

千葉大学『人文研究』第29号（2000年3月）抜刷

〔論 文〕

n グラム統計処理を用いた文字列分析に
よる日本古典文学の研究

— 『古今和歌集』の「ことば」の型と性差—

近 藤 みゆき

n グラム統計処理を用いた文字列分析 による日本古典文学の研究

— 『古今和歌集』の「ことば」の型と性差—

近 藤 みゆき

1 はじめに—検索利用から計量分析へ—

日本古典文学のデータベース環境は、ここ数年、急展開を遂げている。1990年に完成した長瀬真理の「日本語—英語対照「源氏物語」のテキスト・データベース」⁽¹⁾を先駆的業績として、その後、CD-ROMやオンライン⁽²⁾において、各種の古典籍データベースが公開されてきた。特に1999年には、4月に国文学研究資料館による日本古典文学本文データベース（実験版）試験公開の開始、7月に国文学研究資料館データベース古典コレクション『十一代集〔正保版本〕CD-ROM』、『源氏物語（絵入）〔承応版本〕CD-ROM』⁽³⁾の刊行、下半期には『角川古典大観 源氏物語』CD-ROMが刊行⁽⁴⁾と、大型の企画が相次いで完成・公開を見ており、国内の研究者にとっては古典文学研究のためのデータベース環境の基礎は、和歌および仮名散文の分野では、ほぼ固まったといってよい。こうした環境の充実は、近年の研究に着実に反映しており、特に和歌の分野では、1996年の『新編国歌大観CD-ROM版』⁽⁵⁾の刊行以後、歌ことばの研究、あるいは歌人同士の表現摂取の様相などに焦点をあてたような研究は、用例の博搜という点において、精密の度を加えてきた。歌風論・歌人論のいずれにおいてもデータベース化の促進がもたらした成果は計り知れないと言えるだろう。だが、しかし、用例検索を徹底した研究の増加が、一方で、和歌の表現研究にある種のステロタイプ化をもたらしつつあることも、現在、私を含めて多くの研究者が抱く所感に違いない。テキストデータベースの活用が、「用例を検索する」という、いわば研究の補助手段にとどまる限り、表現研究のあり方も固定的になり、