

J-040

子供を対象とする音声対話インタフェースの検討

A Study on the Spoken Dialogue Interface for Children

田島 大嗣†
Hirotsugu Tabata

樽松 明†
Akira kurematsu

白井 克彦†
Katsuhiko Shirai

1. はじめに

子供における音声認識の問題点として、不安定な会話の話し方のペース・不明瞭な発声といった「音響的問題」や不特定な言葉の使い方・多岐にわたる構文の乱れといった「言語的問題」、さらには、ストレスに対する耐性が低い・集中力が散漫などといった、子供が難しいシステムを利用する際に表れる「インタフェース的問題」がある[1][2]。子供を対象としたシステムを構築する際は、これらの問題を解決する必要がある。

本論文では、コンピュータによる音声対話システムには不慣れな子供に対し、如何にして発話音声进行を認識し、より使いやすくより楽しくシステムを利用してもらうかを解決するインタフェースを提案している。具体的には「象はどういう声でなくの？」などという質問を発声すると、これをシステムが認識して合成音声による応答で該当する動物の鳴き声を出力する「鳴き声応答システム」の構築を行った。

2. システムの概要

子供特有の様々な問題に対し、認識精度を上げることも重要ではあるが、子供を対象にするのであれば、前提としてまずは面白くなくては行けない。従って、「インタフェース的問題」を解決することが先決であると考えた。本システムでは飽きさせず面白さを引き出すという観点からインタフェースに重点を置くことにし、使用していても楽しく使いやすい音声対話システムの構築を目指した。例としてシステムの質問開始画面を図1に示す。



図1. 質問開始画面

本システムは、「表示部」・「制御部」・「認識部」で構成されている。構成を図2に、それぞれの概要を以下に示す。

- 表示部：ユーザからのコマンド入力や制御部からの認識結果を受けてブラウザに表示を行う。
- 制御部：ユーザからの音声を音声信号に変え認識部に送る。さらに、認識部から得た認識結果を基に適切な応答内容を解析し表示内容を表示部に送る
- 認識部：制御部から得た音声信号を基に認識を行い、結果を制御部に戻す。

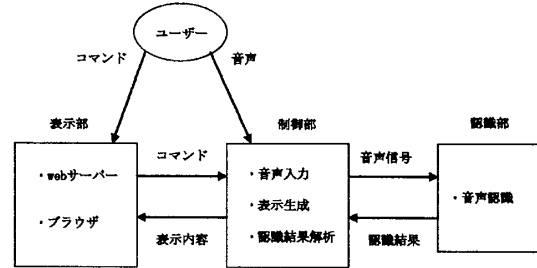


図2. システムの構成

3. 予備実験

3.1 予備実験1

本システムを子供が利用する上での現状の問題点や完成度をタスク達成率の点から探るために、小学生4名を対象に予備実験を行った。この実験において、構文モデルの修正を行った。実験結果を表1に示す。

表1. 構文モデルの変更前後における比較

	タスク達成回数 (全108回)	タスク達成率
修正前	55	50.9
修正後	83	76.9

さらに、観察をしていた上でわかったこととして、音声認識画面で発話終了後から認識結果出力画面に移るまでの間のロード時間が長かったため、システムが正常に認識処理を行っているかわからないといった不安を子供に与えていたことがわかった。従って、発話の録音終了後、認識結果解析中画面を出力することにより不安を軽減させた。

3.2 予備実験2

予備実験1では、ユーザがいつ質問を始めたかわからないといった、発話タイミングの問題があったため、待機画面を出力することにした。

この際、音声案内の必要性を調べるため、成人7名を対象に実験を行った。実験結果を表2に示す。

表2. 音声案内の有無による比較

発話パターン	タスク達成率 (%)
音声アナウンスなし	81.4
音声アナウンスあり	86.4

この実験により、本システムでは、音声認識画面の前に待機画面を挿入し、各画面において音声アナウンスを加える方式を採用することとした。

4. システム運用実験

予備実験の結果を踏まえてシステムを改善した後、現状の課題を利用者の視点から明らかにすること、および本システムが如何に子供を対象としたシステムとして考慮できているかの評価を目的とし運用実験を行った。場所は動物

