

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCION CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”

DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.



V SIMPOSIO INTERNACIONAL DE ESTUDIOS HUMANÍSTICOS

**La medición de la subjetividad como vía para la mejora de los
informes académicos**

Measuring Subjectivity for the Enhancement of Academic Reports

MSc. Humberto Miñoso Machado ¹, Dra. Juana Idania Pérez Morales ², Tania Machado Armas ³

1- MSc. Humberto Miñoso Machado. Universidad Central Marta Abreu de Las Villas, Cuba. humberto@uclv.edu.cu:

2- Dra. Juana Idania Pérez Morales. Universidad Central Marta Abreu de Las Villas, Cuba. juanap@uclv.edu.cu:

3- MSc. Tania Machado Armas. Universidad Central Marta Abreu de Las Villas, Cuba. taniama@uclv.edu.cu:

Resumen: La redacción de informes académicos no solo se limita a la aplicación de un plan o una plantilla preconcebida, sino también a la forma en que las relaciones entre las características objetivas y subjetivas se incluyen en el texto. Se ha observado que no todos los trabajos académicos cumplen con estos requisitos. Por lo tanto, este trabajo es un intento sucinto de aplicar la metodología de lingüística de corpus a la intervención pedagógica práctica dirigida a mejorar la calidad de los informes académicos. Este trabajo también refiere las principales tendencias en la redacción de informes al medir la subjetividad desde tres ámbitos diferentes: documento, oración y cláusula. La metodología de la lingüística del corpus se ha seguido desde una doble dimensión: cuantitativa y cualitativa. El conjunto de datos se ha organizado en tres paneles diferentes sobre la base de los cursos en los que se redactaron los informes. Todos los metadatos se han personalizado utilizando un esquema de definición de tipo de documento (DTD) para

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCION CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”

DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.



que solo se procesen los elementos de información deseados y, al mismo tiempo, se organicen los metadatos para un análisis cualitativo. Los resultados demuestran que la aplicación sistemática de las herramientas lingüísticas de corpus para medir la subjetividad en la redacción académica, puede mejorar la redacción de los informes académicos, ya que los profesores pueden prevenir y, al mismo tiempo, sugerir estrategias apropiadas para escribir no solo informes, sino diferentes géneros académicos en la redacción sobre la base de índices de referencia.

Palabras Clave: Corpus, codificación, escritura, subjetividad, sentimiento.

***Abstract:** Writing an academic report is not only limited to the application of a preconceived plan or template but also the way the relationship of objective and subjective features are included in the text. It has been observed that not all academic works meet these requirements. Thus, this work is a succinct attempt to apply Corpus Linguistics methodology to practical pedagogical intervention aimed at the improvement of the quality of academic reports. This work also refers to the main tendencies in writing reports by measuring subjectivity from three different scopes: document, sentence, and clause. Corpus Linguistics Methodology has been followed from a twofold dimension: quantitatively and qualitatively. The dataset has been organized in three different panels based on the courses in which the reports were written. All the metadata have been customized by using a Document Type Definition (DTD) schema so that only the intended items of information would be processed, and at the same time to organize the metadata for a qualitative analysis. The results prove that the systematic application of corpus linguistic tools to measure subjectivity in academic writing can improve the writing of academic reports since teachers can prevent, and at the same time, suggest appropriate strategies for writing not only reports, but also different academic genres based on benchmarked indexes.*

Keywords: Corpus, encoding, writing, subjectivity, sentiment

1. Introducción

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

**PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCION CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”**

**DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.**



Un tema recurrente en la agenda de las reuniones metodológicas ha sido la calidad de la redacción de los informes académicos escritos por los estudiantes en varios cursos. Los profesores han expresado la necesidad de reestructurar los contenidos de la asignatura Lengua Inglesa III del tercer año en términos de la forma en que se enseñan las habilidades productivas, con especial referencia a la redacción de informes de trabajos de cursos y tesis de grado. Aunque la documentación para redactar los informes académicos propone un formato, se ha observado que los requisitos del formato propuesto no siempre se cumplen. Por lo tanto, en este trabajo se decide como objetivo describir varios conjuntos de informes académicos para identificar dificultades comunes que nos ayuden a establecer una estrategia para mejorar la redacción de informes académicos sobre la base de descriptores de subjetividad a nivel de texto, de oración y de cláusula. Una revisión previa de varios trabajos académicos hizo que los investigadores de este estudio se formularan las siguientes preguntas:

¿Están todos los informes académicos correctamente orientados tanto objetiva como subjetivamente en todas sus partes constituyentes?

