

発話交替マインドを伝えるアバタ会話システム - 人と機械での気持ちの通じる対話を目指して -

湯浅 将英[†] 徳永 弘子[†] 武川 直樹[†]

Turn-taking Avatars Exchanging Utterance Minds - Towards Heart-to-heart Communication between Humans and Computers -

MASAHIDE YUASA[†] HIROKO TOKUNAGA[†] NAOKI MUKAWA[†]

1. はじめに

多人数会話は、「誰かが話し終わったから、次に誰かが話す」という単純なものではなく、話の間にも話し手は「自分の話が終わるので誰かに話させよう」、聞き手は「話し手の話が終わりそうだから次に自分が話そう」あるいは「今は話したくないので他の誰かに話させよう」などと思ながら会話が進む。本研究では、人と機械（エージェント）の間で「話したい／話して欲しい」気持ちが通じる発話交替システムの構築を目指す。今回、人と機械でのシステム構築に先立ち、アバタを介した3人のユーザ間の会話において「話したい／話して欲しい」気持ち（発話マインド）を手のジェスチャや体の動き等で表現することによって、思い通りの発話交替を促せるシステムを作成した。システムを前にしたユーザは、遠隔地にいる相手のアバタに向かって、「話したい」あるいは「話して欲しい」というマインドに対応する仕草をすると、その仕草が遠隔地にある自身のアバタ動作に反映されて相手に「話したい」「話して欲しい」気持ちが伝わる。本システムは、人の観察より得られた発話交替マインドモデル[3]に従ったマインド動作の表出と認識をしており、このアバタシステムで観察される発話交替を検証することで、将来、人と機械での気持ちの通じた自然な発話交替の構築が目指せる。

2. 発話交替マインドモデル

これまでの我々の研究により、人の発話交替には視線、表情、ジェスチャに込められる「発話マインド」

が重要であることがわかった[1]。徳永らは、発話交替前に「発話立候補（自分が話したい）」や「発話懇願（自分は話さないで、是非話して欲しい）」「発話促進」「発話回避」などのカテゴリからなる発話マインドが表出されていることを明らかにするとともに、発話マインドが発話交替に先立って誰が次に発話をするかを定める重要な要素であることを示した[1]。湯浅らは、発話マインドを表出するエージェントを実装し、それらの発話意図がユーザに伝達できることを示した[2]。さらに徳永らは、図1のように発話マインドを「自分の発話する意欲（話したい）」と「相手の発話を制御する意欲（話して欲しい）」の2つの軸に還元するマインドマップモデルを提案し[3]、2軸上に表せるマインドを相互に表出、理解することで、発話が継続できるとした。図1にマインドマップと発話交替マインドカテゴリの関係を合わせて示す。また徳永らは評定者によりラベル付けされた発話マインドの約70%が、表出者の意図する通りの発話交替の結果となることから、マインド表出が有効に機能しているとした[3]。このマインドモデルをアバタシステム、将来的にはエージェントに適用すれば、人と同じように気持ちが通じる発話交替が実現できると考える。

3. システムコンセプト

本システムでは、遠隔地にいる3人を結び、それぞれの人をアバタとして表示し、会話をする。図1のような人物映像からマインドを表す動作を認識し、遠隔地に表示されるCGアバタにマインド表出動作をさせる。「体を伸ばす／縮める」動作が「話したい／話したくない」を、「手を差し出す／手でさえぎる」動作を「相手に話して欲しい／黙って欲しい」に集約した。

