

プレゼンテーション練習支援システム PRESENCE の音声フィードバックと画像フィードバック効果の比較

趙 新博^{†1} 由井 蘭隆也^{†1} 宗森 純^{†2}

概要: プレゼンテーション能力は知識を伝達する社会スキルとして重要である。その中、我々は、発表初心者の実演表現の練習を支援するプレゼンテーション練習支援システム PRESENCE を開発してきた。PRESENCE は発表練習者の発表技能を実時間評価しフィードバックする機能をもつ。そして、そのフィードバックは音声を用いるものと画像を用いるものの2通りである。その練習支援は初心者向けであり、身体表現として顔や体が前を向いているか、音声表現として声が小さくなっていないかを実時間評価する。今回、2通りのフィードバックの効果明らかにするために、フィードバックなしの場合も含む比較実験を行った。その結果、フィードバックがあることによって、身体表現や音声表現が改善すること、そして、音声フィードバックがより効果的であることがわかった。

キーワード: 音声表現, 身体表現, 音声フィードバック, 画像フィードバック

Comparison between Voice Feedback and Visual Feedback by Using Presentation Practice Support System PRESENCE

XINBO ZHAO^{†1} TAKAYA YUIZONO^{†1}
JUN MUNEMORI^{†2}

Abstract: Presentation ability becomes important social skill to transfer knowledge. We have developed a presentation practice support system named PRESENCE to improve the presenter's body and voice expressions. Using the real-time feedback functions, PRESENCE could check undesirable states and give presenters instructions of the desired states in "the vertical face direction", "the horizontal body direction" and "the voice volume". In this study, we compared between effects of the voice feedback function, those of the visual feedback function, and those without feedback function. The results show that (1) both feedback functions could improve the presenter's expression and (2) the voice feedback was more effective than the visual feedback function.

Keywords: Voice Expression, Body Expression, Voice Feedback, Visual Feedback

1. はじめに

近年、プレゼンテーションは知識を広めるための伝達手段として重要となっている。プレゼンテーション技術や経験を教えるテキスト[1], [2]が提供されてきているが、プレゼンテーション技能を習得するためには練習が重要と考えられる。その中、プレゼンテーションの練習を計算機で支援する試みが数多くなされてきている[3]~[7]。

その中、我々はプレゼンテーション練習を支援するプレゼンテーション支援システムを提案、開発してきた（このシステムを PRESENCE と呼ぶ）[6],[7]。これまでに開発されてきたプレゼンテーション練習支援システムの機能はプレゼンテーションの実演練習を支援するものとプレゼンテーション技能の理解を支援するものとに分けることができる。PRESENCE は前者に特徴あるシステムであり、練習者へのリアルタイム・フィードバック機能をもつ。同様なシステムとして、栗原らの「プレゼン先生」[3],[4]、Kopf らのシステム[5]がある。Kopf らのシステムは画像によるフィ

ードバック機能のみであるが、PRESENCE とプレゼン先生は音声によるフィードバック機能をもつ。これらシステムのフィードバック機能は定量的に効果が確認されていないのが現状である。

現在、PRESENCE を用いて、プレゼンテーション初心者をターゲットとしたシステム評価が進められている。具体的に、基礎的なプレゼンテーション技能と考えられる「聴衆を見る」ことと「声を大きくする」ことを支援対象としている。PRESENCE はフィードバック機能として、音声によるものと画像によるものの2通り用意している。一般的に、発表者は聴衆の方を向くことが望ましく、計算機のスクリーンを見ることは望ましくないとされている。従って、PC を見る視覚的フィードバックより音声フィードバックが適切でないかと考えられるが、それも検証するため、今回、我々は PRESENCE を用いた発表練習に音声フィードバック機能と画像フィードバック機能が及ぼす影響を評価・比較することとした。

以下、2. ではプレゼンテーション表現を紹介する。3. では PRESENCE の概要と実現機能、4. では比較評価実験について述べる。5. では実験結果を示す。6. は、おわり

^{†1} 北陸先端科学技術大学院
Japan Advanced Institute of Science and Technology
^{†2} 和歌山大学
Wakayama University

に、である。

2. プレゼンテーション表現

プレゼンテーション支援システム PRECENSE の研究では、プレゼンテーション技術を検討するためにプレゼンテーションに必要な表現とその評価項目を整理した[8]。それらは表1に示す通りであり、プレゼンテーション表現をシナリオ表現、スライド表現、身体表現、音声表現、質疑表現の5つにまとめている。その中、プレゼンテーションの実演と関係が強いものは、身体表現、音声表現、質疑表現である。

PRESENCE は初心者向け機能として、前を向いているかどうか、声が大きいかどうかということをチェックしてフィードバックする機能を実現している。また、両手によるジェスチャパターンや言い淀みの認識機能なども検討している[7]。現在、前者の初心者向け機能を中心にシステム性能の定量的評価を行っている。

表1 プレゼンテーション表現と評価項目[7]

Table 1 Presentation Representations and Evaluation Items.