¿En qué medida las opiniones en las partes constitutivas de los informes académicos están bien distribuidas?

¿Es posible evaluar la calidad de los informes sobre la base de los descriptores de subjetividad que se observan en la estructura de los informes?

Hechos versus opiniones

La información textual generalmente se clasifica en dos tipos principales: hechos y opiniones. Los hechos son expresiones objetivas sobre entidades, eventos y sus propiedades. Las opiniones suelen ser expresiones subjetivas que describen los sentimientos, valoraciones o sentimientos de las personas hacia las entidades, los eventos y sus propiedades. Las opiniones son tan importantes que cada vez que tenemos que tomar una decisión queremos escuchar las opiniones de otros. Esto no solo es cierto para los individuos, sino también para las organizaciones (Liu, 2010).

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCION CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”



DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.

El análisis de sentimientos

El análisis de sentimientos, también conocido como minería de opiniones, es un problema desafiante del procesamiento del lenguaje natural o la minería de textos. Por lo tanto, presentamos la definición y los conceptos básicos relacionados con el análisis de sentimientos.

Clasificación de sentimiento y subjetividad

El análisis de sentimientos se considera un problema de clasificación del texto en el que un documento con opinión se clasifica en las siguientes categorías: documentos que expresan una opinión positiva, negativa o neutral. Además, a nivel de una oración o cláusula, la subjetividad es vista como objetiva o subjetiva.

El análisis de sentimiento basado en características determina los objetivos sobre los cuales se han expresado las opiniones en una oración, y luego determina si las opiniones son positivas, negativas o neutrales. Los objetivos son objetos, y sus componentes, atributos y características (Liu, 2010).

Opiniones proposicionales

Steven Bethard y su equipo definen una opinión proposicional como una opinión que aparece como una proposición semántica, que generalmente funciona como el complemento sentencia de un predicado (Bethard, Yu, Thornton, Hatzivassiloglou, & Jurafsky, 2004). Observen el siguiente ejemplo de proposición tomado de un informe académico, *argumentan que ayudar a los estudiantes a pasar de la narración del conocimiento a un enfoque de transformación del conocimiento a la escritura es un objetivo educativo importante y realista.*

2. Metodología

Se usa una metodología de enfoque mixto la cual combina la perspectiva cuantitativa (cuanti) y cualitativa (cuali) en un mismo estudio, con el objetivo de darle profundidad al análisis de preguntas de investigación complejas. Más que la suma de resultados cuanti y cuali, la metodología mixta es una orientación con su cosmovisión, su vocabulario y sus propias técnicas, enraizada en la filosofía pragmática con énfasis en las consecuencias de

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

**PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCIÓN CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”**

**DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.**



la acción en las prácticas del mundo real. Dentro de esta metodología este trabajo se apoya fundamentalmente en el modelo explicativo o aclaratorio el cual es utilizado para producir datos, que puedan ser usados de manera complementaria o con el fin de aclarar los hallazgos originales. En este tipo de estudio la indagación cuantitativa guía la cualitativa. (Alicia Hamui-Sutton, 2013)

Para los autores, la idea de este enfoque es aportar e integrar desde dos miradas o enfoques metodológicos diferentes el estudio en particular con el objeto de enriquecer y complementar el conocimiento a generar, realizando el análisis de textos desde la lingüística de corpus para desentrañar la subjetividad textual en los informes académicos. Mediante un estudio descriptivo de los datos cuantitativos resultante del procesamiento de los informes de los trabajos de cursos, se realizan inferencias cualitativas válidas para el conjunto de datos, a fin de tomar decisiones para mejorar su redacción.

La muestra consistió de 23 informes de trabajos de cursos que se procesaron en términos de subjetividad a nivel de texto, de oración y de cláusula. Este pequeño número de informes hizo que los investigadores adoptaran un enfoque cualitativo para encontrar respuestas a las preguntas formuladas.

