



## MUESTRA DE UNA METODOLOGÍA DE INVESTIGACIÓN EN DIDÁCTICA DE LAS CIENCIAS

**DOMÍNGUEZ HERRERA, M. (1) y VARELA CALVO, C. (2)**

(1) DIDÁCTICAS ESPECÍFICAS. UNIVERSIDAD DE LA LAGUNA [m\\_cdh@hotmail.com](mailto:m_cdh@hotmail.com)

(2) UNIVERSIDAD DE LA LAGUNA. [CVARELA@ULL.ES](mailto:CVARELA@ULL.ES)

---

### Resumen

El artículo muestra una metodología de investigación que se enmarca dentro de las técnicas de Lingüística Computacional para el Análisis Textual. En general, las técnicas utilizadas para este fin consisten en procedimientos léxicométricos y estudios estadísticos de los resultados obtenidos.

La metodología que se muestra en este artículo aplica este instrumento al ámbito de la Didáctica de las Ciencias y, para ello, se hace uso de una aplicación informática (*PAFE*) con características operativas específicas que lo diferencian de programas similares y que, en síntesis, permite el estudio de conceptos propios de ciencias y sus interrelaciones.

En definitiva, el planteamiento metodológico presenta una forma de articular el método estadístico con el marco teórico utilizado en la Didáctica de las Ciencias en general, y en el estudio sobre las concepciones, en particular.

---

### 1. Objetivos

Los objetivos que se pretenden alcanzar en el artículo se pueden concretar en:

Indicar un nuevo enfoque de las técnicas usuales de análisis textual al aplicarlas al ámbito de la Didáctica y en concreto, a la Didáctica de las Ciencias.

Mostrar el uso del programa informático PAFE, desarrollado como instrumento para aplicar las técnicas de la estadística textual desde esta nueva perspectiva.

## 2. Marco Teórico

En líneas generales, las técnicas de Lingüística Computacional empleadas para análisis textual se identifican con procedimientos léxicométricos seguidos de algún tipo de tratamiento estadístico de los resultados de obtenidos. En la mayoría de los casos, se cuenta con programas informáticos específicos capaces de realizar el recuento léxicométrico y combinar variadas operaciones estadísticas (Llisterri, 2006).

Así, el análisis estadístico de textos es una herramienta interdisciplinar, que integra Estadística, Lingüística o Informática, entre otros campos, y es utilizada tradicionalmente en el ámbito de las Ciencias Sociales (Lebart y Salem, 1994; García Izquierdo, 2000; Gattino y Miglietta, 2004).

En particular, este artículo muestra una metodología de investigación aplicada al ámbito de la Didáctica de las Ciencias. Así, se pueden obtener datos estadísticos de léxicos específicos de ciencias (Bécue, 1991; Etxeberría y otros, 1995; Ceballos y otros, 2000a).

Para este particular análisis, se ha utilizado el programa informático *PAFE* (Programa de Análisis de Frecuencias y Entornos), diseñado por el *GICEC* (Grupo de Investigación sobre Conceptos en la Enseñanza de las Ciencias).

Este grupo de investigación, realiza estudios centrados en el análisis de conceptos propios de cada ciencia (Ceballos y otros, 1998; Ceballos y otros, 2000a, 2000b). Uno de ellos ha concluido en la realización de una Tesis doctoral basada en el análisis de los conceptos propios del Sistema Solar que aparecen en libros de texto de enseñanza básica (Domínguez, 2006) así como en otros trabajos de investigación en la misma línea (Domínguez y Varela, 2005; Domínguez, 2007).

## 3. Metodología

Las técnicas empleadas para el análisis estadístico de textos se identifican con procedimientos orientados a lematizar el contenido, contar las ocurrencias o frecuencias de uso de las unidades léxicas, y después operar con algún tipo de tratamiento estadístico a partir de los resultados de tales recuentos, para establecer órdenes de frecuencias de aparición y relaciones entre unidades (Lebart, Salem y Bécue, 2000).

En los estudios del GICEC, el procedimiento de análisis textual, aunque obedece a los aspectos generales anteriores, presenta características propias, necesarias para utilizar la técnica en el estudio de conceptos en el ámbito de las ciencias.

Después de la *selección* de los textos, se procede a la *preparación* de los mismos. En síntesis, esta etapa inicial consiste en revisar los tiempos verbales (uniendo algunas formas compuestas, según criterios

establecidos), juntar las locuciones para que figuren como unidades semánticas, sustituir los pronombres por el nombre al que representan y diferenciar aquellos términos con distintas acepciones para que puedan ser identificados durante la siguiente fase. Los citados cambios no alteran el contenido del texto sino que, por el contrario, se corresponden con una de las fases más importantes del proceso, pues de su coherencia y rigor depende el grado de optimización de los resultados.

A esta preparación le sigue la fase de *procesamiento* de los documentos, en la que, con ayuda del programa PAFE, se seleccionan y archivan las unidades de la cadena textual en distintas bases de datos (*eliminar, dejar, cambiar a, verbos*) con las que poder realizar las futuras operaciones estadísticas y representaciones gráficas.

