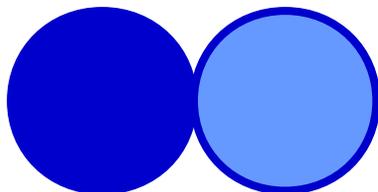
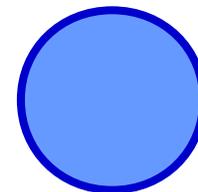


# 1-P-32

## 2モーラ語/ee/におけるパラ言語 情報の認知に対するF0および モーラ継続長の影響

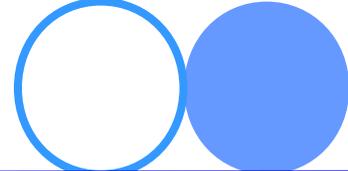
---



竹川弘一(島根大学・総合理工)

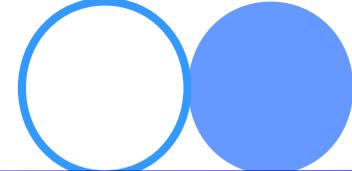
小林聡(島根大学・総情セ)

2010年3月8日



- 声の高さ及び長さの変化から人間はどのような情報(パラ言語情報)を受け取るのかを調査・分析する

# パラ言語情報の定義



- ◆ 音声に含まれる情報は、言語情報、パラ言語情報、非言語情報の3つ

## 藤崎による定義

### パラ言語

(意識的に制御可能)

意図、態度、  
グループ特有のスタイル

### 非言語

(意識的に制御不可能)

個人的特徴、  
身体的状態、

感情等心理的状态

## 本研究の仮定

### パラ言語

(制御が容易)

(対話中の動的変化が容易)

態度、意図、スタイル  
感情等心理的状态

≡ 精神的状态

### 非言語

(制御が困難)

(対話中の動的変化が困難)

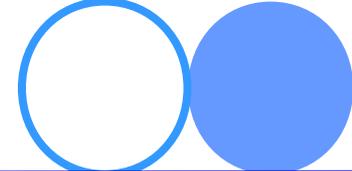
個人的特徴、  
身体的特徴、  
体調

≡ 肉体的状态

- 石井ら、“自然発話に現れる感動詞の発話スタイルと機能の分析” 日本音響学会講演論文集(2008、9) pp. 269-270  
→多くの感動詞に共通した音響的特徴の機能があることに注目
- 高見、“パラ言語情報の認知における声の高さ及びその変化の影響” 島根大学卒業論文(2008)  
→「ええ」の声の高さやその変化から人間が受け取る印象
- 藤原、“「ええ」と「ああ」に対するパラ言語情報の認識におけるF0の影響” 島根大学卒業論文(2009)  
→「ああ」の声の高さやその変化から人間が受け取る印象  
→異なる語彙により人間が受け取るパラ言語情報に差異が生じるかの調査(「ええ」と「ああ」)

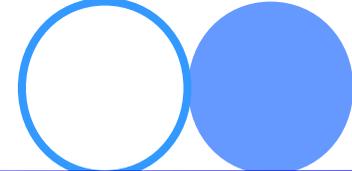
# 実験方法(1/2)

- 基準音声×3＋評価対象音声×1を1セットとし、1人の被験者に60セットを聞いてもらう。実験時間は1人あたり2時間程度
- 基準音声を基準に評価対象音声を聞いた際に受けた印象を28個の印象語について各々1(受け取れない)から5(はっきり受け取れる)の5段階で評価
- 被験者数は30名(大学生及び大学院生)
  - 評価対象音声・・・前後の「え」の高さ及び長さを変えた「ええ」  
(同じ長さの場合もある)
  - 基準音声・・・前後の高さは一般男性の基本周波数120Hzの「ええ」  
前後の長さは1モーラ0.2秒



- 印象語のリスト

悲しみ 無関心 不満 信頼 驚き 恐れ  
落胆 感心 安堵 慎み 怒り 好意 中立  
高揚 満足 疑問 冷静 同情 喜び 軽蔑  
嫌悪 不安 強調 肯定 苛立ち 相槌  
自信あり 自信なし



- 合成音声は「ええ」という2モーラ語

○モーラとは日本語で一音(一拍)と感じる音声の単位

- 音声はフォルマント合成方式を用いたklattの合声器で作成した合成音声を使用

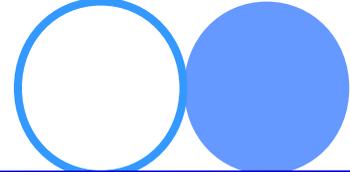
○フォルマントは声道共振周波数で、母音の弁別は第一フォルマントと第二フォルマントによって行うことができる

第1フォルマントと第2フォルマントを求める式

$$F_1 = -0.565 \times F_0 + 580.2$$

$$F_2 = 0.732 \times F_0 + 1712.56$$

# 「ええ」に着目した理由



- 人間同士の対話では、「え」「あ」などの非語彙的な発話が頻繁に現れる。これらの字面には特定の意味はないが、発話の仕方によって何らかの意図、態度、感情などのパラ言語情報を伝達している
- 本研究では対話音声に頻繁に現れ、発話の仕方によって豊富な種類のパラ言語情報が伝達される「ええ」に着目した

# 先行研究で用いた音声

- 先行研究では、120Hzを中心に1オクターブ半幅で作成できる75種類の音声を使用していたが、今回の実験ではその中から5種類の音声を使用する

声の高さが同じ

音程	ファ0	ファイ0	ソ0	サ0	ラ0	チ0	シ0	ド1	デ1	レ1	リ1	ミ1
完全1度	ファ0	ファイ0	ソ0	サ0	ラ0	チ0	シ0	ド1	デ1	レ1	リ1	ミ1

