



Revista Signos

ISSN: 0035-0451

revista.signos@ucv.cl

Pontificia Universidad Católica de Valparaíso  
Chile

Olmos, Ricardo; León, José Antonio; Escudero, Inmaculada; Jorge-Botana, Guillermo  
Análisis del tamaño y especificidad de los corpus en la evaluación de resúmenes mediante el  
LSA: Un análisis comparativo entre LSA y jueces expertos  
Revista Signos, vol. 42, núm. 69, 2009, pp. 71-81  
Pontificia Universidad Católica de Valparaíso  
Valparaíso, Chile

Disponible en: <http://www.redalyc.org/articulo.oa?id=157013773004>

### Resumen

El Análisis Semántico Latente (LSA) es una sofisticada herramienta computacional de análisis semántico capaz de obtener una representación matemática del significado de las palabras o textos. LSA, entre otras aplicaciones, ha demostrado ser eficiente en la evaluación de textos. Esta herramienta adquiere la representación matemática de los textos analizando previamente un corpus lingüístico compuesto por documentos digitalizados. El principal objetivo de este estudio fue analizar qué propiedades han de tener distintos corpus lingüísticos (general, condensado, diversificado, y corpus de base) para que las evaluaciones de los resúmenes efectuadas por el LSA se parezcan lo máximo posible a las realizadas por 4 jueces humanos. Dichos resúmenes fueron elaborados por 390 estudiantes de Primaria, ESO y universitarios españoles. Los resultados indicaron que el tamaño de los corpus no tiene por qué ser tan generales ni tan grandes como los que se utilizan en Boulder (compuesto por millones de textos y más de un millón de palabras), ni tampoco demasiado específicos (menos de 300 textos y 5000 palabras) para que la evaluación que se desee hacer de ellos resulte satisfactoriamente eficiente.

### Palabras clave

Análisis Semántico Latente (LSA), resúmenes, evaluación del discurso, corpus lingüístico, estudiantes universitarios.

- ▶ Cómo citar el artículo
- ▶ Número completo
- ▶ Más información del artículo
- ▶ Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto