

# 自己修復の分析に基づく発話生成モデルに関する考察

2K-3

伊藤雅浩†

佐川雄二†

大西 昇†

杉江 昇‡

†名古屋大学工学部情報工学科

‡名城大学理工学部電気電子工学科

## 1 はじめに

計算機との自然な対話を実現するためには、自由発話理解、実環境への対応、対話のモデル化などが重要な課題である [1]。しかし、話し言葉には省略、倒置、言い直しなど、通常書き言葉にはあまり見られない表現が数多く現れる。

なかでも、音声入力によって書き言葉と異り、誤りなどを含む不適格文が多く入力されることが予想される [2]。Levelt [3] によると、こうした不適格性は話者自身によって直されることが多い。この話者自身による言い直し行為を自己修復 (self-repair) と呼ぶ。

自己修復は発話行動のさまざまな過程で生ずるため、さまざまなものがある。例えば、「秋葉原」と言おうとして「あきはらば」と言うような音素レベルでの誤りを訂正するものから、発話の途中で別の話題に転換してしまうような自己修復もある。

発話を生成する過程というのは大きく分けて3つのステージに分けられる [3, 4]。一つ目で意図 (概念) を構成し、二つ目のステージでは文法構造を構築し、最後のステージで音素が決定され、発話が行なわれる。最近では各ステージの行為が並列的に行なわれていることを示す研究もある [5]。また、Blackmer ら [6] は自己修復に要する時間の分析から、話者は各々のステップでの出力を逐次モニターしているということを報告している。

そこで我々は自己修復を誤りの種類と、修復に要する時間の関係から分析する。得られた結果から、実際の現象に適合する発話生成モデルについて検討する。

## 2 自己修復の分類

本研究で実際に分析を行なった自己修復を、誤りの発生段階に応じて3つのクラスに分類した。1つは概念レベルの修復で意図の変更、および内容の付加 (精緻化) などである (例1参照)。2つ目は文法レベルの修復で単語の選択誤りや、活用語尾の訂正などである (例2参照)。最後は音素レベルの修復で言い間違い (単語そのものは意図と合致している) などである (例3参照)。

例1 あの、わかる、こっちは行きたくないんだろう  
なつてわかるんですけど

例2 そんなともだ、あの、生徒が自分のクラスの中に

例3 とってもいいけんけー、経験をしてるんだな

本研究で分析したコーパス (詳細は3章参照) 中の89例の自己修復の内訳は、概念レベルが35(39.3%)、文法レベルが36(40.4%)、音素レベルが18(20.3%)であった。

また、自己修復文の構造は概ね図1のような形をしている [2]。

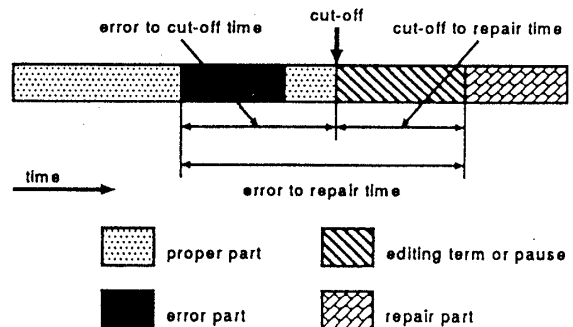


図1: 自己修復文の構造

editing term or pause (冗長語およびポーズ) の時間  
で言い直し部分の準備をし、同時に聞き手にこれから  
自己修復が始まることを知らせる。しかし、error to  
cut-off time の間で言い直しの準備を行なう場合も想

A study on a model of utterance generation based on an analysis of self-repairs  
Masahiro Ito, Yuji Sagawa, Noboru Ohnishi  
Nagoya University  
Furo-cho, Chikusa-ku, Nagoya 464-01, Japan  
Noboru Sugie  
Meijo University  
1-501 Shiogamaguchi, Tenpaku-ku, Nagoya 468, Japan

定され、その場合は誤りの直後に修復が始まることも考えられる。

### 3 コーパスの分析

本研究ではNHKのラジオ番組から録音した音声を生声編集ツールの“xsed”に入力して解析を行なった。対話数は3つで、合計時間は4.4分、内容は教育に関する電話相談である。

