

**GRETEL LIZ DE LA PEÑA SARRACÉN\***  
**ADRIAN FONSECA BRUZÓN\***  
**REYNIER ORTEGA BUENO\***  
**MARÍA PELÁEZ BRIOSO\*\***  
Centro de Estudios de Reconocimiento de Patrones y Minería de Datos\*  
Empresa de Desarrollo de Aplicaciones, Tecnología y Sistemas (DATYS)\*\*  
Santiago de Cuba, Cuba  
{gretel | adrian | renyier.ortega}@cerpamid.co.cu, maria.pelaez@datys.cu

## ***Comprensión del lenguaje natural en sistemas de diálogo***

### **Introducción**

Los sistemas de diálogos, o sistemas conversacionales, son programas informáticos desarrollados con el objetivo de brindar servicios específicos a los seres humanos a través de diálogos coherentes y naturales. De esta forma, facilitan la interacción natural entre una computadora y un usuario [1-2]. Estos sistemas han sido usados en múltiples tareas como la planificación de itinerarios de vuelos [3], la gestión de información en terminales de trenes [4] e información acerca del clima [5]. La principal ventaja radica en que permite el desarrollo de aplicaciones eficientes; donde los diálogos con los usuarios son realizados mediante sistemas de diálogos, mientras que los especialistas humanos se dedican a realizar tareas más complejas [6].

En su funcionamiento, un sistema de diálogo debe realizar un conjunto de acciones básicas que se repiten cíclicamente como respuesta a cada intervención del usuario. En forma general, el sistema extrae el significado del texto, es decir, comprende qué información es útil en el dominio del sistema y luego decide qué acción o acciones deben realizarse como respuesta. Desde el punto de vista de diseño, un sistema de diálogo tiene una estructura modular en la que cada módulo o componente se encarga de un conjunto de determinadas tareas.

Existen diferentes técnicas para desarrollar cada una de las funcionalidades dentro de los componentes de un sistema de diálogo, lo que conduce a diferentes propuestas de sistemas; sin embargo, todos tienen características comunes. Una de las más importantes consiste en que el sistema de diálogo debe alcanzar un cierto grado de comprensión del enunciado recibido por el usuario. Entendiéndose por comprensión a la generación automática de una determinada representación semántica del enunciado. Esta representación constituye la interpretación del texto, que contiene de manera resumida y estructurada la información necesaria por el sistema en la toma de decisiones de forma rápida y sencilla. Para lograr esta característica los sistemas de diálogo son dotados de un componente, conocido como módulo de "Comprensión del Lenguaje Natural" (*NLU*, por sus siglas en inglés), que es el que realiza esta tarea de transformar el texto brindado por el usuario a la representación semántica correspondiente [7].

Las propuestas de diseño para el componente *NLU* se dividen en dos grandes aproximaciones: la comprensión guiada por los datos [8] y la comprensión basada en conocimiento [9]. En el primer caso, el proceso de comprensión utiliza modelos estadísticos estimados a partir de datos (flujos de diálogos etiquetados). Sus mayores ventajas radican en el incremento de la robustez de los sistemas, permitiendo un mayor control sobre sus acciones ante situaciones en el diálogo no previstas; y la facilidad para lograr una gran cobertura del dominio, teniendo en cuenta gran cantidad de situaciones posibles. Sin embargo, para construir los modelos es necesaria una gran cantidad de datos de entrenamiento, introduciéndose la problemática de la escasa disponibilidad de tales datos. Además, ellos deben ser bien representativos para el dominio del problema, y deben contar con una alta calidad de anotación. Por el contrario, en la comprensión basada en conocimiento se definen un conjunto de reglas para la generación de la representación semántica. La generación de tales reglas implica grandes esfuerzos, pero es una variante muy utilizada para componentes *NLU* de sistemas de diálogos de dominios específicos, donde solo se cuenta con algunos documentos de referencia en lugar de muchos datos para entrenar un modelo. Por esta razón, hemos empleado esta aproximación para el diseño y desarrollo del componente de comprensión para un sistema de diálogo para asistir a pacientes que se encuentran en un proceso de pérdida de peso a través de un método de control de nutrientes.

