

ATR自然発話音声・言語データベースSLDB(日英対話) 詳細製品情報		
1. 基本情報	製品名	ATR自然発話音声・言語データベースSLDB(日英対話)
	価格	商用 100万円 研究 18万円（消費税含まず）
	ライセンス区分	商用/研究/アカデミック
	データタイプ	日英対訳・非対面模擬会話音声データおよび音素バランス文音声データ（WAV音声、日英テキスト対訳、話者・発話メタ情報を含む）
	言語	日本語・英語
	利用分野	・音声認識（Automatic Speech Recognition: ASR）技術開発 ・音声合成（Text-to-Speech: TTS）システム開発 ・自然言語処理（Natural Language Processing: NLP） ・対話システム（Dialogue System）開発 ・話し言葉研究（日本語教育、社会言語学など）
	用途	・高精度なタスク志向型対話認識モデルの学習データ 特に電話対応やカスタマーサービスのシナリオに特化した音声認識エンジンの精度向上 ・日本語の自然な話し言葉（フィラー、言い淀みを含む）の分析とモデル化 ・ホテル・旅行業界向けAIオペレーターやチャットボット開発のための対話理解モデルの構築 ・音声合成における自然な対話間のポーズや抑揚のモデル化
	原音声データベース制作者	ATR音声翻訳通信研究所
	原音声データベース制作年	1994年
2. 話者属性	母語	日本語・英語
	居住地域	特に規定せず
	経験/スキル	申込み者役：音声の収録に関して経験のない、もしくは少ない話者 模擬会話の担当者役：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
	年齢/年代	高齢者/子供を除く（年齢情報取得せず）
	性別	男女
3. 定性要件	発話スタイル	模擬会話：自由発話部分には言い直し・間投詞・相づち等を含み、実運用に近い自然な会話表現を収録。非対面自由対話（通訳電話設定） 音素バランス文：自然な音素バランス文（ATR503文）の読み上げ
	発話タスク/ドメイン	ホテル予約、宿泊サービスに関する問い合わせ対応など、宿泊施設の接客場面を想定した日英模擬会話。
	発話文長	短文
	発話方法	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）	77名（37名/40名）
	総発話時間長（男/女別）	模擬会話：18.44時間（3.18時間/15.27時間） ※無音区間を含む 音素バランス文：2.33時間（0.7時間/1.63時間）
	総発話数（男/女別）	模擬会話：15,859発話（2,837発話/13,022発話） 音素バランス文：1,756発話（553発話/1,203発話）
5. 音声収録要件	収録環境	環境雑音の影響が極めて少ない場所
	収録機器	DATレコーダー：PCM-7030 (SONY), PCM-7050 (SONY), DA-P20 (TEAC) 他
	マイクロホン	MKH416 (Sennheiser)：超指向性コンデンサマイクロホン MU-2C (SANKEN)：超指向性コンデンサマイクロホン その他
	サンプリング周波数	48,000Hz (48kHz)
	サンプルサイズ	16bit
6. 音声データ要件	ファイル形式	WAV
	エンコーディング	ヘッダ付きリニアPCM
	無音区間	発話音声の前後に500ミリ秒～1秒程度
	サンプリング周波数	16,000Hz (16kHz) 16kHzにダウンサンプリング
	サンプルサイズ	16bit
	チャンネル構成	モノラル
7. テキストデータ要件	データ形式	プレーンテキスト
	文字コード	UTF-8（BOM無し）
	行末コード	LF (x0A)
8. データ構成	書き起こしデータ	発話文の書き起こしテキストデータ
	収録情報データ	収録条件や話者情報を記載したテキストデータ
	音声データ	発話単位の音声データ
	時刻情報付き音素単位書き起こしデータ	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終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	