

対話コミュニケーションにおける2種類の発話タイミング関連

山本 知仁*¹ 平野 作実*¹ 小林 洋平*²
高野 弘二*³ 武藤 ゆみ子*³ 三宅 美博*³

Two types of correlation of utterance timing in dialogue

Tomohito Yamamoto*¹ Tomonori Hirano*¹ Yohei Kobayashi*²
Koji Takano*³ Yumiko Muto*³ and Yoshihiro Miyake*³

Abstract – Dialogue that consisted of an instruction and a response was analyzed to clarify the relation between cognitive process and timing structure of utterance. Results showed that correlation coefficient between duration of instruction utterance and switching pause had negative value when fluctuation of instruction utterance duration was too small to be recognized. However, the correlation coefficient had positive value when fluctuation of instruction utterance was explicitly large. From these results, we discussed that there are two types of utterance mode in human dialogue.

Keywords : dialogue, utterance timing, correlation analysis, communication

1. はじめに

人間はコミュニケーションを通じて、他者との意思疎通を図ることができる。その際、言語的に表現される情報は不可欠であるが、その一方で、言語的には表現することができない非言語的な情報も重要な役割を果たしている^[1]。この非言語的な情報には、ジェスチャーや視線などの視覚的なものや音声の韻律情報などがあるが、その中でも発話タイミングのような対話の時間的な構造が、コミュニケーションが円滑に行われるための重要な要素として注目されてきた。例えば、Condonらは母子コミュニケーションにおいて、音声リズムと身体リズム間の相互作用が重要な役割を果たしていることを示した^[2]。渡辺らは発話とうなずきのリズムに引き込み現象が観察されることを示し、様々なインタフェースに応用している^[3]。また、MatarazzoらやWebb、長岡らは、発話長、発話速度、反応潜時などが、話者間で同調することを報告している^[4-6]。

しかし、これらの研究はコミュニケーションにおける身体的インタラクションのみを重視したものであり、認知的側面からの影響については十分考慮されていないという問題が残されていた。一方、われわれはリズム音に合わせてタップを行う同期タッピング課題や協

調タッピング課題を用い、インターパーソナルなタイミング共有機構に関する研究を進めてきた^[7,8]。特にタイミング機構が、注意資源を必要としない身体的過程と、注意資源を必要とする認知過程として二重化されており、この「二重性」が人間の共創的コミュニケーションの基盤にあることを明らかにしている。これは対話における交互発話課題においても存在が示唆されている^[9]。そこで本研究では、タイミング機構に関するわれわれの先行研究を踏まえて、認知的過程と身体的過程の両側面から対話コミュニケーションにおける時間的構造の解析を進める。

そのためには、対話における被験者の認知状態を実験的に制御できることが不可欠である。そこで具体的には、指示者と被指示者の対話において、指示者の発話速度の変化が意図的に行われる対話と、それが行われない自然な対話の2つの条件を用意した。前者が相対的に認知的過程を強く反映し、後者がより身体的過程を反映するものと予想される。そして、これらの2種類の対話における時間的特徴量の話者間での相関を分析し比較した。以下、第2章で実験手法について説明し、第3章で結果をまとめ、第4章で本研究において初めて発見された2種類の発話タイミング機構について考察を加える。

2. 実験手法

2.1 実験タスクと被験者

本研究では、指示発話と応答発話からなる対話を実験対象として選び、タスクを次のように設定した。被

*1: 金沢工業大学 工学部 情報工学科
*2: 京都産業大学 経済学部
*3: 東京工業大学大学院 総合理工学研究科 知能システム科学専攻
*1: Kanazawa Institute of Technology
*2: Kyoto Sangyo University
*3: Tokyo Institute of Technology

験者は、Fig.1 に示すような 10 個の同じ形の積み木 (木製, 5cm × 5cm × 2.5cm) が置かれた机に、1.2m 程離れて 2 人が向かい合って座る。そして、以下の 2 つの発話からなる対話を行う。

1. 指示者が「積み木を取ってください」という指示を被指示者に出す

2. 被指示者はその指示に対して「はい」という返事をしてから机の上の積み木を一つ取る

以上の実験タスクにおいて、発話速度の変化が発話行動の時間的な構造に与える影響を明らかにするために、以下 2 つの条件を設定した。

条件 1 指示者の発話速度に制限を加えず、指示と応答を自然に繰り返す

条件 2 指示者の発話速度を意図的に変化させ、指示と応答を繰り返す

条件 1 は被験者が自然な発話速度やタイミングで指示と応答を繰り返す条件であり、この実験によって通常の対話の時間的な構造について調べる。条件 2 では指示者に発話の速度を「はやい」、「ふつう」、「おそい」の三段階で指示し、指示者の発話速度を意図的に変化させたときの対話の時間的な構造について調べる。さらに実験後、各条件における指示者の発話速度の変化について被指示者が認知していたかどうかの評価を行った。評価は指示者の「発話速度は変化したか」という質問に対し、「わからない」、「やや変化していた」、「変化していた」、「非常に変化していた」のいずれかを選択する形式で行った。

