対しては然程重要な要素ではない. そのため, ある感情を分かり易く表現するために, 特定の値まで調音運動を収束させることが可能である. 一方で舌の位置は生成される音声のフォルマントの値を決める重要な要素であることが先行研究により指摘されている. そのため, 舌運動に関しては円唇性と同じ方略で感情表現を行うことはできず, 母音間の距離を相対的に担保しつつ感情表現をおこなっているものだと考察できる.

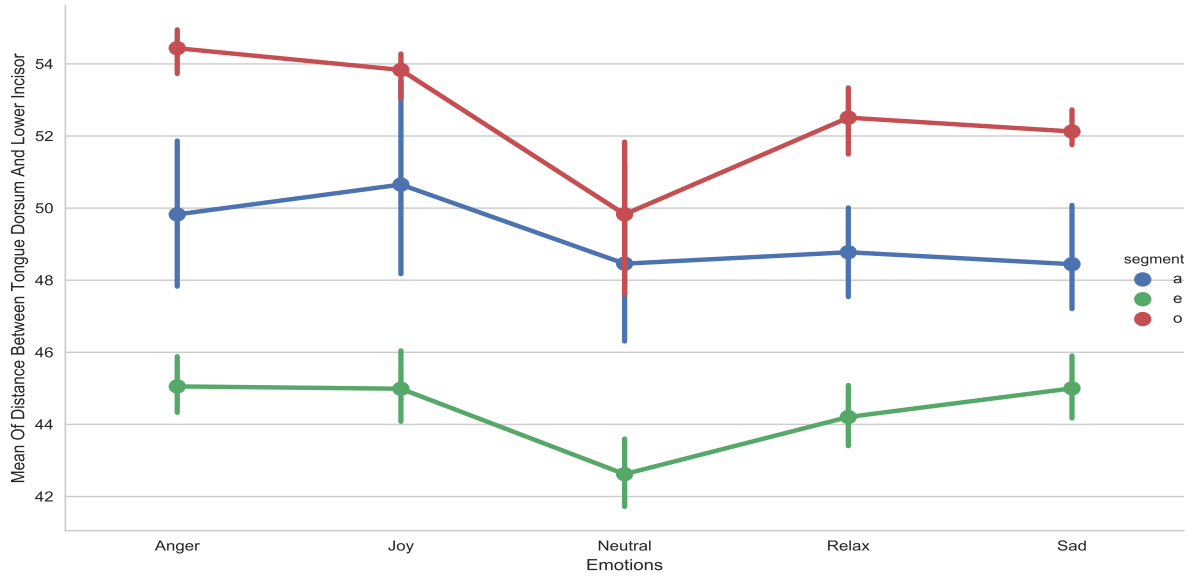


図 3.25: 舌背-下門歯間の平均距離と感情表現及び母音との関係. 縦軸は舌背-下門歯間距離の母音内平均値, 横軸はそれぞれの感情表現を示す. 図中色分けは音声発話時の母音の種類を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

3.7 本章のまとめ

本章では, 活性度および快不快の二軸の空間上に布置される感情表現を行う際の調音運動の変化を開口度, 円唇性, 舌運動の観点から調査した.

3.3 節の解析では, 調音運動が感情表現をどのように反映するのかを調査するために, 母音/a/ を対象に, 5 種類の感情音声発話時の調音運動を比較した. その結果, 表現対象となる感情の活性度は, 開口度や舌の縦方向の位置に影響し, 快不快は円唇性や舌の前後方向の位置に影響することを明らかにした.

3.4 節の解析では, 感情表現の個人差を調音運動の観点から調査した. その結果, 特に舌運動に関しては, 個人差の影響が大きいことを明らかにした. また, 活性度が高い場合の開口度や, 活性度が低い場合の円唇性なども個人差の影響を受けることを発見した. この結果は, 表現対象となる感情の活性度により, 調音運動上の個人差の生じやすさが変化することを示唆する結果である. 一方で, 3.3 節で発見した感情表現が調音運動に与える影響自体は両名の発話者で同じ傾向を示すことも分かった. この結果は, 感情表現には個人による表現の差を越えた一般的な傾向があることを示唆する結果である.

3.5 節の解析では, 感情表現の程度が調音運動にどのような影響を与えるのかを調査した. その結果, 特に Anger を表現している際の開口度は程度の影響を大きく受けることを発見した. 一方で, 円唇性や舌運動などは感情の程度をあまり大きく反映しないことも分かった.

3.6 節の解析では, 複数の母音発話時における調音運動を感情表現ごとに比較した. そ

の結果, 開口度, 円唇性に関しては, 感情表現を行うために, その母音を生成する際の一般的な調音から大きく逸脱することを確認した. 一方で舌の位置に関しては, 母音に対し相対的に感情表現の影響を受け変化することは確認されたが, 母音間の絶対的な位置関係が崩れることはなかった. この結果は, 演技感情音声発話時においては, パラ言語情報の生成と言語情報の生成という二つの発話目的を同時に達成するために, 言語情報の生成に使用する調音運動箇所と感情表現に使用する調音運動箇所とを使い分けていることを示唆する結果である.

第4章　自発感情音声発話時の調音運動

4.1　目的

前章では, 演技感情音声発話時の感情表現と言語的情報との関係を調査した. しかし 1.2 節で指摘した通り, 言語情報の伝達と感情表現の伝達を同時に発話目的とする演技感情音声発話と感情表現の伝達は発話目的にふくまれていないにも拘わらず音声に感情が滲み出てしまっている自発感情音声発話とでは, 同じ感情を表現するとしても異なる方略で調音運動を行っている可能性がある.

そのため本章では, 音声に感情表現がふくまれていない朗読音声発話時と音声に感情が滲み出てしまっている自発感情音声発話時の調音運動を比較することで, 自発的感情表現が音声発話時の調音運動にどのような影響を与えるのかを検討する. 自発感情音声発話時の調音運動の解析には Ericson らの研究があるが, この研究では, 悲しみ音声発話時の調音運動を感情表現にのみ注目して解析しているに留まる [3]. そのため, 本章では, Ericson らの研究で使用されたデータを, 特に言語情報の生成との関連に注目して解析をおこなう. ここで解析の対象としたデータの説明に関しては 4.2 節を参照してほしい.

本章の解析では, まず, 朗読音声発話時と自発感情発話時の調音運動を比較することで, 自発感情発話時にも感情表現が調音運動に影響を与えることを確認する (4.3 節). この解析で得られた知見を元に, 3.4 節では, 母音 /a, e, o/ 発話時の調音運動を比較することで, 自発感情音声発話時に感情表現が言語情報の生成にどのような影響を与えるのかを調査する.

4.2　解析対象

本節では Erickson らの研究 [3] で使用されたデータに関しての概要を説明する. このデータは アメリカ人女性（中西部方言）と日本女性（広島方言）各 1 名の悲しみ音声発話時の調音データが 2D EMA システムを使い収録されている. EMA 収録はフレーム間の約 3s を ブレイクとする 20 s のウインドウ幅で行い, 対話は自然な範囲に留まっていた. この際に発話の音響データも同時に収録を行っている. ただし, 日本人発話データに関しては, 舌背位置の計測に失敗している.

このデータは, 電話での対話が想定し収録されたものである. そのため, 発話者は特定の発話相手に対し, イヤフォンとマイクロフォンを通じて発話を行っている. ここで日本

人発話収録時においては, 発話者は母親を脳動脈瘤のため最近失ったばかりであった. 対話相手は音声収録中にこの状況に気が付き, この話題を中心に音声対話が行われた.

この自発感情音声の収録時に収録された発話の内, /と言ってなったらじゃーもう離婚しなきゃだめかなーと思って/, /何やってもちゃんとできないから/, /救いはだから義理の母が凄く分かってくれるから/ という 3 種類の文章に関しては, 上記条件で収録した自発感情音声発話 (spontaneous) の他に, 録音された自発感情音声を, ヘッドフォンを通して聞いたり, 台本を確認しながら, 単語, フレーズ, イントネーション, 感情を模倣した音声発話 (感情模倣音声発話, imitating) 時と, 感情表現以外を模倣した音声 (読み上げ音声発話: reading) を収録している. この内, 感情模倣音声発話は 3 章における Sad 表現を行っている場合の音声発話に, 読み上げ音声発話は 3 章における Neutral 表現を行っている場合の音声発話に, 相当する発話である.

本研究では上記データの内日本語発話者の音声発話時の調音運動のみを解析の対象とした. Erickson らの研究 [3] で使用されたデータには言語的な情報に対する転記情報は存在しない. そのため, 本研究では収録された全ての発話セットに対し,『日本語話し言葉コーパス』に基づき, 分節音ラベル [23] を転記した. この際に, 特に音素境界が曖昧なものに関しては解析から除外している.

4.3 感情表現の影響

本節では, 自発感情表現が音声発話時の調音運動に与える一般的な影響を整理するために Erickson らの研究で使用されているデータ [3] の内, 母音 /a/ を発話している際の調音データの解析を行った. ここでは, 自発感情音声発話時に感情表現が調音運動に対しどのような影響を与えるのかを検討するために, 自発感情音声発話時 (spontaneous) と読み上げ音声発話時 (reading) の調音運動を比較した. なお解析に使用した母音数は 50 個であった.

感情表現と調音点間の距離との関係を統計的に解釈するために, 調音点間の距離の平均値を T 検定を使い比較した. なお, 統計的検定の有意水準はいずれも .01 と定めた.

4.3.1 開口度

T 検定の結果, 感情音声発話時と読み上げ音声発話時の上下唇 (UL-LL) 間の距離平均値において, 有意差を確認した ($t(21) = 4.66, p < .01$).

UL-LL の距離平均と感情表現との関係を図 4.1 に示す. 図 4.1 を確認すると, 自発感情音声発話時には, 読み上げ音声発話時と比較して口の開きが小くなることが確認された.

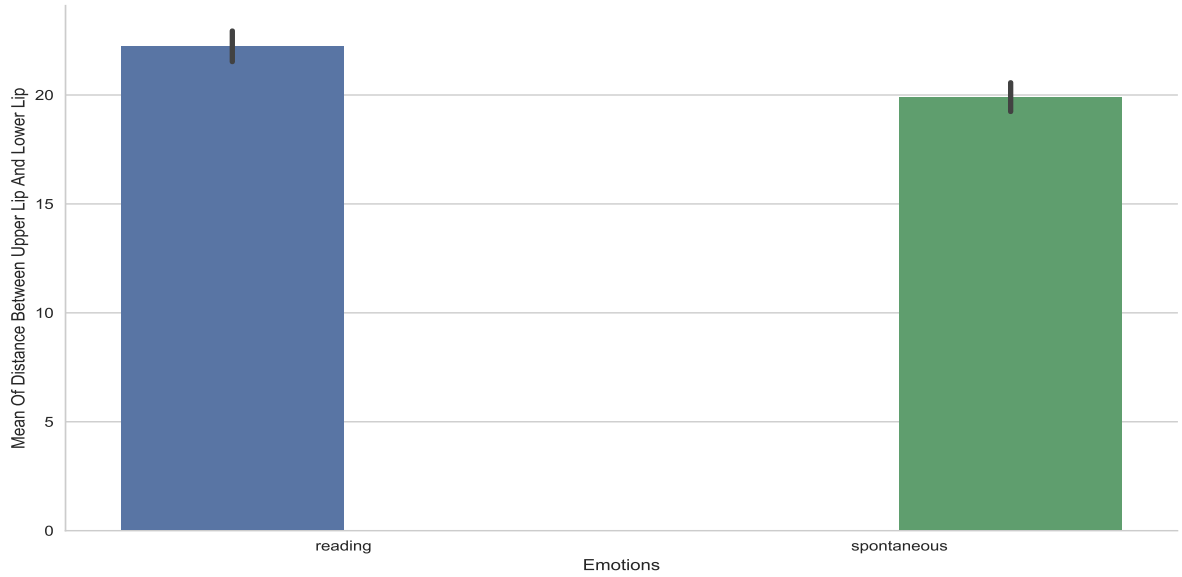


図 4.1: 上唇-下唇間の平均距離と感情表現との関係.

4.3.2 円唇性

T 検定の結果, 感情音声発話時と読み上げ音声発話時の上唇上門歯 (UL-UI) 間, 及び, 下唇下門歯 (LL-LI) 間の距離平均値において, 有意差が確認された (UL-UI: $t(21) = 23.27, p < .01$, LL-LI: $t(21) = 16.28, p < .01$).

UL-UI 及び, LL-LI の距離平均と感情表現との関係をそれぞれ, 図 4.2, 4.3 に示す. 図 4.2, 4.3 を確認すると, 自発感情音声発話時には, 読み上げ音声発話時と比較して唇の突き出しが弱くなることが確認された.

4.3.3 舌運動

T 検定の結果, 自発感情音声発話時と読み上げ音声発話時の舌端上門歯 (TB-UI) 間の距離平均値においてのみ, 有意差が確認された ($t(21) = -0.35, p < 0.73$).

