

## 第二言語での会話構造が視線動作に及ぼす影響についての分析

## Eye Gaze Analyses of Different Interaction Structure between L1 and L2 Conversation

伊集院 幸輝† 石田 充† 堀内 保大† 馬田 一郎† 山本 誠一†  
 Koki Ijuin Mitsuru Ishida Yasuhiro Horiuchi Ichiro Umata Seiichi Yamamoto

## 1. はじめに

話し言葉による円滑なコミュニケーションの成立には、話す内容などの言語情報のみならず、韻律情報や視覚情報などの非言語情報を用いて、会話の参加者が会話の成立のために協調的な共同行為を行うことが不可欠である[1]. Kendon[2]の研究では、会話におけるお互いを見合う相互注視は、共同行為の成立のための重要な役割を果たしていることが示唆されている。また、Jokinen[3]らの研究では、三人会話での発話権の授受における視線動作の重要性が示された。これらの研究を含む先行研究から、コミュニケーションにおける視線動作等の非言語情報の重要性が示唆されてきた[4][5].

近年、輸送技術や情報通信技術の発展によって、第二言語を使う機会が格段に増加している。第二言語におけるコミュニケーションでは、会話参加者の言語運用能力の低さなどにより、非言語情報の役割が母語での会話と異なる可能性がある。Veinottら[6]の研究では、母語話者のペアと第二言語話者のペアにそれぞれ被験者2人で協力し指定されたタスクをこなしてもらい、タスクが終了するまでの時間を計測する実験を行った。その際、被験者間のコミュニケーションを音声のみで行うケースと、音声と相手の姿が見える映像の両方を用いたケースで行った。その結果、母語話者のペアでは2つのケースの実験ではタスク終了までの時間に差はなかったが、第二言語話者のペアでは音声と映像の両方を用いたケースのほうが、有意に早くタスクが終了したことが分かった。このことから、第二言語での非言語情報の役割は、母語よりも重要である可能性が高い。しかし、Veinottらの研究は、遠隔地でのコミュニケーションに着目しており、対面式の会話における視線動作の重要性について十分な解析がされていない。

そこで、筆者らの研究グループでは、非言語情報の一つである視線動作に着目し、筆者らが構築した対面式の3人会話のマルチモーダルコーパスを用いて、円滑なコミュニケーションが成立するためには視線動作がどのような役割を果たしているかについて定量的な解析を行ってきた。その結果、聞き手は発話者を注視する傾向がある事が分かった[7-10]。本稿では、これらの結果を踏まえて、より詳細な視線動作の分析を行うために、3人での会話の中で話者が交替するタイミングに着目し、会話における視線動作の役割の解明を目的とする。

## 2. マルチモーダルコーパス

## 2.1 被験者

本研究には、筆者らが開発したマルチモーダルコーパスを用いた。本マルチモーダルコーパスは、母語を日本語、

第二言語を英語とした日本人大学生を計60人、被験者として用い、各会話は3人で行われるため、計20グループ分の会話データからなる。また、被験者の英語能力を測るためにTOEICスコアを用いて、グループごとにスコアの昇順にRank1, Rank2, Rank3と順位付けを行った。被験者のTOEICスコアの分布は450点から985点で、グループは無作為に選出して構成されている。

## 2.2 収録条件

被験者は円卓に等間隔に座ってもらう。被験者間の距離は約1.5mとなっている。それぞれの被験者の正面にビデオカメラを設置し、各被験者に帽子型の視線追跡装置と接話型マイクを装着してもらうことで、各被験者の正面動画と視線動作、発話音声を別々に収録した。視線追跡装置には、NAC社製のEMR-9を用いた。図1は実際の実験風景である。



図1 収録実験の様子

## 2.3 会話トピック

各会話は、約6分ほどで、使用言語は日本語と英語、会話テーマは自由会話と目的会話の2種類を各言語で行った。そのため、各グループ計4会話行っている。自由会話では、特定のテーマに沿って被験者が所定時間の間に自由に話してもらい、目的会話では、特定のテーマに沿って時間内に被験者3人が協力してある一つの結論を出してもらい。

## 2.4 アノテーション

収録した会話データに注釈付けを行った。注釈内容は被験者の発話区間(DialogAct)、視線動作として左の被験者を見ている区間(GazeAtLeftPerson)、右の被験者を見ている区

† 同志社大学

‡ 情報通信研究機構

間(GazeAtRightPerson), 相づち区間(Backchannel)の動作者, 区間開始時間, 区間終了時間, 区間継続時間である. 今回の注釈付けでは, 200ms未達の視線動作はサックードなどのノイズとみなし, 無視している.