En este estudio se utilizaron herramientas tales como el paquete de Laurence Anthony (Lawrence, 2014), UAM corpus para realizar cálculos y anotaciones (O'Donnell, 2011) y Corpus Presenter 14.5 (Hickey, 2017). Todos los archivos se convirtieron a formatos .txt. Posteriormente, los archivos txt fueron codificados en formato UTF8. Una vez que todos los archivos se codificaron en los formatos requeridos, se cargaron a la herramienta UAM Corpus Tool; por supuesto, fue necesario diseñar un esquema para el análisis del sentimiento en los tres niveles mencionados anteriormente (niveles de documento, oración y cláusula).

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCION CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”

DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.

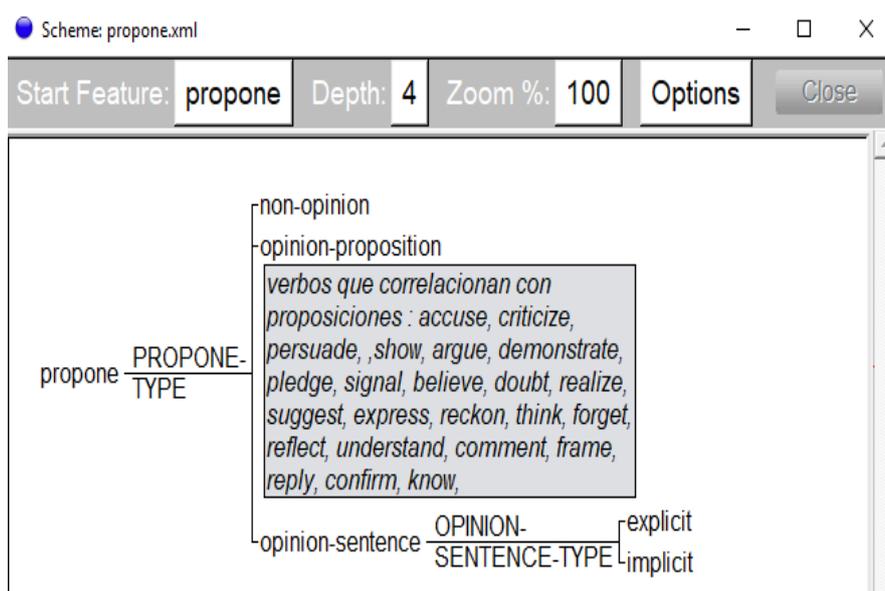


Ilustración 1 UAM Corpus Tool Scheme

Aunque los investigadores pudieron obtener el índice de subjetividad utilizando la herramienta UAM Corpus, también se requirió etiquetar todo el conjunto de datos en términos de las categorías propuestas para calcular la subjetividad por tipo de léxico; a fin de facilitar el análisis cualitativo. Para cumplir este objetivo, se utilizó la herramienta Corpus Presenter 14.5 (Hickey, 2017).

A nivel de texto

El conjunto de datos se ha organizado en cuatro paneles diferentes sobre la base de los cursos en los que se redactaron los informes. Todos los metadatos se excluyeron utilizando un esquema de definición de tipo de documento (DTD) para que solo se procesen los elementos de información deseados y, al mismo tiempo, se organicen los metadatos para facilitar el análisis cualitativo. Solo se procesan los datos a partir de la introducción del trabajo hasta las recomendaciones. Por lo tanto se excluyen del procesamiento la institución, el título del trabajo, los autores, el resumen, las referencias y los anexos.

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCIÓN CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”



DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.

Mediante la herramienta Corpus Presenter 14.5 se realizó el análisis de sentimientos. Esta herramienta brinda a nivel de texto el cálculo de las siguientes categorías de subjetividad para el léxico:

Negativo fuerte, negativo débil, neutral débil, neutral fuerte, positivo débil, positivo fuerte.

A nivel de oración

Los investigadores procesaron las oraciones que expresaban opiniones tanto positivas como negativas ya sea que la subjetividad esta explícitamente expresada o no. Para tal propósito se utilizó la herramienta UAM Corpus Tool. Se diseñó el esquema apropiado para tales cálculos.

A nivel de proposición

Para realizar la determinación de las proposiciones en cada texto fue necesario etiquetar los siguientes verbos que correlacionan con proposiciones: accuse, criticize, persuade, ,show, argue, demonstrate, pledge, signal, believe, doubt, realize, suggest, castigate, express, reckon, think, chastise, forget, reflect, understand, comment, frame, reply, volunteer, confirm, know, scream (Kingsbury, Palmer, & Markus, 2002). Las etiquetas se insertaron utilizando la herramienta Corpus Presenter 14.5.

