CORRECTING CHINESE REPETITION REPAIRS IN SPONTANEOUS SPEECH

Yue-Shi Lee and Hsin-Hsi Chen

Department of Computer Science and Information Engineering

National Taiwan University

Taipei, Taiwan, R.O.C.

e-mail. hh_chen@csie.ntu.edu.tw

Abstract

Disfluencies involving speech repairs pose serious problems for spoken language processing systems. However, which cues in speech signals may facilitate Chinese repair processing is not known. This paper concerns the acoustic and prosodic analysis for correcting Chinese repetition repairs in spontaneous speech. A large spoken corpus is examined in this study. The experimental results show that our method can achieve the precision rate of 93.87% and the recall rate of 90.65%, without using the lexical information.

1 Introduction

Most of the previous acoustic analyses of speech examined data from speakers who carefully pronounce their speech. Natural spontaneous or conversational speech differs from careful or reading speech in several ways. The most obvious difference is the use of speech repairs. In spontaneous speech, people often start talking and then think along the way. This causes spontaneous speech to have a variable speaking rate, and such speech often exhibits speech repairs, which are interrupts in the flow of speech, where the speaker reiterates a portion of the immediately preceding speech, with or without a change.

Heeman and Allen [1] describe that 25% of turns contain at least one speech repair in their corpus. In our study, 17% of turns contain at least one speech repair. Thus, the speech repairs cannot be negligible and have influences to a certain extent. On the one hand, correctly recognizing speech repairs can help automatic speech recognizers to avoid textual errors. In most of the current speech recognition systems, words repeated in a speech repair are simply fed as word hypotheses to the language model of the recognizer. This may cause difficulties in having a proper recognition since the language model is usually trained on fluent text only.

On the other hand, even if all the words in a disfluent segment are correctly recognized, failure to detect a disfluency may lead to interpretation errors during subsequent processing. (1) is an example[1].

(1) 952 ..她應該%-- {A2,2-1,1}
953 ...不應該升工程--- {R1,2-1,2}
954 ..工程師那麼快的.\

There are an addition repair (A) and a repetition repair (R) in (1)[2]. For the first repair, the speaker changes his intention from "應該" to "不應該". If this kind of repair cannot be detected, the system will misunderstand the intention of the speaker.

Recently, text-first approach [1, 4, 5] and speech-first approach [6, 7] have been proposed to touch on repairs in English. The text-first approach assumes the speech

---

1 The transcription system proposed by Bois, *et al.* [2] is used to transcribe the spoken data. The two symbols ... and .. denote a unfilled pause (silence) is medium and short, respectively. The symbol % denotes the glottal stop. The detailed transcribing conventions are shown in Appendix A. Relevant characters in examples are in **boldface** and underline.

2 The annotating scheme of repair can refer to Chou [3].

recognizer could provide a correct transcription. That is, it tries to detect and correct speech repairs automatically using text alone. Hindle [4] adds rules to a deterministic parser to tackle the problem of correcting speech repairs. His parsing strategy depends on the successful disambiguation of the syntactic categories. Although syntactic categories can be determined well by local context, Hindle admits that speech repair disrupts the local context. Bear, *et al.* [5] firstly try to parse the input sentence and then invoke a repair processing when the parsing fails. For repair processing, a simple pattern matcher finds the candidates based on the lexical cues at the first stage. Then the syntactic and semantic processing filters out the impossible candidates. Heeman and Allen [1] present an algorithm that detects and corrects modification and abridged repairs. The algorithm uses some repair patterns to capture potential repairs. These patterns are built based on the identification of word fragments, editing terms[3], and word correspondences between the repaired segment and the repairing segment[4]. The resulting potential repairs are then passed to a statistical filter that judges the proposal as either fluent speech or an actual repair.

The speech-first approach tries to identify speech repairs using acoustic and prosodic cues. Nakatani and Hirschberg [6, 7] investigate the detection of the interruption point of speech repairs based on this line. The cues that they found are the occurrence of a filled pause, the presence of a word fragment, the energy peaks in each word and other features such as accent. However, they did not address the problem of correcting the speech repairs. In other words, they do not determine

---

[3] The editing terms can either be filled pauses (e.g., um, un, er) or cue phrases (e.g., I mean, I guess, well).