### 3. 発話中の視線動作の解析

#### 3.1 分析手法

発話中における発話者の視線動作と聞き手の視線動作に着目して分析を行った. 下記の数式(1)は, 発話中の被験者がほかの聞き手を見ている時間の割合, Speaker's Gazing Ratio(SGR)を, 式(2)は聞き手が発話中の発話者を見ている時間の割合, Listener's Gazing Ratio(LGR)を示しており, (3), (4)はそれぞれ被験者ごとの平均値の算出方法を示している.

$$\text{Speaker's Gazing Ratio} = \frac{DSGL(i)}{D(i)} \quad (1)$$

$$\text{Listener's Gazing Ratio} = \frac{DLGS(i)}{D(i)} \quad (2)$$

$$\text{Average of Speaker's Gazing Ratio} = \frac{\sum_{i=1}^n DSGL(i)}{\sum_{i=1}^n D(i)} \quad (3)$$

$$\text{Average of Listener's Gazing Ratio} = \frac{\sum_{i=1}^n DLGS(i)}{\sum_{i=1}^n D(i)} \quad (4)$$

式中の  $D(i)$  は, 特定の発話者の  $i$  番目の発話長を示している.  $DSGL(i)$  は, 特定の発話者が  $i$  番目の発話をしている間に, その発話者が特定の聞き手を見ている時間を示しており,  $DLGS(i)$  は, 特定の発話者が  $i$  番目の発話をしている間に, 特定の聞き手がその発話者を見ている時間を示している. それぞれ発話者と聞き手を固定して, 各グループそれぞれ6通りのSGR, LGRが算出される.

#### 3.2 分析結果

図2, 3は言語別に各被験者のSGRとLGRの平均を算出したものである. SGRについて, 言語と会話テーマを被験者内要因, 話者と聞き手の英語能力の比較を被験者間要因とした分散分析を行ったところ, 有意な結果はなかった. また, 同様にLGRについて, 言語と会話テーマを被験者内要因, 話者と聞き手の英語能力の比較を被験者間要因とした分散分析を行った結果, 言語の主効果で有意となった( $F(1, 118) = 81.6, p < .01$ ).

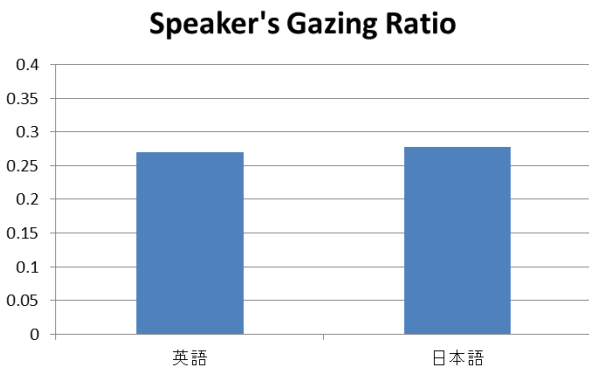


図2 Speaker's Gazing Ratioの平均値

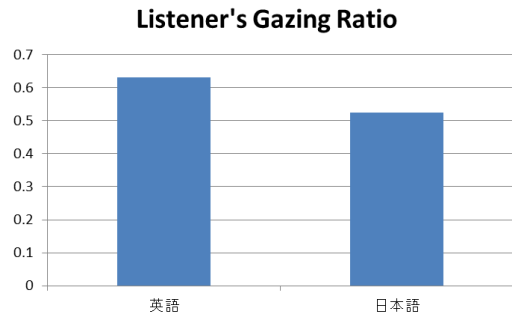


図3 Listener's Gazing Ratioの平均値

#### 3.3 考察

ANOVAの結果より, 発話者は言語の違いによって聞き手を注視している割合に差がないことが分かり, 逆に聞き手は第二言語での会話の場合, 母語のそれと比べてより発話者を注視していることが分かった. このことから, 発話者は言語に関わらず聞き手のある一定量は注視しており, 聞き手は第二言語において何らかの要因から発話者を注視している可能性がある. 聞き手が発話者を注視している要因として考えられる仮説は, 1)言語能力の低さから視覚情報を用いて発話内容を読み取ろうとしている, 2)会話における話者交替が単純である, 等が挙げられる. これらの要因を確かめるために, 聞き手の視線動作に着目し, 詳細な分析を行う.