そこで、各自己修復について error to cut-off time と cut-off to repair time および、error to repair time を計測した。図2に各々の自己修復のカテゴリーごとに計算した平均と標準偏差を示す。

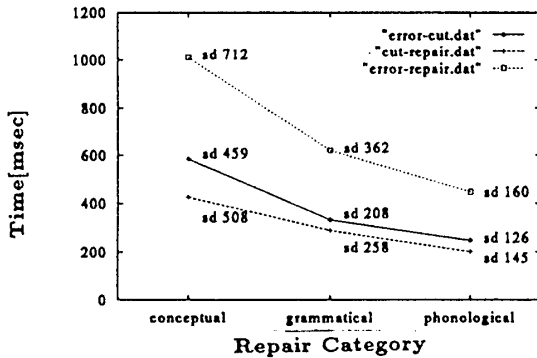


図2: 測定結果 (sd は標準偏差)

Blackmer ら [6] の実験と同様、今回も中断から修復までの時間が 0 ms のものが 13 例 (14.6%) 見られた。このことは各モジュールがパイプラインとして作動し、それぞれの出力が逐次モニターされていることを示している。

また、修復に要する時間は発話生成段階の初期のモジュールで生じた誤りほど長くかかっており、標準偏差も大きい。そこで、誤りが発話生成の後半のモニターで検知されるほど検知に要する時間や修復に要する時間が長くなると考えられる。

その結果、図3のようなモデルを提案する。各モジュールの動作は並行して行なわれ、不適格性はモニターによって検知される。このモデルによれば、概念レベルの誤りが最初のモニターで検知されることもあれば、phonological encoder に付随するモニターで検知され得ようになり、実際の現象とも適合する。

### 4 おわりに

本稿では自己修復における時間的要素の分析結果を示した。結果は、生成の初期の段階で生じた誤りに対

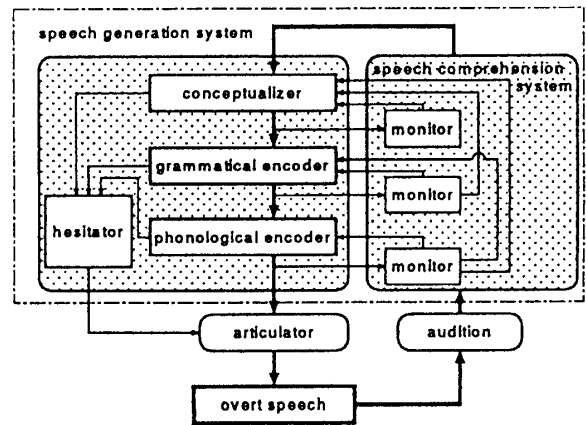


図3: 発話生成モデル

する修復ほど時間がかかっている。そこでモニター処理の部分に注目したモデルを提案した。時間的要素だけでは誤り生成の様子を完全に分析することはできなかったが、こうした分析は自然な発話を扱う対話システムを考える上にも必要であると思われる。

### 参考文献

- [1] 金沢博史, 瀬戸重宣, 新地秀明, 竹林洋一: 音声自由対話システム TOSBURG II におけるデータ収集と評価環境, 信学技報, SP93-114, 1993.
- [2] 佐川雄二, 大西昇, 杉江昇: 自己修復を含む日本語不適格文の分析とその計算機による理解手法に関する考察, 情報処理学会論文誌, Vol.35, No.1 pp.46-52, 1994.
- [3] Levelt, W.: *Speaking: from intention to articulation*, MIT Press, 1989.
- [4] Sagawa, Y., Ito, M., Ohnishi, N. and Sugie, N.: A Model For Generating Self-Repairs, In *Proceedings of International Conference on Spoken Language Processing (ICSLP 94)*, Yokohama, 1994 (to appear).
- [5] Van Wijk, C. and Kempen, G.: A Dual System for Producing Self-Repairs in Spontaneous Speech: Evidence from Experimentally Elicited Corrections, *Cognitive Psychology* 19, pp.403-440, 1987.
- [6] E.R.Blackmer and J.L.Mitton: Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition* 39, pp.173-194, 1991.