[4] A repair is composed of a repaired segment and a repairing segment which immediately follows the repaired segment. A repaired segment denotes the portion of the utterance which is being repaired, and a repairing segment denotes the portion which is accomplishing the repair [8]. That is, the repaired segment is replaced by the repairing segment.

which words are undesired.

These approaches cannot be adopted to deal with Chinese speech repairs for the following reasons. First, a Chinese sentence is composed of a string of characters without any word boundaries. In other words, it is necessary to segment Chinese sentence before tagging and parsing [9, 10]. Repairs make segmentation and text-first approach more difficult. Second, Chinese repairs may not always have an editing terms between a repaired segment and a repairing segment. In other words, editing terms do not have much effect in Chinese repair processing. Third, duplicate constructions (e.g., 幫幫忙, 陸陸續續) in Chinese utterances are used very often, but they do not always initiate a repair. That is, a simple pattern matching mechanism cannot be workable. Forth, the Chinese speech repairs may be initiated at various syntactic environment [11], e.g., before the subject, during the subject, after the subject and before the verb, during the verb, during a direct object, during a prepositional phrase, during subordination, and so on. The variety makes the identification of Chinese speech repairs more troublesome.

Because the identification of repairs in Chinese may not be deferred to the latter modules of the spoken language processing systems, this paper identifies several cues based on acoustic and prosodic analysis of repairs in a spoken corpus and proposes methods for exploiting these cues to correct the repetition repairs. Section 2 defines four major types of speech repairs. Section 3 introduces the spoken corpus. Sections 4 and 5 describe the acoustic and prosodic analysis of repairs. Section 6 is the concluding remarks.

2 Types of Chinese Speech Repairs

Heeman and Allen [1] divide English speech repairs into three types: fresh starts, modifications and abridged. For Chinese speech repairs, Chui [11] classify them into eleven patterns. In this section, we map these eleven patterns into four major types according to their surface forms.

Let $c_1 \ldots c_n\ c_{n+1} \ldots c_{n+m}$ be a sequence of Chinese characters. They may form an utterance or two consecutive utterances. The four major types of speech repairs are described as follows:

(a) Repetition Repair

There exists an i-character string, such that $c_{n-i+1} \ldots c_n$ (the repaired segment) is equal to $c_{n+1} \ldots c_{n+i}$ (the repairing segment). The repetition can range from a portion of a word up to several words. After being repaired, the utterances become $c_1 \ldots c_{n-i}\ c_{n+1} \ldots c_{n+m}$. (2) and (3) show two examples.

(2) 384 ..我--- {R1,1-1,1}
385 ..我是^知道我有這個毛病啊=,-

(3) 667 ..全國的]一起申-- {R1,1-1,1}
668 ..申請的=.\

The repetition repair occurs between utterances 384 (667) and 385 (668). The word "我" and the character "申" are repeated in (2) and (3), respectively. The character "申" is a portion of the word "申請". After being repaired, the utterances become "我是知道我有這個毛病啊" and "全國的一起申請的", respectively.

(b) Addition Repair

There are two types of addition repairs.

(i) There exist a j-character string and a k-character string, such that $c_{n-j+1} \ldots c_n$ (the repaired segment) is equal to $c_{n+k+1} \ldots c_{n+k+j}$. The character string, $c_{n+1} \ldots c_{n+k}\ c_{n+k+1} \ldots c_{n+k+j}$, forms the repairing segment in this case. After being repaired, the utterances become $c_1 \ldots c_{n-j}\ c_{n+1} \ldots c_{n+m}$. That is, $c_{n+1} \ldots c_{n+k}$ are added. (4) shows an example. The speaker's intention is “你們自已應該要”.

(4) 313 ...你們自已要,- {A1,1-1,2}
    314 ^應該要,-

(ii) There exist an i-character string, a j-character string and a k-character string, such that $c_{n-j-i+1} \ldots c_{n-j}$ is equal to $c_{n+1} \ldots c_{n+i}$ and $c_{n-j+1} \ldots c_n$ is equal to $c_{n+i+k+1} \ldots c_{n+i+k+j}$. The character string, $c_{n-j-i+1} \ldots c_{n-j}\ c_{n-j+1} \ldots c_n$, forms the repaired segment and the character string, $c_{n+1} \ldots c_{n+i}\ c_{n+i+1} \ldots c_{n+i+k}\ c_{n+i+k+1} \ldots c_{n+i+k+j}$, forms the repairing segment in this case. After being repaired, the utterances become $c_1 \ldots c_{n-j-i}\ c_{n+1} \ldots c_{n+m}$. (5) shows an example. The desired utterance is “他今天才說”.