### 4. 聞き手の視線動作

#### 4.1 Listener's Gazing Ratioの割合分布

図4は全発話におけるLGRの割合分布を言語別に示したものである. 図中のパーセントは, 左から順に, 全体からLGRが0以上0.05未満である発話が占める割合, LGRが1である発話が占める割合を表記したものである. この表から, 第二言語会話では聞き手が発話者を発話の最初から最後まで見ていることが多いことがわかる. 聞き手が発話の最初から最後まで発話者を注視するためには, 聞き手は発話が始まる前から次の発話者を注視する必要がある.

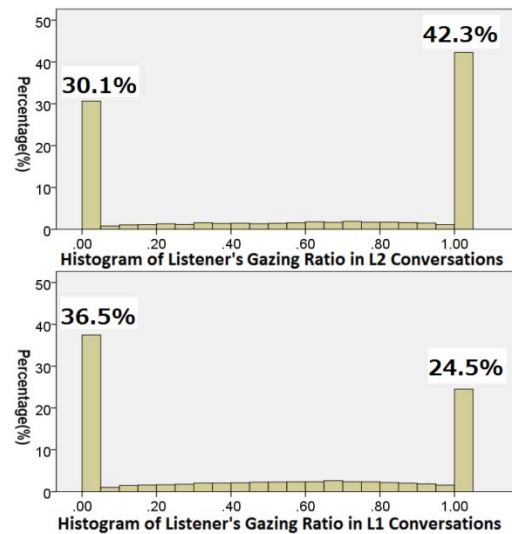


図4 言語別 Listener's Gazing Ratioの割合分布 (L2は第二言語, L1は母語)

## 4.2 聞き手の発話前での視線動作の分析

そこで、発話前の聞き手の視線動作の分析を行った。本研究では先行研究[4]から得られた最小の注視時間を200msとし、発話前の200msの区間における視線動作に着目した。本分析には2つの手法を用いた。一つは、発話前の200msの間に、聞き手が次の発話者を注視している割合を求めた Gazing at Next Speaker Ratio で、もう一つは、発話前の200msの間に聞き手がもう一人の聞き手を注視している割合を求めた Gazing at Other Listener Ratio である。それぞれの算出方法を式(5)、(6)で示した。

$$\text{Gazing at Next Speaker Ratio} = \frac{DLGNS(i)}{200[\text{ms}]} - (5)$$

$$\text{Gazing at Other Listener Ratio} = \frac{DLGOL(i)}{200[\text{ms}]} - (6)$$

式中の DLGNS(i) は発話前の200msに聞き手が発話者を見ている時間で、DLGOL(i) は発話前の200msに聞き手がもう一人の聞き手を見ている時間を表している。それらに対して、被験者内要因を言語、会話テーマ、注視対象の差(発話者か他の聞き手か)とし、被験者間要因を聞き手の英語能力としてANOVAを行った。その結果、言語と注視対象の差の交互作用について有意( $F_{(1,117)} = 115.0, p < .01$ )となった。下位検定を行ったところ、英語について注視対象で有意( $F_{(1,117)} = 226.3, p < .01$ )、日本語について注視対象で有意( $F_{(1,117)} = 38.3, p < .01$ )となった。図5は Gazing at Next Listener Ratio と Gazing at Other Listener Ratio の平均値とANOVAの結果を図示したものである。

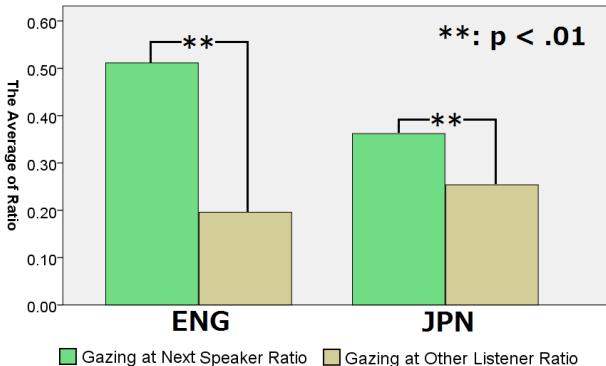


図5 発話前における聞き手の視線動作の割合

## 4.3 考察

以上の結果から、第二言語でのコミュニケーションでは発話の最初から最後まで注視されていることが多いことが分かった。これは、被験者の第二言語能力の低さを補うために、発話者を絶えず注視し続ける可能性があげられる。また、どちらの言語でも発話前から発話者が注視されがちであり、特に第二言語会話ではその傾向が顕著であること

