

『日本語歴史コーパス江戸時代編Ⅲ近松浄瑠璃』の公開

片山久留美, 小木曾智信, 上野左絵

本発表では、2020年3月に国立国語研究所より公開された『日本語歴史コーパス江戸時代編Ⅲ近松浄瑠璃』の概要や特徴等を報告する。

本コーパスは近松門左衛門作の世話物浄瑠璃24作品をコーパス化したもので、およそ25万語の規模となる。『日本語歴史コーパス』の江戸時代編は、これまでⅠとして洒落本、Ⅱとして人情本のサブコーパスを公開してきた。今回江戸時代前期の資料であるⅢ近松浄瑠璃が加わったことで、室町時代末期から幕末期にかけての資料が、細い線ながらもつながることになる。

浄瑠璃の台本は、「語り物」としての性質上、地の文と登場人物の発話部分の境界が明示されないが、本コーパスではテキストを地の文と会話に大別し、会話箇所には性別や身分などの位相も含めた話者の情報を付与した。これにより、話者の位相差を用いる歴史社会言語学的な観点からの研究利用も可能となっている。

また、近松浄瑠璃の特徴の一つに、掛詞・物尽しなどの修辞技法が用いられていることが挙げられる。こうした修辞上の複数の読み（語）についても検索対象とすべく、本文の同一箇所に複数の意味を持たせているものについて形態論情報の多重化（複線化）を行った。インターネット上の検索アプリケーション「中納言」から容易に検索することができる。

江戸時代編のサブコーパスが質・量ともに拡充されたことで、サブコーパス間の比較対照なども可能となってきた。本発表では例として、江戸時代編の3つのサブコーパスのうち、近松浄瑠璃のみに現れる語にどのようなものが見られるかを紹介した。本コーパスの公開、また多重の形態論情報検索などの機能拡張が、今後の日本語史研究の一助となることを願う。