

視線が次話者選択に与える影響：  
擬人化エージェントを用いた検討

**How does the speaker's gaze affect next-speaker selection?:  
An experimental investigation using an anthropomorphic agent**

伝 康晴, 小森 淑子\*  
Yasuharu Den, Yoshiko Komori

千葉大学文学部  
Faculty of Letters, Chiba University  
den@cogsci.l.chiba-u.ac.jp

**Abstract**

In this paper, we investigate the effect of the speaker's gaze on next-speaker selection by using experimental three-party conversations with an anthropomorphic agent. In a target utterance, the agent's gaze was i) put between the two participants, ii) quickly switched from one participant to the other at the end of an utterance, or iii) placed at the other participant than was expected as next speaker in the context. The results showed that these "unusual" gaze behaviors affected next-speaker selection in some remarkable ways; e.g., the next speaker was unbiased between the two participants or contrary to the gaze direction, and the reaction time was increased at the target, and even the following utterances. The timing of gaze-shift was also suggested to be influential.

**Keywords** — Gaze, next-speaker selection, anthropomorphic agent

**1. はじめに**

会話の中で、その時々発言機会において誰が何を話すのかは、様々な要因によって規定される。会話参加者たちは、進行中の発話がどこで終わりそうか、その次は誰が話すのかを1発話ごとに局所的に管理し、長い重複や長い間を作ることなく円滑な話者交替を実現している(Sacks et al., 1974)。次話者の決定に際して、言語的手段以外に視線・身体などの非言語的手段も用いられることが知られている(Kendon, 1967; Duncan, 1972; Duncan & Fiske, 1977)。とくに、視線は次話者選択の有効な手段の1つであり、Kendon (1967)は、話し手が次話者を注視することで話者交替を合図し、次話者が相互注視によってそれを受け入れることで話者交替が成立するとしている。

しかし、会話データを詳細に観察してみると、話者交替の際に、聞き手に視線が向けられなかったり、複数の聞き手の間で視線が移動したり、文

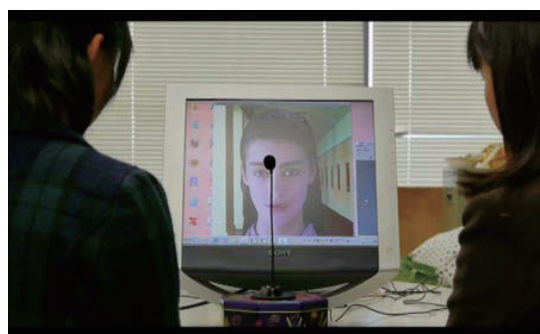


図1 擬人化エージェント (視線方向・左)

脈上の発話の宛て先と視線方向が一致しなかったりする場合が多々見られる。本研究では、このような「典型的でない」視線の使い方が次話者選択に与える影響を、擬人化エージェントとの3人会話場面を設定して実験的に検討する。

**2. 方法**

■実験参加者 18歳から24歳までの大学生ペア10組(男×男:3組、女×女:6組、男×女:1組)計20名  
■刺激と実験計画 フリーソフトの擬人化音声対話エージェント・ツールキット Galatea Toolkit (川本ほか, 2002)を用いて、視線方向(顔の向き)を変えられる擬人化エージェントを作成した(図1)。女性の顔と音声を使用した。「普段の生活」「買い物仕方」「趣味や特技」などをテーマとする、5~7個のやり取りからなる会話を1試行とし、8試行分のシナリオを用意した。個々のやり取りは、[エージェントの質問-参加者の回答-エージェントの承認]の3発話からなり、質問はいずれも「はい」「いいえ」で答える平易なものであった。1つのシナリオは一方の参加者に視線を向けたやり取りから始まり、2~3個のやり取りの後に他方の参加者へ視線を移して会話を続けた。8個中3つのシナリオにおいて、5~6個目のやり取りにターゲット質問を含めた。ターゲット質問ではエージェントの

