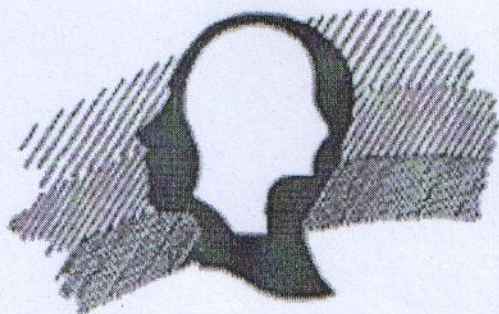


**CENTRO DE LINGÜÍSTICA APLICADA
MINISTERIO DE CIENCIA TECNOLOGÍA, Y MEDIO AMBIENTE
SANTIAGO DE CUBA**



ACTAS - II

**IX SIMPOSIO INTERNACIONAL DE COMUNICACIÓN SOCIAL
SANTIAGO DE CUBA, 24 - 28 DE ENERO DE 2005**

COAUSPICIADORES

- Universidad de
- Universidad Pec
- Dirección Provir
- Centro Cultural
- Teatro Heredia,
- Hotel Meliá – S
- Universidad de
- Knowledge Web
- Universidad del
- Instituto de Ling
- Universidad de
- Agencia de Viaje

COMITÉ ORGANIZADO

Presidenta de Honor:
Dra. Rosa Elena Sime
Ministra de Ciencia, T
República de Cuba

Pedro Beatón Soler
Eloína Miyares Bermú
Vitelio Ruiz Hernández
Leonel Ruiz Miyares
Zaida Valdés Estrada
Ena Elsa Velázquez C
Anton Nijholt
Nancy Cristina Álamo
Celia Pérez Marqués
Celia Álvarez Moreno
Mercedes Cathcart Ro
Ercilia Estrada Estrad
Martha Cordiés Jacks
Miladys Diodene Adan
Dieter Fensel
Iñaki Alegría Loinaz
Arantza Díaz de Ilarraz
Xabier Artola Zubillag
Xabier Arregi Iparragir
Asa Abelin
Jens Allwood
Nicoletta Calzolari
Daniela Ratti
Lucía Marconi
Claudia Rolando
Paola Cutugno
Raúl Ávila
Michael Zock
Sylviane Cardey
Peter Greenfield
Félix Rodríguez
Ruslan Mitkov

Edición y Composición:

Celia Álvarez Moreno
Jorge Pérez Bolaños
Laritz Hernández Rojas
Leonel Ruiz Miyares
Yilian Cortés Gutiérrez

© Todos los derechos reservados
© Sobre la presente edición:
Centro de Lingüística Aplicada, 2005
ISBN: 959-7174-05-7

ÍNDICE

Introducción /545

Índice alfabético de los autores principales de ACTAS-II /547

Lingüística Computacional /549

Anton Nijholt

Human and virtual agents interacting in the virtuality continuum /551

Bas Aarts y Sean Wallis

Recent developments in the syntactic annotation of corpora: a demonstration of ICE-GB and DCPSE /559

Borbála Katalin Benkő

Increasing the syntactical parse efficiency using "strong rules" /562

Bryan Bennett y Babis Theodoulidis

Gathering Together the Strands (Personal ontology in the engineering of group understanding and knowledge development) /567

Choy-Kim Chuah

Specialised Multilingual Databases: Motivation and Construction /571

Chris Reed

Preliminary results from an argument corpus /576

Christian Cave y otros

Un sistema de síntesis de habla en español de Venezuela /581

Evelio Sánchez Solís

"RAPITEX": generación automática de ejercicios de inglés en HTML /583

Gaél Djas y Elsa Alves

Language-Independent Informative Topic Segmentation /588

Germán Bordel y otros

Digital Resources for Automatic Speech Recognition of Broadcast News in Basque and Spanish /592

Helena Morgadinho

El Labelgram español: un sistema para el tratamiento automático de las ambigüedades lingüísticas del español /596

Lucía Marconi y otros

Hemeroteca telemática: instrumento para la organización y la circulación de la información /601

