

話者の注視が次発話者の選択に及ぼす影響に関する母語と第二言語会話での比較分析 Comparison of Effect of Speaker's Gaze on Selection of Next Speaker between Native- and Second-Language Conversations

山下 昂人[†] 伊集院 幸輝[†] 加藤 恒夫[†] 山本 誠一[†]
Takato Yamashita Koki Ijūin Tsuneo Kato Seiichi Yamamoto

1. はじめに

会話における円滑なコミュニケーションには、言語情報に加えて韻律情報や身体動作などの非言語情報も重要な機能を果たしていることが指摘されている[1][2][3]. 非言語情報の中でも視線動作については、親密さの表示[1][2], 発話権の割り当て[3]等の種々の機能に関する研究がなされている. 特に相互注視は、2人での対話での話者交替等の共同行為が成立する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[4]. 複数人の会話参加者からなる会話についても、Vertegaal らの研究では、会話中における会話者への注意の方向を表出するものとして、注視が有用であることが示されている[5]. また、Jokinen らの研究では、3人会話での話者交替時における注視行動の重要性が示されている[6].

近年、輸送技術や情報通信技術の発展に伴い、母語だけでなく第二言語を用いてコミュニケーションを行う機会が増加している. 第二言語でのコミュニケーションでは、第二言語運用能力の低さなどから、非言語情報の役割が母語会話での役割と異なる可能性があることが示唆されている. Veinott らの研究[7]では、母語話者のペアと第二言語話者のペアを被験者とし、協力して指定されたタスクを行う実験を、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い、それぞれのペアにおいて2つのケースでのタスク終了時間を比較した. その結果、母語話者のペアでは2つのケースでタスク終了までの時間に差がなかったが、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. このことは、第二言語における非言語情報は母語会話に比較して、より重要な機能を果たしている可能性が高いことを示している.

筆者らの研究グループでは、母語会話と第二言語会話での非言語情報の機能の違いについて詳細な分析を進めるために、同一被験者3名で構成されるグループでの母語と第二言語会話のマルチモーダルコーパスを開発し[8], 非言語情報の一つである注視行動に着目し、母語会話と第二言語会話における注視行動の違いについて定量的な解析を行ってきた[9][10].

その結果、

- (1) 第二言語会話時の発話者への聞き手の注視時間は、母語会話時よりも有意に長い[8]
- (2) 発話者の二人の聞き手への注視時間を比較すると、母語会話と第二言語会話共に、次発話者への注視時間が他の聞き手への注視時間と比較して有意に高い[10] (二人の聞き手への総注視時間は母語会話と第二言語会話で有意な差はない)
- (3) 現発話者の次発話者への注視時間と他の聞き手への注視時間の差を比較すると、第二言語会話時の注視時間の差は、母語会話時の差よりも有意に長い[10]

ことが明らかとなった. 更に、3人会話中の話者交替のタイミングに着目し、話者交替時の発話と話者交替が起きないそれ以外の発話での注視時間の割合の比較分析を行ったところ、

- (4) 現発話者の次発話者への注視時間と他の聞き手への注視時間の差は、母語会話及び第二言語会話共に、話者交替時の発話ではそれ以外の発話と比較して有意に大きい
- (5) 第二言語会話での、話者交替時とそれ以外の発話の注視時間の差は、母語会話時での差よりも有意に長いことが明らかになった[11].

以上の分析結果から、第二言語会話では母語会話に比べて、現発話者の注視行動は次発話者の発話権の割り当てに一層強い影響を及ぼしていると考えられる. しかし、これら先行研究では、発話中の注視行動に着目した分析をしており、現発話終了から次発話までのポーズ区間における注視行動は考慮されていない. 一方、ポーズ区間中の注視行動の影響については、その注視時間の長さが発話権の割り当てに関連する[12]と示されているため、ポーズ区間での注視時間の長さにより、発話区間とポーズ区間で話者の注視が次発話者の選択に及ぼす影響が異なる可能性がある.

そこで本稿では、話者の注視行動が次発話者の選択に関してどのように影響を及ぼすのかについての分析を目的とし、現発話終了から次発話までのポーズ長に対応して、現発話者から次発話者への注視割合の変化に関して、母語会話と第二言語会話で比較分析を行う. また、現発話終了から次発話までのポーズ長に対応して、実際に現発話者に注視された聞き手が発話する割合の変化に関して、母語会話と第二言語会話で比較分析を行う.

