

# 音声メディア処理による非母語者の会話支援

筑波大学

図書館情報メディア研究科

2016年3月

叶 環

# 目次

第1章	序論	4
1.1	研究背景	4
1.2	研究目的	4
1.3	構成	5
第2章	関連研究	6
2.1	異文化コミュニケーション支援	6
2.2	話速変換における会話支援	6
2.3	話速測定方法	7
第3章	提案システム	8
3.1	録音	9
3.2	発話区間検出	9
3.2.1	均一化处理	9
3.2.2	リエンファシス処理	9
3.2.3	フレームの切り出し	11
3.2.4	発話区間検出	11
3.3	母音検出	14
3.3.1	フレームの切り出し	15
3.3.2	LPC ケプストラム係数算出	15
3.3.3	母音の終了点判断	16
3.4	発話速度の計算	16
3.5	データの保存	17
第4章	動作確認実験	18
4.1	実験概要	18
4.2	実装環境	18
4.3	動作確認実験の結果	18
第5章	評価実験	21
5.1	実験目的	21
5.2	比較条件	21
5.3	実装環境	25
5.4	実験対象	25
5.5	実験手順	25
5.5.1	実験説明	25
5.5.2	発話速度閾値の設定	25

5.5.3 母音標準パターンの作成.....	26
5.5.4 会話実験.....	27
5.5.5 データ取得.....	28
第6章 実験結果.....	31
5.2 主観評価結果.....	31
5.3 客観評価結果.....	35
5.3.1 システムの認識率.....	35
5.3.2 システム平均処理時間.....	35
5.3.3 平均発話速度.....	36
5.3.4 非母語者の発話速度設定値.....	37
5.3.5 成績平均点.....	38
第7章 考察.....	39
謝辞.....	42
参考文献.....	43

## 図表の目次

図 1. システムの流れ .....	8
図 2. 発話区間検出のアルゴリズム .....	13
図 3. 母音検出のアルゴリズム .....	14
図 4. システムのインターフェース .....	17
図 5. 非提示条件条件の配置図 .....	22
図 6. 非提示条件条件の実験様子 .....	22
図 7. システムあり, 話者提示条件する条件の配置図 .....	23
図 8. システムあり片方提示の場合の実験様子 .....	23
図 9. システムあり, 母語者と非母語者両方に提示する条件の配置図 .....	24
図 10. システムあり, 両者提示条件条件の実験様子 .....	24
図 11. インターフェース 1 (発話速度閾値の設定) .....	26
図 12. インターフェース 2 (母音標準パターン作成) .....	26
図 13. 各条件の質問紙項目 1 から 5 (母語者) の比較 .....	31
図 14. 各条件の質問紙項目 1 から 5 (非母語者) の比較 .....	33
表 1. 実験順番 .....	28
表 2. 条件ごとに配る質問紙の項目 .....	29
表 3. 質問紙項目 6 から 10 の結果 .....	34
表 4. 各母語者の平均発話速度 .....	36
表 5. 全実験のゲーム得点 .....	38
付録 1 実験説明 (非母語者) .....	47
付録 2 実験説明 (母語者) .....	48
付録 3 実験説明 (共通) .....	49
付録 4 キーワードメモ用紙 .....	50
付録 5 質問紙 (非母語者) .....	51
付録 6 質問紙 (母語者) .....	53
付録 7 質問紙結果 (母語話者) .....	55
付録 8 質問紙結果 (非母語話者) .....	58

## 第1章 序論

本章では、本研究の背景及び目的を述べ、最後に本論文の構成について説明する。

### 1.1 研究背景

近年、世界中にグローバル化が進んでおり、異なる国の間での交流が活発化している。日本では、訪日外国人数は例年以上に増加し、2015年では1796万人以上に伸びた[1]。異なる母語の人々の間でのコミュニケーションを行う機会がさらに増やすと考えられる。そのため、日本語は母語ではない在日外国人にとって、日本語を共通言語として会話を行う場合、または日本国の母語者と会話を行う場合、非母語者が聞き取れないが、母語者がそのままの速度で話し続けるという会話のやり取りにアンバランスが生じると考えられる。

異文化コミュニケーションの場合、特に異言語間でコミュニケーションを取る場合、遠隔会話と対面会話に関わらず最大の阻害要因は言語であることは知られている。非母語者が聞き取れない時、母語者にその言葉を反復し、あるいは噛み砕いて説明するように頼む行動を取る現象が見られる。しかし、非母語者が母語者に聞き返すことが多くの場合、非母語者にとって心理的に負担となる[2]。非母語者が母語者にゆっくり話すように伝えることは、聞き返す回数より気まづくなくなるためできなくなる可能性もある。また、多くの場合、母語者は無自覚に異文化の人達の前でも、そのままのコミュニケーション行動を実行してしまうことがある[3]。さらに、文化差異の存在より非母語者が聞き返すことに躊躇してできない可能性もある。そのため、母語者は最初、非母語者に対して意識的にやさしくゆっくりして話す[4]、無意識に話速を上げてしまうこともあると考えられる。日常の発話では、発話内で非常に大きく速度が変化する[5]という検証実験の結果がある。

また、非母語者と限定せず、語学力またはコミュニケーション力の差によって会話が成り立たない場面でも同じくアンバランスが生じると考えられる。コミュニケーション力が比較的弱い方をリアルタイムで支援することが大きな意味を持つと考える。

本研究では、音声メディア処理で、このような発話速度にアンバランスな場面を想定し、母語者の話速を検知し、話速が速いと判断した際に Warning を出して母語者に気づかせるウェアネスシステムを提案し、非母語者と母語者の会話を支援する。

### 1.2 研究目的

本研究は異言語コミュニケーション支援を目的とした発話速度の自動検知で早口を母語者に知らせるウェアネスシステムを構築し、システムが非母語者の理解、及び会話にどの影響を与えるかを調べ、会話のアンバランスを解消することを目指す。

### 1.3 構成

本論文は、本章を含め 7 つの章で構成されている。第 1 章では研究背景と研究目的について述べる。第 2 章では関連研究について紹介する。第 3 章では提案システムの設計および実装について述べる。第 4 章では提案システムを使って評価実験とその設計について述べる。第 5 章では評価実験の結果について述べる。第 6 章では考察と今後の課題について述べる。

## 第2章 関連研究

これまで、話を分かりやすくするためには、様々な研究がされてきた。たとえば高齢者や難聴者を対象として、音声を「ゆっくり」、「はっきり」とさせ、低下した言語理解速度を補償する多くの研究がある[6]。

### 2.1 異文化コミュニケーション支援

異文化コミュニケーションを支援するために、音声認識で会話内容を識別して画面で提示するシステム[7]と会話中の名詞のみを認識してネットワークで検索した関連情報を提示するシステム[8]がある。これらの研究については音声認識の精度により、提示する情報の正確さと即時性が十分ではない可能性がある。

そのため、福島らは異言語の討論者同士が言葉を入力できるインターフェースを構築し、多数の異言語をリアルタイムで共通言語に翻訳することができるサーバを開発した[9]。これにより内容の精度を上げることができたが、討論中に非母語者が迅速に討論内容を入力する必要があるまたは第三者が内容を入力する必要があるため、その認知負荷が問題になると考えられる。

本研究は、母語者の話速をリアルタイムで検知し、それを早口であると判断した際にシステムが自動的に母語者にその旨を提示して知らせる。本システムでは、音声の特徴量を利用して音節数を検知することにより話速を検出する。つまり、会話の内容自体は認識しないため、内容認識において生じる誤りがなく、プライバシー保護にもつながる。

### 2.2 話速変換における会話支援

従来、早口音声は若年者と高齢者の聞き取りに影響を及ぼし、特に高齢者に対して、単語理解度[10]は有意に低下する結果があった[11]。また、通常より1.5倍速のニュース音声が高齢者と若年者に対して「疲労感」[12]などの心理成分が検出されるため、早口音声を話速変換処理で会話の理解度を補足するために、多くの話速変換システムが開発されている[13][14]。

しかし、機械的な話速変換は自由会話場面では時間のズレができて者と非同期の違和感を感じることになる[15]。そのため、本研究は機械的に話速を変換せず、話速が速いと検知した際に発者に知らせ、発者が自発的にゆっくり話すようという手法で会話の聞き取り支援を行う。

## 2.3 話速測定方法

音韻論上では、音声を分類し、母音と子音に分けられる。母音は「肺から口や鼻を通過して外に出される息が途中で何の妨害も受けずに発せられる言語音である」。一方子音は母音とは逆に「声道において途中どこかを挟めたり閉じたりする障害を伴う言語音である」と定義されている[16]。母音は英国の音声学学者ダニエル＝ジョーンズが設定した基本母音に基づき、日本語では「a」「i」「u」「e」「o」で示す[17]。それ以外の言葉音(ん以外)は子音と見なす。従って、日本語は「子音+母音」や「母音」だけの組み合わせと見直すことができる。そのため、母音を数えることで発話速度を推定することができると思う。

日本人による日本語単語の話速の評価に用いられる基準は、単位時間当たりの「モーラ数」と「音節数」である。音節は母音と子音の組合せからなり、ひとまとまりと感じられる音連続をいう。モーラ数はリズムを構成する最小単位で、日本語などごく少数の言語にしか認められない[18]。撥音(ん)、促音(小さい「っ」)、長音(ー)は、独立した音節ではなく、音節の一部だが、モーラとしては、1モーラとなる。例えば、「しゃしん」という単語を音節で数えると2音節であり、モーラとすると3モーラである。また、話し言葉音声認識において、「音節」が「モーラ」より音響モデルの単位として適していることを示した研究[19]がある。

従って、本研究では、日本語は子音と母音の組み合わせの特徴から、音響特徴量で母音の検出することで音節数を数える。

ここで、母音と子音を正確に認識するシステムは主に音声認識システムに用いられている。例えば、人間の聴覚前処理部である蝸牛内のリンパ液による音波の減衰を離散ウェーブレット変換による多重解像度解析により模擬して、母音と子音の内容まで正確に認識する研究がある[20]。しかし、この研究は自由発話音声を対象とした認識ではなく、ハードウェアを使って単独の音素の連続マッチングで音素の内容を認識する。さらに設備の準備や複雑な計算が必要である。

また発話速度推定の手法として、異なる狭帯域エネルギーの時間変化から推定し、平均16%の誤差で推定することができる研究[21]がある。しかし、話速が速い時に起こるモーラの連結現象もあり、手法に用いる7つのパラメータを探索する必要もあり、リアルタイムでの自動話速検知性能は知られていない。

なお、会話音声での母音認識方法として、LPC分析を介して得られる音響特徴量と標準パターンの比較照合は最も通常な方法である。その中、音源情報を含まず、スペクトル表現が連続であるLPCケプストラム係数は最も良い結果となる[22]。また、本研究は、リアルタイムで母音の数を検知することを目指し、軽便な非統計的な距離尺度ユークリッド距離を用いる。本研究は、ソフトウェア環境で発話区間を検知し、その中の母音の数を音節数と切り替え、発話速度を計算すると考えている。



### 第3章 提案システム

現在，母語者と非母語者が二人一組での会話場面を想定する．母語者の話速のみが検知され，早口と判断した際に「話すスピードが速い」と Warning 標識を出す．

具体的な流れは図 1 のように考えられる．



図 1. システムの流れ

### 3.1 録音

母語者のみマイクを付け、録音された音声はシステムに入力する。録音する音声の周波数は 10000Hz、ビット数 16bit、モノラル音声である。

### 3.2 発話区間検出

前章で言及したように、日本語は「子音+母音」の組み合わせあるいは「母音」だけだと考えられて、子音あるいは母音をカウントすることで発話の音節数を測定できる。ここで、音声各フレームの零交差数とエネルギーにより発話区間を検出する。

#### 3.2.1 均一化処理

録音による多くのクリップの音量を均一にするために、クリップ全体の最大音量を調べ、その音量を-1dB から 1dB までの範囲に抑えるように均一化処理をかけた。

MATLAB[23]では、以下のコードで均一化処理を行う。

```
newdata = data/ max(abs(data));
```

$\max(A)$  は行列  $A$  から最大値を返すため、ここでは音声データ  $data$  のピークレベルの音量を示す。関数  $\max$  に対して：(1).  $A$  が複素数の場合、 $\max(A)$  は、最大の大きさをもつ複素数を返す。大きさが等しい場合、 $\max(A)$  は最大の大きさと最大の位相角をもつ値を返す。(2).  $A$  がスカラーの場合、 $\max(A)$  は  $A$  を返す。(3).  $A$  が 0 行 0 列の空の配列の場合、 $\max(A)$  も同じく空である。振幅より負値があるため、 $\max$  関数より絶対値を取っている。それから、振幅の割合を音量の制限値をかけてダイナミックレンジの限界まで引き広げることができる。