TB-UI の距離平均と感情表現との関係を図 4.4 に示す. 図 4.4 を確認すると, 感情音声発話時と読み上げ音声発話時と比較して舌の位置が下降することが確認された.

4.3.4 考察

本節の目的は自発感情音声発話時に感情表現が調音運動に与える影響を確認することにあった.

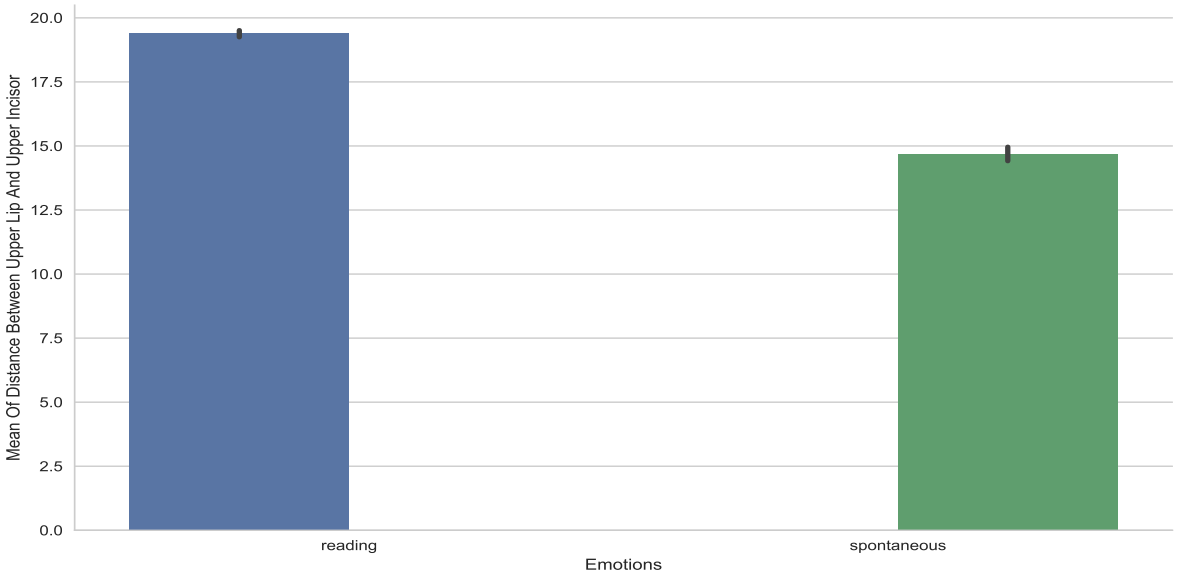


図 4.2: 上門歯-上唇間の平均距離と感情表現との関係.

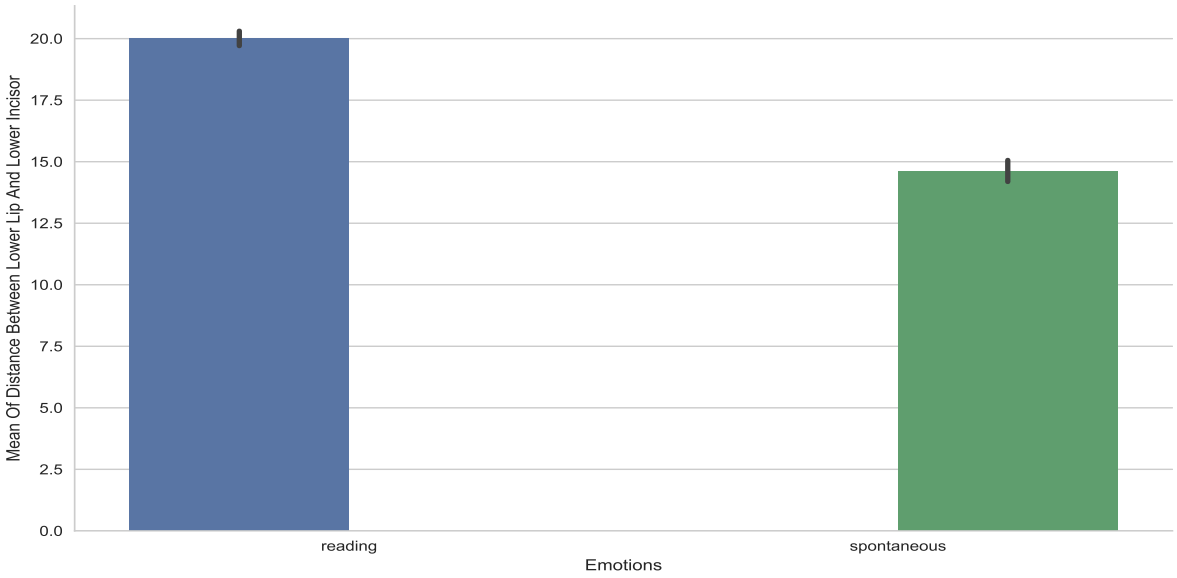


図 4.3: 下門歯-下唇間の平均距離と感情表現との関係.

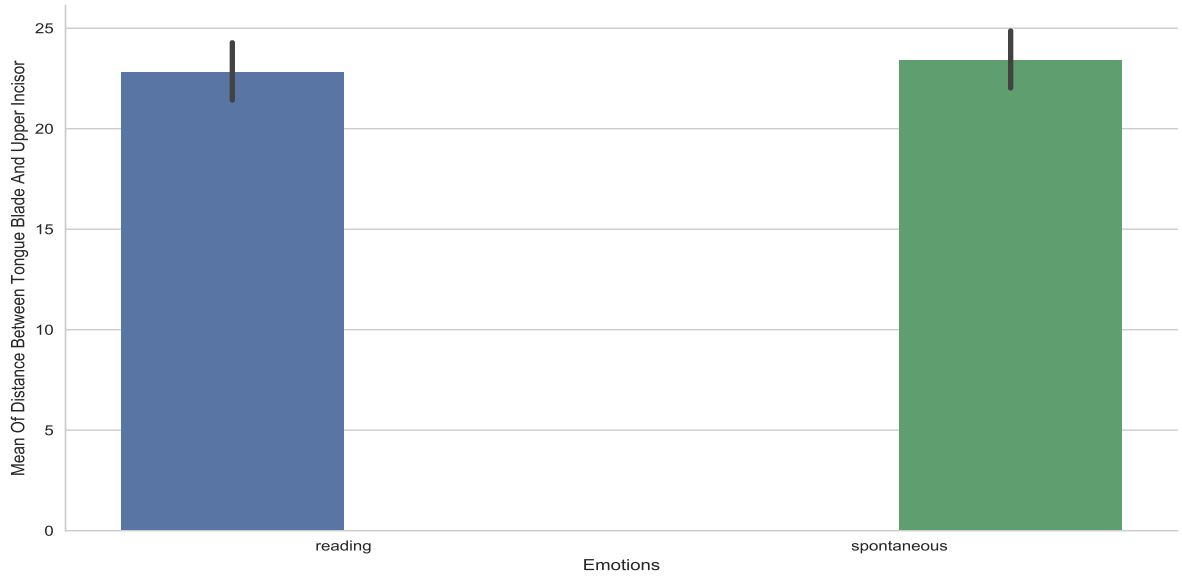


図 4.4: 舌端-上門歯間の平均距離と感情表現との関係.

開口度に関する解析では, 自発感情音声発話時は読み上げ音声発話時と比較して, 口の開きが小くなることが確認された. これは, 自発感情音声発話時に, 開口度を操作することで感情表現を行っていることを示唆する結果である.

円唇性に関する解析では, 自発音声発話時には唇の突き出しが弱くなることが確認された. これは, 自発感情音声発話時に, 円唇性を操作することで感情表現を行っていることを示唆する結果である.

舌運動に関する解析では自発音声発話時には舌端の前後方向の位置が後ろに移動することが確認された. これは, 自発感情音声発話時に, 舌の前後方向の位置を操作することで感情表現を行っていることを示唆する結果である.

4.4 母音の影響

本節では, 自発感情音声発話において感情表現が言語情報の生成に与える影響を整理するために Erickson らの研究で使用されているデータ [3] の内, 母音 /a, e, o/ を発話している際の調音データの解析を行った. ここでは, 自発感情音声発話時の特徴を観察するために, 自発悲しみ音声発話時と読み上げ音声発話時の調音運動を比較した. なお解析に使用したデータは, /a/ が 42 個, /e/ が 15 個, /o/ が 13 個である.

感情表現と各観測点間の距離との関係を統計的に解釈するために, 2 種類の発話条件と 3 種類の母音を独立変数とした 2 要因分散分析を行った. 従属変数には, 母音発話時の調音位置を示す指標として調音点間の距離の平均値を使用した. ここで, 解析に使用した文章内に含まれる母音数はそれぞれ異なる. そのため, 分散分析の解析にはタイプ II 平方和を利用した. なお, 統計的検定の有意水準はいずれも .01 と定めた.

表 4.1: 開口度に対する 2 要因 (感情表現, 母音) 分散分析の結果. 独立変数は音声発話時に指定された感情表現, および, 母音/a, e, o/. 従属変数は表の通り. ただし, 表中 UI, LI はそれぞれ上下門歯を, UL, LL はそれぞれ上下唇を示す.

Dependent Variable	Independent Variable	$d.f.$	F	p
UI-LI (mean)	emotion	1, 69	6.07	$p = .01$
	segment	2, 69	6.27	$p < .01$
	emotion:segment	2, 69	0.04	$p = .95$
UL-LL (mean)	emotion	1, 69	25.47	$p < .01$
	segment	2, 69	53.43	$p < .01$
	emotion:segment	2, 69	0.46	$p < .01$

4.4.1 開口度

開口度に関して, 二要因分散分析の結果を表 4.1 に示す. 分散分析の結果, 上下唇 (UL-LL) 間の距離平均値に対する感情表現の主効果が確認された ($F(1, 69) = 25.47, p < .01$). また, 上下門歯間の距離 (UI-LI) および, 上下唇間の距離の平均値に対する発話母音の種類による主効果が確認された (UI-LI: $F(2, 69) = 6.27, p < .01$, UL-LL: $F(2, 69) = 53.43, p < .01$). くわえて, 上下唇間の距離に対して, 感情表現と発話母音の種類間に交互作用が確認された ($F(2, 461) = 4.84, p < .01$).

それぞれの発話母音の差がどのように開口度に影響を与えるのかを解析するために, UI-UI および UL-LL の距離平均と感情表現との関係を可視化した (図 4.5, 図 4.6). 結果を確認すると, いずれの母音発話時においても自発感情音声発話時には読み上げ音声発話時より開口度が低くなることが確認された.

4.4.2 円唇性

円唇性に関して, 二要因分散分析の結果を表 4.2 に示す. 分散分析の結果, 上唇上門歯 (UL-UI) 間, 及び, 下唇下門歯 (LL-LI) 間の距離平均値において, 感情表現の主効果が確認された (UI-LI: $F(1, 69) = 1006.05, p < .01$, UL-LL: $F(1, 69) = 483.56, p < .01$). また, 両観測点ともに発話時の母音の違いによる主効果が確認された (UI-LI: $F(2, 69) = 78.57, p < .01$, UL-LL: $F(1, 69) = 74.22, p < .01$). くわえて, 感情表現と発話者の間には交互作用が確認された (UI-LI: $F(2, 69) = 6.12, p < .01$, UL-LL: $F(2, 69) = 0.03, p < .01$).

それぞれの個人差がどのように円唇性に影響を与えるのかを解析するために, UL-UI 及び, LL-LI の距離平均と感情表現との関係を可視化した (図 4.7, 図 4.8). 結果を確認すると, いずれの母音発話時においても感情音声発話時には読み上げ音声発話時より唇の突き出しが弱くなることが確認された.