声の高さが上がる

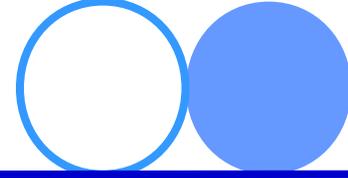
音程	ファ0	ファイ0	ソ0	サ0	ラ0	チ0	シ0	ド1	デ1	レ1	リ1	ミ1
短2度	ファイ0	ソ0	サ0	ラ0	チ0	シ0	ド1	デ1	レ1	リ1	ミ1	ファ1
長2度	ソ0	サ0	ラ0	チ0	シ0	ド1	デ1	レ1	リ1	ミ1	ファ1	ファイ1
短3度	サ0	ラ0	チ0	シ0	ド1	デ1	レ1	リ1	ミ1	ファ1	ファイ1	ソ1
長3度	ラ0	チ0	シ0	ド1	デ1	レ1	リ1	ミ1	ファ1	ファイ1	ソ1	
完全4度	チ0	シ0	ド1	デ1	レ1	リ1	ミ1	ファ1	ファイ1	ソ1		
減5度	シ0	ド1	デ1	レ1	リ1	ミ1	ファ1	ファイ1	ソ1			
完全5度	ド1	デ1	レ1	リ1	ミ1	ファ1	ファイ1	ソ1				
短6度	デ1	レ1	リ1	ミ1	ファ1	ファイ1	ソ1					
長6度	レ1	リ1	ミ1	ファ1	ファイ1	ソ1						
短7度	リ1	ミ1	ファ1	ファイ1	ソ1							
長7度	ミ1	ファ1	ファイ1	ソ1								
完全8度	ファ1	ファイ1	ソ1									

声の高さが下がる

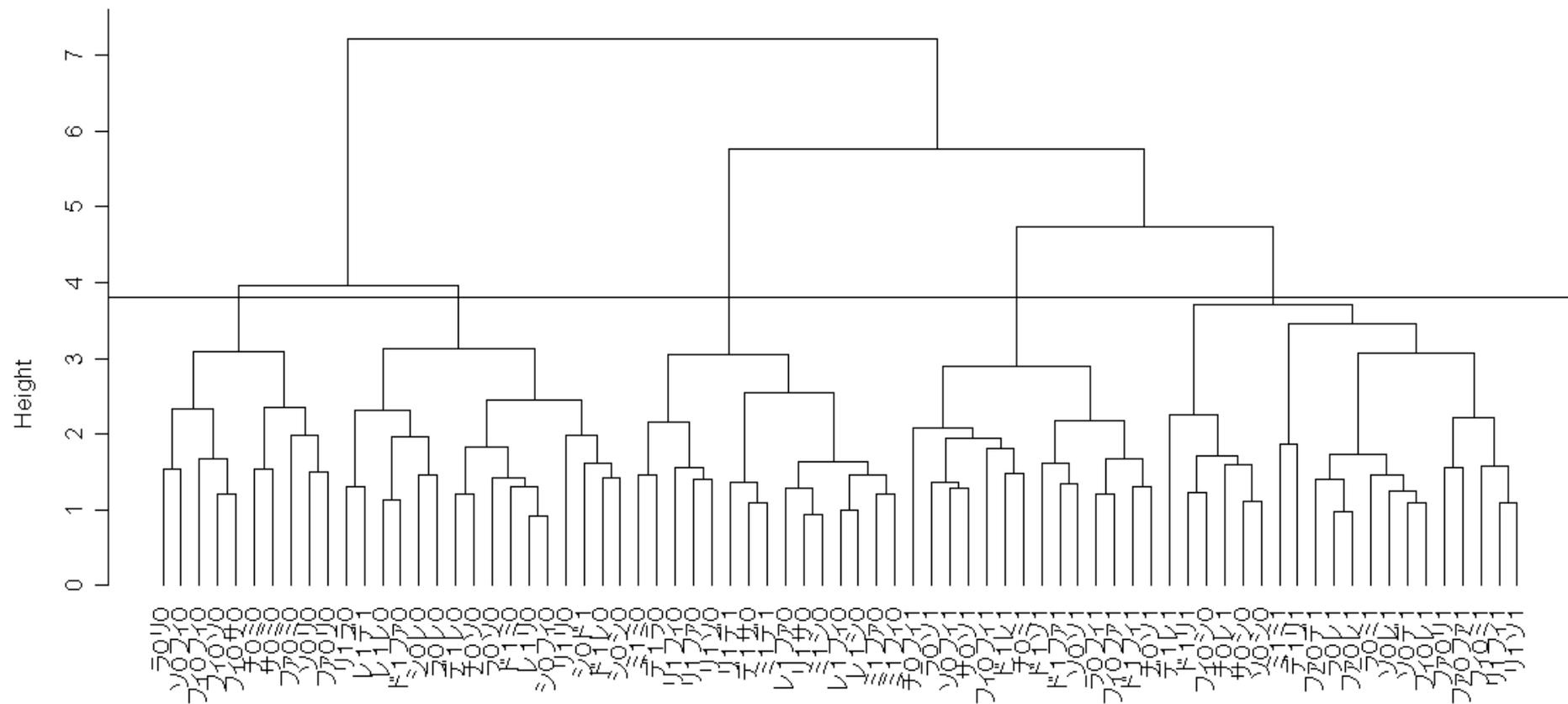
音程	ファ0	ファイ0	ソ0	サ0	ラ0	チ0	シ0	ド1	デ1	レ1	リ1	ミ1
短2度	ミ0	ファ0	ファイ0	ソ0	サ0	ラ0	チ0	シ0	ド1	デ1	レ1	リ1
長2度	リ0	ミ0	ファ0	ファイ0	ソ0	サ0	ラ0	チ0	シ0	ド1	デ1	レ1
短3度	レ0	リ0	ミ0	ファ0	ファイ0	ソ0	サ0	ラ0	チ0	シ0	ド1	デ1
長3度	レ0	リ0	ミ0	ファ0	ファイ0	ソ0	サ0	ラ0	チ0	シ0	ド1	
完全4度		レ0	リ0	ミ0	ファ0	ファイ0	ソ0	サ0	ラ0	チ0	シ0	
減5度			レ0	リ0	ミ0	ファ0	ファイ0	ソ0	サ0	ラ0	チ0	
完全5度				レ0	リ0	ミ0	ファ0	ファイ0	ソ0	サ0	ラ0	
短6度					レ0	リ0	ミ0	ファ0	ファイ0	ソ0	サ0	
長6度						レ0	リ0	ミ0	ファ0	ファイ0	ソ0	
短7度							レ0	リ0	ミ0	ファ0	ファイ0	
長7度								レ0	リ0	ミ0	ファ0	
完全8度									レ0	リ0	ミ0	