以下、第2節では本分析に使用した三人会話マルチモーダルコーパスを紹介し、第3節では話者交替時の現発話者の注視割合の時間的变化、第4節では注視された聞き手が次発話者となる割合の時間的变化、についての比較結果を述べる.

2. 三人会話マルチモーダルコーパス

2.1 収録条件

今回の分析に用いたマルチモーダルコーパスには、図1に示すように、円卓に等間隔に座った被験者3名を1グループとした、日本語と英語のそれぞれ約6分の会話収録されている. 会話の種類は自由会話と目的会話の2種類である. 自由会話は被験者間の所定時間内の自由な会話となっており、目的会話は特定のテーマに沿って時間内に被験者3名が協力して一つの結論を出す会話となっている. そのため、各グループ、日本語と英語での自由会話と目的会話の計4会話を実施している.

被験者間の距離は約1.5mであり、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

[†] 同志社大学 Doshisha University

跡装置と接話型マイクを装着して、各被験者の正面動画と注視行動、発話音声は別々に収録されている。視線追跡装置には、NAC 社製の EMR-9 が用いられている。被験者は母語を日本語、第二言語を英語とする日本人大学生を計 60 人で、計 20 グループ分の会話データが収録されている。



図 1 収録実験の様子

2.2 アノテーション

収録した会話データに対して注釈付けと発話内容の書き起こしが行われている。注釈内容は被験者の発話区間(DialogAct)、視線動作として右の被験者を見ている区間(GazeToRightPerson)、左の被験者を見ている区間(GazeToLeftPerson)、相槌区間(BackChannel)についての動作者、開始時刻、終了時刻、継続時間である。なお、200ms 未満の注視行動はサッケードなどのノイズとみなし、無視している。

3. 話者交替時の発話者の注視割合の分析

3.1 発話区間とポーズ区間での注視

母語会話、第二言語会話共に、話者交替前の発話中には、現発話者は次発話者を他の聞き手に比べて、より注視することが明らかにされている[11]。表 1 に現発話者が発話区間中に次発話者と他の聞き手を各々注視している割合について、話者交替前の発話とそれ以外の発話の比較を示す。ここで、発話区間での注視割合は以下のように定義される。

$$\text{Gazing Ratio during Utterance} = \frac{\text{現発話者が発話中に一方の聞き手を見ている時間}}{\text{発話者の発話継続時間}} \quad (1)$$

表 1 話者交替前の発話とそれ以外の発話における、現発話者の次発話者と他の聞き手、各々への注視割合の比較

	現発話者の次発話者への注視割合	現発話者の他の聞き手への注視割合
話者交替前の発話 (母語)	38.4%	19.3%
それ以外の発話 (母語)	26.8%	20.0%
話者交替前の発話 (第二言語)	46.3%	17.3%
それ以外の発話 (第二言語)	28.8%	16.5%

次に、話者交替時の発話に続くポーズ区間(現発話者の発話終了から次発話者の発話開始までの時間)での注視と発話区間での注視傾向の比較を行うために、発話区間とそれに続くポーズ区間(以下、(発話+ポーズ)区間と略称する)での注視割合の比較を行った。ここで、(発話+ポーズ)区間での現発話者の注視割合は以下のように定義される。

$$\text{Gazing Ratio during Utterance and Pause} = \frac{\text{現発話者が発話とポーズ区間に一方の聞き手を見ている時間}}{\text{現発話開始から次発話までの時間}} \quad (2)$$

表 2 に、発話区間と(発話+ポーズ)区間で、現発話者の一方の聞き手への注視割合を比較した場合に、二人の聞き手の間で注視割合の大小関係が逆転した発話の発生割合を示す。

表 2 発話区間と(発話+ポーズ)区間で、現発話者の一方の聞き手への注視割合の大小関係が逆転した発話の発生割合

	母語	第二言語
0~2.0 秒のポーズが生じた発話数	3165	2342
注視割合の大小関係が逆転した発話数	138	57
注視割合の大小関係が逆転した発話の発生割合	4.36%	2.43%

