

国立国語研究所学術情報リポジトリ

表紙, 目次, 奥付, その他

メタデータ	言語: Japanese 出版者: 公開日: 2017-03-31 キーワード (Ja): キーワード (En): 作成者: メールアドレス: 所属:
URL	https://repository.ninjal.ac.jp/records/1008

電子計算機による国語研究 Ⅱ

——新聞の用語用字調査の処理組織

国立国語研究所

1969

電子計算機による国語研究 Ⅱ

——新聞の用語用字調査の処理組織

国立国語研究所

1969

刊行のことば

昨年、電子計算機による語彙調査に関する第一回の報告書として『電子計算機による国語研究』を刊行した。内容は、きわめて初歩的なものであったが、この方面の研究が全く新しい分野に属するものである関係上、それなりの役目を果たしたものと信じている。

電子計算機で処理する方法により新聞の語彙調査にとりかかったのは昭和41年であった。機械の導入と共に大量のデータを処理する仕事に取りかかったが、この過程で、数々の新課題を解決しなければならなかった。前回の報告書は、そのような課題のいくつかに対する最初の解答であった。

その後のシステム設計のしかたや語彙表作成の諸プログラムについての研究成果、さらに日本語の言語データを能率的に処理するためのいろいろな工夫について、研究の結果がある程度まとまったので、今回『電子計算機による国語研究』のⅡとして刊行する。われわれとしては全く新しい問題の開拓をしているわけで、各方面の方々の指導をお願いしたいと思う。

昭和44年3月20日

国立国語研究所長

岩淵 悦太郎

目 次

林 四郎	新聞語彙調査における層別とその意味……………	1
斎藤秀紀	電子計算機による語彙調査 ——主として長単位処理について……………	16
木村 繁	層別特徴語の判別……………	29
中野 洋	語彙調査の類別語彙表について……………	38
江川 清	「活用形処理」の自動化に関する一方式……………	55
石綿敏雄	COBOLによる漢字索引作成……………	80
野村雅昭	新聞使用漢字の試行的分析……………	89
田中章夫	漢字かなまじり文を全文カナ書き・ローマ字書きに 変換するシステムについて……………	107
石綿敏雄	構文解析自動化の研究 I ——CLからの構文論の見通し……………	139
木村 繁	構文解析自動化の研究 II ——文構造解析のプログラムからプログラム言語へ……………	175

CONTENTS

1. HAYASI Siro: CLASSIFICATION OF LANGUAGE DATA IN NEWSPAPERS, THEIR PRINCIPLES AND THEIR USE. — Four principles of classification: (1) literary genres, (2) topical fields, (3) types according information sources and (4) position of datum in the article. The uses of these classifications.
2. SAITO Hidenori: A SYSTEM OF THE WORD COUNT PROGRAM — for research into newspaper vocabulary by computer.
3. KIMURA Sigeru: DISCRIMINATION OF CHARACTERISTIC WORDS IN VARIOUS FIELDS — a report on method and result to find characteristic words in various fields of newspapers.
4. NAKANO Hiroshi: A TABLE OF THE CLASSIFIED VOCABULARY ON RESEARCH INTO NEWSPAPER VOCABULARY — a program for making vocabulary tables classified by part of speech, and by word origin.
5. EGAWA Kiyoshi: AN AUTOMATIC PROCESSING OF "CONJUGATION": — a report on an algorithm for translating conjugation into the head-word in verb and adjective by utilization of grammatical-information'.
6. ISIWATA Tosio: CHINESE CHARACTER INDEXES MAKING PROGRAM BY COBOL — a program for the research into newspaper vocabulary and Chinese characters.
7. NOMURA Masaaki: AN EXPERIMENTAL STUDY ON USAGE OF KANJI (CHINESE CHARACTERS) IN NEWSPAPER — an analysis from the viewpoint of statistical structure of kanji.
8. TANAKA Aki: A PROGRAM SYSTEM OF TRANSLITERATION, FROM KANJI TO KANA, AND FROM KANJI TO ROMAJI. — In the present day, Kanji (Chinese-characters) are indispensable to the Japanese language, but their number is very great; our computer cannot provide Kanji in on-line system. So, in processing Japanese language data, the kinds of jobs are limited, and efficiency of work is obliged to be low. The program system is one way to resolve these problems in software.

