

# 国立国語研究所学術情報リポジトリ

## 東アジア漢字圏における漢字コード統一のためのパイロット研究

メタデータ	言語: jpn 出版者: 公開日: 2021-06-11 キーワード (Ja): キーワード (En): 作成者: 斎藤, 秀紀, 横山, 詔一, 米田, 純子 メールアドレス: 所属:
URL	<a href="https://doi.org/10.15084/00003329">https://doi.org/10.15084/00003329</a>

# 東アジア漢字圏における漢字コード統一のための パイロット研究

国立国語研究所 情報資料研究部 電子計算機システム開発研究室

室長：斎藤秀紀 e-mail : saito@kokken.go.jp

主任研究官：横山詔一 e-mail : yokoyama@kokken.go.jp

研究員：米田純子 e-mail : jyoned@kokken.go.jp

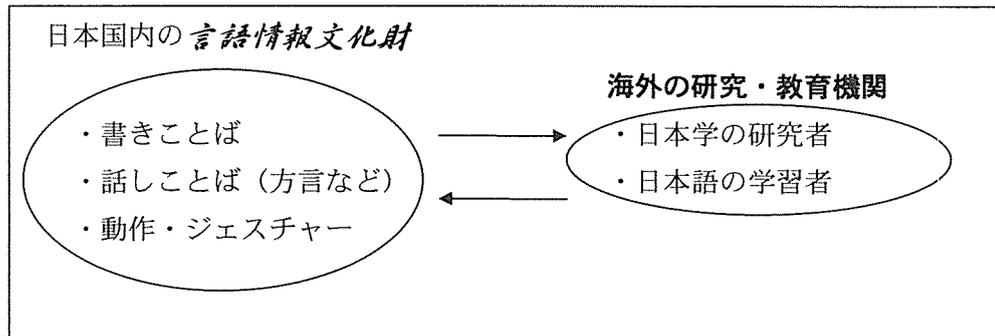
**Why** : どうしてこの研究が必要なのですか？

日本語が読み書きできる海外の友人とインターネットで情報交換してみたいと思ったことはありませんか？「インターネットは英語を喋る」と言われています。しかし、海外在住の日本人や日本語の分かる外国人と情報交換をする場合、やはり日本語を使える情報通信システムがどうしても必要になるでしょう。

**How** : どのようにするのですか？

私たちは、この問題に次の側面から取り組んでいます。

- インターネットに漢字を流通させるための**漢字コードの開発** → 世界共通の漢字符号
- 日本語における**漢字使用の実態調査** (どの漢字がどのくらいの割合で使われているのか？)
- 海外の日本語学習者に日本国内からインターネット経由で学習教材を提供するシステムの確立



**Mission** : 未来に何を残すのですか？

以上の問題を解決することによって、我が国に万葉集の時代から営々と蓄積されてきた日本語による情報を世界に発信することが可能となります。さらに、私たちは、これらの電子化された日本語に関するデータベースを**言語情報文化財**として未来に継承していく使命があると考えています。百年後の科学者が20世紀末の日本語の実態を知りたいと思ったとき、私たちの研究成果がきっと大きな手がかりになると確信しています。

<参考文献>

斎藤秀紀 (1994) 「大漢和辞典の検字番号に基づく構造化4バイトコードの提案」, 『情報処理学会論文誌』, Vol.35, No.6, 1119-1121.

横山詔一・笹原宏之・エリク＝ロング・米田純子ら (1998) 『新聞電子メディアの漢字：朝日新聞 CD-ROM による漢字頻度表』国立国語研究所プロジェクト選書 1, 三省堂