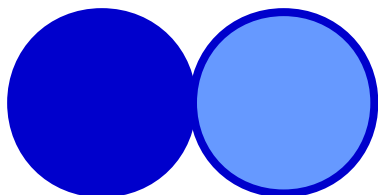
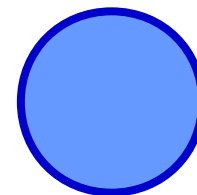


2-P-39

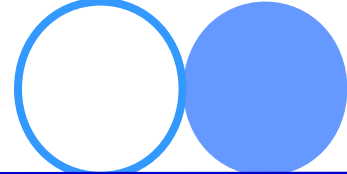
声の高さ・長さ・大きさの パラ言語情報の認知に与える影響



△竹谷雄太(島根大学・総合理工)

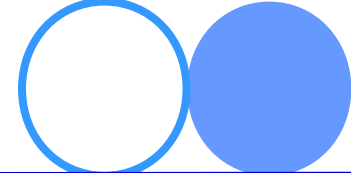
○小林聡(島根大学・総情セ)

2011年3月10日



- 声の高さ・長さ・大き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

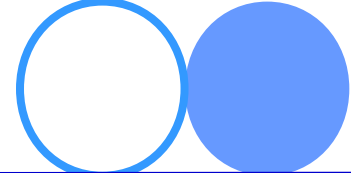
(制御が困難)

(対話中の動的変化が困難)

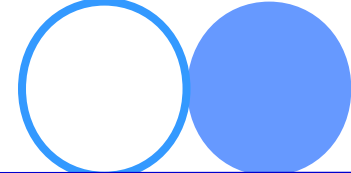
個人的特徴、
身体的特徴、
体調

≡ 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」)
- 竹川 “声の高さ及び長さによるパラ言語情報の認知への影響” 島根大学卒業論文(2010)
→「ええ」の声の高さと長さの変化から人間が受け取る印象
– Onlineで閲覧可能

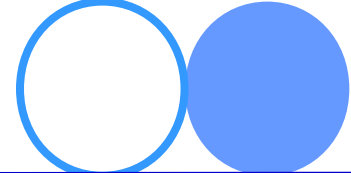


- 基準音声×3＋評価対象音声×1を1セットとし、1人の被験者に50セットを聞いてもらう。実験時間は1人あたり2時間程度。
- 基準音声を基準に評価対象音声を聞いた際に受けた印象を28個の印象語について各々1(受け取れない)から5(はっきり受け取れる)の5段階で評価。
- 被験者数は25名(大学生)。
 - 評価対象音声・・・「え」の高さ・長さ・大きさを変えた「ええ」
(同じ高さ・長さ・大きさの場合もある)
 - 基準音声・・・前後の高さは一般男性の基本周波数120Hzの「ええ」
前後の長さは1モーラ0.2秒、声の大きさは58.5dB



- 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

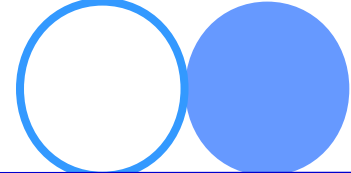
- 音声はフォルマント合成方式を用いたklattの合声器で作成した合成音声を使用

○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

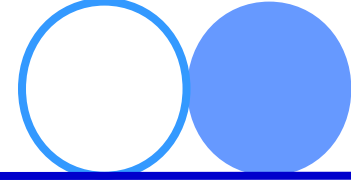
$$F1 = -0.565 \times F0 + 580.2$$

$$F2 = 0.732 \times F0 + 1712.56$$

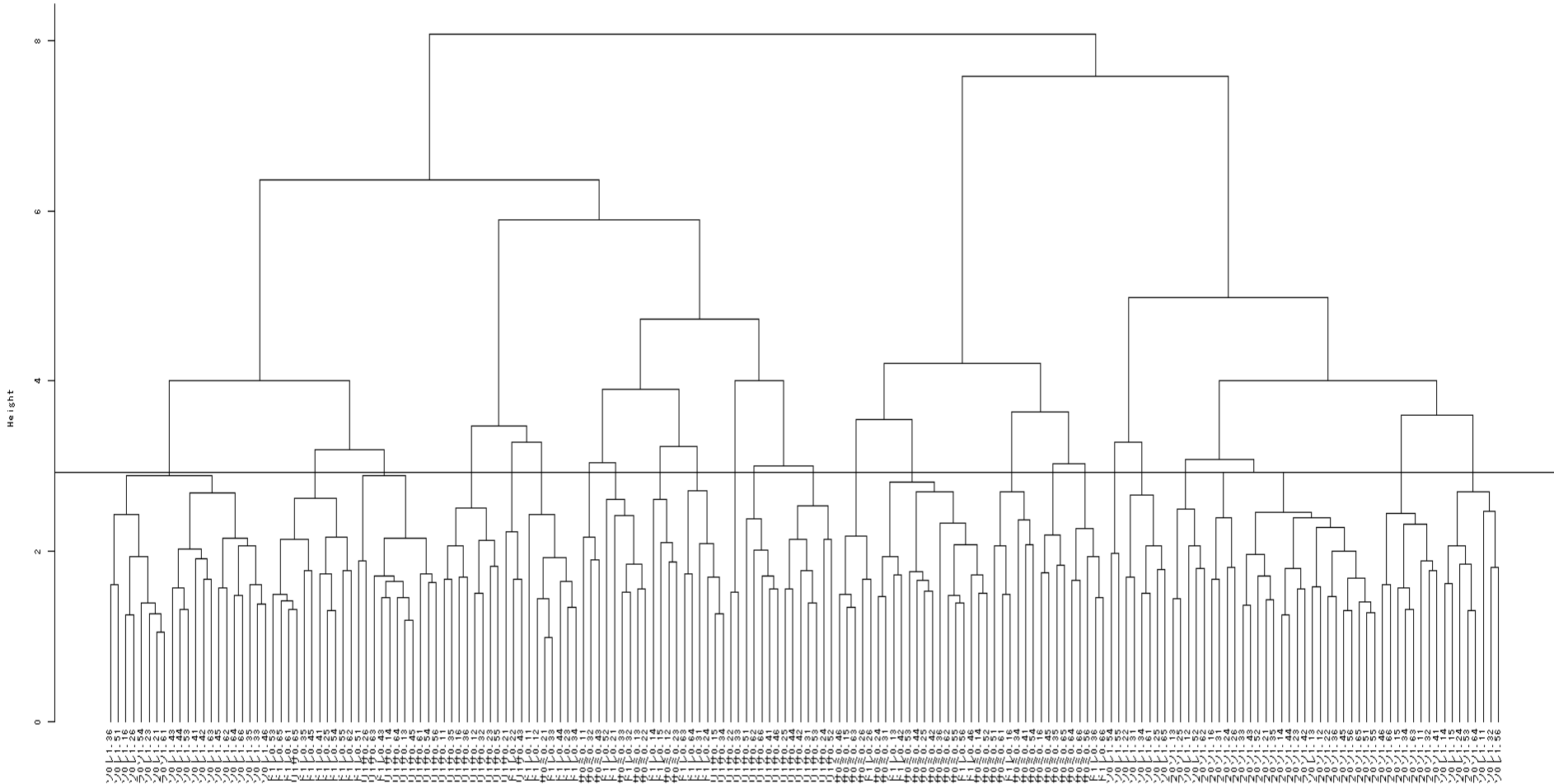


- 先行研究では、声の高さ5パターン×音の長さ36パターンを組み合わせ、180種類の音声を使用していた。
 - 音の高さ
1オクターブ半幅で作成できるすべての音の中から5パターン。
 - 音の長さ
モーラ継続長0.1sec、0.13sec、0.17sec、0.23sec、0.3sec、0.4secを組み合わせ、36パターン。
- 今回はこの中から25パターンを選択し、大きさを変化させて被験者実験を行う。

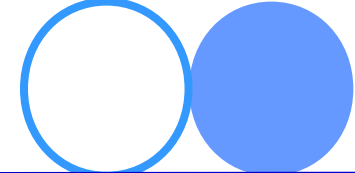
実験で用いる音声(3/6)



Cluster Dendrogram



dist(data)
hclust(%, "complete")

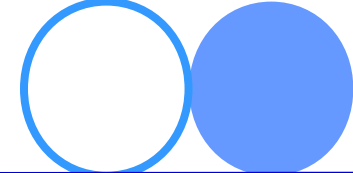


クラスター分析の結果を、MANOVA(多変量分散分析)を適用し検定すると有意差が見られた。

Df	Wilks	approx F	num Df	den Df	Pr(>F)
24	0.12	5.75	672	32023	***

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

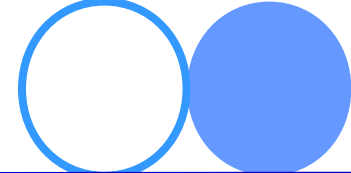
クラスター内で各印象語の平均値を求める。その平均値を基に、各クラスター内の各音声とのユークリッド距離を求め、最も小さい音声を実験で用いる。



- 今回使用する25パターンの音声

(1:「0.1sec」,2:「0.13sec」,3:「0.17sec」,4:「0.23sec」,5:「0.3sec」,6:「0.4sec」,)

音の高さ	音の長さの 組み合わせ
サ0ミ0ー	「12」,「14」,「32」,「45」, 「61」,「63」
ド1レ0ー	「24」,「32」,「33」,「65」, 「66」
リ1サ0ー	「22」,「31」,「36」,「43」, 「45」,「66」
ラ0ソ1+	「16」,「32」,「55」,「64」
ソ0レ1+	「34」,「42」,「52」,「54」



音名とF0周波数との対応表

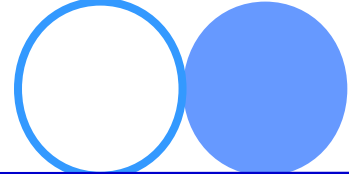
前半1モーラの範囲

音名	F0[Hz]
ミ1(E3)	164.8
リ1(D3#)	155.6
レ1(D3)	146.8
デ1(C3#)	138.6
ド1(C3)	130.8
シ0(B2)	123.5
チ0(A2#)	116.5
ラ0(A2)	110.0
サ0(G2#)	103.8
ソ0(G2)	97.9
フィ0(F2#)	92.5
ファ0(F2)	87.3

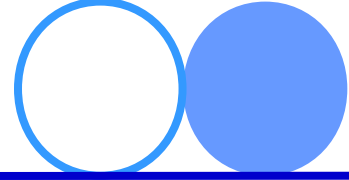
後半1モーラの範囲

音名	F0[Hz]		音名	F0[Hz]
ソ1(G3)	196.0		ソ0(G2)	97.9
フィ1(F3#)	185.0		フィ0(F2#)	92.5
ファ1(F3)	174.6		ファ0(F2)	87.3
ミ1(E3)	164.8		ミ0(E2)	82.4
リ1(D3#)	155.6		リ0(D2#)	77.8
レ1(D3)	146.8		レ0(D2)	73.4
デ1(C3#)	138.6			
ド1(C3)	130.8			
シ0(B2)	123.5			
チ0(A2#)	116.5			
ラ0(A2)	110.0			
サ0(G2#)	103.8			

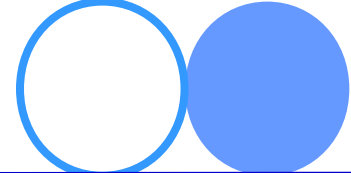
- (理科年表より抜粋) *)同じ音名が存在するため、便宜的に数字を付けている



- 今回は声の大きさとして、52.5dB、55.5dB、58.5dB、61.5dB、64.5dBの5段階を使用した。
- 最終的に実験に用いる音声は、 $25(\text{高さと長さ}) \times 5(\text{大きさ}) = 125$ 種類となる。



- 80デシベル:
地下鉄の車内・電車の車内・ピアノ(正面1m)
- 70デシベル:
ステレオ(正面1m、夜間)・騒々しい事務所の中、騒々しい街頭
- 60デシベル:
静かな乗用車・普通の会話
- 50デシベル:
静かな事務所・クーラー(屋外機、始動時)
- 40デシベル:
市内の深夜・図書館・静かな住宅の昼

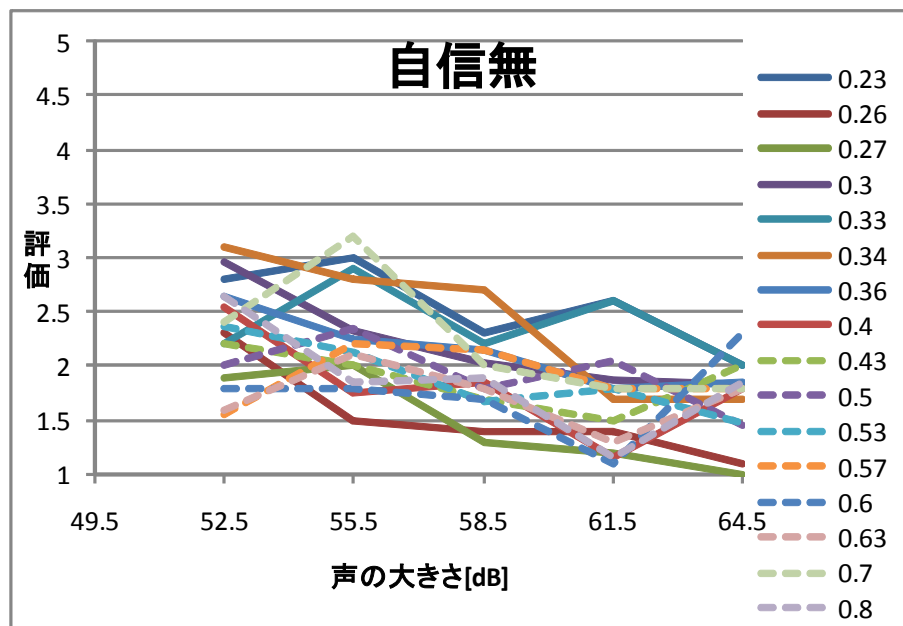
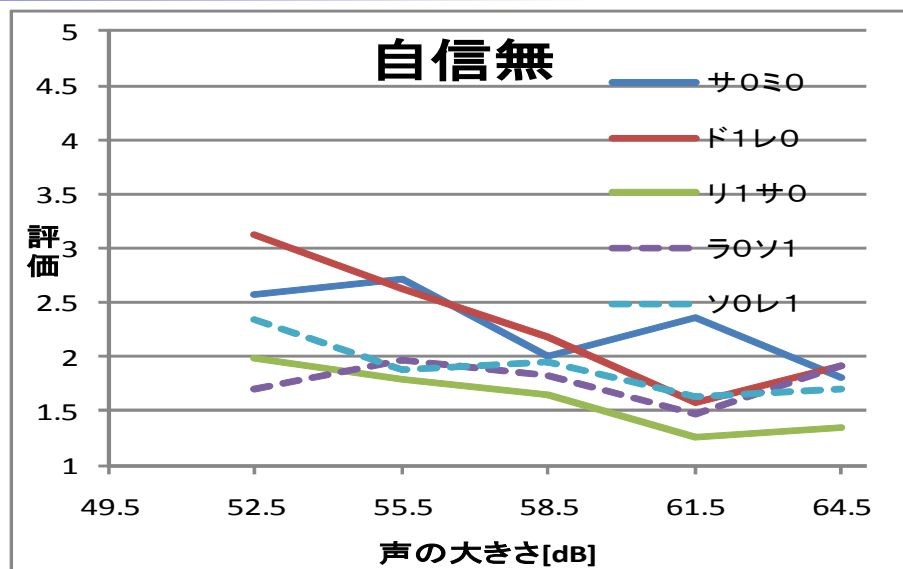


