

研究種目：基盤研究（B）
研究期間：2006～2009
課題番号：18300063
研究課題名（和文） 出力単語の語彙属性を用いた対話韻律制御に基づく音声合成
研究課題名（英文） Communicative speech synthesis using impression attributes of output lexicons
研究代表者
匂坂 芳典(SAGISAKA, Yoshinori)
早稲田大学・国際情報通信研究科・教授
研究者番号：70339737

研究分野：総合領域 情報学
科研費の分科・細目：知覚情報処理・知能ロボティクス 1006A
キーワード：音声合成、韻律制御、対話音声、イントネーション、基本周波数、パラ言語情報

1. 研究計画の概要

本研究では、発話毎の多様性を実現した、自然性の高い対話音声の合成出力を目的とし、対話音声独自の韻律変化とその要因を解明する。また、対話音声合成を可能とする入力情報の特定、抽出を行う。このため、基本周波数制御生成モデルを中心とした韻律制御モデルを導出する。また、実用音声対話システム、他言語への適応を考えた技術の一般化を追求する。

2. 研究の進捗状況

実際の対話音声にみられる多量の一語発話「ん」を分析し、その発話印象は3次元（確信・疑念、肯定・否定、好印象・悪印象）によって近似的に記述でき、韻律制御との対応がとれることが判明した。「確信・疑念」「肯定・否定」に対応する印象はF0の時間変化形状を現す4種のダイナミックパターン（上昇、平坦、下降、上昇後下降）とよく呼応し「好印象・悪印象」に対応する印象は基本周波数F0の高さに反映される。

これらの印象に対応する典型的な語彙からなる対話音声の分析を行い、その語彙自体が有する印象情報と韻律との対応関係を分析した。この結果、「ん」と類似のF0制御特性を確認できた。この実験事実に基づき、語彙が有する印象情報を利用した韻律生成法を新たに提案した。これにより、これまで手付かずだった対話音声合成に新たな方法論を提供することができた。

提案した対話韻律生成法の有効性を評価するため、STRAGHT合成法を用いて典型的な印象を持つ語彙からなる句の対話音声を作成し、その自然性評価試験により提案法の

有効性を確認した。

また、発話印象とF0変化形状パタンの対応関係の言語普遍性を調べため、言語間に渡る対話韻律生成実験を行った。作成した中国語対話合成音声の自然性評価実験の結果、先に提案した印象ベクトル表現に基づく韻律生成法が言語共通に使用できる可能性を確認できた。この実験を通じて発話印象と対話韻律との関係は、二言語に共通であることが判明した。

さらに、より一般的な対話合成への適応を図るため、副詞、形容詞、終助詞からなる句を用い、出力語彙に対応する心理印象量の組み合わせと対話韻律の関係について分析を進めた。この結果、各対話表現の韻律は、表現を構成する各語彙が有する印象属性が与える韻律特性の重ね合わせとして表出されることが確認でき、出力語彙の印象情報の利用可能性が広い語彙の組み合わせに対して確認できた。

3. 現在までの達成度

おおむね順調に進展している。

研究提案時に目標とした、対話音声独自の韻律変化とその要因を解明し、基本周波数制御モデルを導出した。さらに、他言語への適用可能性も確認でき、順調に推移している。

4. 今後の研究の推進方策

対話韻律生成の可能性拡大を図るため、複数の語彙が有する印象情報から韻律制御を行う方法について検討する。また、英語他さらに適応対象を追加し、言語普遍性を追及する。

5. 代表的な研究成果

〔雑誌論文〕(計 1 件)

1) Y. SAGISAKA "Towards Computing Phonetics" 中国語音学報 査読有 Vol. 1 pp. 23-37 2008

〔学会発表〕(計 12 件)

1) Y. Sagisaka, "Corpus-based speech synthesis from reading speech to communicative speech" ISCA Workshop on Spoken Language Technologies for Under-resourced languages, 6 May 2008, Hanoi

2) Y. Sagisaka, Y. Greenberg, K. Li, M. Zhu, M. Tsuzaki and H. Kato "Communicative prosody processing for synthesis and recognition of para-linguistic information" International Conference on Computer Applications Feb. 2008 Yangon

3) Mingzhao Zhu, Ke Li, Y. Greenberg and Y. Sagisaka "Automatic extraction of paralinguistic information from communicative speech" Proc. the 7th Symposium on Natural Language Processing 2007 pp.207-212 Dec.2007 Pattaya

4) K. Li, Y. Greenberg and Y. Sagisaka "Inter-language prosodic style modification experiment using word impression vector for communicative speech generation" Proc. Interspeech 2007 pp.1294-1297 Aug. 2007 Antwerp

5) Y. Greenberg, N. Shibuya, M. Tsuzaki, H. Kato, Y. Sagisaka "A trial of communicative prosody generation based on control characteristic of one word utterance observed in real conversational speech" Proc. Speech prosody pp.37-40 May 2006 Dresden

〔図書〕(計 2 件)

1) K. Li, Y. Greenberg, N. Shibuya, N. Campbell, Y. Sagisaka IOS Press "On the analysis of F0 control characteristics of nonverbal utterances and its application to communicative prosody generation" in NATO Security through Science Series E: Human and Societal Dynamics Vol.8 The Fundamentals of Verbal and Non-verbal Communication and the Biometric Issue

pp.179-183 2007

2) 匂坂、グリーンバーグ、山下 くるしお出版
「文法と音声」所載「語彙情報を用いた会話韻律生成について」 pp.135-145 2006

樣式 C-7-2

自己評価報告書