En este artículo se describe la estructura general del componente *NLU* de un sistema de diálogo cuya única forma de comunicación (entrada y salida) está basada en texto. Además, se definen las características esenciales de las técnicas basadas en reglas, y en particular se detallan las estrategias utilizadas en la propuesta de nuestro *NLU*. Por último, se muestra que los resultados son correctos y alentadores a través de un análisis realizado a la propuesta del componente, y se exponen las conclusiones.

### **Componente para la comprensión de textos en un sistema de diálogo**

El proceso de comprensión en el *NLU*, puede entenderse como un cambio en el lenguaje de representación, del lenguaje natural a una representación semántica, de forma que se mantenga el significado del mensaje. Por ejemplo, en el dominio del control de nutrientes, para el texto: "*¿Por cuál alimento puedo sustituir el chocolate?*", debería obtenerse una representación con los datos más relevantes, en este caso, que se realiza una petición de sustitución de un alimento, en este caso el chocolate.

La definición de la representación semántica se basa, en un gran número de casos, en la utilización del concepto de *frame*, que constituye una estructura que contiene un conjunto de conceptos. Un concepto representa uno de los diferentes tipos de consultas que puede realizar el usuario. Dichos conceptos pueden tener asociados atributos, correspondientes a la información que debe aportar el usuario para completar o modificar la consulta

realizada al sistema. De este modo, tras cada intervención del usuario, el *NLU* recibe un texto, realiza un procesamiento que se centra en la obtención de los conceptos dentro del texto, y por último, genera un *frame* con la representación semántica, que resulta indispensable para el próximo módulo de procesamiento en el sistema de diálogo [11].

En este contexto se define el término acto de diálogo, como un tipo especial de concepto que asigna una clasificación al texto en cuanto a la actitud del hablante [12]. Los actos de diálogo varían entre diferentes sistemas de diálogo de acuerdo al dominio del conocimiento sobre el cual trabajan, pero existen algunos generales que son propios de cualquier diálogo, por ejemplo, aquellos que identifican peticiones, negaciones o afirmaciones. En su definición se trata de cubrir diferentes tipos de textos, de forma que sea posible dar una clasificación a todas las intervenciones de los usuarios. En cada momento se debe poder comprender si se está haciendo una petición, una enunciación, una negación, etc. Tomando como entrada el texto del ejemplo antes mencionado, la representación semántica que se extrae tiene la siguiente información:

**Texto:** ¿Por cuál alimento puedo sustituir el chocolate?

**Acto de diálogo:** SUSTITUCIÓN  
**Alimento:** chocolate

Asumiendo que “SUSTITUCIÓN” es uno de los actos de diálogo definidos en el sistema.

Generalmente, el *NLU* está compuesto por dos tareas principales para la comprensión del texto que recibe: determinar el acto de diálogo y obtener los conceptos y atributos contenidos. La primera de las tareas es tratada, frecuentemente, como un problema de clasificación. Por otra parte, la obtención de conceptos y atributos es modelada como el problema de obtener etiquetas semánticas para cada palabra del texto [11].

Cada una de estas tareas conlleva a la utilización de varias herramientas del Procesamiento del Lenguaje Natural (*PLN*)<sup>1</sup>, introduciendo grades desafíos al módulo *NLU* [14-15]. Dichos desafíos están relacionados con los problemas propios de esta disciplina (Ambigüedad, que puede ocurrir en todos los niveles análisis; Sinonimia; Resolución de Referencias; Implicación Textual; Errores tipográficos y ortográficos; entre otras), que dificultan el proceso de comprensión, principalmente, el proceso de obtención de conceptos.