[†] 東京電機大学情報環境学部
Tokyo Denki University

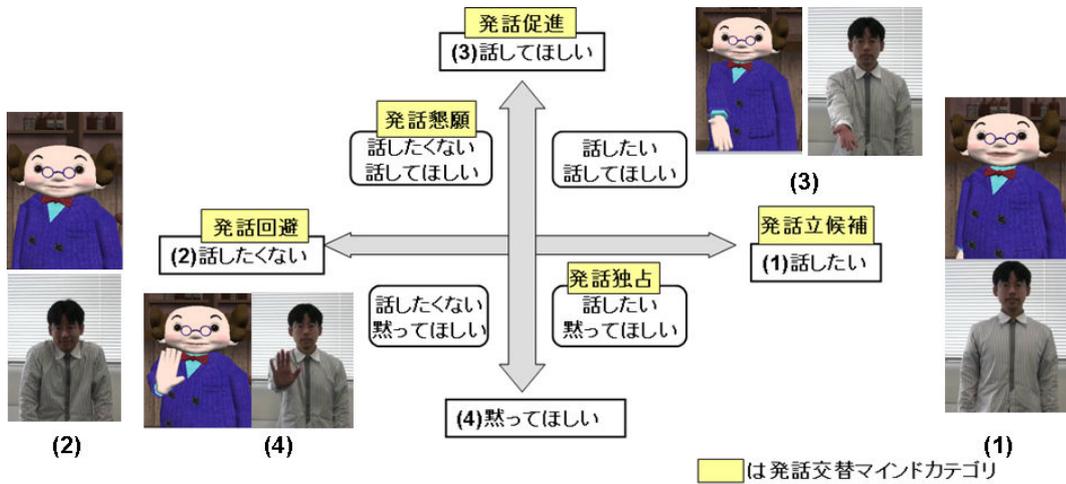


図 1, 2次元発話交替マインドマップと発話交替マインドカテゴリ, およびアバタ[4]とユーザの動作

たとえば、「首や体を伸ばすことで話したい意欲を示す」[2], 「“どうぞどうぞ”と手を差し向ける仕草で発言を促す」など, 人が日常で自然に表出し理解できるものを選んでいいる. 対面の会話における発話マインドは多様かつ微妙な表情, 仕草によって伝えられる[3]が, ここでは, マインドの機能を人にとっても機械にとっても理解しやすい表現にすることを優先している.

4. システム構成

4.1 アバタの表示

1人のユーザの前にディスプレイ2台が配置され, 遠隔のユーザを表す2体のアバタが表示される. 図2左のアバタは遠隔ユーザが実際にこちらを見ており, 右のアバタは遠隔ユーザが左の遠隔ユーザを見ていいることを表し, 会話の空間を正しく構築している.



図 2, 発話交替アバタシステム (1ユーザ分)

4.2 顔, 手の認識

ユーザの顔の向き, 位置, 手の仕草を画像認識し, 3体のアバタが同期動作して会話する. ディスプレイ周辺の WEB カメラにより顔の向きと顔の位置を認識

し, ユーザが目の前にいる2体のアバタのどちら側を見ているかを推定する. その結果を用いて相手側に表示されるアバタの顔向きを反映させる. また, 指に赤いマーカをつけ, 手の動作を推定する.

5. まとめ

人の観察による発話交替マインドモデルに基づきアバタを介した多人数会話システムを構築した. このシステムを用いて人の気持ち (マインド) が通じる発話交替のモデルを評価することが可能である. 現在, ユーザに実際に会話してもらい, 発話交替にマインド表出動作が寄与するかを検証している. 「自分の意見があるときには体が伸びる, 顔を前に近づける」「困ったときには手を差し向ける」など, 有効なマインド動作が観察され, 「音声チャットよりも発言がぶつかることが少なかった」「発言するときの心の準備が伝えられた」との体験者の肯定的な意見があった. 今後, アバタを自律的な擬人化エージェントに置き換えたシステムとすることで人と機械とで話したい気持ちが通じるシステムを構築予定である.

参考文献

- 1) 徳永, 湯浅, 武川, 3人会話の発話交替における視線行動と発話マインド分析, フォーラム顔学 2007, (2007)
- 2) 湯浅, 徳永, 武川: 発話マインドに基づく発話交替モデル, インタラクシオン 2008, (2008)
- 3) 徳永, 湯浅, 武川: 3人会話の視線と発話マインド分析ー聞き手の発話マインドは話者交替に機能するか?ー, HCS 2008, (2008)
- 4) TVML: <http://www.nhk.or.jp/strl/tvml/>