- 被験者25名の実験結果に対して、印象語の5段階での評価の平均をとった値を求める。
- 平均した評価を基に回帰係数を求め、グラフの目視及び回帰係数の値から、声の高さ・長さ・大きさによって各印象語が受け取りやすくなるのか、受け取りにくくなるのかの関係について調べる。

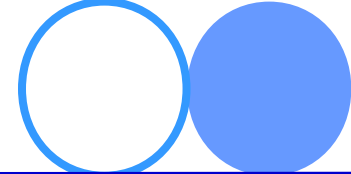
実験結果(1/8)

① 悲しみ、不安、自信無

- F0・音の長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
- グラフから見ても声が小さいほど評価が高くなる傾向
- 例として、「自信無」の図を示す

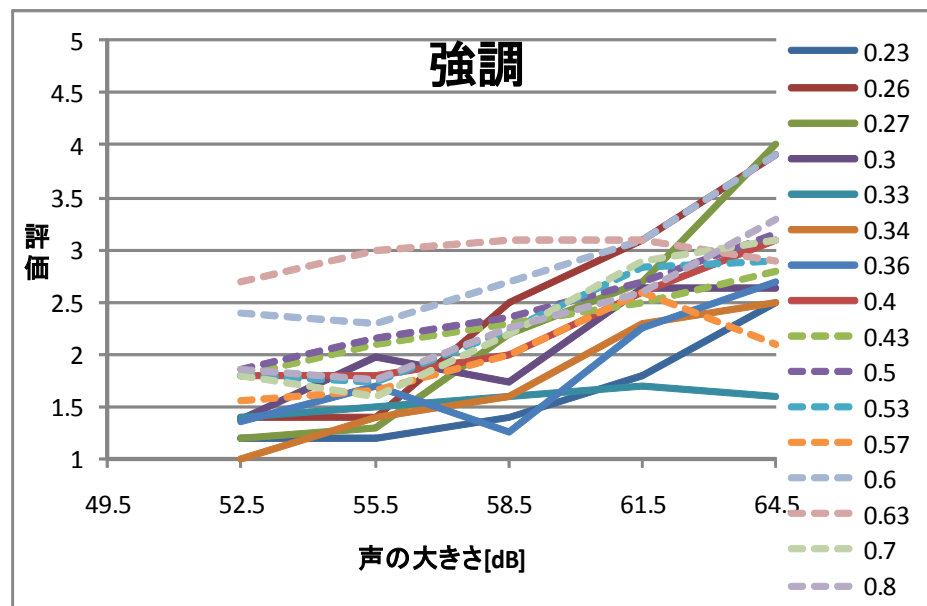
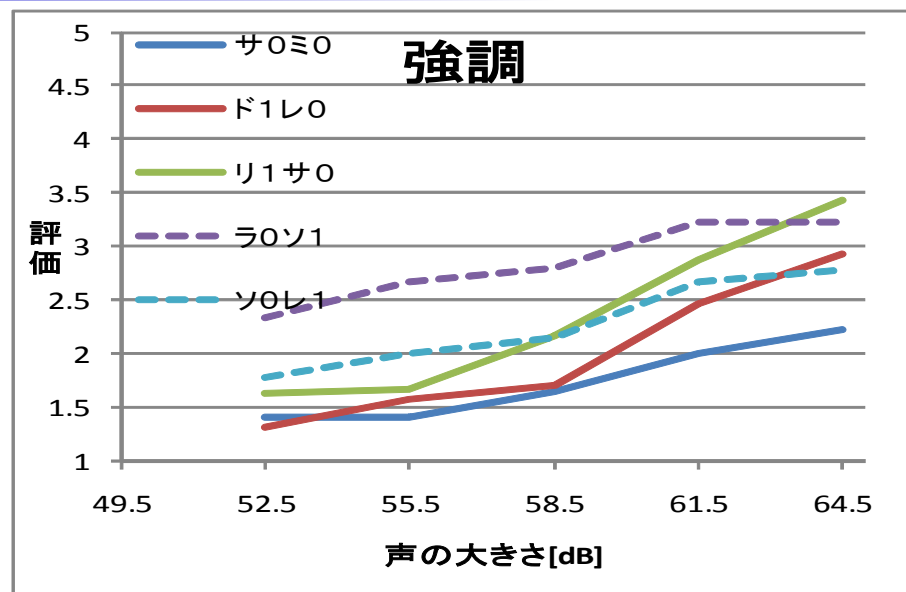


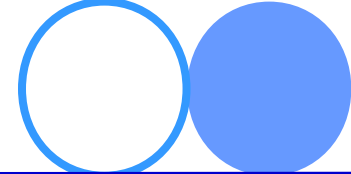
実験結果(2/8)



② 高揚、強調、自信有

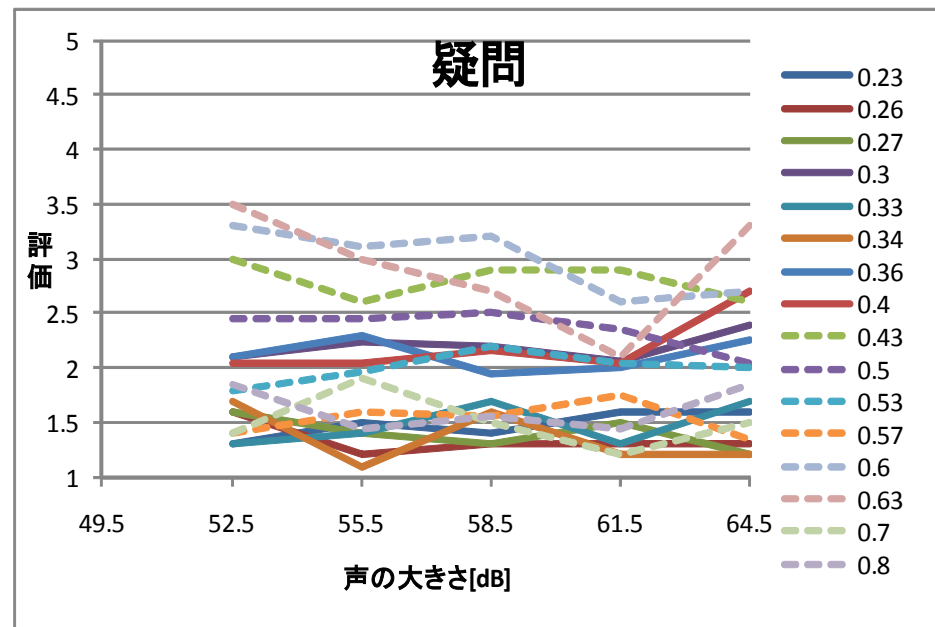
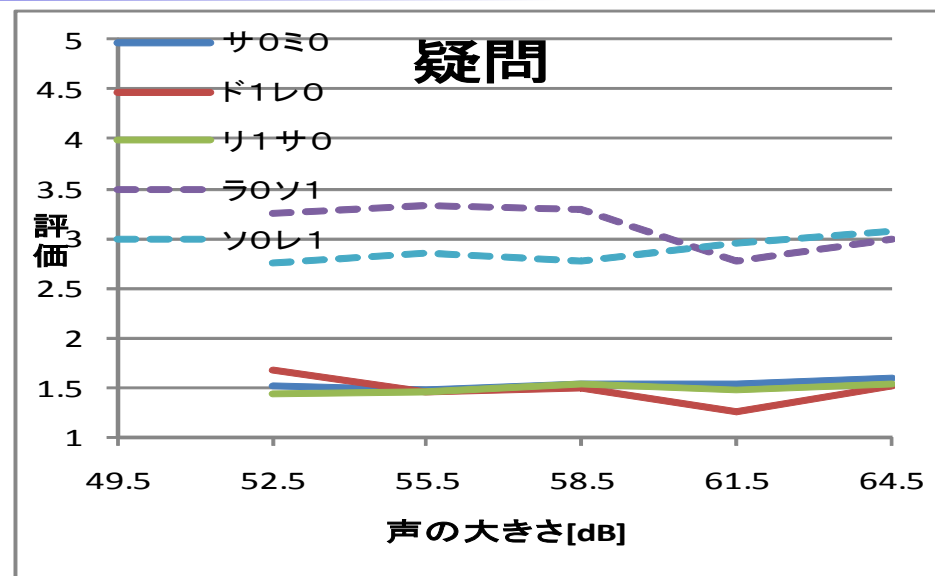
- F0・音の長さの変化に関わらず、多くの回帰係数において正の値の傾向
- グラフから見ても声が大きいほど評価が高くなる傾向
- 例として、「強調」の図を示す

