

名称：筑波大多言語音声コーパス

提供者：筑波大学知能情報生体工学研究室

作成年：2001年（平成13年）

問い合わせ先：NII-SRC

言語：11言語（英語、ドイツ語、フランス語、スペイン語、ロシア語、アラビア語、インドネシア語、タイ語、中国語、韓国語、日本語）

話者・回数：各言語について1～7名、合計98名。話者は主に筑波大学に在籍する外国人留学生および外国人教師である。タイ語については、タイ国立電子工学・計算機工学センターの協力を得た。発声回数1回。A/D条件：16kHz, 16ビット量子化、防音室でヘッドセットマイク収録。

附属情報：音声波形の他に、各言語の「こんにちは」と「おはよう」に相当する言葉の音声波形、スペクトログラム、基本周波数等の分析データを別冊に入れた。

収録

配布媒体・方法：CD-ROM1枚。説明書添付

配布条件：研究目的に限る。

配布費用：無料。郵送料は実費負担をお願いする。

配布実績：これまで、大学等に60セット配布済み。

その他：筑波大学特別プロジェクト「東西言語文化の類型論」（平成9～13年度）で作成。各言語について共通の内容の発話を収録している。発話内容は、世界中で広く一般に使用されているような単語として、数字14語、月の名12語、曜日7語、天気用語4語、挨拶6語、返答3語、時間に関する用語4語、計50語を選んだ。また、連続音声としては、世界で知られていて、資料が用意に入手できる1分間くらいの物語として、イソップ童話の「北風と太陽」を選んだ。各言語の発声テキストは、日本語（英語）を基にして各言語の話者（少なくとも2名）に作成してもらった。音声データは防音室でヘッドセットマイクを使ってDATに収録した。発話時間は1人あたり平均5分くらいである。データはCD-ROM1枚に格納されている。話者の年齢、性別、出身地、母語、居住履歴、両親の出身地等を発声者データとして記入してもらった。