表 2 より、両言語共に、発話区間と(発話+ポーズ)区間の間で、現発話者から聞き手 1, 2 への注視割合の大小関係が逆転した発話の発生割合は低かった。このことは、現発話者は発話中に注視した聞き手をポーズ区間においても、注視する傾向にあるといえる。分析結果と先行研究[11]から、話者交替の際、現発話者は発話中に注視した次発話者をポーズ区間においても注視する傾向にあると考えられる。そこで、ポーズ長に対応して、現発話者の次発話者への注視割合の変化に関して、母語会話と第二言語会話で比較分析を行う。

3.2 現発話者の次発話者への注視割合の変化

母語会話と第二言語会話における、現発話者の次発話者への注視割合の発話区間に続くポーズ区間の長さによる違いを図 2 に示す。図中の縦棒グラフは、グラフは左から各々、発話中の注視割合、発話区間に続くポーズ区間長が 0~0.5 秒の場合の注視割合、発話区間に続くポーズ区間長が 0.5~1.0 秒の場合の注視割合、発話区間に続くポーズ区間長が 1.0~1.5 秒の場合の注視割合、発話区間に続くポーズ区間長が 1.5~2.0 秒の場合の注視割合を示している。

現発話者の次発話者への注視割合については、分析結果から以下のことが明らかになった。

1. (発話+ポーズ)区間の注視割合は発話区間での注視割合と同様に、第二言語会話の方が母語会話よりも高い。
2. 発話区間の注視割合に比べて 0~0.5 秒の短いポーズ区間を含めた(発話+ポーズ)区間での注視割合は両言語ともに高い。

3. 母語会話では、発話区間の注視割合に比べて1.5~2.0秒の長いポーズ区間を含めた(発話+ポーズ)区間での注視割合は低下している。
4. 第二言語会話では、(発話+ポーズ)区間での注視割合は低下せず、発話区間での注視割合と同等以上の値を示している。

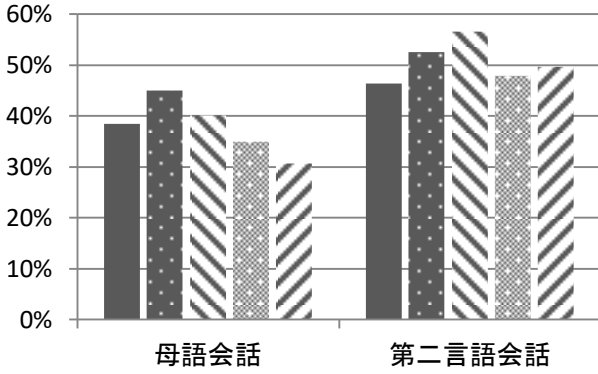


図2 母語、第二言語会話での、現発話者から次発話者への発話中の注視割合と(発話+ポーズ)中の注視割合の比較

以上の事から、母語会話に比べて第二言語会話では、ポーズ区間においても現発話者は次発話者をより注視する傾向にあり、高い注視割合が比較的長く続く傾向を示している。このことから、母語会話に比べて第二言語会話では、現発話者の注視による次発話者の選択の効果が長く持続している可能性が考えられる。

4. 注視された聞き手が次発話者となる割合の変化

発話中の話者の注視が次発話者の選択にどれだけ長く影響を及ぼすのかを分析するために、現発話者から注視された聞き手が次発話者となる割合の時間的変化について、0.2秒で区切られた区間毎の分析を行った。話者交替時の発話に続くポーズ長の累積分布を図3に示す。図3に示されるように、母語会話でのポーズ長は第二言語会話に比べて短く、次発話者は比較的早く応答している。

この0.2秒毎に区切られたポーズ区間に対応して、他の聞き手と比べて現発話者に注視されていた聞き手が次発話者になる割合の時間的変化を図4に示す。図中の点線は各ポーズ区間での割合についての線形回帰直線である。分析結果から以下の事が明らかになった。

1. 母語会話よりも第二言語会話の方が、話者交替の際の平均ポーズ長は長い。
2. 現発話者に注視されていた聞き手が次発話者になる割合は、0~0.2秒程度のポーズが生じる発話の場合両言語で同程度である。
3. 母語会話では、ポーズ長が長くなるにつれて、現発話者によく注視されて聞き手が次話者になる割合は低下する。
4. 第二言語会話では、ポーズ長が長くなっても、現発話者によく注視されていた聞き手が次発話者になる割合は低下しない。