3. Resultados y discusión

El estudio realizado muestra los siguientes resultados de subjetividad:

Resultados a nivel de texto

El estudio de opiniones a nivel de documento denota una orientación positiva (61%), mientras que la orientación negativa de los elementos lexicales del texto constituye tan solo el 39%. El grafico siguiente ilustra el comportamiento de los elementos léxicos a nivel de texto de forma desglosada. El léxico que marca la diferencia en cuanto a opiniones expresadas explícitamente en los informes representa el 8.41%.

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCION CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”



DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.

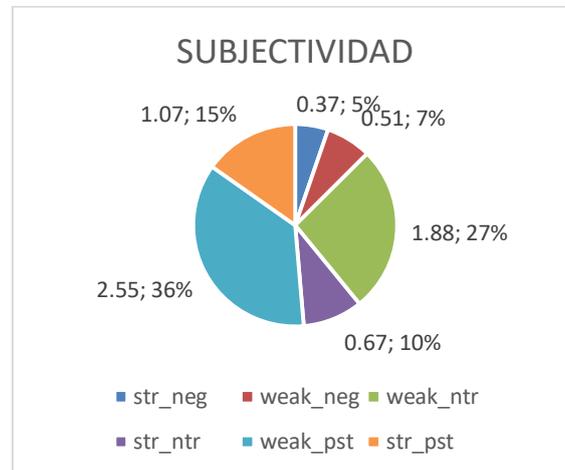


Gráfico1 Subjetividad Lexical desglosada

Resultados a nivel de oración

El muestreo realizado a 163 oraciones que contenían proposiciones ilustra la presencia altamente significativa de 85 oraciones donde las opiniones se expresan en las cláusulas proposicionales. Mientras que en 78 casos las opiniones se expresan a nivel oracional a pesar de la presencia de los verbos que correlacionan con proposiciones. Este resultado como muestra la tabla 2 establece diferencias altamente significativas.

Es evidente que no todos los informes tienen valores cercanos al valor de la media (15.7) y de los 23 informes se dispersan de la media 8. Estos informes fueron objeto de análisis cualitativo.

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCION CIENTIFICA INTERNACIONAL
“II CCI UCLV 2019”



DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.

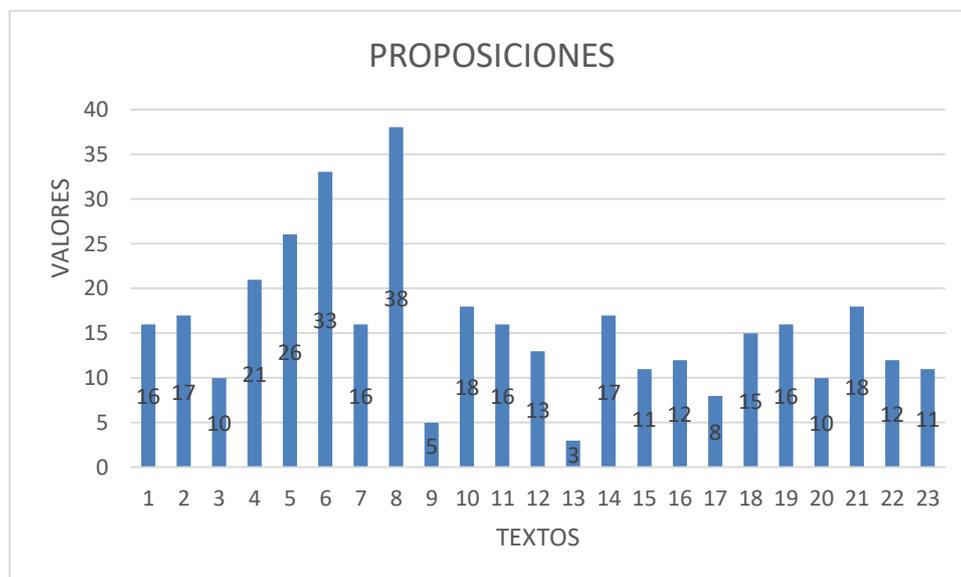


Tabla 1 Frecuencia de proposiciones por informe

	opinion-proposition		opinion-sentence					
Feature	Percent	N	Percent	N	T Stat	Signif.	ChiSqu	Signif.
PROPONE-TYPE	N=85		N=78					
non-opinion	0.0%	0	0.0%	0	0.00		0.00	
opinion-proposition	100.0%	85	0.0%	0	0.00		163.00	+++
opinion-sentence	0.0%	0	100.0%	78	0.00		163.00	+++
OPINION-SENTENCE-TYPE	N=0		N=73					
explicit	0.0%	0	32.9%	24	0.00		0.00	
implicit	0.0%	0	67.1%	49	0.00		0.00	

Tabla 2 Opiniones en las Oraciones y Proposiciones

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCION CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”

DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.



