

第二言語会話における話速の影響を考慮した支援
の研究

筑波大学

図書館情報メディア研究科

2019年03月

LIAO WEI

目次

第1章	はじめに	1
第2章	関連研究	2
2.1	第二言語コミュニケーション支援	2
2.2	NNS リスニングと発話理解	2
2.3	話速が聴解力への影響	3
2.4	話速変換における会話支援	4
2.5	話速測定方法	4
第3章	NNSの「NS発話」の理解困難に関する実験	6
3.1	実験参加者	6
3.2	実験デザイン	6
3.3	実験機材	6
3.3.1	NNS 不理解時点を記録に用いた機材	6
3.3.2	NNS 不理解時点を確認に用いた機材	7
3.3.3	映像の記録に用いた機材	7
3.4	実験環境	7
3.5	実験手順	9
3.6	データ取得	9
3.6.1	映像音声の記録	9
3.6.2	Bluetooth 接続ボタン押下式デバイスで取れた時間ログの記録	9
3.6.3	質問紙調査	10
3.6.4	不理解時点と原因の確認	10
3.7	データ処理	11
3.7.1	NS 発話, NNS 不理解のラベリング	11
3.7.2	NS 話速の計算	14
3.8	結果	15
3.8.1	NNS が感じた不理解回数および原因	15
3.8.2	不理解時の NS 話速変化	17
3.8.3	アンケート結果	18
3.9	検討	19
第4章	Woz 法による話速通知手法有効性の検証	20
4.1	話速通知手法	20
4.2	Woz 法の適用	20
4.3	Woz 法に基づく話速通知システムの実現	20

4.4	録音	21
4.5	発話速度閾値の設定	22
4.6	音声認識	22
4.7	早口判定	22
4.8	データ保存	23
4.9	実験	23
4.9.1	目的	23
4.9.2	実験参加者	23
4.9.3	比較条件	23
4.9.4	実験内容	24
4.9.5	実験機材	25
4.9.5.1	早口通知に用いた機材	25
4.9.5.2	映像音声の記録に用いた機材	25
4.9.6	実験環境	25
4.9.7	手順	27
4.9.8	データ取得	27
4.9.8.1	映像音声の記録	27
4.9.9	データ処理	28
4.9.9.1	NS 発話区間, 通知区間のラベリング	28
4.9.10	結果	29
4.9.10.1	NS の発話速度変化	29
4.9.11	検討	30
第 5 章	話速通知システム有効性の検証	31
5.1	話速通知システムの概要	31
5.2	実験	32
5.2.1	目的	32
5.2.2	実験参加者	32
5.2.3	比較条件	32
5.2.4	実験内容	33
5.2.5	実験機材	33
5.2.5.1	早口通知に用いた機材	33
5.2.5.2	映像音声の記録に用いた機材	33
5.2.6	実験環境	33
5.2.7	手順	35
5.2.8	データ取得	36
5.2.8.1	映像音声の記録	36
5.2.9	データ処理	37
5.2.10	結果	37
5.2.10.1	NS の発話速度変化	37

第 6 章	話速通知手法が非母語話者理解への影響	39
6.1	実験参加者	39
6.2	比較条件	39
6.3	実験デザイン	39
6.4	実験機材	40
6.4.1	NNS 不理解を記録に用いた機材	40
6.4.2	早口通知に用いた機材	40
6.4.3	映像音声の記録に用いた機材	40
6.5	実験環境	40
6.6	実験手順	40
6.7	データ取得	41
6.7.1	映像音声の記録	41
6.7.2	Bluetooth 接続ボタン押下式デバイスで取れた時間ログの記録	41
6.8	結果	41
6.9	検討	41
第 7 章	考察とまとめ	42
	参考文献	44

目次

1	Bluetooth 接続ボタン押下式デバイス Flic(27mm × 27mm × 8mm)	7
2	実験配置図	8
3	実験環境図	8
4	ボタン押しの時間ログ	10
5	取得した不理解時点と原因の例	11
6	無音区間自動認識機能:「無音音量レベル」,「最短の無音時間」および「最短の有音時間」の設定 (A),「有音区間」(B),「無音区間」(C)	12
7	NS 発話区間のラベリング例:会話の映像データ (A), 会話音声データの波形 (B), NS 発話区間のラベリング (書き起こし済み) (C)	13
8	不理解時点とセリフの例	13
9	NNS が不理解を抱いた NS の発話区間例: 日本時間 (D), NNS 不理解区間ラベリング (E), ラベリングした区間のリスト (F)	14
10	不理解直前の発話と不理解発話と flic 押下	15
11	NNS が不理解を感じた時の NS 話速変化 (N=75)	17
12	質問紙調査 (NS: N=15, NNS: N=15)	18
13	フローチャート	21
14	発話速度閾値設定ツール	22
15	音節測定補助キット	23
16	通知画面	24
17	非提示条件の配置図	25
18	非提示条件の実験様子	26
19	提示条件の配置図	26
20	提示条件の実験様子	27
21	通知区間のラベリング	28
22	通知前後の NS 発話区間ラベリング: 通知直前の NS 発話区間 (A), 通知区間 (B), 通知直後の NS 発話区間 (C)	28
23	通知前後の NS 話速 (通知前の NS 話速: N=53, 通知後の NS 話速: N=53)	30
24	自動処理話速通知システム	31
25	通知画面	32
26	非提示条件の配置図	34
27	非提示条件の実験様子	34
28	提示条件の配置図	35
29	提示条件の実験様子	35
30	母音標準パターン作成	36

31 通知前後 NS の話速（通知前の NS 話速: N=81, 通知後の NS 話速: N=81） . 38

第1章 はじめに

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，訪日外国人数は年々増加し，2017年では2869万人以上に伸びた [1]．異なる母語の人々の間で，異言語間の意思疎通のため共通言語を使ってコミュニケーションを行う機会がさらに増加すると考えられる．そのため，日本語母語ではない在日外国人にとって，同じ在日外国人と日本語を共通言語とする会話を行う場合，または日本語母語話者（Native Speaker, NS）と会話を行う場合がある．しかし，日本語を用いて日本語母語話者とコミュニケーションを取る場合，言語能力の差などにより，非母語話者（Non Native Speaker, NNS）の聞き取りと理解困難による会話のアンバランスがよく見られている．

第二言語コミュニケーションにおけるNNSの聞き取りと理解困難の原因は多数の研究で調査されてきた．Bloomfieldら [2] のNNSリスニングに影響する要素に関する調査により，NNSの聞き取りはワーキングメモリ，言語の流暢さといった聞き手要素，言葉の長さ，複雑さ，ポーズ，発話速度といった発話面要素に影響されている．

発話面要素である発話速度に関する調査では，NNSが第二言語リスニングを受ける際，聴解材料の平均発話速度（Speech Rate, SR）は速くなれば，NNSの聞き取りに困難をもたらすことがある [3][4]．しかし，これらの調査ではNSのリアルタイムな話速変化を客観的に調査していなかった．

本研究はNSとNNSによる会話の支援を目的としている．実際の第二言語コミュニケーション場面において，NSの発話速度の増加によりNNSの理解困難が生じることを客観的なデータを通じて検証した上，話速視点から第二言語コミュニケーションを支援すると考え，叶，井上がNSの早口をリアルタイムに知らせ，NSの自発調整を促す手法を提案した [5]．

実際の第二言語コミュニケーション場面におけるこの手法の有効性を検証するため，本研究はWoz法を用い，会話環境を作り，実証実験を行う．結果として早口を知らせることによるNSの自発的話速調整が見られた．

以上の実験結果を踏まえて，話速通知手法の有効性と実用性が期待できると考え，Ye，井上が音声メディア処理で，自動的NSの話速を検知し，話速が速いと判断した際に通知を出し，NSに気づかせる通知システムを構築した [6]．この話速通知システムを使って検証を行う．システムの使用により，NSが自発的に話速調整を行うこと，またNSの話速調整によりNNSの理解度が増加することが期待できる．

第2章 関連研究

これまで、第二言語コミュニケーション支援のために、発話速度やNNSのリスニングに影響する要素などの研究が多数されてきた。

2.1 第二言語コミュニケーション支援

第二言語コミュニケーションを支援するために、様々な手法やシステムが提案されている。

例えば、井上らは会話中のNNSの負担を考慮し、NNSを含む第二言語コミュニケーション中に、NSは会話しながら、会話中の重要な部分や理解しにくい部分をキーボードよりテキスト入力し、それをNNSに表示するというNNS会話支援方法を提案した。この手法により会話中に会話内容に理解度が上がり、入力によって参加者の相互理解も増加する効果があることが指摘された [7]。

異なる国の文化的な背景がもたらす知識の違いを補うために、岡本らは対面型異文化間コミュニケーション支援システム iGengo を開発した。iGengo は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であり、iGengo の評価実験による会話中に発言された名詞に関する画像、母語の情報、説明の文章、関連名詞、関連画像をユーザに提示することは対話を効果的に支援する可能性があると示された [8]。しかし、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。

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NNS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、NSの話速から考え、NS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NSにその旨を提示して知らせる方法で第二言語コミュニケーションを支援する。

2.2 NNSリスニングと発話理解

NSとNNSによる第二言語会話の場面では、NNSの聞き取り問題と理解困難の会話アンバランスがよく見られている。池上がリスニングの一般的なプロセスを説明したように、リスニング過程において、ワーキングメモリは聞いた音声情報を受け、瞬時的な処理を行い、情報の理解に至ることである。その一連の情報処理はNSとNNSによって異なる。母語の場合、音の知覚から単語、語句認知への処理(ボトムアップ・プロセス)が無意識に自

動的になされるに比べ、NNS は その処理を意識的に行い、処理に時間がかかることが指摘された [10].

NNS リスニングを影響する要素に関する調査において、Bloomfield ら [2] は NNS 聴解力へ影響を与える要素について、ワーキングメモリ、言語の流暢さという聞き手方面の要素と段落の長さ、複雑さ、ポーズ、発話速度といった発話面の要素などを指摘した。そして発話速度は他の要素とお互いに影響し合っており、NNS にリスニング上の困難をきたす可能性があることかも示唆された。

また、C. C. Goh.[11] による NNS の聴解に対する聞き取り問題の調査で、NNS はボキャブラリーの欠如、知っている言葉を認識できない、言葉を聞いてイメージ掴めない、会話の流れに乗れない、前の言葉の意味を考えて次の部分を見逃す、予期しない単語の出現が混乱させる、前に問題があったため後続部分も理解できない、聞いたものを忘れるなどの原因で NS の言葉に理解困難を感じていた。

本研究では、主観的データのみならず、NS の話速変化と NNS の理解困難との繋がりを全面的に調査するため、これらの研究で指摘された影響要素に基づき、NNS の不理解と NS の話速についてデータをとった。詳細については四章で述べる。

2.3 話速が聴解力への影響

発話速度の速さは聞き手の聴解に影響を及ぼすことが知られている。特に第二言語会話では発話の速度が重要な要素である。

NNS が第二言語リスニングを受ける際、聴解材料の発話速度 (Speech Rate, SR) は通常より速くなれば、NNS の聞き取りに困難をもたらすことである。Griffiths[3] は fast(200wpm), normal(150wpm), slow(100wpm) の三種の発話速度で録音したテキストを NNS に聞いてもらい、聴解実験を行った。聴解実験により NNS のテキストに関する理解度を調べたところ、fast で録音したテキストの場合、NNS の理解度が大幅に下がったという結果がわかった。話すスピードが比較的速くなれば、NNS の発話理解を妨げたことが示されたが、一般的な速度と緩い速度の場合、その NNS 聴解へ与える影響は大きな違いが見られなかった。Zhao[4] は通常速度の聴解材料、発話速度が異なる聴解材料、発話速度を NNS が自由にコントロールできる聴解材料を用意し、NNS にそれぞれの聴解材料を聞いてもらい、発話速度と NNS リスニング理解度に関する調査を行った。その結果、NNS は聞いた発話速度を自分でコントロールできる場合、よく通常速度より遅いスピードに調整し、高い理解度が示された。

第二言語学習の場面で、NNS にとって通常速度と緩い発話速度は発話理解に有効であるが、発話速度は緩ければ緩いほど聞き取りが良くなるわけではない、NNS によって最も適切な話速が理解に助ける。Hayati[12] は 62 名の英語学習者を 2 グループに分けられ、自然な発話速度の聴解材料と緩い発話速度の聴解材料をそれぞれ 2 グループに配布し、同じ内容の英語授業に参加してもらって調査を行なった。学習者はリスニング開始前、プレテストを受け、2 週後にポストテストをしたことで、聴解力を測った。その結果、2 グループとも聴解力が上がったが、自然な発話速度の聴解材料を受けたグループの方はより顕著な進歩が示された。

また、発話速度と噪音の角度から NNS 聞き取りへの影響に関する調査で、Jones ら [13] は TTS (text-to-speech) 合成システムによって作成した速い発話速度 (155wpm)、通常の

発話速度 (178wpm) の銀行製品の説明文を使い、背景雑音 (6dB) の有無を設定して比較実験を行った。その結果、NNS が正確に内容理解できるかどうかは雑音と関係が少ないことがわかった。また、速い発話速度を聞いてもらうと被験者の内容理解正確率が下がったことが示された。

アクセントのある第二言語に対して、発話速度が NNS リスニングに対する影響について、Matsuura らが調査を行った。なまりが強い英語を聞く場合、日本人の英語学習者は通常速さより緩い発話速度を聞いた方が内容理解に有効であることが示された [14]。

2.4 話速変換における会話支援

早口で話された音声聞き取りにくいと感じられる問題を解決するため、早口音声を話速変換処理で会話の理解度を補足し、話速変換システムが開発された。富士通は音声検出部で受話音に含まれる音声区間と無音区間 (音声が含まれない区間) を判定し、音声区間を伸長しながら無音区間を短縮することで、受話音の音声の話速を遅くする話速変換技術を開発した。全年齢の被験者に対して評価実験を行なった結果、受話音の話速が速い場合に、聞き手の年齢によらず聞きやすさを高める効果があることが確認できた [15]。清山らは無音区間も含めた伸縮による話速変換方式を提案するとともに、高齢者に対する音声放送サービス向上を目指した小型の話速変換器を開発した [16]。

話速変化は聞きやすさを向上する効果があるが、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、非同期の違和感を感じることがある。津村らは話者映像と音声のずれが単語理解度に及ぼす影響を調べるため、4 モーラ単語のうち各モーラにおいて 1 音素違いのペア単語刺激を用意し、「映像のみ」条件、「音声のみ」条件、「音声+映像」条件で単語理解度実験を実施し、各モーラ位置において視聴覚統合が生じているか検討した。その結果、単語音声を 400ms 以上に伸長し、通常速度の映像を付加すると、単語理解度に影響を及ぼす可能性があることが分かった [17]。

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。

2.5 話速測定方法

研究内容によって適当に計測単位を選択するのは重要である。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、母音を中心とする音のまとまりという。モーラ数はリズムを構成する最小単位で、日本語などごく一部の言語にしか認められない [18]。音節を数えるとき、長音、促音、また撥音といった独立した単位とは数えない。

Griffiths[3] は発話速度の聴解力への影響を調査する研究では、単位時間の音節数を発話速度として使用し、その重要性を示した。そして、Griffiths が使った発話速度は 1 分あたりの音節数であり、それが 1 段落などの平均な速度として、ポーズを含んでいる測り方である。しかし、このような研究で計測した発話速度は全体発話の平均な速度であり、発話中の発話速度の変化を考慮していなかった。

従って、本研究では、日本語発話を対象とし、一般的に使われている音節数を話速の計測単位にする。そして数秒単位での話速を計測するため、ポーズを含む計算方法は向いて

いないと考え，ポーズを除いて発話部分だけの単語数或いは音節数を計算した発話の速度
いわゆる調音速度 (Articulation rate) を本研究の発話速度とする。

第3章 NNSの「NS発話」の理解困難に関する実験

関連研究 2.2 に挙げた NNS リスニングを影響する要素に関する調査において、言葉の複雑さやポーズ、発話速度などが NNS にリスニング上の困難をきたす可能性があることが指摘された。しかし、NS と NNS が会話を行う際に、NS のリアルタイムな話速変化が、NNS の理解にどのような影響を与えるかは不明である。

3.1 実験参加者

被験者は初対面の日本語 NS（日本人学生）15 名および日本語 NNS（留学生）15 名のペア 15 組、合計 30 人である。

NS と NNS が二人一組で対面会話実験を行うため、NS は日本語を母語として使う者で、NNS は日本語で日常会話ができるが（N2 以上）、日本人と会話する際に不理解を感じる者が対象である。

NNS に NS と 5 分ほどの自己紹介や挨拶をしてもらい、NNS の日本語レベルを確認する。尾崎に指摘された NNS が困った場合の行動 [19] に基づき、会話中 NS に聞き返しをしたり、NS の発話に対して沈黙したり、困ったような表情や身振りをしたりした NNS だけを募集した。

3.2 実験デザイン

NS1 名と NNS1 名、各ペアに対面による 10 分間の自由会話を 1 回実施する。

会話中、NNS が NS の言葉に対して不理解を感じる際、ボタンを押してもらい、不理解の時間を記録する。

3.3 実験機材

3.3.1 NNS 不理解時点を記録に用いた機材

NNS が不理解を感じた時点を記録するために NNS にのみ手渡す押下式ボタン、Shortcut Labs の Bluetooth 接続ボタン押下式デバイス「Flic」を一台使用した。Flic は Shortcut Labs が提供するソフトウェア開発キットを用いることで、Bluetooth 接続した「The Raspberry Pi 3 Model B」一台上で、テキストファイルに Flic が押下された時刻を記録することができる。

なお、Flic を押下してからその時刻が記録されるまでにかかる遅延時間はほとんどない。



図 1: Bluetooth 接続ボタン押下式デバイス Flic(27mm × 27mm × 8mm)

3.3.2 NNS 不理解時点を確認に用いた機材

Flic が押下された時刻と実験中の時間と照らし合わせ，NNS の不理解を特定する必要があるため，実験後実験ビデオを見て実験中の時間をいつも確認できるように，日本時間を映るスクリーンを使用した．スクリーンは Panasonic TH-L42E60 である（サイズ：幅 957mm × 高さ 619mm × 奥行 217mm，画素数：1920 × 1080）

3.3.3 映像の記録に用いた機材

撮影用のカメラには「SANYO Xacti HD 2000 Full HD 1920*1080」を NS 向けと NNS 向け，そして NS，NNS 向けの合計 3 台使用した．

3.4 実験環境

NS と NNS は実験用の部屋で対面に座る．実験時の配置と環境は図 2 と図 3 に示す．会話ペア相互の対人距離を 120cm とした．

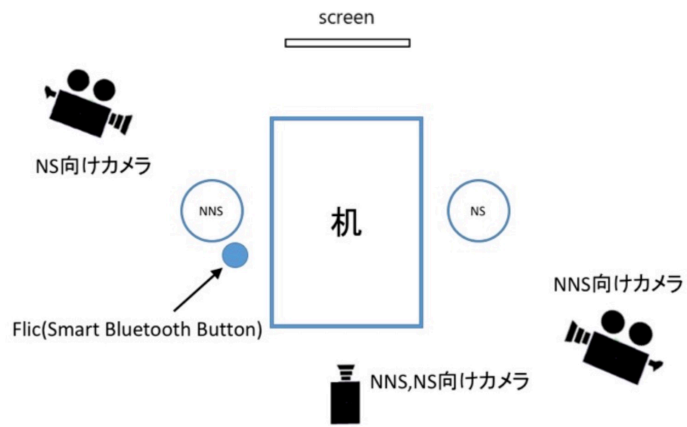


図 2: 実験配置図

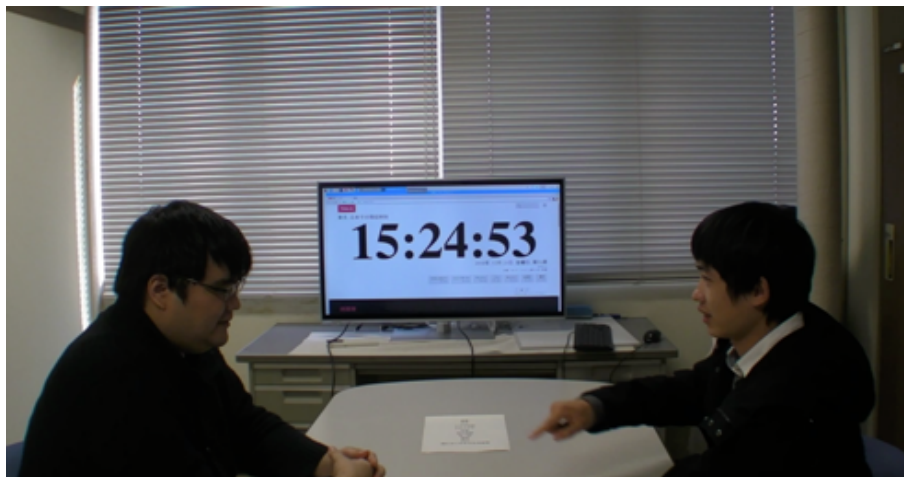


図 3: 実験環境図

3.5 実験手順

- 1) 録画用のスペースの中で、被験者 (NS, NNS 一人ずつ) に実験についての説明を聞いてもらう。
- 2) NS は、音声データを取得するためのマイクを着用する。NS の発話を録音する。
- 3) NNS の母語で NNS にのみ、NS との会話で不理解があった際につき一回 Bluetooth 接続ボタン押下式デバイスを押下するように伝える。そして NS に見えないように机の下で NNS に Bluetooth 接続ボタン押下式デバイスを手渡す。また、正常に押下タイミングが記録されるか確認するため NNS にその場で一回以上試し押しをするよう指示をする。
- 4) 会話をよりスムーズにするために用意した話題用紙について、書かれている話題（授業、ゼミ、クラブ活動、アルバイト、出身、海外経験、食べ物）の使用及び不使用は自由であることを両者に伝える。
- 5) 被験者ペア 10 分間自由に会話してもらう。
- 6) 会話直後に、被験者に会話関するアンケートを記入してもらう。
- 7) アンケート用紙を回収した後 NS を退出させ、NS, NNS 向けカメラで録画された動画を NNS と共に再生し、NNS に対して口頭で押下タイミングおよび不理解原因に関する確認をとり、記録する。

3.6 データ取得

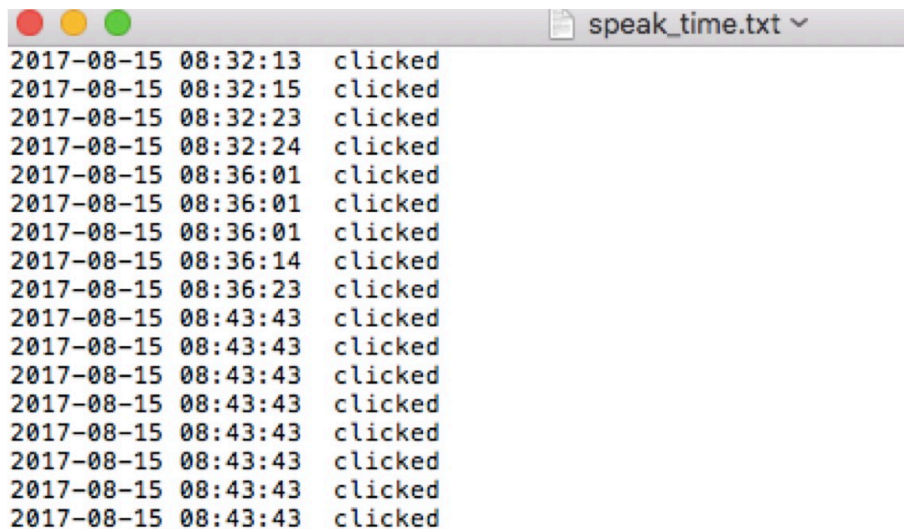
3.6.1 映像音声の記録

実験の様子は 3.3.3. 章に述べた映像記録に用いたカメラで記録する。

3.6.2 Bluetooth 接続ボタン押下式デバイスで取れた時間ログの記録

NNS の不理解時間は 3.3.1 章に述べた NNS 不理解時点を記録に用いた機材に用いた Bluetooth 接続ボタン押下式デバイスで記録する。

NNS が不理解を感じる度ボタンを 1 回押してもらっている。1 回のボタン押しに付き、押す時の日本時間が raspberry pi を通じて記録され、txt として出力される。



```
speak_time.txt
2017-08-15 08:32:13 clicked
2017-08-15 08:32:15 clicked
2017-08-15 08:32:23 clicked
2017-08-15 08:32:24 clicked
2017-08-15 08:36:01 clicked
2017-08-15 08:36:01 clicked
2017-08-15 08:36:01 clicked
2017-08-15 08:36:14 clicked
2017-08-15 08:36:23 clicked
2017-08-15 08:43:43 clicked
2017-08-15 08:43:43 clicked
2017-08-15 08:43:43 clicked
2017-08-15 08:43:43 clicked
2017-08-15 08:43:43 clicked
2017-08-15 08:43:43 clicked
2017-08-15 08:43:43 clicked
2017-08-15 08:43:43 clicked
2017-08-15 08:43:43 clicked
```