By means of table-look-up, we have contrived to transit Kanji

codes produced by Kanji-teleprinter into Kana (Japanese syllabary) codes, and into Romaji (Romanized Japanese). The test-run had a success percentage of 80% and the necessary time was 17 minutes, in processing the statement of a newspaper (8,000 characters). In this test-run, Kanji code (except proper-nouns) is automatically changed into Kana code only.

9. ISIWATA Toshio: A STUDY ON AUTOMATIC SYNTACTIC ANALYSIS OF JAPANESE SENTENCE, I—a synopsis from the point of view of computational linguistics—ÉTUDE SUR L'ANALYSE AUTOMATIQUE SYNTAXIQUE DE LA PHRASE JAPONNAISE, I.—Pour le traitement de l'information linguistique par la machine, il est nécessaire d'établir l'algorithme d'analyse syntaxique de la phrase.

L'auteur a proposé pour l'automatisation d'analyse syntaxique, de traiter non seulement du point de vue des formes linguistiques, mais aussi du point de vue de la sémantique.

L'auteur discute, ensuite, le syntagme, le classement des syntagmes par la fonction (coordination et subordination), la relation entre syntaxe et vocabulaire, la modalité de phrase et l'algorithme d'analyse.

10. KIMURA Sigeru: STUDIES ON AUTOMATIC SYNTACTIC ANALYSIS OF JAPANESE SENTENCE, II — a program for automatic segmentation of Japanese sentences, and its application to the syntactic analysis of computer language.

昭和44年3月

国立国語研究所

東京都北区稲付西山町

電話東京(900)3111(代表)

UDC 495.6:681.142

NDC 810

本書の市販品発行所

東京都新宿区市ヶ谷左内町39 (260)5291

株式会社 秀英出版

989

国立国語研究所刊行書

国立国語研究所年報

1～19 (昭和24年度～昭和42年度)

国立国語研究所報告

- 1 八 文 島 の 言 語 調 査
- 2 言 語 生 活 の 実 態 (秀英出版刊
¥300)
—白河および付近の農村における—
- 3 現 代 語 の 助 詞 ・ 助 動 詞
—用法と実例—
- 4 婦 人 雑 誌 の 用 語
—現代語の語彙調査—
- 5 地 域 社 会 の 言 語 生 活 (秀英出版刊
¥600)
—鶴岡における実態調査—
- 6 少 年 と 新 聞
—小学生・中学生の新聞への接近と理解—
- 7 入 門 期 の 言 語 能 力
- 8 談 話 語 の 実 態
- 9 読 み の 実 験 的 研 究
—音読にあらわれた読みあやまりの分析—
- 10 低 学 年 の 読 み 書 き 能 力
- 11 敬 語 と 敬 語 意 識
- 12 総 合 雑 誌 の 用 語 (前 篇)
—現代語の語彙調査—
- 13 総 合 雑 誌 の 用 語 (後 編)
—現代語の語彙調査—
- 14 中 学 年 の 読 み 書 き 能 力
- 15 明 治 初 期 の 新 聞 の 用 語
- 16 日 本 方 言 の 記 述 的 研 究 (明治書院刊
¥900)
- 17 高 学 年 の 読 み 書 き 能 力
- 18 話 し こ と ば の 文 型 (1)
—対話資料による研究—
- 19 総 合 雑 誌 の 用 字
- 20 同 音 語 の 研 究
- 21 現 代 雑 誌 九 十 種 の 用 語 用 字
—総記および語彙表—
- 22 現 代 雑 誌 九 十 種 の 用 語 用 字
—漢字表—
- 23 話 し こ と ば の 文 型 (2)
- 24 組 横 み の 字 形 に 関 する 研 究
- 25 現 代 雑 誌 九 十 種 の 用 語 用 字
—分析—
- 26 小 学 生 の 言 語 能 力 の 発 達 (明治図書刊
¥2,100)
- 27 共 通 語 化 の 過 程
- 28 類 義 語 の 研 究
- 29 戦 後 の 国 民 各 層 の 文 字 生 活
- 30-1 日 本 言 語 地 図 (1)
- 30-2 日 本 言 語 地 図 (2)
- 30-3 日 本 言 語 地 図 (3)
- 31 電 子 計 算 機 に よ る 国 語 研 究
- 32 社 会 構 造 と 言 語 の 関 係 に つ い て の 基 礎 的 研 究 (1)
—視族語彙と社会構造—
- 33 家 庭 に お け る 子 ども の コ ミ ュ ニ ケ ー シ ョ ン 意 識