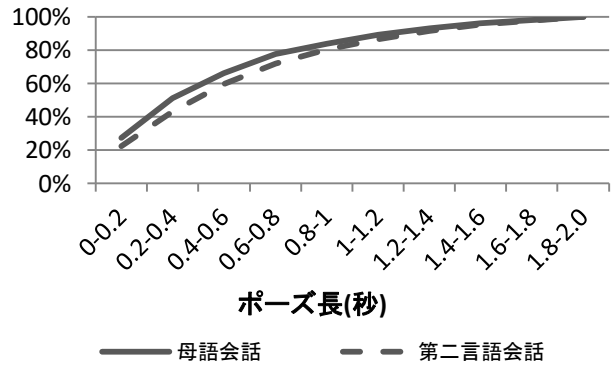
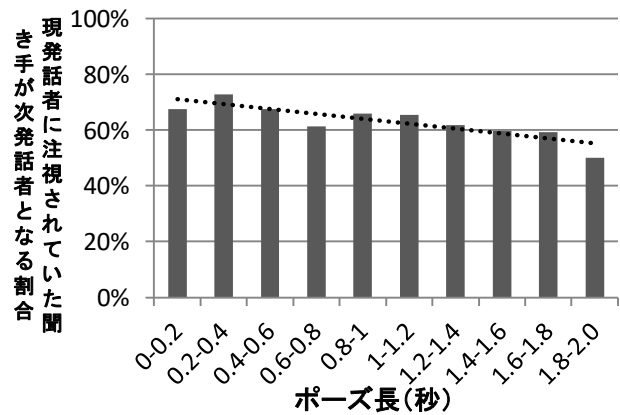
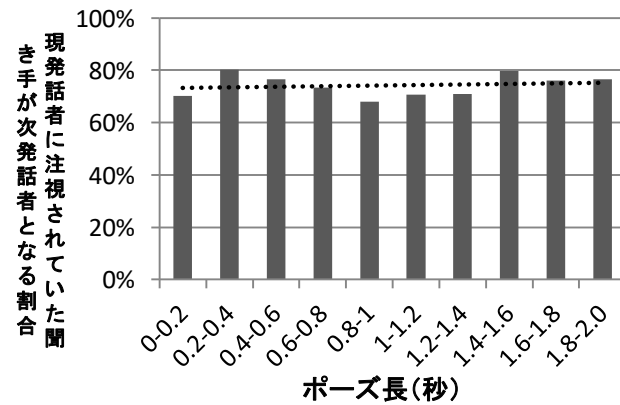


図3 話者交替時のポーズ長の累積分布



(a) 母語会話



(b) 第二言語会話

図4 現発話者に注視されていた聞き手が次発話者となる割合の時間的変化に関する母語会話と第二言語会話の比較。図中の点線は各ポーズ区間での次発話者となる割合についての線形回帰直線

5. 考察

第 3 節, 第 4 節の分析結果から, 母語会話と第二言語会話での注視行動の効果の違いについて以下のように考えられる。

1. 第二言語会話では, 会話参加者の言語運用能力が低く, 直前の発話の理解や応答の生成に時間を要するため, 一般にポーズ長が長くなると考えられる。
2. 0~0.2 秒という短いポーズ後に応答している発話については, 現発話者に注視されていた聞き手が次発話者となる割合について母語会話と第二言語会話での差がほとんどない。これは, 第二言語会話での直前の発話が比較的容易な内容であるため, ポーズをおかずに応答した可能性が高い。このことから, 次発話者が発話するタイミングは直前の発話の難易度や次発話者の言語運用能力に関係していると考えられる。今後, ポーズ長が短い場合の直前の発話の内容を精査し, 発話内容の難易度について考察する予定である。
3. 第二言語会話では, 現発話者は聞き手を注視することで発話権の譲渡先を選択する効果が高いことが示されており, 言語運用能力の低さから, 聞き手は直前の発話を理解することや応答に長い時間を要することで選択効果が長く持続される可能性が考えられる。
4. 母語会話では, ポーズ長が長くなるにつれて, 現発話者に注視されていた聞き手が次発話者となる割合は減少し 50% に近づいていく傾向にある。母語会話では, 聞き手は直前の発話内容を容易に理解できると想定される。そのため, 母語会話において, ポーズ長が長い発話では, 現発話者の注視による次発話者の選択行為以外の種々の要因の影響が大きくなっていると考えられる。