…作成した音声

# 5種類の音声の選択 (1 / 3)



Cluster Dendrogram



dist(da)  
hclust (\*, "complete")

# 5種類の音声の選択(2/3)

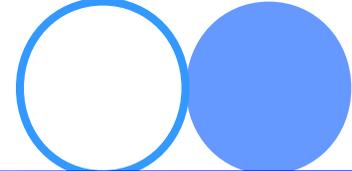
- クラスタ分析の結果を、MANOVA(多変量分散分析)を適用し検定すると有意差が見られた

Df	Wilks	approx F	num Df	den Df	Pr(>F)
4.0	0.28	17.08	96.0	4348.3	***

***	Pr < 0.001	*	Pr < 0.05
**	Pr < 0.01	.	Pr < 0.1

- クラスタ内で各印象語の平均値を求める。  
その平均値を基に、各クラスタ内の各音声とのマンハッタン距離を求め、もっとも小さい音声を実験で用いる

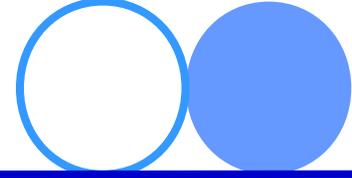




- 音声の1モーラの継続長として、「0.1秒」、「0.13秒」、「0.17秒」、「0.23秒」、「0.3秒」、「0.4秒」の6段階を使用
- モーラ継続長の段階は、前回の実験の長さ「0.2秒」を基準に、1番長いのを2倍、1番短いのを1/2倍になるようにした。段階差が等比になるように以下の計算式より求めた。

$$1/2 \times X^5 = 2 \rightarrow X \doteq 1.3$$

- 最終的に実験に用いる音声は、 $5 \times 6 \times 6 = 180$ 種類とする

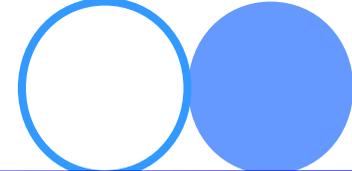


- 被験者30名の実験結果に対して、印象語の5段階での評価の平均をとった値を求める
- 平均した評価を基に回帰係数を求め、目視及び回帰係数の値から、secの増減によって各印象語が受け取りやすくなるのか、受け取りにくくなるのかの関係について調べる
  - \* なお、secとは、「ええ」の全長のことであり、前半1モーラと後半1モーラのモーラ継続長（秒）の合計と定義する

# 回帰係数

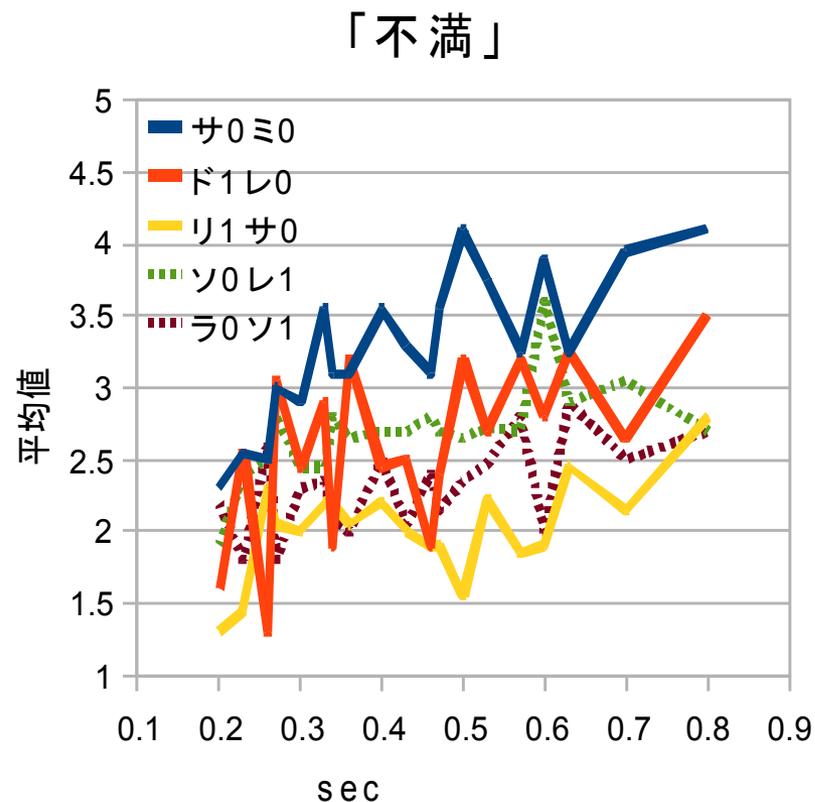
音名	サ0ミ0	ド1レ0	リ1サ0	ソ0レ1	ラ0ソ1
悲しみ	0.09	1.09	0.54	0.06	0.33
無関心	-0.57	-0.23	-0.51	1.05	0.4
不満	2.54	2.03	1.06	1.21	1.07
信頼	-0.38	-1.14	-0.75	-0.56	0.15
驚き	0.22	0.2	0.84	-2.64	-0.59
恐れ	0.15	0.71	-0.09	-0.53	0.67
落胆	1.82	1.59	0.76	1.12	0.45
感心	-0.63	-0.6	0.13	-0.02	-0.72
安堵	-0.75	-0.66	-0.75	0.08	-0.07
慎み	-1.56	-1.81	-0.9	-0.49	-0.26
怒り	0.83	0.48	0.38	-0.09	0.34
好意	-0.4	-0.64	-0.64	0.07	-0.23
中立	-0.81	-0.92	-1.01	-0.46	-0.03
高揚	0.07	0.54	-0.41	-0.11	-0.08
満足	-0.74	-1.42	-0.53	-0.06	0.27
疑問	-0.12	0.95	0.73	-1.46	-1.03
冷静	-2.84	-2.17	-1.48	-0.43	0.03
同情	-0.14	-0.08	-0.24	0.25	0.34
喜び	-0.06	-0.19	-0.25	0.07	-0.26
軽蔑	1.22	1.23	0.19	0.25	0.61
嫌悪	1.49	1.98	0.6	0.68	0.48
不安	0.16	1.5	0.74	-0.21	-0.27
強調	-0.07	0.08	-0.71	-0.37	1.02
肯定	-2.92	-3.27	-1.7	0.25	-0.14
苛立ち	0.73	0.91	1.07	0.22	1.13
相槌	-2.91	-2.86	-1.78	0.53	0.24
自信有	-1.94	-2.33	-0.89	-0.25	0.22
自信無	0.75	1.61	0.29	-0.35	-0.19