### **Propuesta para la comprensión de textos en un sistema de diálogo en el dominio de control de nutrientes**

El componente *NLU* que se presenta en este trabajo ha sido diseñado para integrarse en un sistema de diálogo enmarcado en el dominio de conocimiento del Método ACN [16]. Este es un método para el control de nutrientes que ayuda a la pérdida de peso de manera rápida y segura. Dicho dominio es muy específico, y solo se tienen los conocimientos brindados por el especialista del método. Por tal razón, y al no contar con muestras etiquetadas de diálogos, en este trabajo hemos seleccionado el enfoque basado en conocimiento para el módulo de la comprensión de textos. De esta forma, es posible su diseño y desarrollo solo a partir de la definición de un conjunto de reglas lingüísticas y recursos semánticos. El recurso empleado en la propuesta es una ontología, elaborada para el dominio en cuestión.

El proceso en el componente, de forma general, comienza con la entrada del texto del usuario. Teniendo en cuenta que el texto puede estar formado por más de una oración, la estrategia planteada realiza como primer paso la segmentación por oraciones del texto brindado por el usuario. Cada segmento obtenido representa un texto que es analizado por los dos módulos principales que componen al *NLU*. El primer módulo es el encargado de realizar el preprocesamiento y el segundo el análisis. Como resultado se obtiene una lista de *frames* con los conceptos identificados (un *frame* para cada oración).

#### *Etapas de preprocesamiento*

La etapa de preprocesamiento de los textos consiste en preparar los textos para su análisis, facilitando su manejo en etapas posteriores. En la propuesta se han incluido el siguiente conjunto de tareas del *PLN*:

1. La evaluación de las palabras carentes de información semántica, cuyo significado no es relevante para la tarea (stopwords, en inglés), por ejemplo las preposiciones, artículos, etc.
2. La corrección ortográfica a cada palabra, que permite identificar y corregir automáticamente los errores [17].
3. La lematización, que asigna a cada palabra su raíz léxica. En este sentido, las múltiples derivaciones y/o inflexiones de una palabra son llevadas a una sola, el morfema original. Esto permite normalizar los términos pertenecientes a una misma familia y por tanto próximos en significado, reduciéndolos a una forma común [18].
4. El etiquetado gramatical (Part of Speech, en inglés), que asigna a cada una de las palabras del texto su categoría gramatical (sustantivos, verbos, adjetivos, adverbios, etc.). Este proceso muestra, hasta cierto punto, la estructura de un texto, brindando una gran cantidad de información sobre la palabra y sus vecinas [18].

#### *Etapas de análisis*

En la etapa de análisis se realizan un conjunto de importantes tareas para la obtención de los conceptos y la posterior generación del *frame*. Un concepto puede ser una palabra o una secuencia de palabras (frase) relevante, que se han dividido en tres tipos: actos de diálogos, entidades y acciones.

<sup>1</sup> El Procesamiento del Lenguaje Natural o *PLN* es un campo dentro de las Ciencias de Computación y la Lingüística, que tiene como objetivo fundamental dotar a los sistemas de cómputos de habilidades que le permitan la interacción hombre- máquina mediante el uso del lenguaje natural [13].