Luis A. Pineda e Iván V. Meza

A computational model of the Spanish clitic system /605

Luís Rogério da Silva

Aplicação de métodos estatísticos computacionais para análise de coesão textual /609

María Marilú Parra y Jacinto Dávila

Un modelo computacional para la generación de resúmenes automáticos de artículos científicos en español /613

Nicoletta Calzolari

Language Resources and Content Interoperability. Technical, strategic and political issues for a new generation of Language Resources /617

Octavio Santana Suárez y otros

Una aplicación para el procesamiento de la sufijación en español /623

Rita Marinelli y otros

Metonymic and Metaphorical Uses of Proper Names /630

Rita Marinelli y Adriana Roventini

Some Considerations about the Italian Maritime Lexicon Structuring /635

Stelios Piperidis y otros

Infrastructure for a multilingual subtitle generation system /640

Stephen Taylor

Porting the ARAMORPH arabic morphological system to a relational database /645

Sylviane Cardey y Peter Greenfield

Systemic linguistics with applications /649

Tassadit Amghar y Bernard Levrat

PIAGET (Plateforme Informatique d'Aide à la Génération d'Enoncés Textuels): Présentation générale /654

Ying Ding y Dieter Fensel

Semantic Web Powered Portal Infrastructure /659

Yoelvis González Martínez

Software educativo para el trabajo con la Lengua Española en el 4. grado de la escuela primaria /663