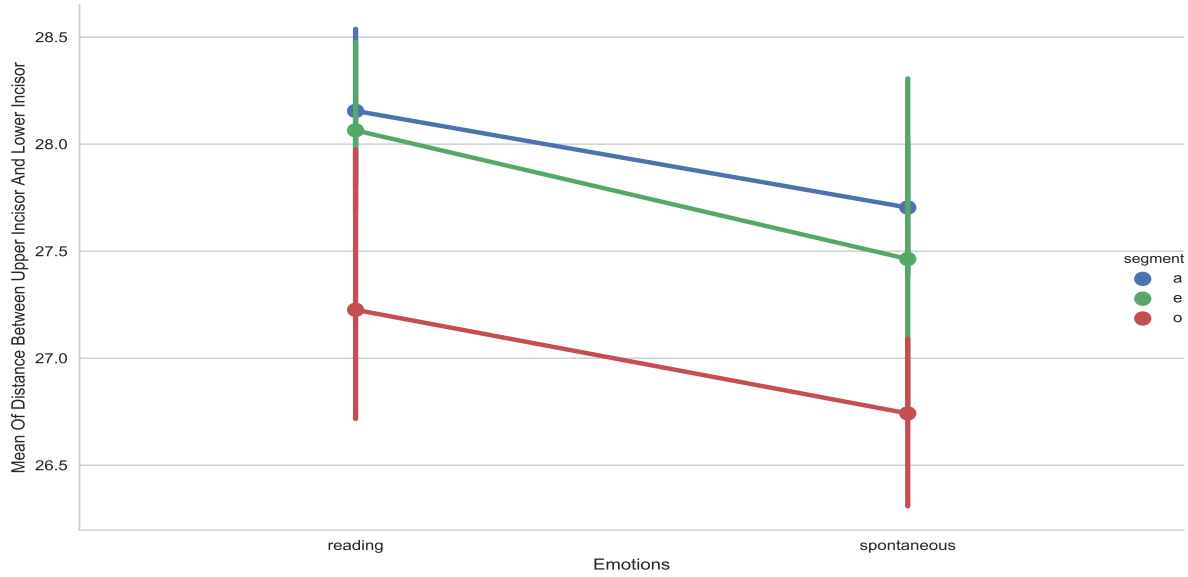


図 4.5: 上門歯-下門歯間の平均距離と感情表現及び母音との関係．縦軸は上唇-下唇間距離の母音内平均値, 横軸はそれぞれの発話条件を示す．図中色分けはそれぞれの音声発話時の母音の種類を表す．図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す．

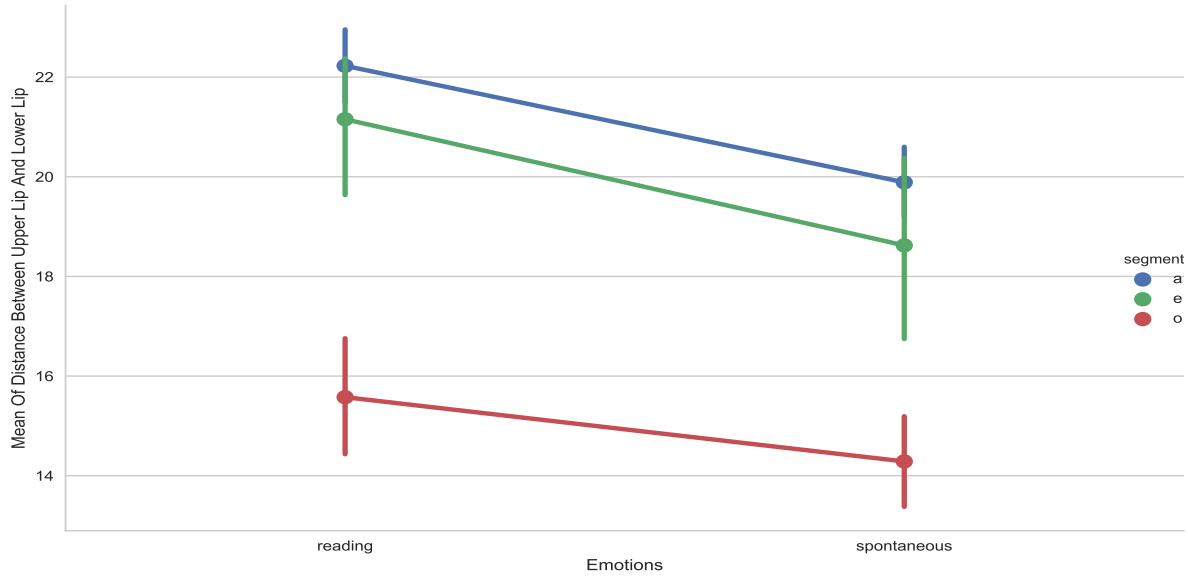


図 4.6: 上唇-下唇間の平均距離と感情表現及び母音との関係．縦軸は上唇-下唇間距離の母音内平均値, 横軸はそれぞれの発話条件を示す．図中色分けはそれぞれの音声発話時の母音の種類を表す．図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す．

表 4.2: 円唇性に対する 2 要因 (感情表現, 母音) 分散分析の結果. 独立変数は音声発話時に指定された感情表現, および母音 /a, e, o/. 従属変数は表の通り. ただし, 表中 UI, LI は上下門歯を, UL, LL はそれぞれ上下唇を示す.

Dependent Variable	Independent Variable	$d.f.$	F	p
UL-UI (mean)	emotion	1, 69	1006.05	$p < .01$
	segment	2, 69	78.57	$p < .01$
	emotion:segment	2, 69	6.12	$p < .01$
LL-LI (mean)	emotion	1, 69	483.56	$p < .01$
	segment	2, 69	74.22	$p < .01$
	emotion:segment	2, 69	0.03	$p < .01$

4.4.3 舌運動

舌運動に関して, 二要因分散分析の結果を表 4.3 に示す. 分散分析の結果, 舌端下門歯 (TB-LI) 間の距離平均値においてのみ, 感情表現の違いによる主効果が確認された ($f(1, 69) = 21.52, p < .01$). 一方で舌端上門歯 (TB-UI) 間の距離へ平均値に対してのみ発話時の母音の種類による主効果が確認された ($f(2, 69) = 7.95, p < .01$). くわえて, 感情表現と発話時の母音の種類の違いの間には交互作用が確認できる観測点は存在しなかった.

それぞれの個人差がどのように舌運動に影響を与えるのかを解析するために, TD-UI 及び, TD-LI の距離平均と感情表現との関係を可視化した (図 4.9, 図 4.10). 図 4.9 を確認すると, 感情音声発話時は読み上げ音声発話時と比べ, 舌の位置は低くなる傾向があることが確認できる. また, この傾向は特に母音 /e/ を発話している際に強くなることが確認された. 一方, 図 4.10 を確認すると, 感情音声発話時は読み上げ音声発話時と比べ, 舌の位置は後ろに下ることが確認された. また, この傾向は特に母音 /a, e/ を発話している際に強くなることが確認された.

4.4.4 考察

本節の目的は, 自発音声発話時の感情表現が調音運動に対し与える影響を母音毎に解析することであった. 解析の結果, 開口度, 円唇性に関しては, すべての母音に対し同様の影響を与えることが示された. 一方で, 感情表現が舌運動に与える影響は特定の母音発話時に変化することが確認された. 舌の高さに関しては, 母音 /e/ を発話している場合に特に感情表現の影響をうけることが確認された. また, 舌の前後方向の位置に関しては, 母音 /o/ を発話している場合に, 感情表現の影響を受けにくく, 読み上げ音声発話時における母音に対する調音運動上の関係を崩すことが確認された.

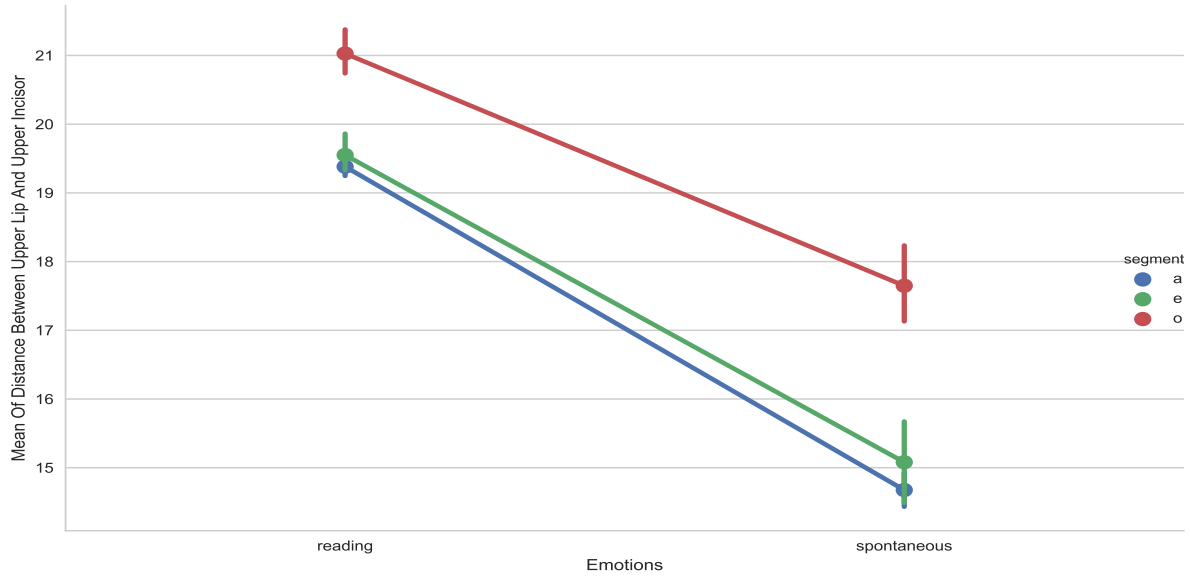


図 4.7: 上唇-上門歯間の平均距離と感情表現及び母音との関係. 縦軸は上唇-上門歯間距離の母音内平均値, 横軸はそれぞれの発話条件を示す. 図中色分けは音声発話時の母音の種類を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

上記のような調音運動に対する感情表現の影響の違いは生成される音声の言語的な性質に影響を与える可能性がある. 本節の解析では, 自発感情音声発話時の母音 /a, e/ と /o/ 間における舌の前後方向の位置関係と読み上げ音声発話時の位置関係とは異なるものであった. 2.2 節で説明したとおり, 母音 /e/ と /o/ の言語的特徴は舌の前後の位置によって決まる. このため, 生成される音声は, 母音の種類を同定しずらい音声になるものと考えられる.

4.5 本章のまとめ

本章では, 自発感情音声発話時に, 感情表現がどのように言語的情報の生成に影響を与えるのかを調査した.

4.3 節の解析では, 調音運動が感情表現をどのように反映するのかを調査するために, 母音 /a/ を対象に, 悲しみ感情を表出している自発音声発話時 (spontaneous) と, 同一の文章を感情を含めず読み上げた音声発話 (読み上げ音声発話, reading) 時の調音運動を比較した. その結果, 自発音声発話時には, 開口度は小くなること, 円唇性が弱くなること, 舌が後ろに移動することが確認された.

この内, 開口度に関する結果は 3 章で得られた演技音声発話時の悲しみ表現の特徴と一致している. 一方で, 円唇性に関しては, 3 章で得られた演技音声発話時の悲しみ表現の特徴とは一致しない. このことは感情表現の伝達を発話目的にふくまない自発感情音声発話時には, 感情表現の伝達を発話目的に含む演技感情音声発話時とは異なる方法で感情を表

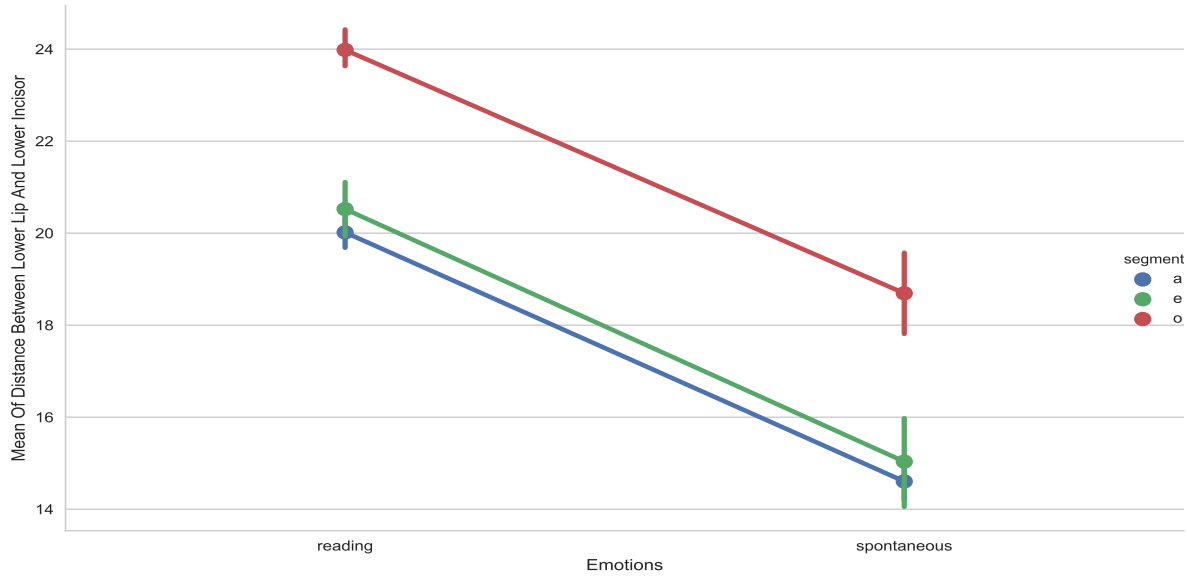


図 4.8: 下唇-下門歯間の平均距離と感情表現及び母音との関係. 縦軸は下唇-下門歯間距離の母音内平均値, 横軸はそれぞれの読み上げ条件を示す. 図中色分けは音声発話時の母音の種類を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

現していることを示唆する. 感情表現の伝達を発話目的にふくまない自発感情音声発話時と感情表現の伝達を発話目的にふくむ演技感情音声発話時の調音運動上の違いに関しては 5 章で検討する.