#### DetECCIÓN DE ACTOS DE DIÁLOGOS

El **acto de diálogo** constituye un tipo especial de concepto que se obtiene para facilitar la toma de decisiones en el manejo del diálogo. En el sistema propuesto se ha definido una jerarquía con un total de 25 actos de diálogos. Los más generales son *Enunciativo\_afirmativo*, *Enunciativo\_negativo* e *Interrogativo*, que permiten guiar al sistema en el tipo de respuesta que debe brindar. El resto de los actos de diálogo asignan una clasificación al texto de forma más detallada, brindando información con un mayor grado de granularidad. El conjunto de los interrogativos es el más extenso, debido a que se consideran grandes cantidades de las interacciones de los usuarios que demandan algún tipo de información. Estas especifican qué es lo requerido, definiéndose los siguientes actos de diálogo: *Definición*, *Caracterización*, *Procedimental*, *Causal*, *Confirmación*, *Si\_No*, *Argumentativo*, *Ubicación*, *Temporal*, *Cantidad*, *Especificación*, *Finalidad* y *Sustitución*. El último de los mencionados se ha definido por la alta probabilidad de ocurrencia de consultas al sistema de diálogo por usuarios para el reemplazo de alimentos sugeridos por el método bajo diferentes situaciones. En el caso de las afirmativas, están representadas por aquellas interacciones del usuario que enuncian una idea afirmativa, y que por tanto, no pueden constituir preguntas. En este grupo se han definido las siguientes: *Apertura*, *Cierre*, *Satisfacción*, *Exhortativo* y *Afirmación*, este último para las interacciones simples del tipo “sí”, “claro”, etc. Por otra parte, las negativas son aquellas interacciones del usuario que enuncian una idea negativa, transmitida mediante el uso de alguna partícula negativa. Las definidas son: *Insatisfacción*, *Dubitativa* y *Negación*, este último para las interacciones simples del tipo “no”, “nunca”, etc.

En el diseño se han definido un conjunto de reglas que identifican cada uno de los actos de diálogo en sus diferentes variantes. Dichas reglas se basan, en muchos casos, en el resultado de la lematización y el etiquetado gramatical; y en otros casos, solo en las palabras originales.

#### DETECCIÓN DE ENTIDADES

Una entidad constituye un tipo importante de concepto, que define el objeto por el cual el usuario realiza una interrogación o brinda alguna información. Incluso, una entidad puede definir a su vez, propiedades de otras entidades que también representan conceptos. Por ejemplo, en el texto “¿Puedo comer carne haciendo la dieta?”, se reconocen dos conceptos de tipo entidad: *carne* y *dieta*. Dentro de este tipo de concepto se han identificado tres tipos particulares: las entidades del dominio, los eventos y las entidades numéricas.

Un concepto de tipo **entidad del dominio** se corresponde con todo aquello cuya existencia es relevante para el sistema, no solo los elementos, sino también las propiedades y las relaciones. Para obtenerlas, en la propuesta se ha creado una ontología que contiene de forma organizada los aspectos de interés que el sistema conoce y está diseñado para tratar. En este recurso se encuentran las diferentes lexicalizaciones (palabras o frases) con las que un concepto puede aparecer, y asociado a ellas el nombre de dicho concepto y su jerarquía correspondiente. La jerarquía se ha definido según el orden intrínseco entre los conceptos. Por ejemplo, todos los elementos correspondientes a las fases del método ACN pertenecen a la clase *Método* y dentro de estos a su vez, hay otras clasificaciones como Fase. En el texto “¿Qué debo hacer en la fase inicial de la dieta?”, aparece la frase “fase inicial de la dieta”; con esta lexicalización, se identifica la jerarquía de conceptos [*Método*, *Fase*, *Fase\_reactivo*], de la que se obtienen los siguientes conceptos:

<b>Lexicalización:</b> <i>fase inicial de la dieta</i>	<b>Método:</b> <i>Fase_reactivo</i>
	<b>Fase:</b> <i>Fase_reactivo</i>

La misma jerarquía, y por tanto los mismos conceptos, se identifican para otras lexicalizaciones como por ejemplo: “la primera de las tres fases”. Este diseño de la ontología permite identificar un mismo concepto a partir de las diferentes lexicalizaciones con las que se conoce y con las que el usuario puede mencionarlo.

Otras jerarquías muy importantes se relacionan con los elementos correspondientes a alimentos. Este es un tipo de concepto muy mencionado al tratarse de un dominio de control de nutrientes. Dichos elementos pertenecen a la clase *Alimentos* y dentro de estos a su vez, hay otras clasificaciones como las Verduras y las Frutas. Por poner un ejemplo, la lexicalización “mango”; se corresponde a la jerarquía de conceptos [*Alimento*, *Fruta*, *mango*].

