

エージェントによる KY とみなされる ノンバーバル行動の識別

Discernment of the nonverbal action it is considered that is KY by an agent

西谷俊輔¹ 大槻怜史² 神田智子²

Syunsuke Nishitani¹, Satoshi Ootsuki², Tomoko Kouda²

¹大阪工業大学情報科学部コンピュータ情報学科

¹ Department of Computer Science, Faculty of Information Science and Technology,
Osaka Institute of Technology

²大阪工業大学情報科学部情報メディア学科

² Department of Media Science, Faculty of Information Science and Technology,
Osaka Institute of Technology

Abstract: KY (“Kuuki-wo yomenai.”) is a term that indicates a situation or person where/who does not read the atmosphere of the place and circumstances. However, the behaviors that generate “KY” situation, or the behaviors that could be regarded as typically “KY”, are not defined clearly. We developed a multi-party interaction scene with virtual agents that exhibit unnatural non-verbal behaviors, i.e., inappropriate gaze, tone of voice, turn-taking, and are examining which non-verbal behaviors are perceived as “KY”.

1. はじめに

人は、その場の雰囲気や状況などを察する（感じる・掴む）さまを「空気を読む」と表現し、自分が関わる会話だけでなく、他者同士の会話からでも「会話が盛り上がっている」や「何か暗い話題を話しているようだ」などといった場の雰囲気を察し、その時自分が何をすべきか、するべきでないかを読み取ることができる。そして場の雰囲気を読み取ることにより失敗し、間違った言動を取ってしまうと周りから「空気が読めない (KY)」などと言って揶揄され、場合によっては敬遠されてしまう[1]。しかし、会話の場の雰囲気を具体的に定義し、雰囲気を感じる方を言葉にして説明することは大変難しく、さらにその雰囲気を感じるメカニズムも明らかとなっておらず科学的定義もない。

内藤[2]は場の空気を読むことは、「場の雰囲気や流れを把握する」、「発言する相手を確認する」、「適切な言葉を選ぶ」、「発話の適切なタイミングを選ぶ」の4つが重要だとし、これらを満足に行わない場合 KY とみなされると述べている。また、大石[3]が行なった調査では、大学生 112 名に対し、学級場面、部活動場面、友人関係場面、アルバイト場面で、「どのような振る舞いをすることを「空気が読めない」こととみなしているのか」を自由記述によって収集した。その結果、「自分ばかりが話す」、「一人だけノリが悪い」、「場の雰囲気に合わないことをいう」、「ま

わりが静かなのに大声で話す」、などの振る舞いが KY であるという意見が多かった。これらの意見をまとめると、KY の振る舞いは内藤[2]が述べた KY とみなされる振る舞い以外に「不適切な声量での発話」も KY の行動であると考えられる。

以上の例を挙げたが、これらの共通項をバーバル・ノンバーバルの行動に分類すると KY だと認識される行動は以下の通りになる。

●バーバル行動

「場の雰囲気に合っていない発言」

●ノンバーバル行動

- (1) 「発言する相手を非凝視」
- (2) 「発話の不適切な交替潜時」
- (3) 「不適切な声量での発話」

しかし、バーバル行動は会話内容が変動しやすく、「空気が読めない」という表現は非常に曖昧なものであるだけに正確な評価が望めない。よって本研究では言語を用いず会話内容に影響されないノンバーバル行動のみに着目する。

本研究の目的は、(1) - (3)のノンバーバル行動から、人が KY だと認識されるノンバーバル行動(以下 KY 行動とする)の識別をする事である。そのためにエージェントを用いた多人数会話から印象評価実験を行う。また実験後にアンケートとして BigFive を元にして作成された TIPI-J (日本語版 Ten Item Personality Inventory) [4]を用い、エージェントの心理的印象も評価する。

2. 実験

2.1. 実験システム

評価実験を行うにあたり、本研究では TVML (TV program Marking Language) を用い会話エージェント 4 体による多人数インタラクションの無意味言葉での会話動画を作成した。TVML で作成したエージェントの多人数インタラクション動画を図 1 に示す。