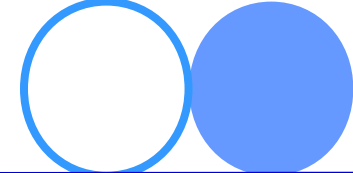




③ 驚き、疑問

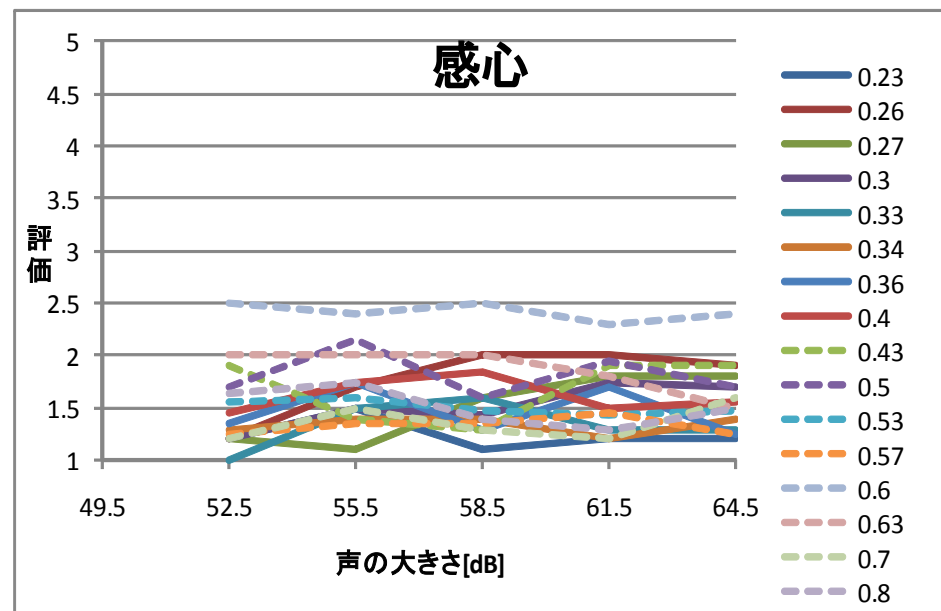
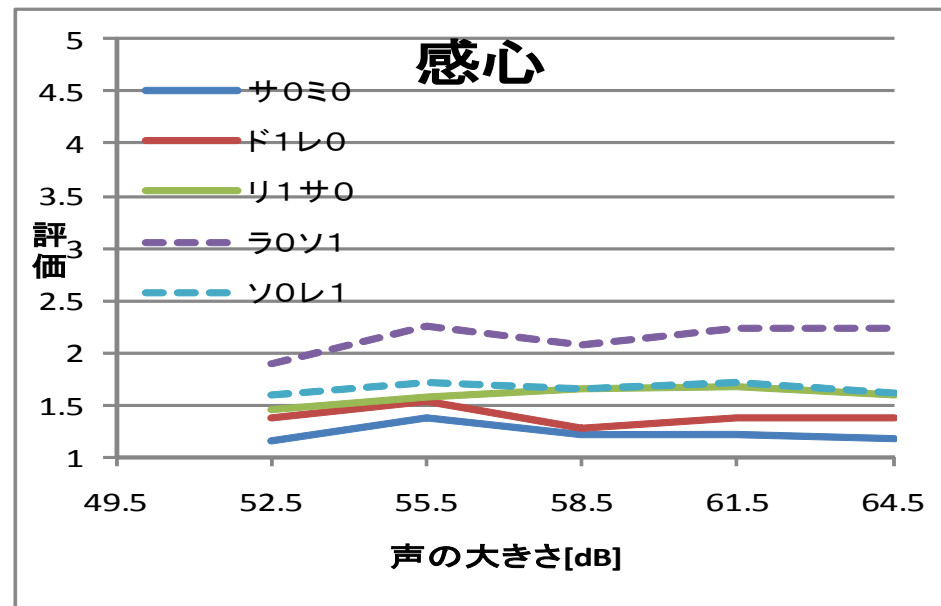
- 回帰係数は0に近い値になる傾向
- グラフから見ると、音が上がる音声と継続長が長い音声は評価が高くなる傾向
- 例として、「疑問」の図を示す

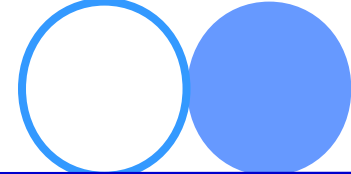




④ 感心

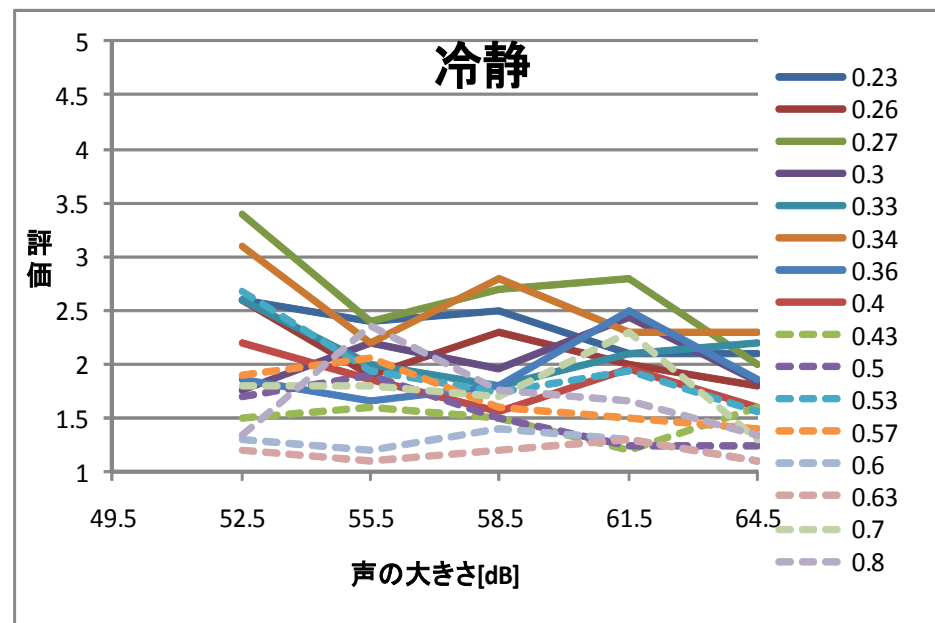
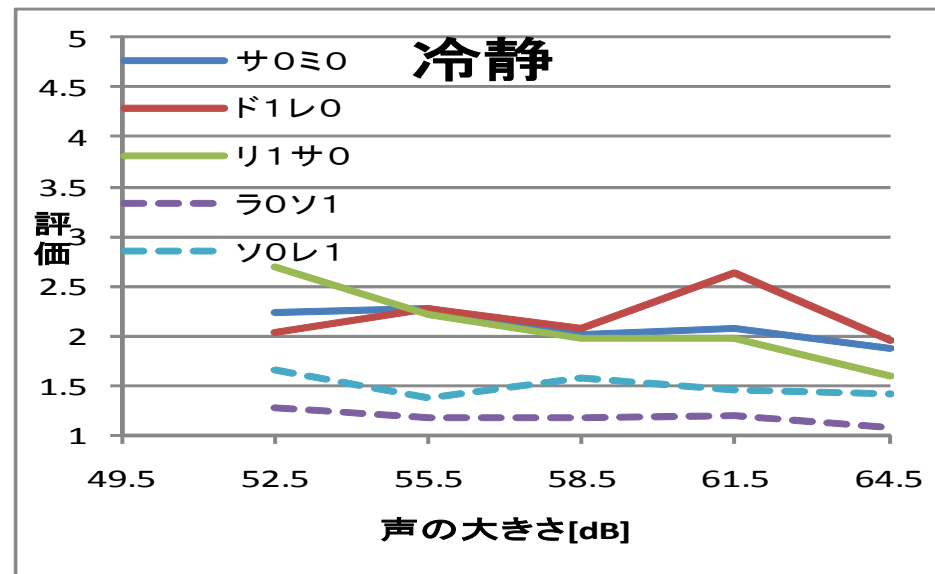
- 回帰係数は0に近い値になる傾向
- グラフから見ると、音が上がる音声は評価が高くなる傾向
- 例として、「感心」の図を示す

