

合成音声の公共施設放送利用に関する検討

(指導教員 世木 秀明 准教授)

世木研究室 0731009 池田 朗

1 はじめに

合成音声を公共施設放送に使用する場合、アナウンサーを雇う必要がなくなるが明瞭性、了解性が問題となる。ここで、音声合成パラメータを検討することで合成音声の明瞭性、了解性の改善が可能であれば経済的に有利であり、緊急時の対応も容易であると考えられる。

そこで本研究では、合成音声ソフトで話速やポーズ長などの合成パラメータを変化させた合成音を作成し、明瞭性、了解性について調査し、合成音声を公共施設のアナウンスに使用できるかについて検討することを目的とした。

2 音声合成システム

本研究で使用した合成音声プログラムは、市販されているプログラムの一つで高品質な音声合成が可能とされている音素片編集型規則合成プログラム VoiceText である。本プログラムは合成音声の話速やポーズ長を任意に設定することが可能である。

3 聴取実験

3.1 実験用刺激

刺激材料として第1文と第2文間に意味的關係がある2文からなる文章を15文章作成し、これをVoiceTextで読み上げた音声を用意した。さらに、第1文と第2文の内容を問う簡単な質問を作成した。

実験用刺激は、刺激材料にマルチトーカー・ノイズを重畳させたものと重畳させないものとした。ここで、ノイズ重畳レベルは、音声とノイズのラウドネスバランス状態からノイズレベルを3dB減じたものとした。

3.2 実験方法

実験は静かな部屋で行い、被験者前方 150 cm、50 cm にそれぞれスピーカとモニタを置き、聴覚刺激、問題文を提示した。被験者には、音声提示後にモニタに表示される質問と聞き取りにくさに関する調査に回答させた。被験者は、健康な聴力を持つ20代男女40名とし、雑音を重畳させていない音声の実験に20名、雑音を重畳させている音声の実験に20名とした。

4 実験結果と考察

雑音のない環境では図1、図2で示すように文章全体の話速を遅くすると正答率が下降する傾向が見られ、ポーズ長を長くすると正答率が上昇する傾向が見られた。雑音のある環境では話速を遅くしても、ポーズ長を長くしても正答率が上昇する傾向が見られた。

また、文章全体の話速を遅くするほど聞き取りにく

いと答える被験者の割合が増える傾向が見られたが、ポーズ長を長くしても雑音環境下では聞き取りにくいと答えた被験者の割合に変化は見られなかった。

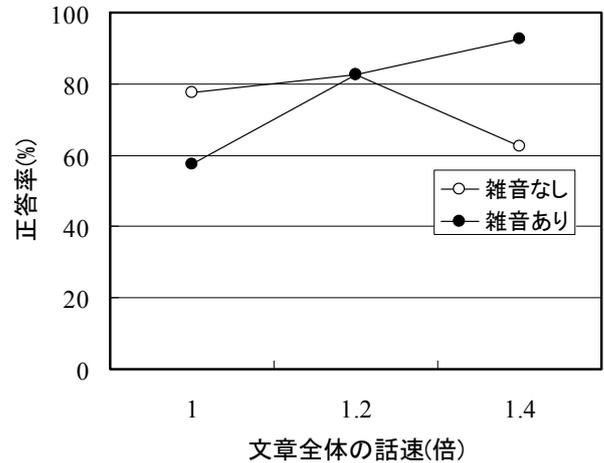


図1 文章全体の話速を変えた時の正答率の変化

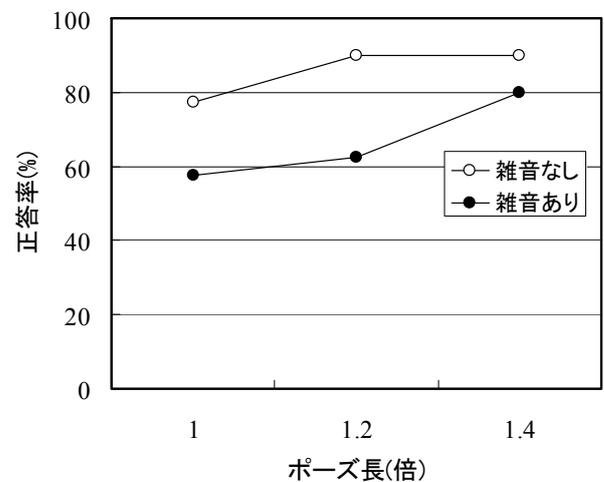


図2 音声のポーズ長を変えた時の正答率の変化

5 まとめ

実験結果より、雑音環境下では話速を遅くしたり、ポーズ長を長くすると正答率が上昇する傾向が見られた。また、話速を遅くすると音声聞き取りにくいと答える被験者が増加する傾向が見られたことから、合成音声において聞き取りやすく、了解性のある音声を作成するにはポーズ長を長くすることが有効であると考えられた。

これは、私たちがポーズ時間中に直前に聞いた情報の理解処理と、次の情報を聞くための準備を行っているという考え方を支持するものであると考えられた。