

(2012年2月1日現在)

既存の日本語教育研究リソース（学習者のデータを含むもの）

話し言葉

- 「日本語学習者による日本語／母語発話の対照言語データベース（発話対照 DB）」（オンライン）
 - ・ 日本語学習者（中国語、韓国語、タイ語母語話者）による日本語の発話と、それとほぼ同内容の母語発話とを収集・整理したもの。また比較のため、同じ課題に基づく日本語母語話者の日本語発話データも収集されている。（一部、日本人とのロールプレイも含まれる。）
 - ・ 国立国語研究所 <http://jpforlife.jp/hatsuwadb.html>

- 「日本語学習者会話データベース」（オンライン）
 - ・ 日本語学習者と日本語母語話者である面接者（インタビュアー・テスター）による、1 データあたり約30分の会話を集めたもの。会話の文字化データ 339 件とその音声データ 215 件。学習者の日本語レベル、年齢、性別、出身国、母語、職業、日本語滞在期間等からデータの検索が可能。インタビューの方法及びレベル判定には、ACTFL（全米外国語教育協会）の面接式口頭能力テスト Oral Proficiency Interview（ACTFL-OPI）の形式を採用。
 - ・ 国立国語研究所 <https://dbms.ninjal.ac.jp/nknet/ndata/>

- 「日本語学習者会話ストラテジーデータ」（オンライン）

英語を母語とするか英語能力の高い外国語母語話者である日本語学習者で日本語能力がある程度ある者を対象に、外国語母語話者であるインフォーマントの日本語使用の実態を把握することを目的としたもの。調査方法は、インフォーマントと日本語母語話者との日本語による実際の会話を録音・文字化し、それをもとに録音日から1週間以内に英語によるインタビューを行い、会話のプロセスを内省報告してもらう、フォローアップインタビューの方法を採用。

 - ・ 国立国語研究所 <https://dbms.ninjal.ac.jp/nknet/ndata/>

- 「BTS による多言語話し言葉コーパス－日本語会話 1（日本語母語話者同士の会話）」（CD-ROM）
「BTS による多言語話し言葉コーパス－日本語会話 2（日本人と学習者の会話）」（CD-ROM）
 - ・ 「日本語会話 1（日本語母語話者同士の会話）」は計 121 会話、約 21 時間。「日本語会話 2（日本人と学習者の会話）」は計 57 会話、約 7 時間。いずれも、「BTSJ（基本的な文字化の法則）」に基づく文字化テキストを収録。
 - ・ 東京外国語大学大学院 総合国際学研究院 宇佐美まゆみ
<http://www.tufs.ac.jp/ts/personal/usamiken/corpora.htm>

- 「インタビュー形式による日本語会話データベース」（CD-ROM）
 - ・ 日本語母語話者（54 名）と非母語話者（56 名）の発話パターンの比較分析と、日本語教育向けの基礎資料となるコーパス構築を目的として収集されたもの。会話モードとロールプレイモードの二つの要素によって構成される日本語 OPI（Oral Proficiency Interview）の文字化テキスト及び音声を収録。
 - ・ 北九州市立大学 国際環境工学部情報メディア工学科 上村隆一研究室

- 「KY コーパス」（CD-ROM）
 - ・ 90 人分の OPI を文字化したもの。90 人の被験者の母語別内訳は中国語・英語・韓国語、各 30 人。30 人の OPI の判定結果別の内訳は、それぞれ初級 5 人・中級 10 人・上級 10 人・超級 5 人。
 - ・ 実践女子大学 文学部 山内博之

書き言葉

- 「日本語学習者による日本語作文と、その母語訳との対訳データベース（作文対訳 DB）」（オンライン）
 - ・ 日本語作文 1,565 編（データ収集国は 21 か国、日本人による日本語作文 66 編を含む）及び作文執筆者本人による母語訳、日本語教師等による添削情報からなる。
 - ・ 国立国語研究所 <http://jpforlife.jp/regist/regist.cgi>

- 「日本語学習者言語コーパス」（オンライン）
 - ・ E-ラーニングシステムを活用した日本及び海外（漢字圏）の大学生の作文コーパスと、誤用コーパス作成を主眼とした、海外（非漢字圏）の大学生の作文コーパスからなる。品詞タグのタグ付けも行われている。
 - ・ 東京外国語大学大学院 総合国際学研究院 海野多枝
<http://cbllc.tufs.ac.jp/llc/ja/index.php?menulang=ja>

- 「JLPTUFS 作文コーパス」(CD-ROM)
 - ・ 東京外国語大学「全学日本語プログラム」(JLPTUFS)において書かれた作文（作文データ数：1,515 編 学習者の出身国・地域数：55 か国）が収録されている。作文は、学習者の執筆した作文をスキャナーで取り込んだ PDF 形式のものと、それをテキスト形式で入力したものの 2 種類が収録されている。
 - ・ 東京外国語大学留学生日本語教育センター 鈴木智美・中村彰

- 「日本・韓国・台湾の大学生による日本語意見文データベース」(オンライン)
 - ・ 日本語を母語とする大学生（134 名）と日本語を学ぶ大学生（台湾 57 名、韓国 55 名）が日本語で執筆した意見文を収録したもの。日本語学習者データには、作文執筆者本人による母語への翻訳（意識）も含まれる。作文データは PDF 形式、テキスト形式、エクセル形式の 3 種類が収録されている。
 - ・ 東京外国語大学留学生日本語教育センター 伊集院郁子
http://www.tufs.ac.jp/ts/personal/ijuin/koukai_data1.html

その他

- 「寺村誤用例集データベース」(オンライン)
 - ・ 寺村秀夫（1990）『外国人学習者の日本語誤用例集』に、品詞や国籍等による様々な検索機能を付加したデータベース
 - ・ 国立国語研究所 <http://teramuradb.ninjal.ac.jp/db/>

- 「言語行動意識調査データ」(オンライン)
 - ・ ビデオで言語行動場面を提示し、それに関する質問を行う面接調査方式で、外国語母語話者と日本語母語話者の言語意識の違いについて調査したデータ。「頼みごとへの返事（頼んだ側の用件のきりだし、頼まれた側の曖昧な返答）」のデータ（在外日本人 211 人、在日外国人 159 人、国内日本人 60 人）を公開。
 - ・ 国立国語研究所 <https://dbms.ninjal.ac.jp/nknet/ndata/>