(5) 1953 Z:[<F^他說,- {A1,1-2,4}
    1954 ..他今天]才說,-

(c) Replacement Repair

There are five types of replacement repairs.

(i) There exist an i-character string, an h-character string and a k-character string, such that $c_{n-h-i+1} \ldots c_{n-h}$ is equal to $c_{n+1} \ldots c_{n+i}$. The character string, $c_{n-h-i+1} \ldots c_{n-h}\ c_{n-h+1} \ldots c_n$, forms the repaired segment and the character string, $c_{n+1} \ldots c_{n+i}\ c_{n+i+1} \ldots c_{n+i+k}$, forms the repairing segment in this case. After being repaired, the utterances become $c_1 \ldots c_{n-h-i}\ c_{n+1} \ldots c_{n+m}$. (6) shows an example. The final utterance is "你一定沒有講出來吧".

(6) 733 Z:[你一定沒有]^唸%-- {P1,1-3,3}
734 ...沒有講出來吧=,-

(ii) There exist a j-character string, an h-character string and a k-character string, such that $c_{n-j+1} \ldots c_n$ is equal to $c_{n+k+1} \ldots c_{n+k+j}$. The character string, $c_{n-j-h+1} \ldots c_{n-j}\ c_{n-j+1} \ldots c_n$, forms the repaired segment and the character string, $c_{n+1} \ldots c_{n+k}\ c_{n+k+1} \ldots c_{n+k+j}$, forms the repairing segment in this case. After being repaired, the utterances become $c_1 \ldots c_{n-j-h}\ c_{n+1} \ldots c_{n+m}$. (7) shows an example. The word "很多" is replaced by "非常多" and the utterances become "非常多人過來我們這邊買東西".

(7) 2612 Y:..很多,- {P2,2-1,2}
2613 ..^非常[多人過來我們這邊]買東西,-

(iii) There exist an i-character string, a j-character string, an h-character string and a k-character string, such that $c_{n-j-h-i+1} \ldots c_{n-j-h}$ is equal to $c_{n+1} \ldots c_{n+i}$ and $c_{n-j+1} \ldots c_n$ is equal to $c_{n+i+k+1} \ldots c_{n+i+k+j}$. The character string, $c_{n-j-h-i+1} \ldots c_{n-j-h}\ c_{n-j-h+1} \ldots c_{n-j}\ c_{n-j+1} \ldots c_n$, forms the repaired segment and the character string, $c_{n+1} \ldots c_{n+i}\ c_{n+i+1} \ldots c_{n+i+k}\ c_{n+i+k+1} \ldots c_{n+i+k+j}$, forms the repairing segment in this case. After being repaired, the utterances become $c_1 \ldots c_{n-j-h-i}\ c_{n+1} \ldots c_{n+m}$. (8) shows an example. The word "差一滴" is substituted by "差幾滴".

(8) 528 ...(1.1)eh 我這還差一滴-- {P1,3-1,3}
529 ...eh 差幾滴 ei=.\

(iv) There exist an i-character string, a j-character string and an h-character string, such that $c_{n-j-h-i+1} \ldots c_{n-j-h}$ is equal to $c_{n+1} \ldots c_{n+i}$ and $c_{n-j+1} \ldots c_n$ is equal to $c_{n+i+1} \ldots c_{n+i+j}$. The character string, $c_{n-j-h-i+1} \ldots c_{n-j-h}\ c_{n-j-h+1} \ldots c_{n-j}\ c_{n-j+1} \ldots c_n$, forms the repaired segment and the character string, $c_{n+1} \ldots c_{n+i}\ c_{n+i+1} \ldots c_{n+i+j}$, forms the repairing segment in this case. After being repaired, the utterances become $c_1 \ldots c_{n-j-h-i}\ c_{n+1} \ldots c_{n+m}$. (9) shows an example.