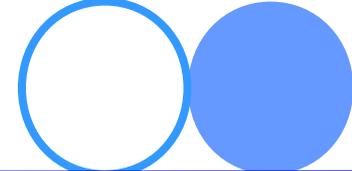
- 回帰係数が、0.3以上を傾きが正、-0.3以下を傾きが負、その他は0と定義した。それを基に決めた分類グループの定義は以下のとおりである
  - 正の傾向の条件・・半分以上が正の値かつその他は0の値
  - 負の傾向の条件・・半分以上が負の値かつその他は0の値
  - 0に近い値の傾向・・正の傾向の条件、負の傾向の条件を満たさないかつ半分以上が0の値
- \* 半分以上とは、「サ0≒0」、「ド1≒0」、「リ1サ0」の中から2つ以上または「ソ0≒1」、「ラ0ソ1」の中から1つ以上のこと



## ① 悲しみ、不満、落胆、怒り、軽蔑、嫌悪、苛立ち

- F0の高さの変化に関わらず、多くの回帰係数において正の値の傾向
- 例として、「不満」の図を示す

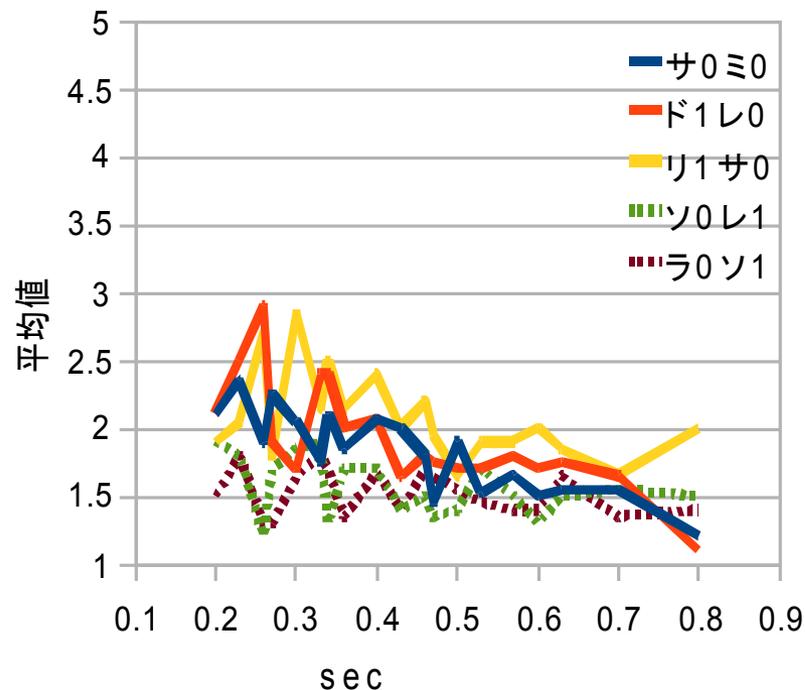


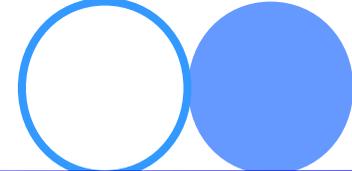