#### 3.2.2 プリエンファシス処理

プリエンファシス処理は声帯の基本振動、空気中の伝搬の影響を除き、口腔内の形状による周波数特性を強調する処理である[24]。

プリエンファシスの式は式 (1) のように示す。

$$y(m) = x(m) - px(m-1) \quad (1)$$

$x(m)$ はサンプル数  $m$  番目のデジタル音声データで、 $p$  はプリエンファシス係数、 $y(m)$  はプリエンファシス処理をしたデータである。ここで  $p$  は 0.97 である。1 サンプルはその前の値にたいして、大きい変化する場合つまり高周波数の場合、値がそのまま保存される。値の変化が小さかった場合つまり低周波数の場合、その値は減じられると高域強調を行う。プリエンファシス処理が Matlab では関数 `filter[25]` を使い、以下のコードより示す。

```
newData = filter([1 -0.97], 1, data);
```

$y = \text{filter}(b, a, x)$  は、分子係数と分母係数  $b$  および  $a$  で定義される有理伝達関数を使用して入力データ  $x$  をフィルター処理である。 $x$  は元のデータ `data` で、 $y$  はプリエンファシス処理をした `newData` である。 $a$  は分母係数のため、非ゼロでなければならない。関数 `filter` に対して：(1).  $x$  がベクトルの場合、`filter` はフィルター処理されたデータを  $x$  と同じサイズのベクトルとして返す。(2).  $x$  が行列の場合、`filter` は最初の次元に沿って機能し、各列のフィルター処理されたデータを返す。(3).  $x$  が多次元配列の場合、`filter` は、サイズが 1 でない最初の配列次元に沿って機能する。

有理伝達関数は、次の差分方程式で表すこともできる。

$$a(1)y(m) = b(1)x(m) + b(2)x(m-1) + \dots + b(m_b+1)x(m-m_b) - a(2)y(m-1) - \dots - a(m_a+1)y(m-m_a) \quad (2)$$

$n_a$  はフィードバック フィルターの次数、 $n_b$  はフィードフォワード フィルターの次数である。ここで、

```
b = [1 -0.97];
```

```
a = 1;
```

のため、式 (3) で表すことができる。

$$a(1)y(m) = b(1)x(m) + b(2)x(m-1) \quad (3)$$

### 3.2.3 フレームの切り出し

連続音声の音声分析を行うため、音声データを短い時間単位で切り出す処理が必要である。一定時間単位で切り出されたデジタル信号化したデータをフレームとする。音声フレームの長さ[26]について、信号が弱定常状態と仮定できる範囲が 20-40 msec である。また、音声分析用の窓関数のエネルギーは、矩形窓長で約  $1/3 - 1/2$  のエネルギーと同一になるため、本研究では、フレーム長は 25msec で、シフト長は 10msec にした。

### 3.2.4 発話区間検出

フレームを切り出された後に発話区間検出処理を行う。そのため、フレームごとの音声特徴量である STE (Short-Time Energy : 声帯振動より発生するエネルギー) 式 3 と ZCR (Zero Crossing Rate : 波が振幅中央軸より上(正)から振幅中央軸より下(負)に変わる頻度, つまり零交差数) 式 5[27][28]を抽出する。

$$E_n = \sum_{m=0}^{N-1} x_n^2(m) \quad (4)$$

$$Z_n = \frac{1}{2} \sum_{m=0}^{N-1} |\text{sgn}[x_n(m)] - \text{sgn}[x_n(m-1)]| \quad (5)$$

(3)の中,  $\text{sgn}$  は符号関数, つまり

$$\text{sgn}(x) = \begin{cases} -1, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

と示す。

その中,  $x_n(m)$ はサンプル数  $m$  を変数とする第  $n$  フレームの音声データである。  $E_n$  は第  $n$  フレームの STE の値で,  $Z_n$  は第  $n$  フレームの ZCR の値である。また,  $N$  はフレーム長を示す。STE は MATLAB では以下のコードより表す。

```
ste = sum((abs(enframe(x, FrameLen, inc))).^2, 2);
```

ここでは `enframe` 関数はフレーム長 `FrameLen` とシフト長 `inc` で音声  $x$  をフレームとして切り出す関数である。切り出されたフレームは行ごとに保存されて行列として出力する。

`sum` 関数[29]は配列要素の和を計算する関数である。 `sum(A, dim)` に対して,  $A$  が行列の場合, `sum(A, 2)` は各行の和をもつ列ベクトルになる。また,  $^2$  は二乗と示す。ここでは `sum` はフレームごとの振幅より算出したエネルギーで, `ste` はフレームの STE 値を保存する配列である。

また, ZCR は MATLAB では以下のように示す.

```
tmp1 = enframe(x(1:end-1), FrameLen, inc);  
tmp2 = enframe(x(2:end) , FrameLen, inc);  
sgn = (tmp1.*tmp2)<0;  
zcr = sum(sgn, 2);
```

tmp1 と tmp2 は相次ぐフレームを表す. sgn は相次いでいるフレームの零交差率を算出し, zcr はフレームの ZCR 値を保存する配列である.

演算子.\*[30]は要素単位の乗算で,  $A.*B$  は  $A$  と  $B$  の要素単位の積を求めることになる. 演算子コロンは for ループの反復を意味し,  $j:k$  は  $[j, j+1, j+2, \dots, k]$  と同じ結果になる. ここでは tmp1 と tmp2 は 1 サンプル数違いでのフレームを分けになり, 隣接する音声データを示す. sgn はその要素単位の乗算の積の中, 負数のみを行列に保存する. zcr は各行の負数の数, すなわち各フレームの零交差率である.

一定のレベルを越える STE が検出されるとき, 音声始端として認識処理を開始する. そして, STE と ZCR の値が一定以下になったときに, 音声の終端としてそこで入力を区切って 1 発話分の認識を終了する判定であり, 音声認識システムの開発・研究のためのオープンソースの高性能な汎用大語彙連続音声認識エンジン Julius もこの方法を用いて音声の区間を検知している[31].

本研究では, 先ず母音の開始点として STE の高い閾値  $S1$ , また母音の終了点として STE 低い閾値  $S2$  と ZCR の低い閾値  $Z1$  を設定する. ここで, 低い閾値は音声の沈黙または騒音の部分より算出する.  $S1$  と  $Z1$  両方を超えない時は音声あるいは騒音と見なすことができる. また, 母音の開始点を検出するための高い閾値  $S1$  はガウス混合分布モデル[32]などの音響モデルで算出できるが, 複雑な計算および個人差が大きいため, 本研究は音響モデルを省略し, 全フレームの最大 STE 値より  $1/4$  として使用する.

図 2 は図 1 のフレームごとに音節検出処理のアルゴリズムである.

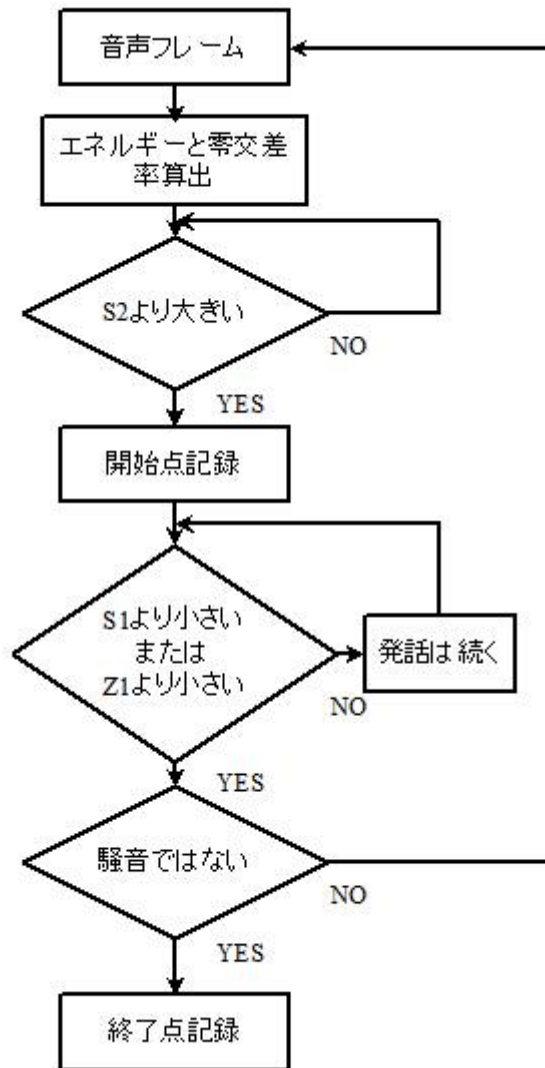


図 2. 発話区間検出のアルゴリズム

フレームの  $ste$  値が高い閾値  $ste\_max$  より高い時は母音があると判断する. しかしここで騒音の可能性があるので, 母音の持続時間を計る. その後処理するフレームの  $ste$  値は  $ste\_min$  と比較し,  $zcr$  値は  $zcr\_min$  と比較し, 高いと判断した時に音節はまだ続いているとする. 逆に, 両方を比較して低いとなった時は音声が終わりにしていると判断する. しかし, ここでは音声短いすぎる場合, 騒音と判断し, 音節が見つかる前の状態に戻る. 騒音閾値について, 自然発話の最小音節持続時間は 0.083 秒[33]だが, 話速が速い時もあると考えられるため, やや短くにして 0.05 秒にした. 音節の長さはフレーム数で数えられ,  $c$  とすると, 騒音判断の式は式 (6) のように表す.

$$c*100/f_s < 0.05 \quad (6)$$

ここで、100 は重複するではないフレーム長で、 $f_s$  は音声周波数の 10000 とすると、 $c$  は 5 フレームになる。つまり、数えられたフレーム数は 5 フレームより小さいと、騒音と判断され、音節とはカウントされないことになる。騒音ではないと判断した時では、音声の発話区間として認識し、発話区間の開始点と終了点及びその間の持続時間を記録する。

### 3.3 母音検出

検出された発話区間に母音検出処理を行う。事前に母音の標準パターンを作成し、各母音の LPC ケプストラム係数  $L1$  を算出する。先程検出された区間の各フレームの LPC ケプストラム係数  $L2$  を算出する。 $L1$  と  $L2$  のユークリッド距離が一定値以下となる場合は、母音が存在すると判断できる。一定区間の中に幾つの母音が存在するかを判断するために、母音終了判断も行う。アルゴリズムは以下の図 3 のように表す。

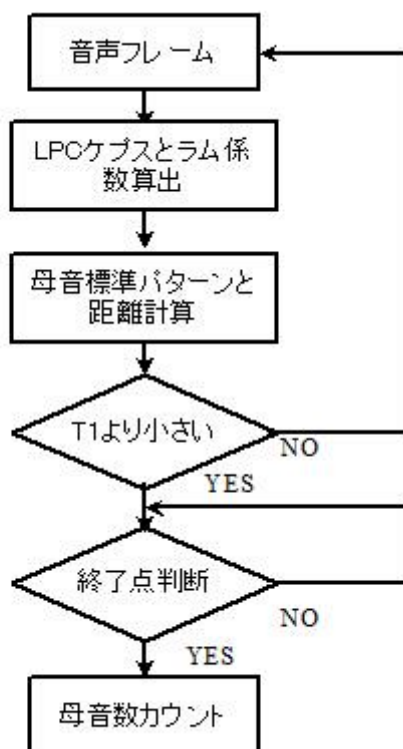


図 3. 母音検出のアルゴリズム

### 3.3.1 フレームの切り出し

発話区間検出と同じく，区間内の音声は均一化処理とプレエンファシス処理をかけてからフレームの切り出し処理を行う．母音検出のフレームにハミング窓を使われる．ハミング窓の係数は以下の式で表す[34]．

$$w(n) = 0.54 - 0.46 \cos\left(2\pi \frac{n}{N}\right), 0 \leq n \leq N \quad (7)$$

ここで， $w$  はハミング窓関数による結果の列ベクトルで， $n$  はフレーム数， $N$  はウィンドウの長さである．MATLAB では  $w=\text{hamming}(N+1)$  として引用することができる

### 3.3.2 LPC ケプストラム係数算出

フレームごとの予測線形係数 (LPC) を計算するため，自己回帰 (AR) モデリングの自己相関法を使用して，フィルター係数を求める[35]．式は以下のように表す．

$$\overline{x(n)} = -a(2)x(n-1) - a(3)x(n-2) - \dots - a(p+1)x(n-p) \quad (8)$$

$p$  は予測フィルター多項式  $a=[1 \ a(2) \ \dots \ a(p+1)]$  の次数で，本研究は 18 次まで用いる． $x(n)$  は線形予測子の係数列である．MATLAB では以下のコード[36]で LPC 係数を計算する．

`a=lpc(x, p);`

`lpc` 関数は過去のサンプルに基づいて実数値時系列  $x$  の現在値を予測する， $p$  次の線形予測子 (FIR フィルター) の係数列  $a$  を出力する．

