

ATR中規模自然会話音声データベース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中規模自然会話音声データベース48k 400万円（消費税含まず） 商用 音声データ（WAV 48kHz/16bit/モノラル）、書き起こしテキスト、時刻情報付き音素単位書き起こしテキスト、収録情報データ、形態素データ 日本語 音声認識、音声合成、教育、医療、対話AI等 音声認識モデルの学習・評価、会話型AIの応答精度向上、自然発話に基づくTTS合成、話者特徴の抽出と適応、話し言葉の文字起こし、発話スタイル分析および応対訓練教材の開発など ATR音声翻訳通信研究所 1994～1997年
2. 話者属性	母語 居住地域 経験/スキル 年齢/年代 性別	日本語 特に規定せず 音声の収録に関して経験のない、もしくは少ない話者：模擬会話の申込み者役 ホテルのフロントでの業務経験者、模擬会話の収録経験者：模擬会話の担当者役 高齢者/子供を除く（年齢情報取得せず） 男女
3. 定性要件	発話スタイル 発話タスク/ドメイン 発話文長 発話方法	話し言葉特有のフィラー（えー、あのー）、言い淀み、言い直し、間投詞、ポーズを含む自然な話し方 旅行に関する模擬会話形式による対話想定発話（非対面収録） 短文 会話プロットに基づき、読み上げではなく、話者が自由に発話内容を即興的に構成
4. 定量要件	話者数（男/女別） 総発話時間長（男/女別） 総発話数（男/女別）	562名（218名/344名） 47.28時間（16.93時間/30.35時間） ※無音区間を含む 38,544発話（13,346発話/25,198発話）
5. 音声収録要件	収録環境 収録機器 マイクロホン 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：DA-30<TASCAM>, DA-30mkII<TASCAM>他 C-355（SONY）：単一指向性コンデンサマイクロホン MU-2C（SANKEN）：単一指向性ダイナミックマイクロホン その他 48,000Hz（48kHz） 16bit
6. 音声データ要件	ファイル形式 エンコーディング 無音区間 サンプリング周波数 サンプルサイズ チャンネル構成	WAV 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 48,000Hz（48kHz） 16bit モノラル
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	書き起こしデータ 収録情報データ 音声データ 時刻情報付き音素単位書き起こしデータ	発話文の書き起こしテキストデータ 収録条件や話者情報を記載したテキストデータ 発話単位の音声データ 音声データに対応した音素表記書き起こしラベルデータ。発話およびポーズの始端時刻情報がミリ秒（ms）単位で付与
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	