

## 話者による注視の発話ターン制御機能と 発話の基盤化アクトについて

### Speech Turn Organization Function of Gaze and Grounding Acts in Triad Communication

馬田 一郎<sup>†</sup>, 伊集院 幸輝<sup>‡</sup>, 加藤 恒夫<sup>#</sup>, 山本 誠一<sup>#</sup>  
Ichiro Umata, Koki Ijuin, Tsuneo Kato, Seiichi Yamamoto

<sup>†</sup>KDDI総合研究所, <sup>‡</sup>産総研, <sup>#</sup>同志社大学  
Cognitive Science University, JCSS Corporation  
jcass@jcass.gr.jp

#### 概要

発話ターン制御に注視行動が果たす機能について、発話の基盤化アクト毎に注視行動の分析を行った。言語能力が与える影響を詳細に比較する目的で、母語 3 人対話と第 2 言語 3 人対話の注視行動を比較した結果、両言語条件に共通して *ack* 発話では現話者から次話者への注視量は他の基盤化アクトに比べ低くなることが観察された。これは *ack* 発話では、基盤化だけを行いそれ以外の新規情報を提示しないため、現話者の次話者の選択権限が低くなるためと考えられる。

キーワード: 視線, 発話ターン, 基盤化アクト, 多人数対話, 言語能力

#### 1. はじめに

注視行動は、マルチモーダルインタラクションにおいて感情の表出、社会的コントロールの行使、発話の情報構造のハイライティング、発話ターンの調整など、重要な役割を果たすことが観察されている(Argyle, Lallijee, & Cook, 1968; Duncan 1972; Holler & Kendrick 2015; Kendon 1967)。中でも、発話ターンとの関係については早くから分析が行われており、母語での二人対話における発話ターン調整機能(Kendon 1967; Argyle, Lallijee, & Cook, 1968; Duncan 1972)を示した研究や、多人数対話において同様の機能を確認した研究(Rosano, 2013; Jokinen, Furukawa, Nishida, & Yamamoto, 2013; Holler & Kendrick, 2015; Yamamoto, Taguchi, Ijuin, Umata, & Nishida 2015; Ishii, Otsuka, Kumano, Matsuda & Yamato, 2013; Ishii, Otsuka, Kumano & Yamato, 2016; , Auer, 2018; Ijuin, Umata, Kato & Yamamoto, 2018)が行われている。さらに、注視の発話ターン調整機能は、聞き手から話者への注視量や現話者から次話者への注視量が第 2 言語でより多いことから、第 2 言語でより顕著であることも示されている (Yamamoto, Taguchi, Umata, Kabashima & Nishida, 2013; Umata, Yamamoto, Ijuin, Nishida, 2013; Yamamoto, Taguchi, Ijuin, Umata & Nishida, 2015; Ijuin, Umata, Kato & Yamamoto, 2018)。

しかし一方で、注視がコミュニケーションで果たす役割は、会話の場面によって異なってくることも示されている (Kleinke 1986; Holler & Kendrick, 2015; Kendrick & Holler, 2017; Rosano, 2013; Rossano, Brown, & Levinson, 2009; Stivers & Rossano, 2010)。注視行動はコミュニケーションの文脈の影響を受けるため、発話中の注視行動を分析するには発話のコミュニケーション機能を考慮することが重要である。本研究では、基盤化アクトの観点に基づく発話の機能分類に着目し、発話の機能と注視の発話ターン調整機能との関係、およびこの関係と言語能力との関係について分析を行う。

#### 2. 分析手法

本研究では、母語および第 2 言語による 3 人対話コーパス(Yamamoto, Taguchi, Ijuin, Umata & Nishida, 2015 参照)に、基盤化アクトタグを付与したデータを用いた日本人学生 3 人 20 組(全 60 人、年齢: 18-24 歳)による対話データについて分析を行った。本分析では、好きな食べ物について話すなど、参加者間で共通の結論をまとめる必要のない自由対話と、「無人島に何を持っていくか」などを決める、いわゆるサバイバル課題会話とを対象とし、日本語会話と第 2 言語である英語会話における注視傾向を比較分析した。データ収集と処理に際して、注視行動の計測にはナックイメージテクノロジー社製 EMR-9 視線検出装置を用い、発話区間切り出しおよび発話機能ラベリングにはアノテーションツール ELAN を用いた。基盤化アクトに基づく発話の機能分類については、Traum の体系(Traum, 1994)に基づいた下記の基盤化アクト分類に従い、訓練を受けた作業員 1 名がアノテーションを行った(図 1 参照)。ただし、データ中には *ack* と *init* の両方の機能を 1 発話で兼ね備える場合が多く見られたため、このような発話の基盤化アクトラベルを *ack\_init* とした。