4.4 節の解析では, 複数の母音発話時における調音運動を自発感情音声発話時と読み上げ音声発話時間で比較した. その結果 , 開口度, 円唇性に関しては, 感情表現はどの母音に対しても同様の影響を与えることが示された. 一方で舌運動に関しては, 母音の種類により, 感情表現が調音運動に対する影響が変化することが確認された. 特に自発感情音声発話時の母音 /a, e/ と /o/ 間における舌の前後方向の位置関係と読み上げ音声発話時の位置関係とは異なるものであった. このため, 自発感情音声発話時に生成される音声は, 母音の種類を同定しずらい音声になると予想された. これらの結果から, 自発感情音声発話時には, 言語情報の生成に重要な調音運動も感情表現に対する影響を大きく受けることが示唆された.

表 4.3: 舌運動に対する 2 要因 (感情表現, 母音) 分散分析の結果. 独立変数は音声発話時に指定された感情表現, および 母音/a, e, o/. 従属変数は表の通り. ただし, 表中 TT は 舌先, TB は舌端, TD は 舌背 を示し, UI, LI はそれぞれ上下門歯を示す.

Dependent Variable	Independent Variable	$d.f.$	F	p
TT-UI (mean)	emotion	1, 69	1.98	$p = .16$
	segment	2, 69	5.21	$p < .01$
	emotion:segment	2, 69	0.15	$p = .85$
TT-LI (mean)	emotion	1, 69	0.02	$p = .86$
	segment	2, 69	2.04	$p = .13$
	emotion:segment	2, 69	0.07	$p = .92$
TB-UI (mean)	emotion	1, 69	0.74	$p = .39$
	segment	2, 69	7.95	$p < .01$
	emotion:segment	2, 69	0.23	$p = .79$
TB-LI (mean)	emotion	1, 69	21.52	$p < .01$
	segment	2, 69	0.12	$p = .88$
	emotion:segment	2, 69	0.63	$p = .53$

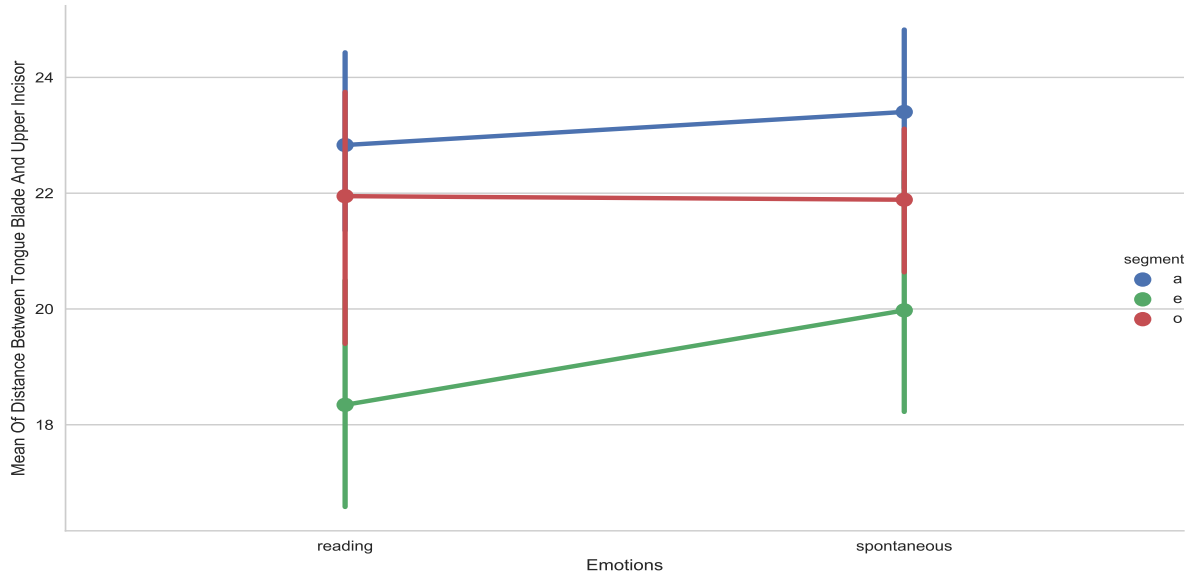


図 4.9: 舌端-上門歯間の平均距離と感情表現及び母音との関係. 縦軸は舌端-上門歯間距離の母音内平均値, 横軸はそれぞれの発話条件を示す. 図中色分けは音声発話時の母音の種類を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

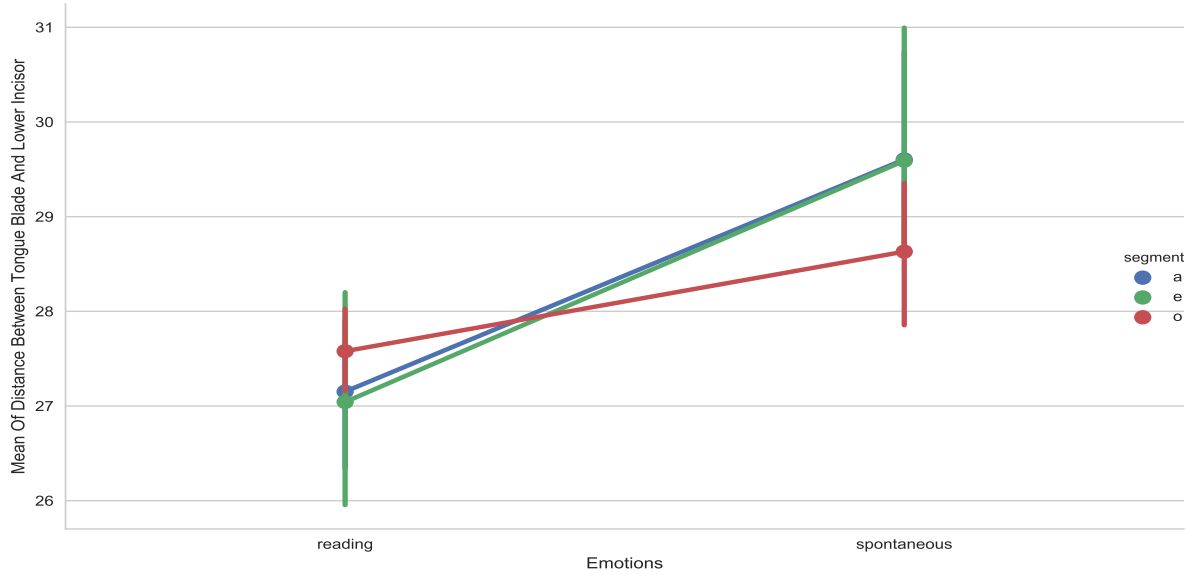


図 4.10: 舌端-下門歯間の平均距離と感情表現及び母音との関係. 縦軸は舌端-下門歯間距離の母音内平均値, 横軸はそれぞれの発話条件を示す. 図中色分けは音声発話時の母音の種類を表す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

第5章　演技/自発音声発話時の調音運動の比較

5.1　目的

4 章では, 自発感情音声発話時の感情表現と言語的情報との関係を調査した. ここで得られた知見と 3 章で得られた知見を比較すると, 演技感情音声発話時と自発感情音声発話時とでは, 感情を表現する調音運動に差があることが示唆される. しかし, 3 章の解析と 4 章の解析とでは音声発話を行っている発話者が異なるため, 演技音声発話時と自発音声発話時の調音運動上の特徴を直接比較することはできない.

そのため本章では, Ericson らの研究で使用されたデータセットの内, 感情模倣音声発話時 (imitation) の調音運動と自発感情音声発話時 (spontaneous) の調音運動とを比較を試みる. ここで解析の対象としたデータの説明に関しては 4.2 節を参照してほしい.

まずは, 感情模倣音声発話時と自発感情音声発話時間において, 感情表現が調音運動に与える影響がどのように異なるのかを検討するために, 感情模倣音声発話時と自発感情音声発話時の調音運動を比較した (5.2 節). 続いて, 演技感情音声発話時と自発感情音声発話時間において, 感情表現が言語情報の生成に与える影響はどのように異なるのかを検討するために, 母音 /a, e, o/ 発話時の調音運動をそれぞれの発話条件ごとに比較した (5.3 節).

5.2　感情表現が調音運動に対し与える影響の比較

本節では, 演技感情音声発話時と自発感情音声発話時間における, 感情表現が調音運動に与える影響の違いを検討する. そのために, Erickson らの研究で使用されているデータ [3] の内, 母音 /a/ を発話している際の調音運動をそれぞれの発話条件 (自発感情音声発話, 感情模倣音声発話, 読み上げ音声発話) 毎に比較した. なお解析に使用した母音数は 71 個であった.

自発感情発話時と感情模倣音声発話時の調音運動上の違いを統計的に解釈するために, それぞれの発話時の条件三種類（自発感情音声, 模倣感情音声, 読み上げ音声）を独立変数とする 1 要因分散分析を行った. 従属変数には, 母音発話時の調音位置を示す指標として調音点間の距離の平均値を使用した. なお, 統計的検定の有意水準はいずれも .01 と定めた.

表 5.1: 開口度に対する 1 要因 (発話条件) 分散分析の結果. 独立変数は自発感情音声発話, 感情模倣音声発話, 読み上げ音声発話の 3 水準. 従属変数は表の通り. ただし, 表中 UI, LI はそれぞれ上下門歯を, UL, LL はそれぞれ上下唇を示す.

Dependent Variable	$d.f.$	F	p
UI-LI (mean)	2, 70	19.40	$p < .01$
UL-LL (mean)	2, 70	12.21	$p < .01$

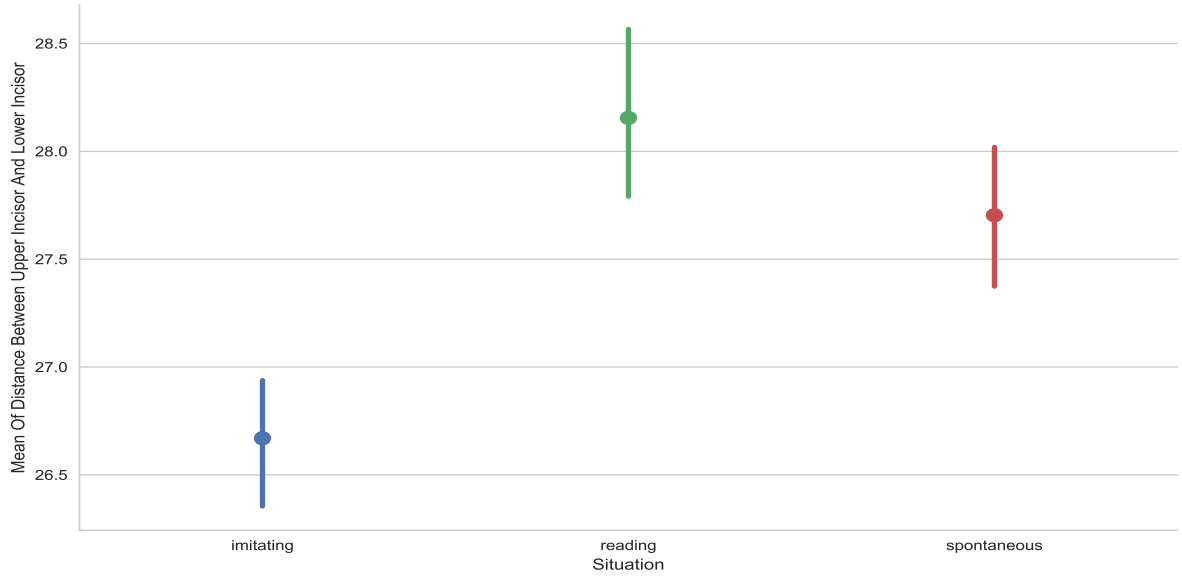


図 5.1: 上門歯-下門歯間の平均距離と発話条件との関係. 縦軸は上門歯-下門歯間距離の母音内平均値, 横軸はそれぞれの発話条件を示す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

5.2.1 開口度

開口度に関して, 分散分析の結果を表 5.1 に示す. 分散分析の結果, 上下門歯 (UI-LI) 間, 及び, 上下唇 (UL-LL) 間の距離平均値において有意差が確認された (UI-LI: $F(2, 70) = 19.40, p < .01$, UL-LL: $F(2, 70) = 12.21, p < .01$).

UI-LI 及び, UL-LL の距離平均と発話条件との関係を図 5.1, 5.2 に示す. それぞれの図を確認すると, いずれの観測点に関しても, 読み上げ音声時と比較して, 自発感情音声発話時, 感情模倣音声発話時ともに開口度が低くなることが確認された. それぞれの発話条件間における感情表現が開口度に与える影響の違いを確認すると, 上下門歯間の距離に関しては感情模倣音声発話時に開口度はより小くなるのに対し, 上下唇間の距離に関しては自発音声発話時の方に開口度がより小くなることが確認された.

