

中国英语学习者的语音数据分析：基于语音语料库的建设 [An Analysis of Chinese English Learners' Speech Data: Based on the Construction of the Speech Corpus]

Zhu Lin, Beijing International Studies University

超音段音位的学习对中国英语学习者来说是一个难点。为了得到中国英语学习者的语音数据，建设对比本族发音人和中国发音人的语音语料库成为了重要的研究项目。语音语料库的建设过包括以下几个步骤：制定研究目标、与发音人签订法律文书、发音语料的确定、录制过程、以及语音标注等。该语音语料库重点是用于语音识别和研究中国人说英语的节奏模式。通过对该语音语料库的数据提取，得出中国发音人的英语节奏模式不是以重音为支点的节奏模式，而是以音节等时长的节奏模式。

[English Phonetics is a difficult problem for Chinese, but also reflected on the suprasegmental phonemes. In order to study the data of the Chinese learners, there is need to build a corpus with native speakers (including American speakers and British speakers) and Chinese speakers in English. There are some processes of English speech corpus, such as the establishment of a research target, recording, the people who need legal documents signed, determining the materials for corpus, and labeling ect. The speech corpus is mainly used for speech synthesis and research the rhythm of Chinese English. The corpus points out that the rhythm pattern of Chinese English is syllable timed rather than stress timed pattern like native speakers.]