また，フレーム  $x(n)$  に予測線形係数から LPC ケプストラム係数に変換するため以下の式が使われる[37]．

$$c_0 = \log_e P \quad (9)$$

$$c_m = -a_m + \frac{1}{m} \sum_{k=1}^{m-1} [-(m-k) \cdot a_k \cdot c_{(m-k)}], 1 \leq m \leq p \quad (10)$$

$$c_m = \sum_{k=1}^p \left[ \frac{-(m-k)}{m} \cdot a_k \cdot c_{(m-k)} \right], p < m < n \quad (11)$$

ここで， $a$  は LPC 係数で， $c$  は LPC ケプストラム係数である． $m$  は処理される多項式係数列の項数で， $n$  は音声のフレーム数である．MATLAB は以下のコードで LPC ケプストラム係数



を求める [38].

```
hlp=dsp.LPCToCepstral;  
c=step(hlp,a);
```

LPC ケプストラム係数間のユークリッド距離を計算するため、以下の式が使われる [39].

$$d_{st}^2 = (L2 - L1)(L2 - L1)' \quad (12)$$

ここでは、L2 は会話する時の録音フレームの LPC ケプストラム係数で、L1 は各母音標準パターンの LPC ケプストラム係数である。MATLAB では、 $D = \text{pdist2}(L2, L1)$  より計算することができる。録音フレーム L2 の LPC ケプストラム係数と各母音標準パターンの LPC ケプストラム係数 L2 とのユークリッド距離 D を算出し、最も小さい距離を求める。この距離が閾値 T1 小さい場合は、母音が存在することが判断できる。閾値 T1 は環境騒音と母音標準パターンとのユークリッド距離を算出し、騒音が誤認識されないように、最も小さい距離より 0.1 より小さい距離尺度にした。

### 3.3.3 母音の終了点判断

母音の数をカウントするために、母音が存在することを確定される後、母音の終了点を検出する必要がある。本研究は、以下の 3 条件の場合では、母音の終了点と認識し、母音の数をカウントする：(1). 前フレームと現フレームとも母音を検出され、且つ異なる種類の母音である。(2). 前フレームは母音を検出され、現フレームでは母音を検出されない。(3). 録音音声区間の最後のフレームに母音が存在する、且つ前のフレームとの母音の種類が異なる。以上の条件が満たす場合は、母音の数がカウントすることができる。

## 3.4 発話速度の計算

母音検知で測定される母音数は音節数と引き換えに、発話区間検知で測定される発話時間を除することで、発話速度を計算することができる。速度が閾値 T2 に超えると、システムに警告の文字が現われる。ここで、我々はユーザーに、つまり非母語者に事前に適切な速度を選択するように、T2 を決める。母語者の発話速度は非母語者が選んだ速度より上回ると、母語者が早口と判断し、赤色背景の文字「T00 FAST!!」を表示し、超えていない時は緑色背景の文字「GOOD」と表示する。そのインターフェースは図 4 のように表す。



図 4. システムのインターフェース

### 3.5 データの保存

データを分析処理するため、録音された音声データ、カウントされた母音数および発話区間と発話速度は excel ファイルに書き出し処理を行う。

録音音声は検知処理が終わる次第 wav ファイルとして保存する。また、検知された音節数、発話区間や速度は検知処理が終了するまでのシステム処理時間はテンポラリデータとして保存され、会話が終わる時に外部 excel ファイルに保存する。その上、母語者が母音を発声した音声および母音標準パターンも同じ excel ファイルに保存する。

## 第4章 動作確認実験

提案手法を実装し、発話検知機能の動作確認実験を行った。

### 4.1 実験概要

実験では、スピーカで録音された音声を流し、システムに音節を数えてもらった。使った音声データは高橋研究室が作った同じ人が同じ文章を異なる話速（6.73 モーラ/秒、8.00 モーラ/秒、自然発話速度 8.74 モーラ/秒、9.51 モーラ/秒、11.31 モーラ/秒、13.45 モーラ/秒）で読み上げた音声である。その 5 つの話速の音声を逐一でスピーカで流し、システムに録音させて音節をカウントする動作について確認した。ここで、文章は読売新聞の「編集手帳」（630 字）で、文章を朗読する人は発話プロフェッショナル（PM00）である。

また、システムはシングルプロセスのため、音節をカウントする処理の時間には録音はしていないため、処理時間の長さはシステムリアルタイム性には大きな影響を与えると考えられる。録音音声データは大きいほど、処理時間も長くなるため、最も短い処理時間を見つけるため、録音時間を 1 秒、2 秒、3 秒、5 秒、10 秒で上記音声データを流した。システムがそれぞれカウントした音節数、録音が終わるからカウント処理が終わるまでの処理時間を記録した。

### 4.2 実装環境

本実験で使用するソフトウェア環境として、OS は Windows 7、数値計算言語には MathWorks 社の MATLAB R2015a で、スピーカの型番は Logicool multimedia speakers Z150、マイクは SteelSeries Siberia Microphone white 50002 である。

使われた音声データの周波数は全部 44100Hz で、ステレオ音声である。話速 6.73 モーラ/秒の音声の長さは 162 秒で、話速 8.74 モーラ/秒の音声は 128 秒、話速 9.51 モーラ/秒の音声は 115 秒、話速 11.31 モーラ/秒の音声は 72 秒、話速 13.45 モーラ/秒の音声は 82 秒である。

実験は 29dB の騒音レベルの静かな環境で行われた。スピーカはマイクの直前に置き、流れる音声の音量は 70dB である。

### 4.3 動作確認実験の結果

動作確認実験の結果は表 1 より示す。表 1 では各話速・各録音時間における検出音節数と正解音節数（検出音節数/正解音節数）、各話速・各録音時間における一秒の平均誤差（音節/秒）と、各話速・各録音時間での録音回数での録音回数や各話速・各録音時間における音声データを処理するためにかかる平均処理時間（秒）を示している。

検出音節数はシステムがカウント処理による検出した音節数で、正解音節数はシステムで

録音した音声データを聞いて台本を見ながら手で数えた音節数である。ここで、4.2 節で言及したように、音節をカウントする処理の時間には録音はしていないため、処理時間を除く、録音された内容のみの正解音節数である。

平均誤差とは各話速・各録音時間の（正解音節数-検出音節数）の平均値で、1 秒あたりでの平均誤差とは、その平均値から録音時間で割るものである。つまり  $(|(\text{正解音節数}-\text{検出音節数})|/\text{録音時間})$  の平均値である。

録音回数とは、録音データを流し、実際に録音された回数である。

処理時間とは録音が終わるからデータが保存されるまで経つ時間で、システム内部の時間関数より測定される。一回作動にかかる平均処理時間は各話速・各録音時間で録音する際に、データを処理するにかかる処理時間の平均値である。

右側は全話速が各録音時間での平均値（検出音節数/正解音節数）、一秒の平均誤差（音節/秒）、録音回数（回）、一回録音にかかる平均時間（秒）を表す。

表の見方として、話速 13.45 モーラ/秒の音声データが 1 秒の録音時間で録音すると、検出した音節数は 337 音節であるが、録音データを手動で数えた正解音節数は 430 音節である。1 秒当たりの平均誤差は 2.1 音節/秒で、音声データを録音した回数は 62 回、時間処理にかかるシステム処理時間の平均値は 0.4 秒である。また、1 秒で録音する時に、全話速での平均検出音節数は 440 音節で、録音された音声を手で数えると 445.5 音節数があった。そのため、1 秒あたりの平均誤差は 1.7 音節/秒、平均録音回数は 89.8 回、1 秒の録音データを処理するためにかかる平均時間は 0.4 秒である。

表 1. 動作確認実験の結果

録音時間・各項目		話速(モーラ/秒)						全話速 平均
		6.73	8	8.74	9.51	11.31	13.45	
1	検出音節数/正解音節数	515/456	487/462	450/440	455/443	396/442	337/430	440/445.5
	誤差(音節/秒)	1.5	1.5	1.7	1.6	1.9	2.1	1.7
	録音回数(回)	121	103	94	86	73	62	89.8
	処理時間(秒)	0.34	0.34	0.36	0.34	0.35	0.36	0.348
2	検出音節数/正解音節数	422/488	404/490	343/474	365/487	350/471	289/454	362.2/477.3
	誤差(音節/秒)	1	1.2	1.7	1.4	1.6	2.6	1.6
	録音回数(回)	67	56	52	47	39	33	49.0
	処理時間(秒)	0.42	0.44	0.44	0.45	0.48	0.48	0.45
3	検出音節数/正解音節数	384/510	363/516	334/482	343/490	323/500	280/503	337.8/500.2
	誤差(音節/秒)	1	1.3	1.5	1.6	2.2	3.2	1.8
	録音回数(回)	46	39	36	32	27	23	33.8
	処理時間(秒)	0.52	0.54	0.56	0.59	0.59	0.57	0.56
5	検出音節数/正解音節数	356/510	315/488	289/503	318/488	288/474	258/478	304.0/490.2
	誤差(音節/秒)	1.1	1.5	1.9	1.8	2.3	3.3	2.0
	録音回数(回)	27	23	22	19	16	13	20.0
	処理時間(秒)	0.92	0.92	0.94	1.05	1.15	1.31	1.05
10	検出音節数/正解音節数	284/460	263/455	228/452	245/412	238/422	222/431	246.7/438.7
	誤差(音節/秒)	1.4	1.7	2.2	2.1	2.6	3.5	2.3
	録音回数(回)	13	11	10	8	7	6	9.2
	処理時間(秒)	2.46	2.66	2.64	3.38	3.84	4.10	3.18

## 第5章 評価実験

### 5.1 実験目的

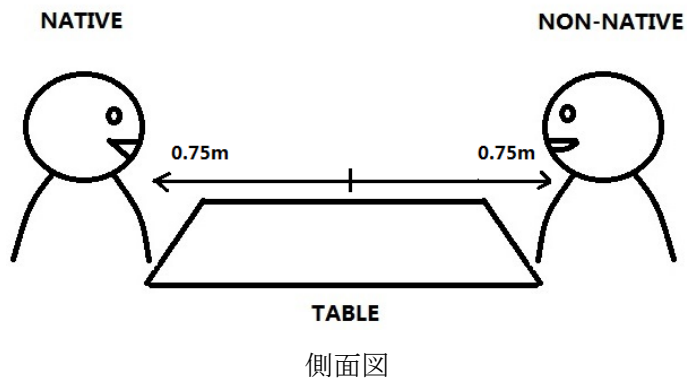
本研究では、非母語者と母語者の自由対面会話場面に、提案する発話自動認識で警告標識を提示することが非母語者の聞き取りやすさについて有効であるかを確認するために、上記機能を実現したシステムを構築し、実験を行った。

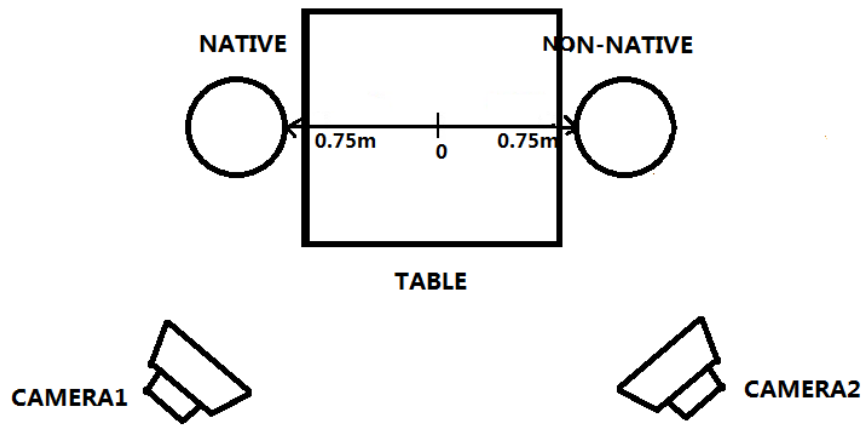
### 5.2 比較条件

システムの有効性を確かめるために、非提示条件とシステムありの条件を分けて2条件を設けた。また、システムありの条件ではシステムを話者提示条件とシステムを母語者と非母語者両方の両者提示条件2パターンを設置した。非提示条件条件、話者提示条件と両者提示条件総計3条件である。

条件	配置	提示対象	配置位置
非提示条件		なし	なし (図5)
話者提示条件		母語者	非母語者の後方、頭より上 (図6)
両者提示条件		母語者と非母語者	母語者と非母語者の間 (図7)

