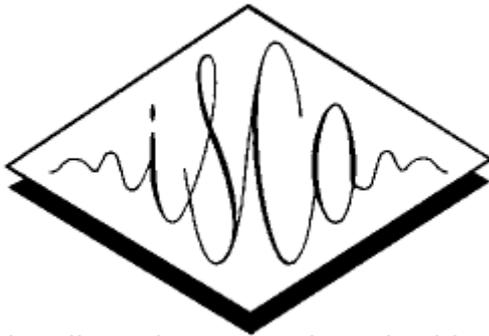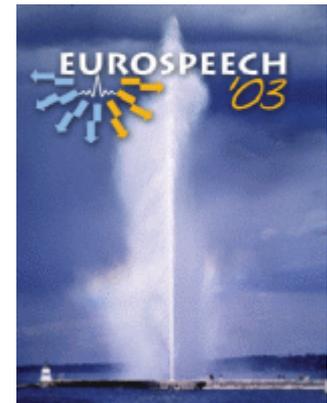## ISCA Archive
http://www.isca-speech.org/archive

EUROSPEECH 2003 - INTERSPEECH 2003
8th European Conference on Speech Communication and Technology

Geneva, Switzerland
September 1-4, 2003

Named Entity Extraction from Japanese Broadcast News

Akio Kobayashi (1), Franz J. Och (2), Hermann Ney (3)

(1) NHK Science & Technical Research Laboratories, Japan
(2) University of Southern California, USA
(3) RWTH Aachen, Germany

This paper describes a method for named entity extraction from Japanese broadcast news. Our proposed named entity tagger gives entity categories for every character in order to deal with unknown words and entities correctly. This character-based tagger has models designed by maximum entropy modeling. We discuss the efficiency of the proposed tagger by comparison with a conventional word-based tagger. The results indicate that the capability of the taggers depends on the entity categories. Therefore, the features derived from both character and word contexts are required to obtain high performance of named entity extraction.

Full Paper

Bibliographic reference. Kobayashi, Akio / Och, Franz J. / Ney, Hermann (2003): "Named entity extraction from Japanese broadcast news", In *EUROSPEECH-2003*, 1125-1128.