## ② 信頼、感心、慎み、中立、冷静

- F0の高さの変化に関わらず、多くの回帰係数において負の値の傾向
- 例として、「慎み」の図を示す

「慎み」

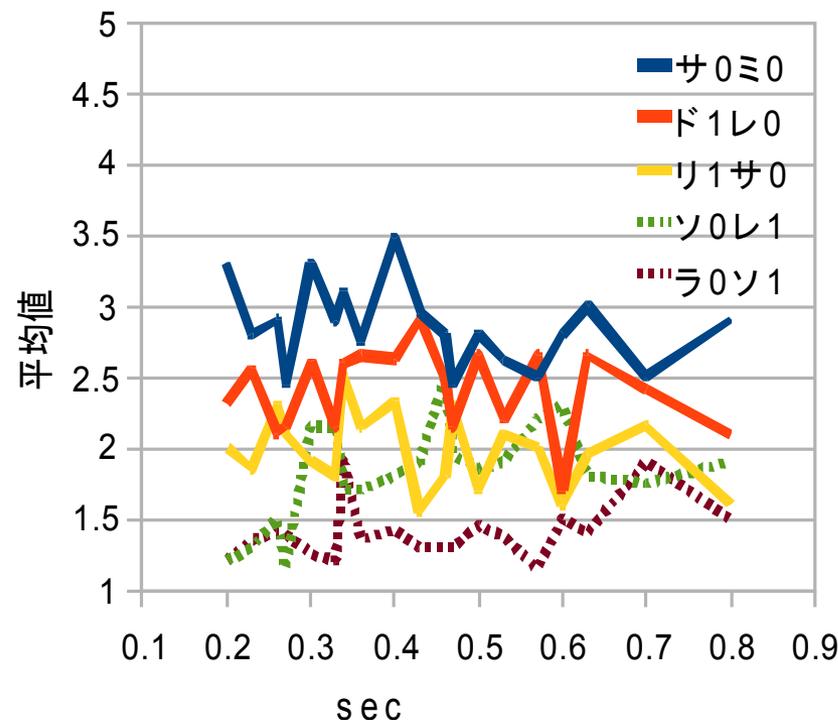


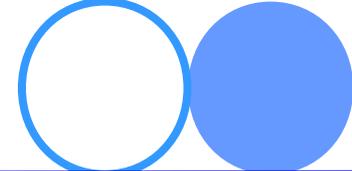


## ③ 無関心、相槌

- F0の高さの変化が正の音声 (点線)は、回帰係数が正の値になる傾向
- F0の高さの変化が負の音声 (実線)は、回帰係数が負の値になる傾向
- 例として、「無関心」の図を示す

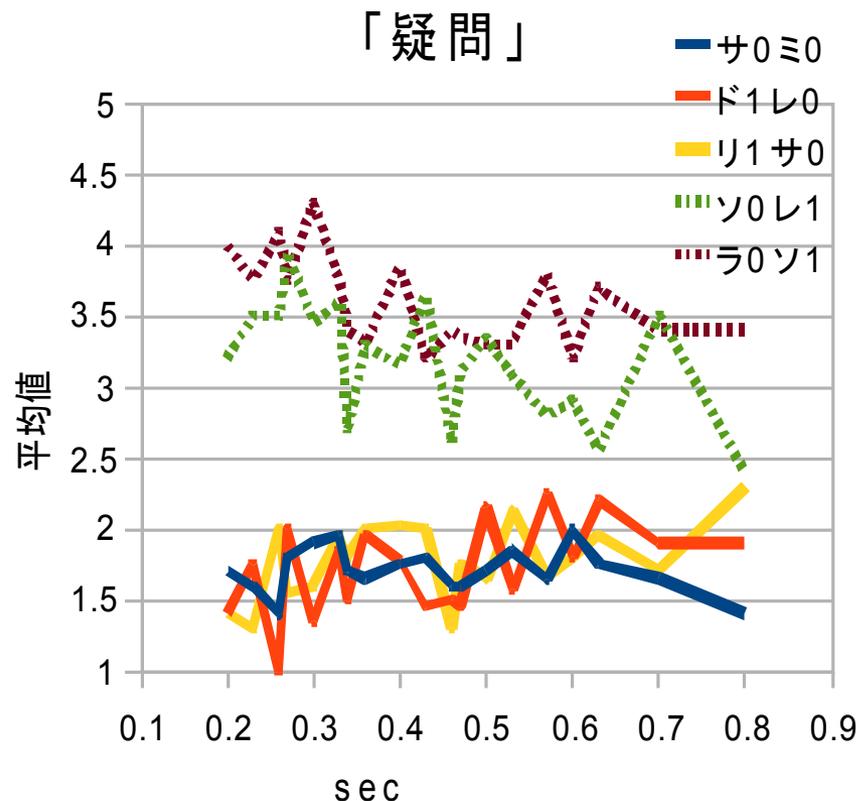
「無関心」

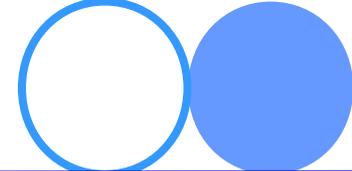




## ④ 疑問、自信なし

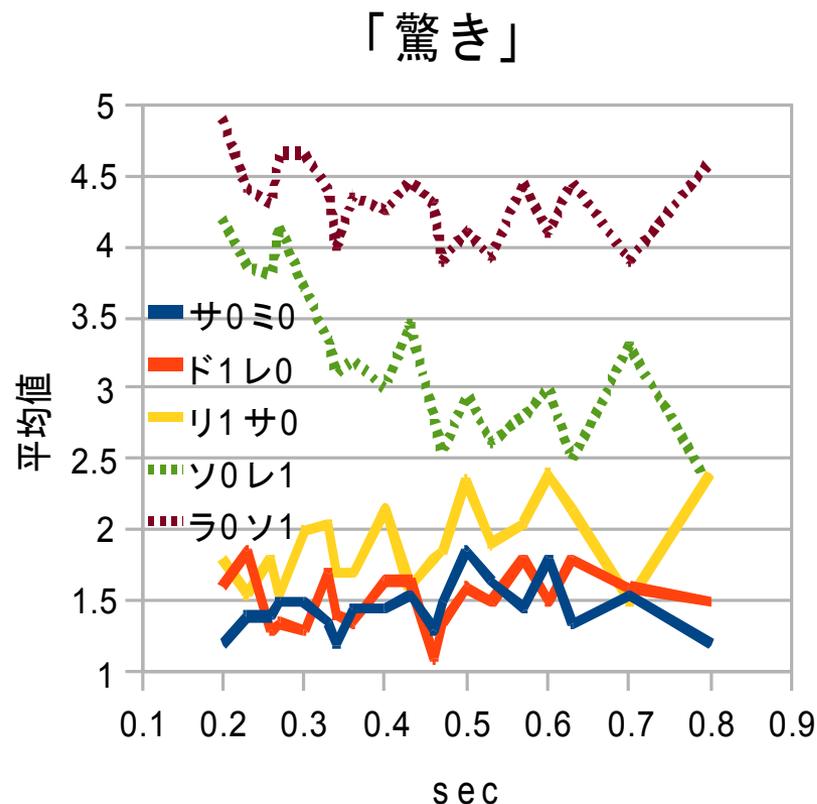
- F0の高さの変化が正の音声 (点線) は、回帰係数が負の値になる傾向
- F0の高さの変化が負の音声 (実線) は、回帰係数が正の値になる傾向
- 例として、「疑問」の図を示す