(9) 406 [我就是--- {P1,3-1,2}
407 ..我是-]--

(v) Different from the above replacement repairs, the repaired segment and the repairing segment in this type do not match any characters. (10) shows an example.

(10) 659 ..他們-- {P1,2-1,1}
660 ...她唸的,-

(d) Abandon Repair

The original utterance is discarded and a new utterance is initiated. (11) shows an example.

(11) 54 H:...(.9)那就不要一堆-- {B1,6-1,1}
55 ..大家坐在一起幹嘛啊?/

After being repaired, the utterances become “大家坐在一起幹嘛啊”.

3 Spoken Corpus

The spoken corpus analyzed in this paper consists of two commonplace, everyday conversations among friends. Each is about forty-minute long. There are four and five speakers in these two conversations, respectively. It is originated from Professor Kaiwai Chui at National Chengchi University [11]. In total, this corpus contains 5395 utterances, 22409 words and 2602 turns. There are totally 440 self-repairs[5]. On the average, 17% of turns contain at least one repair. Tables 1 and 2 list the frequency distribution of each type of repairs in the two conversations.

Table 1. Frequency Distribution of Repairs in Conversation 1

Speaker	Repetition	Addition	Replacement	Abandon
L	23	8	3	1
H	54	12	9	4
Z	35	3	3	2
O	10	0	1	0
Total	122	23	16	7

Table 2. Frequency Distribution of Repairs in Conversation 2

Speaker	Repetition	Addition	Replacement	Abandon
L	39	5	4	4
W	53	9	6	8
Y	61	9	8	12
Z	44	2	4	1
J	2	1	0	0
Total	199	26	22	25

5 The speech repairs discussed in this paper are all self-repairs. That is, only the repairs accomplished by the same speaker are considered. This is because this kind of repairs is the most common form of repairs. Nevertheless, the present study includes repairs placed across different turns.

In the above statistics, we find that the repetition repairs form the majority (72.62% in conversation 1 and 73.16% in conversation 2) of the repairs. Addition (Replacement) repairs have 13.69% (9.52%) and 9.56% (8.09%) in conversations 1 and 2, respectively. The rest (4.17% in conversation 1 and 9.19% in conversation 2) are the most complex type of repairs, i.e., Abandon. Because this paper corrects repairs based on acoustic and prosodic cues, the Chinese characters in the spoken corpus are converted into the corresponding syllables manually.

4 Basic Analysis Method

4.1 Simple Pattern Matching Mechanism

Because the repetition repairs form the majority, we focus on the repetition repairs in this paper. Although the repetition repairs have the simple surface form, correcting such a kind of speech repairs is not trivial. That is, a simple pattern matching mechanism cannot work perfectly. Table 3 explains this phenomenon. A repair is proposed when a string of syllables repeats within an utterance or between two consecutive utterances.

Table 3. The Experimental Results Using Simple Pattern Matching Mechanism

Conversation	Total Repairs	Proposed	Correct
1	122	243	118
2	199	412	196
Total	321	655	314

Columns 2, 3 and 4 denote the total repetition repairs, the number of repairs proposed by the simple pattern matcher and the number of correct proposed repairs, respectively.

For example, 243 repairs are proposed by the simple pattern matcher in conversation 1, but only 118 of them are correct. That is, there are 125 false alarms. Since there are 122 repetition repairs in conversation 1, 4 repetition repairs are not captured. Some of them are listed below.

(12) 794 L:...那假如<L2 ^**Mac** L2>-- {R1,1-1,1}
795 ...<L2 **Mac** L2>的<L2 set up L2>不起來的.\

(13) 1541 ...連<L2 **Genni**--- {R1,1-1,1}
1542 ...**Genni**fer L2>三個=.\

(14) 1835 ...它<L2 **supermar**--- {R1,1-1,1}
1836 ...**super**]**mar**ket L2>也是很多那種,-

Because only Chinese speech repairs are considered, English repairs are lost. Although this technique can achieve recall rate of 97.82%, it has a relatively low precision rate, i.e., 47.94%.

Since the simple pattern matching mechanism cannot solve this problem properly, two basic analyses are firstly considered in the next two subsections: the length of the repeated syllable string and the number of inter-utterances.

4.2 The Length of the Repeated Syllable String

One of the most important things that we want to know is "how many syllables are repeated in the repetition repairs?"

