

Word Co-occurrence Analysis with Utterance Pairs for Voice Interface Robot for Elderly Care

Yuka Kobayashi^{#1}, Daisuke Yamamoto^{#2}, Miwako Doi^{#3}

[#]Corporate Research & Development Center, Toshiba Corporation
Komukai Toshiba-cho, Saiwai-ku, Kawasaki-shi, Kanagawa 1, Japan

¹ yuka3.kobayashi@toshiba.co.jp

² daisuke2.yamamoto@toshiba.co.jp

³ miwako.doi@toshiba.co.jp

Abstract—We have been developing a robot to serve as an active listener with elderly people throughout the days. A human active listener misses a part of words but can catch the appropriate keywords from the interlocutor's utterance to continue the conversation based on dialogue context. A voice interface robot also cannot recognize the interlocutor's utterance word by word because of speech recognition errors. We propose using a word co-occurrence analysis with utterance pairs in order to obtain the appropriate keywords. Dialogue context appears as the high co-occurrence with words in utterance pairs. The robot reacts to the recognized words that have high co-occurrence with words in the robot's last utterance.

I. INTRODUCTION

We have been developing a robot as an active listener with elderly people though days using our robot ApriPocoTM[1]. A voice interface robot cannot recognize what their interlocutor says word by word because of speech recognition errors.

We suggest a way of identifying the correct words based on semantic co-occurrence from recognition results based on the hypothesis that the word co-occurrences between correct words are higher than ones between incorrect words and correct words. The robot calculates co-occurrence rates between words in the robot's last utterance and words in the recognition result of the user's utterance and uses the words with high co-occurrence for the robot's next utterance. We call this method word co-occurrence analysis (WCA) with utterance pairs. There are many researches extracting appropriate words with co-occurrence between words in recognition results [2]-[5]. They use only words in recognition results, which are uncertain information.

II. WORD CO-OCCURRENCE ANALYSIS (WCA) WITH UTTERANCE PAIRS

Fig. 1 shows the system components. The robot makes a fixed utterance first. After the user talks and the recognition result is obtained, the robot makes the next utterance using WCA. The robot divides the recognition result and the robot's utterance into words, calculates a co-occurrence rate S_p for each word in the recognition result with any word in the robot's utterance and uses the word with the highest S_p in the next utterance. The recognition result may include errors, but because these words didn't have high co-occurrence with the robot's words, they are removed from outputs. We use Cosine Coefficient for calculating co-occurrence rate.

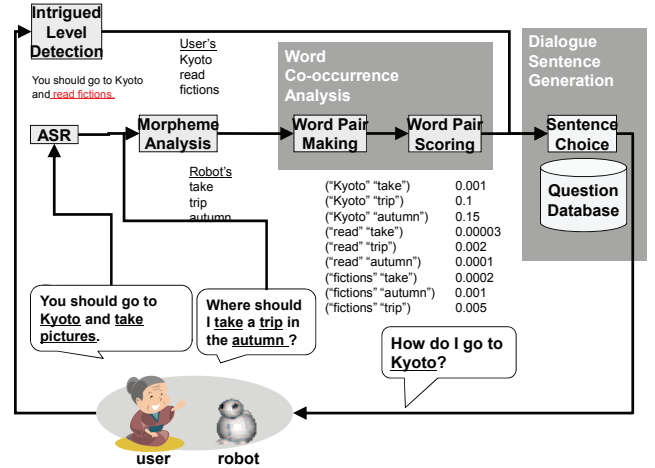


Fig. 1. System Overview of Word Co-occurrence Analysis with Utterance Pairs.

III. CONCLUSION

We conducted an experiment and the result was that the precision rate was 39% in the original recognition system and it was improved to 61% with the WCA. We intend to develop this system with a view to realizing better human-robot interaction.

ACKNOWLEDGMENT

This research was supported by the Ministry of Internal Affairs and Communications.

REFERENCES

- [1] Y. Kobayashi, D. Yamamoto, T. Koga, S. Yokoyama and M. Doi, "Design targeting voice interface robot capable of active listening", in *Proc. HRI*, pp.161-162, 2010.
- [2] S. Cox, S. Dasmahapatra, "High-level approaches to confidence estimation in speech recognition", *Speech and Audio Processing, IEEE Transactions on*, vol.10, no.7, pp. 460- 471, Oct. 2002.
- [3] Q. Gao, X. Lin and X. Wu, "Just-in-time latent semantic adaptation of language model for Chinese speech recognition using web data", *Spoken Language Technology Workshop, IEEE.*, 2006.
- [4] J. R. Bellegarda, "Large Vocabulary Speech Recognition with Multispan Statistical Language Models", *Speech and Audio Processing, IEEE Transactions on*, vol.8, no.1, pp.76-84, 2000.
- [5] R. Kuhn and R. De Mori, "A Cache-based natural language model for speech recognition", *Pattern Analysis and Machine Intelligence, IEEE.*, vol.12, no.6, pp.570, 1990.