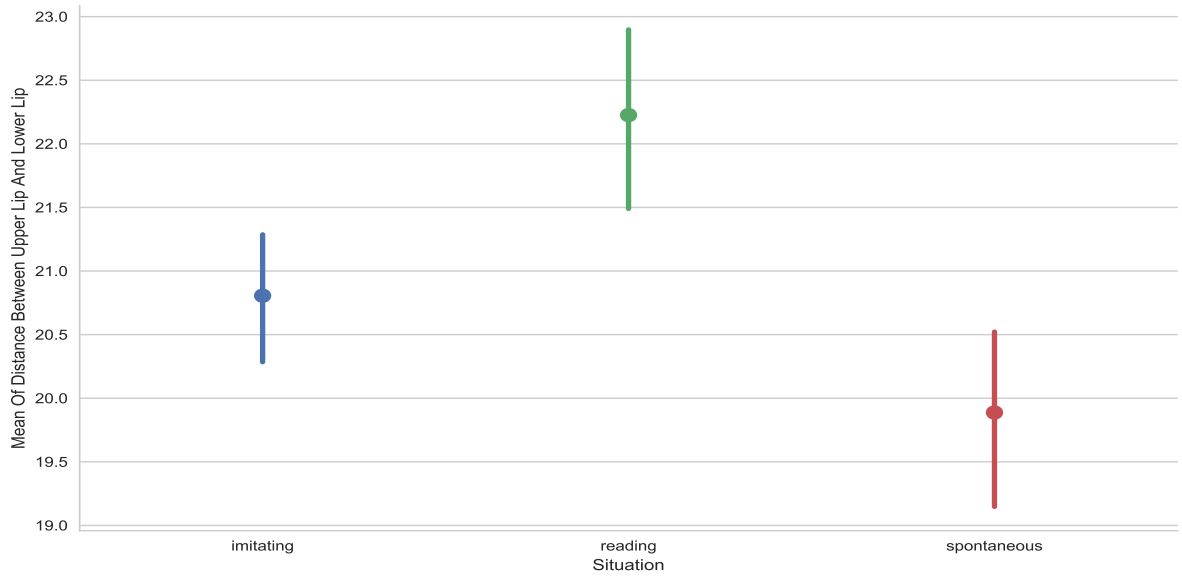


図 5.2: 上唇-下唇間の平均距離と発話条件との関係. 縦軸は上門歯-下門歯間距離の母音内平均値, 横軸はそれぞれの発話条件を示す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

表 5.2: 円唇性に対する 1 要因 (発話条件) 分散分析の結果. 独立変数は自発感情音声発話, 感情模倣音声発話, 読み上げ音声発話の 3 水準. 従属変数は表の通り. ただし, 表中 UL, LL はそれぞれ上下唇を, UI, LI は上下門歯を示す.

Dependent Variable	$d.f.$	F	p
UL-UI (mean)	2, 70	525.23	$p < .01$
LL-LI (mean)	2, 70	286.26	$p < .01$

5.2.2 円唇性

円唇性に関して, 分散分析の結果を表 5.2 に示す. 分散分析の結果, 上唇上門歯 (UL-UI) 間, 及び, 下唇下門歯 (LL-LI) 間の距離平均値において, 有意差が確認された (UL-UI: $F(2, 70) = 525.23, p < .01$, LL-LI: $F(2, 70) = 286.26, p < .01$).

UL-UI 及び, LL-LI の距離平均と発話条件との関係をそれぞれ, 図 5.3, 5.4 に示す. それぞれの図を確認すると, どちらの図においても, 読み上げ音声発話時と比べ, 自発感情音声発話時には, 唇の突き出しが弱くなるのに対し, 感情模倣音声発話時には, 唇の突き出しが強くなることが確認された.

図 5.3: 上門歯-上唇間の平均距離と発話条件との関係. 縦軸は上門歯-上唇間距離の母音内平均値, 横軸はそれぞれの発話条件を示す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

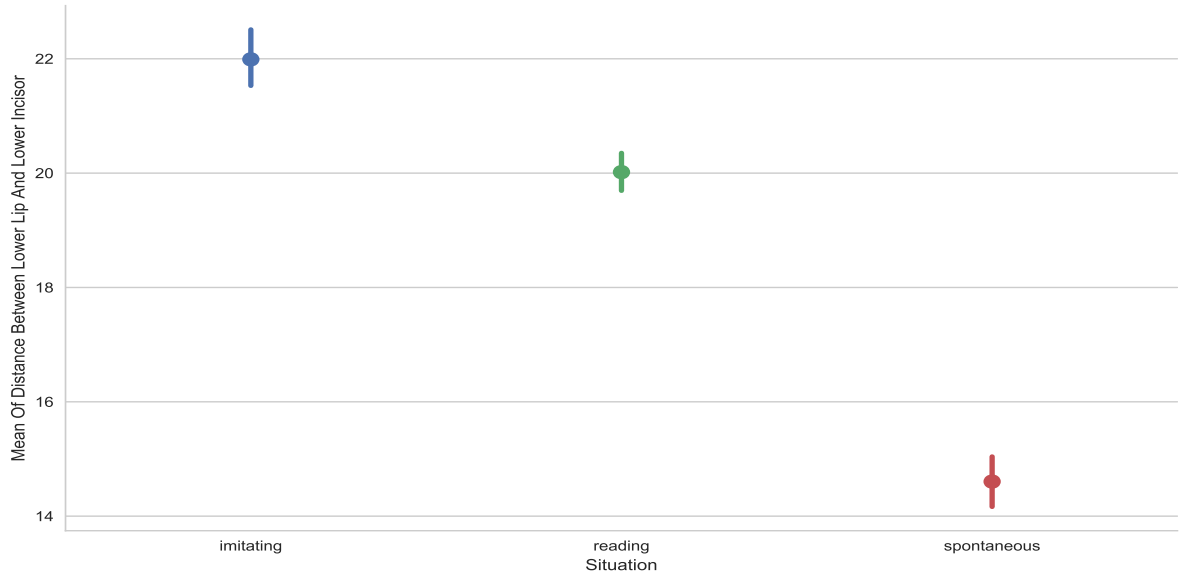


図 5.4: 下門歯-下唇間の平均距離と発話条件との関係. 縦軸は下門歯-下唇間距離の母音内平均値, 横軸はそれぞれの発話条件を示す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

表 5.3: 舌運動に対する 1 要因 (発話条件) 分散分析の結果. 独立変数は自発感情音声発話, 感情模倣音声発話, 読み上げ音声発話の 3 水準. 独立変数は自発音声, 演技音声, 読み上げ音声のいずれか. 従属変数は表の通り. ただし, 表中 TT は 舌先, TB は舌端, TD は 舌背 を示し, UI, LI はそれぞれ上下門歯を示す.

Dependent Variable	$d.f.$	F	p
TT-UI (mean)	2, 70	2.50	$p = 0.08$
TT-LI (mean)	2, 70	3.20	$p = 0.04$
TB-UI (mean)	2, 70	2.65	$p = 0.07$
TB-LI (mean)	2, 70	40.39	$p < .01$

5.2.3 舌運動

舌運動に関して, 分散分析の結果を表 5.3 に示す. 分散分析の結果, 舌端下門歯 (TB-LI) 間の距離平均値においてのみ, 発話条件で有意差が確認された ($F(2, 70) = 40.39, p < .01$).

TB-LI の距離平均と発話条件との関係を図 5.5 に示す. 図を確認すると, 読み上げ音声発話時と比べ, 自発音声発話時には, 舌の位置が後ろに移動しているのに対し, 感情模倣音声発話時には, 舌の位置が前に移動していることが確認された.

5.2.4 考察

本節の目的は演技感情音声発話時と自発感情音声発話時間における, 感情表現が調音運動に与える影響の違いを検討することであった. この目標を達成するために, 母音 /a/ を発話している際の調音運動をそれぞれの発話条件毎に比較した.

解析の結果, 開口度は自発感情音声発話時にも感情模倣音声発話時にも, 読み上げ音声発話時に対し, 低い値を示すことが確認された. これは, 自発感情音声発話時にも, 演技音声発話時にも感情表現は開口度に対し, 同様の影響を与えることを示唆する結果である.

一方, 円唇性, 舌運動に関しては, 自発音声発話時と感情模倣音声発話時とでは, 感情表現が調音運動に対し与える影響は異なった. 読み上げ音声発話時と比べ, 自発感情音声発話時では唇の突き出しが強くなり, 舌が後退するのに対し, 感情模倣音声発話時では唇の突き出しが弱くなり, 舌が前に移動することが判明した. これらの結果は, 自発感情音声発話時と演技音声発話時とでは, 感情表現がこれらの調音運動へ与える影響は異なることを示唆する.

ただし, 4.2 節の解析により, 舌運動に関しては音声生成時の母音の影響を受けやすいことが明らかになっているため, この影響の解釈には, 異なる母音発話時の比較が必要である.

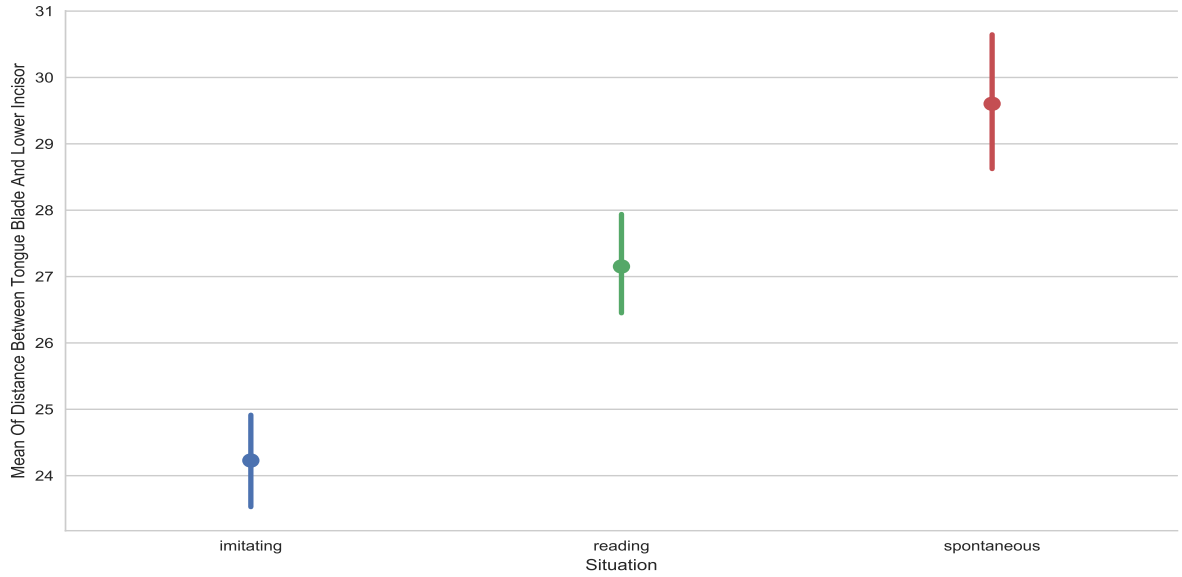


図 5.5: 舌端-下門歯間の平均距離と発話条件との関係. 縦軸は舌背-下門歯間距離の母音内平均値, 横軸はそれぞれの発話条件を示す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

5.3 感情表現が言語情報の生成に与える影響の比較

本節では, 演技感情音声発話時と自発感情音声発話時間において, 感情表現が言語情報の生成に与える影響はどのように異なるのかを検討する. そのために Erickson らの研究で使用されているデータ [3] の内, 母音 /a, e, o/ を発話している際の調音運動をそれぞれの発話条件 (自発感情音声発話, 感情模倣音声発話, 読み上げ音声発話) 毎に比較した. なお解析に使用したデータは, /a/ が 73 個, /e/ が 22 個, /o/ が 21 個である.

それぞれの発話条件と各調音点間の距離との関係を統計的に解釈するために, 3 種類の発話条件と 3 種類の母音を独立変数とした 2 要因分散分析を行った. 従属変数には, 母音発話時の調音位置を示す指標として調音点間の距離の平均値を使用した. ここで, 解析に使用した文章内に含まれる母音数はそれぞれ異なる. そのため, 分散分析の解析にはタイプ II 平方和を利用した. なお, 統計的検定の有意水準はいずれも .01 と定めた.

