CONSTRUCTING PRONUNCIATION LEXICON WITHOUT PHONEME INVENTORY

A. Guntario Sukma Cahyani[*]

Institut Teknologi Bandung
Departement of Informatics
Jl. Ganeca no. 10 Bandung, Indonesia

B. Sakriani Sakti, C. Satoshi Nakamura[†]

Nara Institut of Science and Technology
Departement of Information System
8916ł5, Takayamacho, Ikoma, Nara, Japan

ABSTRACT

Pronunciation lexicon plays an important role in automatic speech recognition (ASR) system. Often times, pronunciation lexicon uses canonical form i.e. the correct pronunciations as agreed by experts, but in some cases, the language does not have known phoneme inventory, as in the case of low resource language, or the available phoneme inventory i.e. the canonical form of the pronunciation does not represent the surface pronunciation that are uttered by the target speaker, as in the case of non-native and code-switching ASR. For such cases, pronunciation lexicon cannot be built manually, since the phonemes are practically unknown. In this paper, we investigate several methods of lexicon building for the case of low-resource language: data-driven mapping which generated pronunciations of low-resource language words using rich-resource language phonemes inventory, unsupervised learning that discovered new acoustic units, and we proposed a method that is a combination of the two. We carried out the experiment in two general scenarios: first scenario that served as an analysis, using Indonesian language as a low-resource language and English as rich-resource language , and the second scenario as an implementation using Indonesian Ethnic language as the low-resource language and Indonesian as the rich resource language. We found that the best method for building lexicon without phoneme inventory is the method we proposed, and we implement it to the real low-resource language ASR.

Index Terms— pronunciation lexicon, automatic speech recognition, lexicon building, low-resource language

1. INTRODUCTION

[*]
[†]