\*所属は投稿時

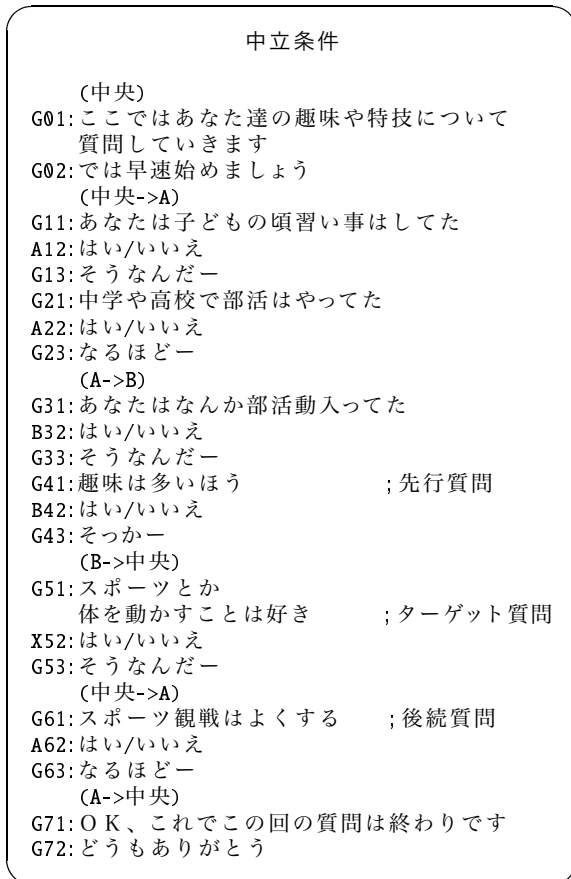


図2 中立条件のシナリオ

視線方向が以下のいずれかに設定された。

**中立条件** 両参加者の中央に視線を移動し、発話を開始する。

**移動条件** 発話の開始時に一方の参加者を見ていた視線を発話末で他方の参加者に移動する。

**不整合条件** 文脈上連続した内容の発話をそれまでの参加者とは異なる参加者に視線を移動して発話する。

中立条件のシナリオを図2に示す。左端のG, A, Bは発話者を示し、Gはエージェント、AとBは想定される回答者を表わす。ターゲット後の回答者がXとなっているのは、この回答がAとBのいずれによってなされる可能性もあることを示している。( )内はエージェントの視線方向がその時点で移動することを示している。この試行では、最初のやり取り(G11~G13)はAに視線を向けて行ない、その後、3番目のやり取り(G31~G33)からはBに視線を向けている。ターゲット質問はG51であり、中立条件では、この質問は視線を両参加者の中央に移動してなされる。これに対する回答X52が参加者AとBのいずれによってなされるかを見ること

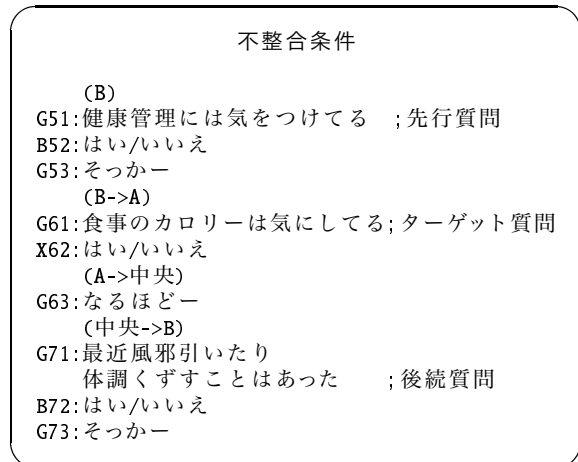
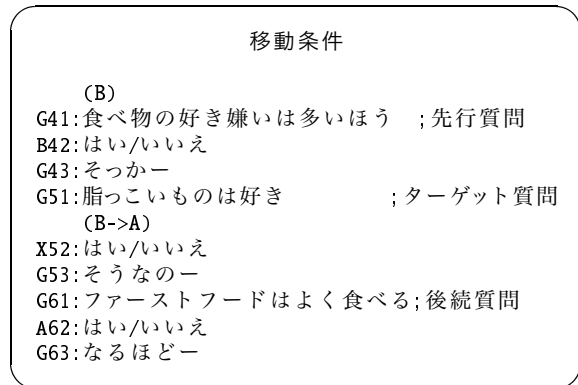