表2. システムの配置





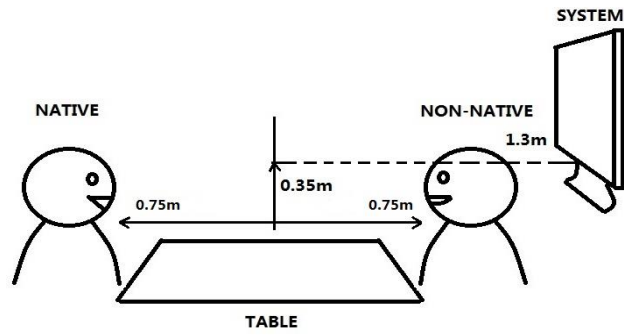
上面図

図 5. 非提示条件の配置図

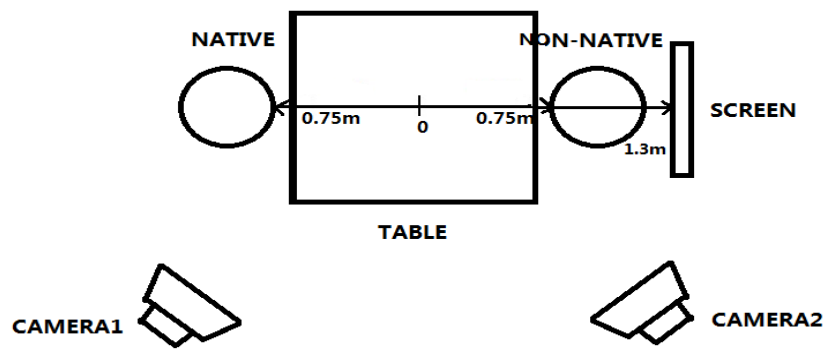
非提示条件の実験様子は図 6 より示す。



図 6. 非提示条件の実験様子



側面図



上面図

図 7. 話者提示条件の配置図

話者提示条件の実験様子は図 8 より示す。



図 8. 話者提示条件の実験様子



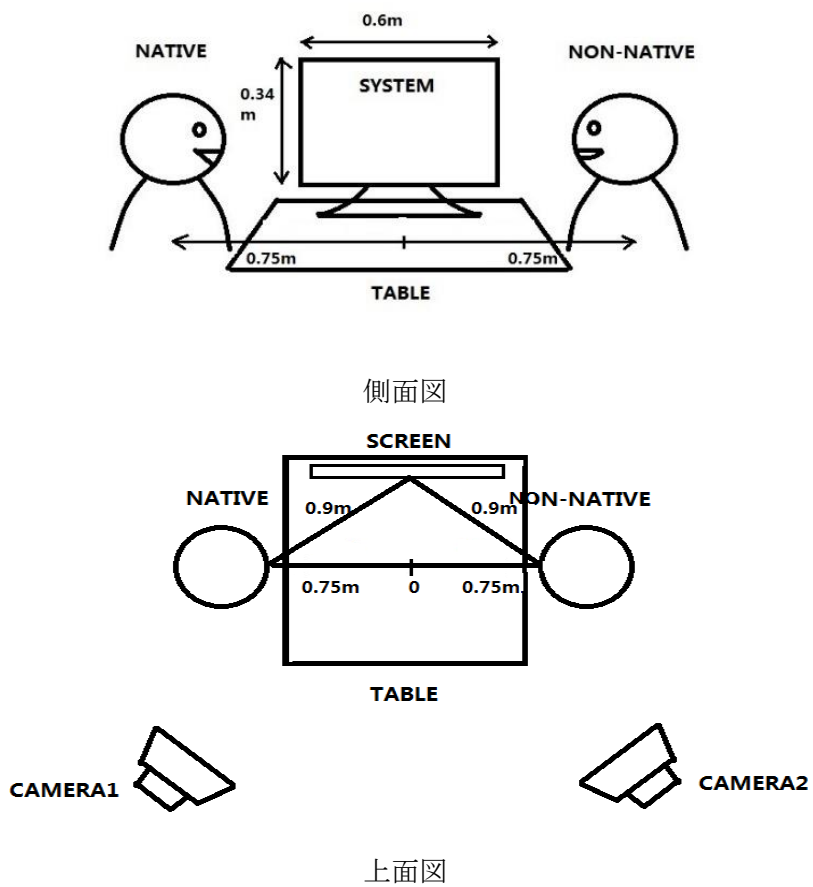


図 9. 両者提示条件の配置図

両者提示条件の実験様子は図 10 より示す。



図 10. 両者提示条件の実験様子

## 5.3 実装環境

本実験で使用するシステム環境として、OSはWindows 7、数値計算言語にはMathWorks社のMATLAB R2015bを用いる。録音は10000Hzのモノラル音声である。

非母語者が適切な発話速度を選ぶ際に使われるスピーカーの型番はLogicool multimedia speakers Z150、スピーカーは非母語者の横にある床に置き、机の高さで聞こえる音声は70dBになるように設定した。実験は29dBの騒音レベルの環境で行われた。母語者が実験中に録音される際に使われるマイクはSteelSeries Siberia Microphone white 50002である。

警告標識を提示するためのモニターはProLite XB2783HSU XB2783HSU-B1 27インチで、画面の解像度は1920×1080である。撮影機械について、2台の型番ともDMX-HD2000、撮影する画像の比率は4:3で、解像度は1280×720である。配置図より母語者と非母語者の側後から撮影する。

## 5.4 実験対象

母語者と非母語者が二人一組で対面会話実験を行うため、母語者は全員日本語を母語として使う者で、非母語者は日本語能力試験N2程度、日本語で日常会話ができる者が対象である。

## 5.5 実験手順

本実験では、提案するシステムが作動する条件（話者提示条件と両者提示条件含む）および対照条件となる非提示条件の両者対面会話の条件で非母語者の聞きやすさと話しやすさに関する影響を調べるために、各実験の条件で会話実験を行った。

### 5.5.1 実験説明

実験の手順として、先ず実験者に実験説明を行った。付録の実験説明用紙1から3を配り、実験内容とシステムの動作を詳しく説明をした。また、実験参加同意書を読み上げ、同意を得てから記入をしてもらった。

### 5.5.2 発話速度閾値の設定

非母語者に適切な発話速度を設定ように、事前に作られた音声を聞かせる。インターフェースは以下の図10のように表す。左のボタンAから右のボタンJを押すと3音節/秒から12音節/秒までの1音節/秒ずつ発話速度が増加するように音声サンプルを作成した。音声サンプルの基は高橋研究室[40]より発話プロフェッショナル (PM00) が8モーラ/秒話速で

朗読した音声である。各ボタンを押して異なる音声を再生できるように、まずは 12 音節/秒のサンプルを作成するために該当音声サンプルを 8 つの部分に分け、120 音節の内容をカットし、発音時間を 10 秒に変更した。同様に、8 部分の音声を 110 音節の部分を取り、発音時間を 10 秒に変更し、11 音節/秒のサンプルを 8 つ作成した。このような方法で、3 音節/秒から 12 音節/秒までの 1 音節/秒ずつ発話速度の音声サンプルを計 80 つの音声サンプルが用意した。また、同じボタンを複数回押した時に乱数を生成し、同じ音声サンプルを再生しないように設定した。

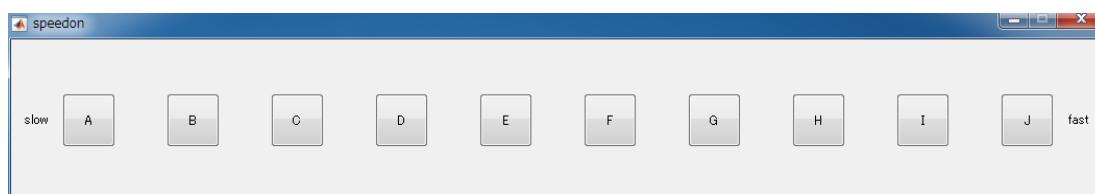


図 11. インターフェース 1 (発話速度閾値の設定)

本実験は、非母語者に自由にボタンを押させ、速いと思った発話速度を判断して警告提示の閾値としてシステムの中に入力した。

#### 5.5.3 母音標準パターンの作成

母音標準パターンを作るために、母語者に母音「a」「i」「u」「e」「o」を三回発音してもらい、第 3 章の方法より各母音の LPC ケプストラム係数を計算し、外部 excel ファイルに保存した。インターフェースは図 12 のように表す。REC ボタンを押すと 15 秒間録音を始め、母音者の発音を録音することができる。録音が完了したときには右の空欄に「OK」という文字が現れる。また、DO ボタンを押すと録音された音声データを使って母音標準パターン、つまり各母音の LPC ケプストラム係数を計算し、計算が完了する時には同様に右の空欄に「OK」と文字が現れる。作られた母音標準パターンは、本人の会話実験中の音声フレームの LPC ケプストラム係数とのユークリッド距離を算出ために使われることになる。

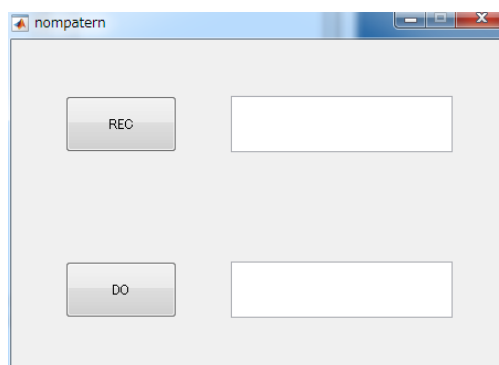


図 12. インターフェース 2 (母音標準パターン作成)

#### 5.5.4 会話実験

上記設定を終わると、対面会話実験を始める。本実験では、母音者が自然会話中に発話速度をあげるような場面を想定し、時間制限と目標が設けられたお題当てゲームを会話内容として実験を行った。ゲームのルールは1969年NHK番組の連想ゲーム[41]より参照し、キーワード1個を当てると10点を獲得、違反ヒントを出すと5点が減点されるとした。150点以上獲得できると謝礼金を増えることを伝えた。5分間に高い得点を獲得するために、母語者は非母語者にできる多くのキーワードを当てさせるために、発話速度が上げる可能性を見込んでいる。また会話ゲームのため、ジャスチヤは禁止とし、全部言葉で表現することで発話速度が非母語者の理解への影響を分析する。

本実験は提案システムが会話の影響を調べるため、お題当てゲームを3回行った。実験への参加順は表1のようにランダムにし、順序効果を相殺したためである。表の見方として、1組の実験者は先ず両者提示条件条件で実験を行った、次に非提示条件条件で行った。最後では話者提示条件条件で実験を行った。なお、システムあり話者提示条件条件及び両者提示条件条件は母語者の発話速度を検知するため、母語者のみがマイクを装着させた。

対話実験の前では、母語者にキーワードを記載されたメモ用紙を渡し、1分間準備をさせる時間を設けた。それから5分間のお題当てゲームを行った。キーワードメモ用紙には一枚24つのキーワードが記載されていて、母語者はゲーム中にもメモを見ることができ、各条件には異なるメモ用紙でゲームをする。キーワードは全部日本語の教科書「みんなの日本語」初級[44]から記載された名詞の単語リストである。全部で4枚のメモ用紙が作られ、実験ではその中にランダムで3枚が使われる。その詳細は付録4に示している。

ゲーム中に、実験実施者は同一室で単語リストを見ながら、正解の内容と数を記録し、採点を行った。非母語者がキーワードを当てた時に10点を加点し、母語者が違反ヒントを出すと5点を減点し、各条件の成績を記録した。また、ビデオ分析するため、お題当てゲームの5分間のみを撮影した。

各条件が終りに、母語者と非母語者に質問紙を渡して記入してもらった。

表 1. 実験者参加順

条件 参加者ペア	非提示条件	話者提示条件	両者提示条件
1	1	2	3
2	3	2	1
3	1	3	2
4	1	2	3
5	2	1	3
6	1	3	2
7	2	1	3
8	3	2	1
9	2	3	1
10	2	1	3
11	3	2	1
12	3	1	2
13	2	1	3
14	3	1	2
15	1	3	2
16	2	3	1
17	2	3	1
18	3	1	2

#### 5.5.5 データ取得

##### (1) 主観評価

従来、異文化コミュニケーションに関するシステムについての主観評価は主に、発者が内容を把握できた度合い、発者が言いたいことを発話できた度合い及び会話が盛り上がる度についてのアンケート項目[43][44]、またはシステムの精度、システムの関与性及び情報の提示方法についてのアンケート項目[45]を設けられている。

本研究は、支援対象の非母語者に関する項目だけではなく、システムの影響による母語者の意識とストレスも分析対象とし、非母語者と類似する各実験条件の会話についての感想、発話のしやすさと聞き取りやすさおよびシステムに関する印象と有効性について質問項目を設けた。非提示条件の条件では、会話についての質問項目1「会話は楽しかった」、

項目 2「発話速度にストレスを感じた」、項目 3「相手の話をよく聞き取れた」、項目 4「自分はよく話せた」と項目 5「自分の話が速いと感じたことがあった」の 5 つの項目を設けている。

システムありの条件では、上記の質問以外にシステムに関する項目 6「話が速すぎるという表示に気がついた」、項目 7「話が速すぎるという表示のタイミングは適切だった」、項目 8「話が速すぎるという表示によって自分はゆっくり話すようになった」、項目 9「システムは会話の妨げになった」、項目 10「システムの画面は見やすかった」5 項目を加えた。

また、非母語者に渡す質問用紙には、項目 5 は「相手の話が速いと感じたことがあった」、項目 8 は「話が速すぎるという表示によって相手はゆっくり話すようになった」に母語者の発話速度に関する非母語者の意識について質問を変えている。