	評価項目
シナリオ表現	アウトライン
	発表構成
	発表目的
	発表時間
	聴衆考慮
スライド表現	タイトルやスライドタイトル
	簡単明瞭な文章
	文字や図表のデザイン
身体表現	アイコンタクトと体の向き
	ジェスチャ
	顔の表情
音声表現	声の大きさ・強弱
	話すスピード
	聞き取りやすさ
質疑表現	質問への回答
	質問への準備

3. プレゼンテーション練習支援システム PRESENCE

3.1 システム概要

プレゼンテーション練習支援システム PRESENCE の概要について述べる。PRESENCE は、プレゼンテーション表現の基礎である身体表現の評価項目「アイコンタクト」と「体の向き」、音声表現の評価項目「声の大きさ」を支援対

象としている。なお「アイコンタクト」については直接支援しているわけではなく、「顔の上下方向」について支援している。また「体の向き」は左右方向で判断しているため、「体の左右方向」と呼ぶ。システム構成を図1に示す。

システムのハードウェア構成について説明する(図2)。発表練習者が参照するメイン画面をノートPCに、スライド内容をモニタ1とモニタ2に表示する。スライド内容の表示はノートPCの横に設置したモニタ及び液晶プロジェクタを介してスクリーンに表示している。発表練習者の画像・音声を取得するためにKINECT センサ (Microsoft 社) を使用している。発表練習に音声指示を与える音声フィードバックには Bluetooth イヤホンを使用している。

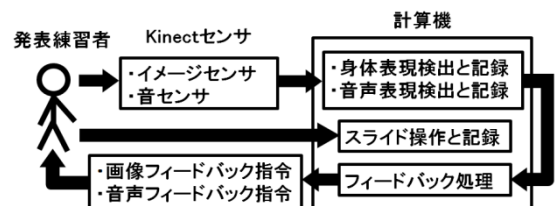


図1 システム構成

Figure 1 System Configuration.



図2 システムのハードウェア構成例

Figure 2 Hardware Configuration of System.

3.2 基本機能

基本機能はシステムによる情報表示を行う部分とスライドの操作・記録を行う部分である。

図3左側のシステム表示ウィンドウでは、KINECT センサから得た発表練習者の画像を表示している。その画像に対応する時系列で取得した骨格ノードの三次元座標(上半身:頭と首, 左右の腕首, 肘, 肩)を、色付きの円で表示している。また図3右側のスライド操作を用いて発表スライドを操作できる。その際、操作時間と画面表示されたスライド番号を記録している。またスライド操作は「次に進める」、「前に戻る」の2種類であり、マウス操作または矢印キー操作を用いて行える。



図 3 システム画面

Figure 3 Snapshot of System Screen.

3.3 発表練習者の状態について

プレゼンテーション練習者の発表状態を理解するために PRESENCE が測定データを用いて判断する状態について説明する。

3.3.1 身体表現の状態

身体表現の判定のために「顔の上下方向」に対して“聴衆を見る状態”と“パソコンを見る状態”，「体の左右方向」に対して“前方を向く状態”と“スクリーンへ向く状態”を設定している。これら4状態は発表練習者の身長や立ち位置によって変動する。そこで，発表練習を実施する前に図4に示す閾値角度を取得・設定している。

「顔の上下方向」において，顔の俯角が角度 α_1 より小さい場合，“聴衆を見る状態”としている。また，顔の俯角が角度 α_2 より大きい場合，“パソコンを見る状態”としている。次に「体の左右方向」において，体の左右方向の角度が角度 β_1 より小さい場合かつ， $-\beta_1$ より大きい場合，“前方を向いている状態”とする。そして，同角度が角度 β_2 より大きい場合，“スクリーンへ向く状態”としている。

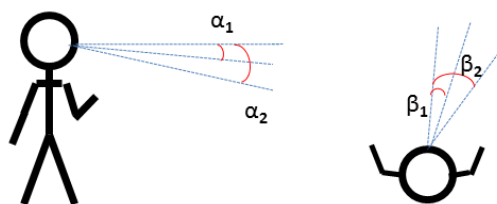


図 4 身体表現の閾値角度

Figure 4 Threshold Angle for Body Representation.