図 4: ボタン押しの時間ログ

3.6.3 質問紙調査

NSには、コミュニケーションの状況及び会話への印象を評価するため、NNSへの配慮に関する書面調査で用いられた項目3問[20]と自分の話速変化に対する主観的印象の5問、全8問の項目を利用した。NNSに対しては、会話内容を把握できた度合いと言いたいことを発話できた度合いに関する3問[21]とNSの話速変化に対する主観的印象の5問、全8問を利用した。全ての項目は「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」、「どちらとも言えない」、「やや当てはまる」、「当てはまる」、「よく当てはまる」を17に割り当てる7段階の評定尺度で評価してもらった。

3.6.4 不理解時点と原因の確認

Bluetooth 接続ボタン押下式デバイス Flic が押されたとしてテキストファイルに記録されたタイミングと不理解時点が合致しているかを確定するため、そして NNS はどのような原因で NS の言葉に理解困難を感じたのかを知るため、実験者は NNS 本人と一緒に撮影したビデオを再生しながら、NNS に対して口頭で押下タイミングおよび不理解原因に関する確認をとる。

確認する内容は、

1. システムに記録された時点に間違いがないか
 2. 不理解を感じたが Bluetooth 接続ボタン押下式デバイス Flic を押下し忘れたケースの有無
 3. 不理解を感じた際に、相手のどのセリフに、どのような理由で不理解を感じたのか
- 以上の三つの内容である。ここで間違え押しを除外し不理解時点と認めない。また、押下忘れがあった場合は押下タイミングとして追加する。

関連研究にあげられたように、BloomfieldらはNNS聴解力へ影響を与える要素を聞き手方面、話し手発話面、2つに分けられている[2]。C. C. Goh.によるNネイティブスピーカーの聴解に対する聞き取り問題の調査で、話し手方面の理解へ影響する要素は段落の長さ、複雑さ、ポーズ、発話速度などがあると指摘された。聞き手方面要素に関しては、NNSはボキャブラリーの欠如、知っている言葉を認識できない、言葉を聞いてイメージ掴めない、会話の流れに乗れない、前の言葉の意味を考えて次の部分を無視する、予期しない単語の出現が混乱させる、前に問題があったため後続部分も理解できない、聞いたものを忘れるなどの原因でNSの言葉に理解困難を感じていたことがわかった[10]。従って、本実験も聞き手方面、話し手発話面要素を軸にNNSに対して不理解理由を取得した。

取得した不理解時点と原因の例を図5に示す。

不理解番号	不理解時間	flic押下状況	分からなかったセリフ	理由	
13ペア	1回目	16:06:47	ボタンを押しなかったが、実験後にNNSに確認できた	「地層」分からなかった	ボキャブラリーの欠如（単語知らない）
	2回目	16:07:09	実験中でNNS自分でボタンを押した	「情報系」分からなかった	母語話者の話速早すぎた
	3回目	16:07:36	ボタンを押しなかったが、実験後にNNSに確認できた	「縦が...」分からなかった	母語話者の話速早すぎた
	4回目	16:07:47	ボタンを押しなかったが、実験後にNNSに確認できた	「お口に合いませんか」分からなかった	母語話者の話速早すぎた
	5回目	16:07:56	実験中でNNS自分でボタンを押した	「四川省」分からなかった	知っている言葉への認識や反応が遅かった（言葉は知ってる、その時点は反応できなかったが、後すぐわかった）
	6回目	16:09:21	ボタンを押しなかったが、実験後にNNSに確認できた	「リーゼント」分からなかった	ボキャブラリーの欠如（単語知らない）
	7回目	16:09:34	実験中でNNS自分でボタンを押した	「構って欲しい」分からなかった	ボキャブラリーの欠如（単語知らない）
	8回目	16:10:46	実験中でNNS自分でボタンを押した	「学校の授業に水泳が入っていますか」分からなかった	母語話者の話速早すぎた

図 5: 取得した不理解時点と原因の例

3.7 データ処理

3.7.1 NS発話、NNS不理解のラベリング

2.3章の先行研究より、発話速度が速い場合、NNSの理解度が下がったことが明確であるが、考察された発話の速度は基本的に1段落あるいは1文章の平均な速度を指した。NNSはNSと会話する過程で、NSの発話に理解しにくい場合、リスニング上の問題が出た直前に、NS発話速度の変化を調べる必要がある。会話中でNNSは発話の理解が難しい場面で、NSのリアルタイムな話速変化を分析するために、不理解が起きる前のNSの発話区間をえることが必要である。

そのため、記録された15組各10分間、合計150分間の音声付きの実験ビデオデータを編集せずに用い、NS発話区間をラベリングした。

また、実験後に確認できた不理解時間および不理解を感じたNSセリフにより、15組合計115回の実験に対して、NNSが不理解を抱いたNSの発話区間をラベリングした。手続きとして、

1) マルチモーダルデータのアノテーションツールとして有名なELANを用いて会話全体の発話区間を切り出した。有音区間単位は、会話分析の研究[22][23]でよく使われている統

語的要素と独立して客観的に定義できる，無音で区切られた単位 (Inter Pausal Units: IPU) を用いることとし，ここでは 200 ms 以上の無音によって区切られた IPU を分析単位とした。

2) ELAN の無音区間自動認識機能 (図 6) を用いて，書き出した音声データに対し，無音区間と有音区間の自動分別を行った。メディアファイルを ELAN に読み込んだ後，「音声認識」タブをクリックすると，「Recognizer」のプルダウンメニューから実験オーディオファイルを「Silence Recognizer MPI-PL」として選択できる。「無音音量レベル」(実験環境騒音により調整，30dB か 40dB 程度)，「最短の無音時間」(200ms) および「最短の有音時間」(0ms) を設定し，「開始」ボタンをクリックすると，図 6 中に x から s までが「有音区間」と認識され，s から x までが「無音区間」と認識される。「有音区間」をラベリングするため，「注釈の詳細設定」で，「表示する注釈」の「s」のチェックを外し，「注釈層を作成する」をクリックすると自動認識による有音区間がラベリングされた。

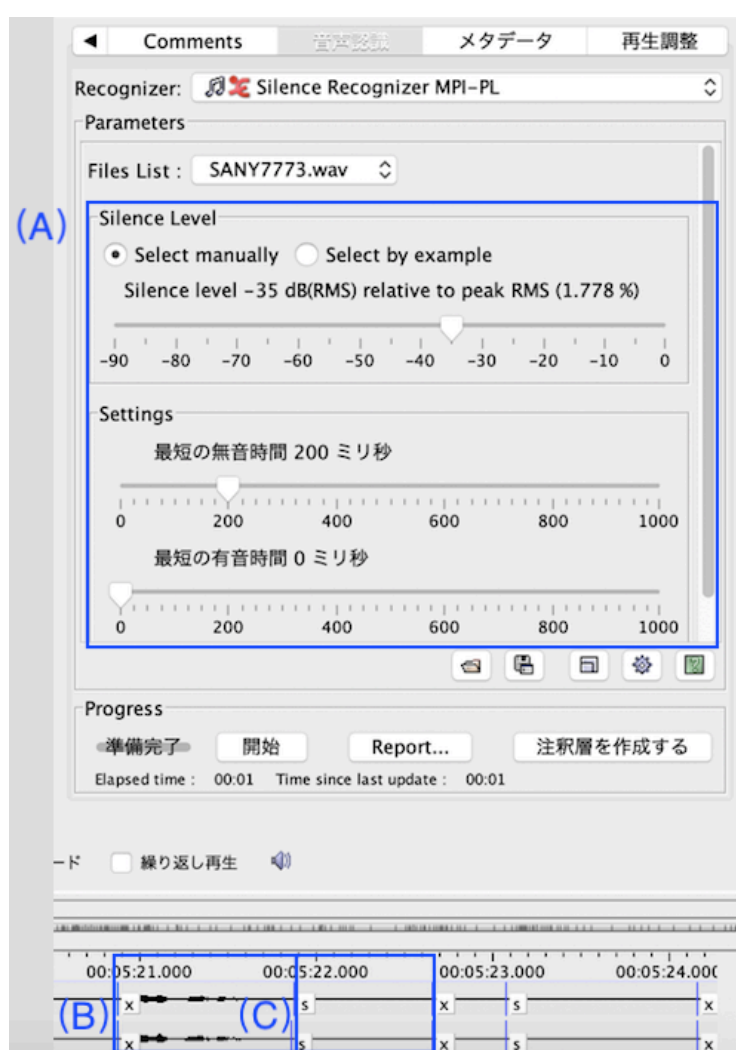


図 6: 無音区間自動認識機能：「無音音量レベル」，「最短の無音時間」および「最短の有音時間」の設定 (A)，「有音区間」(B)，「無音区間」(C)

3) ELAN による自動処理後に、無音区間、有音区間の分別に問題がないことを確認し、自動分割が正確でない箇所を手で修正した。また、笑い声や咳、溜息などの N 言語有音区間は発話区間の対象外にした。

4) NS の発話速度を分析のため、筆者が会話を聞きながら有音区間の中から NS だけの発話区間を残した。

5) 区間ごとに NS の発話セリフを書き起こした (図 7)。

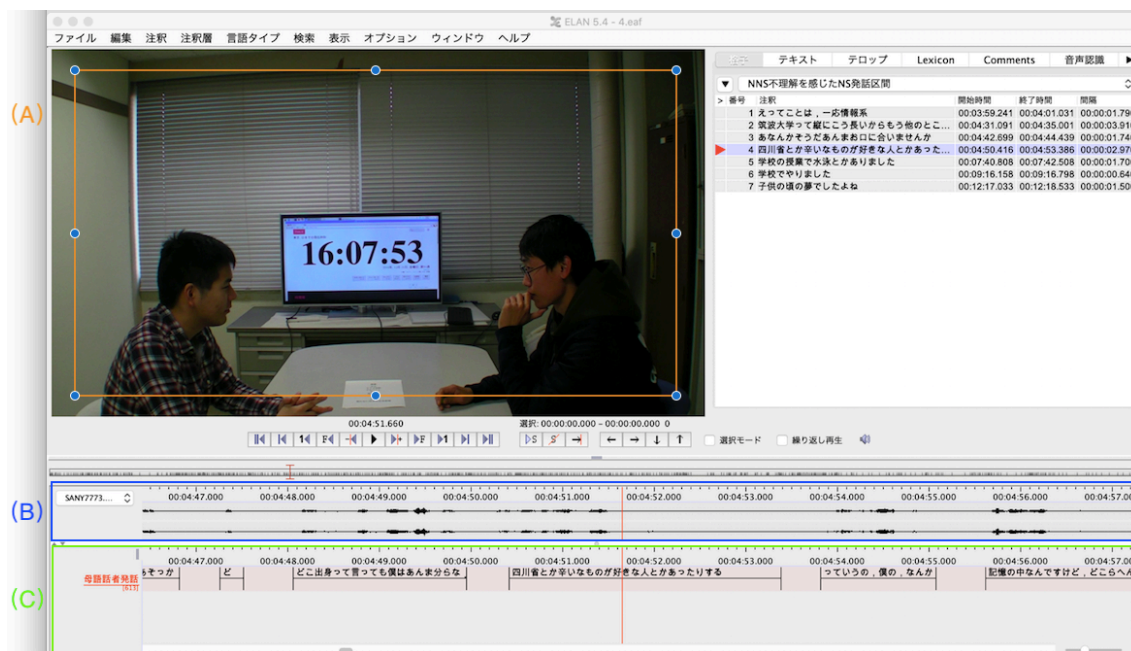


図 7: NS 発話区間のラベリング例:会話の映像データ (A), 会話音声データの波形 (B), NS 発話区間のラベリング (書き起こし済み) (C)

6) 3.6.4 章で述べた、記録できた NNS の不理解タイミングとスクリーンに映った日本時間と照らし合わせることにより、NNS が不理解を感じたところの位置を確定し、記録した NNS が不理解を感じた NS セリフを使い、NNS が不理解を抱いた NS の発話を特定できた。例として、不理解時点とセリフを確認し (図 8), NNS は日本時間 16:07:56 に flic を押下したことにより、実験ビデオに映ったスクリーンに 16:07:56 が表示されているところが特定できる。16:07:56 数秒程度前のビデオをチェックし、「四川省」という言葉が入っている発話を不理解発話として特定する (図 9)。

