

「日本語情報処理特集号」の編集にあたって

長尾 真† 田中 穂積††

最近日本語情報処理に対する期待と関心が高まってきている。こうした動きは、計算機が我が国の社会に滲透するにつれて、我々の日常生活で生み出される多量の日本語情報を、直接「生」の形で計算機に投入、処理し、出力したいという要求の高まりと、近年の計算機技術の急速な進展に無縁ではないだろう。

現時点で、日本語情報処理という言葉は、幅広いスペクトルをもっている。漢字入出力の問題から、日本語理解システムや機械翻訳システムの問題に至るまで、さまざまなレベルの問題がある。本特集号では、できるだけ広い範囲をカバーする立場から、日本語情報処理の多様な側面と話題とを取り上げ、解説と展望を与えることを試みた。

本特集号で取り上げた話題は、三つに分けられる。第一は、日本語情報処理に関する比較的基礎的な問題と、その理論的枠組みを与えるための解説である。はじめに言語学の立場から、日本語をどのようにとらえるかを「日本文法の仕組み」で、また、意味構造をどのようにとらえるかを「言語理論からみた意味表現」で考察する。これら言語学者の論考は、情報処理技術者にとっても示唆に富むものであると思われる。膠着語の日本語では、形態素の認定と自動分割の問題が重要になる。「日本語の形態素分析」では、そのための手法、理論、問題点が指摘される。「日本語の構文解析」では、構文解析アルゴリズムに焦点をあてて、日本語構文解析手法の概要と指針とが与えられる。より高度な日本語理解システムを実現するためには、計算機科学の立場からの意味構造の解明と、一文を越えた談話理解の問題が重要になる。これらは「計算機から見た意味表現」、「談話理解の構造」で説明される。

第二の話題としては、実際に我が国で開発された日本語情報処理システムの実例紹介と、日本語情報処理応用システムの動向と問題点を探るために、「日本語によるデータベース照会」、「情報検索のための日本語解析」、「カナ漢字変換」、「JICST の実用的全自動漢字-カナ変換システム、K-KACS について」を取り上げた。そこでは、実用化を目指すための工学的な手

法と工夫とが説明されるだろう。機械翻訳は、自然言語処理のさまざまな側面を包含した総合的な応用システムの一つである。これまでの研究の概観と展望は、「機械翻訳」で説明される。

その他の話題として、「ドキュメンテーションと用語」、「日本語入出力装置」、「日本語テキストエディタ」、「日本語分析資料およびツールの調査」がある。漢字入出力手法については、漢字出力をも含めて、我が国で重要な問題になってきている。これは「日本語入出力装置」で説明される。この問題とも関連するが、日本語のテキストエディタは、欧米のそれと比較し歴史も浅く問題点も多い。これについては「日本語テキストエディタ」で説明される。情報管理、情報検索の立場からは、「ドキュメンテーションと用語」で実例、考え方が示されるだろう。最後の「日本語分析資料およびツールの調査」では、日本語情報処理を行うために開発、公開されているプログラムやデータが整理されている。蓄積性ある立場から研究を進める場合に、これらは役立つと思われる。

以上のように言語の理論的側面、計算機科学的側面、およびかなり実際の、実用的側面からの解説を行ったが紙頁の都合で本特集号に収録できなかった話題もいくつかある。たとえば、意味に近いレベルの情報から日本語文を生成する問題や、大量の言語データを分析するための計量言語学的手法の説明などである。

編者は、本特集号を機会として、日本語情報処理に対する基礎的研究の重要性がさらによく認識されるとともに、実用的な応用システムの実現を目指す努力が、我が国で一層活発化することを願っている。日本語を含む、いわゆる自然言語情報処理技術の延長には、将来の計算機技術にインパクトを与える要素が含まれている。日本語情報処理は本質的に長期の研究テーマである。しかも新しい研究分野である。今後この分野に多くの研究者の関心が高まることを期待したい。

最後に、本特集号に御多忙中にもかかわらず、短時間で原稿執筆を引き受けて下さった方々、および企画するに当たって御討議御協力をいただいた中田育男前編集担当理事と榎本肇編集担当理事をはじめとする編集委員の方々に厚くお礼申し上げる。(昭和54年9月6日)

† 京都大学工学部
†† 電子技術総合研究所