ÍNDICE ALFABÉTICO DE LOS AUTORES PRINCIPALES DE ACTAS-II

- Aarts, Bas; 559
 Agulló Ortiz, María; 846
 Alba Bermúdez, Ania M.; 1024
 Alves, Elsa; 588
 Amghar, Tassadit; 654
 Aranda Cintra, Belkis Luisa; 886
 Báez Viladón, Mercedes; 958
 Bejerano Llanes, Caridad; 890
 Beltrán Hurtado, Rolando; 974
 Benkő, Borbála Katalin; 562
 Bennett, Bryan; 567
 Betancourt Pulsan, Anselma; 881
 Bordel, Germán; 592
 Brooks Fuentes, Maritza; 1012
 Cabrales Mercerón, Reina Carmina; 1035
 Calzolari, Nicoletta; 617
 Cardey, Sylviane; 649
 Carrazana Piñero, Maikel; 954
 Carrera Morales, Mercedes; 1016
 Carvajal Jiménez, Vivian; 991
 Castillo Morales, Fidelina; 915
 Cave, Christian; 581
 Chen, Zhiyuan; 809
 Chuah, Choy-Kim; 571
 Cisneros Garbey, Susana; 978
 Coello Tissert, Juana Lidia; 728
 Correoso Céspedes, José A.; 725
 Cruz Rizo, Lorna; 743
 Da Silva, Luís Rogério; 609
 Dávala, Jacinto; 613
 Dias, Gaël; 588
 Díaz Jiménez, Jorge; 717
 Díaz Reyes, María Fidelia; 958
 Ding, Ying; 659
 Diodene Adame, Miladys; 776
 Domínguez Odio, Aníbal; 1047
 Escalona Vázquez, Irelis; 874
 Felipe de Souza, Angela Maria; 677
 Fensel, Dieter; 659
 Fernández Kindelán, Consuelo; 902
 Ferrer Miyares, Virgen Arelis; 806
 Ferrera Vaillant, Juan Ramón; 842
 Figueroa Corrales, Eufemia; 911
 Fonseca Arranz, Yaquelin; 920
 Fonseca Ramis, Josefina; 1073
 Franco Noa, Miguel Ángel; 917
 Frescura, Marina; 1053
 Gabardo, Maristella; 768
 García Arias, Nemis; 784
 Garner, Mark; 945
 Garriga Hernández, María Eugenia; 752
 Gell Labañino, Adia; 871
 Giannini, Antonella; 849
 Gómez Sánchez, Dario; 823
 Gonçalves, Eliane; 709
 González, Jorge; 720
 González Heredia, Katia; 739
 González López, Iselys; 714
 González Martínez, Juan; 950
 González Martínez, Yoelvis; 663
 Greenfield, Peter; 649
 Guerra Mercado, Gloria; 1040
 Guilarte Columbié, Herma; 926
 Hamze Ruiz, Isabel Luisa; 942
 Hendges, Danhela Cristina; 819
 Hernán Gómez Prieto, Beatriz; 680
 Herrera García, Ana Zoila; 1007
 Herrera Medina, Crispulo; 688
 Herrera Santana, Diana J.; 706
 Jurado Spuch, Adelaida; 669
 Kageura, Kyo; 861
 Lasagabaster, David; 906
 Leiva Armas, Daniel; 692
 Levrat, Bernard; 654
 Li, Han Z.; 1058
 López Carvajal, Haylín; 924
 López de Luzuriaga Gamboa, Ana; 734
 López Forés, Soraya María (2 trabajos); 1035; 1040
 Lorenzo García, Ana María; 877
 Luna Hernández, Alberto; 743
 Machado Simões, Adriana; 815
 Mantero, Miguel; 772
 Marconi, Lucia; 601
 Marinelli, Rita (2 trabajos); 630; 635
 Mármol Soñora, León; 1050
 Matos Columbié, Zulema de La Caridad; 996
 McKenna, Gwénaëlle; 712
 Mckenny, John; 945
 Meléndez de Vivas, Maritza; 961
 Merino Flores, Carlos Luis; 894
 Mesa Alonso, Milvia Ángela; 963
 Meza, Iván V.; 605
 Miyares Bermúdez, Eloína; 981
 Mora Joa, Juan José; 954
 Morelli, Mara; 747
 Morgadinho, Helena; 596
 Morgan Scott, Silvia Miriam; 792
 Muñoz González, Idalmi; 934
 Nijholt, Antón; 551
 Nocado Fernández, Pilar; 858
 Núñez Guzmán, Héctor; 1032
 Orozco Hechavarría, Oílda; 1032
 Ortiz, Delia Hilda; 701
 Parra, María Marilú; 613
 Paz Domínguez, Irela; 938
 Peña García, Alina; 874
 Peña Sánchez, Hortensia; 929
 Pérez Rodríguez, Juana María; 732
 Pettorino, Massimo; 849
 Pichs García, Rosalba Yeanet; 858
 Pineda, Luis A.; 605
 Piñeiro Maceiras, María Luisa; 756
 Piperidis, Stelios; 640
 Polanco Torres, Adelfa; 867
 Pupo Hernández, Nivia; 1028
 Quesada Lima, Adela M.; 674
 Quevedo Armaiz, Ned; 780
 Rams de La Cruz, René Antonio; 970
 Rayas Pérez, Maribel Eugenia; 1007
 Recino Pineda, Uvaldo; 797
 Reed, Chris; 576
 Rey Rivas, Pura de la Caridad; 732
 Ricalo Borrero, Vicente; 802
 Ruiz Hernández, Vitelio; 981
 Ruiz Rivero, José Vitelio; 1047
 Rodríguez Rodríguez, Beatriz; 1003
 Roventini, Adriana; 635
 Sánchez Avendaño, Carlos; 899
 Sánchez-Castro, Olga; 788
 Sánchez Cots, María de los M.; 762

Un modelo computacional para la generación de resúmenes automáticos de artículos científicos en español