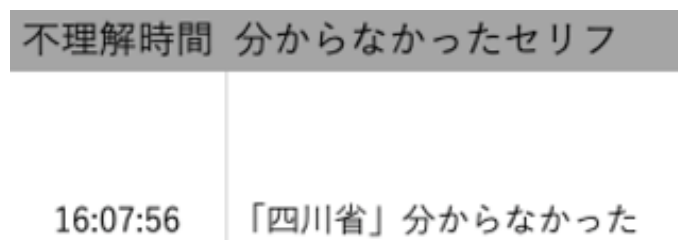


図 8: 不理解時点とセリフの例

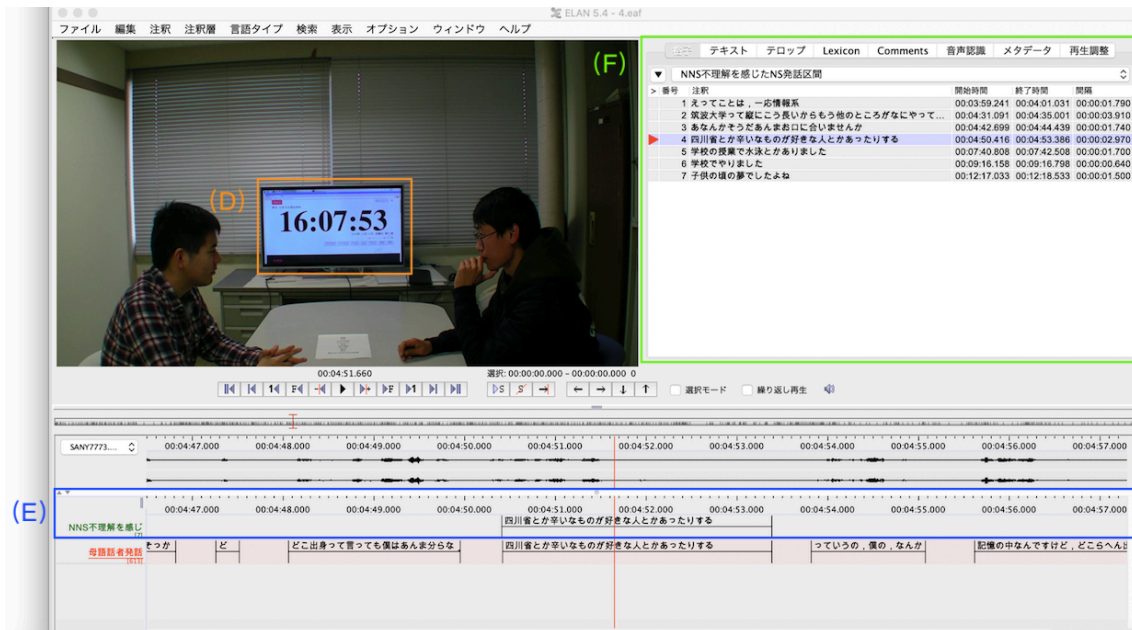


図 9: NNS が不理解を抱いた NS の発話区間例：日本時間 (D)，NNS 不理解区間ラベリング (E)，ラベリングした区間のリスト (F)

3.7.2 NS 話速の計算

発話はずっと同じ速度で行うわけではなく、丸島 [24] は発話中で速度の変化を測ったところ、自然発話の速度が常に変化していることを示した。会話中の NNS が困ったところの NS 話速を対象とし、リアルタイムの話速変化を調査するため、NS 発話全体の平均な話速ではなく、1 つ 1 つの発話区間に対しての調音速度 (Articulation rate)、ポーズ時間を除いた時間長で算出される話速を計算する。

本実験では、会話中不理解時点の記録および会話後のビデオ再生とインタビューによる不理解時点の確認を通じ、NNS が会話中感じた全部の不理解を確認できた。NS の 1 つの発話の長さ、不理解の単語が 1 つの発話における位置、NNS の判断時間などにより、NNS の flic 押下は不理解発話中、または不理解発話後に発生することが見られた (図 10)。不理解時の NS 話速を調査するため、不理解発話とその直前の発話の 2 つ発話区間の NS の話速計算し、変化を調べた。

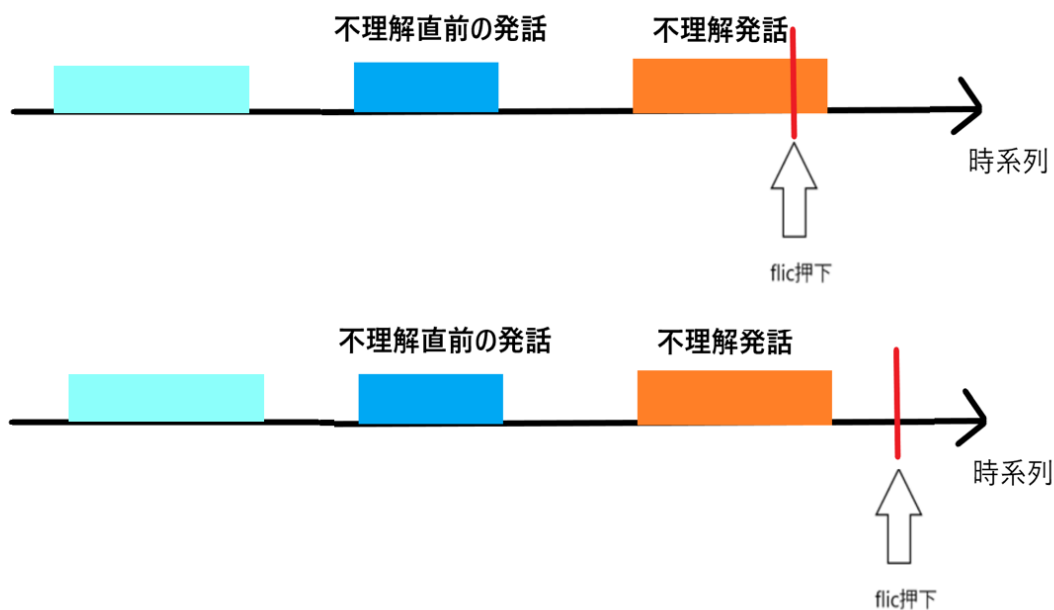


図 10: 不理解直前の発話と不理解発話と flic 押下

話速（音節数/秒）の定義としては NS の 1 発話の音節数（個）をその発話の時間長（秒）で割った数値である。音節数を数えた際に、ELAN で書き起こした NS の発話を平仮名に変更し、音節の定義に基づき、長音「ー」、促音「っ」、撥音「ん」を抜きで計算した。例えば、「おんせつ」という単語を音節で数えると 3 音節である。発話時間は ELAN が切り出した発話区間の時間長である。

3.8 結果

3.8.1 NNS が感じた不理解回数および原因

NNS と一緒に会話ビデオを見ながら、C. C. Goh. に挙げられた話し手方面の理解へ影響する原因：言葉が長い、言葉が複雑、ポーズが短い、発話速度が早いと、聞き手方面原因：ボキャブラリーの欠如、知っている言葉を認識できない、言葉を聞いてイメージ掴めない、会話の流れに乗れない、前の言葉の意味を考えて次の部分を無視する、予期しない単語の出現が混乱させる、前に問題があったため後続部分も理解できない、聞いたものを忘れるの原因 [10] を NNS に提示し、選択してもらい、インタビューをしながら不理解の原因を確認した。また同様に聞き手要素および話し手発話面要素の 2 要素に分類をした。

結果として NNS は主に「元々言葉は知らなかった」、「知ってる言葉への認識や反応が遅かった」、「会話の流れに乗れない」、「前の言葉の意味を考えて次の部分に反応できない」、「相手の発音の大きさ、語調などに影響される」、「相手の話速早かった」、「聞いていなかった」の理由で NS の発話に不理解を感じた。また、1 つの不理解に対して複数の理由があっ

たと答えたケースも見られた。それに、理由わかりませんという回答もあった。不理解原因とその原因で不理解を感じた回数を表1にまとめた。

表1 不理解原因

ペア番号	NS設定した発話速度の閾値
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
19	8
20	10
21	10
22	8
23	9
24	7
平均値	8.625

15組各10分間、合計150分間の自由会話において、NNSが不理解を感じた事例は115個のうち、単なる「NSの話速が早かった」が原因で不理解を感じたのが18回で、「複数の原因で不理解を感じた」と答えた不理解事例のうち、「NSの話速が早かった」を含んでいたのは6回であるため、NSの話速が早いという原因があったのは「単なる「NSの話速が早かった」が原因で不理解を感じた」の18回+「「複数の原因で不理解を感じたがそのうち「NSの話速が早かった」のも含まれた」の6回、合計24回である。トータル115回不理解の中、直接話速を原因としてNNSの不理解を起こした回数が24回で、割合は24/115、約21%である。

2割超のNNS理解困難問題もNSの早口がもたらしたことは、NSの早口問題を解決すれば、2割超のNNS理解困難問題が解消できると考えられる。

また、「知っている言葉への認識や反応が遅かった」、「会話の流れに乗れない」の理由づけにも話速の影響している可能性があるため、この2つの原因を加えて割合を計算すると、トータル115回不理解の中、直接話速を原因としてNNSの不理解を起こした回数24回、「知っている言葉への認識や反応が遅かった」25回、「会話の流れに乗れない」2回、合計51回で、割合は51/115、約44%である。割合がさらに大きくなったことで、話速の影響が大きいと言える。

3.8.2 不理解時のNS話速変化

Bloomfieldらに指摘された発話速度は他の要素とお互いに影響し合ってNNSにリスニング上の困難をきたす可能性が指摘されている[2]。そのことを考え、115個NSが不理解を感じた事例から、話速と関係の無い不理解原因が「元々言葉は知らなかった」(37回)と「聞いていなかった」(1回)である事例を除外した。そして不理解前の発話がないため分析できない「会話の一番最初のセリフに不理解を感じた事例」(1回)と「筆者が聞き取れないため書き起こしできなかった事例」(1回)を除き、NS話速変化の分析対象とする不理解事例は75個である。不理解事例に対して、NNSが不理解を感じたNS発話と不理解直前のNS発話の話速を計算、比較した。その結果は図11に示す。

NNSが聞き取れなかったNSの発話の話速は平均7.15音節/秒(SD=2.14)、その不理解直前のNS発話の話速は平均7.85音節/秒(SD=1.96)であった。

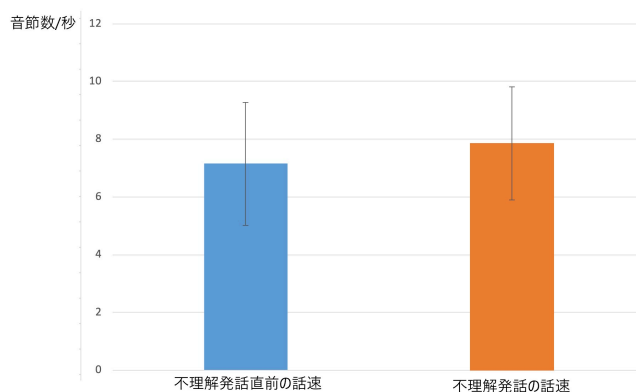


図 11: NNS が不理解を感じた時の NS 話速変化 (N=75)

Kolmogorov-Smirnov の正規性の検定において有意確率が.05 以上であるため、データは正規分布していると考えられ、対応のあるサンプルの t 検定を行った。結果として有意差が認められた ($t(74)=2.231, p<.05$)。

この結果から、NNS が不理解を感じた直前に NS の発話速度が速くなったことが分かった。

3.8.3 アンケート結果

アンケートの評価は図 12 に示す。

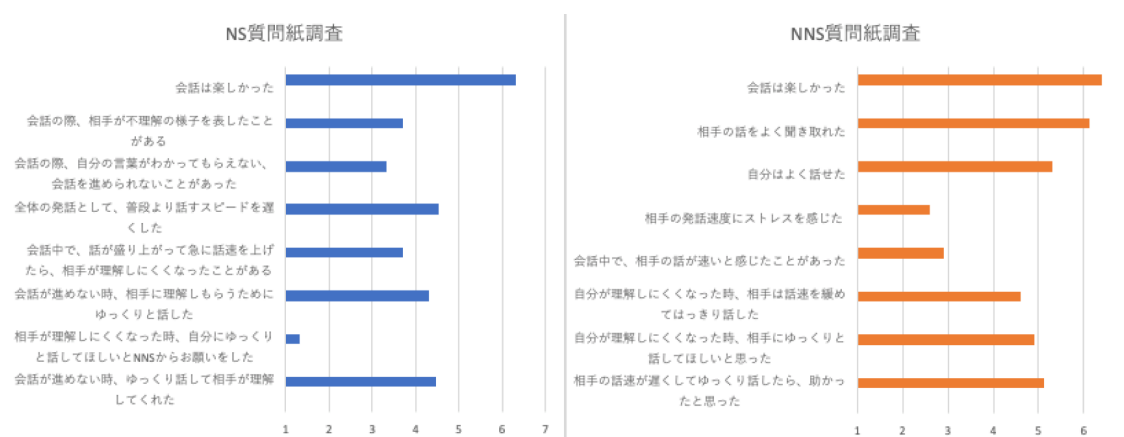


図 12: 質問紙調査 (NS: N=15, NNS: N=15)