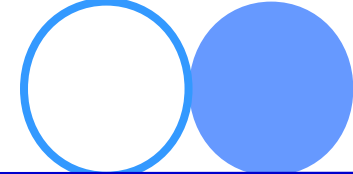




⑤ 冷静

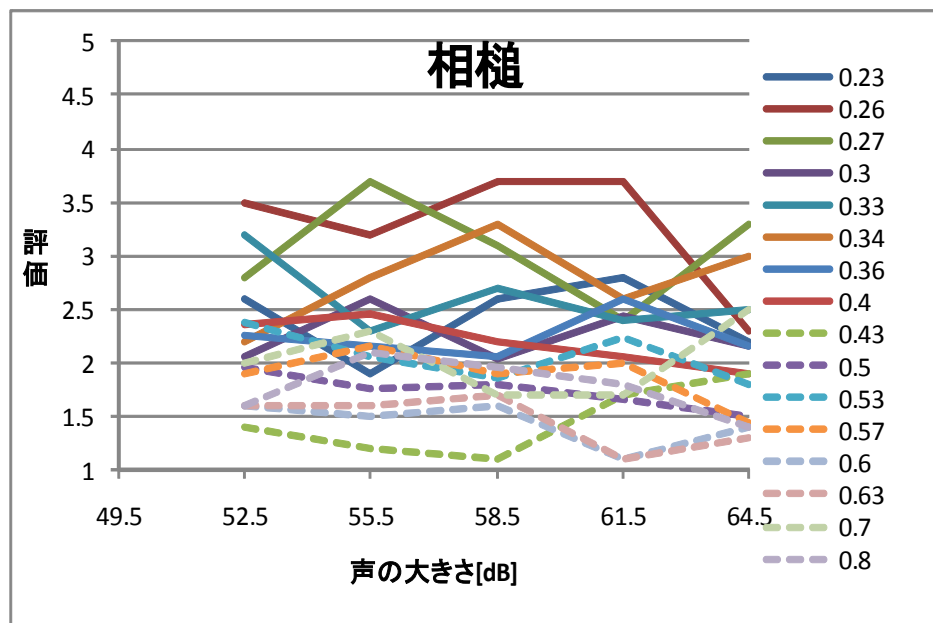
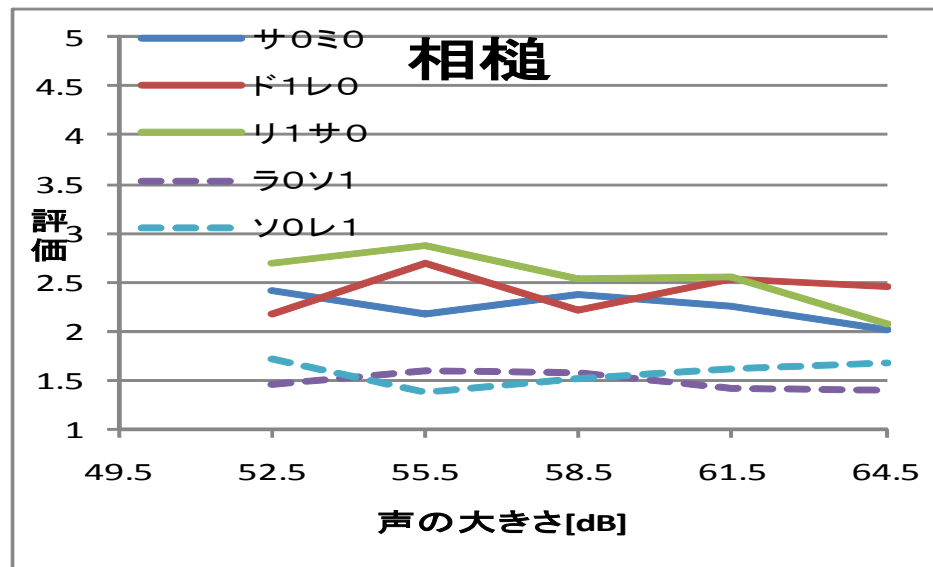
- F0の変化が負の時、回帰係数が負の値になる傾向
(音が下がる音声の時、小さい声の影響を受けやすい)
- グラフから見ると、音が下がる音声と継続長が短い音声は評価が高くなる傾向
- 例として、「冷静」の図を示す



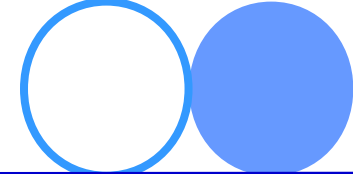


⑥ 肯定、相槌

- 回帰係数は0に近い値になる傾向
- グラフから見ると、音が下がる音声と継続長が短い音声は評価が高くなる傾向
- 例として、「相槌」の図を示す

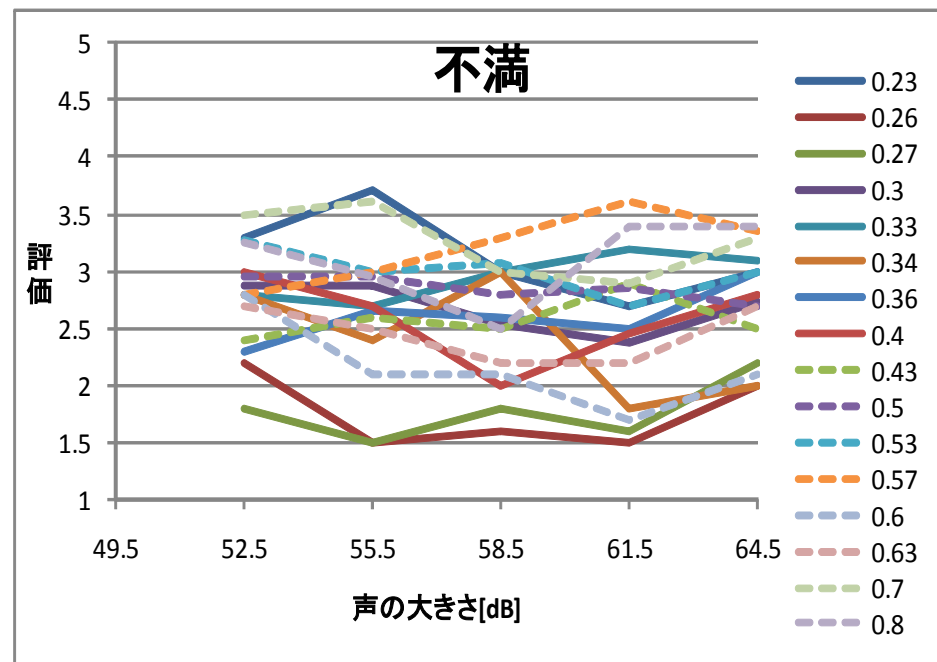
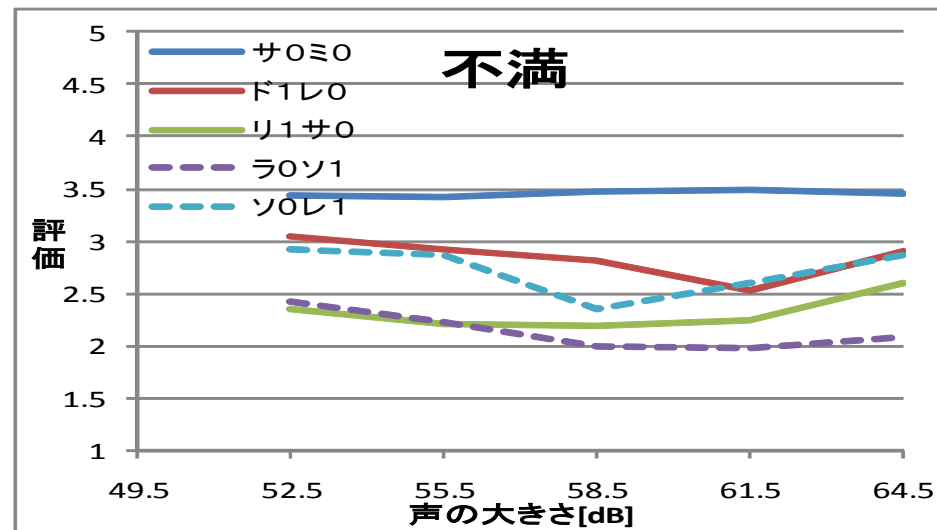