各項目には「全く当てはまらない」、「当てはまらない」、「あまり当てはまらない」、「どちらともいえない」、「ややあてはまる」、「当てはまる」と「よく当てはまる」7 段階尺度で選択肢が設けられている。条件ごとに配る質問紙の項目は以下の表 2 のように表す。

表 2. 条件ごとに配る質問紙の項目

条件 \ 対象	母語者	非母語者
非提示条件	項目 1 から項目 5	項目 1 から項目 5
話者提示条件	項目 1 から項目 10	項目 1 から項目 5
両者提示条件	項目 1 から項目 10	項目 1 から項目 10

## (2) 客観評価

本実験では、質問紙調査の主観評価以外、客観評価も合わせてシステムと会話を評価する。

システムの認識率として、録音された音声データを再生し、発話内容を書きだして母語者が発声した正解音節数を数えた。誤差率は（正解音節数-認識音節数）/正解音節数から計算した。

また、システムの影響より母語者の発話速度の変化について分析するため各条件の母語者の発話速度の平均を求めた。

システムの存在が会話の理解に影響を与えるかを分析するため、お題当てゲームの成績の平均も求めた。

## 第6章 実験結果

母語者と非母語者が二人一組総計 18 組 (36 人, 男性 17 人, 女性 19 人) で会話実験を行った。全ての参加者は初対面である。全員の平均年齢は 23.1 歳で, 標準偏差は 2.6 である。

非母語者の平均在日時間は 11.0 ヶ月である。実験に参加した謝礼として最低 500 円分のアマゾンカードを配布した。

### 5.2 主観評価結果

各条件母語者と非母語者に質問紙を記入してもらった結果, 計 108 部の質問紙を回収した。前章より記述したように質問項目は 7 段階で設けている。それぞれ 1 から 7 点を対応付け, 母語者と非母語者別で質問紙の平均を求めた。質問紙の結果は付録 7 より添付した。

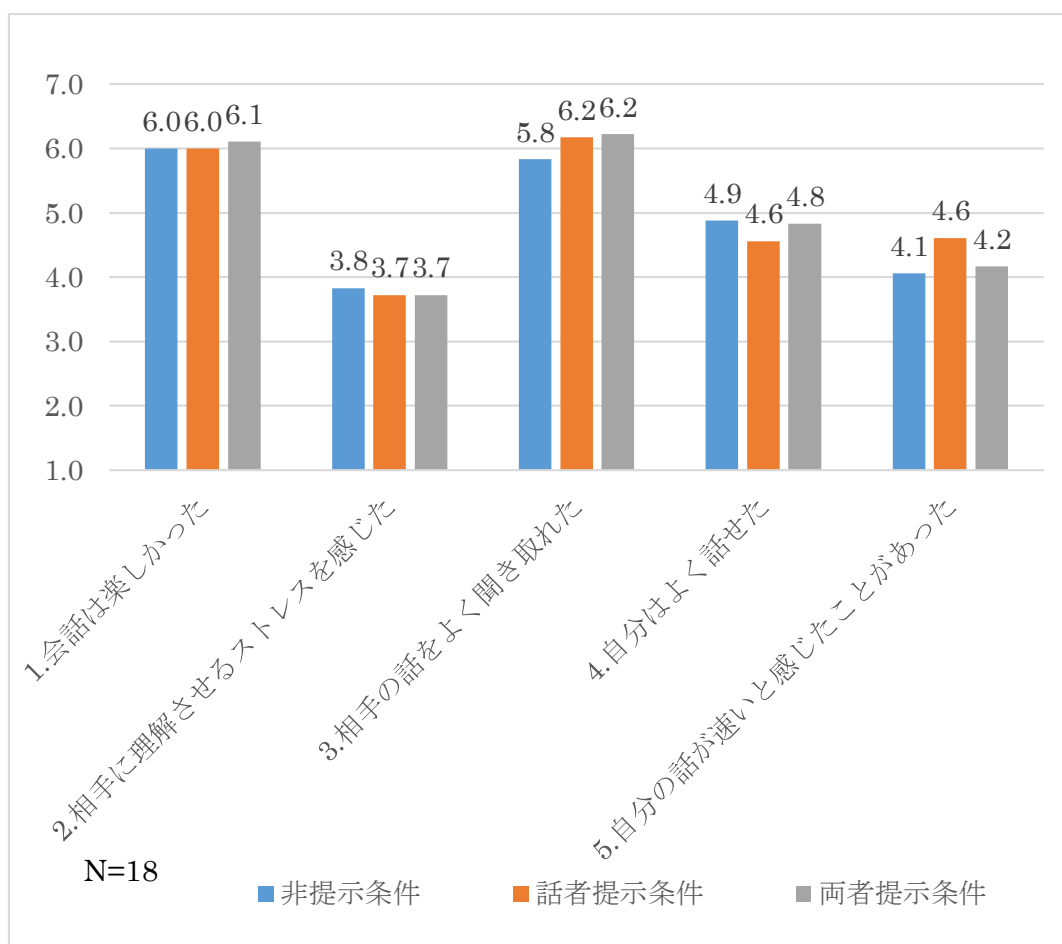


図 13. 各条件の質問紙項目 1 から 5 (母語者) の比較



母語者が記入した質問紙項目 1 から項目 5 の平均結果は図 13 のように表す。図の見方として、縦軸は質問紙項目の平均点で、横軸は各質問紙の項目内容である。

結果の例として、質問紙項目 1 「会話は楽しかった」に対して非提示条件では、母語者が選んだ平均点は 6.0 点で、話者提示条件の場合、平均点は 6.0 である。両者提示条件の場合、平均点は 6.1 である。三条件を一元配置分散分析で分析した結果、有意差は見られなかった。

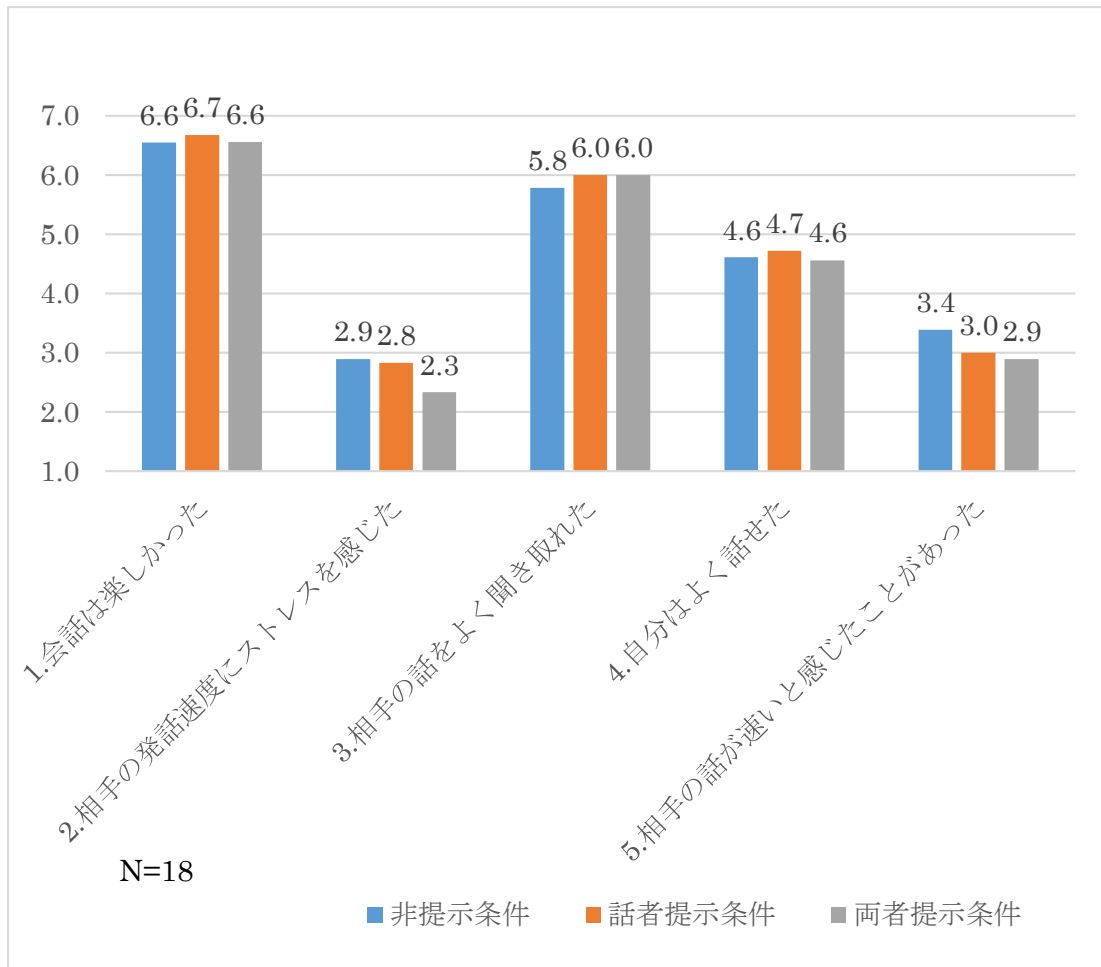


図 14. 各条件の質問紙項目 1 から 5 (非母語者) の比較

非母語者が記入した質問紙項目 1 から項目 5 の平均得点は図 14 のように表す。縦軸は質問紙項目の平均点で、横軸は各質問紙の項目内容である。

結果の例として、質問紙項目 1 「会話は楽しかった」に対して非提示条件では、非母語者が選んだ平均点は 6.6 点で、システムあり話者提示条件の平均点も 6.7 である。システムあり両者提示条件の平均点は 6.6 である。三条件を一元配置分散分析で分析した結果、有意差は見られなかった。

非提示条件では母語者が記入した質問紙に項目 5 から 10 がないため、話者提示条件と両者提示条件の平均点を求めた。また、非提示条件および話者提示条件では非母語者が記入した質問紙に項目 5 から 10 がないため両者提示条件のみの平均点を求めた。結果は表 3 のように表す。

表 3. 質問紙項目 6 から 10 の結果

実験対象・条件 質問紙項目内容	母語者		非母語者
	話者提示条件	両者提示条件	両者提示条件
6. 話が速過ぎるという表示に気が付いた	5.3	5.5	4.7
7. 話が速過ぎるという表示のタイミングは適切だった	4.6	4.8	4.7
8. 話が速過ぎるという表示によって自分はゆっくり話すようになった (母語者)	5.2	5.6	
8. 話が速過ぎるという表示によって相手はゆっくり話すようになった (非母語者)			5.3
9. システムは会話の妨げになった	3.2	3.2	2.7
10. システムの画面は見やすかった	5.9	5.7	5.6

(n=18)

表の見方として、左側は質問紙項目内容で、右側はそれぞれ母語者と非母語者が選んだ項目 6 から 10 の結果である。

## 5.3 客観評価結果

### 5.3.1 システムの認識率

本研究は、システムの警告画面が適切に表示できたかを評価するためにシステムの認識率を計算した。システム認識率は  $1 - \text{誤差率}$  より計算し、誤差率は  $|( \text{正解母音数} - \text{認識母音数} )| / \text{正解母音数}$  で計算する。認識母音数は実験中にシステムが自動的に保存したもので、正解母音数は録音データを聞いて、手動で数えた母音数である。その結果、本実験の認識率は 54.5% (N=12) である。

### 5.3.2 システム平均処理時間

システムの警告画面は適切なタイミングで表示できたかを評価するために、システムの平均処理時間を求めた。図 15 のように、リアルタイムで標識を表示できるように、システムは図 1 のアルゴリズムのように録音と処理を繰り返して作動する。システムが繰り返して作動するには録音時間と処理時間を含んでいる。処理時間は (作動時間 - 録音時間) で計算する。また、録音時間は動作確認実験で得られた最も高い認識率の 2 秒を用いた。その結果、本実験の平均処理時間は 0.22 秒である (N=18)。

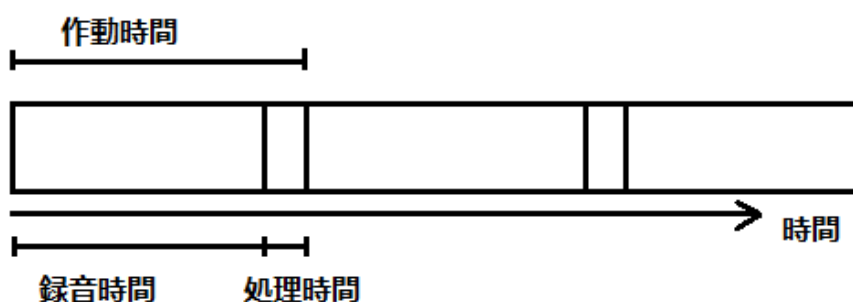


図 15. システム処理時間

### 5.3.3 平均発話速度

母語者の発話速度は（音節数/発話区間時間）で計算しているため，母語者の発話をビデオより記録されるため，本研究はビデオの音声を聞き，母語者と非母語者の発話を書き出して EXCEL に保存した．ビデオ分析の作業は ELAN[46]を使って会話内容と発話区間をマッチングし，発話区間の統計を求めた．ここで，2.3 節で記述したように，音節は撥音（ん），促音（小さい「っ」），長音（ー）は一音節として見られないため，口頭発話にあるあいづちやフィラー，意味のない発話は発話音節数から除外した．これで正確な音節数と一致したと考える．母語者の発話速度は母語者が発声する音節数を発話区間を除するものである．