Las proposiciones a nivel de las partes constituyentes del informe se manifiestan de forma irregular lo que presupone un análisis cualitativo de los datos. Tan solo el 21.70% de las proposiciones ocurren en las secciones de resultados y discusión.

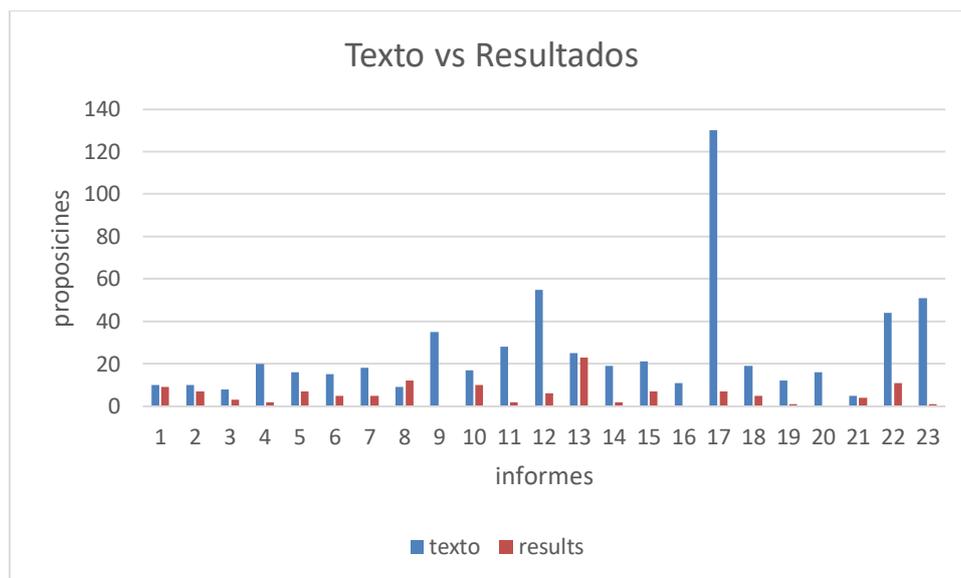


Tabla 3 Comparación de las proposiciones en las partes del informe

Discusión

No todos los informes académicos están correctamente orientados tanto objetiva como subjetivamente en todas sus partes constituyentes. Como se puede observar en la tabla 3, la mayoría de los informes muestran las proposiciones de opiniones en las secciones de resultados y discusión en una frecuencia muy baja con respecto al valor total en los textos. Lo ideal es que se comportaran como lo muestran los informes 1 y 2; o aún mejor en el informe 13 donde el número de proposiciones es equivalente en ambas partes del informe, tanto en el marco teórico como en el análisis de los resultados. Ejemplos de opiniones expresadas mediante proposiciones le brindamos a continuación y que aparecen en el marco teórico de los informes académicos:

On the other hand, others argue_PROP_VERB that although L1 and L2 writing share some common writing strategies, they are fundamentally different from each other (Grabe, 2001).

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCION CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”



DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.

These two authors argue PROP VERB that the choice of macrostructure is one of the most important decisions to be made by a specialized lexicographer... (Bergenholtz & Tarp, 1965).

Las opiniones están bien distribuidas en el marco teórico en la mayoría de los informes. Sin embargo, no todos los informes poseen la misma distribución en las secciones de resultados y discusión. Es importante señalar que las opiniones de oración prevalecen por encima de las opiniones de proposiciones en ambas secciones de resultados y discusión; lo que contradice el resultado a nivel de informe como se puede apreciar en la tabla 2. Es notorio ilustrar como están expresadas las opiniones de proposiciones en los informes mejor elaborados:

The statistics of the survey show PROP VERB that all of them need to use specialized dictionaries for reception purposes.

The students also suggest PROP VERB that professors should be more motivating...

33.3 % think PROP VERB academic writing plays an important role in foreign language learning and teaching.