Resumen¹

Los **propósitos generales del lenguaje escrito**² son básicamente los mismos independientemente de la lengua en que se genere. Así como la necesidad de ser comprendido por **otros es universal**, las sociedades necesitan transmitir su herencia de ideas y conocimientos a través del lenguaje. En este trabajo definimos el conocimiento como un flujo mixto de experiencia, valores, información contextualizada y visión experta que provee un marco de referencia para evaluar e **incorporar nuevas experiencias e información**. Desde este punto de vista, el conocimiento es producto de un proceso dinámico y como tal, se fundamenta en gran medida de la transmisión de la información. En la actualidad la gran cantidad de información que se genera en forma de texto, y que está disponible gratuitamente en **Internet** se incrementa día a día. Esto hace que se encaminen esfuerzos avocados a la búsqueda de soluciones en el procesamiento del lenguaje natural, lo que ha contribuido a **impulsar la investigación y el desarrollo de técnicas y aplicaciones que combinan tecnología con conocimiento lingüístico**, monolingüe y multilingüe, dando lugar a la llamada ingeniería lingüística. El objetivo principal de esta área de estudio es la aplicación del conocimiento de la lengua al desarrollo de sistemas informáticos capaces de reconocer, comprender, interpretar y generar el lenguaje humano en todas sus formas. Estas técnicas permiten "entender" el **texto** o el habla en lenguaje humano y desarrollar tareas que requieren de tal comprensión. El dictamen principal que rige nuestra conceptualización de lo que hemos definido como proceso de comprensión textual, se fundamenta en que los textos deben abordarse no sólo como **un conjunto de oraciones**, sino como un todo con sentido completo.

La formalización del proceso de comprensión está orientada al **reconocimiento de la superestructura**, específicamente al reconocimiento de las secciones retóricas de carácter persuasivo para obtener como resultado un resumen automático del texto. Para ello fue necesario **la delimitación del contexto**. Dicha limitación estuvo dirigida a la identificación de la estructura de Artículos de Investigación Científica (AIC), con el objeto de **enfocarnos en aspectos específicos de este tipo de literatura**. El modelo que explicita la superestructura de los AIC es representado a través del **modelo IMRD (introducción, métodos, resultados, discusión)** y el modelo Swales [Swales, 1990] para las introducciones. **Con la aplicación de estos modelos teóricos y con la ayuda de nuestro corpus de estudio que sirvió de banco de pruebas** fue posible sistematizar el proceso que siguen los humanos en la producción y consumo de textos especializados. A través de **la sistematización** y haciendo uso de lenguajes formales fue posible formalizar el proceso de reconocimiento de superestructura de los AIC. Sin embargo, aún se presentan **ciertas limitaciones en estos intentos por teorizar las estructuras textuales**. Al igual que las técnicas usadas en el Procesamiento del Lenguaje Natural (PLN), estos esfuerzos están encaminados a facilitar y mejorar la comunicación humana, por lo que se hace necesario la interdisciplinariedad para alcanzar este objetivo. Sin embargo, los avances en este campo están aún muy lejos de lograr **resultados** óptimos o equivalentes al rendimiento humano.

Al final de este trabajo planteamos una propuesta para extender las bondades de este programa, la cual se enfoca en la **formalización de las macrorreglas** semánticas para generar un resumen automático producto de la inferencia de las proposiciones explícitas en el texto.

El problema investigado

El texto **Resumen** que Ud. acaba de leer fue generado automáticamente a partir de un extenso documento (más de 120 páginas) dispuesto con una estructura documental similar a la de un artículo científico. Se trata de un ejercicio extremo de la lingüística computacional del texto que hemos conducido con la intención de validar un modelo de la superestructura de artículos científicos en Español [Parra, 2004].

La proliferación de información en formato digital motiva el trabajo. A menudo, para poder hacer uso eficiente de la información del texto, se requiere que la misma esté puesta en una cierta clase de formato estructurado. Extraer la información requerida de entre grandes volúmenes de documentos es, generalmente, un proceso manual costoso; por lo tanto, el texto en formato digital crea una necesidad de métodos de procesamiento para extraer la información automáticamente [Jacobs, 2000].

En este trabajo ahondamos en el problema de la formalización del proceso de comprensión para la generación de resúmenes, ajustado a un contexto específico de textos escritos con cierto propósito y estilo y con la intención de construir una plataforma teórica, computacionalmente factible, para el análisis de textos.

Objetivo de la investigación

El propósito general de este proyecto fue crear un modelo de la estructura de un texto que podamos usar luego para guiar un sistema de cómputo que aproveche esa estructura para analizar y sintetizar (resumir) un texto. Ese modelo puede entenderse como un refinamiento o una formalización de la teoría de estilos para escribir textos claros y comprensibles propuesta por J. Williams [Williams, 1990]. Ese trabajo de formalización comenzó con un

¹ Este resumen fue generado automáticamente (nivel 0) de una versión del documento en formato HTML.

