

ATR子供日本語音声データベース (小学校低学年/高学年)

Japanese Elementary School Children Speech Database

概要

本データベースは、日本語を母語とする小学校低学年(1~3年生)と小学校高学年(4~6年生)の子供にあらかじめ準備した絵本などの台詞を読み上げてもらい、その音声をスマートフォンで収録しています。

特徴

- ◆ 日本人子供(6~13歳)の411人(男子175名 女子236名)による音声を収録
- ◆ 話者毎の書き起こしデータと音声ファイルにより音声認識開発(モデル学習/評価)に利用可能

話者数/発話文リスト/発話量

	低学年	高学年	合計
話者数 (男子/女子)	179名 (71/108)	232名 (104/128)	411名 (175/236)
発話数	平均約325/話者	平均約385/話者	平均約359/話者
総発話数	58,214	89,454	147,668
発話時間 ^[1]	約116.6時間	約145.4時間	約262.0時間

[1] 発話前後の無音区間を含む。

- ◆ 発話文リスト(低学年：1019文 高学年：1018文)
(例)L0088：オオカミは クマと おなじように、 けっこんしきに いくのをやめて かえりました。
H0307：日本は国をあげて、省エネ技術(ぎじゅつ)の開発に取り組んできました。

データ構成

- ◆ 発話音声：WAV形式、16kHz、16bit、モノラル(スマートフォンで収録)
- ◆ 読み上げ原稿 } TSV(タブ区切り)形式、文字コードUTF-8(BOM無し)、改行コードLF
- ◆ 書き起こしデータ }
- ※ 納品メディア：DVD又はBD
(低学年：約12.9GB、高学年：約16.2GB、合計：約29.1GB)

価格(消費税抜)

商用利用：(低学年) ¥6,700,000 研究利用：(低学年) ¥4,020,000
(高学年) ¥8,400,000 (高学年) ¥5,040,000

※ 年齢別のサブセットもご用意しています。詳しくはお問合せ下さい。

※ アカデミック価格についてはお問合せ下さい。

ご利用方法

株式会社ATR-Promotionsより用途に応じた条件によるライセンス(利用許諾)をいたします。
ホームページまたは下記よりご連絡ください。

●本データベースの仕様/価格は予告なしに変更されることがあります。

株式会社 **ATR-Promotions**

<https://www.atr-p.com>

〒619-0288 京都府相楽郡精華町光台2-2-2

TEL:0774-95-1300 FAX:0774-95-1191

E-mail:voice@atr-p.com