本実験ではこれら 2 つの条件において、各被験者の組ごとに続けて 3 回の試行を行った。個々の試行は上記の対話を 10 回繰り返すものとする。ただし、条件 1, 2 を行う順番は被験者組ごとにランダムに決めた。実験には、20 代 (平均年齢: 22.8 歳) の健康な男子大学生 4 人が被験者として参加し、全ての 2 人の組み合わせ (指示者と被指示者の入れ替えも含めた) 12 組分の実験を行った。

2.2 実験システム

本研究では、Fig.2 に示すように実験の様子を映像と音声で記録した。記録にはビデオカメラ (SONY 社: DCR-PC300) を使用し、音声はヘッドセット (Audio Technica 社: PROSHEW/P) により収録した。このとき同一試行内での対話行動がリズム化してしまうのを防止するために、指示者に 2~5 秒のランダムに設定された待機時間だけ待つように PC のディスプレイに提示した。また、条件 2 において、指示者に発話速度を指示する際には「はやい」、「ふつう」、「おそい」のいずれかを同 PC を用いて提示し、指示者はこれらの提示を確認した後、意図的に発話速度を変えて発話させるようにした。なお、それぞれの提示回数は試行



図 1 実験の様子

Fig. 1 A picture of experiment

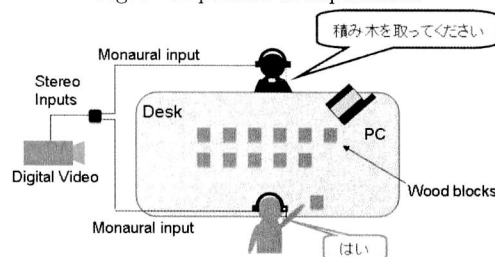


図 2 実験システム

Fig. 2 Experiment system

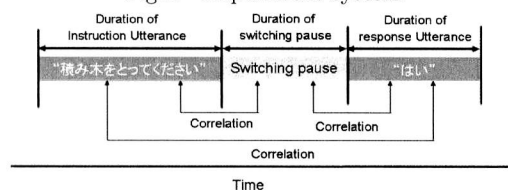


図 3 時間的特徴量

Fig. 3 Indices of dialogue

全体でバランスをとった。

2.3 実験データの解析

本研究では、対話の時間的な構造の解析を行うために録音された音声データの時間的な側面だけに注目した。実験で用いた対話における時間的特徴量は Fig.3 に示すように、指示者の指示発話長 (Duration of instruction utterance), 被指示者の応答発話長 (Duration of response utterance), 交替潜時長 (Duration of switching pause) の 3 つであり、これらの間の相互相関を解析した。それぞれの時間長は、録音された WAV データの有音部を解析することより算出した。

3. 実験結果

3.1 時間的特徴量の変化について

最初に、条件 1, 2 における指示、応答発話及び交替潜時の時間長の結果について示す。Table 1, 2 は、それぞれ条件 1, 2 における各時間長の平均値、及び標準偏差を示したものである。これらは、(積み木 10 個) × (3 試行) × (被験者組み合わせ 12 組) の 360 個のデータから算出している。Table 1 と 2 の結果を比較すると、各時間的特徴量の平均値に関しては大きな差が

表1 条件1における各時間的特徴量の平均と標準偏差

Table 1 Mean and S.D. of indices in condition 1

時間的特徴量 (msec)	指示発話	交替潜時	応答発話
平均値	1120	362	136
標準偏差	71.3	101	27.7

表2 条件2における各時間的特徴量の平均と標準偏差

Table 2 Mean and S.D. of indices in condition 2

時間的特徴量 (msec)	指示発話	交替潜時	応答発話
平均値	1170	383	144
標準偏差	173	108	79.0

ないことがわかる。それに対し、条件2において指示発話長の標準偏差が条件1に比べ約2.5倍大きくなっている (F-test; $p < 0.001$) ことから、指示者が発話速度を変化させていたことがわかる。また、交替潜時長は条件間で標準偏差が大きく変わらないが (F-test: $p > 0.1$)、応答発話長は、条件1に比べ条件2の方が3倍弱大きくなっている (F-test: $p < 0.001$)。これは、被指示者の応答発話速度も指示発話速度の影響を受け変化していることを示唆している。

