

国立国語研究所学術情報リポジトリ

目次

メタデータ	言語: Japanese 出版者: 研究社 公開日: 2026-07-17 キーワード: 作成者: メールアドレス: 所属:
URL	https://repository.ninjal.ac.jp/records/2000623

This work is licensed under a Creative Commons
Attribution-NonCommercial-ShareAlike 4.0
International License.



「シリーズ言語資源学」の刊行にあたって	iii
はじめに	v

第1章 | 学習者コーパスの意義と役割 (迫田 久美子) 1

1 はじめに：学習者コーパスの重要性	1
2 教育研究における意義と役割	2
2.1 教師の文法と学習者の文法	2
2.2 「わかる」と「できる」の乖離	3
3 言語研究における意義と役割	4
3.1 日本語の指示詞「コ・ソ・ア」の使い分け	4
4 習得研究における意義と役割	7
4.1 タスクの違いが習得に与える影響	7
4.2 横断的データと縦断的データの特性と役割	10
4.3 W-CoLeJaに見る言語変化	12
5 おわりに	15

第2章 | 日本語学習者コーパス開発の歴史 (石黒 圭・呉 楚琦) 18

1 学習者コーパスの特徴	18
2 学習者コーパスの展開	19
2.1 学習者コーパスが出現するまでの研究状況	19
2.2 技能から見た学習者コーパスの展開	20
2.3 母語から見た学習者コーパスの展開	22
2.4 習熟度から見た学習者コーパスの展開	24
2.5 そのほかの個性的な学習者コーパスの展開	25
3 総合学習者コーパス CoLeJa の誕生	27

第3章 | 学習者コーパスの設計 (石黒 圭) 31

1 学習者縦断コーパスを構想する	31
1.1 CoLeJa 計画の背景	31
1.2 S-CoLeJa の概要	32
1.3 調査協力者の代表性	33
1.4 調査協力校の選定	35
1.5 世界の学習者コーパスの規模	36
1.6 W-CoLeJa の規模	38
2 データを収集する	39
2.1 作文のテーマ	39
2.2 作文の文字数	41

2.3	作文執筆の方法	41
2.4	日本語能力の測定	42
2.5	背景情報の把握	44
3	データを加工・公開する	45
3.1	匿名化作業	45
3.2	形態素解析	45
3.3	誤用分析	46
3.4	印象評価	46
3.5	コアデータ	47
3.6	全文検索システム	47
3.7	公開とメンテナンス	48
4	まとめ	48

第4章 | 調査協力者に調査を依頼する

— 協力校との連携と学習者のサポート —

(烏日哲) 49

1	はじめに	49
2	調査協力校および現地教員との連携体制の整備	50
2.1	適切な調査協力校の選定	50
2.2	現地教員との丁寧なやり取り	51
2.3	連携を「見える化」する書面の作成	52
2.4	適切なデータ管理体制	53
3	学習者との信頼関係の構築	54
3.1	信頼関係の証になる同意書の作成	54
3.2	調査担当者の顔が見える調査説明会の実施	55
4	学習者への教育的還元	56
4.1	日本語教師による作文添削サービス	58
4.2	現役大学生との会話機会の提供	59
4.3	専門家による日本語学習支援	59
4.4	就活をサポートする参加証明書の発行	60
5	まとめ	61

第5章 | 入力インターフェイスの開発

— EssayLoggerとMoodle —

(井上 雄太) 64

1	分析しやすい執筆プロセスの記録	64
2	EssayLoggerの仕様	66
2.1	EssayLoggerの記入欄	67
2.2	多言語表示	68
2.3	文字数と注意	69
2.4	編集機能	69
3	収集できるデータ	70

4	協力者とのオンラインでのやりとりの方法	70
4.1	メールとクラウドストレージを用いたやりとり	72
4.2	告知 & アップロードを行うプラットフォームとしてのLMSの利用	72
4.3	Moodleと連携するWebアプリ版EssayLogger	74
5	まとめ	75