コミュニケーションの状況及び会話全体への印象に関して、「会話は楽しかった」の項目に対して、NS と NNS 療法とも高い評価をしてくれた。しかしながら、「相手の話をよく聞き取れた」項目に対して、NNS は 6.1 で「当てはまる」以上の評価を得た一方、「会話の際、相手が不理解の様子を表したことがある」項目に対して、NS は 3.7 の「どちらとも言えない」寄りの評価をえた。

NS の話速に関する評価項目において、「全体の発話として、普段より話すスピードを遅くした」項目に対して、NS から平均 4.5 の「どちらとも言えない」と「やや当てはまる」中間の評価をえた。NNS と会話する際に、配慮をして普段よりゆっくり話す NS がいた。そして NNS も話し手の発話速度から受けたストレスが少なかった。NS がゆっくりして話すことで NNS が助かるかどうかを調査する項目、「会話が進めないとき、ゆっくり話して相手が理解してくれた」「相手の話速が遅くしてゆっくり話したら、助かったと思った」に対して、NS と NNS 両方ともに 4~5 点程度の評価を得られ、ゆっくり話すことは NNS に望まれ、会話を進めるのに一定の効果があることがわかった。一方、「自分が理解しにくくなったとき、相手にゆっくりと話してほしいと思った」項目に対して、NNS から 4.9 の「やや当てはまる」よりの評価をえたが、「相手が理解しにくくなったとき、自分にゆっくり話してほしいと NNS からお願いをした」項目に対して、NS に 1.3 「全く当てはまらない」よりの評価を得られた。これは NNS が不理解を感じ、相手にゆっくり話してほしいと思っても口に出さない可能性があると考えられる。

3.9 検討

本実験により，第二言語会話の場面では，NNSはNSの話速が早くなるときに理解困難を感じたケースが見られ，不理解を感じた直前にNSの話速が速くなったことがわかった。それに，NNSに対するインタビューにより，ボキャブラリーの欠如と知っている言葉への認識や反応が遅いとこの2つの理由より少なかったが，NNSの不理解の原因の2割超はNSの話速が原因とのことがわかった。またアンケートにより，会話中に発話に集中し話速に気づけなかった時が多く，気づいてもNSに「ゆっくり話してほしい」と言い出せないことがわかった。

これらの結果から，第二言語コミュニケーションを支援するために，NNSに代わり，NSが早口になった際にそのことを知らせる必要があると考えた。

第4章 Woz法による話速通知手法有効性の 検証

4.1 話速通知手法

NSの早口はNNSの理解困難を起こす要因の一つであることに対して、第二言語会話における話速の影響を考慮し、NSの早口をリアルタイムに知らせ、NSの自発調整を促す手法が叶らによって提案されている [5].

本手法は共通言語を用いるNSとNNSとのコミュニケーションで使うことを想定して、異言語コミュニケーション場面で使用することを前提とする。日本語NSと日本語NNSが円滑なコミュニケーションを可能とするために、NNSの理解困難である原因のNSの早口をリアルタイムに検知し、NSに知らせる話速通知手法である。

本手法が想定している応用場面では、参加者数はNNS1名とNS1名、そのうちNNSは共通言語で日常会話ができる程度の能力を持つ人である。

早口をNSに知らせることによりNSに自発的話速調整を促し、異言語コミュニケーション支援することはこの手法の目的である。言語能力の差が大きい異言語コミュニケーションにおいては、会話アンバランス、いわゆるNNSの不理解を引き起こす原因の一つであるNSの早い話速、をNSに知らせば、NSの自発的調整が期待でき、会話のアンバランスを改善すると考えられる。

4.2 Woz法の適用

Wizard of Oz (Woz) 法 [25] は、wizard と呼ぶ人間がシステムのふりをし、完全には開発されていないシステムの効果的なシミュレーションをする手法であり、ヒューマンコンピュータインタラクションの領域において、教示エージェントや対話エージェントのシミュレーション等にこの手法が用いられている [26].

本研究においては、話速通知手法を被験者実験で評価する際、実験阻害要因となる笑う声、雑音の認識ミスなどを排除し理想的なシステム動作を達成するためにつかわれる。

4.3 Woz法に基づく話速通知システムの実現

話速通知手法を以下の2つの設計方針をもとに、Woz法により実現した。

- (1) NSの話速をリアルタイムに計測する。
 - (2) NSにとって会話をしながら見やすいインタフェースを設計する。
- 理由としては、

(1) 本システムは、外国人留学生など簡単な日常会話を日本語でできる日本語 NNS と NS が共通言語の日本語を用いてコミュニケーションを取るときに利用し、NNS の不理解を起こす原因になる NS の早口を検知し、そのことを NS に通知することを目的とする。そのため、NS の話速をリアルタイムに計測することが必要である。

(2) 本システムは、対面環境において、会話をしながら利用されることを想定している。システムを利用することが会話の妨げになってはならない。一方、「今早口になったよ」という通知が見えることが必要である。

したがって、本実験で使ったシステムは「録音」、「話速閾値設定」、「音節認識」、「早口判定」、「データ保存」の5つの機能が実現されていた。

システムの構成を図 13 に示す。4.4 章、4.5 章、4.6 章、4.7 章に、それぞれの構成要素について述べる。

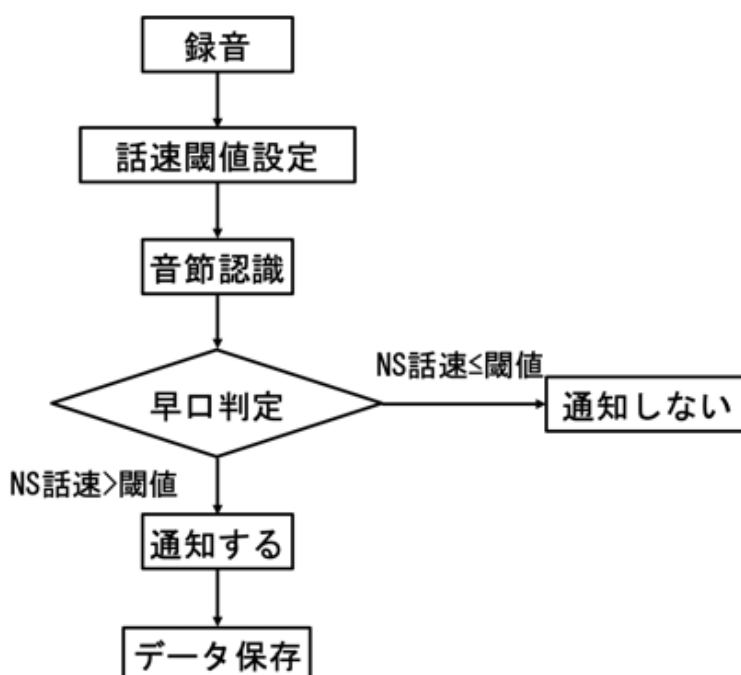


図 13: フローチャート

4.4 録音

録音に使っているマイクは SteelSeries Siberia Microphone white 50002 である。NS のみマイクを付け、録音された音声はデータ分析のためシステムに入力する。録音する音声の周波数は 44100Hz、ビット数 16bit、モノラル音声である。

4.5 発話速度閾値の設定

早口判断の際に、早口の閾値が必要となる。本研究では、同じNNSとしても語学力またはコミュニケーション力の差により、早口だと感じる話速が違ふと考え、事前にNNS一人一人に事前に作られた音声を聞かせ、適切な早口閾値を設定する。

発話速度閾値設定ツールはmatlabで作られていた。インターフェースは以下の図14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な音声サンプルが再生する。1つのボタンを押すと、8つの同じ話速の音声サンプルから1つをランダムに再生する。

音声サンプルを作成した際に、高橋研究室[33]が作った音声データベースの中から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を基にした。作成方法を説明するため、例を挙げる。例え12音節/秒のサンプルを作成する際に、該当音声サンプルを8つの部分を分け、各部分から120音節の内容をカットし、発音時間を10秒に変更した。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つの音声サンプルを用意した。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。

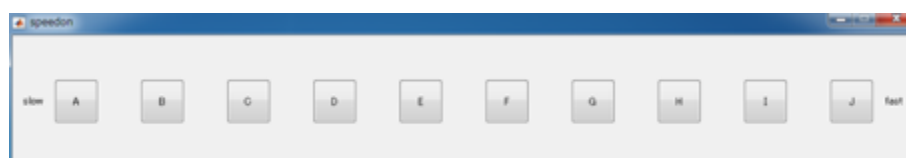


図 14: 発話速度閾値設定ツール

4.6 音声認識

NSの話速をリアルタイムに計測するため、図15のような音節測定補助キットを使う。音節測定補助キットはc#で作られた。

Wizardと呼ばれる実験者がNSの発話を認識し、話速を判断する。Wizardが1つの音節を聞くと、図14のtextboxにキーボードを1回押す。そして、システムは1秒ごとにtextboxの中身をクリアしながら、1秒ごとにタイピングされた文字列の長さを計算する。システムのこの動きにより、1秒ごとのキーボード押し回数が得られる。話速の定義によると、1秒間の音節数を数えるとその1秒間の話速が得られるため、1秒ごとに記録されたタイピング数は1秒間の音節数、つまりその1秒間の話速である。

4.7 早口判定

NNSに事前に選択した閾値を基準とし、NSの早口を判断する。NSの話速が閾値より上回ると、早口と判断し、通知画面が赤色になり、「FAST」との文字が表示される。超えていない時の通知画面は緑色に「GOOD」の文字で維持する。通知画面の変化は図16のように示す。

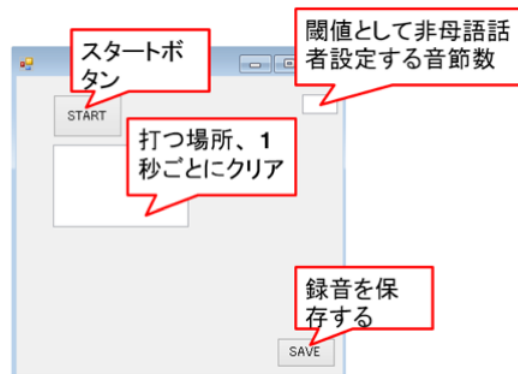


図 15: 音節測定補助キット

4.8 データ保存

データを確認、分析するため、録音された音声データ、カウントされた音節数が保存される。

4.9 実験

Woz 法に基づく話速通知システムを用い、話速通知手法の有効性を検証する。

4.9.1 目的

話速通知手法を使うことにより、NS が自分の話速を調整するかどうかを調査する。

4.9.2 実験参加者

実験参加者は初対面の日本語 NS20 名と日本語 NNS20 名、20 ペアである。

NS と NNS が二人一組で対面会話実験を行うため、NS は日本語を母語として使う者で、NNS は日本語で日常会話ができるが (N2 以上)、日本人と会話する際に不理解を感じる者が対象である。

NNS に NS と 5 分ほどの自己紹介や挨拶をしてもらい、NNS の日本語レベルを確認する。尾崎に指摘された NNS が困った場合の行動 [19] に基づき、会話中 NS に聞き返しをしたり、NS の発話に対して沈黙したり、困ったような表情や身振りをしたりした NNS だけを募集した。

4.9.3 比較条件

システムの有効性を確かめるために、非提示条件と提示条件に分けて 2 条件を設けた。

1) 非提示条件：提案システムが会話中に動作しない



図 16: 通知画面

2) 提示条件：提案システムが会話中に動作し、NS の話速が早すぎるときスクリーンに警告を表示する。スクリーンは NS と NNS 両者に見られるように机の上に設置する
各条件への参加順序はカウンターバランスをとり、各ペアに 5 分間の実験を 2 回行わせた。

4.9.4 実験内容

NS1 名と NNS1 名による対面会話である。各ペアに話速通知システムを使う使わない 2 条件に分け、条件毎に 1 回ずつ 5 分間の会話実験を行う。

会話内容は時間制限と目標が設けられたお題当てゲームである。ゲームのルールは NHK 番組の連想ゲーム [27] より参照し、キーワード 1 個を当てると 10 点を獲得でき、違反ヒントを出すと 5 点減点されるとした。150 点以上獲得できると点数により謝礼金が増えることを被験者に伝えた。5 分間に高い得点を獲得するために、NS は NNS 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、ジャスチャは禁止とし、全部言葉で表現することにした。

お題当てゲームで使われたキーワードは全部日本語の教科書「みんなの日本語」初級 [28] 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。

