ISCA Archive
http://www.isca-speech.org/archive

3rd International Conference on
Spoken Language Processing (ICSLP 94)
Yokohama, Japan
September 18-22, 1994

A MODEL FOR GENERATING SELF-REPAIRS

Yuji Sagawa *Masahiro Ito* *Noboru Ohnishi* *Noboru Sugie*

Dept. of Information Engineering, Nagoya University
Furo-cho, Chikusa-ku, Nagoya, 464-01 Japan

ABSTRACT

Self-repairs give us a key to understand how a human generates utterances. In this paper, we report a result of our analysis of timing factor of self-repair. We categorize an error repaired by a stage it is generated, and investigate relation between a category and timing factor of the repair. The result shows that it takes more time to repair an error generated at earlier stage of generation. From the result, we develop a model of utterance generation. Finally, we try to account for some characteristics of self-repair with the model.

1 INTRODUCTION

Self-repairs give us a key to understand how a human generates utterances, because they show us how s/he plans what to say and how to say it, what kind of error occurs and how s/he detects and correct it.

In this paper, we report a result of our analysis of timing factor of self-repairs. We analyze when a speaker makes an error, when s/he detects it and interrupt his/her utterance and when s/he starts a repair are investigated, using a record of spontaneous dialogues taken from radio programs. We measure error-to-cut-off time, cut-off-to-repair time and error-to-repair time.

A process of generating an utterance is divided into three stages roughly, generating an intended meaning, formulating a grammatical structure (including lexical choice), and articulating. Recent studies on speech production support an idea that these steps run in pipeline fashion (processing a message begins before the previous message has not fully uttered)[1].

Blackmer and Mitton[2] report the result of their analysis of timing of self-repair that supports pipeline generation and propose a model for generating utterances. In their model, a speaker monitors outputs of each step, and if an error is detected, s/he starts a self-repair. Our analysis also support this idea.

Moreover, we investigate relation between a category of an error repaired and timing of the repair. We categorize an error by a stage it occurs in, i.e. conceptualizing errors, formulating errors, articulating errors. The result shows that it takes more time to repair an error generated at earlier stage of generation. To account for this fact, we modify Blackmer and Mitton's model.

In our model, each stage has a monitor and the monitor monitors output of the stage as well as Blackmer and Mitton's model. But our monitor can detect an error which occurs in an earlier stage, that is, monitor of articulator can detect a formulating error or a conceptualizing error as well as an articulating error.

Finally, we try to account for some characteristics of self-repair with the model.

2 SELF-REPAIR AND UTTERANCE GENERATION MODEL

When a speaker detects an error or a less felicitous expression, s/he interrupts his/her utterance and repair it.

One question about self-repairs is how a speaker detects them. Levelt[3] points out that a speaker monitors his/her inner speech as well as his/her overt speech. Blackmer and Mitton[2] found that there exist repairs made 0 ms after interruption, and conclude that the fact supports Kempen and Hoenkamp's concept of incremental processing[1] and inner speech is monitored at each stage of utterance generation.

Recently Carletta *et.al*[4] propose a generation model that accounts for hesitations, filled pauses and so on as well as self-repairs.

Based on these two models we propose a preliminary version of generation model. Figure 1 shows the model.

The model contains a speech comprehension system because monitoring needs the system. A speech generation system consists of a conceptualizer, a grammatical encoder, a phonological encoder and a hesitator.

A conceptualizer generates "what to say" and

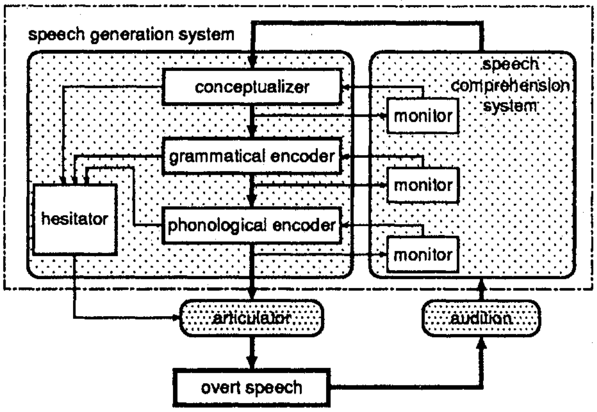

Figure 1: A Utterance Generation Model (preliminary version)

passes it to the next stage, grammatical encoder. A grammatical encoder access lemmas and build a syntactic structure of the message. This process includes lexical selection, so an output of this stage is an ordered sequence of words and it is passed to phonological encoder.

A phonological encoder accesses phonetic plan of each word and build a phonetic plan of the utterance. An articulator accepts the phonetic plan and executes an articulation program by the musculature of the laryngeal.

As well as an overt speech, an output of each stage is monitored by each monitor. If a monitor detects something troublesome, it reports to the stage it monitors and repair starts. We will discuss process of interruption and repair in the next chapter in detail.

If some generating process will take much time than usual, a hesitator is invoked and an editing term such as "anoh" or "eh" is generated to maintain the right to speak.