Los **eventos** son casos particulares de entidades del dominio que se corresponden con situaciones especiales a las que el sistema trata de forma diferenciada. Estas situaciones se corresponden con las frases en que el usuario plantea alguna incomodidad con el método ACN o algún malestar físico. Por ejemplo, en el texto “tengo náusea desde que comencé la dieta”, se identifica como entidades *dieta* y *tener\_náusea*, resultando esta última a su vez un evento; y la jerarquía correspondiente en la ontología es: [*Evento*, *Malestar\_físico*, *tener\_náusea*].

Las **entidades numéricas** son aquellas frases que identifican cantidades. Estas son expresiones que tienen un carácter general, y por tanto no dependen del dominio, pero que juegan un papel muy importante en la comprensión. Por ejemplo, en el texto “Tengo 24 años”, se obtiene el concepto de tipo entidad: *24 años*, que logra resumir convenientemente el contenido en la estructura que procesará el sistema.

En el diseño se han definido varios tipos de entidades numéricas tomando en consideración las clasificaciones fundamentales de las cantidades. Esto permite aumentar el nivel de comprensión de lo que el usuario dice, y por tanto, mejora el proceso de toma de decisiones del sistema. Los tipos de entidades numéricas son: *Masa*, *Longitud*, *Capacidad*, *Tiempo*, *Cantidad*. El último caso es asociado a las entidades detectadas como numéricas que no correspondan con las otras clasificaciones. Teniendo en cuenta esta clasificación, en el ejemplo antes analizado, el concepto obtenido se corresponde con el tipo *Masa*.