4.9.5 実験機材

4.9.5.1 早口通知に用いた機材

警告標識を提示するためのモニターは ProLite XB2783HSU-B1 27 インチで、画面の解像度は 1920 × 1080 である。

4.9.5.2 映像音声の記録に用いた機材

NS 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。録音は 10000Hz のモノラル音声である。

撮影機械について、一台は SANYO Xacti HD 2000 Full，撮影する画像の比率は 4:3 で、解像度は 1920*1080 であり、二台は SONY HDR-CX560v，撮影する画像の比率は 4:3 で、解像度は 1440*1080 である。

4.9.6 実験環境

対面に座る 2 者が会話する。1 台のビデオカメラで全景を撮影し、1 台のビデオカメラで NS に向けて上半身を撮影すし、1 台のビデオカメラで NNS に向けて上半身を撮影する。図 17, 18, 19, 20 が実験配置と環境様子を示している。会話ペア相互の対人距離を 120cm とする [17]。

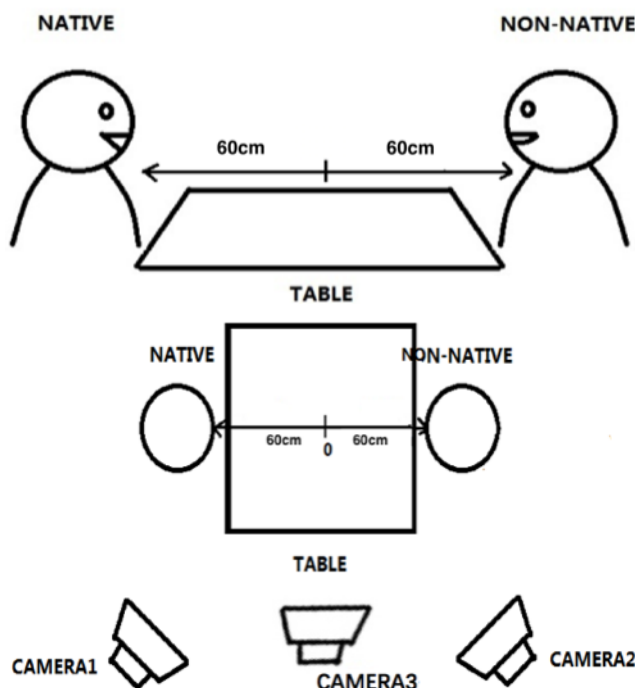


図 17: 非提示条件の配置図

非提示条件の実験様子は図 18 より示す。



図 18: 非提示条件の実験様子

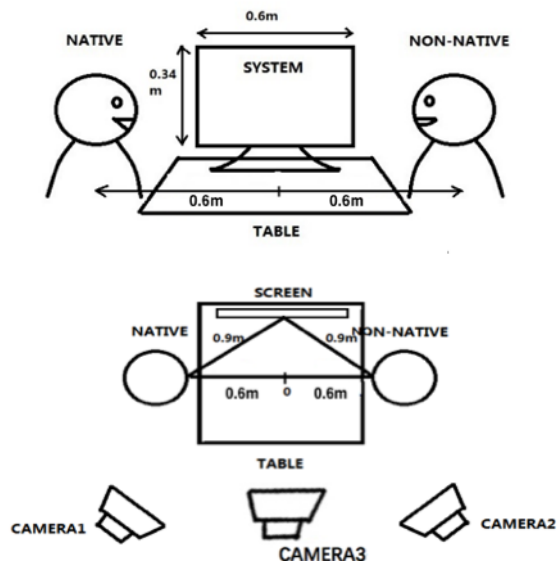


図 19: 提示条件の配置図

提示条件の実験様子は図 20 より示す。



図 20: 提示条件の実験様子

4.9.7 手順

- 1) NSが入室する前に、実験部屋の中で、4.5章に言及した話速閾値設定ツールを用い、NNSに適切な発話速度を設定してもらおう。事前に作られた違う発話速度の音声サンプルを聞かせ（NNSに自由に音声サンプルを流すボタンを押させる）、一番理解しやすいと思った発話速度を教えてください、警告提示の閾値としてシステムの中に入力した。
- 2) 実験部屋の中で、被験者(NS, NNS一人ずつ)に実験についての説明を聞いてもらう。
- 3) 同意書を読み上げ、被験者全員の同意を得てから署名してもらう。
- 4) NSは、音声データを取得するためのマイクをそれぞれ着用する。NSの発話を録音する。
- 5) NSにキーワードを記載されたメモ用紙を渡し、1分間の準備時間を与える。
- 6) 提示条件、非提示条件の2条件で、条件毎に一回ずつ5分間のお題当てゲームをする。各条件には異なるメモ用紙を使う。各条件への参加順序はカウンターバランスをとる。
- 7) 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。NNSがキーワードを当てた時に10点を加点し、NSが違反ヒントを出すと5点を減点し、各条件の成績を成績表に記録する。
- 8) 会話直後に、被験者に会話に関するアンケートを記入してもらう。

4.9.8 データ取得

4.9.8.1 映像音声の記録

実験の様子は4.9.5.2章に述べた映像音声の記録に用いた機材、カメラとピンマイクで記録する。

4.9.9 データ処理

4.9.9.1 NS 発話区間，通知区間のラベリング

NS に早口を通知することにより，NS が自ら話速を調整してくれるかどうかを確認するため，早口通知が出る前後の NS の発話に対して調査することは必要である。

手続きとして，

1) 実験中に使ったシステムスクリーンの表示変化を ELAN で照らし合わせることにより，通知区間を特定し，ラベリングする（図 21）。

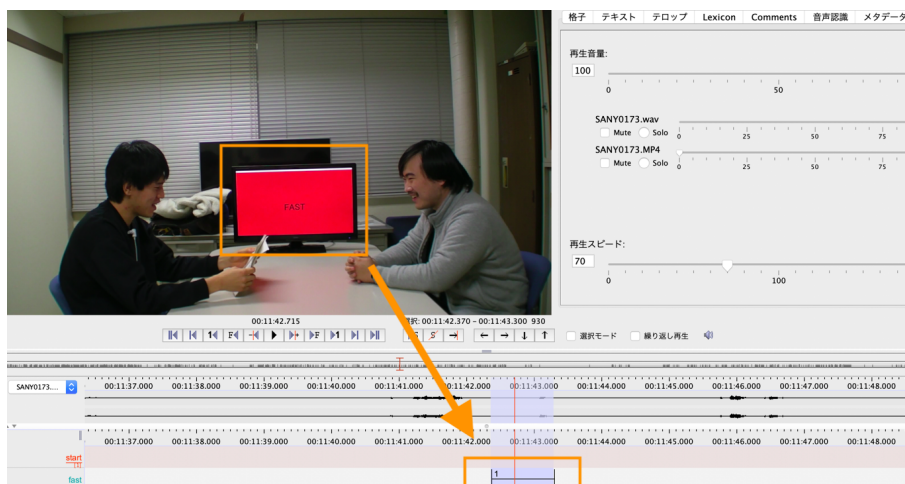


図 21: 通知区間のラベリング

2) 3.7.1 のように，ELAN を用いて NS の発話区間をラベリングし，通知前後の NS 発話区間に対してセリフを書き起こす（図 22）。

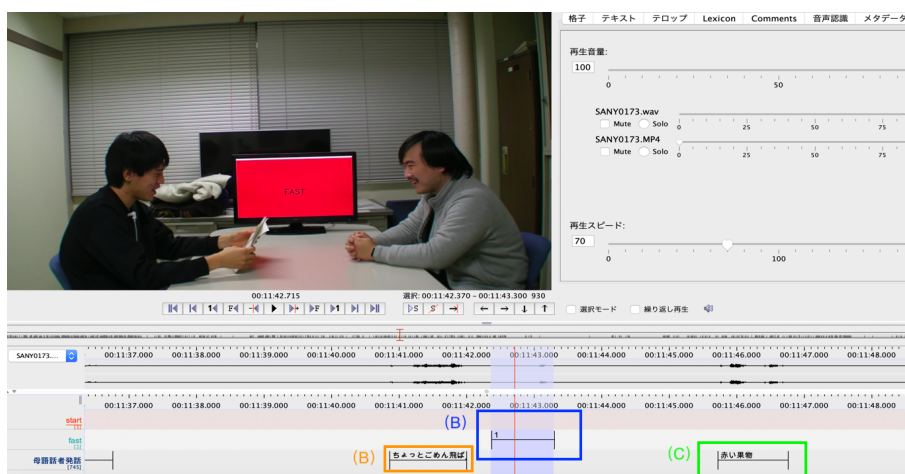


図 22: 通知前後の NS 発話区間ラベリング：通知直前の NS 発話区間 (A)，通知区間 (B)，通知直後の NS 発話区間 (C)

4.9.10 結果

4.9.10.1 NS の発話速度変化

早口通知がある提示条件の 20 組データ，つまり 100 分間の会話を用いて NS の話速を分析した．合計 63 回の通知事例に対し，通知直前の発話と通知直後の発話に対して NS の話速を計算した．

ここで認識判定の誤りがあり，早口の閾値を達していないのに早口通知を出した事例を除外するため，ペアごとに NNS に選定してもらった閾値と通知直前の発話区間の話速と比較し，閾値を超えていなかったデータ 10 回を除いた．NNS が設定した発話速度の閾値を表 2 にまとめた．

表2 NNSが設定した発話速度の閾値

ペア番号	NS設定した発話速度の閾値
1	7
2	8
3	7
4	6
5	7
6	7
7	7
8	6
9	7
10	7
11	7
12	7
13	7
14	7
15	7
16	7
17	7
18	7
19	7
20	7
平均値	6.95

残りの 53 回の通知に対し，通知直前の発話区間と通知直後の発話区間の話速を比較した結果，通知直前の NS 話速は平均 9.37 音節/秒 (SD=3.35)，通知直後の NS 話速は平均 6.87 音節/秒 (SD=2.18) であった (図 23)。

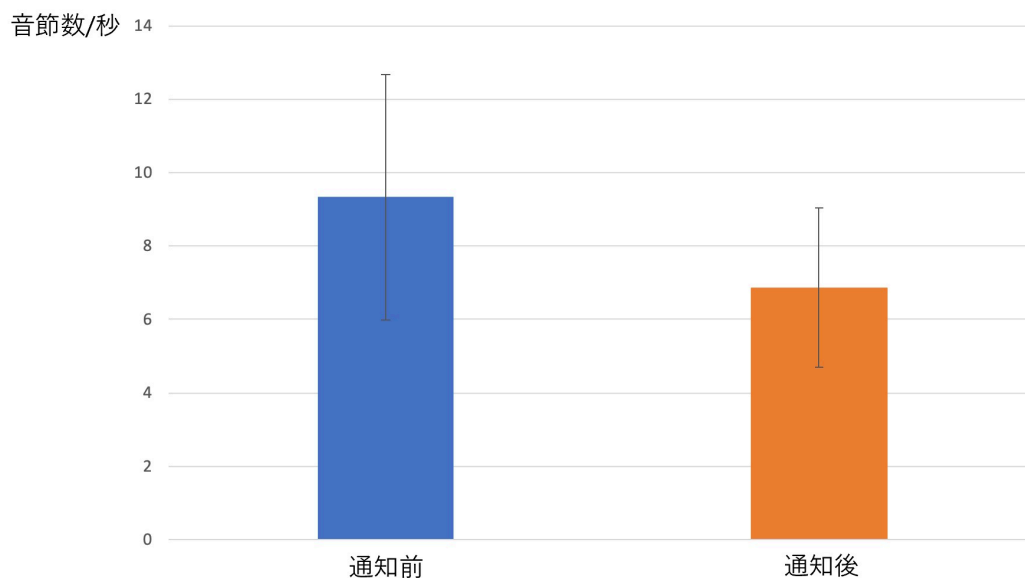


図 23: 通知前後の NS 話速 (通知前の NS 話速: N=53, 通知後の NS 話速: N=53)

Kolmogorov-Smirnov の正規性の検定において有意確率が .05 以下であるため，データは正規分布していないと考えられ，通知前後の NS 話速を Wilcoxon 符号付順位検定を用いて検定した結果，有意差が認められ ($Z=-5.759$, $p=0.000<.05$)，通知が出た直前より通知直後の NS の発話速度が遅くなったことが分かった。

4.9.11 検討

Woz 法に基づく話速通知システムを用い，NS が早口になった際にそのことを気させる話速通知手法の有効性を検証した。通知前後の NS 話速を比較したところ，通知前の発話速度より通知後が遅くなった結果がわかった。この結果から NS は通知画面の変化に気づき，自発的に発話調整を行ったと考えられ，話速通知手法は第二言語会話支援に有効であると判断した。

しかし，Woz 法に基づく話速通知システムは常に第三者である wizard での操作が必要で，あくまでもシステム開発前のシミュレーションであるため，話速通知手法を具現化した自動システムを開発する必要性がある。

