

Ngôn ngữ học Máy tính tiếng Việt và các Ứng dụng
Đình Điền, ddien@fit.hcmus.edu.vn

Tóm tắt

Cùng với sự tiến bộ vượt bậc của ngành Công nghệ Thông tin trong thời đại cách mạng khoa học công nghệ 4.0 hiện nay, ngành Ngôn ngữ học Máy tính (computational linguistics) – một liên ngành giữa Ngôn ngữ học và Tin học, ngày càng phát triển và có nhiều ứng dụng trong đời sống kinh tế, xã hội, giáo dục, khoa học - công nghệ, an ninh - chính trị,... Trong bài trình bày này, chúng tôi sẽ giới thiệu về Ngôn ngữ học Máy tính tiếng Việt (Vietnamese Computational Linguistics) và các ứng dụng của nó trong lĩnh vực giảng dạy và nghiên cứu về ngôn ngữ, cụ thể là: giảng dạy ngoại ngữ cho người Việt, dạy tiếng Việt cho người nước ngoài, biên soạn giáo trình, biên soạn từ điển, đánh giá năng lực ngôn ngữ, phát hiện đạo văn, nghiên cứu tiếng Việt, so sánh đối chiếu tiếng Việt với các ngoại ngữ, ...

Bằng cách khai thác các kho ngữ liệu (corpora) cùng với các công cụ tìm kiếm, so sánh, hiển thị, đánh giá trực quan của ngôn ngữ học máy tính, chúng ta có thể khảo sát, thống kê để tìm ra các quy luật trong mỗi ngôn ngữ, cũng như đối chiếu cách dùng từ theo ngữ cảnh trong ngôn ngữ đó với các ngoại ngữ khác. Từ đó, chúng ta có thể xây dựng giáo trình, từ điển với vốn từ và cấu trúc câu phù hợp cho từng cấp độ của người nước ngoài học tiếng Việt hoặc người Việt học ngoại ngữ. Ngoài ra, qua việc quan sát trực quan cách dùng trong thực tế từ các kho ngữ liệu này, người học sẽ tự “nghiệm” ra cách phát âm, cách dùng từ theo ngữ cảnh, các quy luật của tiếng Việt, quy luật chuyển ngữ, cách dùng từ phù hợp trong tiếng Việt ứng với một từ/ngữ trong ngoại ngữ tương ứng, nắm bắt được những tri thức, những kỹ năng ngôn ngữ hiệu quả hơn so với các cách tiếp cận truyền thống hiện nay.

Từ khóa: ngôn ngữ học máy tính, ngôn ngữ học tính toán, giảng dạy ngôn ngữ, nghiên cứu ngôn ngữ, xử lý ngôn ngữ tự nhiên.

Thông tin tác giả: PGS. Đình Điền (www.clc.hcmus.edu.vn/ddien/), chuyên ngành Tin học, Ngôn ngữ học, hiện là giảng viên cao cấp tại khoa Công nghệ Thông tin và kiêm giám đốc Trung tâm Ngôn ngữ học - Tính toán (Computational Linguistics Center) của Trường ĐH Khoa học Tự nhiên thuộc ĐHQG Tp.HCM.