国立国語研究所資料集

- 1 国語関係刊行書目(昭和17~24年)
- 2 語彙調査
—現代新聞用語の一例—
- 3 送り仮名法資料集
- 4 明治以降国語関係刊行書目 (秀英出版刊
¥300)
- 5 沖縄語辞典 (大蔵省印刷局刊
¥2,500)
- 6 分類語彙表 (秀英出版刊
¥900)

国立国語研究所論集

- 1 こ と ば の 研 究
- 2 こ と ば の 研 究 第2集
- 3 こ と ば の 研 究 第3集

国語年鑑

- (昭和29年版) (秀英出版刊
¥450)
- (昭和30年版) (秀英出版刊
¥600)
- (昭和31年版) (秀英出版刊
¥450)
- (昭和32年版) (秀英出版刊
¥480)
- (昭和33年版) (秀英出版刊
¥480)
- (昭和34年版) (秀英出版刊
¥500)
- (昭和35年版) (秀英出版刊
¥550)
- (昭和36年版) (秀英出版刊
¥800)
- (昭和37年版) (秀英出版刊
¥500)
- (昭和38年版) (秀英出版刊
¥950)
- (昭和39年版) (秀英出版刊
¥980)
- (昭和40年版) (秀英出版刊
¥1,100)
- (昭和41年版) (秀英出版刊
¥1,100)
- (昭和42年版) (秀英出版刊
¥1,100)
- (昭和43年版) (秀英出版刊
¥1,100)

高校生と新聞 国立国語研究所共著 (秀英出版刊
¥280)
日本新聞協会

青年とマスコミュニケーション 日本新聞協会共著 (秀英出版刊
¥280)
国立国語研究所

STUDIES IN COMPUTATIONAL LINGUISTICS

- HAYASHI Siro: Classification of Language Data in Newspapers, Their Principles and Use.
- SAITO Hidenori: A System of the Word Count Program.
- KIMURA Sigeru: Discrimination of Characteristic Words in Various Fields.
- NAKANO Hiroshi: A Table of the Classified Vocabulary on Research into Newspaper Vocabulary.
- EGAWA Kiyosi: An Automatic Processing of "Conjugation".
- ISIWATA Tosio: Chinese Character Indexes Making Program by COBOL.
- NOMURA Masaaki: An Experimental Study on Usage of Kanji (Chinese Characters) in Newspaper.
- TANAKA Akio: A Program System of Transliteration, from Kanji to Kana, and from Kanji to Romaji.
- ISIWATA Tosio: A Study on Automatic Syntactic Analysis of Japanese sentence I.
- KIMURA Sigeru: Studies on Automatic Syntactic Analysis of Japanese Sentence II.

1969

THE NATIONAL LANGUAGE RESEARCH INSTITUTE

1969

THE NATIONAL LANGUAGE RESEARCH INSTITUTE