園の中の室内、被験者は小学生43名である。実験の様子を図3に示す。



図3. 実験の様子

4.1 実験結果

実験結果を図4, 図5, 表3, 表4に示す。

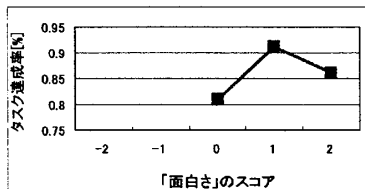


図4. 「面白さ」と「タスク達成率」の関係

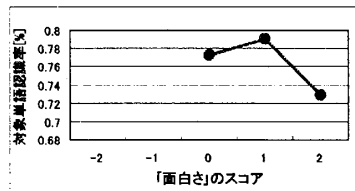


図5. 「面白さ」と「対象単語認識率」の関係

表3. 「面白さ」と「簡単さ」の関係

	とてもむずかしい	むずかしい	ふつう	かんたん	とてもかんたん
とてもつまらない	-	-	-	-	-
つまらない	-	-	-	-	-
ふつう	-	-	1	1	2
おもしろい	-	1	4	10	3
とてもおもしろい	-	1	-	11	9

表4. 「面白さ」と「年齢」の関係

	6	7	8	9	10	11	12
とてもつまらない	-	-	-	-	-	-	-
つまらない	-	-	-	-	-	-	-
ふつう	1	-	1	-	1	1	-
おもしろい	2	2	4	4	1	4	-
とてもおもしろい	-	9	1	4	3	2	2

4.2 考察

アンケート結果において、「おもしろい」・「とてもおもしろい」が多く、すべて「ふつう」以上であったことから、システムにおける「面白さ」は実現できたと言える。

タスク達成率は被験者43名において87.8%と高く、システムとしては充分であった。ここで、「面白さ」と「タスク達成率」の関係を表した図4を見ると、「面白さ」によらず、タスク達成率は80%以上の高い評価を得ており、両者に相関はないことが伺えた。これらのことから、タスク達成率は「面白さ」の直接的要因にはならないことがわかった。

対象単語認識率は被験者43名において、子供を対象と

したシステムにおいては75.8%と高かった。ここで、「面白さ」と「対象単語認識率」の関係を表した図5を見ると、「面白さ」によらず、対象単語認識率は75%以上の高い評価を得ており、両者に相関はないことが伺えた。従って、対象単語認識率においても、「面白さ」の直接的要因にはならないことがわかった。

また、「面白さ」と「年齢」の関係は、表4に示したとおり、各年齢において「とてもつまらない」・「つまらない」の評価はおらず、人数の最も多いことを示す網掛け部分に「おもしろい」・「とてもおもしろい」の評価が集まっている。従って、子供であれば年齢を問わず本システムが有効であることがわかった。

5. まとめ

本論文では、飽きさせず面白さを引き出すという観点からインタフェースに重点を置くことで、子供にとってより使いやすくより楽しい音声対話システムの構築を試みた。システムの内容は、動物の鳴き声応答を題材とした音声対話システムである。

予備実験の結果に基づいて、以下のようなインタフェースの工夫を行った。

- ・動物画像の配置による、見た目の面白さ
- ・音声入力と単純な操作による、シンプルな操作性
- ・発話タイミング明確化による、ユーザビリティの向上
- ・音声案内による、ユーザビリティの向上

これらの工夫を行ったシステムを用いて、現実的な環境に近い上野動物園でシステム運用実験を行った。結果、子供であれば年齢を問わず本システムは「使いやすさ」という面において有効であることがわかり、全体として「面白い」という評価を得た。しかし、すべてのユーザが最高評価をつけたわけではなく、今後更に向上させていく余地があることがわかった。

また、システム運用実験ではユーザから動物の鳴き声以外の情報も聞きたいといった意見があった。従って、将来的には「面白さ」の評価を向上させる工夫として、動物の鳴き声のみではなく種類・特徴や動物園の情報などにも対応させることが考えられる。これについては、現在のように発話文中から動物名のみを抽出して認識を行うのではなく、認識部で発話文中におけるコンテンツの種類もキーワードとして利用できるように改善することで実現させる。

謝辞

本研究の一部は、早稲田大学理工学研究所の研究課題「自発的コミュニケーション機構を有するマルチモーダルヒューマンインタフェースの研究」、平成19年度科学研究費基盤研究(B)課題番号173000666によるものである。

また、恩賜上野動物園と共同研究の関係にあったことでスムーズに研究や実験を行うことができた。恩賜上野動物園の関係者に深く感謝する。

参考文献

- [1] 西村竜一, 西原洋平, 鶴身玲典, 李晃伸, 猿渡洋, 鹿野清宏, “音声対話エージェントによる生駒市コミュニティセンターの案内システム,” 情報処理学会, 第65回全国大会.
- [2] Shrikanth Narayanan, Senior Member, IEEE, and Alexandros Potamianos, Member, IEEE. “Creating Conversational Interfaces for Children,” IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, vol.10, no.2, Feb. 2002.