A Close Ring Structure of Speech Recognition and Understanding*

Fu Qiuliang Chen Fang Lin Biqin and Yuan Baozong
Institute of Information Science, Northern Jiaotong University
Beijing, China, 100044

ABSTRACT—Speech recognition and understanding is an active and important research subject of information science. Now, continuous speech recognition is studied by a lot of scientists, and some man-machine interface systems which have good performance are delivered recently. In this paper, a new approach for Chinese speech(language) recognition and understanding is proposed, which is called a close ring structure with information feedback for speech recognition and understanding, and it is realized that a system of Chinese sentence understanding based on this new approach.

I. INTRODUCTION

As we know, Chinese speech recognition system may detect the result of all syllable of Chinese, but, as an integrated speech recognition (includes speech understanding) system, it must implement the task of Speech to Text(STT). It means that the integrated Chinese speech recognition system should translate a Chinese syllable string into a Chinese char, and have a capacity of translating Chinese syllable matrix into Chinese sentence.

Recently, researchers of speech processing and natural language processing have done much work on this subject area. Now, there have been two traditional approaches mainly to achieve this type of translation. They are (1) the approach of speech recognition and understanding with the knowledge database of Chinese grammar searched manually and (2) the approach of speech recognition and understanding with the database of the collocation probabilities between a word and the near word with statistic from a great number of Chinese language material by computer.

By analyzing for both traditional approaches of speech recognition and understanding, we could discover that there is not any essential difference in the two approaches, because this kind of system adopt an open ring structure, this kind of system has not any processing of information feedback, thus the system could not correct the former error in later processing. The simplified block diagram of these traditional approach is illustrated in Figure 1. It is very clear that this system only adopt an open ring structure, it has not any information feedback.

Because the system of speech recognition and understanding adopt the open ring structure, the output of speech recognition system is the input of the speech understanding system, there is not any information feedback from understanding system to the recognition system, so, the former result of this kind of system can influence directly the rear processing of speech recognition and understanding, but the former mistake couldn't be corrected in the rear speech processing.

Because the system of speech recognition and understanding adopt the open ring structure, the system has not any capacity of language generation, thus the speech understanding is not active, but we think that the speech understanding of human being is an active processing, this trait must be shown in the approach of speech understanding.

In this paper, we'll propose a new approach for speech recognition and understanding, it adopt a close ring structure with information feedback and has a capacity of activly understand speech (it has a module of language generation), It could correct the former mistake in the rear speech processing by information. Finally, we will do an

* This work was supported by the National Natural Science Foundation

0-7803-4253-4/97/$10.00 © 1997 IEEE

experiment on Chinese sentence understanding, give test data of the experiment and analysis to the result in detail.

II. CLOSE RING STRUCTURE FOR SPEECH UNDERSTANDING

After analyzing the speech processing of human being, this paper will propose a new approach for Chinese speech recognition and understanding, which can actively generate language and has a capacity of correcting by itself, it is a speech processing with information feedback, and the speech recognition and understanding system adopt a close ring structure(Reference[1]). We think that there are three facts in the speech recognition and understanding processing of human being, which supports the new approach.

(1)For human being, the capability of human being speaking is the precondition of that he could understand language, it is also said that the human being is able to speak firstly, then he can recognize and understand this kind of language secondly.

(2)The processing of oral language generation is that of information feedback. When man is speaking, the speaking word(or sentence) being heard by himself, and the word could be the guide for his speaking the other words. It is also said that the former words is very important heuristic information for the processing of the rear word generation, if we have known the former word, we can generate the rear words more easily, we could said that the processing of language generation is that of information feedback.

(3) The processing of speech understanding is that of information feedback also, and it includes the processing of language generation too. As we know, when a person is listening to the other, he is also predicting the words that will be spoken by the other man at the same time, it is also said that the listening person is doing work of predicting simultaneously language generation when he hears from the other. In the above speech processing, the listening person can compare the result of his understanding on speaking language with the predictive language by himself, extract the accurate sentence. In the processing of human being understanding, his predictive capable make him to actively understand natural language.

According to the above analysis, we have known that a speech recognition and understanding processing by human being is close ring processing of information feedback, and it also include the processing of language generation. It runs in two steps, first, the system compares the generating language with the understanding result, corrects the mistake by itself in the basis of linguistic rule set, and extract the final sentence, second, the system could get a few of information from the correcting processing, that can be controlled to the next processing of understanding, and may influence the next processing, that it includes basic sentence understanding, language generating and correcting by itself. This speech understanding includes information feedback is illustrated in figure2. From the figure, we could get that it includes a module of *Language Generation*(Reference[2]) and a module of *Speech Understanding*, the former module insure that the system has a capacity of active understanding speech, and the later module that the system has a capacity of correcting mistake by itself altogether with the module of *Correcting by Itself.*

III. THE EXPERIMENT OF SENTENCE UNDERSTANDING

By the above new approach of close ring structure speech understanding with information feedback, we design a system for Chinese sentence understanding. This system is based on both system, which are a *Chinese word recognition system* and a *Chinese phrase recognition and understanding* system.

The *Chinese word recognition system* could recognize the Chinese word, as we have been described in reference[3], This system finished the

translation from the Chinese syllable to Chinese char, and to Chinese word finally. It is the basis of Chinese phrase recognition and understanding system and the Chinese sentence understanding system.