3.2 時間的特長量間の相関関係

次に、各時間的特徴量間の相互相関について示す。Table 3, 4は、指示発話長と交替潜時長の相関係数の全データであり、列が指示者、行が被指示者を表している。相関係数の値は同じ被験者の組み合わせ内の3試行分のデータを合わせて算出した。Fig.4はTable 3, 4に示した指示発話長と交替潜時長の相互相関係数のヒストグラムである。実線が条件1を、破線が条件2を示している。

図より、条件1の場合には指示発話長と交替潜時の間の相関係数が負の方に分布が偏っているのがわかる。値の中には、Fig.5に示すような、相関係数が-0.572のような例も見られた。それに対し、条件2では値の分布が正の方にかなり偏っている。Fig.6は条件2における相関関係 (相関係数: 0.771) の一例である。また、条件1と条件2の分布の平均値について検定を行った結果、統計的に有意な差が見られた (Mann-Whitney U-test; $p < 0.001$)。

また、今回の実験においては指示発話長と応答発話長の間、交替潜時と応答発話長の間、相関関係に関しては、条件1と条件2の平均値の間に統計的に有意な差は見られなかった (Mann-Whitney U-test: $p > 0.25$)。

3.3 発話速度変化に対する認知状態

最後に、指示者の発話速度の変化に対する被指示者の認知の状態について示す。Fig.7は、条件1, 2それ

表3 条件1における指示発話長と交替潜時の相関係数

Table 3 Correlation coefficient between instruction utterance and switching pause in condition 1

相関係数	被験者 A	被験者 B	被験者 C	被験者 D
被験者 A		0.119	-0.182	-0.444
被験者 B	-0.0810		-0.572	-0.159
被験者 C	-0.181	-0.0905		-0.167
被験者 D	-0.167	-0.00981	0.180	

表4 条件2における指示発話長と交替潜時の相関係数

Table 4 Correlation coefficient between instruction utterance and switching pause in condition 2

相関係数	被験者 A	被験者 B	被験者 C	被験者 D
被験者 A		0.526	0.320	0.502
被験者 B	0.343		0.287	-0.164
被験者 C	0.699	0.771		0.0840
被験者 D	0.346	0.544	0.345	

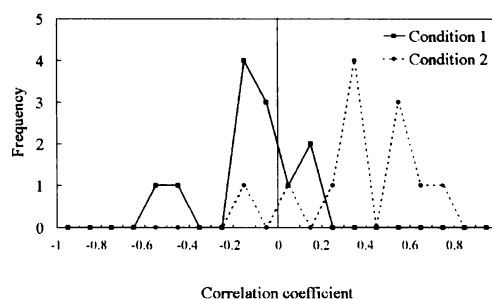


図4 指示発話長と交替潜時長の相関係数の分布

Fig. 4 Distribution of correlation coefficient between instruction utterance and switching pause

ぞれにおいて、「発話速度は変化したか」という質問に対し、「わからない」と答えた被験者と「やや変化した」、「変化した」、「非常に変化した」と答えた被験者の比率である。条件1では12人中11人が、「わからない」と答えたのに対し、条件2において「わからない」と答えたのは12人中、3人であった。これらの結果は、指示発話速度の変化が小さい条件1においては被指示者は発話速度の変化を認知しにくい、指示発話長の変化が大きい条件2では被指示者は指示者の発話速度の変化を認知しやすいことを意味している。

4. 考察

本研究では、対話における認知的過程が、対話の時間的な構造に与える影響について調べた。その結果として、指示者の発話速度の変化が小さく被指示者がその変化を認知しにくい条件では、指示発話長と交替潜時の相関係数が負に偏った分布をとるのに対し、指示者の発話速度の変化が大きく、その変化を被指示者が

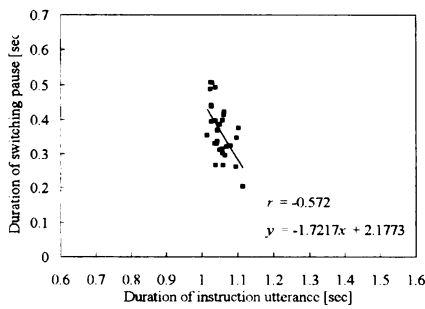


図5 条件1における指示発話長と交替潜時長の相関関係

Fig.5 A correlation between instruction utterance and switching pause in condition 1