- ・ *init* (始動): 談話ユニット(DU) を構成する最初の発話
- ・ *cont* (継続): 同じ話し手によってなされる直前の発話の継続
- ・ *ack* (承認): 直前の発話の理解を主張する発話・*repair* (修理): DU の内容の変更
- ・ *reqRepair* (修理要求): 他者による修理の要求
- ・ *reqAck* (承認要求): 直前の発話に対して承認を要求
- ・ *cancel* (キャンセル): *init* したが *ack* を待たず、その DU を放棄
- ・ *ack\_init* (承認\_始動): *ack* と *init* 両方の機能を同時に持つ発話

作業者は録画データと発話書き起こしを参照し、ELAN を用いて全ての発話に基盤化アクトタグを付与した。以下では、出現頻度の極めて少ない基盤化アクト分類を除外し、*init*, *cont*, *ack*, *ack\_init* の4カテゴリについて、量的分析を行った。

### 3. 分析: ターン推移前の発話における、現話者から次話者への注視

発話ターン推移直前の発話における現話者から次話者への注視量について、言語能力要因として母語(L1)/第二言語(L2)、基盤化アクト要因を *init*, *cont*, *ack*, *ack\_init* の4カテゴリとし、以下の仮説に基づき分散分析を行った。

**仮説:** 基盤化だけを行いそれ以外の新規情報を提示しない *ack* 発話では、現話者の次話者の選択権限が低くなり、その結果現話者から次話者への注視量は他の基盤化アクトに比べ低くなる

分析の結果、言語と基盤化アクトの両要因において有意な主効果がみられた(言語: Language:  $F_{(1, 18)} = 25.441$ ,  $p = .000$ ,  $SS = 0.262$ ,  $MS = 0.262$ ; 基盤化アクト:  $F_{(3, 54)} = 10.355$ ,  $p = .000$ ,  $SS = 0.350$ ,  $MS = 0.117$ )が、交互作用は有意ではなかった。多重比較の結果、*ack* と *cont* ( $p = .002$ , 平均の差:  $-.078$ )、および *ack* と *ack\_init* ( $p = .000$ , difference in averages:  $-.133$ )の間で有意な差がみられた。