3.3.2 音声表現の状態

音声表現「声の大きさ」において望ましくない状態として“声が小さい状態”を判定している。1m 離れた話者による会話音声の等価騒音レベルは音圧レベルが 56dB であることが知られている[8]。そこで，この値より小さい 55dB に対応する KINECT センサの音量値以下の場合，“声の大

きさ」として望ましくない状態とした。KINECT センサから取得された 16bits の音声データは，0~1 までの音量値に変換される。現在，音圧レベル 55dB に相当する値を測定実験から求めており，KINECT センサの音量値が 0.70 以下の場合，声の大きさが望ましくない状態としている。

3.4 フィードバック機能について

PRESENCE は発表練習者の望ましくない状態を判定し，その状態を改善するための指示を発表者にフィードバックする機能をもつ。望ましくない状態は 3.3 で述べたように，身体表現については“パソコンを見る状態”，“スクリーンへ向く状態”，音声表現については“声が小さい”状態である。

聴衆を対象とするスピーチにおいて，“不誠実な話し手”であると印象を与えるスピーカのアイコンタクト使用割合は 20.8% であることが知られている[9]。そこで，本研究では，身体表現「顔の上下方向」，“体の左右方向”，音声表現「声の大きさ」に関するデータを 30 秒間測定し，その間の望ましくない状態の割合を求める。もし，その割合が 20.8% 以下である場合，次に述べるフィードバックを与える。なお本研究では中国人留学生を実験参加者としており，実験参加者の母国語である中国語を用いたフィードバックを実装している。

最初に，音声フィードバックについて述べる。「顔の上下方向」については中国語で“画面前方”（日本語で“前を見る”という意味）と話した音声ファイル，“体の左右方向”については中国語で“朝向前方”（日本語で“前方を向く”という意味）と話した音声ファイル，“声の大きさ”については中国語で“大声一点”（日本語で“大きい声にしろ”という意味）と話したファイルが用意されている。望ましくない状態に対応した音声ファイルがフィードバック時に再生される。いずれの音声ファイルも再生時間は 1 秒である。よって，音声フィードバック時間は 1 秒である。

次に画像フィードバックについて述べる。図 5 に示す画像がシステム画面ウィンドウ上にポップアップ・ウィンドウを用いて 1 秒間表示される。「顔の上下方向」に対して図 5(a)，「体の左右方向」に対して画像（図 5(b)），「声の大きさ」に対して画像（図 5(c)）が用いられ，それぞれの画像に書かれた文字は音声フィードバックと同じ中国語で書かれている。また，各画像にはアイコンも表示している。

4. 評価実験

プレゼンテーション練習支援システム PRESENCE の評価実験について述べる。

4.1 実験参加者と実験環境

実験参加者は，北陸先端科学技術大学院大学の中国人修士学生 30 名である。実験は 3 通りであり，音声フィードバックあり，画像フィードバックあり，フィードバックなし，である。各実験にランダムに 10 人の実験参加者が割り当て

られた。実験参加者は修士学生であるが、研究会や国際会



(a)顔の上下方向“前を見る”

(b)体の左右方向“前方を向く”

(c)声の大きさ“大きい声にしろさい”

図 5 画像フィードバック

Figure 5 Visual Feedback.

議などの学会発表経験がないものであった。

実験環境は図 2 と同じである。ノート PC とモニタはレクチャ卓に並べて配置される。レクチャ卓の高さは 0.9m である。KINECT センサの高さは 1.6m であり、レクチャ卓から真正面の水平距離 2.4m の場所に設置している。

実験手順について述べる。最初に、実験参加者は発表練習支援システムについて説明を受ける。そして、実験参加者は 15 分かけて発表準備をする。その際、前もって準備された発表スライドを見ることが出来る。次に、3.3.1 で述べた閾値角度を測定する作業を行った。そして、実験参加者は約 5 分かけてシステムの操作練習をする。その後、実験参加者はそれぞれの支援機能を使用した発表練習を行った。すべての実験において、3.2 で説明した基本機能を用いることができた。そして、音声フィードバックありの場合、3.4 で説明した音声フィードバックが、画像フィードバックありの場合、3.4 で説明した画像フィードバックが与えられた。

