

ATR多数話者音声データベース音素バランス文 48k 詳細製品情報

1. 基本情報	製品名 価格 ライセンス区分 データタイプ 言語 利用分野 用途 原音声データベース制作者 原音声データベース制作年	ATR多数話者音声データベース音素バランス文 48k 600万円(消費税含まず) 商用 音声 日本語 音声合成(TTS)、自動音声認識(ASR)、音声品質評価など 音声認識エンジンの改良、トレーニングデータ、ノイズ耐性やエコーキャンセリングなど ATR音声翻訳通信研究所 1995年度～1998年度
2. 話者属性	母語 居住地 経験/スキル 年齢/年代 性別	日本語 全国各地 音声の収録に関して経験のない、もしくは少ない一般人 14～65歳 男女
3. 定性要件	発話スタイル 言語パターン 発話タスク/ドメイン 発話モード 発話文長 発話方法	指定原稿の読み上げ発話 日本語 ATR音素バランス503文(*2) 連続発話 短文 普通の速さ、音量
4. 定量要件	総話者数(男/女別) 総発話数 総発話時間(*1)	3,771名(1,380名/2,391名) ※但し2名(男女各1名)タイトル間重複話者あり 112,560発話 127.46時間
		(*1) 音声データから発話前後の無音区間を除き実効発話区間のみで計算
		(*2) 言語的に区別される音声の最小単位を「音素」という。日本語ではローマ字表記のアルファベット1文字にほぼ相当する。日本語の音素(子音と母音)の連なり(音素連鎖)のパターンがバランスよく出現するよう考慮された文セットを「音素バランス文」と呼ぶ。「ATR音素バランス503文」は(株)国際電気通信基礎技術研究所(ATR)の音声翻訳研究成果であり、「ATR503文」としてひろく知られ音声研究・開発の標準データとして活用されている
5. 音声収録要件	収録環境 収録機器 マイクロホン データ形式 サンプリング周波数 サンプルサイズ	環境雑音の影響が極めて少ない場所 DATレコーダー C-355(SONY)：単一指向性コンデンサマイクロホン リニアPCM 48kHz 16bit
6. 音声データ要件	データ形式 無音区間 サンプリング周波数 サンプルサイズ	WAV形式(ヘッダ付きリニアPCM) 発話音声の前後に500ミリ秒～1秒程度 48kHz 16bit
7. テキストデータ要件	データ形式 文字コード 行末コード	プレーンテキスト UTF-8(BOM無し) LF(x0A)
8. データ構成	収録情報データ ATR503文テキストデータ 音声データコメント 音声データ ラベルデータ	発話セット単位で収録条件や話者情報を記載したテキストデータ 基本形(漢字かな)と読み(かな)からなるATR503文テキストデータ 読み間違いのコメントや音声エラーを記載したテキストデータ 1文ごとにファイル(WAV:サンプリング周波数48kHz)化され話者・発話文単位でディレクトリに格納 音素表記(ATR音節表に基づく)の書き起こし文字列とポーズ区切りの音声区間の始端終端時刻情報データ(時間単位:ミリ秒)
9. 知的財産権		本データベースの知的財産権(著作権)はすべて株式会社国際電気通信基礎技術研究所(ATR)、およびに株式会社ATR-Promotionsに帰属する 株式会社ATR-Promotionsは、本データベースの独占的な利用許諾権を有している
10. 問い合わせ		○問合せ先：株式会社ATR-Promotionsコーパス事業部 ・Email: corpus@atr-p.com ・TEL: 0774-95-1300 FAX: 0774-95-1191 ・住所: 〒619-0288 京都府相楽郡精華町光台2丁目2番地2 ○問い合わせフォーム: https://www.atr-p.com/form1/form1.html