5.3.1 開口度

開口度に関して, 分散分析の結果を表 5.4 に示す. 分散分析の結果, 上下門歯 (UI-LI) 間, 及び, 上下唇 (UL-LL) 間の距離平均値に対して発話条件要因の主効果が確認された (UI-LI: $F(2, 107) = 27.79, p < .01$, UL-LL: $F(2, 107) = 13.98, p < .01$). また, 両観測点ともに発話時の母音の種類による主効果が確認された (UI-LI: $F(2, 107) = 7.45, p < .01$, UL-LL: $F(2, 107) = 83.53, p < .01$). くわえて, 両観測点ともに発話条件要因と母音の種類との間

表 5.4: 開口度に対する 2 要因 (発話条件, 母音) 分散分析の結果. 独立変数は自発感情音声発話, 感情模倣音声発話, 読み上げ音声発話の 3 水準. 従属変数は表の通り. ただし, 表中 UI, LI はそれぞれ上下門歯を, UL, LL はそれぞれ上下唇を示す.

Dependent Variable	Independent Variable	$d.f.$	F	p
UI-LI (mean)	type	2, 107	27.79	$p < .01$
	segment	2, 107	7.45	$p < .01$
	type:segment	4, 107	0.19	$p < .01$
UL-LL (mean)	type	2, 107	13.98	$p < .01$
	segment	2, 107	83.53	$p < .01$
	type:segment	4, 107	0.51	$p < .01$

に交互作用が確認された (UI-LI: $F(4, 107) = 0.19, p < .01$, UL-LL: $F(4, 107) = 0.51, p < .01$).

UI-LI 及び, UL-LL の距離平均と発話条件との関係を図 5.6, 5.7 に示す. それぞれの図を発話時の母音の種類と発話条件の違いに注目するして観察すると自発感情音声発話時, 感情模倣音声発話時ともに, どの母音においても, 読み上げ音声発話時より, 開口度が低くなることが判明した.

ただし, 図 5.6, 5.7 ともに演技感情音声発話時には, 母音間の開口度同士の距離が狭くなるのに対し, 演技音声発話時には, 広くなった. また, 上下唇の距離に関しては, 母音 /o/ に関しては発話条件の違いによる影響を受けにくいことが確認された.

5.3.2 円唇性

円唇性に関して, 分散分析の結果を表 5.5 に示す. 分散分析の結果, 上唇上門歯 (UL-UI) 間, 及び, 下唇下門歯 (LL-LI) 間の距離平均値に対して, 発話条件要因の主効果が確認された (UL-UI: $F(2, 107) = 27.79, p < .01$, LL-LI: $F(2, 107) = 13.98, p < .01$). また, 両観測点において, 発話時の母音の種類の違いによる主効果が確認された (UL-UI: $F(2, 107) = 7.45, p < .01$, LL-LI: $F(2, 107) = 83.53, p < .01$). くわえて, 発話条件要因と母音の種類の違いとの間に交互作用が確認された (UL-UI: $F(4, 107) = 0.19, p < .01$, LL-LI: $F(4, 107) = 0.51, p < .01$).

UL-UI 及び, LL-LI の距離平均と発話条件との関係をそれぞれ, 図 5.8, 5.9 に示す. それぞれの図を発話時の母音の種類と発話条件の違いに注目するして観察すると図 5.8, 5.9 ともに読み上げ音声発話時と比べ, すべての母音発話時に, 自発感情音声発話時には, 唇の突き出しが弱くなっているのに対し, 感情模倣音声発話時には, 唇をより突き出していることが確認された. この傾向はどの母音に対しての同等の影響を与えるようであるが, UL-UI, LI-LI ともに感情模倣音声発話時には, 読み上げ音声発話時と比較して, 母音間の計測値間の距離が狭くなることが確認された. また, UL-UI に関しては自発感情音声発話

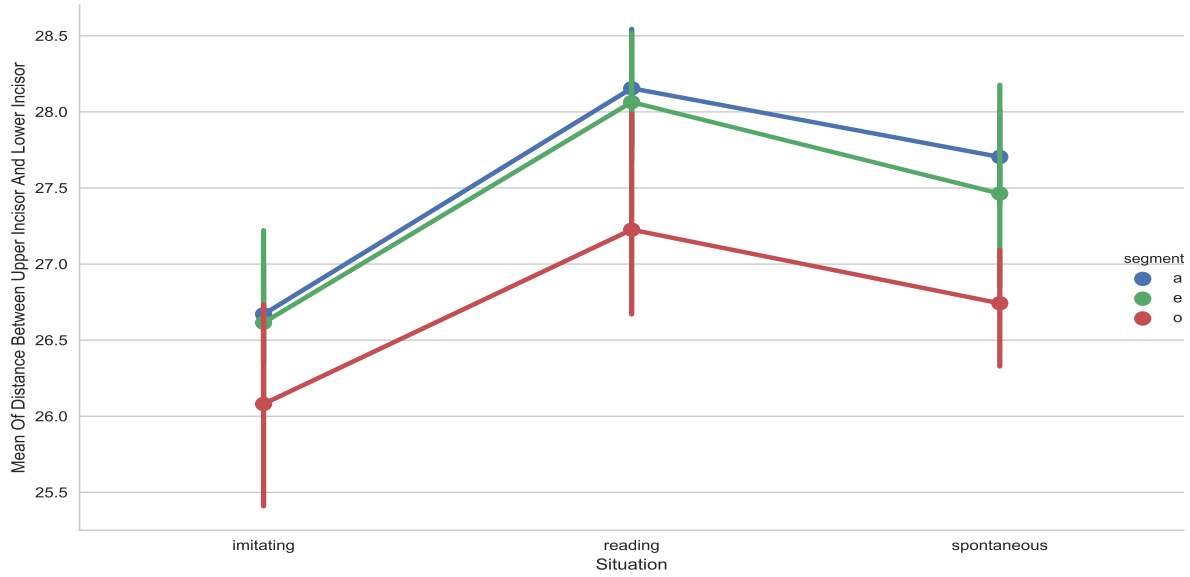


図 5.6: 上門歯-下門歯間の平均距離と発話条件との関係．縦軸は上門歯-下門歯間距離の母音内平均値, 横軸はそれぞれの発話条件を示す．図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す．

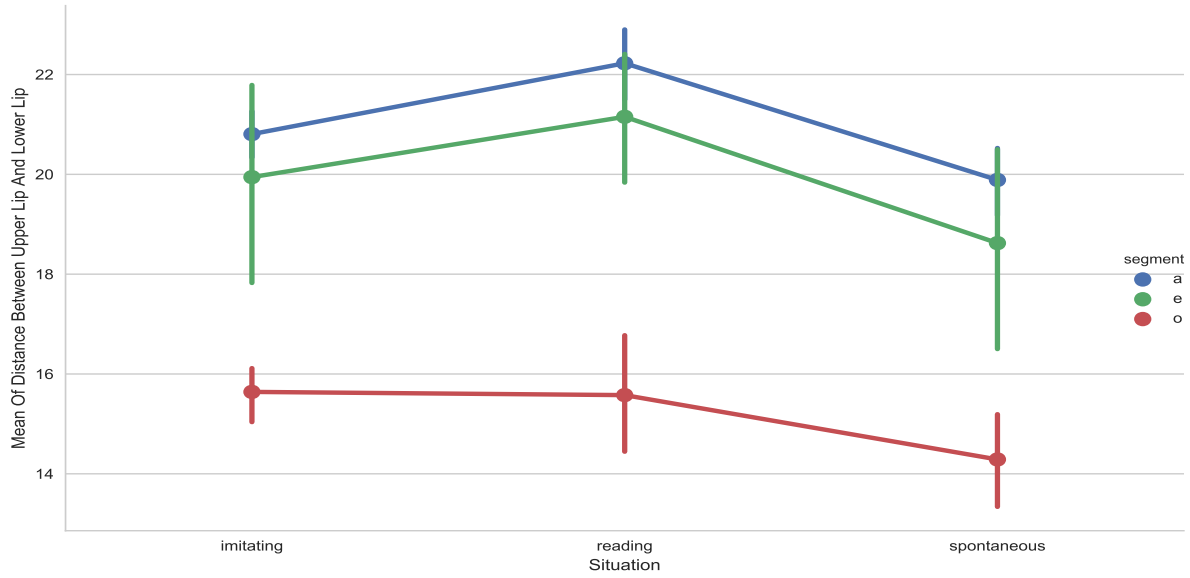


図 5.7: 上唇-下唇間の平均距離と発話条件との関係．縦軸は上門歯-下門歯間距離の母音内平均値, 横軸はそれぞれの発話条件を示す．図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す．

表 5.5: 円唇性に対する 2 要因 (発話条件, 母音) 分散分析の結果. 独立変数は自発感情音声発話, 感情模倣音声発話, 読み上げ音声発話の 3 水準. 従属変数は表の通り. ただし, 表中 UL, LL はそれぞれ上下唇を, UI, LI は上下門歯を示す.

Dependent Variable	Independent Variable	$d.f.$	F	p
UL-UI (mean)	type	2, 107	27.79	$p < .01$
	segment	2, 107	7.45	$p < .01$
	type:segment	4, 107	0.19	$p < .01$
LL-LI (mean)	type	2, 107	13.98	$p < .01$
	segment	2, 107	83.53	$p < .01$
	type:segment	4, 107	0.51	$p < .01$

時に, 母音間の計測値間の距離が広くなることが確認された.

5.3.3 舌運動

舌運動に関して, 分散分析の結果を表 5.6 に示す. 分散分析の結果, 舌端下門歯 (TB-LI) 間の距離平均値に対してのみ, 発話条件要因の主効果が確認された ($F(2, 107) = 74.19, p < .01$). また, 舌端下門歯 (TT-UI), 舌端下門歯 (TB-UI), そして舌端下門歯 (TB-LI) 間の距離平均値に対して発話時の母音間の違いによる主効果が確認された (TT-UI: $F(2, 107) = 7.59, p < .01$, TB-UI: $F(2, 107) = 11.30, p < .01$, TB-LI: $F(2, 107) = 0.15, p < .01$). くわえて, 舌端下門歯 (TB-LI) に関しては発話条件要因と母音の種類との間に相互作用が確認された ($F(4, 107) = 0.39, p < .01$).

TB-LI の距離平均と発話条件との関係を図 5.10 に示す. 図を確認すると, どの母音発話時においても, 読み上げ音声発話時と比べ, 自発感情音声発話時には, 舌の位置が後ろに移動しているのに対し, 感情模倣音声発話時には, 舌の位置が前に移動していることが確認された. この傾向は自発感情音声発話時には特に母音 /a, e/ に対して顕著に確認され, 感情模倣音声発話時には母音 /u/ に対して顕著に確認された. 読み上げ音声発話時における母音間の調音運動上の関係は自発音声時にも感情模倣音声時にも破壊されることが確認された. また, 自発感情音声発話時と感情模倣音声発話時の母音間の調音運動上の距離を比較すると, 感情模倣音声発話時には母音間距離が狭くなるのに対し, 自発感情音声発話時には母音間距離が広くなることが確認された.

5.3.4 考察

本節の目的は, 演技感情音声発話時と自発感情音声発話時間において, 感情表現が言語情報の生成に与える影響はどのように異なるのかを検討することであった. そのために母音 /a, e, o/ を発話している際の調音運動をそれぞれの発話条件毎に比較した.