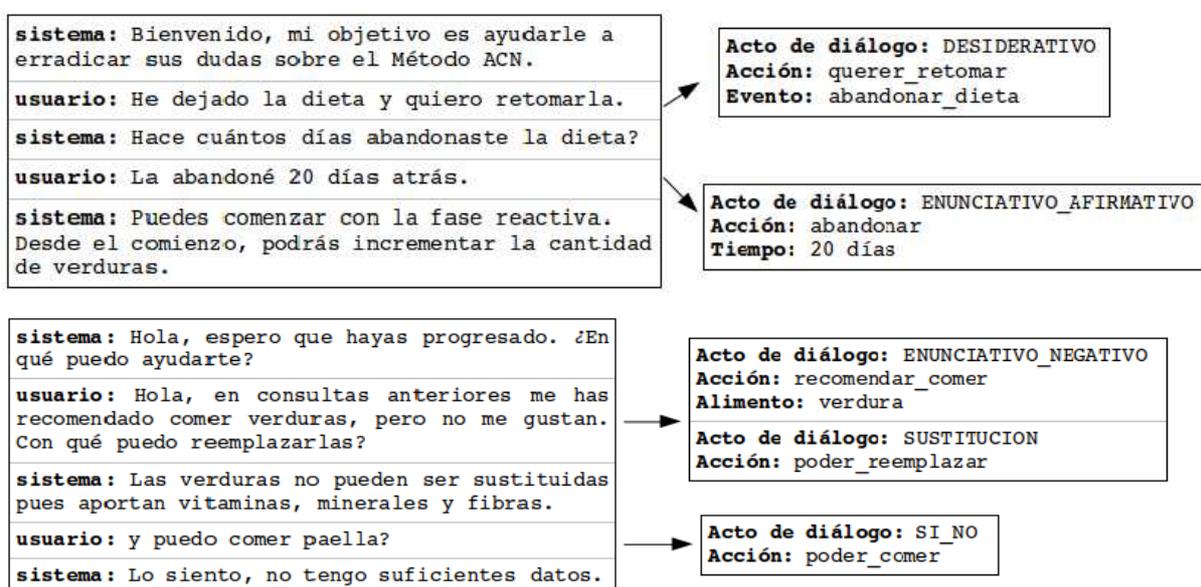
### Detección de la acción

Una información importante en muchos casos es la **acción**, correspondiente a las formas verbales, que aparecen en el texto. Este es un dato que ayuda a disminuir el grado de ambigüedad que puede ocurrir en el proceso de comprensión. Por ejemplo, los textos “¿puedo comer chocolate?” y “¿el chocolate es bueno?”, generan los mismos datos con los tipos de conceptos analizados hasta ahora. Ambos textos tienen el mismo acto de diálogo: *Si\_No*, y de ellos se obtiene el concepto de tipo entidad: chocolate. Sin embargo, los dos casos no corresponden exactamente a la misma interrogante, ya que puede que el chocolate en forma general sea bueno, y no lo sea para una persona en particular que no debería comerlo. Por tanto, estas preguntas requieren de tratamientos diferentes. Esta situación se maneja con la incorporación de otro tipo de concepto: la acción.

El proceso de obtención de la acción se basa en los resultados del etiquetador y un conjunto de reglas lingüísticas. Por una parte, el etiquetador identifica las palabras que constituyen verbos y por otra parte, las reglas determinan los diferentes tipos de construcciones verbales. En el ejemplo mencionado, la acción del primer texto sería *poder\_comer*, mientras que en el segundo sería *ser*, lo que ayuda a obtener una mejor comprensión de dichos textos.

### Resultados

En la evaluación del componente se analizó un conjunto de diálogos con el sistema, comprobando la comprensión de los textos de los usuarios a partir de la verificación de la coherencia de las respuestas del sistema. En cada caso se evaluó la extracción de los diferentes tipos de conceptos. A continuación se muestra una representación de este análisis.



En el primer ejemplo se generaron los *frames* esperados para cada texto del usuario. Estas estructuras contienen en efecto los datos necesarios para una correcta comprensión. En el caso del segundo fragmento de diálogo, los resultados no fueron los mismos. Para el texto de la primera intervención del usuario, formado por dos oraciones, se obtuvo dos *frames*, uno por cada oración. En la primera, la acción detectada no es la principal, pues la esencia de lo que el usuario dice es que no le gustan las verduras. Esto se debe a que se trata de una oración compuesta, y es un problema en la comprensión automática detectar su verbo principal. Sin embargo, este no es uno de los conceptos más importantes en la comprensión y el sistema logra dar una respuesta coherente. El mayor problema está dado en los elementos que no aparecen en la ontología construida. En el segundo texto, el usuario menciona el concepto paella, que no ha sido detectado y por ende, el sistema ha sido incapaz de brindar una respuesta. Por lo que se hace necesario para tratar esta problemática, la incorporación de una ontología complementaria con datos sobre un gran número de alimentos y sus relaciones.

### Conclusiones

En este artículo se realizó una descripción del componente para la comprensión de textos de usuarios en un sistema de diálogo sobre el dominio de control de nutrientes. El diseño se basa en el enfoque basado en conocimiento, que permitió el desarrollo del componente para un dominio específico, del cual solo se tenía un libro de referencia, y no se contaban con muestras de diálogos etiquetados. El módulo procesa el texto del usuario y genera un *frame*, estructura que contiene los conceptos que resumen apropiadamente la información, y que puede ser utilizado por el sistema para la toma de decisiones. Los conceptos se componen de diferentes tipos de elementos extraídos del texto: actos de diálogos, entidades y acciones. El proceso para su obtención se dividió en dos etapas: preprocesamiento, basado en tareas del *PLN*; y el análisis, basado en reglas y el uso de una ontología construida para el dominio. Como resultado, en la evaluación del componente se logró generar, para los textos empleados, las estructuras correspondientes con la información relevante de forma resumida y manejable por el sistema de diálogo.