実験で使用する発表スライドは所属大学の紹介であり、実験参加者が身近な内容とした(図 6)。また、実験参加者の母国語に合わせて中国語で書かれたスライドを用意した。スライドは 8 枚であり、その内容は順番に、表紙、目次、大学概要、本学の特徴、その大学を選ぶ理由(研究、生活、進路)であった。

実験終了後、実験参加者はアンケートに回答する。すべての種類の実験において、システム全体、情報表示画面、スライド操作画面が役に立ったかどうかについて 5 段階評価する。また、フィードバックを用いた実験の場合、使用したフィードバックが身体表現や音声表現に役立ったか



図 6 実験用スライドの内容(中国語)

Figure 6 Experimental Slide in Chinese.

を 5 段階評価する。5 段階評価の選択項目は「1. とても役に立たなかった, 2. 役に立たなかった, 3. どちらともいえない, 4. 役に立った, 5. とても役に立った」である。最後に、システムに対する意見・感想を自由記述できる欄がある。

4.2 フィードバックのデータ解析方法

過去の我々の研究では、2 秒間のサンプリング時間を用いて音声フィードバックの効果を解析してきた。よって、本研究においても、フィードバック効果の解析に、2 秒間のサンプリング時間を用いることとした。また人間工学において音声や視覚情報に対する反応時間は平均的に 300ms であることが知られている[10]。よって、フィードバック情報が提示され、発表練習が気づいて身体動作に反映するために 300ms かかるとした。

フィードバックの前後データ解析に用いる時間を図 7 に示す。フィードバック開始時間を t とする。なお身体表現「顔の上下方向」、「体の左右方向」については角度値を、音声表現「声の大きさ」については音量値を数値データとして用いる。また 3.4 で述べたように、すべてのフィードバックは再生・提示時間は 1.0 秒である。

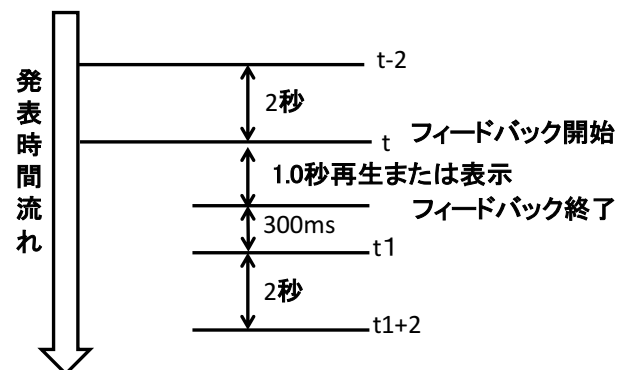


図 7 フィードバック解析

Figure 7 The Analysis of Feedback.

フィードバック前の値は、フィードバック前の時間 $t-2$ から時間 t までの 2 秒間の数値データの平均値とする。フィードバック後の角度値や音量値は、フィードバックの提

示が完了した時間から反応時間300ms経った時間をt1とし、時間t1から時間t1+2sまでの2秒間にある数値データの平均値とする。フィードバックなしの場合も、フィードバックのタイミングをログデータとして記録しているため同じ方法で値を求めることができる。

最後に、フィードバックによってプレゼンテーション表現が改善されたかどうかの判定について述べる。フィードバック後の値が3.3で述べた望ましい状態となれば、改善されたと判断する。望ましい状態は、身体表現については“聴衆を見る状態”，“前方を向く状態”であり、音声表現については“声が小さい状態”ではない状態である。

5. 実験結果

5.1 定量的評価

実験結果を表2に示す。フィードバックの「回数」、フィードバック前後の測定値差を「変化量」、そして、フィードバック後の測定値が改善した回数を「改善回数」を求めた。ここで3種類の実験群間を比較するために、一元配置の分散分析を用い、有意差が見られた場合、Tukey-Kramer法による多重比較検定を行った。

表 2 実験結果
 Table 2 Results of Experiment.