第5章 話速通知システム有効性の検証

5.1 話速通知システムの概要

前章記述したように，Woz法に基づく話速通知システムを用い，NSが早口になった際にシステムの通知画面に気づくことにより，自発話速調整が見られた．この結果を踏まえ，自動的話速通知システムの有効性と実用性が期待できると考え，Ye，井上は音声認識を使った自動処理できる話速アウエネスシステムを構築した [6]．

システムは音声メディア処理で，NSの話速を自動的に検知し，話速が速いと判断した際に通知を出してNSに気づかせる機能が実現されている．システム流れは図24のように示す．システムはWindows7をソフトウェア環境として使い，数値計算言語にはMathWorks社のMATLAB R2015aで構築された [6]．



図 24: 自動処理話速通知システム

通知画面は4章に述べたWoz法に基づく話速通知システムと同様である．

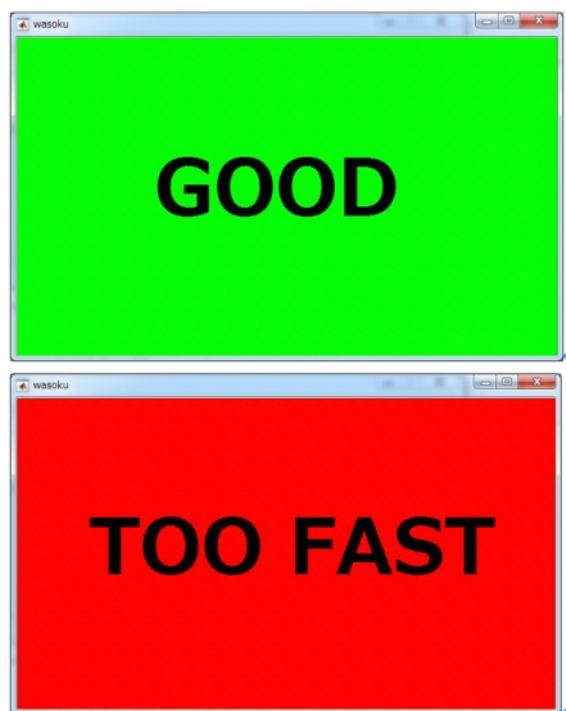


図 25: 通知画面

5.2 実験

5.2.1 目的

話速通知システムを使うことにより，NSが自分の話速を調整するかどうかを調べるのが目的である。

5.2.2 実験参加者

実験参加者は初対面の日本語 NS24名と日本語 NNS24名，24ペアである。

NSとNNSが二人一組で対面会話実験を行うため，NSは日本語を母語として使う者で，NNSは日本語で日常会話ができるが（N2以上），日本人と会話する際に不理解を感じる者が対象である。

NNSにNSと5分ほどの自己紹介や挨拶をしてもらい，NNSの日本語レベルを確認する。尾崎に指摘されたNNSが困った場合の行動[19]に基づき，会話中NSに聞き返しをしたり，NSの発話に対して沈黙したり，困ったような表情や身振りをしたりしたNNSだけを募集した。

5.2.3 比較条件

システムの有効性を確かめるために，非提示条件と提示条件に分けて2条件を設けた。

- 1) 非提示条件：提案システムが会話中に動作しない
 - 2) 提示条件：提案システムが会話中に動作し、NSの話速が早すぎるときスクリーンに警告を表示する。スクリーンはNSとNNS両者に見られるように机の上に設置する
- 各条件への参加順序はカウンターバランスをとった。各ペアに5分間の実験を2回行わせた。

5.2.4 実験内容

本実験は4.5.4章と同様に、NS1名とNNS1名による対面会話である。各ペアに自動処理ができる話速通知システムを使う使わない2条件に分け、条件毎に1回ずつ会話実験を行う。そして会話タスクは同じく5分間のお題当てゲームにした。

5.2.5 実験機材

5.2.5.1 早口通知に用いた機材

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。

5.2.5.2 映像音声の記録に用いた機材

NS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。録音は10000Hzのモノラル音声である。

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。

5.2.6 実験環境

対面に座る2者が会話する。1台のビデオカメラで全景を撮影し、1台のビデオカメラでNSに向けて上半身を撮影すし、1台のビデオカメラでNNSに向けて上半身を撮影する。図26, 27, 28, 29が実験配置と環境様子を示している。会話ペア相互の対人距離を120cmとする。

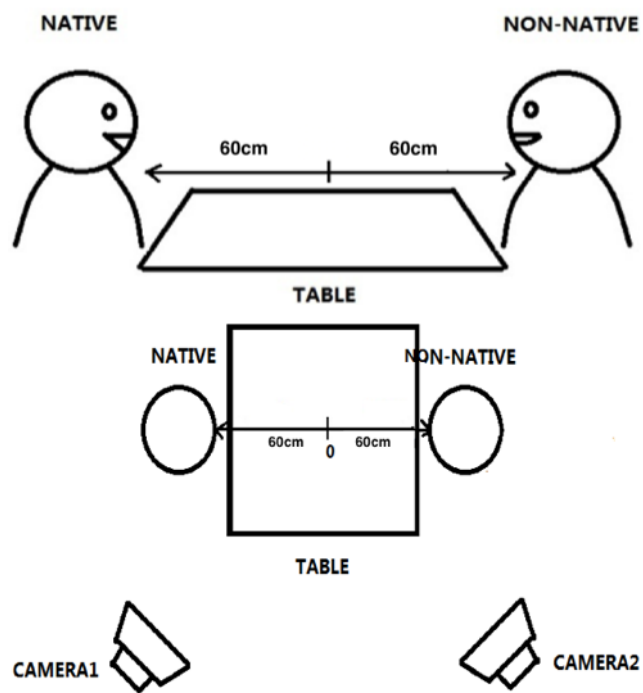


図 26: 非提示条件の配置図

非提示条件の実験様子は図 27 より示す。



図 27: 非提示条件の実験様子

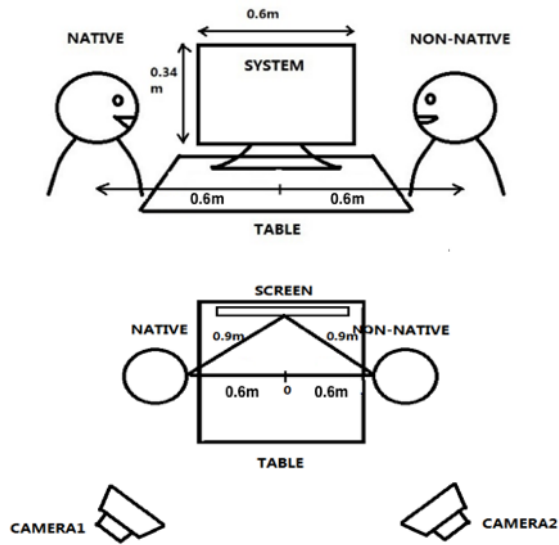


図 28: 提示条件の配置図

提示条件の実験様子は図 29 より示す。



図 29: 提示条件の実験様子

5.2.7 手順

1) NS が入室する前に、録画用のスペースの中で、4.4.3.1 章に言及した話速閾値設定ツールを用い、NNS に適切な発話速度を設定してもらう。事前に作られた違う発話速度の音声サンプルを聞かせ (NNS に自由に音声サンプルを流すボタンを押させる)、一番理解しやすいと思った発話速度を教えてもらい、警告提示の閾値としてシステムの中に入力する。

2) 録画用のスペースの中で、被験者 (NS, NNS 一人ずつ) に実験についての説明を聞いてもらう。

3) 同意書を読み上げ、被験者全員の同意を得てから署名してもらう。

4) NS の発音を認識するため母音標準パターン作成する。NS に母音「a」「i」「u」「e」「o」を三回発音してもらい、外部 excel ファイルに保存する。インターフェースは図 30 のように表す。REC ボタンを押すと 15 秒間録音をし始め、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するために使われることになる。この手続きは 5 分程度かかる。

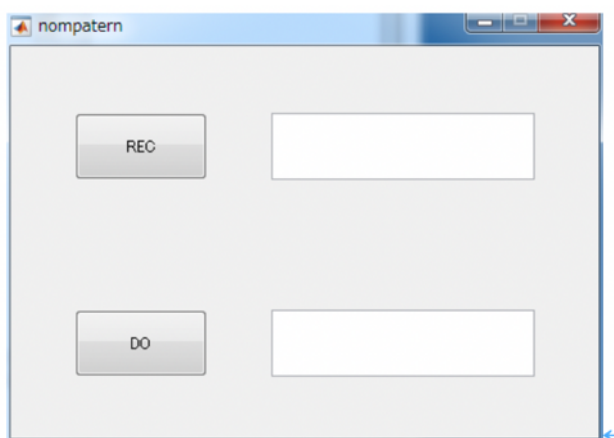


図 30: 母音標準パターン作成

5) NS のみにマイク (SteelSeries Siberia Microphone white 50002) を付けてもらい、録音をする。

6) NS にキーワードを記載されたメモ用紙を渡し、1 分間の準備時間を与える。

7) 提示条件、非提示条件の 2 条件で、条件毎に一回ずつ 5 分間のお題当てゲームをする。各条件には異なるメモ用紙を使う。各条件への参加順序はカウンターバランスをとる。

8) 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。NNS がキーワードを当てた時に 10 点を加点し、NS が違反ヒントを出すと 5 点を減点し、各条件の成績を成績表に記録する。

9) 会話直後に、被験者に会話関するアンケートを記入してもらう。

5.2.8 データ取得

5.2.8.1 映像音声の記録

実験の様子は 5.2.5.2 章に述べた映像音声の記録に用いた機材、カメラとピンマイクで記録する。

5.2.9 データ処理

24組各10分間、合計240分間の音声付きの実験ビデオデータが取得できたが、まだ分析途中のため、現時点では18組各10分間、合計180分間のデータを編集せずに用い、6.7.8章と同様にELANを用いてNS発話区間と通知区間をラベリングし、NS発話を書き起こした。

5.2.10 結果

5.2.10.1 NSの発話速度変化

早口通知がある提示条件の24組データ、つまり120分間の会話を用いてNSの話速を分析した。通知回数が合計336回であり、NS発話、通知区間をラベリングした後、通知直前の発話区間と通知直後の発話区間に対してNSの話速を計算した。

ここで認識判定の誤りがあり、早口の閾値を達していないのに早口通知を出したケースを除外するため、ペアごとにNNSに選定してもらった閾値と通知直前の発話区間の話速と比較し、閾値を超えていなかったデータ(255回)を除いた。NNSが選定した閾値を表3にまとめた。

表3 NNSが設定した発話速度の閾値

ペア番号	NS設定した発話速度の閾値
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
19	8
20	10
21	10
22	8
23	9
24	7
平均値	8.625

残りの 81 回の通知に対して、通知直前の発話区間と通知直後の発話区間の話速を比較した結果、通知直前の発話区間の速度は平均 9.85 音節/秒 (SD=1.87)、通知直後の発話の速度は平均 6.83 音節/秒 (SD=1.88) である (図 31)。

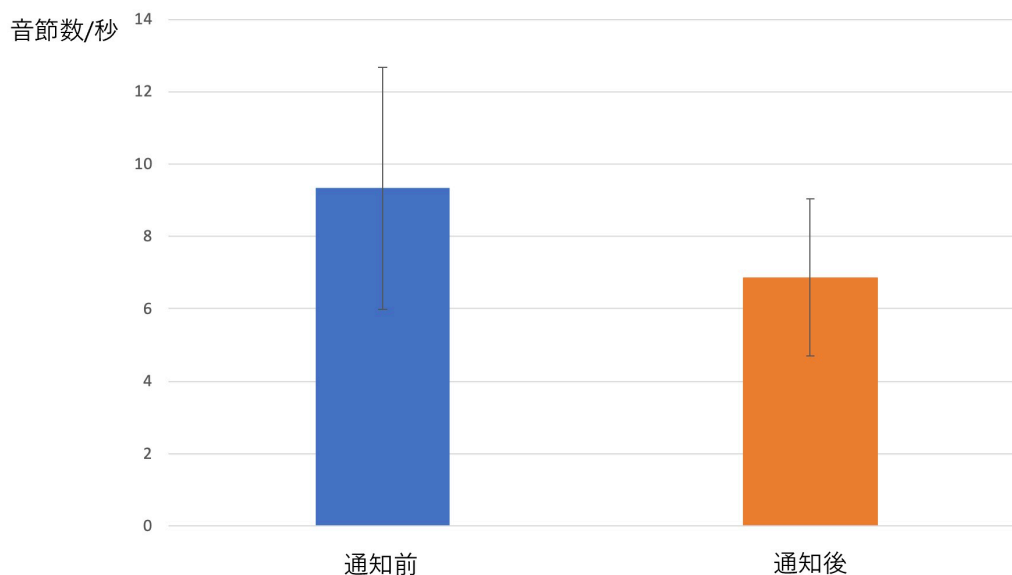


図 31: 通知前後 NS の話速 (通知前の NS 話速: N=81, 通知後の NS 話速: N=81)

