

音声対話におけるエージェントの態度が人間に与える影響*

7R-3

八木正紀 平栗 覚 伊賀聡一郎 安村通晃†

慶應義塾大学環境情報学部‡ 慶應義塾大学総合政策学部§

慶應義塾大学大学院 政策・メディア研究科¶

近年、音声認識・音声合成を利用したエージェントベースの音声対話システムの研究が多くみられるようになってきた。エージェントに様々な表情を持たせることで新しいヒューマンインターフェイスを提供している。しかし、エージェントに様々な表現力を持たせる反面、それらの人間に与える影響が考慮されていない。

そこで、本研究では人間とコンピュータ内のエージェントが互いに音声を用いてやりとりする上で、エージェントからの情報提示の変化による人間側の発話に与える影響を調査するために、Wizard of Oz(以下WOZ)方式による模擬音声対話システムを構築し、その実験の結果について考察を加えることとする。

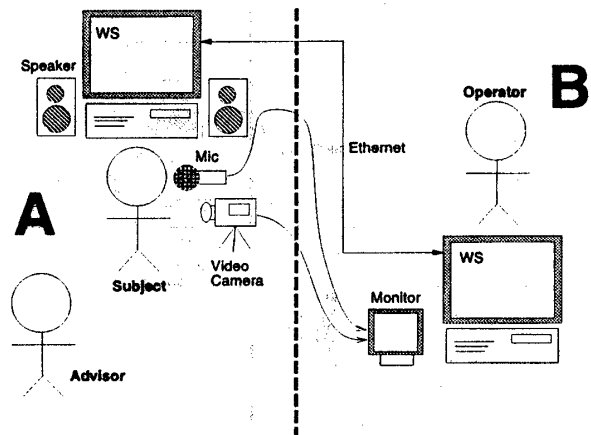


図 1: WOZ 実験システム構成

1 実験システムについて

実験システムは、コンピュータから提示される音声によるメッセージが人間の発話に及ぼす影響を与えるかという点に注目し、WOZ方式を用いた模擬的な音声対話システムである。図1にシステムの構成を示す。

部屋Aに被験者とアドバイザー、部屋Bにオペレーターがおり、被験者の実験の様子と音声はビデオカメラとマイクによりオペレーターに送られる。被験者にはオペレーターの存在は告げられない。オペレーターは被験者がマイクで発話した音声を元に返答を選択し、部屋Aの端末へ送信する。送信された返答は音声合成ソフトウェアによる合成音声として被験者に提示される。実験のパラメータを対話情報だけに限定するため、ディスプレイには簡単な説明があるのみで、グラフィカルな情報は提示していない。アドバイザーは実験の説明と問題が生じたときの手助けをするのみで、実験中は被験者とのインタラクションは行なわないようにした。

*Human attitude effected by the attitudes of agents in a speech dialogue system

†Masanori YAGI, Satoru HIRAGURI, Soichiro IGA, Michiaki YASUMURA

‡Keio Univ. Faculty of Environmental Information

§Keio Univ. Faculty of Policy Management

¶Keio Univ. Graduate School of Media and Governance

2 実験の方法

本実験に先立って、同様のタスクを用いて予備実験を行なった。その際の対話例を見てみると、多くの被験者側が単語に近いレベルでの発話しかしていなかった。

そこで、エージェントの態度を変化させることにより、ユーザの発話に影響が表れるのではないかと考え、本実験を行うに至った。ここでは、エージェントの態度がより人間に近いほど、それにつられて人間の発話で情報伝達に必要な単語(以下、不要語)が増加するのではないかと、という仮説を検証したい。

被験者には、当大学湘南藤沢キャンパス(SFC)内の教員に関する情報を音声対話によって引き出してもらうという課題を与える。

エージェントには返答の仕方が異なる3つのモードを用意した。提示する情報自体は同一であるが、冗長性や口調はモードにより異なっている(表1)。

被験者には、3つのモードにつき最低一人以上の教員について検索を行なってもらう。また、モードの順番による影響を減らすため、被験者によってモードの順番を変更した。各モードでの実験が終わるたびに被験者から簡単な感想を聞く。また実験終了後に、システムに対する評価についてのアンケートに答えてもらう。

表 1: モードによる変化

モード	冗長性	口調
1	高い	丁寧
2	低い	丁寧
3	低い	ぞんざい

被験者は、比較的音声対話システムに関する知識の少ない SFC の学生 (8 名) である。

3 対話例

対話例の一部を各モードごとに示す。

(被:被験者の発話 シ:システムの発話 を示す)

(モード 1)

シ: あなたの知りたいのはあいだかずお先生についてですか。
 被: はい、そうです。
 シ: あなたの知りたい情報は何でしょう。
 被: あいだ先生の今期の担当科目はありますか。
 シ: あいだ先生の担当の講義には、企業と会計、経済管理と会計情報があります。あいだ先生について他に知りたい情報があれば教えてください。

(モード 2)

シ: えとうじゅん先生ですか。
 被: はい。
 シ: 知りたい情報は何でしょう。
 被: 研究会のテーマについて教えてください。
 シ: 翻訳論です。他に知りたい情報があれば教えてください。

(モード 3)

シ: ふるかわこういち先生のことかな。
 被: そうです。
 シ: 何を聞きたいんだね。
 被: 先生のオフィスアワーを教えてください。
 シ: それなら、月曜十五時から十六時三十分だ。他に知りたいことがあったらほしい。

4 本実験の課題と考察

前述の仮説に基づき、発話内の全単語のうち、被験者別・モード別に不要語に占める割合を求め、平均を取ったものが表 2 である。

これを見る限り、本実験においてはモード間の差異よりも被験者間の個人差の方が大きいように見える。またアンケートの項目の中の、「どのモードが最も人に近いと思うか」「どのモードが最も対話しやすかったか」という質問の答とも相関があるとはいえない。

本実験において考えられる課題を以下に挙げておく。

表 2: 全語数における不要語の占める割合 (%)

被験者	mode1	mode2	mode3	平均
1	64.3	62.5	63.1	63.3
2	45.5	27.3	50.0	40.9
3	68.3	60.0	69.4	65.9
4	47.6	75.4	48.7	57.2
5	65.2	63.6	20.0	49.6
6	77.0	82.4	79.4	79.6
7	58.3	44.4	37.5	46.7
8	13.6	18.2	16.0	15.9
平均	55.0	54.2	48.0	52.4

(1) タスクについて

今回行った実験のタスクでは、エージェントとのやりとりは事実上単語に近いレベルの発話ですんでいる可能性がある。より複雑な文レベルの発話を起こしやすいタスクの採用を検討したい。

(2) モードについて

実験後のアンケート等で被験者から多く指摘を受けていたことに、「モード 1 とモード 2 にあまり違いが感じられない」というものがあった。モード間の差異をさらに明確にする必要がある。また、人間同士で同様のタスクを行った場合の対話を収集して比較することにより、共通点・相違点がはっきりしてくると思われる。

(3) 被験者の数

個人間の差異の影響を減少させるため、被験者の数を増やす必要があると思われる。

以上の課題を踏まえてさらに実験を継続し、エージェントの態度が人に与える影響を明らかにしたい。

参考文献

- [1] 安村通見, 伊賀聡一郎, “マルチモーダル・インターフェイスの試み, 情報処理学会第 35 回プログラミング・シンポジウム, pp.93-100, 1994.
- [2] Hauptmann, A., and Rudnicky, A., “Talking to Computers: An Empirical Investigation.”, *International Journal of Man-Machine Studies*, 28(6), pp.583-604, 1988.