図3 移動・不整合条件のターゲット周辺のやり取り

が実験の目的である。ターゲット後の最後のやり取り(G61~G63)は視線をAに移動してなされている。以下では、ターゲットの1つ前のやり取りの質問(G41)を「先行質問」、1つ後のやり取りの質問(G61)を「後続質問」とよぶ。

他の2つの条件、移動・不整合条件における先行質問から後続質問までのやり取りを図3に示す。中立・移動・不整合条件に用いられたテーマはそれぞれ「趣味や特技」「食生活」「健康管理」であった。

■手続き 2人の参加者は擬人化エージェントを表示したディスプレイに対峙し、横に並んで着座した。エージェントとの対話のためのマイクは両参加者の中央に設置した。エージェントの視線方向は、「左の参加者」「中央」「右の参加者」が明確に区別できるように調整した。実験中の音声と映像をデジタルカメラで収録し、エージェントの質問に対する回答の発話者と発話潜時を記録した。

各ペアは8つの試行をランダムな順序で行なった。ただし、ターゲット質問を含む3つのシナリオは最初の2試行には現れないようにした。また、各シナリオにおける参加者AとBを左右いずれに着座した参加者に割り当てるかはペア間でカウンターバランスした。

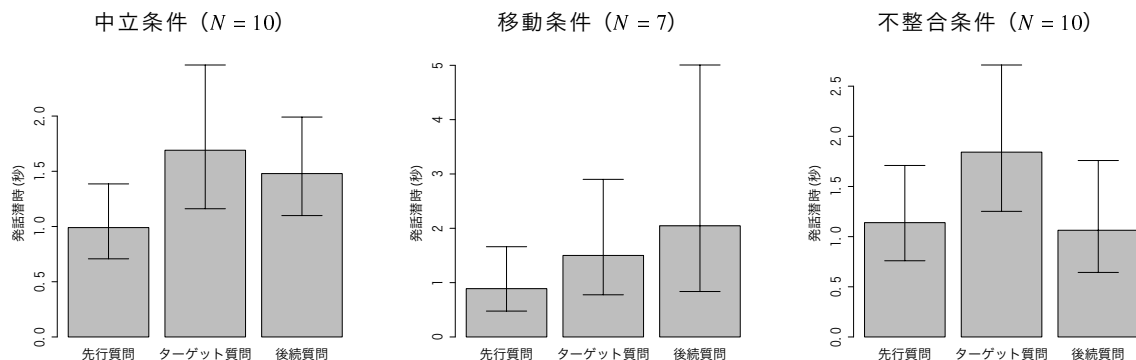


図5 先行・ターゲット・後続質問への回答の発話潜時（エラーバーは95%信頼区間）

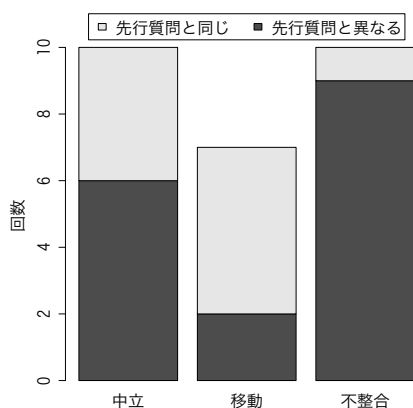


図4 ターゲット質問の回答者

### 3. 結果

#### 3.1 視線の向け先と回答者

全参加者に対するターゲット質問を含む全30試行について、ターゲット質問以外への回答・計160件におけるエージェントの視線の向け先と回答者との一致傾向を調べた。視線の向け先以外の参加者が回答したのは2回(1.25%)のみであり、他はすべて視線の向け先が回答者となっていた。なお、視線の向け先と回答者が一致しない2回は、いずれも移動条件における後続質問に対する回答であった。