1 ペアから 12 ペアまでの母語者の発話速度は表 4 のように表す．表の見方として，1 ペア目の母語者非提示条件での平均発話速度は 5.4 音節/秒で，話者提示条件の平均発話速度は 7.0 音節/秒．両者提示条件では平均発話速度は 6.0 音節/秒である．

表 4. 各母語者の平均発話速度

参加者 ペア \ 条件	非提示条件	話者提示条件	両者提示条件
1	5.4	7.0	6.0
2	6.5	5.9	5.8
3	4.1	3.7	3.7
4	4.8	4.6	4.8
5	5.2	5.0	5.1
6	5.7	5.6	5.7
7	4.8	5.2	4.6
8	5.4	5.4	5.5
9	5.0	6.3	4.8
10	5.2	6.0	4.8
11	6.1	5.8	6.0
12	4.5	4.8	4.2

(音節/秒)

### 5.3.4 非母語者の発話速度設定値

非母語者の実験設定部分で事前に準備した音声サンプルを聞き、速いと思った発話速度は表5のように表した。その平均値は8.6音節/秒である。

表 5. 非母語者が設定した発話速度の閾値

参加者 ペア	条件	母語話者設定発 話速度の閾値
	1	9
	2	8
	3	11
	4	8
	5	8
	6	8
	7	6
	8	9
	9	7
	10	10
	11	7
	12	9
	13	8
	14	10
	15	9
	16	9
	17	11
	18	8
	平均値	8.6

(音節/秒)

### 5.3.5 成績平均点

システムが会話にどのような影響を与えるかを分析するために、各条件の会話ゲームの成績を記録した。結果は表5のように表す。表の見方の例として、1組は非提示条件条件での点数は40点で、話者提示条件の条件では100点に上った。さらに、システムあり両者提示条件の条件では125点を獲得した。各条件の成績は正規分布とは言えないデータであったため、三条件の結果の比較に一元分散分析検定を用いた。この結果、各条件の成績には有意差が認められなかった。

表 6. 全実験のゲーム得点

条件 ペア	非提示条件	話者提示条件	両者提示条件
1	40	100	125
2	80	120	65
3	110	80	70
4	130	100	50
5	80	150	120
6	40	80	75
7	100	135	120
8	90	75	60
9	150	165	155
10	155	155	130
11	35	100	110
12	140	135	130
13	90	100	120
14	220	140	175
15	140	90	50
16	50	50	80
17	70	60	80
18	95	40	120
平均得点	100.8	104.3	101.9

## 第7章 考察

本研究は異言語コミュニケーション支援を目的とした発話速度の自動検知で早口を母語者に知らせるウェアラブルシステムを構築し、システムが非母語者の理解、及び会話にどの影響を与えるかを調べるために、評価実験を行った。

評価実験の結果では、分析より有意差は認められなかった。その原因は二つがあると考えられる。一つは、母語者は発話速度が速いという表示を気づかない。話者提示条件の質問項目6「話が速過ぎるという表示に気づいた」について、18人の母語者の内、5人が1「全く当てはまらない」、2「当てはまらない」、3「あまり当てはまらない」を選択した。両者提示条件の質問項目6「話が速過ぎるという表示に気づいた」について、18人の母語者の内、3人が1「全く当てはまらない」、2「当てはまらない」、3「あまり当てはまらない」を選択した。また、表示を気づいた母語者は質問項目8「話が速過ぎるという表示によって自分はゆっくり話すようになった」について、話者提示条件では2人が1「全く当てはまらない」、2「当てはまらない」、3「あまり当てはまらない」を選択した。両者提示条件では1人が1「全く当てはまらない」、2「当てはまらない」、3「あまり当てはまらない」を選択した。表示をきかない原因としては、被験者に何う結果、ゲームに集中して高得点を得られるため、あまり画面を見ていなかったという意見が多かった。

二つ目の原因は非母語者は自分の日本語能力に過剰評価したため、速い発話速度を閾値として設定した。会話中母語者の発話速度は閾値を越えないため、速いという警告は表示しない。閾値は被験者各自の日本語能力によるため、同一の値ではないが、平均的に8.6音節/秒の結果があった。しかし、10音節/秒以上を設定した非母語者は4人いる。例えば、第17ペアの非母語者は11音節/秒を閾値として設定した。しかし、日本人の最も聞きやすい閾値は8.8モーラ/秒[47]。音節はそれ以下になると考える。その値より遙か超えている。そのため、表示は表さない。非母語者が非提示条件と提示条件で質問項目5「相手の話が速いと感じたことがあった」について6「当てはまる」と選択した。適切な発話速度を設定する方法を見直す必要があると考える。

しかし、システムが母語者と非母語者両方に発話速度が速いを提示する場合、非母語者は母語者の発話速度に感じるストレスがシステムが非提示する場合の平均点より軽減する結果が見られた。また、母語者と非母語者はシステムが非提示する場合より提示する場合のほうが相手の話をよく聞き取れるについて平均点も高い結果があった。

一方、非母語者は「相手の話が速いと感じたことがあった」についての平均点と「相手の話が速いと感じたことがあった」についての平均点はシステムの存在より低下することが見られた。

質問紙項目の平均点から、内容を把握できた度合いについての項目3「相手の発話をよく聞き取れた」に関しては、母語者と非母語者も平均点6「当てはまる」以上を選び、よく相



手の話聞き取れた結果があった。ここで、非提示条件と提示条件では、非母語者と母語者両方の聞き取れる度合いが上がったことが見られた。発話者が言いたいことを発話できた度合いについての項目4「自分はよく話せた」に関して、母語者と非母語者の得点はおよそ5「やや当てはまる」になった。今回の実験では母語者が言葉を説明し、非母語者に理解させるかの会話場面を想定して設定したが、非母語者も母語者と同様によく話せたという結果があった。

システムに関する評価について、質問紙項目6「話が速過ぎるという表示に気が付いた」に関して、母語者の回答も平均5「やや当てはまる」以上に上回り、母語者はよく警告標識を気が付いたと考える。一方、非母語者は平均4.7点で、5「やや当てはまる」には至らなかった。これに関して、標識を気づいていない非母語者に聞いた結果、会話内容に集中し、相手の顔を見ているため、横にあるシステム画面は気にしなかったとのことである。項目7「話が速過ぎる表示のタイミングは適切だった」について、母語者と非母語者両方も平均的に4「どちらでもない」以上を選び、およそ5「やや当てはまる」結果があった。システムの認識率は評価できなかった。標識が気づいてないより評価はできない。または標識を気づくがタイミングがよく分からなかった意見があった。母語者の項目8「話が速過ぎるという表示によって自分がゆっくり話すようになった」と非母語者の項目8「話が速過ぎるという表示によって相手がゆっくり話すようになった」について、両方とも5「やや当てはまる」を選ぶ、場合によってゆっくり話すようにするが、時々無視することもある結果があった。項目9「システムは会話の妨げになった」について、母語者と非母語者も3「あまり当てはまらない」、システムの存在は妨げではないという結果があった。最後に画面の見やすさの項目10「システムの画面は見やすかった」について、母語者と非母語者もおおよそ6「当てはまる」を選び、システムの設定位置は見やすかったという結果である。

システムの認識率について、二矢田らの実験では音声パターン認識の距離尺度を比較した結果、LPC ケプストラム係数のユークリッド距離は騒音環境の場合、認識率は60%[48]以下であることに一致した。この結果は質問紙7「話が速過ぎるという表示のタイミングは適切だった」に母語者と非母語者が選んだ4「どちらでもない」から5「やや当てはまる」にあたり、認識率はあまり高くないが、やや当てると結果と一致した。母音の認識率は自由発話かつ騒音によって低下するほか、母音の無声化現象[49]に影響される。その母音の無声化とは、音韻論的に考えれば母音があるはずのところ、当該母音が声帯の振動を伴わずに発音される現象である。その上、発話速度が速い時に、「アエ」、「アイ」、「アウ」、「アオ」、「ウエ」、「ウイ」、「オエ」、「オイ」、「オウ」など二重母音の認識率も低いため、今後、自由発話の場面かつ騒音がある場合でも頑丈な認識率で認識することができる方法を改善すると考える。

システムの処理時間について、システムは動作確認実験で最も誤差が小さい録音時間。2秒録音をし、平均処理時間は約0.2秒で、合計で2.2秒でリフレッシュ動作をする。人類の

知覚極限は0.2秒であり、無視することができると思う。

この実験では、母語者の母音標準パターンは本人の声で作っており、不特定者では対応できない。また、不特定者の認識に当たって最低でも50人から60人[48]の母音標準パターンを集める必要がある。今後は、必要な数の母音標準パターンを集めて不特定者の発話認識に対応する予定である。

以上の結果から、本実験で提案した話速アウェアネスシステムは母語者の発話速度が速過ぎる時に標識を表示することができた。しかし、システムの認識率は高くないため、騒音に頑丈な認識方法を改善する必要があると思う。また、早口の標識によって母語者が確実に対応を応じてもらえる画面の設計も期待している。その上、適切な発話速度を閾値を設定するため、非母語者は会話の途中でも発話速度の閾値を調整できるようにすることや、母語者の発話速度の変化を検知し、速度が速過ぎる前に警告画面を表示することも期待できると考える。

## 謝辞

本修士論文は、筆者が筑波大学大学院図書館情報メディア研究科図書館情報メディア専攻博士前期課程において、井上研究室において行った研究をもとめたものである。本研究に関して終始ご指導ご鞭撻を頂きました指導教員井上智雄教授に心より感謝致します。また、本論文をご精読頂き有用なコメントを頂きました副指導教員三河正彦准教授に深謝致します。

また、最後まで一緒に頑張って来た同研究室の皆様、及び実験に協力していただいた方々に心より感謝しております。ありがとうございました。

## 参考文献

- [1] 2015 年 訪日外客数（総数） 出典：日本政府観光局（JNTO） <[http://www.jnto.go.jp/jpn/reference/tourism\\_data/visitor\\_trends/](http://www.jnto.go.jp/jpn/reference/tourism_data/visitor_trends/)>
- [2] 柳町智治, 長野克則, 繪内正道, 馬場直志, “外国人留学生による日本語コミュニケーションにおける問題点と改善策”, 工学教育 61(4), pp. 43-47, Jan. 2013.
- [3] 佐久間重, “異文化コミュニケーションの様々な側面: 言語以外の要素について”, 名古屋文理大学紀要 3, pp. 13-21, Apr. 2003.
- [4] 一二三朋子, “母国語者と非母国語者との会話における母国語者の意識的配慮の検討”, 日本教育心理学会教育心理学研究 vol. 43, pp. 277-286, Sep. 1995.
- [5] 丸島歩, “速度変化をとまなう音声の速度感とその規定要因”, 実験音声学・言語学研究 (Research in Experimental Phonetics and Linguistics) 4, pp. 1-21 2012.
- [6] 伊福部達, “高齢難聴者のための音声情報変換方式”, 電子情報通信学会誌 84(5), pp. 325-328, May. 2001.
- [7] 後藤真孝, 北山広治, “音声スポット: 人間同士の会話中に音声認識が利用可能な音声入力インターフェース”, 情報処理学会論文誌 48(3), pp. 1274-1283, Mar. 2007.
- [8] 岡本健吾, 吉野孝, “会話中の名詞の関連情報を用いた対面型異文化間コミュニケーション支援システムの構築と評価”, 情報処理学会論文誌 52(3), pp. 1213-1223, Mar. 2011.
- [9] 福島拓, 吉野孝, 喜多千草, “共通言語を用いた対面型会議における非母語者支援システム Panelive の構築”, 電子情報通信学会論文誌. D, 情報システム J92-D(6), pp. 719-728, Jun. 2009.
- [10] 東北大学 電気通信研究所:  
<<http://www.ais.riec.tohoku.ac.jp/lab/wordlist/index-j.html>>.
- [11] 平田 真, 森本 政之, 佐藤 逸人, “音声の発話速度が高齢者の聴き取りにくさに及ぼす影響”, 日本建築学会近畿支部研究報告集. 環境系 (47), 41-44, May. 2007.
- [12] 磯野 春雄, 山口 淳平, “ニュース音声の話速に対する心理・生理的反応”, 聴覚研究会資料 39(1), 43-48, Feb. 2009.
- [13] 外川 太郎, 大谷 猛, 鈴木 香緒里, “受話音を聞きやすくする音声強調技術/話速変換技術”, 電子情報通信学会誌 96(11), 874-881, Nov. 2013.
- [14] 清山 信正, 今井 篤, 三島 剛 [他], 都木 徹, 宮坂 栄一, “高品質リアルタイム話速変換システムの開発”, 電子情報通信学会論文誌. D-II, 情報・システム, II-パターン処理 J84-D-II(6), 918-926, Jun. 2001.

- [15]津村 光美 , 田中 章浩 , 坂本 修一 [他] , 鈴木 陽一, “話速変換単語音声聴取に非同期者映像がもたらす影響”, 電子情報通信学会技術研究報告. HIP, ヒューマン情報処理 105(479), 103-108, 2005-12-08.
- [16]斎藤純男, “音声学入門”, 三省堂; 改訂版, ISBN-10: 4385345880(2006-03).
- [17]大辞林特別ページ, 日本語の世 5 日本語の音, 大辞林第三版: <<http://daijirin.dual-d.net/extra/nihongoon.html> (2015-08-17) >.
- [18]木田章義, “音節とモーラ”, 音声研究 2(1), pp. 41-49, Apr. 1998.
- [19]緒方淳, 有木康雄, “日本語話し言葉音声認識のための音節に基づく音響モデリング”, 電子情報通信学会論文誌. D-II, 情報・システム, II-パターン処理 J86-D-II(11), pp. 1523-1530, Nov. 2003.
- [20]岡本佳太, 荻島祐一, 田向権 [他], 関根優年, “音声認識システムにおける音素前処理回路”, 電子情報通信学会技術研究報告. CAS, 回路とシステム 112(418), pp. 143-148, Jan. 2013.
- [21]吉原 亨 , 蔦木圭悟 , 高橋 弘太, “音声の高速再生のための話速推定法と高速発話時の特性解析 話速バリエーション型データベースの活用例”, 情報処理学会研究報告 音声言語情報処理 (SLP) 2008(123(2008-SLP-074)), 233-238, Dec. 2008.
- [22]二矢田 勝行 , 藤井 諭 , 森井 秀司, “不特定者を対象とした音声認識のための特徴パラメータと距離尺度に関する考察”, 電子通信学会論文誌 A 69(5), p628-636, May. 1986.
- [23]MATLAB R2015a, MathWorks 社:<<http://jp.mathworks.com/>>
- [24]梢奇方, 島村徹也, 高橋淳一, 鈴木誠史, “線形予測分析に基づくホルマント周波数抽出の雑音耐性の改善”, 電気情報通信学会誌, Vol. J84-A, No. 6, pp. 745-758, 2001.
- [25]MATLAB filter 関数: 1 次元のデジタル フィルター  
<<http://jp.mathworks.com/help/matlab/ref/filter.html>>
- [26]HARK クックブック, 窓長とかシフト長の適切な値が知りたい: <[http://winnie.kuis.kyoto-u.ac.jp/members/okuno/Lecture/13/SpeechProcAdv/usb\\_root.20131204/html/hark-cookbook-ja/index.html](http://winnie.kuis.kyoto-u.ac.jp/members/okuno/Lecture/13/SpeechProcAdv/usb_root.20131204/html/hark-cookbook-ja/index.html)>
- [27]M. H. SAVOJI, “A robust algorithm for accurate endpointing of speech signals”, Speech Communication 8 pp. 45-60, 1989.
- [28]J. Sangeetha, S. Jothilakshmi, “Robust Automatic Continuous Speech Segmentation for Indian Languages to Improve Speech to Speech Translation”, International Journal of Computer Applications(0975-8887) Volume 53-No.15, September 2012.
- [29]MATLAB sum 関数: 配列要素の和  
<<http://jp.mathworks.com/help/matlab/ref/sum.html>>

- [30]MATLAB 配列演算と行列演算 :  
 <[http://jp.mathworks.com/help/matlab/matlab\\_prog/array-vs-matrix-operations.html](http://jp.mathworks.com/help/matlab/matlab_prog/array-vs-matrix-operations.html)>
- [31]Julius, 大語彙連続音声認識エンジン: <<http://julius.osdn.jp/>>
- [32]MATLAB 混合ガウス モデルによるクラスタリング :  
 <<http://jp.mathworks.com/help/stats/gaussian-mixture-models.html>>
- [33]佐藤虎男, ” 音節の持続時間と文法” , ISSN: 02882159. NCID: AN00039331. 学大国文.  
 巻: 34., pp1-18, Feb. 1991.
- [34]Oppenheim, Alan V., Ronald W. Schaffer, and John R. Buck. Discrete-Time Signal Processing. Upper Saddle River, NJ: Prentice Hall, 1999, p. 468.
- [35]Jackson, L. B. Digital Filters and Signal Processing. 2nd Edition. Boston: Kluwer Academic Publishers, 1989, pp. 255-257.
- [36]MATLAB lpc 関数: 線形予測フィルター係数  
 <<http://jp.mathworks.com/help/signal/ref/lpc.html>>
- [37]Papamichalis, Panos E. Practical Approaches to Speech Coding. Englewood Cliffs, NJ: Prentice Hall, 1987.
- [38]MATLAB LPC to/from Cepstral Coefficients  
 <<http://jp.mathworks.com/help/dsp/ref/lpctofromcepstralcoefficients.html>>
- [39]MATLAB 観測値の 2 つの集合間のペアワイズ距離:  
 <<http://jp.mathworks.com/help/stats/pdist2.html>>
- [40]高橋研究室, 電気通信大学, 話速バリエーション型音声データベース: <<http://www.it.ice.uec.ac.jp/SRV-DB/>>
- [41]NHK クイズ番組連想ゲーム (1969~91年) :  
 <<https://www.nhk.or.jp/archives/hakkutsu/features/201409/>>
- [42]スリーエーネットワーク 編, 「みんなの日本語初級 I 本冊」, ISBN : 978-4-88319-102-4
- [43]山下 直美, エチエニーケ アンディ, “国際電話会議の負担を軽減する手法の提案と評価”, 情報処理学会論文誌 Vol. 54 No. 6 1794-1806, 2013. 07
- [44]仲嶺 真, 大坊 郁夫, “2 者間会話における非言語行動と場の活性度認知との関連”, 信学技報, 223-228, May. 2013
- [45]岡本健吾, 吉野孝, “会話中の名詞の関連情報を用いた対面型異文化間コミュニケーション支援システムの構築と評価”, 情報処理学会論文誌 Vol. 52 No. 3, 1213 - 1223, Mar. 2011
- [46]ELAN : <https://tla.mpi.nl/tools/tla-tools/elan/>
- [47]磯野 春雄, 山口 淳平, ニュース音声の話速に対する心理・生理的反応, 電子情報通信

学会技術研究報告. WIT, 福祉情報工学 108(435), 43-48, 2009-02-13(聴覚研究会資料 39(1), 43-48, 2009-02-20

- [48] 二矢田 勝行, 藤井 諭, 森井 秀司, “不特定者を対象とした音声認識のための特徴パラメータと距離尺度に関する考察”, 電子通信学会論文誌 A 69(5), p628-636, May. 1986
- [49] 三松国宏, 福盛貴弘, 菅井康祐, 宇都木昭, “日本語の母音の無声化について”, 一般言語学論叢第2号, pp. 73. 100, Mar. 2000

## 付録 1 実験説明 (非母語者)

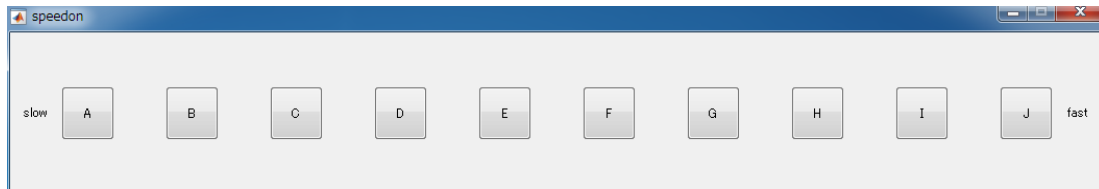
本日は実験に協力していただき、誠にありがとうございます。

この実験は、母語者と非母語者の会話場面を分析するためのもので、こちらは支援システムの使用について説明いたします。

### 1, 手順 1

まず、デフォルト話速値の設定画面を説明します。

話速値の設定画面では 10 つのボタンがあります。ボタンを押したら 10 秒程度の音声流されます。左側のボタンで流れる音声の速度は遅いが、右側に行くと流れる音声の速度は速くなります。ボタンを自由に押しただいて、話の速度が速すぎると思うところのボタンの番号を教えてください。



### 2, 手順 2

システムは自動で母語者の発話速度を検知し、速いと判断した時は赤い「TOO FAST」と表示します。

逆に普段では緑の「GOOD」と表示します (右図より)。

但し書き：ソフトウェアが異常な時、実験者を呼んで下さい。





## 付録 2 実験説明 (母語者)

本日は実験に協力していただき、誠にありがとうございます。

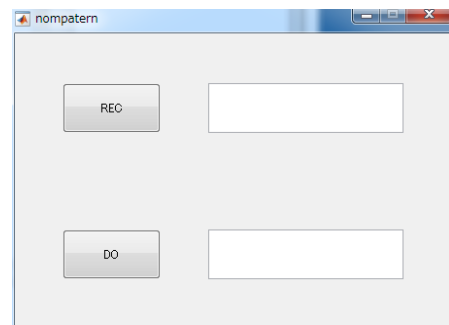
この実験は、母語者と非母語者の会話場面を分析するためのもので、こちらは支援システムの使用について説明いたします。

### 1, 手順 1

まず、マイクを装着してください。口に近い場所をお願いします。

### 2, 手順 2

母語者設定画面について、実験者は「REC」ボタンを押したら、「あ」「い」「う」「え」「お」を三回発声してください。



### 3, 手順 3

実験中では画面は変わります。

普段では画面に緑の「GOOD」と表示します。



もしあなたの話速が速いと検知された時は、画面に赤い「TOO FAST」が表示します。



### 付録 3 実験説明 (共通)

これからお題当てゲームの説明を行います。

このゲームでは、いくつかの単語が記されたお題のカードを母語者にお渡しします。非母語者は、母語者に渡されたカードの単語を答えることが目的です。カードの単語を非母語者に答えてもらうために、母語者は非母語者にヒントを与えます。

まずゲームの開始前に、母語者は1分間、お題のカードを見ながら出すヒントを考えてください。その後、5分間ゲームを行います。この間に母語者はヒントを出し、非母語者はヒントを聞いてお題と思う単語を回答します。

母語者は、お題カードに記されているどの単語についてヒントを出しても構いません。時間内にできるだけ多くの単語を当てさせることを目指してください。

ただし、以下のヒントは違反になります。

- 1、**答えの漢字と1字でも重なるヒント**
- 2、**正解と同じ意味の言葉または反対意味の言葉を使ったヒント**
- 3、**ジャスチャを使ったヒント**

例として、お題「富士山」のヒントとして「日本の山」と言うてはいけません。

非母語者がお題を当てたときには10点を獲得します。

母語者が違反ヒントを出してしまうと5点減点となります。

できるだけ高得点を目指してください。総合得点が150点を超えた場合、2人の謝礼金がそれぞれ100円増えます。

ここまで質問はありますか。

では、母語者にお題のカードをお渡しますので、頑張ってください。

## 付録 4 キーワードメモ用紙

<p>手帳(てちょう) カレーライス 時計(とけい) 救急車(きゅうきゅうしゃ) 布団(ふとん) 会議室(かいぎしつ) 蜜柑(みかん) 自転車(じてんしゃ)</p> <p>ミルク 体育館(たいいくかん) 学生(がくせい) アイスクリーム 会社員(かいしゃいん) 病院(びょういん) 帽子(ぼうし) 喫茶店(きっさてん)</p> <p>ロボット 食堂(しょくどう) 地下鉄(ちかてつ) 部屋(へや) 階段(かいだん) コート 椅子(いす) 宿題(しゅくだい)</p>	<p>電話(でんわ) 銀行(ぎんこう) 学校(がっこう) 美術館(びじゅつかん) サラリーマン 飛行機(ひこうき) ポケット 辞書(じしょ)</p> <p>ノート 新幹線(しんかんせん) ペット 玄関(げんかん) 切符(きっぷ) ローン 教室(きょうしつ) 醤油(しょうゆ)</p> <p>ビール 野球(やきゅう) 歌舞伎(かぶき) 子供(こども) 爆弾(ぼくだん) 風呂(ふろ) 郵便局(ゆうびんきょく) 地震(じしん)</p>
1	2
<p>風邪(かぜ) パン 図書館(としょかん) 眼鏡(めがね) 刺身(さしみ) 駐車場(ちゅうしゃじょう) ドライバー 花火(はなび)</p> <p>ハンカチ 弁当(べんとう) 運動会(うんどうかい) 柔道(じゅうどう) カメラ 漫画(まんが) 動物園(どうぶつえん) マラソン</p> <p>警察(けいさつ) 親子丼 カップ 交番(こうばん) 人形(にんぎょう) お土産 台風(たいふう) 洗濯機(せんたくき)</p>	<p>冷蔵庫(れいぞうこ) 公園(こうえん) 温泉(おんせん) 本屋(ほんや) 紅茶(こうちゃ) 会社(かいしゃ) 財布(さいふ) 水道(すいどう)</p> <p>乗り場 林檎(りんご) 鉛筆(えんぴつ) ベッド 靴下(くつした) 葉書(はがき) サンドイッチ 切手(きって)</p> <p>祭り 試験(しけん) トイレ 寿司(すし) 天ぷら 市役所(しやくしょ) ボタン 神社(じんじや)</p>
3	4