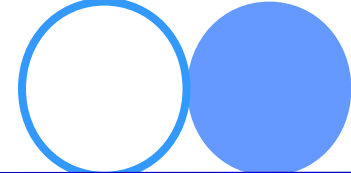
実験結果(7/8)



⑧ 無関心、嫌悪、 不満、落胆、苛立ち

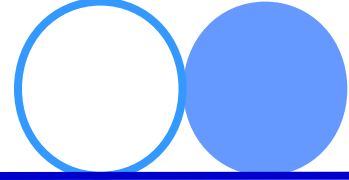
- 回帰係数での変化は小さく、
グラフでの評価は中程度
- 例として、「不満」の図を示す



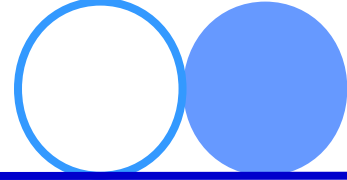
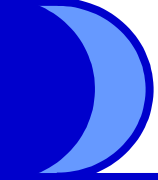


・印象語の定性的分類

	印象語	F0について	継続長について	声の大きさについて
1	悲しみ、不安、自信無	声の大きさが小さいほど評価が受けやすくなる傾向		
2	高揚、強調、自信有	声の大きさが大きいほど評価が受けやすくなる傾向		
3	驚き、疑問	F0の変化が正の音声について大きくなる傾向	継続長が長いほど大きくなる傾向	—
4	感心	F0の変化が正の音声について大きくなる傾向	—	—
5	冷静	F0の変化が負の音声について大きくなる傾向	継続長が短いほど大きくなる傾向	F0の変化が負の時、負の値の傾向
6	肯定、相槌	F0の変化が負の音声について大きくなる傾向	継続長が短いほど大きくなる傾向	—
7	中立	—	継続長が短いほど大きくなる傾向	—
8	無関心、嫌悪、不満、落胆、苛立ち	回帰係数での変化は小さく、グラフでの評価は中程度		
9	怒り、恐れ	回帰係数での変化は小さく、グラフでの評価は低い		
10	軽蔑、喜び	F0・継続長・声の大きさによらず、中程度の評価		
11	安堵、慎み、好意、満足、同情、信頼	F0・継続長・声の大きさによらず、低い評価		

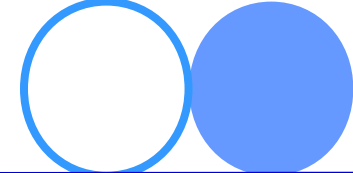


- 今回得られたデータのより詳細な分析を行なう。
- 他の語の場合について検討する。
- 表情と音声の組み合わせを検討する。



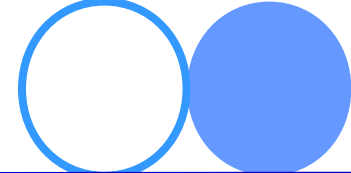
ご清聴ありがとうございました

クラスター分析(1/4)



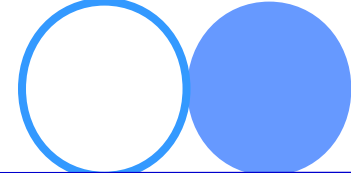
音声 クラスター1		音声 クラスター2	音声 クラスター3	音声 クラスター4	音声 クラスター5	音声 クラスター6	音声ク ラスター7
ソ0レ1+16	ソ0レ1+45	ド1レ0-25	ド1レ0-43	リ1サ0-11	ド1レ0-22	サ0ミ0-21	サ0ミ0-11
ソ0レ1+21	ソ0レ1+46	ド1レ0-35	ド1レ0-51	リ1サ0-12	リ1サ0-21	ド1レ0-11	サ0ミ0-32
ソ0レ1+23	ソ0レ1+51	ド1レ0-41	リ1サ0-13	リ1サ0-16	リ1サ0-43	ド1レ0-12	サ0ミ0-43
ソ0レ1+26	ソ0レ1+53	ド1レ0-45	リ1サ0-14	リ1サ0-23		ド1レ0-23	
ソ0レ1+33	ソ0レ1+62	ド1レ0-53	リ1サ0-26	リ1サ0-32		ド1レ0-33	
ソ0レ1+35	ソ0レ1+63	ド1レ0-54	リ1サ0-45	リ1サ0-35		ド1レ0-34	
ソ0レ1+36	ソ0レ1+64	ド1レ0-55	リ1サ0-54	リ1サ0-36		ド1レ0-44	
ソ0レ1+41	ソ0レ1+66	ド1レ0-61	リ1サ0-56	リ1サ0-55			
ソ0レ1+42	ラ0ソ1+54	ド1レ0-62	リ1サ0-61				
ソ0レ1+43	ラ0ソ1+61	ド1レ0-65	リ1サ0-63				
ソ0レ1+44		リ1サ0-65	リ1サ0-64				

クラスター分析(2/4)



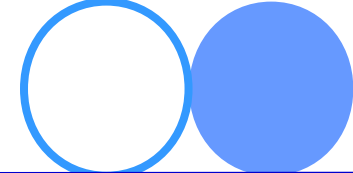
音声 クラスター8	音声 クラスター9	音声 クラスター10	音声 クラスター11	音声 クラスター12	音声 クラスター13	音声 クラスター14
サ0ミ0-13	サ0ミ0-12	ド1レ0-24	リ1サ0-22	リ1サ0-41	リ1サ0-24	サ0ミ0-15
サ0ミ0-22	サ0ミ0-23	ド1レ0-31	リ1サ0-33	リ1サ0-46	リ1サ0-25	サ0ミ0-26
サ0ミ0-33	ド1レ0-14	ド1レ0-63		リ1サ0-51	リ1サ0-31	サ0ミ0-46
ド1レ0-21	ド1レ0-15	ド1レ0-64		リ1サ0-62	リ1サ0-42	サ0ミ0-63
ド1レ0-32		リ1サ0-15		リ1サ0-66	リ1サ0-44	ド1レ0-26
ド1レ0-52		リ1サ0-34			リ1サ0-52	
					リ1サ0-53	

クラスター分析(3/4)