4 体のエージェントの立ち位置は仮想空間内で、1 辺 2m の正方形の頂点上である。体の向きはそれぞれ対角線とし、水平方向の頭の動きは対角線（エージェントから見て正面）を基準値 0° とした左右のエージェントを見る角度 $\pm 45^\circ$ を限界値とする。垂直方向の頭の動きは一方向（下向き）のみとした。実験では以下で述べる仮定と TIPI-J をもとに印象評価を行い、KY 行動を識別する。

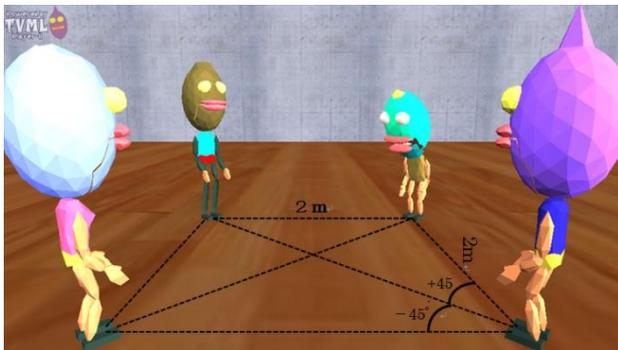


図 1. TVML で作成したエージェントの多人数インタラクション動画の画像例

2.2. 仮定

1 章で述べた KY 行動をエージェントに実装することにより、実験参加者が人間に対する認識と同じくエージェントにも以下の認識を持つと仮定する。

●仮定 1: 人間同士のコミュニケーションでは、発言する相手を確認することが重要とされている[2]。このことから、会話エージェントによる多人数インタラクションにおいて、エージェントが非凝視行動（うつむき、視線逸らし）を行うことにより実験参加者が KY と認識する。

●仮定 2: 前節では「発話の不適切な交替潜時」、「不適切な声量での発話」の 2 つが KY 行動であると述べている。このことから会話エージェントによる多人数インタラクションにおいて、エージェントの不適切な交替潜時、不適切な声量で発話を行うことにより実験参加者は KY と認識する。

2.3. 実験条件

2.2 節の仮定をふまえ、事前実験により以下の実験条件を設定した。

事前実験では、実験で使用する各 KY 行動の基準を設けるため、各行動における動作の程度の上限、下限を測定した。理系大学生 4 名（男性 2 名、女性 2 名）に KY 行動を含んだ動画を評価してもらい、以下の (1) ~ (3) の各 2 種類ずつの値を設けた。

(1) 「発言する相手を非凝視」

(1)に関しては、非凝視行動を、顔の角度を下方向に変化させる「うつむき」・水平方向に変化させる「視線逸らし」の 2 動作に分類した。

(1-a) 「うつむき（話し手・聞き手）」

うつむきの角度に関しては、 $20\sim 60^\circ$ の角度でエージェントがうつむく動画を 5° 刻みで作成し、事前実験でエージェントがうつむいていると判断する判断下限および上限を、調整法を用いて測定した。

その結果うつむきの度合により KY 度に差が出るかを比較するために以下の 2 種類を実験条件とした。

● 30° : 事前実験より、うつむきの下限とした角度。

● 40° : 事前実験より、うつむきの上限とした角度。

(1-b) 「視線逸らし（聞き手）」

視線逸らしについては、「聞き手と聞き手の間を見る場合」と「話し手とは別の聞き手を見る場合」と定めた。視線逸らしの度合により、KY 度に差が出るかを比較するために以下の 2 種類を実験条件とした。