El programa permite ejecutar dos aplicaciones consecutivas: *Análisis de Frecuencias* y *Estudio de Entornos*. Con la primera se pueden obtener listados de términos ordenados en función de su frecuencia absoluta y relativa, así como el número de relaciones entre cada uno de ellos. Gracias a la segunda aplicación se determinan y miden las relaciones entre unidades léxicas, considerando el llamado *entorno* de las mismas.

Este *entorno* está integrado por  $n$  unidades a cada lado de la seleccionada, atribuyéndoles el programa un valor diferente, o peso, de acuerdo a su proximidad. De esta manera, se determina el valor de la relación de cada unidad con aquella que le precede y su inmediata posterior ( $\text{entorno}=1$ ); con las dos que le anteceden y las dos que le suceden ( $\text{entorno}=2$ ) o, de la misma manera, con las tres de su entorno ( $\text{entorno}=3$ ).

A continuación, se realizan los gráficos a cada fragmento de texto procesado en los entornos requeridos por el estudio a realizar. En estos entramados conceptuales en dos dimensiones, las unidades léxicas aparecen con distintos colores y unidas por diferentes tipos de líneas, en función de su frecuencia absoluta y valor de relación entre ellas, respectivamente.

Debe destacarse que el proceso descrito es cíclico, por lo que, en función de los resultados obtenidos, pueden efectuarse algunos cambios prefijados y repetir el proceso (pases) hasta advertir que estas modificaciones no produzcan variación en los resultados obtenidos.

#### **4. Conclusiones**

Este nuevo enfoque de las técnicas de análisis textual se convierte en una forma de aplicar el método estadístico dentro la Didáctica de las Ciencias para indagar, en última instancia, en el estudio sobre las concepciones. Asimismo, se comprueba que el programa informático PAFE representa un recurso de gran utilidad y adaptado a este tipo de estudios.

La gran versatilidad de esta metodología investigadora permite desarrollar distintas posibilidades de trabajo como, por ejemplo, centrar la atención en conceptos relevantes que aparecen en los textos de ciencias analizar el vocabulario empleado en ellos, así como observar las relaciones que se establecen entre los términos; el análisis de textos de libros antiguos y su comparación con otros actuales de la misma temática para determinar los conceptos obsoletos; establecer comparaciones entre textos de de distintos niveles de la enseñanza básica de una misma editorial; estudiar el lenguaje utilizado en las actividades planteadas en

algún tema de ciencias, etc.

### Referencias bibliográficas.

Becue, M. (1991). *Análisis estadístico de datos textuales: Métodos de análisis y algoritmos*. París: Cisia.

Ceballos, J. P., Galotti A., Pérez, D., López, C. Álvarez, J. M., Vadillo, A. y López, A. (2000b). PAFE –Programa de Análisis de Frecuencias y Entornos, en *Hacia el tercer milenio: cambio educativo y educación para el cambio. Actas XII Congreso Nacional y I Iberoamericano de Pedagogía*, 2, pp. 51-52.

Ceballos, J. P., Galotti, A., Santana, F. y Varela, C. (2000a). Waves in secondary education. *Physics Teacher Education beyond 2000. (PHYTEB) Internacional Conference*, p. 111.

Ceballos, J. P., Galotti, A., Varela, C. y Leal, P. (1998). Análisis de textos de Enseñanza Secundaria italianos y españoles: El sonido. *Actas 8º Encontro Ibérico para o Ensino da Física, Sociedade Portuguesa de Física*, pp. 174-175.

Domínguez, M. C. (2006). *Análisis de conceptos en el estudio de textos de enseñanza básica sobre el Sistema Solar. Contribuciones a la Didáctica de las Ciencias*. Tesis doctoral. Universidad de La Laguna.

Domínguez, M. C. (2007). Análisis conceptual en textos de enseñanza básica: El día y la noche. *Curriculum*, 20, pp. 133-146.

Domínguez, M. C. y Varela, C. (2005). Un ejemplo de análisis de textos de Enseñanza Primaria: El concepto del día y la noche. *Actas XXX Reunión Bienal de la Real Sociedad Española de Física*, pp. 256-257.

Etxeberría, J., García, E., Gil, J. y Rguez., G. (1995). *Análisis de datos y textos*. Madrid: Ra-ma.

García Izquierdo, I. (2000). *Análisis textual aplicado a la traducción*. Valencia: Tirant Lo Blanch.

Gattino, S. y Miglietta, A. (2004). Entre la marginación y la integración. Un estudio sobre el prejuicio étnico, orientación política y el empleo del lenguaje. *Boletín de Psicología*, 80, pp. 37-57.

Lebart, L. y Salem, A. (1994). *Statistique Textuelle*. París: Dunod.

Lebart, L., Salem, A. y Bécue, M. (2000). *Análisis Estadístico de Textos*. Madrid: Editorial Milenio.

Llísterri, J. (2006). *Herramientas para el análisis textual*. En: Universidad Autónoma de Barcelona. Disponible en:

. [Con acceso el 21-01-2008].

## CITACIÓN

DOMÍNGUEZ, M. y VARELA, C. (2009). Muestra de una metodología de investigación en didáctica de las ciencias. *Enseñanza de las Ciencias*, Número Extra VIII Congreso Internacional sobre Investigación en Didáctica de las Ciencias, Barcelona, pp. 1954-1958  
<http://ensciencias.uab.es/congreso09/numeroextra/art-1954-1958.pdf>