² Las frases con este estilo de edición (negrita e itálica) son los tópicos identificados por el resumidor durante la generación del resumen automático.

modelo para resúmenes de textos, [Dávila-Contreras, 2002], que fue convertido en un programa lógico, [Contreras, 2002], para un computador que, de hecho, resume a través de una interfaz web en Internet³. En esta ocasión, la nueva aplicación tiene como propósito extender las capacidades de ese primer resumidor para la construcción de resúmenes en textos completos tales como los artículos científicos. Es decir, nos proponemos la formalización de los aspectos lingüísticos-semánticos que, presumimos, intervienen en el proceso de comprensión de textos completos dirigido a generar una versión sintetizada de la información que exponga los conceptos clave del tema a través de la construcción de secuencias de contenido.

La solución computacional que ofrecemos permitirá probar la potencialidad del modelo a partir del prototipo implementado. El lenguaje de desarrollo elegido, así como la metodología para el diseño, permite que a partir del uso del sistema, el usuario tenga facilidades para conseguir "respuestas" sobre la relevancia de los textos en el menor tiempo posible.

Basados en nuestro corpus de estudio, nuestro objetivo se centra en el hecho de representar el "valor estructural" [Van Dijk, 1983], es decir, la organización de las macroestructuras relacionadas con la superestructura y posiblemente con otras unidades cognitivas como el interés del usuario.

Superestructura

Representación formal que da cuenta de la coherencia textual. Una superestructura es un tipo de esquema abstracto que establece el orden global de un texto y que se compone de una serie de categorías, cuyas posibilidades de combinación se basan en reglas convencionales. Proporcionan una base de conocimiento relativamente invariable, que facilita la realización de inferencias causales, durante la lectura y comprensión de textos orales y escritos. Desde el punto de vista del productor, las superestructuras determinan jerárquicamente algunas de las decisiones iniciales de la planificación de sus discursos [Trujillo, 2000], se trata de esquemas prototípicos que han sido internalizados por los productores y receptores, y que "guían" el proceso de producción y comprensión.

Modelo IMRD

La estructura IMRD, refleja el patrón formal del artículo científico que fue adoptado en 1960, y aceptado por más de 350 revistas científicas en lengua inglesa. Cada sección responde a un objetivo particular. Por esta razón presentan diferentes funciones retóricas, lo que requiere de diferentes recursos lingüísticos para su realización. Es importante aclarar que nuestra intención no es realizar un análisis exhaustivo de los constituyentes retóricos de cada sección. El objetivo es identificar la estructura del modelo y algunos constituyentes retóricos que nos permitirán aplicar las reglas lógicas para la identificación de la estructura. La **sección introducción** es la sección retórica que motiva el estudio. Normalmente incluye una revisión de las investigaciones previas. En el aspecto lingüístico, se observa una alta frecuencia de construcciones nominales. En cuanto a la retórica, se observa una cuidadosa retórica para la versión final de la sección, ilustrada por la oración de "cierre", que diferencia y delimita el contenido del bloque. Similarmente la **sección discusión** es la sección retórica cuya función primordial es dar a conocer los resultados estadísticos en lenguaje no estadístico. Se hace referencia a los resultados de la investigación, por lo cual se observa una alta frecuencia en el uso de las construcciones nominales. Estas dos secciones están directamente relacionadas: en la primera es donde se abordan las preguntas del estudio y el objeto de la investigación, así como la revisión de la literatura y en la sección discusión/conclusiones se intenta dar respuesta a las preguntas formuladas en la introducción. Se exponen, discuten, argumentan, se contrastan y comparan las similitudes y divergencias de los hallazgos y se plantean los trabajos pendientes.

