

コーパスを使った方言研究の開拓 — 『日本語諸方言コーパス (COJADS) モニター版』を使って— 大槻知世, 上村健太郎, カルリノ・サルバトーレ, 佐藤久美子, 中澤光平, 木部暢子

近年、各国における言語研究では、大量のデータの整備と大規模コーパスの構築が進み、新たな発展を見せている。一方、国内の方言研究においては、個別の方言ではいくつか音声が開示されている（日本放送協会（1959～1972）『全国方言資料』、杉藤美代子代表（1989～1992）科研費特定領域研究「日本語音声」研究成果 CD、国立国語研究所「日本の消滅危機言語・方言」データ公開のページ <http://kikigengo.ninjal.ac.jp/ほか>）。しかし、全般としてデータ整備が遅れており、大規模の通方言的な方言コーパスはこれまで存在しなかった。これに対し、発表者は諸方言を横断的に検索できる『日本語諸方言コーパス (Corpus of Japanese Dialects: COJADS)』の構築を進め、2019年3月にモニター版を公開した。本発表では、COJADSを活用することで、方言研究における新展開が期待できることを述べる。COJADSの基となるデータは、文化庁の『各地方言収集緊急調査』（1977～1985年に収録された自然談話資料）である。データは音声、方言テキスト、標準語テキストから構成され、時間情報により相互に紐付けされている。モニター版は、全国48地点、各地点約30分、合計24時間のデータよりなるコーパスである（2021年度には計75時間以上のデータを公開する予定）。検索方法は、①共通語から方言を検索する方法、②方言の文字列を検索する方法、③タグによる検索の3種類がある。発表では COJADS モニター版による検索結果をもとにした研究方法を提示するとともに、COJADSによるデモンストレーションを行う。

調査票調査（エリシテーション）は方言の体系的記述には欠くことができない。しかし、次のような欠点があることも指摘されている。①調査票の枠から漏れた形式を見つけることが難しい。②実際の使用場面に即した言語運用の情報が漏れることが多い。これを補うのがコーパスを含むテキストである（下地理則（2011）「文法記述におけるテキストの重要性」『日本語学』30(6): 46-59）。本発表では、音声、方言テキスト、標準語テキストを備えた COJADS モニター版を活用した研究実例として、4つの例：形態音韻論的研究、統語論的研究、フィラー研究、音調研究、を挙げる。