	内容	音声フィードバック	画像フィードバック	フィードバックなし
回数	顔の上下方向*	1.1回	2.9回	3.4回
	体の左右方向	0.3回	0.2回	0.5回
	声の大きさ*	1.1回	1.2回	3.6回
変化量	顔の上下方向	-3.3度(n=11)	-1.5度(n=29)	-1.9度(n=34)
	体の左右方向	-9.6度(n=3)	4.7度(n=2)	3.8度(n=5)
	声の大きさ*	0.08(n=11)	0.01(n=12)	0.03(n=36)
改善数	顔の上下方向	1.0回	2.1回	2.0回
	体の左右方向	0.2回	0.0回	0.3回
	声の大きさ	1.0回	0.7回	1.9回

一元配置分散分析：*p<.05

回数についてみると、身体表現「顔の上下方向」(p<.05)、音声表現「声の大きさ」(p<.05)において実験群間に有意差が見られた。身体表現「顔の上下方向」においては音声フィードバックがある場合は、フィードバックなしの場合と比べて有意差があり(p<.05)、フィードバック回数が少ない結果となった。音声表現「声の大きさ」については画像フィードバックがある場合は、フィードバックなしの場合と比べて有意差があり(p<.05)、フィードバック回数が少ないという結果となった。変化量については、音声表現「大きい声」において実験群間において有意差が見られた(p<.05)。そして、音声フィードバックがある場合、フィードバックなしの場合と比べて有意差があり、声の大きさの変化量が

大きいことがわかった。改善数については実験群間に有意差はみられなかった。

次に、フィードバック前後の測定値を比較した結果を表3に示す。その前後比較には対応のあるt検定を用いた。音声フィードバックを用いた場合、身体表現「顔の上下方向」に有意差があり(p<.05)、前を向く角度に近づくこと、音声表現「声の大きさ」に有意差があり(p<.05)、声が大きくなる結果となった。また画像フィードバックを用いた場合、身体表現「顔の上下方向」に有意差があり(p<.01)、前を向く角度に近づくことがわかった。

表 3 フィードバック時の前後データ比較
 Table 3 Comparison between pre-data and post-data at feedbacks.

	フィードバック内容	前	後
音声フィードバック	顔の上下方向(n=11) *	14.0度	10.8度
	体の左右方向(n=3)	32.6度	22.9度
	声の大きさ(n=11) *	0.54	0.62
画像フィードバック	顔の上下方向(n=29) **	14.6度	13.1度
	体の左右方向(n=2)	7.9度	12.5度
	声の大きさ(n=12)	0.56	0.56
フィードバックなし	顔の上下方向(n=34)	14.3度	12.5度
	体の左右方向(n=5)	39.3度	43.1度
	声の大きさ(n=36)	0.55	0.57

対応のあるt検定：*p<.05 **p<.01

以上より、音声フィードバックが最も発表練習者に影響を与えており、声の大きさの変化量が大きいとともに、顔の上下方向は望ましくない状態が少ないことがわかった。また画像フィードバックがある場合、声の大きさが小さい状態が少ないことがわかった。

5.2 アンケート評価

システムに対する5段階アンケートの結果を表4に示す。その結果、音声フィードバックや画像フィードバックがある場合、システム全体、情報表示機能、スライド操作機能のすべての項目において、最頻値が4または5であり、高評価であることがわかる。一方、フィードバックなしの実験では情報表示機能の最頻値が2と低くなるという結果になった。しかしながら、Kruskal-wallis検定を用いて3つの実験群を比較したが有意差は得られなかった。

フィードバックがプレゼンテーション表現の支援に役立つかどうかのアンケート結果を表5に示す。その結果、音声フィードバックはすべての表現において最頻値が4であり、役に立つと評価されている傾向があった。一方、画像フィードバックについては、いずれの項目も最頻値は3であり、どちらともいえないという結果となった。しかしながら、Mann-WhitneyのU検定で両フィードバックを表現ごとに比較したが、いずれも有意差はみられなかった。

表4 システムに関する5段階アンケート結果

Table 4 Results of Five-Scaled Questionnaire on System.

評価	音声フィードバック					画像フィードバック					フィードバックなし				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
システム操作	0	0	0	3	<u>7</u>	0	0	0	<u>8</u>	2	0	0	1	3	<u>6</u>
情報表示	1	1	1	<u>6</u>	1	0	2	1	<u>7</u>	0	1	<u>4</u>	2	3	0
スライド操作	0	0	2	<u>6</u>	2	0	0	1	<u>6</u>	3	0	0	2	<u>7</u>	1

下線太字: モード

表5 フィードバックに関するアンケート結果

Table 5 Results of Five-Scaled Questionnaire on Feedbacks.