Table 4. The Distribution of the Length of the Repeated Syllable String

Conv.\Length	**1**	**2**	**3**	**4**
1	71	40	6	1
2	107	72	15	2
Total	178	112	21	3

Table 4 lists the distribution of the length of the repeated syllable string in the repair. The length ranges from 1 to 4. That is, when a string of syllables repeats and the length of this string is greater than 4, we do not regard it as a repetition repair.

4.3 The Number of Inter-Utterances

Basically, most of the repetition repairs occur within an utterance or between two consecutive utterances of one speaker without interrupting by other speakers. However, if enough utterances pronounced by other speakers are inserted between two utterances of one speaker, the repetition repairs usually do not occur between them. (15) is an example.

(15) 2445 L:...哦=,-
2446 .那不[是%]-\
2447 Y:[三]個多<@ 月 @>.\
2448 ...[[三個半月]].\
2449 J:[[好過份啊]],-
2450 Z:(0)對啊=?/
2451 ...我覺得?/
2452 Y:...我^這次我十月=?/
2453 ...十五號啊,-
2454 ..<P 十五號十六[號 P>.\
2455 Z:[huh huh],-
2456 J:...你那個時候訂了.\
2457 就沒有貨了=是不是.\
2458 ...還是說%,-
2459 Y:...對啊.\
2460 他的意思[是%]--
2461 L:[不是].\

Although there is a matched string "不是" between utterances 2446 and 2461 for speaker L, it is neither a repaired segment nor a repairing segment. This is because 14 utterances interrupt the flow of thought of the speaker L. After analyzing the spoken corpus, some statistics are shown below.

(1) Total 13.69% of repetition repairs occur in the same utterance.

(2) Total 71.66% of repetition repairs occur between two consecutive utterances without interrupting by other speakers.

(3) Only 0.32% of repetition repairs occur across more than 3 utterances issued by other speakers.

According to these analyses, when more than 3 utterances pronounced by other speakers interrupt the speech of a speaker, we do not check whether there is a repetition repair or not.

5 Advanced Analysis Method

5.1 Unfilled Pause (...)

Observing the spoken corpus, we find that there is a significant unfilled pause (silence) between a repaired segment and a repairing segment for repetition repairs[6], whereas actual or intended repeated characters (syllables) usually do not have any unfilled pauses between them. Some typical examples are shown below.

(16) 505 ...(.8)不是說[你%-- {R1,1-1,1}
506 ..你..感覺到已經],-

(17) 606 Z:[那我--- {R1,2-1,2}
607 Z:..那我還=謝謝]]你們<F 啊 F>]?/
608 H:...(.7)^當然你要謝謝我們啊=,-

In the above examples, actual repeated characters (syllables) "謝謝" do not have any unfilled pause, whereas there is a unfilled pause between utterances 505 (606) and 506

6 Because the filled pauses such as um, un and er do not occur frequently in the spoken corpus, the effects of filled pauses are not demonstrated in this paper.

(607). Based on the basic analysis and the unfilled pause information, the experimental results for two conversations are listed below.

Table 5. The Experimental Results Using Unfilled Pause

Conversation	Total Repairs	Proposed	Correct
1	122	99	86
2	199	191	158
Total	321	290	244

The experimental results show that the precision rate is increased to 84.14%, and the recall rate is decreased to 76.01%.

5.2 Glottal Stop (%)

Glottal stop has the similar functions to unfilled pause. That is, a glottal stop may occur between the repaired segment and the repairing segment for the repetition repairs, whereas actual repeated characters usually do not have such a marker between them. (16) is an example. Based on the basic analysis and the glottal stop information, the experimental results for two conversations are listed below.

Table 6. The Experimental Results Using Glottal Stop

Conversation	Total Repairs	Proposed	Correct
1	122	31	31
2	199	85	82
Total	321	116	113

From Table 6, we find that glottal stop is a more reliable cue than unfilled pause, but it does not occur as frequently as unfilled pause. These points are verified by the high precision rate (97.41%) and the low recall rate (35.20%). When the basic analysis, the unfilled pause and glottal stop information are applied together, the experimental results for two conversations are listed in Table 7.

Table 7. The Experimental Results Using Unfilled Pause and Glottal Stop

Conversation	Total Repairs	Proposed	Correct
1	122	110	97
2	199	204	169
Total	321	314	266

Both the precision rate (84.71%) and the recall rate (82.87%) are all better than those in the former models.

