

中研院具音段標記之中文對話語音資料庫

本院覽號

24T-1031221

公告日期

智財權狀態

know-how

摘要

中研院具音段標記之中文對話語音資料庫(The Sinica Phone-aligned Chinese Conversational Speech Database)為Sinica MCDC_8語音資料庫的一部分。包含3.5小時，16人的帶音段標記之對話語音資料。語音檔以wav格式儲存(.wav);音節與音段時間標記以PRAAT格式儲存 (.TextGrid)。

技術優勢

本語料庫音節語音段時間標記的文字檔案已PRAAT格式儲存(.TextGrid)。所有標記邊界皆經過人工驗證。

應用範圍

語音辨識、語言學訓練、語料庫之應用。

創作人

曾淑娟



中央研究院
ACADEMIA SINICA