

文構造を考慮したオーバーラップを含む発話タイミングの評価

○吉田奎一郎 大村卓矢 山本知仁（金沢工業大学） 三宅美博（東京工業大学）

概要 これまでわれわれは音声対話システムにおいて、正の反応潜時における適切な発話タイミングについての調査を行ってきた。一方、実際の対話では反応潜時が負になるオーバーラップが起きることがある。この負の反応潜時における予備的な実験では、文末付近でのオーバーラップが好印象を与えることが示唆されている。本研究では、文構造を考慮したオーバーラップの印象評価実験を行い、さまざまな発話タイミングがどのような印象を与えるかを調査した。

キーワード: 音声対話システム, 反応潜時, オーバーラップ, 文構造

1. はじめに

音声対話システムは近年、Apple の「Siri」のようなサービスとして実用化され、その有用性が広く認知されている。しかし、現在の音声対話システムと人間とのコミュニケーションでは、発話の間や声の大きさといった非言語情報が十分に考慮されていないため、利用者が違和感を持ってしまうことがある。

このような問題を解決するために、これまでわれわれは発話タイミングを制御する音声対話システムを構築し、反応潜時が正の場合の適切な発話タイミングについて評価を行ってきた¹⁾。一方、人間同士の対話では反応潜時が負になり、相手の発話に割り込むオーバーラップが発生することが知られている。これは対話が盛り上がった時や、親しい間柄での対話で起きやすいと考えられており、オーバーラップを考慮した音声対話システムを実現することは、より自然な人間との対話を実現することにつながると考えられる。

そこでわれわれは、音声対話システムにオーバーラップしながら応答する機能を実装し、予備的な評価実験を行った²⁾。結果として、発話開始後すぐのオーバーラップは悪印象を与えるが、文末付近でのオーバーラップは好印象を与えることが明らかになった。この結果はオーバーラップが起こる位置によって、応答の印象が大きく変わることを意味し、その位置は文章の構造に大きく依存することが推測される。

そこで本研究では、文脈などを考慮しない単純な発話を文節で区切り、オーバーラップを行うタイミングを言語情報によって指定する。続いて、この発話を用いてシステムと対話を行い、各オーバーラップを含む応答がどのような印象を与えるかを調べる。

2. 実験手法

2.1 実験概要

実験は、音声認識エンジン Julius を用いて構築した音声対話システムを用いて行った。Fig.1 に実験環境を示す。本実験では被験者が発話する内容を「明日の天気はどう」、「あなたは英語を話せますか」というものとし、それに対してシステムが「明日の天気は晴れです」、「はい、話せます」と応答するものとした。このときオーバーラップのタイミングは、被験者の

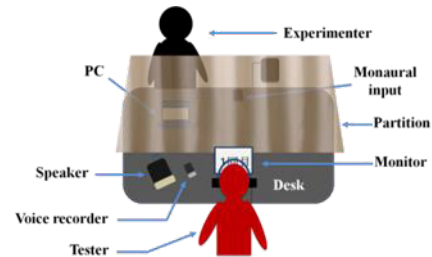


Fig. 1: Experiment environment

発話各 5 回ずつをあらかじめ録音しておき、それぞれの文節までの時間を計測して、その平均値より算出した。具体的なオーバーラップのタイミングとしては 2 つの質問文を主語、述語、修飾語の文節に分け、以下に示す発話開始後の時間とした。

(Timing 1-1)「明日の」：250ms 後

(Timing 1-2)「明日の天気は」：700ms 後

(Timing 1-3)「明日の天気はどう」：950ms 後

(Timing 2-1)「あなたは」：250ms 後

(Timing 2-2)「あなたは英語を」：650ms 後

(Timing 2-3)「あなたは英語を話せますか」：1350ms 後

2.2 実験手順

実験は以下の手順で行った。

1. 被験者に実験の概要を説明し練習対話を行う
2. 比較する 2 つの発話タイミングをランダムに選びそれぞれ 2 回発話を行う
3. 対話システムの応答の印象に関する評価（7 項目）を質問紙に記入する
4. 異なるタイミングの組合せ（全てで 3 通り）で手順 2, 3 を行う

実験において発話速度は被験者にとって自然な速度とした。また実験には、20 代の男性 10 名（平均年齢 22.1 歳）が参加した。各応答の評価は一対比較法（中屋の変法）を用いて、手順 2-3 に示すように比較する 2 つのタイミングを 2 回ずつ交互に発話してもらい、そのときの印象を質問紙に記入することで評価を行った。評価項目として「自然か」「好きか」「丁寧か」「親しみやすいか」「速いか」「活発か」「話しやすいか」の 7 項目を用意し、それぞれ 5 段階で評価を行った。

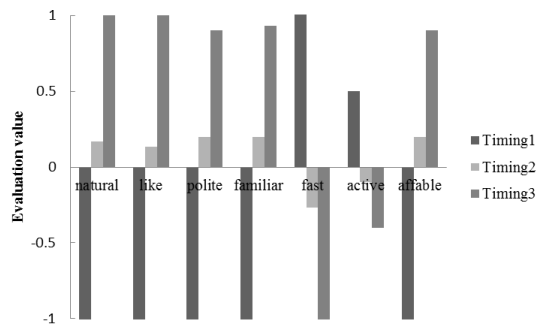


Fig. 2: Result of "What's the weather like tomorrow?"