第6章 | 母語対訳データの意義と活用法 (呉 丹) 77

1	はじめに	77
2	母語対訳に関する先行研究	78
3	W-CoLeJaにおける母語対訳データの概要	79
4	母語対訳データの意義	82
5	母語対訳データを用いて可能となる分析	85
5.1	表記の分析—母語の表記習慣が与える影響を探索	85
5.2	語彙習得の分析—学習者の真意を把握するために	85
5.3	文法習得の分析—母語との対照から誤用の背景を探索	86
6	まとめ	89

第7章 | 日本語能力測定試験の作成とその妥当性の検証 —「かんたん日本語テスト」— (市江 愛・吉田 暁) 91

1	はじめに	91
2	先行研究と課題整理	92
2.1	日本語能力を測るための既存テストと日本語教育研究	92
2.2	新たな low-stakes test に求められる特徴	94
3	「かんたん日本語テスト」の開発	94
3.1	テスト設計方針と構成	94
3.2	表記規則と制限時間	95
3.3	CBT化と運用環境の整備	95
4	テスト開発のための予備的調査	96
4.1	予備的調査の概要(対象・実施環境)	96
4.2	項目困難度・弁別力分析	96
4.3	G-P分析	97
5	テストの多版化と信頼性・妥当性に関する調査	99
5.1	テストの多版化	99
5.2	調査の概要(対象・実施環境)	100
5.3	版間の同等性の検証	100
5.4	信頼性と妥当性の検証	101
6	まとめ	103
6.1	「かんたん日本語テスト」の利点	103
6.2	今後の課題と展望	103

第8章 | 学習者の背景情報の収集法

(田 佳月) 106

1 学習者の背景情報の収集意義	106
2 先行研究	107
3 W-CoLeJaにおけるフェイスシートの開発過程と概要	111
3.1 フェイスシートの開発過程	111
3.2 フェイスシートの概要	111
4 W-CoLeJaにおける学習者アンケートの開発過程と概要	112
4.1 学習者アンケートの開発過程	112
4.2 学習者アンケートの概要	114
5 W-CoLeJaにおける作文執筆情報シートの開発過程と概要	118
5.1 作文執筆情報シートの開発過程	118
5.2 作文執筆情報シートの概要	119
6 まとめ	120

第9章 | 学習者の学習環境と学習リソースを把握する

(田 昊) 123

1 学習環境及び学習リソースを把握する意義	123
2 先行研究	124
3 北京語言大学で実施された作文調査	124
4 教室内の学習リソース	125
4.1 授業カリキュラム	125
4.2 教材	128
4.3 予習・復習時間数	129
5 教室外の学習リソース	129
5.1 物のリソース	129
5.2 人のリソース	131
5.3 場のリソース	132
6 作文データから見る学習リソースの役割	133
6.1 作文の主題に与える影響	133
6.2 作文の文体と表現に見られるリソースの役割	133
6.3 作文の構造に見られるリソースの役割	134
7 作文執筆における学習リソースの使用	135
7.1 作文執筆におけるリソース利用割合	136
7.2 利用した辞書リソースの種類	138
7.3 参考にしたウェブサイトの種類	138
8 まとめ	139

第10章 | 学習者作文における形態素解析の課題と対応

(李 琦・朱 雅蘭・工藤 隆弘・横野 光) 141

1 はじめに	141
2 学習者作文における誤解析の実態	144

2.1	辞書未登録語による誤解析	144
2.2	ひらがな表記の多用による誤解析	145
2.3	誤用による誤解析	146
2.4	誤解析の影響	147
3	問題を解決するための工夫	148
3.1	形態論情報修正作業のワークフロー	148
3.2	ユーザ辞書構築による対応	152
3.3	辞書構築の課題	153
4	パターンマッチングに基づく自動修正	153
5	おわりに	154