#### 3.2 ターゲット質問の回答者

3組のペアに用いた刺激で、移動条件におけるエージェントの視線移動のタイミングが正しく設定できていなかった(発話末ではなく、発話冒頭で移動していた)ことが判明したため、これら3組の移動条件のデータを分析から除外した<sup>1</sup>。

各視線条件においてターゲット質問と先行質問

<sup>1</sup> これら3組の移動条件におけるターゲット質問の回答者は、いずれも移動後の視線の向け先の参加者であった。

の回答者が同じである/異なる回数を図4に示す。両参加者で回答の確率が異なるかを視線条件ごとに二項検定で検討した。

■中立条件 10組中4組(40%)で先行質問と同じ回答者が回答した。二項検定の結果、両参加者の回答確率が有意に異なるとはいえなかった( $p = .75$ )。

■移動条件 7組中5組(71%)で先行質問と同じ回答者が回答した。二項検定の結果、両参加者の回答確率が有意に異なるとはいえなかった( $p = .45$ )。

■不整合条件 10組中1組(10%)で先行質問と同じ回答者が回答した。二項検定の結果、両参加者の回答確率は有意に異なり( $p < .05$ )、先行質問と異なる回答者(移動後の視線の向け先)が回答することのほうが多かった。

#### 3.3 先行・ターゲット・後続質問への回答の発話潜時

各視線条件における先行・ターゲット・後続質問への回答の平均発話潜時(対数値で算出し、線形スケールに逆変換)を図5に示す。これらの値(対数値)を視線条件ごとに分散分析で比較した<sup>2</sup>。

■中立条件 平均発話潜時の質問位置による違いは有意であった(先行質問:0.99秒、ターゲット質問:1.69秒、後続質問:1.48秒;  $F(2, 18) = 6.40, p < .01$ )。ライアン法による多重比較の結果、先行質問への回答の平均発話潜時と比べて、ターゲット・後続質問への回答の平均発話潜時は有意に長かった(ターゲット質問:  $p < .01$ 、後続質問:  $p < .05$ )。

■移動条件 平均発話潜時の質問位置による違いは有意でなかった(先行質問:0.88秒、ターゲット質問:1.50秒、後続質問:2.05秒;  $F(2, 12) = 1.57, p = .25$ )。

<sup>2</sup> 先行質問以前のすべての回答の個人ごとの平均発話潜時、および、直前の質問との回答者の変化の有無を共変量として考慮した線形混合効果モデルによる分析でも、概ね同様の結果が得られた。ただし、中立条件では、ターゲット質問への回答と後続質問への回答の平均発話潜時の間にも有意差が見られた。

■不整合条件 平均発話潜時の質問位置による違いは有意であった（先行質問：1.14秒、ターゲット質問：1.84秒、後続質問：1.06秒； $F(2, 18) = 4.26, p < .05$ ）。ライアン法による多重比較の結果、先行・後続質問への回答の平均発話潜時と比べて、ターゲット質問への回答の平均発話潜時は有意に長かった（先行質問： $p < .05$ 、後続質問： $p < .05$ ）。

#### 4. 議論

ターゲット質問以外への回答の約99%において、エージェントの視線の向け先が回答者であった。参加者たちはエージェントの視線を次話者選択の手掛かりとして用いていたと考えられる。

文脈上の連続性と視線方向が一致しない不整合条件では、90%のペアにおいて、視線の向け先の参加者が回答を行なった。このことから、文脈よりも当該発話の視線方向のほうが次話者選択により強い影響を与えるといえる。一方、平均発話潜時はターゲット質問への回答で先行・後続質問への回答よりも有意に長く、視線方向が文脈と不整合であることが次話者選択への迷いをもたらした可能性を示唆する。実際、「こっちかな」といった逡巡や、両参加者が顔を見合わせ、指差しや頷きで確認するといった行動が観察された。