Kolmogorov-Smirnov の正規性の検定において有意確率が.05 以下であるため、データは正規分布していないと考えられ、通知前後の NS 話速を Wilcoxon 符号付順位検定を用いて検定した結果、有意差が認められ ($Z=-7.137$, $p=0.000<.05$)、通知が出た直前より通知直後の NS の発話速度が遅くなったことが分かった。

第6章 話速通知手法が非母語話者理解への影響

話速通知手法を利用する際に、会話がどのように変化しているかを調査するため、実際の第二言語会話の場面を設け、話速通知手法を使う条件と使わない条件に分けて非母語話者の不理解を記録した。

6.1 実験参加者

実験参加者は初対面の日本語 NS3 名と日本語 NNS3 名、3 ペアである。

NS と NNS が二人一組で対面会話実験を行うため、NS は日本語を母語として使う者で、NNS は日本語で日常会話ができるが (N2 以上)、日本人と会話する際に不理解を感じる者が対象である。

NNS に NS と 5 分ほどの自己紹介や挨拶をしてもらい、NNS の日本語レベルを確認する。尾崎に指摘された NNS が困った場合の行動 [19] に基づき、会話中 NS に聞き返しをしたり、NS の発話に対して沈黙したり、困ったような表情や身振りをしたりした NNS だけを募集した。

6.2 比較条件

システムの影響を確認するため、非提示条件と提示条件に分けて 2 条件を設けた。

- 1) 非提示条件：提案システムが会話中に動作しない
 - 2) 提示条件：提案システムが会話中に動作し、NS の話速が早すぎるときスクリーンに警告を表示する。スクリーンは NS と NNS 両者に見られるように机の上に設置する
- 各ペアに 5 分間の実験を 2 回行わせた。

6.3 実験デザイン

NS1 名と NNS1 名、各ペアに対面による 5 分間のお題当てゲームを提示条件と非提示条件に分けて 1 回ずつ実施する。

本実験では 4 章で言及した WOZ 法に基づく話速通知システムを用い、母語話者の早口を知らせる。

ゲーム中、NNS が NS の言葉に対して不理解を感じる際、ボタンを押してもらい、不理解の時間を記録する。

6.4 実験機材

6.4.1 NNS 不理解を記録に用いた機材

3.3.1 章と同じである。

6.4.2 早口通知に用いた機材

4.9.5.1 章と同じである。

6.4.3 映像音声の記録に用いた機材

4.9.5.2 章と同じである。

6.5 実験環境

4.9.6 章と同じである。

6.6 実験手順

1) NS が入室する前に、実験部屋の中で、4.5 章に言及した話速閾値設定ツールを用い、NNS に適切な発話速度を設定してもらおう。事前に作られた違う発話速度の音声サンプルを聞かせ（NNS に自由に音声サンプルを流すボタンを押させる）、一番理解しやすいと思った発話速度を教えてもらい、警告提示の閾値としてシステムの中に入力した。

2) 実験部屋の中で、被験者 (NS, NNS 一人ずつ) に実験についての説明を聞いてもらおう。

3) 同意書を読み上げ、被験者全員の同意を得てから署名してもらおう。

4) NS は、音声データを得るためのマイクをそれぞれ着用する。NS の発話を録音する。

5) NNS の母語で NNS にのみ、NS との会話で不理解があった際につき一回 Bluetooth 接続ボタン押下式デバイスを押下するように伝える。そして NS に見えないように机の下で NNS に Bluetooth 接続ボタン押下式デバイスを手渡す。また、正常に押下タイミングが記録されるか確認するため NNS にその場で一回以上試し押しをするよう指示をする。

6) NS にキーワードを記載されたメモ用紙を渡し、1 分間の準備時間を与える。

7) 提示条件、非提示条件の 2 条件で、条件毎に一回ずつ 5 分間のお題当てゲームをする。各条件には異なるメモ用紙を使う。

8) 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。NNS がキーワードを当てた時に 10 点を加点し、NS が違反ヒントを出すと 5 点を減点し、各条件の成績を成績表に記録する。

9) 会話直後に、被験者に会話関するアンケートを記入してもらおう。

6.7 データ取得

6.7.1 映像音声の記録

実験の様子は 6.4.3 章に述べた映像記録に用いたカメラで記録する。

6.7.2 Bluetooth 接続ボタン押下式デバイスで取れた時間ログの記録

3.6.2 章に述べたように、ボタン押しの時間ログが記録に用いた機材に用いた Bluetooth 接続ボタン押下式デバイスで取得できる。時間ログと実験時間を照らし合わせることにより NNS が各条件の実験でボタンを押した回数が数えられる。

6.8 結果

提示条件と非提示条件の 2 条件に分けて非母語話者がボタンを押した回数を表 4 にまとめた。提示条件の場合、NNS が不理解を感じてボタンを押した回数が平均 5.67 回であり、非提示条件の場合平均 6.33 回より少なかったことがわかった。

表4 非母語話者がボタンを押した回数

ペア番号	提示条件	非提示条件
1	13	13
2	2	3
3	2	3
平均回数	5.67	6.33

6.9 検討

話速通知手法の評価に、NNS の不理解に与える影響について追加実験を実施した。NNS が NS の言葉に不理解を感じた際にボタンを押す手法で NNS の不理解を記録し、早口提示条件と非提示条件の NNS 不理解を比較した。

その結果、話速通知手法を応用した提示条件と比べ、話速通知手法を応用しない非提示条件の方において、非母語話者がボタンを押した回数が高くなる傾向があり、つまり、話速通知手法を用いることにより NNS が不理解を感じた回数が減少する傾向があると考えられる。この結果は実験 3 組で得られたもので、今後は実験数を増やして検討することが望まれる。

第7章 考察とまとめ

本研究は母語話者に自発的話し速調整を促すことにより、母語話者と非母語話者による会話の支援を目的としている。本研究では2つのポイントがある。1つ目は Bloomfield らに指摘された第二言語コミュニケーションの場合 NNS の理解度を影響する要素に基づき、NS のリアルタイムな話し速変化が NNS の理解困難を起こす原因の1つであることを客観的に証明したこと。2つ目はその知見から提案された話し速通知手法の有効性を Woz 法および実機により調査を行なったこと。

調査実験にはまずブルーテラス接続ボタンを用い、NNS の不理解時間を記録し、NNS の不理解が発生前の NS 発話を特定し、調査をした。また、会話後に NNS と共に録画されたビデオを見直し、不理解原因をインタビューをした。結果としては、NNS が不理解を感じた前に、NS の話し速が早くなっていたことと、NNS の不理解の原因の2割超が NS の早口であったことが分かった。

この知見から、NS の話し速を検測し、早口になった際にそのことを知らせる手法、いわゆる話し速アウエネス手法が井上らに提案された。

自動認識システムを構築する前に、実験阻害要因となる笑う声、雑音の認識ミスなどを排除し、理想的なシステム動作を達成するために Wizard of Oz (Woz) 法 [] を用い、話し速アウエネス手法の有効性を調査した。通知前後の NS 話し速を比較するところ、通知直前より NS の話し速が遅くなった。つまり、NS は通知に気づき、自発的に発話速度の調整をしたと考えられる。

そのことを踏まえ、井上らは音声メディア処理で、NS の話し速を自動的に検知し、話し速が速いと判断した際に通知を出し、NS に気づかせる機能が搭載された自動処理話し速アウエネスシステムを実装した。実際の第二言語コミュニケーション場面において、NS と NNS に会話させ、通知前後の NS 発話速度を調査した結果、通知直前より NS の話し速が遅くなった。つまり、NS は自動処理話し速アウエネスシステムの通知にも気づき、自発的に発話速度の調整をしたと考えられる。

謝辞

最後に、本研究を進めるにあたって、始終ご指導ご鞭撻等を頂きました指導教員の井上智雄教授に心より感謝を申し上げます。また、本研究に非常に有意なご意見をくださった副指導教員の叶少瑜助教に深く感謝いたします。また、三年の間、最初から最後まで共に頑張っておられる同研究室の皆様、および実験にご協力をくださった参加者の皆様に心から感謝しております。ありがとうございました。

参考文献

- [1] 2017年訪日外客数（総数）出典：JTB総合研究所＜
<https://www.tourism.jp/tourism-database/stats/inbound/>>
- [2] Bloomfield, Amber, et al, 「What makes listening difficult? Factors affecting second language listening comprehension」, MARYLAND UNIV COLLEGE PARK, 2010.
- [3] Griffiths, Roger, 「Speech rate and NNS comprehension: A preliminary study in time - benefit analysis.」, Language Learning, Vol.40, No.3, 1990, pp.311-336.
- [4] Zhao, Yong, 「The effects of listeners' control of speech rate on second language comprehension.」, Applied linguistics, Vol.18, No.1, 1997, pp.49-68.
- [5] 叶環, 井上智雄, 「会話弱者のための話速アウェアネスシステムの提案」, 研究報告グループウェアとネットワークサービス (GN), Vol.17, 2015, pp.1-6.
- [6] Ye Jing, Inoue Tomoo, 「A speech speed awareness system for non-native speakers」, Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion (CSCW'16 Companion), 2016, pp.49-52.
- [7] Hanawa Hiromi, Song Xiaoyu, Inoue Tomoo, 「Keyword generation by native speaker is quick and useful in conversation between native and non-native speaker」, Proceedings of the 2017 IEEE 21th International Conference on Computer Supported Cooperative Work in Design (CSCWD 2017), 2017, pp.145-150.
- [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, Vol.52, No.3, 2011, pp.1213-1223.
- [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), 2009, pp.719-728.
- [10] 池上真人, 「英語リスニングの習得段階を測る試み-ポーズと発話速度を変化させた内容把握問題を用いて.」, 言語文化研究, Vol.32, No.1-1, 2012, p59-88.
- [11] Goh, C. C., 「A cognitive perspective on language learners' listening comprehension problems」, System, Vol.28, No.1, 2000, pp.55-75.
- [12] Hayati, A, 「The effect of speech rate on listening comprehension of EFL learners」, Creative Education, Vol.1, 2010, pp.107-114.

- [13] Jones, Caroline, Lynn Berry, and Catherine Stevens, 「Synthesized speech intelligibility and persuasion: Speech rate and non-native listeners.」, *Computer Speech & Language*, Vol.21, No.4, 2007, pp.641–651.
- [14] Matsuura, Hiroko, et al, 「Accent and speech rate effects in English as a lingua franca.」, *System* 46, (2014), pp.143–150.
- [15] 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*, Vol.96, No.11, 2013, pp.874–881.
- [16] 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), 2001, pp.918–926.
- [17] 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*, Vol.105, No.479, 2005, pp.103–108.
- [18] 木田章義, 「音節とモーラ」, *音声研究*, Vol.2, No.1, pp.41–49.
- [19] 尾崎明人, 「接触場面における在日ブラジル人の「聞き返し」とその回避方略」, *社会言語科学*, Vol.4, No.1, 2001, pp.81–90.
- [20] 俵山雄司, 「留学生対象の会話における講師の言語調整行動と意識との関連—留学生向けの教養・専門科目会話の方法を検討するために」, *群馬大学留学生センター論集*, Vol.8, 2008, pp.13–29.
- [21] 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, 2013, pp.1794–1806.
- [22] メイナード泉子, 「会話分析」, くろしお出版, 1993.
- [23] Koiso, H., Horiuchi, Y., Tutiya, S., Ichikawa, A., and Den, Y., 「An analysis of turn-taking and backchannels based on prosodic and syntactic features in Japanese map task dialogs」, *Language and Speech*, Vol.41, No.3-4, 1998, pp.295–321.
- [24] 丸島歩, 「速度変化をとも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)*, Vol.4, 2012, pp.1–21.
- [25] N. M. Fraser, G. N. Gilbert, 「Simulating Speech Systems」, *Computer Speech and Language*, Vol.5, No.1, 1991, pp. 81–99.
- [26] Maulsby, D., Greenberg, S., Mander, R., 「Prototyping an Intelligent Agent through Wizard of Oz」, *Proceedings of International conference on Human Factors in Computing Systems (INTERCHI-93)*, 1993, pp.277–284.
- [27] NHK クイズ番組連想ゲーム (1969 91 年):
<https://www.nhk.or.jp/archives/hakkutsu/features/201409/>
- [28] スリーエーネットワーク編, 「みんなの日本語初級I本冊」.