● 22.5° : 話し手を中心として、水平方向に 22.5° 視線が逸れている。

● 45° : 話し手を中心として、水平方向に 45° 視線が逸れている。

(2) 「不適切な交替潜時（話し手）」

河添[5]は、人間同士のコミュニケーションで通常の交替潜時を 0.36 秒と定義している。これにより、交替潜時が 0.36 秒より短い場合を不適切な交替潜時とした。しかし、0.1 秒や 0.2 秒の交替潜時では不適切な交替潜時と認識できなかったため、交替潜時の値がマイナスとなる発話割り込みを実験条件とした。 $-0.5\sim -3.0$ 秒の交替潜時でエージェントが発話割り込みを行う動画を 0.5 秒刻みで作成し、事前実験で、発話割り込みと認識できる秒数を、調整法を用いて測定した。

発話割り込みの会話が重なる時間の長短より、KY 度に差が出るかを比較するために以下の 2 種類を実験条件とした。

● -1.0 秒: 発話割り込みの下限を、交替潜時 -1.0 秒とした。

● -2.0 秒: 発話割り込みの上限を、交替潜時 -2.0 秒とした。

(3) 「不適切な声量での発話」

人間が耳で聞いて約 2 倍大きいと感じる声量は、原音から+10 dB, 約 1/2 倍と感じる声量は、原音から-10 dBとされている[6]. その場に不適切な声量, 声大きい, 小さい, の 2 条件で KY 度に差が出るかを比較するために以下の 2 種類を実験条件とした.

- +10 dB : 4 体中 1 体だけ声量が約 2 倍.
- 10 dB : 4 体中 1 体だけ声量が約 1/2 倍.

2.4. 実験手順

会話エージェントによる多人数インタラクションの動画を用いて KY 度とエージェントに対する印象を調べる (TIPI-J) アンケート評価を実施する. 実験手順は以下のとおりである.

1. 実験参加者に KY 行動を含まない動画を見てもらい, 「この動画を KY 度 4 (どちらでもない) として下さい」と伝える.
2. 2.3 節で示した KY 行動の内, いずれか 1 つ含んだ動画を 8 個見てもらい KY 度と TIPI-J の比較アンケートに記入してもらう. このとき, 何の KY 行動を含んだ動画なのかは実験参加者に伝えない. どのエージェントに注目してもらうかをあらかじめ実験参加者に伝えておく.
3. KY 行動を含んだエージェントの提示順は, 順序効果を相殺するため, ランダム順とする.
4. アンケートの評価尺度はそれぞれ 7 段階とし, KY 度は「7:非常に KY と感じる-1:全く KY と感じない», TIPI-J は「7:強くそう思う-1:全く違うと思う」とする.

3. おわりに

本研究では, 会話エージェントによる多人数インタラクションの動画を用い KY 行動の識別実験を行う. 実験結果により, KY だと認識されるノンバーバル行動をあきらかにすることで, 逆に会話の雰囲気や空気の読めるエージェントの実現に寄与できることが期待できる.

参考文献

- [1] Yahoo!辞書「空気を読む」:
<http://dic.yahoo.co.jp/dsearch?enc=UTF-8&p=%E7%A9%BA%E6%B0%97%E3%82%92%E8%AA%AD%E3%82%80&dtype=0&dname=0na&stype=0&pagenum=1&index=20473300> (2012 年 10 月 19 日所得)
- [2] 内藤 誼人: 『「場の空気」を読む技術』, サンマーク出版 (2004)

- [3] 大石千歳: 「空気が読めない」とはどうゆうことか?: 社会的スキルの欠如という観点からの検討, 東京女子体育大学・東京女子体育短期大学紀要 第 44 号 (2009)
- [4] 小塩真司, 阿部晋吾, カトローニ ピノ: 日本語版 Ten Item Personality Inventory (TIPI-J) 作成の試み, パーソナリティ研究, 21, 40-52 (2012)
- [5] 河添麻衣子, 北村泰彦: 発話タイミングを考慮したマルチエージェント対話システム, 電子情報通信学会技術報告書, 53-55 (2007)
- [6] 人が聴覚上の感じる音量差とデシベル(dB)の関係:
<http://www.ones-will.com/blog/dtm/music-theory/8848/>
(2012 年 10 月 4 日取得)