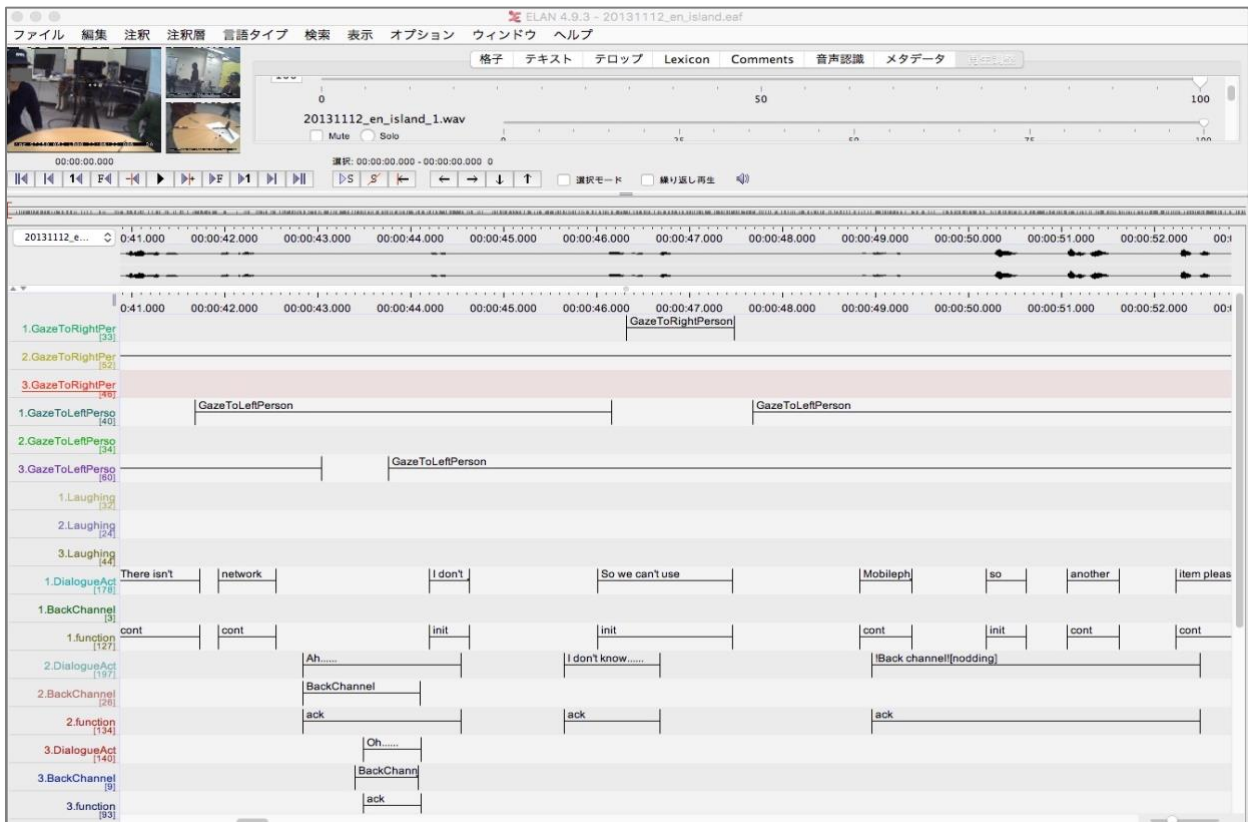


図1 ELANのアノテーション画面

#### 4. 結果と考察

基盤化アクトについての仮説については、*init* 発話と *ack* 発話の間を除く全ての基盤化アクト間で支持された。*ack Int* に比べ *ack* で 現話者から次話者への注視が両言語に共通して少ないのは、自然な結果と言える。*ack* 発話は、前話者の発話内容を基盤化する発話であり、基盤化した上で新たな発話を行う *ack Int* の場合と比べ、現話者の次話者選択における権限が低くなっていると考えられる。*cont* と比べても *ack* で現話者から次話者への注視が少ないのは、発話ターンを長く維持していることにより、現話者の次話者選択権限が上がっている、と考えられる。*init* と *ack* の間で有意差が見られなかった原因は明らかではないが、新たな情報が提示されたばかりの *init* 発話ではまだ会話の流れが確立されていないため、次話者選択の権限もそれほど高くなっていない、という可能性が考えられる。

また、現話者から次話者への注視量は L2 の方が L1 よりも全体に大きいという先行研究での結果(Ijuin, Umata, Kato & Yamamoto, 2018)は再確認されたものの、言語要因と基盤化アクト要因の交互作用が有意でなかった。

このことから、次話者選択において注視が果たす役割は *ack* において *cont* および *ack init* よりも減少していること、および対象とした 4 つ全ての基盤化アクトカテゴリにおいて母語では第二言語より低くなっていることが示されているが、*ack* 発話において注視の次話者指定機能が低下する程度については、両言語条件で同様であると考えられる。

上記の分析の結果、注視の発話ターン調整機能は発話のコミュニケーション機能に影響される可能性が示された。このことは、例えば注視情報を含むマルチモーダルデータに基づく次話者の推定や、インタラクティブインタフェースにおけるエージェントやロボットの設計などにおいても、コミュニケーション機能の影響を考慮する必要があることを示唆しており、今後のインタラクション支援研究などにも重要な知見であると考えられる。

#### 5. まとめと今後の課題

本研究では、注視行動が発話ターン制御に果たす機能について、言語能力要因を考慮しつつ、発話の基

盤化アクト毎に注視行動の分析を行った。分析の結果、言語能力の高低に関わらず、基盤化だけを行いそれ以外の新規情報を提示しない *ack* 発話では、現話者から次話者への注視量は他の基盤化アクトに比べ低くなることが示された。

なお、本研究では会話データにおける現話者・次話者の注視行動を事後的に分析している。従って、現話者の意図した受け手が次話者となっているのかは明らかではない。現話者による次話者選択と発話の受け手、という視点を分析に取り入れていくことは今後の課題である。