5.3 Two Consecutive Equal Utterances

If two consecutive utterances are equal, repetition repairs usually do not occur within and between them when the length of the utterance is long enough. (18) is an example. The matched string is "是有這種" which denotes an emphasis. Thus, utterances 894 and 895 do not form a repair.

(18) 892 L:...(1.4)啊=,-
893 ..[對對對.\
894 ...是有這種,-
895 ..是有這種,-
896 對對 .\

Based on our spoken corpus, when the equal utterance length is greater than 2, no repetition repairs occur. This cue can eliminate some implausible repairs, so that the precision rate can be increased.

5.4 Cue Patterns

To increase the precision rate, another method is proposed. We collect the wrong proposed repairs that satisfy the criteria of basic analysis, unfilled pause and glottal stop. For the generalization, only the first syllable of each wrong proposed repair is concerned. The syllables whose frequency is larger than 1 is regarded as the type I cue pattern. By this way, six type I cue patterns, i.e., ㄧㄡˇ, ㄕˋ, ㄏㄠˇ, ㄛˊ, ㄚ˙ and ㄉㄨㄟˋ, are selected. (19) and (20) are two examples. Although the matched strings in these two examples satisfy the criteria of basic analysis, unfilled pause and glottal stop, they are not repairs.

(19) 69 J:[這什麼意思啊]?/
70 Y:...他是說他喜歡^當那隻,-
71 他說^錯了.\
72 J:...(1.3)^啊?/((WHAT?))

(20) 3362 Z:...(1.6)哦,\
3363 W:...其他的事情你們要...處理.\
3364 Y:(0)@@@=
3365 Z:...(1.3)哦真的啊,-

Because the first two patterns, ㄧㄡˇ and ㄕˋ, do not have the actual benefits in experiments, they are discarded. This is because only the negative examples (wrong proposed repairs) are used to generate this kind of patterns. Thus, a repair is proposed when a string of syllables repeats, satisfies criteria of basic analysis, unfilled pause and glottal stop, and the first syllable of the string does not belong to one of the four type I cue patterns.

Similarly, another kind of patterns is considered to increase the recall rate. Those repetition repairs that do not satisfy the criteria of unfilled pause and glottal stop are collected. The similar procedures for type I cue patterns are adopted to generate type II cue patterns. Finally, four such patterns are selected, i.e., ㄕㄜˊ, ㄋㄧˇ, ㄨㄛˇ and ㄋㄚˋ. (21) and (22) are two examples. Although the matched strings in these two examples do not satisfy the criteria of unfilled pause and glottal stop, they are all repairs.

(21) 464 O:[[你那-]]-- {R1,2-1,2}
465 [[你那-]]-- {R1,2-1,2}
466 L:[[那^吐]]出來=.\
467 O:(0)你那時--- {R1,3-1,3}
468 你那時候已經,-

(22) 1251 Z:[那就]-- {R1,2-1,2}
1252 (0)那就表示他不想你[[^買=]].\

Based on type II patterns, some additional repairs can be proposed when a string of syllables repeats, it does not satisfy the criteria of unfilled pause and glottal stop, but the first syllable of the string belongs to one of the four type II cue patterns.

Based on the method described in Section 5.2, the equal utterances information and the cue patterns, the experimental results are listed below.

Table 8. The Experimental Results

Conversation	Total Repairs	Proposed	Correct
1	122	120	111
2	199	190	180
Total	321	310	291

The experimental results show that the precision rate of 93.87% and the recall rate of 90.65% can be achieved. Besides, we also test another spoken corpus. The corpus has 504 utterances. It is about ten-minute long. It is originated from Professor Shuanfan Huang at National Taiwan University [12]. There are four speakers and totally 19 repetition repairs in this corpus. Table 9 lists the experimental results of the simple pattern matching and the method used in this section.

Table 9. The Experimental Results

Method	Total Repairs	Proposed	Correct
Pattern Matching	19	45	19
Our Method	19	21	18

Because the glottal stop is not annotated in this corpus, this cue is not used in this experiment. Apparently, our method (precision: 85.71%, recall: 94.74%) is better than simple pattern matching (precision: 42.22%, recall: 100%).