5段階評価	音声フィードバック (人数)					画像フィードバック (人数)				
	1	2	3	4	5	1	2	3	4	5
顔の上下方向	0	1	2	<u>7</u>	0	0	1	<u>6</u>	3	0
体の左右方向	0	1	3	<u>6</u>	0	0	1	<u>5</u>	4	0
声の大きさ	0	1	1	<u>7</u>	1	0	1	<u>5</u>	4	0

下線太字: モード

アンケートの自由記述欄について述べる。音声フィードバックについては、「音声提示はよい」、「フィードバックは発表を妨害するところもあるが、効果的である」という意見が1名ずつ回答されていた。画像フィードバックについては、「フィードバックの時間間隔がよい」、「(身体表現の)フィードバック情報が区別しにくい、音量の改善(声の大きさをどうすればよいか?)がわからない」、「発表を邪魔する場合がある」、「フィードバック時間が少し短い」という意見が1名ずつ回答されており、音声フィードバックと比較して問題点が指摘されていた。両フィードバックを通じてシステム操作に対して3名の好意的意見があることに對して、システム画面をきれいにしたいという指摘が2名よりあった。フィードバックなしの実験では、「情報表示画面にある人物が小さくて、目立たない」という意見に加えて「PowerPointで直接操作したかった」、「次のスライド内容を提示できればよい」という意見が1名ずつ回答されていた。

6. おわりに

プレゼンテーション練習支援システム PRESENCE は、発表練習者の身体表現と音声表現を実時間チェックし、望ましい状態に向けたフィードバックを与える機能をもつ。今回、PRESENCE がもつ音声フィードバック機能と画像フィードバック機能を比較する実験を行った。また合わせてフィードバックなしの実験も行った。その結果、得られた知見は次の通りである。

(1) 音声フィードバックが最も発表練習者の実演に影

響を与えており、身体表現「顔の上下方向」に問題がある状態が少なくなると共に、身体表現「顔の上下方向」と音声表現「声の大きさ」が改善する方向に練習できる。

(2) 画像フィードバックは音声表現「声の大きさ」に問題がある状態が少なくなると共に、身体表現「顔の上下方向」が改善する方向に練習できる。

今後、身体表現におけるジェスチャに対する理解支援機能[7]を使用することなどにより、より高度なプレゼンテーション支援を検討していく予定である。

謝辞 本研究にご協力頂いた北陸先端科学技術大学院の皆様、謹んで感謝の意を表す。

参考文献

- [1] A. Wallwork. English for Presentations at International Conferences. Springer, 2010, 196p.
- [2] ロバート・R・H・アンホルト. 理系のための口頭発表術. 株式会社講談社, 2013, 229p.
- [3] 栗原一貴, 後藤真孝, 緒方淳(他). プレゼン先生: 画像情報処理と音声情報処理を用いたプレゼンテーションのトレーニングシステム. WISS 第14回インタラクティブシステムとソフトウェアに関するワークショップ論文集, 2006, p.59-64.
- [4] K. Kurihara, M. Goto, J. Ogata, Y. Matsusaka & T. Igarashi. Presentation sense: a presentation training system using speech and image processing. Proc. of ACM ICMI International Conference on Multimodal Interfaces, 2007, p.358-365.
- [5] S. Kopf, D. Schon, B. Guthier, R. Rietsche, & W. Effelsberg. A Real-time Feedback System for Presentation Skills. Proc. of AACE EdMedia: World Conference on Educational Media and Technology, 2015, vol.2015, no.1, p.1686-1693.
- [6] 趙新博, 由井菌隆也. ノンバーバル表現に注目したプレゼンテーション支援システムの提案. 研究報告グループウェアとネットワークサービス(GN), 2016, vol.91, no.42, p.1-6.
- [7] 趙新博, 由井菌隆也, 宗森純. ノンバーバル表現に注目したプレゼンテーション支援システムの開発. 研究報告グループウェアとネットワークサービス(GN), 2015, vol.94, no.6, p.1-6.
- [8] 白石君男, 神田幸彦. 日本語における会話音声の音圧レベル測定. Audiology Japan, 2010, vol.53, no.3, p.199-207.
- [9] Mark L. Knapp, Judith A. Hall. Nonverbal communication in human interaction 4th ed. Harcourt Brace College Publishers, 1997, p.496.
- [10] 野呂影勇(編). 図説エルゴノミクス. 1990, 日本規格協, 666p.