



Universidad de Concepción  
Facultad de Humanidades y Artes  
Programa de Magíster en Lingüística

# Evaluación automática de coherencia textual en noticias policiales

Tesis para optar al grado de Magíster en Lingüística

Candidato: Sergio Hernández Osuna.  
Profesora guía: Dra. Anita Ferreira Cabrera.  
Co-tutor: Dr. Bernardo Riffo Ocares.

## Resumen

El siguiente trabajo aborda el tema de la evaluación automática de la coherencia textual. Su foco de estudio son noticias aparecidas en medios de prensa escritos, específicamente, las pertenecientes al dominio conocido como “Periodismo policial” o “Crónica roja”.

Para cumplir con los objetivos propuestos, se diseñó una herramienta, a nivel de prototipo, capaz de llevar a cabo la tarea de evaluación automática de la coherencia textual. Dicha herramienta se construyó en base a aplicaciones ya disponibles y de código abierto.

El prototipo utilizó para la medición automática de la coherencia textual la técnica de *Análisis Semántico Latente* y, como corpus de entrenamiento, empleó un conjunto de 1505 textos pertenecientes al dominio en cuestión.

Con el objeto de validar el prototipo se compararon las evaluaciones realizadas por la máquina con las efectuadas por evaluadores humanos. Para este fin se utilizó un conjunto de seis textos, extraídos desde los mismos medios de prensa desde los cuales se obtuvieron los textos que conforman el corpus de entrenamiento.

Como resultado se obtuvo que la evaluación realizada por el prototipo resultó ser equivalente a la efectuada por evaluadores humanos. Estos últimos se dividieron en dos grupos: uno formado por periodistas expertos en el dominio señalado; y otro formado por profesores de español con estudios de postgrado en lingüística, a quienes se considera como expertos de la lengua.