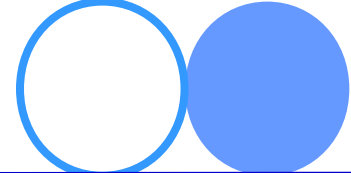


音声 クラスター15		音声 クラスター16	音声 クラスター17	音声 クラスター18	音声 クラスター19	音声 クラスター20	音声 クラスター21
サ0ミ0-14	ド1レ0-13	サ0ミ0-34	サ0ミ0-16	サ0ミ0-56	ソ0レ1+54	ソ0レ1+22	ソ0レ1+12
サ0ミ0-24	ド1レ0-42	サ0ミ0-41	サ0ミ0-35	サ0ミ0-64	ソ0レ1+55	ソ0レ1+25	ソ0レ1+52
サ0ミ0-25	ド1レ0-46	サ0ミ0-51	サ0ミ0-45	サ0ミ0-66		ソ0レ1+31	ラ0ソ1+13
サ0ミ0-31	ド1レ0-56	サ0ミ0-54	サ0ミ0-65	ド1レ0-36		ソ0レ1+34	ラ0ソ1+25
サ0ミ0-36		サ0ミ0-61		ド1レ0-66		ソ0レ1+61	ラ0ソ1+62
サ0ミ0-42		ド1レ0-16				ソ0レ1+65	
サ0ミ0-44							
サ0ミ0-52							
サ0ミ0-53							
サ0ミ0-55							
サ0ミ0-62							

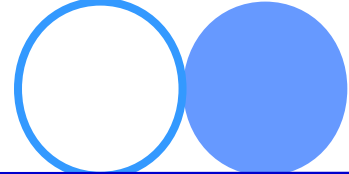
クラスター分析(4/4)



音声 クラスター22	音声 クラスター23	音声 クラスター24	音声 クラスター25	
ラ0ソ1+16	ソ0レ1+13	ラ0ソ1+44	ラ0ソ1+11	ソ0レ1+11
ラ0ソ1+24	ラ0ソ1+12	ラ0ソ1+45	ラ0ソ1+15	ソ0レ1+14
ラ0ソ1+26	ラ0ソ1+14	ラ0ソ1+51	ラ0ソ1+32	ソ0レ1+15
ラ0ソ1+31	ラ0ソ1+21	ラ0ソ1+52	ラ0ソ1+34	ソ0レ1+24
	ラ0ソ1+22	ラ0ソ1+55	ラ0ソ1+41	ソ0レ1+32
	ラ0ソ1+23	ラ0ソ1+56	ラ0ソ1+46	ソ0レ1+56
	ラ0ソ1+33	ラ0ソ1+65	ラ0ソ1+63	ラ0ソ1+53
	ラ0ソ1+35		ラ0ソ1+66	ラ0ソ1+64
	ラ0ソ1+36			
	ラ0ソ1+42			
	ラ0ソ1+43			

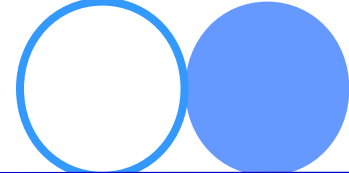


- フォルマントは声道共振周波数
— 声帯音源が声道（声帯～唇）で共鳴することにより
特定帯域ごとに倍音が増幅される。この増幅された
成分の塊、ピークをフォルマントと呼ぶ
- 母音はこのフォルマント周波数の低いほうから2つ「第1フォルマント」、「第2フォルマント」
で聞き分けることができる



- 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



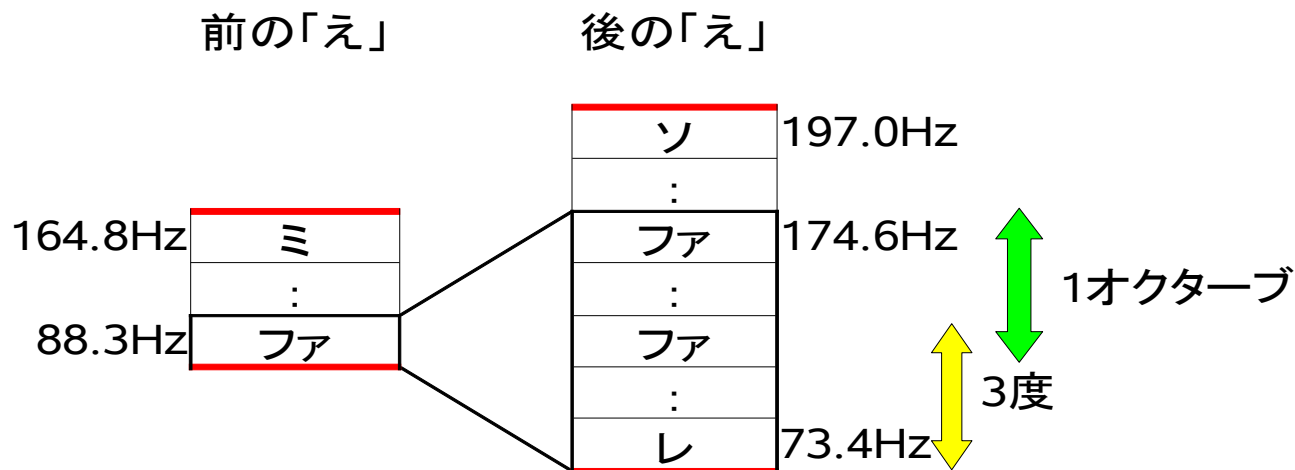
全音

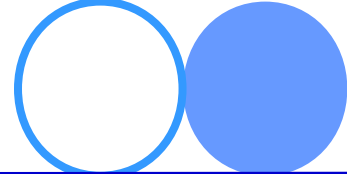
半音

☆ミファ(短2度)

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





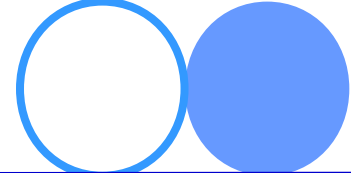
- 人間同士の対話では、「え」「あ」などの非語彙的な発話が頻繁に現れる。これらの字面には特定の意味はないが、発話の仕方によって何らかの意図、態度、感情などのパラ言語情報を伝達している
- 本研究では対話音声に頻繁に現れ、発話の仕方によって豊富な種類のパラ言語情報が伝達される「ええ」に着目した