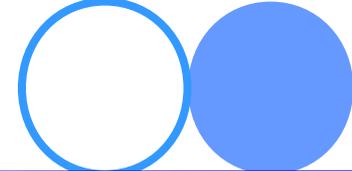




## ⑥ 驚き

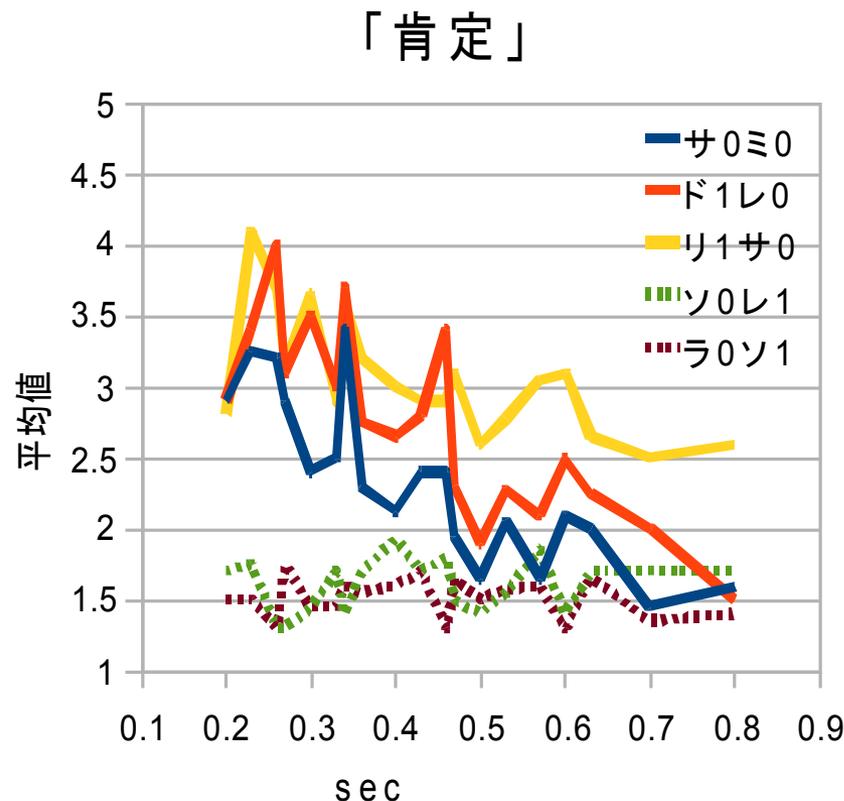
- F0の高さの変化が正の音声 (点線) は、回帰係数が負の値になる傾向
- F0の高さの変化が負の音声 (実線) は、回帰係数が0に近い値になる傾向
- 例として、「驚き」の図を示す

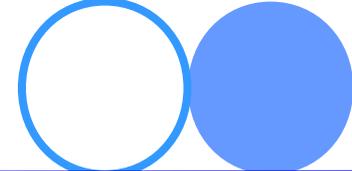




## ⑧ 安堵、好意、満足、肯定、自信あり

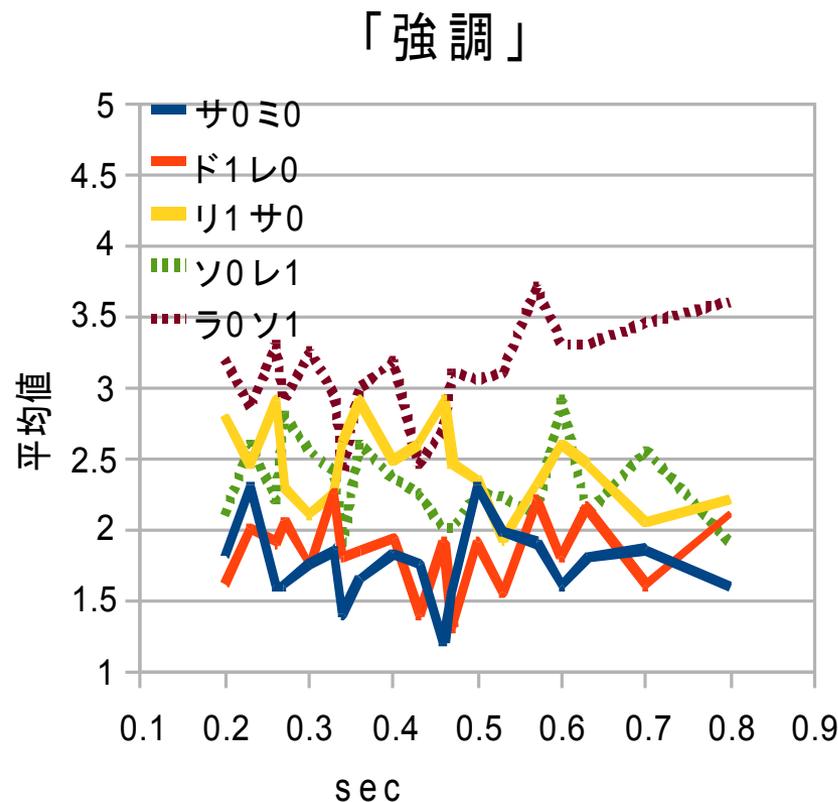
- F0の高さの変化が正の音声 (点線) は、回帰係数が0に近い値になる傾向
- F0の高さの変化が負の音声 (実線) は、回帰係数が負の値になる傾向
- 例として、「肯定」の図を示す

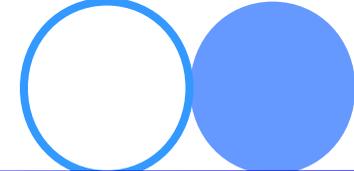




## ⑩ 恐れ、高揚、強調

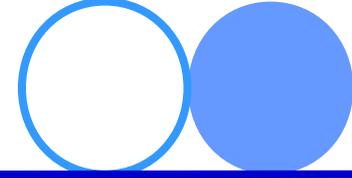
- F0の高さの変化と回帰係数の関係に特徴的な点は見られなかった
- 例として、「強調」の図を示す