視線が特定の参加者に向けられない中立条件では、次話者の分布がほぼ半々（4:6）になり、ターゲット質問への回答で平均発話潜時が有意に長くなった。これらのことは、視線が次話者選択のリソースとなっていることをさらに支持する。一方、後続質問への回答でも先行質問への回答と比べて平均発話潜時が有意に長かった。追加的な分析の結果、この傾向は、ターゲット質問において先行質問とは異なる回答者が回答した場合に顕著であることがわかった（異なる回答者：1.81秒、同じ回答者：1.09秒； $t(8) = 2.30, p < .06$ ）。これらの事例では、エージェントの視線がターゲット質問の時点で中央にとどまった後、さらに参加者のほうに移動したことに気づいて戸惑ったり、見つめあったり、笑いあったりといった反応が典型的に観察された。発話潜時の長さはその影響と考えられる。このことは、次話者選択の手掛かりとして参加者たちは、視線方向という静的な情報よりも、視線移動そのものという動的な情報により敏感であることを示唆する。

発話末で視線が移動する移動条件では、統計的には有意でないものの71%のケースで先行質問と同じ回答者がターゲット質問に回答した。ターゲット質問は先行質問の回答者に視線を向けた状態で始まっており、視線の移動は発話終了とほぼ同時であった。このため、視線移動を確認する前に、

先行質問と同じ回答者がマイクに近づく動作を開始することが多かった。これらの事例では、他方の参加者もマイクに近づくそぶりを見せたり、後続質問への回答の際に発話者を指差しなどで確認する行動が観察された。一方、回答前に視線移動に気づいたペアでは、先行質問の回答者が続けて回答しようとして言い止めたり、手で他方の参加者に合図したりする行動が見られた。他方の参加者は視線が自分に向いていることに気づいて驚いた表情を見せたり、「俺か」とつぶやいたりしながら、回答を始めていた。このため、先行質問と同じ回答者が回答する場合よりも異なる回答者が回答する場合のほうが発話潜時は長くなっていた（同じ回答者：1.05秒、異なる回答者：3.68秒）。

これらのことから、視線移動のタイミングが重要であることがわかる。聞き手は発話途中で発話末を予測しながら次発話開始のタイミングをはかっており（Sacks et al., 1974; 榎本, 2007）、多くの場合、発話が実際に終了する時点では次話者は発話開始に取りかかりつつある。そのため、発話末での視線移動は次話者選択に必ずしも有効でないかもしれない。また、何らかの理由で次発話開始が遅れた場合には、このような視線移動は参加者たちを混乱させるもととなる。3.1で触れたように、移動条件では、後続質問において視線の向け先以外の参加者が回答したケースが2件見られた。これらはいずれも、先行質問の回答者がターゲット質問に回答し、そのまま後続質問にも回答したものである。これは、混乱のもととなった視線移動を無視した行動であるように思える。

以上のように、発話者の視線は次話者選択に強い影響を与えるが、同時にそのタイミングが重要であることがわかった。

#### 参考文献

- Duncan Jr., S., & Fiske, D. W. (1977). *Face-to-face interaction: Research, methods, and theory*. Hillsdale, NJ: Lawrence Erlbaum.
- Duncan Jr., S. (1972). Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, **23**, 283-292.
- 榎本美香(2007). 日本語におけるターン構成単位の認知メカニズム. *社会言語科学*, **9**(2), 17-29.
- 川本真一・下平博・新田恒雄・西本卓也・中村哲・伊藤克亘・森島繁生・四倉達夫・甲斐充彦・李晃伸・山下洋一・小林隆夫・徳田恵一・広瀬啓吉・峯松信明・山田篤・伝康晴・宇津呂武仁・嵯峨山茂樹(2002). カスタマイズ性を考慮した擬人化音声対話エージェントツールキットの設計. *情報処理学会論文誌*, **43**, 2249-2263.
- Kendon, A. (1967). Some functions of gaze direction in social interaction. *Acta Psychologica*, **26**, 22-63.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, **50**, 696-735.