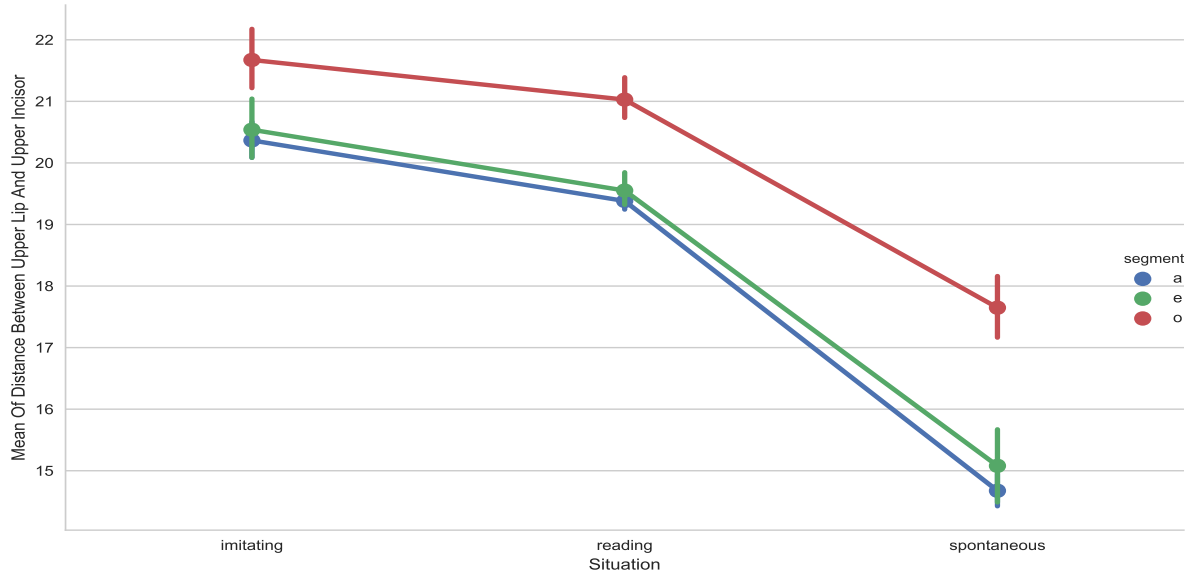


図 5.8: 上門歯-上唇間の平均距離と発話条件との関係. 縦軸は上門歯-上唇間距離の母音内平均値, 横軸はそれぞれの発話条件を示す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

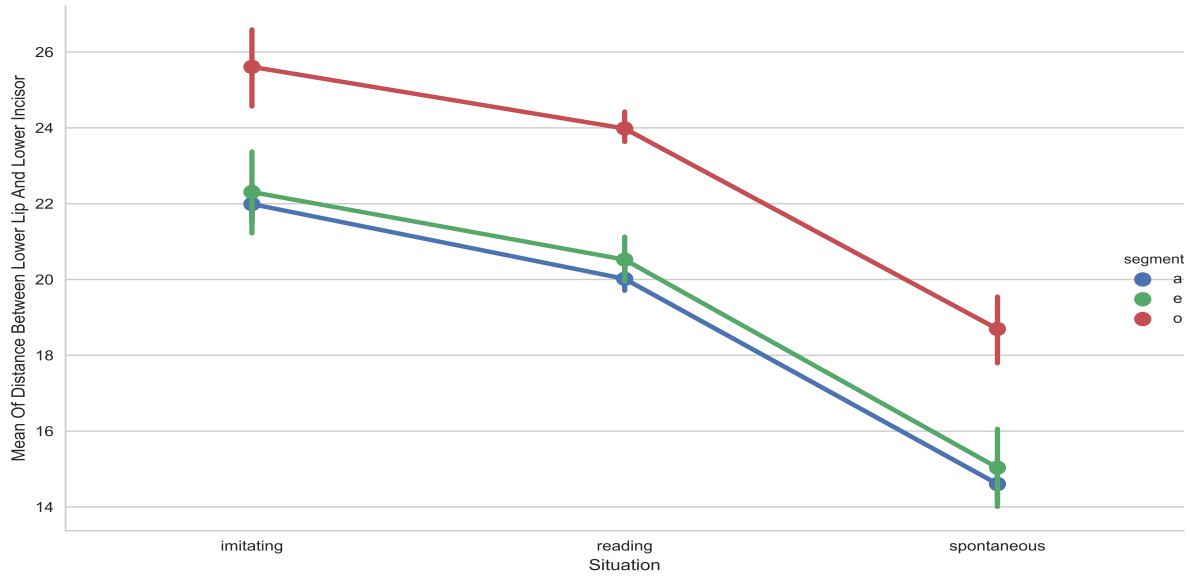


図 5.9: 下門歯-下唇間の平均距離と発話条件との関係. 縦軸は下門歯-下唇間距離の母音内平均値, 横軸はそれぞれの発話条件を示す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

表 5.6: 舌運動に対する 2 要因 (発話条件, 母音) 分散分析の結果. 独立変数は自発感情音声発話, 感情模倣音声発話, 読み上げ音声発話の 3 水準. 従属変数は表の通り. ただし, 表中 TT は舌先, TB は舌端, TD は 舌背 を示し, UI, LI はそれぞれ上下門歯を示す.

Dependent Variable	Independent Variable	$d.f.$	F	p
TT-UI (mean)	type	2, 107	2.89	$p = .05$
	segment	2, 107	7.59	$p < .01$
	type:segment	4, 107	0.18	$p = .94$
TT-LI (mean)	type	2, 107	4.65	$p = .01$
	segment	2, 107	2.63	$p = .07$
	type:segment	4, 107	0.15	$p = .96$
TB-UI (mean)	type	2, 107	4.67	$p = .01$
	segment	2, 107	11.30	$p < .01$
	type:segment	4, 107	0.11	$p = .97$
TB-LI (mean)	type	2, 107	74.19	$p < .01$
	segment	2, 107	0.15	$p < .01$
	type:segment	4, 107	0.39	$p < .01$

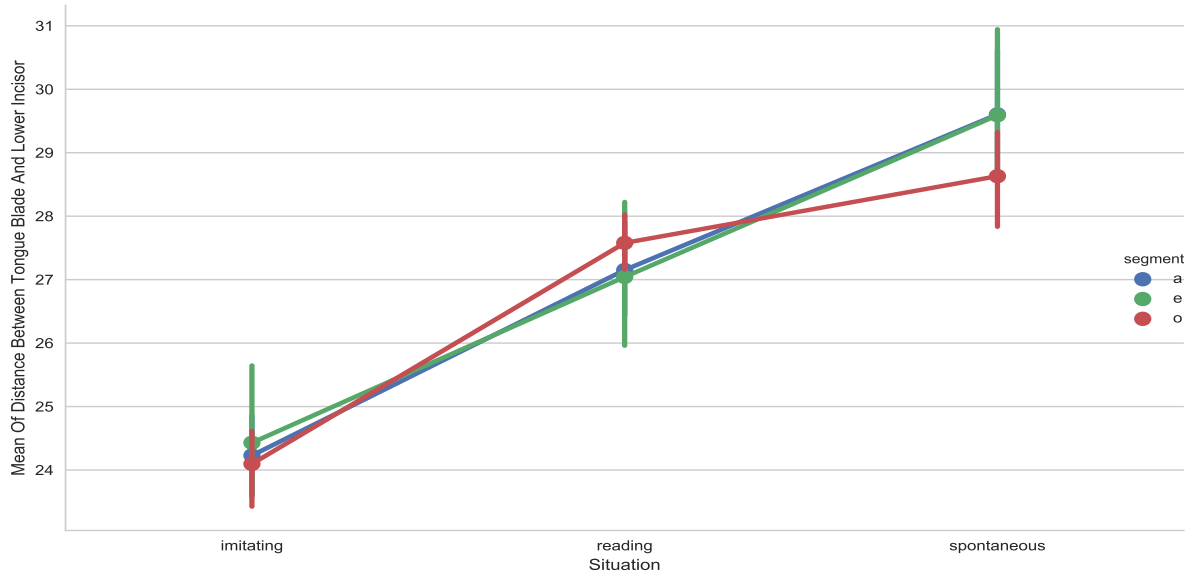


図 5.10: 舌端-下門歯間の平均距離と発話条件との関係. 縦軸は舌背-下門歯間距離の母音内平均値, 横軸はそれぞれの発話条件を示す. 図中それぞれの点は従属変数の平均値を示し, 実線は 95 % 信頼区間を示す.

解析の結果, 開口度, 円唇性に関しては, 自発感情音声発話時, 感情模倣音声発話時ともにすべての母音に対して, 感情表現の影響は一定であり, 読み上げ音声発話時の調音運動上の母音間の位置関係は保持されていることが判明した. 一方, 舌運動に関しては母音によって感情表現の影響の大きさが変化し, 読み上げ音声発話時の調音運動上の母音間の位置関係は保持されないことが判明した. 特に母音 /o/ を発話している場合, 自発感情音声発話時には感情表現が舌の前後方向の位置与える影響が大きくなるのに対し, 感情模倣音声発話時には小くなることが確認された. 2.2 節で説明したとおり, 母音 /e/ と /o/ の言語的特徴は舌の前後の位置によって決まる. このことを踏まえると上記の結果は, 感情表現は, 自発感情音声発話時において, 感情模倣音声発話時より, 言語的情報の生成を阻害する傾向が強くなることを示唆する.

また, 開口度, 円唇性, 舌運動のいづれの調音運動も感情模倣音声発話時には調音運動上の母音間距離が狭くなり, 自発感情音声発話時には広くなることが確認された. これは, 感情模倣音声時には, ある感情表現を行っている場合には調音運動を感情表現毎に, 一定の範囲に収束させるのに対し, 自発感情音声発話時には, 広い範囲に分散させることを示唆する結果である.

5.4 本章のまとめ

本章では, 感情模倣音声発話時 (imitation) と 自発感情音声発話時 (spontaneous) の調音運動の違いを明らかにするために, 悲しみ音声を対象に, 自発感情音声発話時, 感情模倣音声発話時, 読み上げ音声発話時の 3 種類の発話条件間の調音運動の比較を行った.

5.2 節の解析では, それぞれの発話条件間で, 感情表現が調音運動に対し与える影響がどのように違うのかを調査した. その結果, 開口度は自発感情音声発話時でも, 感情模倣音声発話時でも同様の方法で感情表現を行っていることが判明した. 一方で, 円唇性や舌運動に関しては, 自発感情音声発話時と感情模倣音声発話時とでは, 感情表現を行うための調音運動が異なることが示された.

これらの結果からある感情を意図的に表現しようとする場合に使用する調音運動と, 実際に, ある強い感情が音声に滲み出てしまっている場合に感情表現の影響を受ける調音運動は必ずしも一致しないことが示唆された.

5.3 節の解析では, ぞれぞれの発話条件間で, 感情表現が言語情報の生成に与える影響がどのように違うのかを調査した. 解析の結果, 感情模倣音声発話時には, 読み上げ音声発話時の調音運動上の母音間の位置関係が保持される傾向にあるのに対し, 自発音声発話時には, 特に舌の前後方向の位置において調音運動上の母音間の位置関係は保持されないことが判明した. また, 感情模倣音声発話時には, 感情表現を行う場合には調音点間の距離を一定の範囲に収束させるのに対し, 自発感情音声発話時には, 広い範囲に分散させることが示された.

これらの結果からある感情を意図的に表現しようとする場合に使用する調音運動と, 実際に, ある感情が音声に滲み出てしまっている場合に感情表現が言語情報の生成へ与える

影響は異なることが示唆された.

感情模倣音声発話時には, 特定の感情表現に対し, 調音運動が一定の範囲に収束させることから, どのような母音を発話している場合にも似たような音響的特徴をもつ音声が生成されると予測される. ただし, 読み上げ音声発話時の調音運動上の母音間の位置関係が保持される傾向にあることから, 聞き手が言語音としての各母音が異なる音であると知覚できれば, その音がどの母音であるのかを同定することは可能であると予測される. つまり, 言語情報の伝達と感情の伝達の両方が発話目的に含まれる感情模倣音声発話時には, 言語音の同定に必要な音響特徴を生成するための調音運動を保持しつつ, 伝達する感情表現に対し, 一定の調音運動を行うことが考察される.

一方で, 感情表現の伝達は発話目的にふくまれない自発感情音声発話時には, その音声に滲みでてしまう感情表現の影響から, 読み上げ音声発話時の調音運動上の母音間の位置関係は保持されない. そのような条件で言語情報の伝達という発話目的を達成するために, 調音運動上の母音間距離が広げることで, 各母音が言語音として異なる音であると弁別しやすい音声の生成を促進することが予測される.

第6章　全体考察

本研究の目標は, 言語情報の伝達と感情情報の伝達という発話目的の違いに注目し, これらの発話目的の違いがどのように音声生成時の調音運動に影響を与えるのかを解明することであった. そのために, 発話目的が異なる音声発話時の感情表現と言語情報の生成との関係を調音運動の観点から調査した.

以下に本研究で明らかになった調音運動における感情表現が言語的情報に与える影響に関する知見をもとに発話目的の違いがどのように音声生成時の調音運動に影響を与えるのかを考察する.

解析対象とした音声発話時の発話目的

本研究で解析したそれぞれの音声発話の発話目的をそれぞれの発話収録条件から考察していく.

3 章では, プロの声優による演技感情音声発話を解析した. ここで, 発話者は特定の文章を教示された感情毎に演じ分けながら発話することが求められた. そのため, 言語情報の伝達と感情表現とを同時に達成することを目的とした音声発話であると考えられる. ただし, 上記発話セットの内, Neutral 音声発話時に関しては, 感情表現を行わないことを目的とした音声発話である.

4 章で解析した自発感情音声発話に関しては, 発話者は発話収録時の直前に母親を亡くしており, かつ, その話題を話している際の発話である. そのため, 特に感情表現を行っている音声発話であると考えられる. ただし, この音声発話においては, 発話者は母親を亡くした時の様子を発話相手に説明することが求められており, 明示的にその際の感情を表現することを指示されてはいないことに注意する必要がある. このことを考慮すると発話者は言語情報の伝達を目的として発話を行っているものと考えられる.

最後に, 5 章に登場した感情模倣音声に関しては, 発話収録時に発話者は上記の自発感情音声発話を聞き, その音声と同じ内容を同じような感情表現を行いながら発話することが求められている. このため, 発話目的としては, プロの声優による演技感情音声発話と同様, 言語情報の伝達と感情表現とを同時に達成することを目的とした音声発話であると考えられる. だたし, 発話者は音声表現の素人であり, 特に感情表現の再現を求められている. このことを考慮すると発話者は言語情報の伝達よりも感情表現の伝達を目的とした音声発話であること考えられる.

感情表現の伝達を目的とした音声発話

以下に感情表現の伝達という発話目的が調音運動にどのような影響を与えるのかを考察する.

まず, 感情が表現されている音声発話においては, 感情表現を目的とするか否かに拘わらず, 感情が表現されていない音声発話時とは調音運動上の特徴が異なる. 3.3 節の解析では, 演技音声発話時には, 表現する感情毎に, 開口度, 円唇性, 舌運動が, 一定の方向に変化することを示した. 続く, 3.4 節の解析では, 二名の発話者の調音運動を感情表現毎に比較し, 上記の感情表現が調音運動に与える影響は発話者の個人性を越えたものであることを示した. また, 5.2 節の解析では, 感情模倣音声発話時と読み上げ音声発話時とでは, 開口度, 円唇性, 舌運動の特性が異なることを示した. これらの結果から, 音声に感情が表出されている場合には, 調音運動は感情により一定の方向に変化すると考えられる.

一方で感情表現の伝達が明示的に目的にしている場合と, 明示的には目的とされていない場合とでは, 感情を表現するための調音運動が異なることも示された. 例えば, 3.3 節 と 5.2 節の円唇性に対する解析では, 演技感情音声発話時にも, 感情模倣音声発話時にも悲しみを表現する際に, 唇の突き出しが強まっているが, 自発感情音声発話時では唇の突き出しは弱くなる (図 3.6, 図 5.3). この結果は, 感情を伝達しようという発話目的を持っている場合とそうでない場合とで感情を表現する調音運動が異なることが示唆された.

感情表現の伝達を発話目的とすることを明示的に求められた音声発話の場合には, 音声発話時に調音パラメータを一定の値にする傾向が確認された. 3.6 節の解析では, 母音 /a, e, o/ いずれの母音発話を行っている場合にも, 円唇性が感情表現に対し一定の値を示すことが判明した (図 3.22). また, 5.3 節の解析では, 感情模倣音声発話発話時には, 舌運動が感情表現に対し一定の値を示すことが確認された (図 5.10). これらの結果から, 感情表現の伝達という発話目的は調音運動を表現する感情に対し, 一定の値に収束させる性質をもつと考えられる.

言語情報の伝達を目的とした音声発話

以下に言語の伝達という発話目的が調音運動にどのような影響を与えるのかを考察する. 本研究で解析した音声発話データはいずれも感情表現がなされたものであるが, 言語情報の伝達という目的も同時にもつ音声発話である. 収録時の教示を考慮すると, 演技感情音声発話時と自発感情音声発話時には明確な目的付けがなされているのに対し, 感情模倣音声発話では, 目的づけが弱いと考えられる.

感情表現の伝達と言語情報の伝達という二つの発話目的を持った音声発話の場合, 母音毎の調音運動上の位置関係は担保されることが示された (図 3.22 及び 図 3.25).

一方で, 感情表現の伝達は発話目的に含まれていないにも拘わらず音声に感情表現が表出してしまう音声発話の場合, 音声に表出されてしまう感情表現の影響で, 母音

毎の調音運動上の位置関係を担保することが困難になるものと考えられる. 4.4 節の解析では, 言語情報の生成に重要な舌の前後方向の母音間の位置関係が, 自発感情音声発話時と読み上げ音声発話時とで異なることを示した (図 4.3). だたし, この解析において, 自発感情音声発話時と読み上げ音声発話時とでは調音運動上の母音間の距離は広がることも確認されている. このような調音運動からは, 母音の同定は困難であれど, 言語音として異なる音であることを判別可能な音声が生成されるものと予測される.

これらの結果から, 言語情報の伝達という発話目的は母音毎の位置関係と母音間の距離という二つの調音運動上の特徴に影響を与えるものであると考えられる. また, 特に, 発話者の強い感情が, 発話目的と関係なく, 調音運動に影響してしまう場合には, 母音間の距離を拡大することで, 言語上, 異なる音であること知覚できる音声を生成しようとすると考えられる.

上記の知見から感情表現と言語情報との関連について考察する. まず, 冒頭に挙げた, 言語的情報の生成に重要な調音運動以外で感情表現を行うことで, パラ言語情報の生成と言語情報の生成を同時に達成しているのではないかという仮説に関しては限定的には正しいものであると考えられる.

一方で, 強く感情表現が滲み出てしまっている音声である自発感情音声発話時には, 調音音声学の観点から音韻の対立に重要であると見なされる調音運動も感情によって影響をうけ, 音韻の同定にとって致命的に思われる調音運動の変化を起こすことが確認された. だたし, このような音声発話時においては, 調音運動上の母音間距離が拡大することが確認されている. この結果から, 言語的情報の生成に重要な調音運動以外で感情表現を行うという方法がとれない場合には言語上, 異なる音であることは強調する調音運動を行うことで, 言語的情報の生成とパラ言語的情報の生成を両立している可能性が示された.

第7章　まとめと結論

7.1　本研究の成果

本研究では言語情報の伝達と感情表現の伝達という発話目的の違いに注目し, 感情音声発話時の感情表現が言語的情報の生成にどのような影響を与えるのかを解明することであった. そのため, 本研究では発話目的が異なると考えられる演技・自発感情音声発話時の調音運動の特徴を調査した. 以下に本研究で明らかになった演技感情音声発話時と自発感情音声発話時の感情表現が言語的情報の生成に与える影響をまとめる.

演技感情音声発話時の調音運動

演技感情音声の場合, 基本的にそれぞれの調音点を一定の範囲に収束させることで感情表現を行う. 特に音声表現になれたプロの声優の場合には, 言語情報の伝達と感情表現の伝達という発話目的を同時に達成するために, 言語情報の生成に重要な調音運動と, そうではない調音運動とを使い分けることが分かった.

自発感情音声発話時の調音運動

自発感情音声発話時には言語情報の生成に重要な調音運動に関しても感情表現の影響をうける. その結果, 調音音声学上の音韻対立を覆すような調音運動を行う場合もある. ただし, 感情音声発話時のそれぞれの調音運動の母音間の距離は自発感情音声発話では広くなる. このようにすることで, 生成される音声が言語上, 異なる音であることを強調することができるものだと考えられる.

そして, 分析結果から発話目的の違いがどのように音声生成時の調音運動に影響を与えるのかを考察した. また, 感情音声発話においてはどのように, パラ言語情報の生成と言語情報の生成を同時に達成しているのかについて論じ, 特に音声に表出する感情表現により, 言語情報の生成が阻害されやすい音声発話においては, 生成される音声が言語上, 異なる音であることを強調することで, パラ言語情報の生成と言語情報の生成を同時に達成している可能性を示した.

7.2　今後の課題

本研究で挙げた考察の妥当性を確認するためには, 演技感情音声と自発感情音声の知覚的側面を調査するために聴取実験を行う必要がある. まず, 同じ感情表現がされた音声に

関して, それがどの程度感情表現を行っているのかを客観的に比較する. その上で, 同程度の感情表現がなされた演技・自発感情音声間の言語的特性を調査する. ここでは特に生成された音声の音韻の同定課題および, 聞き分け実験をおこなうことにより, 自発感情音声発話時の調音運動でみられた母音空間の拡大がどのような意味を持つのかを知覚的な側面から意味付けることができると考える. この際に, 収録する音声の発話者は音声による感情表現に慣れているプロの声優であることが望ましい. そして, 本研究での演技自発感情音声発話時の調音運動の比較は, 悲しみ音声のみであったが, より多様な種類の感情表現を分析する必要がある. 特に, 舌運動に関しては感情表現を目的とした調音運動に, 個人差の影響が大きいことが確認されたため, より発話者を増やした検討も行う必要がある.

謝辞

本研究を行うにあたり, 終始多大なるご指導・御鞭撻を賜りました北陸先端科学技術大学院大学情報科学研究科赤木 正人 教授, 党 建武 教授, 鵜木 祐史 教授に深く感謝の意を表します. 並びに本研究の課程において多くの御協力を賜りました, 本学の教官の皆様に心より感謝致します.

本研究において用いた演技感情音声発話データの収録にご協力頂いた札幌保健医療大学の末光厚夫氏並びに, 自発音声発話データの提供・利用許可して頂いた Haskins Laboratories の Donna Erickson 氏に深謝致します.

参考文献

[1] Marc Schröder. *Speech and Emotion Research: An overview of research frameworks and a dimensional approach to emotional speech synthesis*. Institut für Phonetik, 2004.

[2] Hiroya Fujisaki. "Prosody, models, and spontaneous speech". In: *Computing Prosody*. Springer, 1997, pp. 27–42.

[3] Donna Erickson et al. "Exploratory study of some acoustic and articulatory characteristics of sad speech". In: *Phonetica* 63.1 (2006), pp. 1–25.

[4] Dimitrios Ververidis and Constantine Kotropoulos. "Emotional speech recognition: Resources, features, and methods". In: *Speech communication* 48.9 (2006), pp. 1162–1181.

[5] Carl E Williams and Kenneth N Stevens. "Emotions and speech: Some acoustical correlates". In: *The Journal of the Acoustical Society of America* 52.4B (1972), pp. 1238–1250.

[6] Chul Min Lee and Shrikanth S Narayanan. "Toward detecting emotions in spoken dialogs". In: *IEEE transactions on speech and audio processing* 13.2 (2005), pp. 293–303.

[7] Serdar Yildirim et al. "An acoustic study of emotions expressed in speech." In: *INTERSPEECH*. 2004.

[8] Carlos Busso, Sungbok Lee, and Shrikanth Narayanan. "Analysis of emotionally salient aspects of fundamental frequency for emotion detection". In: *IEEE Transactions on audio, speech, and language processing* 17.4 (2009), pp. 582–596.

[9] Osamu Fujimura, Shigeru Kiritani, and Haruhisa Ishida. "Computer controlled radiography for observation of movements of articulatory and other human organs". In: *Computers in biology and medicine* 3.4 (1973), 371IN1379–378IN3384.

[10] Maureen Stone. "A guide to analysing tongue motion from ultrasound images". In: *Clinical linguistics & phonetics* 19.6-7 (2005), pp. 455–501.

[11] Khalil Iskarous et al. "The interaction between prosodic boundaries and accent in the production of sibilants". In: *ISCA Proceedings of the 5th International Conference on Speech Prosody, Chicago*. Citeseer. 2010, pp. 1–4.

[12] Joseph S Perkell et al. “Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements”. In: *The Journal of the Acoustical Society of America* 92.6 (1992), pp. 3078–3096.

[13] Shrikanth S Narayanan, Abeer A Alwan, and Katherine Haker. “An articulatory study of fricative consonants using magnetic resonance imaging”. In: *The Journal of the Acoustical Society of America* 98.3 (1995), pp. 1325–1347.

[14] Shrikanth Narayanan et al. “An approach to real-time magnetic resonance imaging for speech production”. In: *The Journal of the Acoustical Society of America* 115.4 (2004), pp. 1771–1776.

[15] 前川喜久雄 et al. “調音運動から見たパラ言語情報の生成”. In: **日本音響学会研究発表会講演論文集** 1999.2 (1999), pp. 257–258.

[16] Donna Erickson, Caroline Menezes, and Akinori Fujino. “Some articulatory measurements of real sadness.” In: *INTERSPEECH*. 2004.

[17] Jangwon Kim et al. “A kinematic study of critical and non-critical articulators in emotional speech production”. In: *The Journal of the Acoustical Society of America* 137.3 (2015), pp. 1411–1429.

[18] Catherine P Browman and Louis Goldstein. “Articulatory gestures as phonological units”. In: *Phonology* 6.02 (1989), pp. 201–251.

[19] 田窪行則 et al. **岩波講座言語の科学**. Vol. 2. 岩波書店, 1998.

[20] Thurid Vogt, Elisabeth André, and Johannes Wagner. “Automatic recognition of emotions from speech: a review of the literature and recommendations for practical realisation”. In: *Affect and emotion in human computer interaction*. Springer, 2008, pp. 75–91.

[21] 赤木正人. “音声に含まれる感情情報の認識 : 感情空間をどのように表現するか”. In: **日本音響学会誌** 66.8 (Aug. 2010), pp. 393–398. ISSN: 03694232.

[22] 平賀 裕 et al. “音声に含まれる感情情報抽出の一検討”. In: **電子情報通信学会技術研究報告**. *HC,* **ヒューマンコミュニケーション** 93.439 (Jan. 1994), pp. 1–8.

[23] 前川喜久雄. “『日本語話し言葉コーパス』の分節音ラベリング Version1.0”. In: (2004).

学会発表リスト

[24] 浅井拓也, 末光厚夫, and 赤木正人. “EMA データに基づいた演技感情音声発話時の調音運動制御の分析”. In: **日本音響学会研究発表会講演論文集** (2017).

[25] 浅井拓也, 末光厚夫, and 赤木正人. “Articulatory Characteristics of Expressive Speech in Activation-Evaluation Space”. In: *NCSP* (2017).