También es necesario ejemplificar los casos donde las opiniones se manifiestan a nivel de oraciones:

*They **lack** the necessary socio-cultural knowledge to **fully** understand the texts they read. This **lack** of knowledge on several topics and **particularly** about personalities is **definitely** one of the factors that **directly** affect students when trying to understand and translate a given text.*

Sin embargo, como se resalta en los ejemplos dentro de las oraciones hay elementos léxicos que marcan la polaridad. Como lo han expresado Breck, Choi and Cardie es necesario considerar la intensidad con que se expresa una opinión además de quién o cuál es la fuente o portador de la opinión (Breck, Choi, & Cardie, 2007)

Definitivamente, el estudio demuestra que si es posible evaluar la calidad de los informes sobre la base de los tres descriptores de subjetividad medidos en este estudio en las partes constitutivas de los informes académicos. No obstante, resulta una limitación del estudio

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

**PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCION CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”**

**DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.**



en primer lugar el número limitado de informes (23) y en segundo lugar que no se consideró como parámetro de medición los conceptos de expresiones subjetivas directas (DSE) y el concepto de elementos de expresión subjetiva para los casos en que las opiniones no se expresaban en cláusulas proposicionales. (ESE) (Breck, Choi, & Cardie, 2007). En futuras acciones es recomendable incluir además mediciones sobre cohesión como las propuestas por la comunidad de investigadores de Cohmetrix (Grasser, MacNamara, Louwerse, & Zhiqiang, 2004).

4. Conclusiones

El estudio demuestra que la redacción de un informe académico no implica solamente la observación de requisitos estructurales de organización del texto sino que requiere también de un pensamiento sobre la expresión en el texto, considerando los factores de objetividad y subjetividad.

Las herramientas de la lingüística de corpus usadas para medir la subjetividad en la escritura académica, pueden convertirse en instrumentos para la intervención educativa pues contribuyen a mejorar la redacción de los informes académicos, y a su vez permiten a los profesores prevenir posibles errores, sugerir estrategias apropiadas para la escritura no solo de informes, sino de otros géneros académicos sobre la base de índices de referencia previamente concebidos.

Los resultados del estudio indicaron patrones de subjetividad desde tres niveles que constituyeron un paso de avance, pero es necesario implementar un análisis aún más minucioso de la polaridad en los textos. Esto último debe constituirse en prioridad de futuras acciones.

5. Referencias bibliográficas

- O'Donnell, M. (febrero de 2011). UAM CorpusTool Version 2.7. U.K. Obtenido de <http://wagsoft.com/CorpusTool>
- Alicia Hamui-Sutton. (2013). Un acercamiento a los métodos mixtos de investigación en educación médica. *Investigación en Educación Médica*, 211-216.

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCION CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”



**DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.**

- Bethard, S., Yu, H., Thornton, A., Hatzivassiloglou, V., & Jurafsky, D. (2004). Automatic Extraction of Opinion Propositions and their Holders. 1-8.
- Breck, E., Choi, Y., & Cardie, C. (2007). Identifying Expressions of Opinion in Context. *IJCAI-07: Proceedings of the Twentieth International Joint Conference on Artificial Intelligence*.
- Grasser, A. C., MacNamara, D., Louwerse, M. M., & Zhiqiang, C. (2004). *Cohmetrix: analysis of text on cohesion and language*. Memphis: University of Memphis.
- Hickey, R. (October de 2017). Corpus Presenter. Processing software for language analysis. (J. Benjamins, Ed.) Amsterdam:. Obtenido de <http://www.uni-due.de/~lan300/HICKEY.htm>
- Kingsbury, P., Palmer, M., & Markus, M. (2002). Adding Semantic annotation to the Penn TreeBank. *Proceedings of the Human Language Technology Conference*.
- Lawrence, A. (2014). AntConc (Version 3.4.3w). *[Computer Software]*. Tokyo, Waseda University, Japan. Obtenido de <http://www.antlab.sci.waseda.ac.jp/>
- Liu, B. (2010). Sentiment Analysis and Subjectivity. En N. Indurkha, F. J. Damerau, & (editors: N. Indurkha and F. J. Damerau) (Ed.), *Handbook of Natural Language Processing, Second Edition, (editors: N. Indurkha and F. J. Damerau), 2010* (págs. 1-38). Chicago: Department of Computer Science University of Illinois at Chicago.

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu