

## Presentación

---

La lingüística aplicada ha realizado un uso cada vez más extensivo de la computadora como herramienta para el análisis de datos. Con todo, la relación que existe entre la computación y la lingüística va más allá, como pretende mostrar este número temático de *Estudios de Lingüística Aplicada*, en donde ya no sólo constituye una herramienta más, sino un componente esencial e incluso un objetivo de aplicación misma para el procesamiento y análisis del lenguaje.

El volumen reúne nueve trabajos en el ámbito de una interdisciplina entre la lingüística y la computación, que según los presupuestos, los matices, los objetivos y los fines, se le ha denominado procesamiento de lenguaje natural, lingüística computacional o ingeniería lingüística. La selección de estos trabajos se concentra en el ámbito de la lingüística de textos.

Como punto de partida, en el contexto de la morfología, Alfonso Medina y Cristina Buenrostro abren con un trabajo conjunto de la aplicación de lingüística cuantitativa a una lengua indígena. En esencia, se trata del cálculo de las medidas de contenido de información y de economía de signos para la segmentación automática de palabras a un corpus pequeño de una lengua maya, en particular para determinar los afijos de flexión verbal.

Le siguen dos artículos en el aspecto sintáctico. En el primero, César A. Aguilar muestra los beneficios del empleo de un sistema de cómputo e intercambio de datos especializado en la transcripción y extracción de datos lingüísticos enfocados a fenómenos del lenguaje infantil. Su objetivo se concentra en la localización y conteo de frases nominales y pronominales en función de correferencia anafórica con respecto a los actores principales de un evento.

Por su parte, Luis Pineda e Iván Meza realizan un proyecto de inteligencia artificial enfocado al desarrollo de sistemas conversacionales en lenguaje natural hablado en español. Como una parte central de dicha investigación, en este artículo presentan algunos de los aspectos de la gramática del español formalizada para el proyecto. En particular, se centran en el modelo computacional de la perífrasis y de su interacción con el sistema de pronombres clíticos.

En cuanto al terreno de la semántica, encontramos dos artículos relacionados. El primero, de Gabriel Castillo y Gerardo Sierra, busca obtener de manera automá-

tica grupos de palabras relacionadas semánticamente que pueden ser sustituidas unas por otras en un contexto sin cambiar su sentido. Revisan un algoritmo que inicia del alineamiento de pares de definiciones extraídas de diccionarios y que llega a la identificación de las palabras que guardan una relación sinonímica con base en una medida de similitud.

El segundo artículo de esta misma línea, de Héctor Jiménez, tiene como objetivo determinar diversas relaciones léxicas, principalmente las de sinonimia e hiper/hiponimia, entre acepciones de palabras que se encuentran en un corpus del español. Para ello, con base en las propiedades de los conceptos formales, utiliza una medida de la proporción de las características de los contextos de una palabra que aparecen en las características de los contextos de otra palabra.

Como una aplicación concreta en el ámbito de la lexicografía, Alexander Gelbuck y Grigori Sidorov trabajan en la construcción de diccionarios explicativos. Muestran el empleo de una herramienta computacional creada explícitamente para ayudar al lexicógrafo tanto para la construcción automática de muchos de los elementos de este tipo de diccionarios, como para el control de su calidad, a la vez de servir como medio para el desarrollo interactivo del mismo.

Por su parte, en el campo de la terminología, María Pozzi se ocupa de la identificación y análisis de términos mediante el uso de sistemas de extracción automática en textos de especialidad. Con base en medidas estadísticas y en el análisis de concordancias ordenadas alfabéticamente y por frecuencias, se concentra en identificar la inconsistencia en el manejo de términos y en corregirlos, evitando el uso de sinónimos y de formas polisémicas

Más adelante, como parte del análisis del discurso, Rodrigo Alarcón y Gerardo Sierra realizan un análisis de los patrones léxicos recurrentes que utiliza un autor para resaltar visual y gráficamente la presencia de un término que define. De esta manera, sientan las bases para desarrollar un posterior sistema de recuperación automática de los posibles conceptos, esto es, los términos y sus definiciones, en textos de especialidad.

Finalmente, en la misma línea de extracción de información de corpus, Luis Villaseñor, Aurelio López, Manuel Montes y Claudia Vásquez describen su aportación a los procesos de tratamiento automático de textos en español, y en especial de aquellos obtenidos por Internet. Ejemplifican con su trabajo realizado en la búsqueda de información relevante en periódicos sobre los desastres ocurridos en México.

En las conclusiones del último artículo, y como cierre de la presentación de este volumen de *Estudios de Lingüística Aplicada*, se enfatiza, a manera de resumen, que se cuenta con un campo interdisciplinario en el que se comparten las experiencias y conocimientos de los lingüistas y los informáticos. De esta unión resultan tres objetivos principales: la construcción de recursos lingüísticos útiles para ambas disciplinas, el diseño de herramientas para validar teorías lingüísticas propias del español de México, y la construcción de sistemas computacionales de uso práctico en los que sea un componente básico el lenguaje humano.

Cabe señalar que el contenido de los artículos es responsabilidad de los autores y que estos sólo siguieron las observaciones dadas por los dictaminadores. Debido a que se trata de un área nueva de investigación en México y que la mayoría de la literatura se encuentra en inglés, se llegan a utilizar extranjerismos, préstamos lingüísticos y, en algunas ocasiones, se proponen neologismos que a la fecha no se han estandarizado. Por ello, se puede encontrar que, por ejemplo, el segundo término del concepto *precision and recall* se traduce, en este volumen, como *evocación* o *recuerdo*.

Sólo resta agradecer a la editora responsable, Natalia Ignatieva, por abrirnos este espacio de diálogo y reflexión dentro de *Estudios de Lingüística Aplicada*. Asimismo, a los dictaminadores especializados en este tema: Albert Álvarez (Universidad de Sonora), Rosa Estopá (Universitat Pompeu Fabra), Fabio Rinaldi (Universidad de Zurich), Carlos Rodríguez Penagos (Universitat Pompeu Fabra), Randy Sharp (University of British Columbia), Arturo Trujillo (CRCE) y Elia Yuste (Universidad de Zurich). Finalmente, al mismo equipo técnico y de apoyo para hacer posible este número especial de ELA.

Gerardo Sierra  
Editor invitado