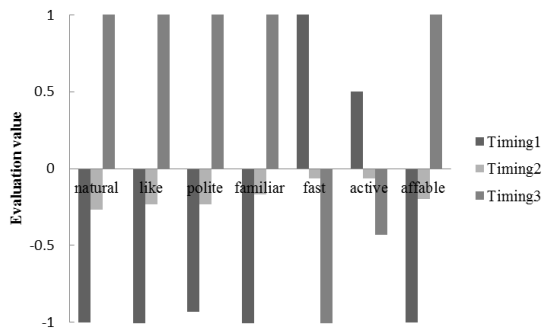


Fig. 3: Result of "Can you speak English?"

3. 実験結果

発話内容が「明日の天気はどう」の結果を Fig.2 に、「あなたは英語を話せますか」の結果を Fig.3 に示す。まず「明日の天気はどう」の結果に関して、「活発か」を除く 6 項目で主効果に有意差が見られた($p < .05$)。傾向としては発話終了直後に応答する Timing3 が良い印象を与えており、文頭でのオーバーラップである Timing1 が多くの項目で強い悪印象を与えているのがわかる。Timing2 のオーバーラップに関しては Timing1 と 3 の中間ぐらいであるが、全体的には、ややよい印象を与えているのがわかる。

次に「あなたは英語を話せますか」に関しては、先の結果と同様に「活発か」を除く 6 項目で主効果に有意差が見られた($p < .05$)。また、全体的な傾向も先と同じく Timing3 が良い印象を与えており、Timing1 は悪い印象を与えているのがわかる。Timing2 のオーバーラップに関しては先の結果と異なり、全体的にやや悪い印象を与えているのがわかる。

4. 考察

本研究では音声対話システムにおける、文構造を考慮したオーバーラップを含む応答の印象について調査した。結果として、システムのオーバーラップが行われる文節の位置によって、応答の印象が変化することが明らかになった。

まず、本研究では発話終了直後に応答を行う条件、いわばオーバーラップが 0 秒となる条件を実験に用意したが、結果として、この発話タイミングが最もよい印象を与えた。オーバーラップは対話が盛り上がった

ときや、親しい間柄での対話で表れやすいものであり、新たな情報を取得する場面ではあまり表れない³⁾。今回の実験内容は、必ずしも先のような状況ではないため、基本的には発話が終了した時点で応答した方が、よい印象を与えたのではないかと考えられる。

以上の結果を踏まえた上で、本実験では文頭に大きくオーバーラップする応答が、両発話内容において共に悪い印象を与えた。この結果は、まだ発話が十分に行われていない時点で、いわば発話を遮る形で音声対話システムが応答することが、発話者に不快な印象を与えることを意味している。

一方、文中のオーバーラップに関しては「明日の天気はどう」ではやや良い印象を与えたものの、「あなたは英語を話せますか」ではやや悪い印象を与えていた。まず、「明日の天気はどう」では述語部分である「どう」が短く、省略しても発話の意図を伝えることができる。そのため、述語に被る形で応答しても会話に自然さや親しみやすさが現れることが考えられる。一方で、「あなたは英語を話せますか」では述語部分の「話せますか」が長い文節であり、発話における意図を含んでいる。そのため、発話者が意図を伝えていないにも関わらず、オーバーラップして応答が返ってくると、話を遮られる形になり印象が悪くなると考えられる。このように、応答のオーバーラップが与える印象は文構造に依存することが考えられ、今後、さまざまな発話内容や文構造を考慮した実験を行い、より詳細な解析を行う必要があると考えられる。

5. おわりに

本研究では、音声対話システムにおける、文構造を考慮したオーバーラップを含む応答の印象評価を行い、オーバーラップが行われる文節の位置によって印象が変化することを明らかにした。

今後はオーバーラップが起りやすい発話内容や文構造を考慮した実験を行い、オーバーラップが与える印象をより明確にすることを考えている。また、オーバーラップを含めた応答を行う音声対話システムの実現可能性についても検討を行う。

参考文献

- 1) 小林, 大村, 山本: 音声対話システムにおける適切な発話タイミング生成に関する考察, ヒューマンインタフェース学会研究報告集, vol.15, No.9 pp.23-28 (2013)
- 2) 大村, 小林, 山本: 音声対話システムとの質問-応答型対話におけるオーバーラップを含む発話タイミングの評価, 平成 26 年度電気関係学会北陸支部連合大会, pp.27 (2014)
- 3) 藤原, 伊藤, 荒木: タスク指向対話における相互の対話意図を考慮した対話リズムの分析, 言語・音声理解と対話処理研究会, SIG-SLUD-A701, pp.45-50 (2007)