6. まとめ

本研究では, 現発話者の注視が次発話者の選択に関してどのような影響を及ぼすのかについての分析を目的とし, 現発話終了から次発話までのポーズ区間を含めた区間での, 現発話者から次発話者への注視割合の時間的変化に関して, 母語会話と第二言語会話で比較分析を行った。その結果, 母語会話に比べて第二言語会話では, 現発話者は発話中に注視した次発話者をポーズ区間においても長く注視することがわかった。

次に, 現発話終了から次発話までのポーズ長に対応して, 現発話者に注視された聞き手が発話する割合の時間的変化に関して, 母語と第二言語会話で比較分析を行った。その結果, 母語会話では, ポーズ長が長くなるにつれて, 注視されていた聞き手が次発話者となる割合は低下した。一方, 第二言語会話では, ポーズ長が長くなっても, 注視されていた聞き手が次発話者となる割合は低下しなかった。このことから, 母語会話に比べて第二言語会話の方が, 話者の注視が次発話者の選択に及ぼす効果はより長く持続するといえる。一方, 0~0.2 秒という短いポーズ区間の後に, 注視されていた聞き手が次発話者となる割合は両言語で同程

度であった。このことは, 短いポーズ後に応答している発話については, 直前の発話が比較的容易な内容であるため, ポーズを置かずに応答した可能性が高く, 注視による次発話者の選択の効果は聞き手の言語運用能力に関係していることを示唆している。今後, 発話内容や会話参加者の言語運用能力に着目した注視行動の次発話者選択に対する影響に関する分析を行う予定である。

7. 謝辞

ご議論頂いた馬田一郎博士 (KDDI 総合研究所) に感謝します。なお, 本研究の一部は JSPS 科研費 15K02738 の助成を受けたものです。

参考文献

- [1] A. Mehrabian and M. Wiener, "Decoding of inconsistent communications", *Journal of Personality and Social Psychology*, Vol. 6, no. 1, pp. 109-114, 1967.
- [2] A. Mehrabian and S. R. Ferris, "Inference of attitudes from nonverbal communication in two channels", *Journal of Consulting Psychology*, Vol. 31, no. 3, pp. 248-251, 1967.
- [3] H. H. Clark, "Using Language", *Cambridge University Press*, 1996.
- [4] A. Kendon, "Some functions of gaze-direction in social interaction," *Acta Psychologica*, vol. 26, pp. 22-63, 1967.
- [5] Vergeal, R., Slagter, R., Verr, G., & Nijholt, A. (2001). Eye gaze patterns in conversations: there is more to conversational agents than meets the eyes. In CHI '01 proceedings of the SIGCHI conference on human factors in computing systems (pp. 301-308), ACM Press, Seattle, USA.
- [6] Jokinen, H. Furukawa, M. Nishida, and S. Yamamoto, "Gaze and Turn-taking Behavior in Casual Conversational Interactions," *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 3, no. 2, Article 12, 12:1-30, 2013.
- [7] E.S. Veinott, J. Olson, G.M. Olson, and X. Fu, "Video helps remote work: Speakers who need to negotiate common ground benefit from seeing each other," *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pp. 302-309, ACM, 1999.
- [8] S. Yamamoto, K. Taguchi, K. Ijuin, I. Umata, and M. Nishida, "Multimodal corpus of multiparty conversations in L1 and L2 languages and findings obtained from it," *Language Resources and Evaluation*, vol. 49, no. 4, pp. 857-882, 2015.
- [9] K. Ijuin, K. Taguchi, I. Umata, and S. Yamamoto, "Eye gaze analyses in L1 and L2 conversations: From the perspective of listeners' eye gaze activity," *Proceedings of the 2014 workshop on Understanding and Modeling Multiparty, Multimodal Interactions*, pp. 33-37, 2014.
- [10] K. Ijuin, Y. Horiuchi, I. Umata, and S. Yamamoto, "Eye gaze analyses in L1 and L2 conversations: Difference in interaction structures," *International Conference on Text, Speech, and Dialogue*, pp. 114-121, Springer, 2015.
- [11] K. Ijuin, I. Umata, T. Kato, and S. Yamamoto, "Difference in Eye Gaze for Floor Apportionment in Native- and Second-language Conversations", *Journal of Nonverbal Behavior* (in press).
- [12] A. Kalma, "Gazing in triads: A powerful signal in floor apportionment", *British Journal of Social Psychology*, 31 (1), pp. 21-39, 1992.