Modelo Swales

Normalmente las investigaciones en lengua inglesa abordan la sección introducción como un medio retórico de carácter persuasivo más bien que descriptivo, de gran riqueza léxica. Refiere a una reconstrucción distanciada del análisis original razonado para el estudio. Se ubica entre la tradición y la innovación. Es el medio donde se manifiesta una especie de negociación comunicativa entre el autor, el lector y el conocimiento para definir los términos del texto. En esta sección se encapsula el objetivo "problema-solución", característica de los textos argumentativos. Investigadores en sus escritos necesitan continuamente dirigirse a un contexto intelectual. Más específicamente, los investigadores desarrollan sus metas, capacidad, problemas y criterios de evaluación, como procesos que derivan y funcionan dentro de una disciplina. Esta división interna de la sección introducción permite al autor dirigir su investigación a un área particular, ubicándola en un marco contextual a través de la revisión de la literatura.

Swales concluyó que las introducciones usualmente consisten en una combinación de movimientos que responden a la noción de género. Así crea el modelo CARS (Create a Research Space) [Swales, 1990], para ubicar la investigación en un contexto dado. A través de movimientos y pasos entre movimientos, el modelo permite detallar información específica en la introducción de los AIC. La retórica que se observa siguiendo este modelo, facilita la identificación de unidades de información independientemente del campo, grupo, escuela de investigación o comunidad académica dada.

El concepto de persuasión discursiva de una comunidad representa una importancia potencial como medio que orienta la investigación, e influye en los receptores a través de una elaborada retórica.

El nuevo género discursivo propuesto por Swales, ha discutido la estructura de pirámide invertida para la sección introductoria, cuya finalidad es la organización de la información en orden de importancia descendiente, de lo general a lo particular. Corresponde a la segmentación del texto en partes, definiendo unidades funcionales para los diferentes niveles del texto.

³ Página con interfaz web para el resumidor <http://chama.cecalc.ula.ve/~hyelitza/resumidor.html> o <http://cesimo.ing.ula.ve/INVESTIGACION/PROYECTOS/GIL/resumidor.html>

Para
espa
inter
Desp
rigen
léxic
argu
ingle
Proc
Para
llega
de m
El m
forma
Estos
texto
lector
En el
cuen
mode
La fig
comp

La fig
partic
?- mo
|---m
|

A = r
L = [

Para efectos de este proyecto se revisará en detalle el movimiento3 de los artículos de investigación escritos en español, más específicamente la presencia/ausencia del paso 1B. Este movimiento mantiene una estructura interna de apertura y cierre. Nos concentramos en el paso de cierre de dicho movimiento.

Después de analizar AIC escritos en español, provenientes de diferentes áreas de estudio, concluimos que se rigen por los mismos parámetros retóricos y esquemáticos de los AIC en lengua inglesa. El uso terminológico y léxico difiere dependiendo del campo de investigación. Sin embargo, se mantiene la conceptualización lógico-argumentativa del modelo IMRD y el modelo Swales, modelos planteados para las producciones en lengua inglesa.

Procesamiento computacional

Para modelar y simular el proceso de comprensión nos apoyamos en la teoría que considera dicho proceso. Para llegar a este punto fue de gran valor el corpus de estudio. Constituyó la principal fuente de recuperación de datos de manera sistemática, convirtiéndose en nuestro "banco de pruebas" de hipótesis o análisis lingüísticos.

El modelo general consta de cuatro componentes teóricos claves: modelo textual IMRD, el modelo Swales, formato HTML y el resumidor.

Estos componentes intentan representar los pasos que siguen los humanos en el proceso de comprensión de textos. En este caso, recordemos que estamos considerando textos especializados, donde suponemos que el lector cuenta con experiencia previa en el reconocimiento de la organización textual.

En el PLN se establece una relación entre el lenguaje, sistema simbólico y estructuras de pensamiento. Para dar cuenta de estas relaciones identificamos las microestructuras, la macroestructura y la superestructura en el modelo a través de los componentes que conforman el sistema.

