

TR-I-0143

大規模音声データベースに基づく音声合成

Automatic Rule Derivation for Speech Synthesis
from Large-scaled Speech Database

武田 一哉
Kazuya TAKEDA

1990.2

内容更概

データと規則の分離という観点から、新しい規則合成システムの構築を検討した。即ち、大規模な音声データベースの構築を行い、種々の発声の変動を統計的にとらえてそのモデル化を行う。さらに、合成単位に関する種々の問題を、最適な合成素片の選択という統一的な見地からとらえ、構築したデータベースから最適な合成素片を得る方法を検討する。

1 はじめに

規則による音声合成システムの開発においては、多くの精力が種々の規則の開発に注がれてきた。合成単位の作成、音韻変形規則の作成等の試行錯誤によるシステム開発には、膨大なデータと時間が必要である。これは、数多くの実用システムが開発され、いくつかの優れたシステムが構築されている現在でも、依然として残る問題である。

このようなシステム開発の非効率の原因に、これまでの合成方式が用いている知識(規則)が、合成方式から独立した一般的なものでなかったことがあげられる。予備実験により、高い品質の合成音声を得られながら、実際に任意語彙の合成を行った結果、十分な品質が得られないといったことは、知識の一般化が不足しているからに、ほかならない。

そこで、このような規則合成システム開発の問題を解決するための検討を、以下の2点の特徴を持つ合成システムの構成を通じて行った。

- (1) 大規模な音声データベース中に含まれる発声のバリエーションの出現を、統計的なモデルを用いて、抽出する。
- (2) 単位接続型の合成システムを前提とし、最適な単位の選択を通じて、音声単位の決定、単位の作成といった問題を扱い、単位に関する規則と単位抽出用データベースとを分離する。

以下本稿では、これらの研究の結果を概説する。

2 大規模音声データベースの構築

音声合成のための知識抽出の基盤として、大規模な音声データベースを構築した。このデータベースの特徴は、

- (1) 特定話者の大規模(8500語相当)な発声を収録していること
 - (2) 視察に基づく、高精度な音韻ラベルが付与されていること。
 - (3) データベースの管理・利用のための管理システムが構築されていること
- である。本データベースの構築作業は現在も継続して進められており、不特定話者、連続音声データベースに発展している。

3 特殊モーラ規則の抽出

単語音声データベースから、母音の無声化規則を抽出することを試みた。データベース中の母音無声化サンプルと、それらの音韻環境とを調べた結果、従来指摘されてきた関係とは異なる傾向が発見され、単語発声特有の無声化生起規則の存在が示唆された。

大規模な発声中に含まれる無声母音生起の揺らぎを、周囲の音韻環境などのカテゴリカルな要因を用いて予測するためのモデルとして、数量理論II類を用いた。分析の結果、後続音節の種類、アクセント核の位置等が、母音無声化に支配的な要因であることが確認された。分析を通じて得られた予測式により、80%の精度で母音の無声化が予測可能であることが、明らかになった。

4 韻律規則の抽出

前節同様数量化の手法を用いて、音韻の継続長と音韻環境等の関係を分析しモデル化した。分析の結果、音韻継続長に支配的な要因として先行音韻の種別が指摘され、これまで指摘されてきたC-V間での音韻長の補償現象、即ち日本語の持つ mora-timed な特性が確認された。さらに、分析を通じて得られたモデルは、単語発声中の音韻継続長を予測するモデルとして優れた性能を示した。

この分析により得られたモデルを文発声に適用し、誤差の解析を行った結果、以下の文レベルでの音韻継続長の変動が明らかになった。

- (1) 呼気段落末における音韻長の伸長は、単語末におけるそれよりも大きいこと。
- (2) いわゆる文節末で、継続長が若干短縮されること。
- (3) 朗読発声では、文末において大幅な音韻長の短縮が見られたにもかかわらず、会話発声ではその傾向が認められなかったこと。

これらの知見に基づき、単語レベルでの予測モデルに文内位置での補正を加えることにより、文発声に対する音韻長の良好な予測モデルが得られると同時に、会話発声に関しての分析を重ねる必要があることが明らかになった。

5 種々の音韻接続を選択的に用いる規則合成システムの構築

音声データベースから任意の音声素片を選択し、これを合成単位として用いる規則合成システムの構築を行った。このシステムの構築には、音声データベースに含まれる発声の特徴を、選択時に十分把握する必要があるため、素片検索辞書の構成に際し、音韻ラベルのみでなく、音声事象、異音化ラベル等の情報を用いることが、重要であることを明らかにした。

6 合成単位選択アルゴリズムの検討

前節で構築したシステムにおける、単位選択上の以下に示す問題を明らかにした。

- (1) 種々の選択基準を統合した尺度を用いているため、各々の選択基準の適切な評価が困難である。
- (2) 単位系列の適性を適性値の総和で評価しているため、局所的に大きく適合度が劣化する場合がある。

これらの問題を解決するため、トップダウンに単位候補を絞り込み、得られた候補を種々の観点から(時には物理パラメータを用いる)、評価し単位選択を行う方法を検討した。この方法を導入した結果、従来に比べ主観評価で65%(プリファランススコア)、単語明瞭度で、99%(従来法で、95%)と高い品質の合成音を得られた。

7 合成単位の抽出規則の検討

前節で構築したシステムに基づき、単位接合型の規則合成における、単位の使用基準を考察した。即ち、選択的な単位使用により得られた合成音声の品質を、音節了解度で評価し、これを用いられた単位音声素片の音韻的な属性に関係づけた。分析により得られた知見は、以下の通りである。

- (1) 合成単位の長さが長くなるに従い、品質が向上する。特に単音節単位と、それより長い単位の間大きな品質の差が認められた。

- (2) 単位の抽出環境と使用環境との差異に着目すると、置換・脱落誤りは、比較的差異が小さな環境でも生じる。さらに差異が大きくなるに従い、付加・挿入誤りが増える。
- (3) 了解誤りは、単音節>単位頭>単位末>単位中の順でおこりやすく、抽出環境と使用環境の差異は、左側の環境をより重視して評価すべきである。
- (4) 局所的であっても、大きく環境の異なる単位が使われた場合、以降の単位にも影響が及ぶ場合がある。
- (5) 接続音韻環境に着目すると、誤りは、母音/無声摩擦<母音定常部<母音/無声破裂<母音/有声破裂<その他、の順で低く、この順が単位接続点としての適性を反映していると考えられる。

8 まとめ

規則合成において、データから独立したより一般性の高い規則を得るための枠組みを検討した。特に、母音無声化規則、音韻長制御規則、合成単位選択規則等を確立し、十分な規模のデータベースが与えられれば、これを用いて音声合成が可能となることを示した。

発表文献リスト

以下に本研究に関して行った発表文献を示す。

- (1) 音声データベース構築のための音韻ラベリング
武田、匂坂、片桐
音響学会全国大会, 昭62年3月, 2-5-10
- (2) 音声データベース管理システムの構築
武田、森川、桑原
情報処理学会全国大会, 昭62年後期, 2H-4
- (3) 音韻ラベルを持つ日本語音声データベースの構築
武田、匂坂、片桐、桑原
音声研究会資料SP87-19, (1987-6)
- (4) 研究用日本語音声データベースの構築
武田、匂坂、片桐、桑原
音響学会誌44(10), pp.747-754, (1988-10)
- (5) 音声データベース管理システムの構築
武田、桑原、森川
音声研究会資料sp87-116, (1988-1)
- (6) Acoustic-phonetic Labels in a Japanese Speech Database
Takeda, Sagisaka, Katagiri
Proc. Euro.Conf. Speech Tech., Vol.2, pp.13-16, (1987-9)
- (7) 母音無声化の要因分析と予測手法の検討
武田、桑原
音響学会全国大会, 昭62年10月, 3-3-8
- (8) 文発声と単語発声における音韻継続長の分析
武田、匂坂、桑原
音響学会全国大会, 昭63年3月, 2-1-6
- (9) On sentence-level factors governing segmental duration in Japanese
Takeda, Sagisaka and Kuwabara
J. Acoust. Soc.Am., 86(6), pp.2081-2087(1989-12)
- (10) 入力音韻系列に応じた音声合成素片選択法の改良
武田、安部、匂坂、桑原
音響学会全国大会, 昭63年10月, 1-2-19
- (11) Adaptive Manipulation of Non-uniform Synthesis Units using Multi-level Unit Transcription
Takeda, Abe, Sagisaka and Kuwabara

- (12) 合成単位の選択における音韻的知識と韻律的属性の処理の統合
武田、匂坂、桑原
音響学会全国大会, 平1年10月, 3-P-16
- (13) エキスパートシステムを用いた単位選択の検討
武田、安部、匂坂
音声研究会資料sp89-113, (1990-1)
- (14) 合成単位の音響的属性と合成音声の品質との関係
武田、安部、匂坂
音響学会全国大会, 平2年3月, 1-4-14