The *Chinese phrase recognition and understanding system* has a capacity of recognizing and understanding Chinese phrase .It is a basis of Chinese sentence understanding system also, on the other hand, it is based on the Chinese word recognition system. We could discover the detail of this system in reference[4]

Based on both systems of Chinese language unit(Chinese word and phrase) recognition and understanding, we could design and implement a system for Chinese sentence understanding. It is illustrated in figure 3. It exists an information feedback between SU (Sentence Under-standing) and WR(Word Recognition), as same as the information feedback between PRU (Phrase Recognition and Understanding) and WR. At the same time, it exists the information feedback between SU and PRU too.

IV. CONCLUSION

Based the scheme of our Chinese sentence understanding(Figure3), We do an experimental system on the Chinese sentence, which could understanding 48 kinds of Chinese sentence structure. The whole system is based on a Chinese syllable recognition system. The system is in charge of translating the syllable matrix into legal sentence. The final result of this experimental system on Chinese sentence understanding is shown in table 1.

From the above table of result, we will know that the Chinese syllable recognition could recognize single Chinese syllable, that its first syllable correct rate is 69.57%, and the whole candidate syllable correct rate is 93.95%; but, we extract that the char correct rate in this sentence understanding system is 96.6%, it is more than 93.95%, because this system is adopted a technology of information feedback processing, it can feed the mistake of recognition back to the rear speech processing. It may say that this close ring structure speech understanding with information feedback has a capacity of understanding speech intangibly, this scheme could be the improvement of the performance of speech recognition system.

Table1: The result of experiment(Test Date: 1996-6-15)

First syllable correct *rate* of SR*	69.57%
All candidate syllable correct *rate* of SR	93.95%
Chinese char correct *rate* of SU**	96.60%
Sentence correct *rate* of SU	81.25%

* SR: Speech Recognition system
** SU: Speech Understanding system

V. REFERENCE

[1] Fu Qiuliang, The Study of Chinese Speech recognition and Understanding with Close Ring Structure, Ph.D. thesis, Northern Jiaotong University, 1996.

[2] Fu Qiuliang, Lin Biqin and Yuan Baozong, Language Generation and Active Speech Understanding, ICCPOL'97, HongKong, April, 1997.

[3] FuQiuliang, et.al, Chinese Word Recognition and Understanding with Information Feed-back, ICSP'96, Beijing, Oct., 1996.

[4] Fu Qiuliang, et.al, Chinese Structural Phrase Recognition with Information Feedback, The Journal of Northern Jiaotong University, To be submitted.

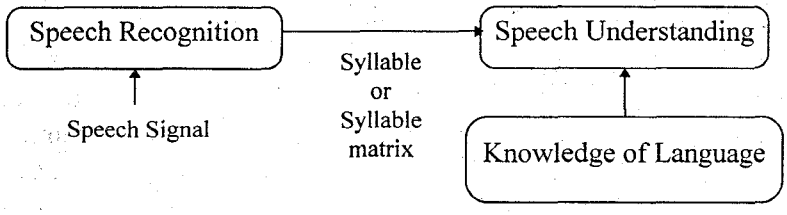

Figure 1. The traditional approach of speech recognition and understanding

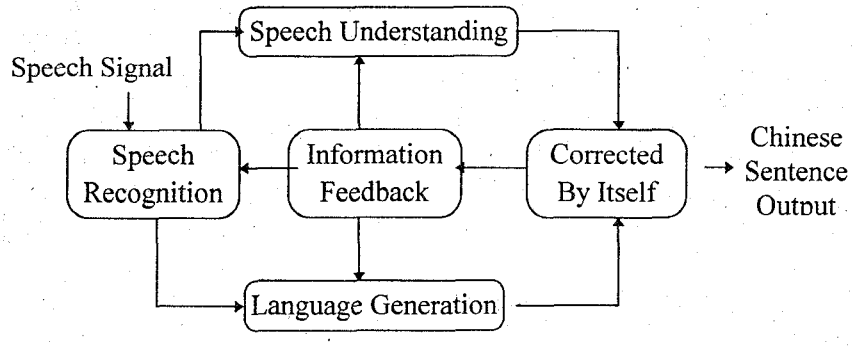

Figure 2. The Speech Understanding with Information Feedback

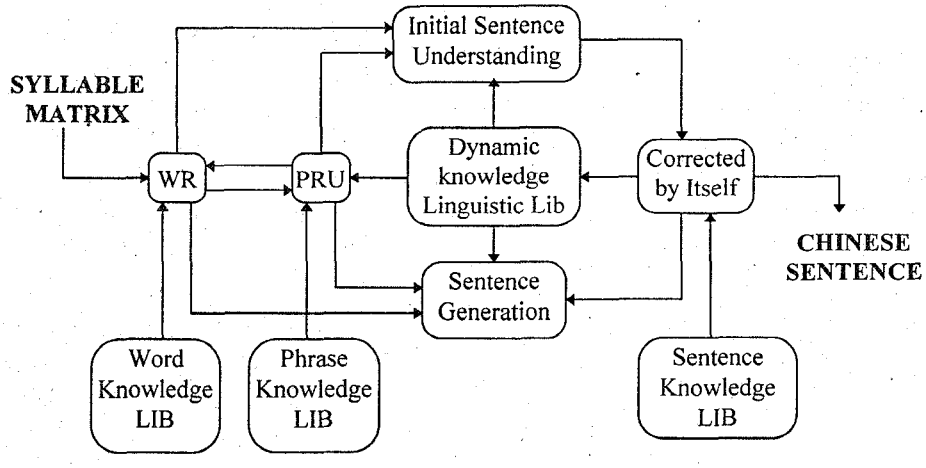

Figure 3. The Diagram of Chinese Sentence Understanding