3 TIMING OF SELF-REPAIR

Figure 2 shows a typical structure of a repaired utterance[5]. In some cases, planning of repair starts in error to cut-off time, in other cases, it starts in cut-off to repair time. In the former cases, there can be a repair done immediately after error (see figure 3 (a), (b) is the other case) as Blackmer and Mitton[2] point out.

In our preliminary model, each monitor detects an error that occurs in only one stage, i.e. monitor of grammatical encoder can detect only grammatical error and report it only to grammatical encoder. If so, timing of repair is determined only by time

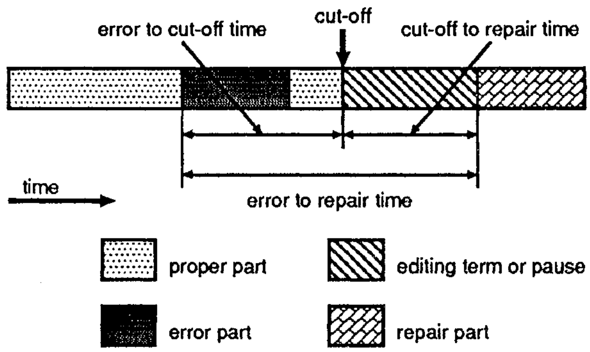

Figure 2: Typical Structure of a Utterance Including a Repair

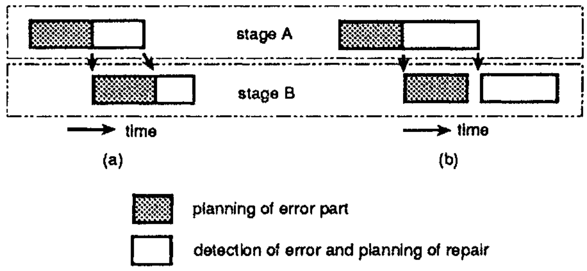

Figure 3: Timing of Repair

for monitoring and planning of repair.

If a monitor can also detect an error that occurs at earlier stages, time for planning of repair may be longer. If a monitor of phonological encoder detects an error that occurs at conceptualizer, the monitor must report it to conceptualizer and planning of a repair must be started from conceptualization.

4 CLASSIFICATION OF SELF-REPAIRS

To solve the problem introduced at the end of the last chapter, we must classify self-repairs by the stage it occurs. Conceptually based repair (CBR) is to repair "what to say", such as elaboration[6] and fresh start. Grammatically based repair (GBR) is to repair syntactic structure or lexical selection. Phonologically based repair (PBR) is to repair such as slip of tongue and so on.

It is difficult to classify all self-repairs into these three classes. In the experiments below, first two authors classify repairs independently. On repairs classified into different class by two authors, we try three analysis, based on classification of one author and the other and classification excluding them. They are not so different as a whole. We show the result based on first classification in the next chapter.

5 EXPERIMENT

5.1 Method

We use a set of recordings of consulting dialogues over telephone broadcasted as radio programs by NHK (Japan Broadcasting Corporation). Table 1 shows the detail of the data.

Table 1: The Detail of the Data

No. of dialogues	3
No. of participants	2 / dialogue
Total time	44 min.
Contents	consultation on education
No. of self-repairs	89

The data recorded with stereo cassette deck is input to Sony NEWS workstation and analyzed with sound editing tool "xsed".

We measure **error to cut-off time(ECtime)**, **cut-off to repair time(ERtime)** and **error-to repair time(ERtime)** for self-repairs (see figure 2).

5.2 Results

There are 89 self-repairs in the data. Table 2 shows the classification of self-repairs.

Table 2: The Classification of Self-Repairs

Repair class	No.
CBR	35 (39.3%)
GBR	36 (40.4%)
PBR	18 (20.3%)

Figure 4 shows the averages and standard deviations of them for each repair class.

While mean time of each time is longer at repairs of earlier stage errors, standard deviation is greater for such repairs. That means shorter repair and longer repair are involved in such repairs.

As Blackmer and Mitton's experiment[2], there exists 13 self-repairs (14.6% [1]) whose cut-off to repair time are 0 ms. This supports their model in the point that generation processes run in pipeline fashion and outputs from each process are monitored.

[1] 12.4% in Blackmer and Mitton's data.

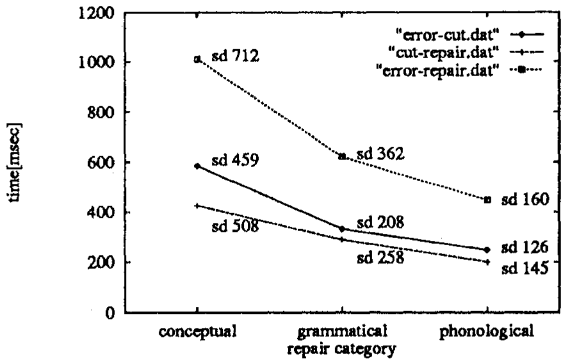

Figure 4: Average of ECtime, CRtime, ERtime (sd is standard deviation)[msec]

6 discussion

The fact that repairs of earlier stage errors involves both long and short repair suggests that such errors can be detected by monitors of another (later) stages as discussed in chapter 3. If an error detected by later monitor, repair of it takes much time. If not, repair takes less time.