La figura 1 muestra una forma simplificada (pero muy similar a la que usa el programa) del modelo gramatical-computacional del llamado movimiento 3 en el modelo lingüístico de Swales:

```

sentencia --> movimiento3
movimiento3 --> frase_de_proposito, copula, verbo_proposito.
frase_de_proposito --> articulo, proposito, de, trabajo.
articulo --> [el].
proposito --> [proposito].
de --> [de, este].
trabajo --> [trabajo].
copula --> [es].
verbo_proposito --> [analizar].
  
```

Fig. 1. Gramática de movimiento3

La figura 2, por su parte, ilustra el funcionamiento del programa que construye un árbol identificando un caso particular del movimiento 3 en una oración.

?- movimiento3(A,[el,proposito,de,este,estudio,es,investigar],L), a(A,0).

```

|---mv3
  |---fdp
    |---art
      |---el
        |---pro
          |---proposito
            |---d
              |---de
                |---este
                  |---trab
                    |---estudio
  
```

A = mv3(fdp(art(el), pro(proposito), d(de, este), trab(estudio)))

L = [es, investigar]

Fig. 2. Salida del programa Prolog para movimiento3 paso1B

Resultados de los experimentos con el modelo computacional.

Hemos conducido una serie de experimentos comparando resúmenes producidos por los autores de artículos científicos con los producidos por el programa. La figura 3 muestra un caso como ejemplo.

Resumen del autor	Resumen automático
OBJETIVO DEL TRABAJO: Cuantificar el papel que la variable temperamento imprime en la involución (proceso degenerativo de cada ser), y cualificar los niveles de ansiedad y depresión, y sus implicaciones en la construcción de la realidad de las personas).	OBJETIVO DEL TRABAJO: De esta manera se ha considerado importante en este trabajo establecer los niveles de ansiedad, depresión, temperamento la distinta adaptación de la mujer en este período vital en función del tipo de menopausia, además de establecer nuevos planteamientos a la hora de atender a las pacientes que se encuentren en este período.

Figura 32. Objetivo Experimento3-Medicina

Conclusiones

Asumiendo un modelo textual (IMRD) y construyendo sobre un resumidor de párrafos, produjimos un modelo computacional que identifica la superestructura y extrae los bloques retóricos que se consideran importantes para la creación de resúmenes (en la comunidad científica de habla hispana). Ese modelo constituye un aporte transdisciplinario y de unificación de esfuerzos en el estudio del lenguaje escrito.

Referencias

- Contreras, 2002] Contreras, H. *Una Técnica para la Extracción Automática de Resúmenes Basada en una Gramática de Estilo*. Tesis de Grado de Maestría. Universidad de Los Andes. Mérida-Venezuela. 2002.
- Contreras-Dávila, 2001] Contreras, H. y Dávila, J. *Procesamiento del Lenguaje Natural Basado en una "Gramática de Estilos" para el Idioma Español*. Memorias de la Conferencia Latinoamericana de Informática, CLEI. Mérida-Venezuela. 2001.
- Dávila-Contreras, 2002] Dávila, J. y Contreras, H. *Una gramática de estilos para resumir textos en Español*. Memorias del XVIII Congreso de la Sociedad Española para el Procesamiento del Lenguaje Natural, SEPLN. Septiembre 2002.
- Parra, 2004] Parra, Marilú. *Un Modelo Computacional para la Generación de Resúmenes Automáticos de Artículos Científicos en Español*. Tesis de Maestría. Postgrado en Modelado y Simulación. Universidad de Los Andes. Venezuela. Julio 2004.
- Swales, 1990] Swales, John M. *Genre Analysis English in Academic and Research Settings*. Cambridge University Press. 1990.
- Trujillo, 2000] Trujillo Sáez, Fernando. *Propuesta de Utilización de Modelos Textuales Definidos Culturalmente para la Enseñanza de la Lectura en Inglés y Español*. Facultad de Educación y Humanidades de Ceuta, Departamento de Didáctica de la Lengua y la Literatura. Universidad de Granada. 2002.
- Van Dijk, 1983]. Dijk, Teun van. *Texto y Contexto*. Cátedra. Madrid. 1983.
- Willians, 1990] Williams, J. *Style: Toward Clarity and Grace*. The University of Chicago Press. Chicago and London. 1990.