#### 6. 謝辞

実験およびデータ処理にご協力いただいた同志社大学情報システムデザイン学科、インテリジェント情報工学科の学生諸氏、示唆に富むコメントをいただきました査読者の方々に深く感謝致します。

#### 7. 参考文献

- [1] Argyle, M., Lallijee, M. & Cook, M. (1968). The effects of visibility on interaction in a dyad, *Human relations*, 21, pp. 3-17.
- [2] Auer, P. Gaze, addressee selection and turn-taking in three-party interaction. (2017). In: *InLiSt - Interaction and Linguistic Structures*, No. 60, August 2017, URL: <http://www.inlist.uni-bayreuth.de/issues/60/index.htm>
- [3] Duncan, S. (1972). Some Signals and Rules for Taking Speaking Turns in Conversations. *Journal of Personality and Social Psychology* 23, pp. 283–292.
- [4] ELAN. <http://www.lat-mpi.eu/tools/elan>
- [5] Holler, J., & Kendrick, K. H. (2015). Unaddressed participants' gaze in multi- person interaction: optimizing reciprocity. *Frontiers in psychology* 6, Article 98 (Feb.2015), 14 pages. <https://doi.org/10.3389/fpsyg.2015.00098>
- [6] Ijuin, K., Umata, I., Kato, T., & Yamamoto, S. (2018). Difference in eye gaze for floor apportionment in native- and second-language conversations. *Journal of Nonverbal Behavior* 42, pp. 113–128.
- [7] Ishii, R., Otsuka, K., Kumano, S., Matsuda, M., &

- Yamato, J. (2013). Predicting next speaker and timing from gaze transition patterns in multi-party meetings. *Proceedings of the 15th ACM on International conference on multimodal interaction (ICMI '13)*. ACM, New York, NY, USA, pp. 79-86. DOI: <https://doi.org/10.1145/2522848.2522856>
- [8] Ishii, R., Otsuka, K., Kumano, S., & Yamato, J. (2016). Prediction of Who Will Be the Next Speaker and When Using Gaze Behavior in Multiparty Meetings. *ACM Trans. Interact. Intell. Syst.* 6, 1, Article 4 (May 2016), 31 pages. DOI: <https://doi.org/10.1145/2757284>
- [9] Jokinen, K., Furukawa, H., Nishida, M., & Yamamoto, S. (2013). Gaze and turn-taking behavior in casual conversational interactions. *ACM Transactions on Interactive Intelligent Systems* 3, 2, 12.
- [10] Kendon, A. (1967). Some functions of gaze-direction in social interaction, *Acta Psychologica*, 26, 22- 63.
- [11] Kendrick, K. H. & Holler, J. (2017). Gaze direction signals response preference in conversation. *Research on Language and Social Interaction* 50, 12–32.
- [12] Kleinke, C. L. (1986). Gaze and eye contact: a research review, *Psychological Bulletin*, 100, 78-100.
- [13] Rossano, F. (2013). Gaze in conversation. *The handbook of conversation analysis*. Stivers, T. Sidnell, J. (Ed.). Wiley-Blackwell, Malden, MA, pp. 308–329. <https://doi.org/10.1002/9781118325001.ch15>
- [14] Rossano, F., Brown, P., & Levinson., S. C. (2009). Gaze, questioning, and culture. Cambridge University Press, pp. 187–249. <https://doi.org/10.1017/CBO9780511635670.008>
- [15] Stivers, T. & Rossano, F. (2010). Mobilizing response. *Research on Language & Social Interaction* 43, 3–31.
- [16] Traum, D. (1994). A computational theory of grounding in natural language conversation. PhD thesis, University of Rochester.
- [17] Umata, I., Yamamoto, S., Ijuin, K. & Nishida, M. (2013). Effects of language proficiency on eye-gaze in second language conversations: toward supporting second language collaboration, *Proceedings of the international conference on multimodal interaction (ICMI2013)*, (pp. 413-419).
- [18] Yamamoto, S., Taguchi, K., Ijuin, K., Umata, I., & Nishida, M. (2015). Multimodal corpus of multiparty conversations in L1 and L2 languages and findings obtained from it, *Language Resources and Evaluation*, 49, 857-882, DOI: 10.1007/s10579-015-9299-2.