To account for the fact, we revise our preliminary model as figure 5. It is different from the preliminary version in the following two points.

- Each monitor can detect earlier stage error
- When a monitor detects errors, it is reported to the stage which generate the error.

Alternative account for longer repair time is that it takes more time to plan repair part in the case of repair of earlier stage error. But this does not account for why it takes longer ECtime for repairs of earlier stage errors. All ECtime is not consumed to detection[2],but longer ECtime cannot be account for without monitoring model above.

The other characteristics of self-repair which can be accounted for with this model is within-word repair of word selection. Within-word repair is a self-repair with within-word interruption as in (1).

(1) kono kensin de ketue, ano, ketunyou ga aru

The speaker first try to say a word "ketueki"(blood) but interrupt within-word ("ketue"), and re-select a word "ketunyou" (bloody urine).

[2] Planning of repair may begin before cut-off. But it is impossible to detect exactly when it begins so far.

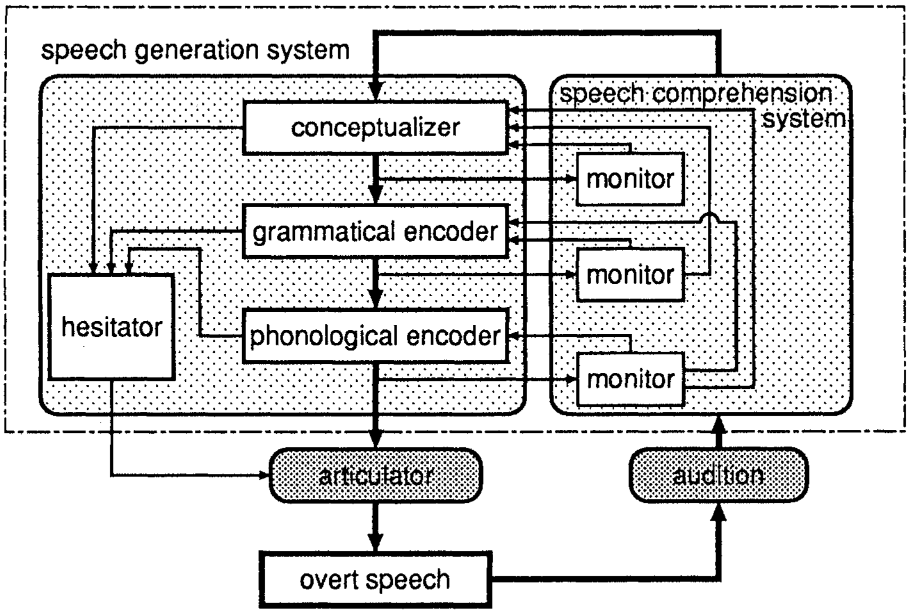

Figure 5: A Model for Generating Speech (Revised Version)

A repair of word selection is conceptually based or grammatically based repair. But within-word interruption can be done only after phonological plan is determined. So, it is more natural to think this error is detected by the phonological monitor and the monitor interrupts utterance within the word and report the error to grammatical encoder (or conceptualizer) to repair.

7 CONCLUSIONS

In this paper, we report a result of our analysis of timing factor of self-repair. The result shows that it takes more time to repair an error generated at earlier stage of generation. From the result, we develop a model of utterance generation, especially a model of monitoring process.

How a listener can understand self-repaired utterance is also an important problem. It can partially be solved with some linguistic cues[7], but some prosodic information seems to be useful. We cannot conclude that we can identify the stages that generates error by length of pause, but analyses like we did are necessary for spontaneous dialogue system.

References

[1] G. Kempen and E. Hoenkamp. An incremental procedural grammar for sentence formulation. *Cognitive Science 11*, pages 201–258, 1987.

[2] E. R. Blackmer and J. L. Mitton. Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition 39*, pages 173–194, 1991.

[3] W. J. M. Levelt. *Speaking: From Intention to Articulation*, chapter 12, pages 458–499. The MIT Press, Cambridge, MA, 1988.

[4] J. Carletta, R. Caley, and S. Israd. A system architecture for simulating time-constrained language production. Research Paper RP-43, Human Communication Research Centre, University of Edinburgh, 1993.

[5] Yuji Sagawa, Noboru Ohnishi, and Noboru Sugie. Repairing self-repairs in Japanese. In *Proceedings of Natural Language Processing Pacific Rim Symposium (NLPRS '93)*, pages 191–198, Fukuoka, 1993.

[6] M. Okada and H. Otsuka. Incremental elaboration in generating spontaneous speech. In *Proceedings of International Symposium on Spoken Dialogue '93*, pages 49–53, 1993.

[7] Yuji Sagawa, Noboru Ohnishi, and Noboru Sugie. A parser coping with self-repaired japanese utterances and large corpus-based evaluation. In *Proceedings of the 15th International Conference on Computational Linguistics, COLING-94*, Kyoto, 1994 (to appear).