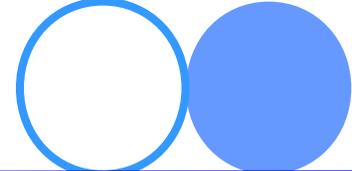


## 印象語の定性的分類

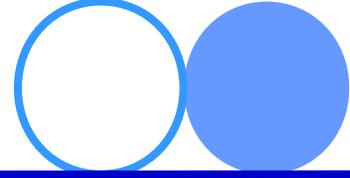
	印象語	F0の高さの変化 (正)の音声の 回帰係数	F0の高さの変化 (負)の音声の 回帰係数
1	「悲しみ」, 「不満」, 「落胆」, 「怒り」, 「軽蔑」, 「嫌悪」, 「苛立ち」	多くの回帰係数が正の値であった。 負の値の場合でも0に近い値である。	
2	「信頼」, 「感心」, 「憤り」, 「中立」, 「冷静」	多くの回帰係数が負の値であった。 正の値の場合でも0に近い値である。	
3	「無関心」, 「相槌」	正の値の傾向	負の値の傾向
4	「疑問」, 「自信なし」	負の値の傾向	正の値の傾向
5	「同情」	正の値の傾向	0に近い値の傾向
6	「驚き」	負の値の傾向	0に近い値の傾向
7	「不安」	0に近い値の傾向	正の値の傾向
8	「安堵」, 「好意」, 「満足」, 「肯定」, 「自信あり」	0に近い値の傾向	負の値の傾向
9	「喜び」	0に近い値の傾向	
10	「恐れ」, 「高揚」, 「強調」	—	—



- 他の語や声の大きさなどを考慮して、パラ言語情報の伝達・語による印象の傾向や程度の違いについて調べていきたい

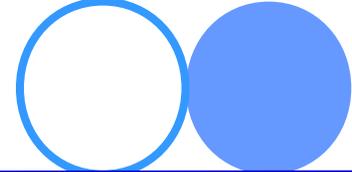


- フォルマントは声道共振周波数  
— 声帯音源が声道（声帯～唇）で共鳴することにより  
特定帯域ごとに倍音が増幅される。この増幅された  
成分の塊、ピークをフォルマントと呼ぶ
- 母音はこのフォルマント周波数の低いほうから2つ「第1フォルマント」、「第2フォルマント」  
で聞き分けることができる



- モーラ……言語における音の単位の一つ。  
拍とも呼ぶ
  - 母音
  - 子音 + 母音、子音 + 子音 + 母音 (拗音「きゃきゅきょ」など)
  - 長音「ー」
  - 促音「っ」
  - 撥音「ん」 を1モーラとして数える
- 短歌や俳句はモーラを単位として、五・七・五……のように計られる
- 「サッカー」は「サッ」「カー」で2音節だが、モーラで考えると「サ」「ッ」「カ」「ー」で4モーラ

# 音程について



- 音程・・・ふたつの音の高さの隔たり
- シャープやフラットのつかない音を幹音といい、この幹音間の音程は、完全系の度数と長短系の度数に分かれる。1, 4, 5, 8度が完全系の度数で、2, 3, 6, 7度が長短系の音程。長短系は半音の数で決定される

☆ドレ (長2度)

ド レ ミ ファ ソ ラ シ ド

☆ミファ (短2度)

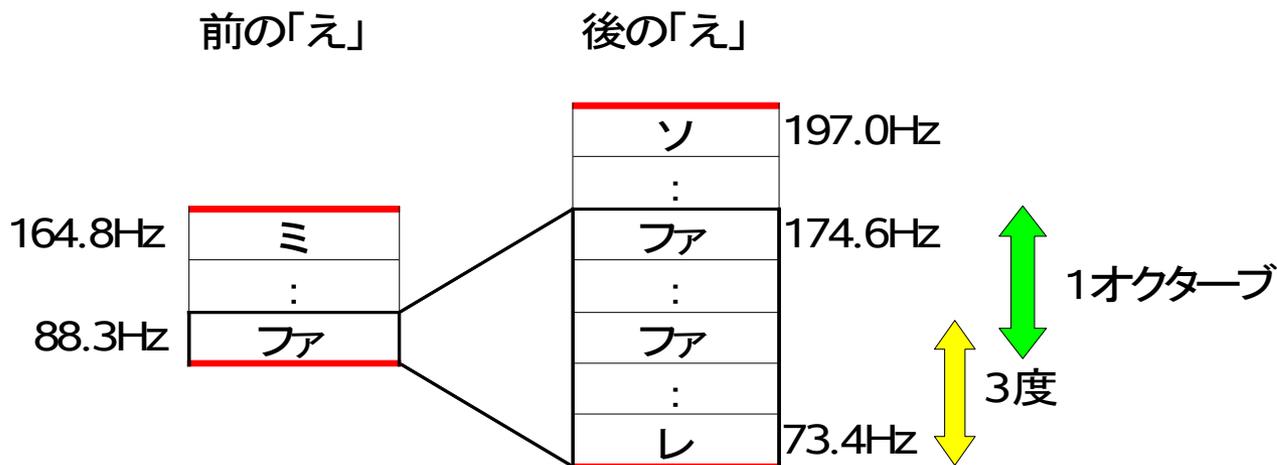


全音

半音

# 実験で用いる音声(補足)

- 先行研究では、前半の「え」は120Hzを中心に1オクターブ幅、後半の「え」は1オクターブ半幅の範囲で作成し、後半の「え」は前半の「え」に対して、1オクターブ幅の範囲内で変化させた