が分かった。このことから、第二言語会話では何らかの理由で次の発話者を予測しやすい環境である可能性が考えられる。よって、次に発話者が交替するタイミングに着目し、会話参加者がどのような視線動作を行っているかを分析する。

## 5. 役割別視線動作

### 5.1 分析手法

本項では、三人会話の中で、発話者が切り替わるタイミングに着目し、被験者を「現発話者」、「次発話者」、「聞き手」と分類して、現発話者の発話中の視線動作と注視対象について分析を行い、式(7)を用いた。

$$\text{Role Gazing Ratio} = \frac{DR_kGR_l(i)}{D_{cs}(i)} - (7)$$

式(7)中における、 $D_{cs}(i)$  は、現発話者の  $i$  番目の発話区間長を示しており、 $DR_kGR_l(i)$  は現発話者の  $i$  番目の発話中に役割  $R_k$  (現発話者、次発話者、聞き手のうちのどれか) の被験者が役割  $R_l$  ( $R_k$  以外の役割) の被験者を注視している区間長を表している。これによって、現発話者が発話中に、それぞれの役割の被験者がどの役割の被験者を注視しているかがわかる。

### 5.2 分析結果

図6は、Role Gazing Ratio の平均と標準偏差を表したものである。図中の C は現発話者、N は次発話者、L は聞き手を表しており、CtoN の場合現発話者が次発話者を注視している、ということを示している。また、表1はそれらの平均の値を用い、注視者の役割の平均ごとに言語、会話テーマ、注視対象を被験者内要因としてANOVAを行った結果である。また、注視者が現発話者のとき、言語と注視対象の交互作用について有意となったため、下位検定を行った。その結果、第二言語会話での注視対象について有意( $F_{(1,19)} = 125.3, p < .01$ )、母語会話での注視対象について有意( $F_{(1,19)} = 39.3, p < .01$ )、次発話者を注視している場合での言語について有意( $F_{(1,19)} = 14.5, p < .01$ )、聞き手を注視している場合での言語について有意( $F_{(1,19)} = 17.7, p < .01$ )となった。

表1 Role Gazing Ratio における、言語、注視対象、会話テーマを被験者内要因としたANOVA結果

	ANOVAの結果			
	言語	注視対象	言語×注視対象	言語×注視対象×会話テーマ
現発話者	有意なし	$F_{(1,19)} = 151.0, p < .01$	$F_{(1,19)} = 22.4, p < .01$	$F_{(1,19)} = 4.0, p < .1$
次発話者	$F_{(1,19)} = 10.6, p < .01$	$F_{(1,19)} = 492.3, p < .01$	$F_{(1,19)} = 30.0, p < .01$	有意なし
聞き手	$F_{(1,19)} = 12.7, p < .01$	$F_{(1,19)} = 297.7, p < .01$	$F_{(1,19)} = 18.8, p < .01$	$F_{(1,19)} = 4.8, p < .05$

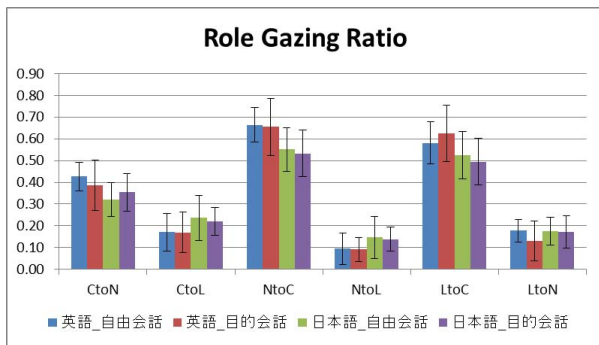


図6 Role Gazing Ratioの平均値と標準偏差

### 5.3 考察

ANOVAの結果から、現発話者は、英語、日本語のどちらの会話についても、次発話者をよく注視していることが分かり、また英語会話の方が日本語会話と比べてF値が大きいため、その傾向が顕著であることが分かった。Speaker's Gazing Ratioではそのような言語間の差異は見られなかった。また、第二言語会話では、会話参加者の言語能力の低さから、より視線動作の役割が重要となり、その傾向がより顕著になっている可能性がある。また、第二言語での会話では、言語運用能力の低さから、話者交替が単調になっており、現発話者が会話における主導権を握っている可能性も考えられる。