「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%

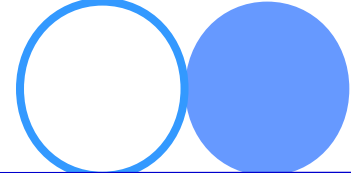
「え」「えと」「あの」「あ」「ま」の5種類で全体の92%をカバーしている



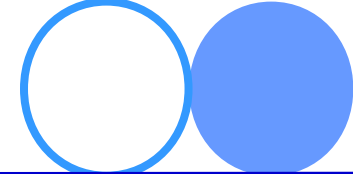
- F0の場合

回帰係数が、0.03以上を傾きが正、-0.03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

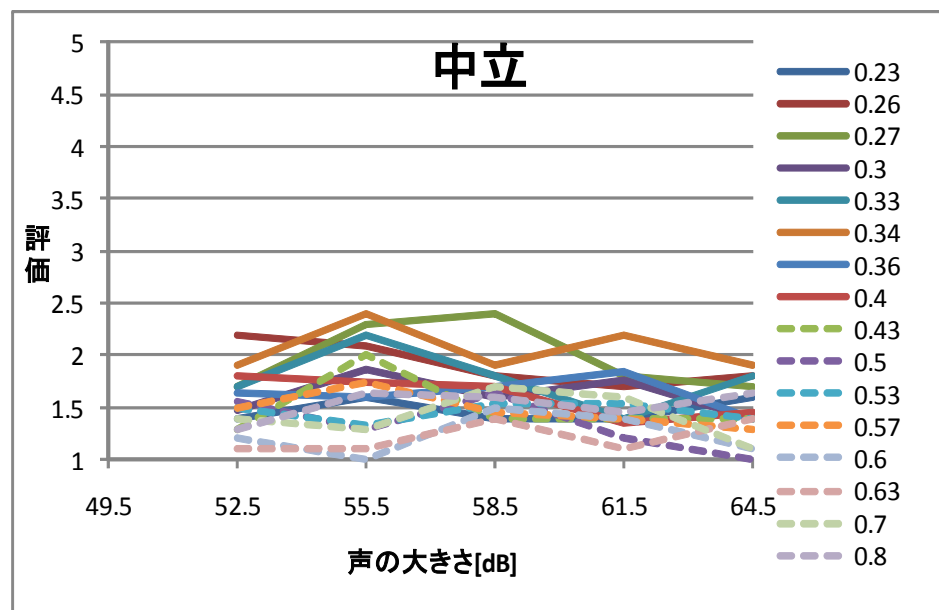
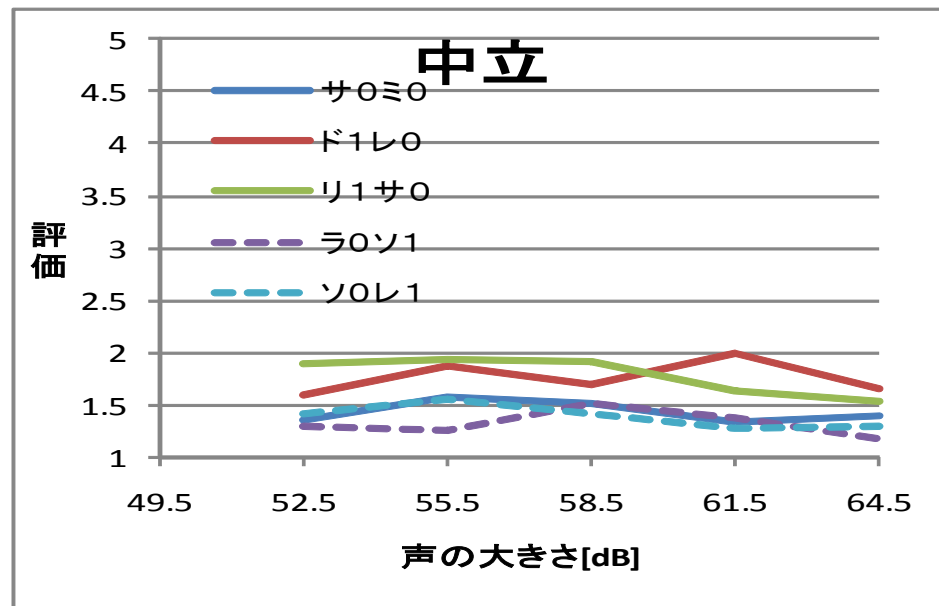


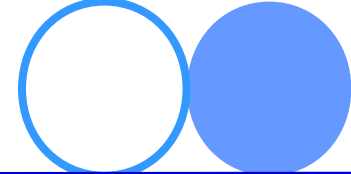
- 継続長の場合
回帰係数が、0.04以上を傾きが正、-0.04以下を傾きが負、その他は0と定義した。それを基に決めた分類グループの定義は以下のとおりである
- 正の傾向の条件・・半分以上が正の値かつその他は0の値
- 負の傾向の条件・・半分以上が負の値かつその他は0の値
- 0に近い値の傾向・・正の傾向の条件、負の傾向の条件を満たさないかつ16個の内8個以上が0の値
 - * 半分以上とは、音の長さの前半8つの中から4つ以上または後半8つの中から4つ以上のこと



⑦ 中立

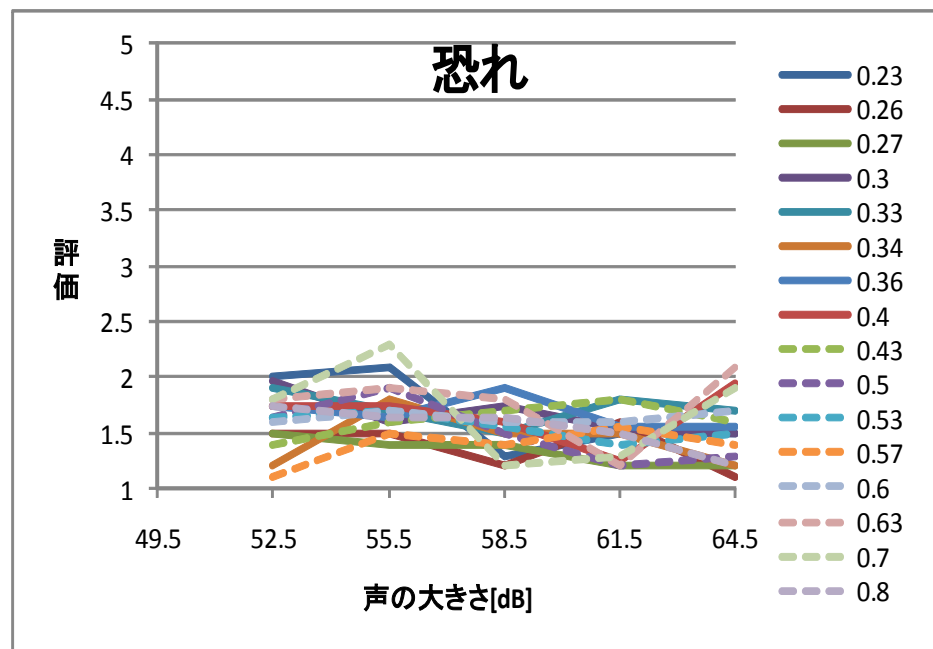
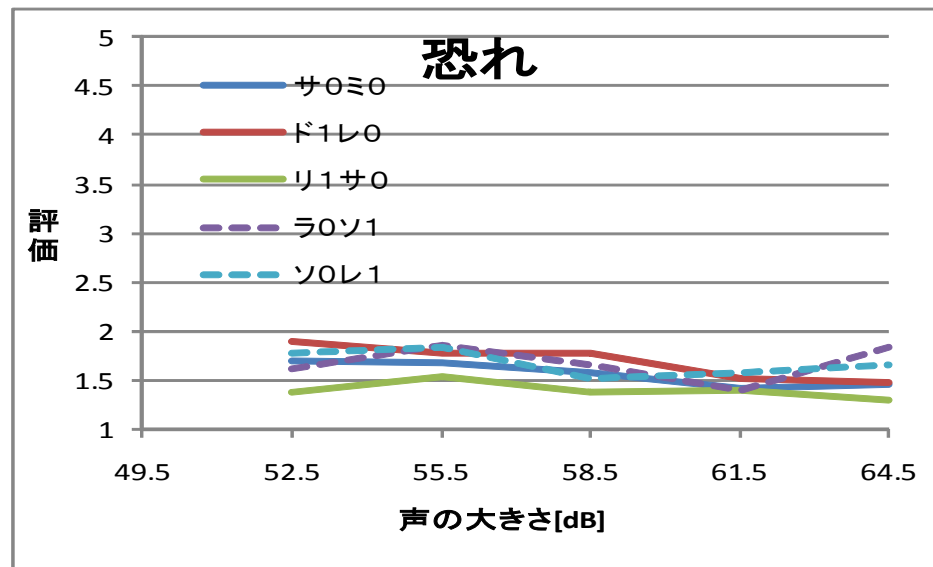
- 回帰係数は0に近い値になる傾向
- グラフから見ると、継続長が短い音声は評価が高くなる傾向
- 例として、「中立」の図を示す

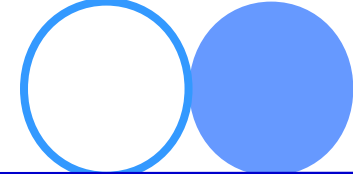




⑨ 怒り、恐れ

- 回帰係数での変化は小さく、グラフでの評価は低い
- 例として、「恐れ」の図を示す





⑩ 軽蔑、喜び

- F0・継続長・声の大きさによらず、評価は中程度
- 例として、「軽蔑」の図を示す

