

FUNDAMENTOS Y LÍMITES DE LOS SISTEMAS DE VERIFICACIÓN AUTOMÁTICA DE LA SINTAXIS Y EL ESTILO

Javier Gómez Guinovart

Universidad de Vigo

uvifejgg@cesga.es

Tesis doctoral presentada el 28 de junio de 1996 en la Facultad de Filología de la Universidad de Santiago de Compostela. Director: Dr. Juan Carlos Ruiz Antón. Ponente: Dr. Guillermo Rojo Sánchez. Tribunal: Dr. Joseba Abaitua Odriozola, Dra. Irene Castellón Masalles, Dr. José María García-Miguel Gallego, Dr. Leonardo Gómez Torrego, Dr. Kepa Sarasola Gabiola.

Resumen:

En esta tesis se presentan los fundamentos y límites lingüísticos e informáticos de la verificación automática de la sintaxis y el estilo, prestando una atención especial a los problemas de corrección lingüística de los textos redactados en lengua castellana. El objetivo de la tesis consiste en establecer las bases teóricas, el ámbito de aplicación y las diversas estrategias de implementación informática de los sistemas de verificación lingüística desarrollados en el área del procesamiento del lenguaje natural.

La tesis consta de tres partes principales. La primera parte está dedicada a la caracterización del error gramatical y a su procesamiento informático. En

ella, se revisan los fundamentos lingüísticos teóricos de la distinción entre gramaticalidad y agramaticalidad, se examinan diversos criterios de caracterización y clasificación de los errores gramaticales, y se analizan las diferentes estrategias informáticas para su tratamiento automático, centrándose en los procedimientos elaborados por la lingüística de corpus para la detección de los errores sintácticos a partir del etiquetado de los textos, en las técnicas de atenuación de condiciones gramaticales propias del análisis gramatical automático resistente, y en los métodos de verificación sintáctica mediante reconocimiento de patrones.

La segunda parte de la tesis está dedicada a caracterizar la impropiedad del estilo y a investigar las posibilidades de su tratamiento informático. En ella, tras analizar las principales aproximaciones teóricas al estudio del estilo, se establece una caracterización de la impropiedad estilística ajustada a las posibilidades de su procesamiento automático, y se examinan las diversas estrategias informáticas aptas para esta aplicación, centrándose en los métodos lingüístico-cuantitativos empleados para valorar el nivel de dificultad de comprensión del sentido de un texto, en las estrategias de verificación fundadas en tipologías textuales estilísticas, y en las técnicas de detección basadas en la gradación de los rasgos estilísticos y en la determinación de sus umbrales de impropiedad.

Por último, en la tercera parte de la tesis, se presenta una aportación a la metodología de evaluación de los componentes de reconocimiento y corrección de errores de los sistemas de verificación de la sintaxis y el estilo.