### 5 まとめ

本研究は、第二言語での3人会話における視線動作の解析を行った。その結果、条件を設定しない場合、発話者が聞き手を注視している割合は母語での会話と第二言語での会話で差異は見られなかった。一方、聞き手は母語での会話と比べて、第二言語での会話でよく発話者を注視していることが分かった。聞き手の視線動作を詳細に分析すると、聞き手は第二言語会話において、よく発話の最初から最後まで注視していることが示された。

また、発話前の視線動作を分析することで、第二言語会話では聞き手は発話者が話し始める前から発話者を注視していることが多いということが分かった。これらのことから、第二言語会話では聞き手は何らかの要因で次の発話者を予測できているのではないかと考えた。

次に会話参加者を複数の役割に分類して視線動作の解析を行った。その結果、第二言語において、現発話者は次の発話者を注視しがちである事が分かった。これは、現発話者の第二言語運用能力が母語のそれと比べて低いため、次の発話者にターンを円滑に渡すために視線動作を用いている可能性が挙げられる。また、母語会話と比べて、第二言語会話では三人の言語能力の低さから、三人による複雑な会話ではなく、質問者と回答者といった単純な発話ターンの受け渡しが多くなされている可能性も考えられる。これらの可能性は、第二言語会話における現発話者が会話における主導権を握っているという可能性に準拠する。

これらの結果から、第二言語での会話は、母語での会話とは会話の構造が異なっている可能性が高く、その構造の差が円滑なコミュニケーションの成立を阻害している可能性が高いことが明らかになった。

### 6 今後の予定

本研究は、3人会話を視線動作に着目して定量的な分析を行った。その結果、第二言語会話では会話の構造が異なっており、それが原因で円滑なコミュニケーションが成立していない可能性が高いことが分かった。これらの分析は、発話内容を考慮せずに、統計的解析をもとに行ったものである。今後は、発話内容を考慮した会話構造の比較分析を行うことで、言語間の差異や、会話参加者の第二言語運用能力の差異が、視線動作にどのような影響を及ぼすか解明する。また、現在英語母語話者と第二言語話者による会話を収録しており、それらのデータを用いてより現実に即した会話状況での分析を行っていく。

### 参考文献

- [1] S. Garrod et al. (2004). "Why is conversation so easy?", *TRENDS in Cognitive Sciences*, 8, pp.8-11.
- [2] A. Kendon. (1967). "Some functions of gaze-direction in social interaction", *Acta Psychologica*, 26, pp.22-63.
- [3] K. Jokinen et al. (2013). Gaze and Turn-Taking Behavior in Casual Conversational Interactions, *ACM Transactions on Interactive Intelligent Systems: vol 3, No.2*, pp. 12:1-30.
- [4] H. H. Clark and S. A. Brennan. (1991). *Grounding in Communication, Perspectives on Socially Shared Cognition*, American Psychological Association, pp. 222-233.
- [5] M. Argyle, J. Dean. (1965). Eye-contact, distance and affiliation. *Sociometry* 28 pp. 289-304
- [6] E. Veinott, et al. (1999). Video helps Remote Work: Speakers Who Need to Negotiate Common Ground Benefit from Seeing Each Other, *Proceedings of the Conference on Computer Human Interaction. CHI'99*, ACM Press, PA, USA, pp. 302-309.
- [7] K. Jokinen et al. (2013). Gaze and Turn-Taking Behavior in Casual Conversational Interactions, *ACM Transactions on Interactive Intelligent Systems: vol 3, No.2*, pp. 12:1-30.
- [8] I. Umata. (2013). Effects of language proficiency on eye gaze in second language conversations: toward supporting second language collaboration. In: *ICMI 2013*. (2013) 413 - 420
- [9] K. Ijuin et al. (2014). Eye gaze analyses in l1 and l2 conversations: From the perspective of listeners' eye gaze activity. In: *Understanding and Modeling Multiparty, Multimodal Interactions - Workshop at ICMI2014*.
- [10] Seiichi Yamamoto et al. (2015). Multimodal corpus of multiparty conversations in L1 and L2 languages and findings obtained from it, *Language Resources and Evaluation*, (DOI: 10.1007/s10579-015-9299-2)

以上