## 付録 5 質問紙 (非母語者)

名前： \_\_\_\_\_ 年齢： \_\_\_\_\_ 在日時間： \_\_\_\_\_ 日本語能力： N1 N2 N3

設問に対して当てはする箇所（7箇所のいずれか）に○をしてください。

1. 会話は楽しかった

全く当てはまらない   当てはまらない   あまり当てはまらない   どちらともいえない   やや当てはまる   当てはまる   よく当てはまる

2. 相手の発話速度にストレスを感じた

全く当てはまらない   当てはまらない   あまり当てはまらない   どちらともいえない   やや当てはまる   当てはまる   よく当てはまる

3. 相手の話をよく聞き取れた

全く当てはまらない   当てはまらない   あまり当てはまらない   どちらともいえない   やや当てはまる   当てはまる   よく当てはまる

4. 自分はよく話せた

全く当てはまらない   当てはまらない   あまり当てはまらない   どちらともいえない   やや当てはまる   当てはまる   よく当てはまる

5. 相手の話が速いと感じたことがあった

全く当てはまらない   当てはまらない   あまり当てはまらない   どちらともいえない   やや当てはまる   当てはまる   よく当てはまる

6. 話が速過ぎるという表示に気がついた

全く当てはまらない	当てはまらない	あまり当てはまらない	どちらともいえない	やや当てはまる	当てはまる	よく当てはまる
-----------	---------	------------	-----------	---------	-------	---------

7. 話が速過ぎるという表示のタイミングは適切だった

全く当てはまらない	当てはまらない	あまり当てはまらない	どちらともいえない	やや当てはまる	当てはまる	よく当てはまる
-----------	---------	------------	-----------	---------	-------	---------

8. 話が速すぎるという表示によって相手はゆっくり話すようになった

全く当てはまらない	当てはまらない	あまり当てはまらない	どちらともいえない	やや当てはまる	当てはまる	よく当てはまる
-----------	---------	------------	-----------	---------	-------	---------

9. システムは会話の妨げになった

全く当てはまらない	当てはまらない	あまり当てはまらない	どちらともいえない	やや当てはまる	当てはまる	よく当てはまる
-----------	---------	------------	-----------	---------	-------	---------

10. システムの画面は見やすかった

全く当てはまらない	当てはまらない	あまり当てはまらない	どちらともいえない	やや当てはまる	当てはまる	よく当てはまる
-----------	---------	------------	-----------	---------	-------	---------

(このページはシステムあり条件のみに配布する)

## 付録 6 質問紙 (母語者)

名前： \_\_\_\_\_ 年齢： \_\_\_\_\_ 日本在住年数： \_\_\_\_\_ 出身地： \_\_\_\_\_

設問に対して当てはまる箇所（7箇所のいずれか）に○をしてください。

1. 会話は楽しかった

\_\_\_\_\_

全く当てはまらない    当てはまらない    あまり当てはまらない    どちらともいえません    やや当てはまる    当てはまる    よく当てはまる

2. 相手に理解をさせるストレスを感じた

\_\_\_\_\_

全く当てはまらない    当てはまらない    あまり当てはまらない    どちらともいえません    やや当てはまる    当てはまる    よく当てはまる

3. 相手の話をよく聞き取れた

\_\_\_\_\_

全く当てはまらない    当てはまらない    あまり当てはまらない    どちらともいえません    やや当てはまる    当てはまる    よく当てはまる

4. 自分はよく話せた

\_\_\_\_\_

全く当てはまらない    当てはまらない    あまり当てはまらない    どちらともいえません    やや当てはまる    当てはまる    よく当てはまる

5. 自分の話が速いと感じたことがあった

\_\_\_\_\_

全く当てはまらない    当てはまらない    あまり当てはまらない    どちらともいえません    やや当てはまる    当てはまる    よく当てはまる

6. 話が速過ぎるという表示に気がついた

全く当てはまらない	当てはまらない	あまり当てはまらない	どちらともいえない	やや当てはまる	当てはまる	よく当てはまる
-----------	---------	------------	-----------	---------	-------	---------

7. 話が速過ぎるという表示のタイミングは適切だった

全く当てはまらない	当てはまらない	あまり当てはまらない	どちらともいえない	やや当てはまる	当てはまる	よく当てはまる
-----------	---------	------------	-----------	---------	-------	---------

8. 話が速すぎるという表示によって自分はゆっくり話すようになった

全く当てはまらない	当てはまらない	あまり当てはまらない	どちらともいえない	やや当てはまる	当てはまる	よく当てはまる
-----------	---------	------------	-----------	---------	-------	---------

9. システムは会話の妨げになった

全く当てはまらない	当てはまらない	あまり当てはまらない	どちらともいえない	やや当てはまる	当てはまる	よく当てはまる
-----------	---------	------------	-----------	---------	-------	---------

10. システムの画面は見やすかった

全く当てはまらない	当てはまらない	あまり当てはまらない	どちらともいえない	やや当てはまる	当てはまる	よく当てはまる
-----------	---------	------------	-----------	---------	-------	---------

(このページはシステムあり条件のみに配布する)

付録 7 質問紙結果 (母語話者)

非提示条件					
質問項目 参加者 ペア	1.会話は楽 しかった	2.相手に理 解させるスト レスを感じた	3.相手の話 をよく聞き取 れた	4.自分はよく 話せた	5.自分の話 が速いと感 じたことがあ った
1	7	1	7	7	7
2	6	3	5	5	5
3	7	3	7	7	2
4	6	2	6	6	2
5	6	5	5	6	6
6	7	3	7	6	5
7	6	4	4	5	3
8	7	1	7	7	1
9	3	7	3	3	7
10	6	5	7	5	5
11	5	6	7	3	5
12	6	3	6	5	3
15	7	4	6	4	5
16	6	3	6	5	3
18	6	5	5	4	3
19	6	4	7	5	3
21	5	5	6	3	3
22	6	5	4	2	5
平均	6.0	3.8	5.8	4.9	4.1



話者提示条件										
質問項目 参加者ペア	1. 会話は楽しかった	2. 相手に理解させるストレスを感じた	3. 相手の話をよく聞き取れた	4. 自分をよく話せた	5. 自分の話が速いと感じたことがあった	6. 話が速過ぎるという表示に気が付いた	7. 話が速過ぎるという表示のタイミングは適切だった	8. 話が速過ぎるという表示によって自分はゆっくり話すようになった	9. システムは会話の妨げになった	10. システムの画面は見やすかった
1	7	1	7	7	6	3	6	3	1	5
2	6	2	6	3	5	7	4	6	2	7
3	7	2	6	7	2	2	4	4	4	7
4	6	2	7	4	5	7	6	6	5	7
5	6	3	6	7	6	7	4	3	2	6
6	4	7	7	5	7	7	7	7	2	7
7	6	5	4	3	5	6	3	6	4	6
8	6	2	7	5	2	6	2	6	6	6
9	7	1	7	7	1	7	7	7	1	7
10	6	2	7	7	3	5	3	5	3	4
11	6	5	7	4	5	6	5	4	5	5
12	6	5	6	3	6	3	4	6	4	6
15	7	3	6	4	5	5	5	6	2	6
16	6	5	6	3	6	3	4	6	4	6
18	6	6	5	2	6	6	4	6	5	5
19	6	6	6	5	2	6	4	2	2	6
21	5	5	6	3	5	2	4	3	3	4
22	5	5	5	3	6	7	7	7	2	7
平均値	6.0	3.7	6.0	4.6	4.6	5.3	4.6	5.2	3.2	5.9

両者提示条件										
質問項目  参加者ペア	1. 会話は楽しかった	2. 相手に理解させるストレスを感じた	3. 相手の話をよく聞き取れた	4. 自分をよく話せた	5. 自分の話が速いと感じたことがあった	6. 話が速過ぎるという表示に気が付いた	7. 話が速過ぎるという表示のタイミングは適切だった	8. 話が速過ぎるという表示によって自分はゆっくり話すようになった	9. システムは会話の妨げになった	10. システムの画面は見やすかった
1	6	1	7	5	5	5	5	4	2	7
2	7	3	6	2	5	6	5	6	3	2
3	7	2	7	6	3	5	5	2	3	6
4	6	5	7	4	6	6	6	6	3	7
5	6	5	6	6	7	7	5	4	3	5
6	7	5	7	7	7	7	5	7	5	7
7	5	5	4	3	5	1	4	5	4	6
8	6	5	7	5	1	6	2	7	6	7
9	7	1	7	7	1	7	5	7	1	7
10	6	4	7	7	3	7	6	6	5	7
11	5	5	7	4	3	6	4	5	5	5
12	6	3	6	5	5	3	4	5	3	4
15	7	4	6	5	3	6	6	6	2	6
16	6	3	6	5	5	3	4	5	3	4
18	5	5	5	3	5	6	4	6	3	5
19	6	3	6	5	2	6	4	7	2	6
21	6	4	7	4	3	5	5	5	3	6
22	6	4	4	4	6	7	7	7	2	5
平均値	6.1	3.7	6.2	4.8	4.2	5.5	4.8	5.6	3.2	5.7

付録 8 質問紙結果（非母語話者）

非提示条件					
質問項目 参加者 ペア	1.会話は楽 しかった	2.相手の発 話速度にス トレスを感じ た	3.相手の話 をよく聞き取 れた	4.自分はよく 話せた	5.相手の話 が速いと感 じたことがあ った
1	7	1	6	6	6
2	6	1	7	4	4
3	7	1	7	4	1
4	6	6	7	6	6
5	7	7	5	5	7
6	6	1	6	5	1
7	7	6	6	5	6
8	7	1	5	5	2
9	6	6	3	5	6
10	7	2	6	2	1
11	7	1	7	6	1
12	7	1	7	4	1
15	7	3	5	4	4
16	7	1	7	4	1
18	5	3	5	6	4
19	5	4	5	3	4
21	7	6	5	4	5
22	7	1	5	5	1
平均	6.6	2.9	5.8	4.6	3.4

話者提示条件					
質問項目 参加者 ペア	1.会話は楽 しかった	2.相手の発 話速度にス トレスを感じ た	3.相手の話 をよく聞き取 れた	4.自分はよく 話せた	5.相手の話 が速いと感 じたことがあ った
1	7	1	6	6	6
2	7	1	6	5	2
3	7	1	7	4	1
4	6	6	6	6	6
5	7	7	5	5	7
6	5	1	5	5	1
7	7	7	6	5	6
8	7	1	7	6	1
9	6	2	6	6	2
10	7	1	6	5	2
11	7	1	7	6	1
12	7	1	7	4	1
15	7	5	7	5	4
16	7	1	7	4	1
18	5	2	6	3	1
19	7	5	4	5	4
21	7	6	5	2	5
22	7	2	5	3	3
平均	6.7	2.8	6.0	4.7	3.0

両者提示条件										
質問項目  参加者ペア	1. 会話は楽しかった	2. 相手の発話速度にストレスを感じた	3. 相手の話をよく聞き取れた	4. 自分をよく話せた	5. 相手の話が速いと感じたことがあった	6. 話が速過ぎるという表示に気が付いた	7. 話が速過ぎるという表示のタイミングは適切だった	8. 話が速過ぎるという表示によって相手はゆっくり話すようになった	9. システムは会話の妨げになった	10. システムの画面は見やすかった
1	6	1	6	6	6	6	6	5	1	6
2	7	1	6	5	5	6	6	6	2	7
3	7	1	7	4	1	4	4	4	1	4
4	5	6	5	5	6	6	6	6	6	6
5	7	7	5	5	7	7	7	7	7	5
6	5	1	6	5	1	7	3	3	1	6
7	7	6	5	4	5	5	6	6	2	6
8	7	1	7	6	1	6	4	7	4	7
9	6	2	6	6	2	6	3	5	1	5
10	7	1	7	4	1	5	3	6	4	7
11	7	1	6	4	1	1	4	4	1	4
12	7	1	7	4	1	5	5	7	2	6
15	7	2	6	4	3	4	4	5	4	4
16	7	1	7	4	1	5	5	7	2	6
18	5	2	6	2	1	5	5	5	2	7
19	7	2	5	5	3	4	6	4	2	4
21	7	5	5	4	6	1	4	4	1	6
22	7	1	6	5	1	1	4	5	5	5
平均値	6.6	2.3	6.0	4.6	2.9	4.7	4.7	5.3	2.7	5.6