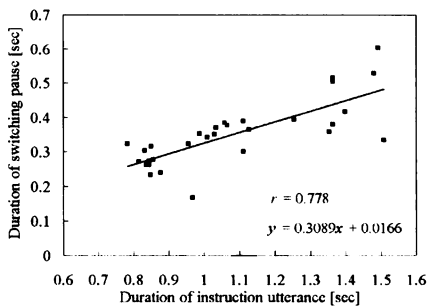


図6 条件2における指示発話長と交替潜時長の相関関係

Fig.6 A correlation between instruction utterance and switching pause in condition 2

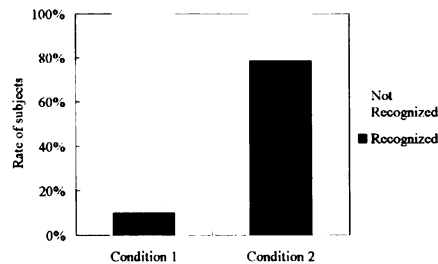


図7 発話速度変化に対する被指示者の認知状態

Fig.7 Recognition to change of instruction utterance speed

認知しやすい条件では、指示発話長と交替潜時の相関係数の分布が大きく正の方に偏ることが示された。

指示発話速度が明示的に変化する条件2において、指示発話長と交替潜時の間に正の相関が表れるような形で被指示者の発話タイミングが変化した結果は、Matarazzoらや長岡ら^[4-6]が示している同調傾向のひとつと見做すことができる。同調傾向とは、コミュニケーションにおいて相手と発話行動や身体動作などが同調してくる傾向であり、円滑なコミュニケーションの指標として捉える研究もある。

それに対し、指示者の発話速度に制限を加えず、また、被指示者も発話速度の変化を認知しにくい条件1においては、指示発話長と交替潜時長の相関係数の分布は負の方に偏ることが明らかになった。この結果は、同調傾向とは逆の傾向であり、先行研究においても注目されていない結果である。このような発話行動にどのような機能的な意味があるのかは明らかではないが、対話コミュニケーションにおいて発話速度変化に対する認知の差がこのような2つの発話行動のモードを生み出しているという結果は大変興味深いといえる。

われわれは既に、同期タッピング課題を用いてインターパーソナルなタイミング機構が二重化されていることを報告してきた^[8,9]。そこでは、注意資源を必要としない身体的過程と注意資源を必要とする認知的過程の両方が、共創的コミュニケーションに不可欠であることを明らかにした。本研究で得られた上記の結果が、この「二重性」と、どのような関係にあるのかは今後の研究を待たなければならないが、発話速度の変化の認知されやすさに依存して異なる相関関係が得られたことは、対話と二重性との関係を強く示唆するものである。このような作業仮説のもとで、今後さらなる研究を進める予定である。

参考文献

- [1] 大坊郁夫: しぐさのコミュニケーション—人は親しみをどう伝えあうか—; セレクション社会心理学 14, サイエンス社 (1998).
- [2] Condon, W.S., Sander L.W.: Neonate Movement is Synchronized with Adult Speech, Science; Vol.83, pp.99-101, (1974).
- [3] 渡辺富夫, 大久保雅史, 中茂睦裕, 檀原龍正: InterActorを用いた発話音声に基づく身体的インタラクションシステム; ヒューマンインタフェース学会論文誌, Vol.2, No.2, pp.21-29, (2000).
- [4] Matarazzo, J.D., Weitman, M., Saslow, G., Wiens, A.N.: Interviewer influence on durations of interviewee speech; Journal of Verbal Learning and Verbal Behavior, vol.1, pp.451-458, (1963).
- [5] Webb, J.T.: Interview synchrony: An investigation of two speech rate measures in an automated standardized interview; In B. Pope and A.W. Siegman (Eds.), Studies in dyadic communication New York: Pergamon, pp.115-133 (1972).
- [6] 長岡千賀: 対人コミュニケーションにおける非言語行動の2者間相互影響; 対人社会心理学研究, Vol.6, pp.101-112, (2006).
- [7] 三宅美博, 大西洋平, エルンスト・ペッペル: 同期タッピングにおける2種類のタイミング予測機構; 計測自動制御学会論文集, Vol.38, No.12, pp.1114-1122, (2002).
- [8] 今 啓, 三宅美博: 協調タッピングにおける相互同調過程の解析とモデル化; ヒューマンインタフェース学会論文誌, Vol.7, No.4, pp.61-70, (2005).
- [9] 三宅美博, 辰巳勇臣, 杉原史郎: 交互発話における発話長と発話間隔の時間的階層性; 計測自動制御学会論文集, Vol.40, No.6, pp.670-678, (2004)