## Referencias Bibliográficas

1. **Cheongjae Lee, Sangkeun Jung, Kyungduk Kim, Donghyeon Lee, and Gary Geunbae Lee.** "Recent approaches to dialog management for spoken dialog systems". *Journal of Computing Science and Engineering*, 4(1):1–22, 2010.
2. **Suket Arora, Kamaljeet Batra, and Sarabjit Singh.** "Dialogue system: A brief review". 2013.
3. **Alexander I Rudnicky, Eric H Thayer, Paul C Constantinides, Chris Tchou, R Shern, Kevin A Lenzo, Wei Xu, and Alice Oh.** "Creating natural dialogs in the carnegie mellon communicator system". In *Eurospeech*, 1999.
4. **Norman M Fraser.** "The sundial speech understanding and dialogue project: results and implications for translation". In *Aslib proceedings*, volume 46, pages 141–148. MCB UP Ltd, 1994.
5. **Victor Zue, Stephanie Seneff, James R Glass, Joseph Polifroni, Christine Pao, Timothy J Hazen, and Lee Hetherington.** "Jupiter: a telephone-based conversational interface for weather information". *Speech and Audio Processing*, IEEE Transactions on, 8(1):85–96, 2000.
6. **Joaquim Llisterri.** "Introducción a los sistemas de diálogo. Los sistemas de diálogo". Bellaterra-Soria: Universidad Autónoma de Barcelona, Publicaciones-Fundación Duques de Soria (Manuales de la Universidad Autónoma de Barcelona, Lingüística, 45), p. 11–21, 2006.
7. **David Griol.** "Desarrollo y evaluación de diferentes metodologías para la gestión automática del diálogo". PhD thesis, Departamento de Sistemas Informáticos y Computación. Universidad Politécnica de Valencia, 2007.
8. **Oliver Lemon and Olivier Pietquin.** "Data-Driven Methods for Adaptive Spoken Dialogue Systems: Computational Learning for Conversational Interfaces". Springer, November 2012. ISBN 978-1461448020.
9. **Eric Jackson, Douglas E Appelt, John Bear, Robert C Moore, and Ann Podlozny.** "A template matcher for robust *nl* interpretation". In *HLT*, volume 91, pages 190–194, 1991.
10. **Wayne Ward and Sunil Issar.** "Recent improvements in the cmu spoken language understanding system". In *Proceedings of the workshop on Human Language Technology*, pages 213–216. Association for Computational Linguistics, 1994.
11. **Ye-Yi Wang, Li Deng, and Alex Acero.** "Semantic Frame-Based Spoken Language Understanding". *Spoken Language Understanding: Systems for Extracting Semantic Information from Speech*, pages 41–91, 2011.
12. **Andreas Stolcke, Klaus Ries, Noah Coccaro, Elizabeth Shriberg, Rebecca Bates, Daniel Jurafsky, Paul Taylor, Rachel Martin, Carol Van Ess-Dykema, and Marie Meteer.** "Dialogue act modeling for automatic tagging and recognition of conversational speech". *Computational linguistics*, 26(3):339–373, 2000.
13. **López, Julio C. Torres, Christian Sánchez-Sánchez, and Esaú Villatoro-Tello.** "Laboratorio en línea para el procesamiento automático de documentos". Universidad Autónoma Metropolitana, Unidad Cuajimalpa, D.F., México.
14. **Pierre Lison.** "Structured Probabilistic Modelling for Dialogue Management". PhD thesis, University of Oslo, 2013.
15. **Bates, Madeleine.** "Models of natural language understanding". *Proceedings of the National Academy of Sciences* 92(22):9977-9982, 1995.
16. **Rodolfo Arruáran.** "El método ACN: Una manera saludable de adelgazar". 2014. ISBN 978-8494211683.
17. **Castro, Daniel Castro, C. Aurora Pons Porrata, and Ms C. Yunior Ramírez Cruz.** "Métodos para la corrección ortográfica automática del español". Diss. Master's thesis, Universidad de Oriente. Facultad de Matemática y Computación, Santiago de Cuba, 2012.
18. **Moya, Lisette García, C. Aurora Pons Porrata, and C. Leonel Ruiz Miyares.** "Un etiquetador morfológico para el español de Cuba". Diss. Master's thesis, Universidad de Oriente. Facultad de Matemática y Computación, Santiago de Cuba, 2008.