第11章 | 日本語教師の誤用訂正はどこまで一致するのか

(中村 有里・江澤 実紀・高橋 香菜・Lai Thanh Hoa・石黒 圭) 156

1	学習者コーパスと誤用	156
2	添削タグの必要性	157
3	添削タグ付与における課題	158
3.1	他の学習者コーパスにおける添削タグの付与	158
3.2	添削タグを付与するうえでの課題	159
4	添削タグの付与に向けたパイロット調査	162
4.1	パイロット調査の概要	162
4.2	添削タグ付与データの収集方法	162
4.3	パイロット調査の結果	164
5	パイロット調査結果の分析	165
5.1	判断の揺れが生じる観点	165
5.2	判定者による判断の一致度	167
6	認定の揺れが生じやすい誤用	168
6.1	課題①誤用判断(誤用／正用)の揺れ	168
6.2	課題②誤用種類の判断の揺れ	169
6.3	課題③修正後表現の判断の揺れ	171
7	まとめ：解決法、今後の展望	172

第12章 | 生成AIはどのように学習者作文の誤用を訂正するのか

—誤用分析のアノテーション付与のための一考察— (加藤 敬介) 175

1	生成AIによる誤用訂正と学習者コーパス	175
2	先行研究	177
2.1	長単位と解析ツール	177
2.2	自動作文評価とレベル判定	177
2.3	生成AIによる訂正フィードバック	177
3	生成AI支援下での誤用訂正のステップ	177
3.1	ChatGPTのモデル選択	177
3.2	プロンプト設計の方針	178

3.3	ステップ1 長単位分割	178
3.4	ステップ2 日本語レベルの判定	181
3.5	ステップ3 制約付き訂正フィードバック	183
3.6	制約解除リライト	185
4	まとめ	187

第13章 | 印象評価情報の付与—クラウドソーシングの活用—

(本多 由美子・井伊 菜穂子・石黒 圭) 189

1	印象評価情報付与の意義	189
2	先行研究と印象評価情報付与における課題	190
2.1	先行研究	190
2.2	印象評価情報の付与における課題	191
3	印象評価調査の設計に向けたパイロット調査	191
3.1	パイロット調査の概要	191
3.2	検討課題①評価者間の評価尺度のずれを小さくできるか	192
3.3	検討課題②何名の評価者に依頼をするか	193
3.4	検討課題③項目別評価の評価項目に問題がないか	193
3.5	検討課題④総合評価の数値をどのように設定するか	194
3.6	検討課題⑤評価理由をどのように記述してもらうか	195
3.7	誤解析につながる表記・活用の誤りの事前修正	196
4	本調査で収集した印象評価データの概要	196
4.1	評価対象の作文	197
4.2	評価者数およびデータの収集方法	197
4.3	共通作文の評価	198
4.4	収集したデータ	199
5	印象評価情報の活用の可能性	202
6	まとめ	203

第14章 | W-CoLeJa 活用環境の構築

—全文検索システム『ひまわり』を用いて—

(山口 昌也) 205

1	はじめに	205
2	全文検索システム『ひまわり』	206
2.1	『ひまわり』の概要	206
2.2	XMLによるテキストデータの構造化	208
2.3	『ひまわり』の基本機能とXMLデータの構造との関係	210
3	活用環境の設計	210
3.1	W-CoLeJaの配布データのファイル構成	210
3.2	設計方針	211
3.3	XMLデータの構造の設計	212
4	活用環境の実現	213
4.1	『ひまわり』へのインポート	213

4.2 実現した機能の実行例	214
5 活用環境の利用例	218
5.1 日本語作文とその母語作文の文字数の関係	218
5.2 漢字率の経時変化	219
6 おわりに	220
コラム テキストファイルを形態素解析して、『ひまわり』にインポート	(山口 昌也)

第15章 | 学習者コーパスの公開と公開後の運用 (須賀 和香子) 223

1 はじめに	223
2 日本語学習者コーパス公開の意義	224
3 コーパス公開前の留意点	224
3.1 調査やコーパスの概要説明書	225
3.2 公開する学習者の産出データ	225
3.3 公開する関連データ	227
3.4 データの公開方法	228
3.5 権利関係の処理	231
4 公開後の責任	233
4.1 コーパス構築者の役割	233
4.2 コーパス利用者の心構え	236
5 おわりに	236
おわりに	239
索引	242
編者・執筆者紹介	247