6 Concluding Remarks

Any spoken language systems will not perform well without treating speech repairs. Correcting speech repairs make more reliable environments for the subsequent processing. This paper identifies several cues based on acoustic and prosodic analysis of repairs in a large spoken corpus and proposes methods for exploiting these cues to correct the repetition repairs. The experimental results show that our method can achieve the precision rate of 93.87% and the recall rate of 90.65%, without using lexical information. O'Shaughnessy [13] claims that most speech repairs do not have lengthening prior to the hesitation pause. If this cue is used in our model, it can slightly increase the precision rate (95.37%), but the recall rate (76.95%) is greatly decreased.

Although our method can perform well in repetition repairs, other kinds of repairs such as addition, replacement and abandon repairs are not addressed in this paper. They have more complex surface forms and should be investigated further.

Acknowledgments

We are grateful to Professor Shuanfan Huang and Professor Kawai Chui for their kindly providing their spoken corpora to us.

References

[1] P. Heeman and J. Allen (1994) "Detecting and Correcting Speech Repairs," *Proceedings of ACL*, 1994, pp. 295-302.

[2] D. Bois, *et al.* (1992) "Discourse Transcription," *Santa Barbara Papers in Linguistics*, Vol. 4, 1992.

[3] M.L. Chou (1996) *Detecting and Correcting Chinese Speech Repairs*, Master Thesis, Department of Computer Science and Information Engineering, National Taiwan University, 1996.

[4] D. Hindle (1983) "Deterministic Parsing of Syntactic Nonfluencies," *Proceedings of ACL*, 1983, pp. 123-128.

[5] J. Bear, J. Dowding and E. Shriberg (1992) "Integrating Multiple Knowledge Sources for Detection and Correction of Repairs in Human-Computer Dialog," *Proceedings of ACL*, 1992, pp. 56-63.

[6] C. Nakatani and J. Hirschberg (1993a) "A Speech-First Model for Repair Detection and Correction," *Proceedings of EUROSPEECH*, 1993a, pp. 1173-1176.

[7] C. Nakatani and J. Hirschberg (1993b) "A Speech-First Model for Repair Detection and Correction," *Proceedings of ACL*, 1993b, pp. 46-53.

[8] B.A. Fox and R. Jasperson (forthcoming) "A Syntactic Exploration of Repair in English Conversation," *Descriptive and Theoretical Models in the Alternative*

Linguistics, Davis P. (Ed.), forthcoming.

[9] K.J. Chen and S.H. Liu (1992) "Word Identification for Mandarin Chinese Sentences," *Proceedings of COLING*, 1992, pp. 101-107.

[10] R. Sproat, *et al.* (1994) "A Stochastic Finite-State Word-Segmentation Algorithm for Chinese," *Proceedings of ACL*, 1994, pp. 66-73.

[11] K. Chui (1995) "Repair in Chinese Conversation," *Proceedings of the Second International Symposium on Language in Taiwan*, 1995, pp. 75-96.

[12] 黃宣範 (1996) "漢語口語語料庫的建立," 語言學門專題計畫研究成果發表會, 台北, 南港, 1996.

[13] D. O'Shaughnessy (1992) "Recognition of hesitation in Spontaneous Speech," *Proceedings of ICASSP*, 1992, pp. 521-524.

Appendix A The Transcribing Conventions of the Corpus

Units

{Carriage Return}	Intonation Unit
--	Truncated Intonation Unit
{Space}	Word
-	Truncated Word

Speakers

:	Speaker Identify / Turn Start
[ ]	Speech Overlap

Transitional Continuity

.	Final
,	Continuing
?	Appeal

Terminal Pitch Direction

\	Fall
/	Rise
_	Level

Accent and Lengthening

^	Primary Accent
=	Lengthening

Pause

...(N)	Long
...	Medium
..	Short
(0)	Latching

Vocal Noises

(H)	Inhalation
%	Glottal Stop
@	Laughter

Quality

< @ @ >	Laugh Quality
< Q Q >	Quotation Quality
< F F >	Fast Tempo
< PP PP >	Very Soft
< MRC MRC >	Each Word Distinct and Emphasized

Specialized Notations

< L2 L2 >	Code Switching from Mandarin to English
(( ))	Transcriber's Comment