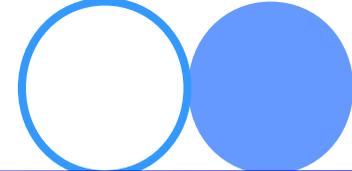


# 「ええ」について

## 10対話1052文中の間投詞の出現頻度

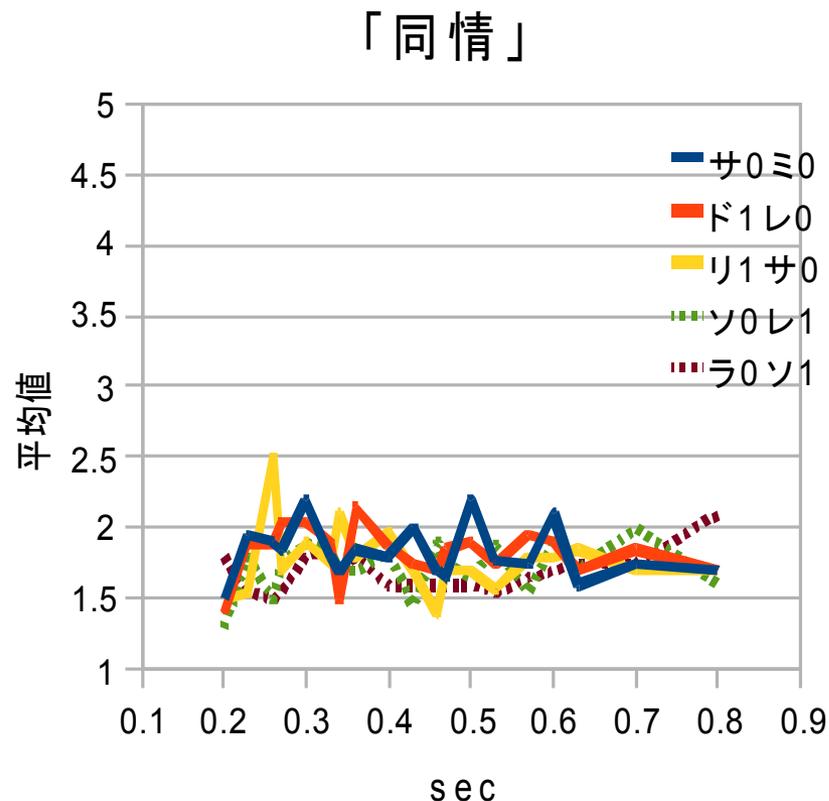
間投詞	出現数	出現頻度	累積
え	238	22.8%	22.8%
えと	230	22.0%	44.8%
あの	207	19.8%	64.6%
あ	162	15.5%	80.1%
ま	130	12.4%	92.5%
えとですね	33	3.2%	95.7%
ん	9	0.9%	96.6%
そですね	9	0.9%	97.5%
その	7	0.7%	98.2%
その他	7	0.7%	100.0%

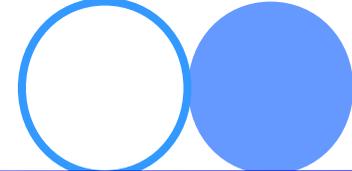
「え」「えと」「あの」「あ」「ま」の5種類で全体の92%をカバーしている



## ⑤ 同情

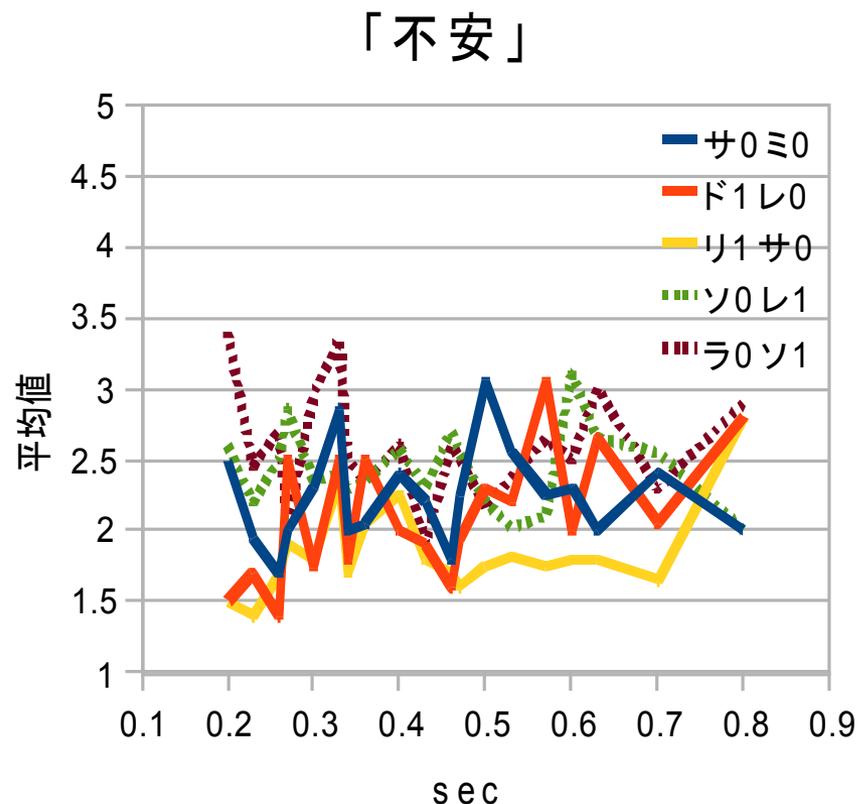
- F0の高さの変化が正の音声（点線）は、回帰係数が正の値になる傾向
- F0の高さの変化が負の音声（実線）は、回帰係数が0に近い値になる傾向
- 例として、「同情」の図を示す

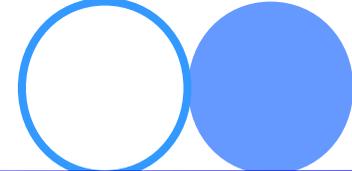




## ⑦ 不安

- F0の高さの変化が正の音声（点線）は、回帰係数が0に近い値になる傾向
- F0の高さの変化が負の音声（実線）は、回帰係数が正の値になる傾向
- 例として、「不安」の図を示す





## ⑨ 喜び

- F0の高さの変化に関わらず、多くの回帰係数において0に近い値になる傾向
- 例として、「喜び」の図を示す

