

# EL EXPERIMENTO MAGO DE OZ Y SUS APLICACIONES: UNA MIRADA RETROSPECTIVA

## WIZARD-OF-OZ EXPERIMENT AND ITS APPLICATIONS: AN OVERVIEW

CARLOS M. ZAPATA

*Grupo de Investigación en Ingeniería de Software. Facultad de Minas, Universidad Nacional de Colombia-Medellín.*

NICOLÁS CARMONA

*Grupo de Investigación en Ingeniería de Software. Facultad de Minas, Universidad Nacional de Colombia-Medellín.*

Recibido para revisar 26 de Enero de 2006, aceptado 26 de Septiembre de 2006, versión final 15 de Octubre de 2006

**RESUMEN:** En la Lingüística Computacional los corpus han sido uno de los recursos más utilizados para estudiar los diferentes fenómenos lingüísticos. La recopilación de corpus en los cuales se presenta interacción entre el humano y el computador presenta como problema fundamental los errores que se pueden inducir por parte del humano, cuyo lenguaje se simplifica en presencia del computador. Para realizar la recopilación de estos corpus se han empleado los Experimentos Mago de Oz (MDO). En este artículo se muestra una mirada retrospectiva del Experimento Mago de Oz, clasificando sus principales aplicaciones.

**PALABRAS CLAVE:** Corpus, Mago de Oz, Lingüística Computacional, Reconocimiento del habla.

**ABSTRACT:** In Computational Linguistics, corpora have been ones of the most useful resources for several linguistic phenomena studies. Corpus collection in Human-Computer Interaction has a main problem: humans simplify their language in presence of a computer and they can introduce errors in collection. For achieving corpus collection, Wizard-of-Oz Experiments have been used. In this paper, we show an overview of Wizard-of-Oz Experiment, and we classify its main applications.

**KEY WORDS:** Corpus, Wizard-of-Oz, Computational Linguistics, Speech recognition.

### 1. INTRODUCCIÓN

La forma en que se realiza la transmisión de los mensajes entre los seres humanos y su replicación a la relación hombre-máquina han sido del interés de la Lingüística Computacional desde sus inicios. Uno de los recursos utilizados por esta ciencia desde entonces es el corpus, una gran cantidad de datos recolectados a partir de diálogos, textos u otras fuentes, con el fin de realizar estudios de tipo lingüístico. Generalmente, se procura estudiar alguna característica que pueda ser posteriormente transferible a los sistemas que interactúan con los humanos. Entre los usos identificados de los corpus se encuentran: desambiguación de textos, extracción de términos, construcción de bases de conocimientos, evaluación de sistemas de procesamiento del lenguaje y anotación

automática a partir de anotación manual (McEnergy, 2003).

En el contexto de este artículo, uno de los usos más significativos de los corpus tiene que ver con su recolección para el diseño y evaluación de sistemas de diálogo. Los corpus que se emplean en estos sistemas suelen ser de dos tipos: persona-persona o persona-computador. En la recolección de la información para los corpus persona-computador se encuentran un obstáculo que es aparentemente contrario a la intuición, pues las personas que se someten a hablar o interactuar con un computador tienden a utilizar un lenguaje más sencillo que cuando hablan con otras personas, como si estuvieran comunicándose con un niño (Whittaker y Stenton, 1989). De esta forma, tales personas pueden introducir errores en la recolección del

corpus y hacerlo inútil para el objetivo que se busca lograr. Para superar esta limitación, desde inicios de la década de los ochenta se usa un tipo de experimento denominado “Mago de Oz”, que permite la recolección de corpus de diálogos persona-computador, simulando el computador con una persona oculta que realiza todas o algunas de las funciones que realizará el computador en el sistema de diálogo definitivo (Dahlbäck et al., 1993).

En este artículo se presenta una mirada retrospectiva del Experimento Mago de Oz y sus aplicaciones, enfatizando en los sistemas de diálogo, pero tomando en consideración otras aplicaciones de la técnica en diferentes entornos.

Este artículo está organizado de la siguiente manera: a continuación de esta introducción, en la Sección 2 se define la Lingüística de Corpus, como área en la cual se realizan los Experimentos Mago de Oz; el experimento Mago de Oz se describe en la Sección 3 y una mirada retrospectiva a sus principales aplicaciones se presenta en la sección 4. Finalmente, en las Secciones 5 y 6 se muestran las conclusiones y los trabajos futuros respectivamente.

## 2. LINGÜÍSTICA DE CORPUS

Un corpus es un cuerpo de evidencia lingüística compuesta típicamente por usos probados del lenguaje, tales como conversaciones diarias, publicidad radial, escritos publicados, etc. Los corpus adquieren su real utilidad cuando se encuentran en un formato legible por máquina, y en ese caso se convierten en colecciones bien organizadas de datos, recolectadas mediante un marco de ejemplos diseñado para permitir la exploración de cierta característica lingüística mediante los datos recogidos (McEnery, 2003). Tognini-Bonelli (2001) ejemplifica un acercamiento a la Lingüística de Corpus que toma evidencia de este recurso como la fuerza primaria detrás de las descripciones lingüísticas.

Meyer (2002) ofrece una guía teórica combinada con un manejo práctico para el diseño,

construcción, anotación y explotación de un corpus en inglés. Llisterri et al. (2003) y Llisterri (2003) presentan un resumen interesante de las Tecnologías del habla, con un apartado especial para el manejo de corpus, incluyendo la técnica Mago de Oz, la cual se describe en la sección siguiente.

En Aijmer y Altenberg (2004) se presentan algunas de las tendencias actuales en adquisición, anotación e interpretación de corpus. Otros ejemplos de nuevas tendencias en el uso de corpus son las siguientes:

- Jönsson y Dahlbäck (2000) plantean un método para la reescritura (destilación) de diálogos naturales, con el fin de elicitar el diálogo que podría ocurrir si uno de los participantes del diálogo es un computador.
- Segarra et al. (2002) presentan un enfoque para el desarrollo de sistemas de comprensión del lenguaje, el cual se validó en cuanto a precisión utilizando un corpus de diálogos telefónicos de itinerarios de trenes en España.
- Freddi (2005) propone el uso de corpus de los capítulos introductorios de 10 libros de texto en lingüística para estudiar la forma de argumentación de los autores y para discutir sus peculiaridades estilísticas.
- Wang (2005) realiza un análisis del fenómeno de reduplicación y repetición de palabras en un discurso, empleando un corpus construido específicamente para ese fin.
- Hyland y Tse (2005) utilizan un corpus académico para explorar los usos de la palabra inglesa “That” en escritos académicos.
- Meurers (2005) utiliza un corpus del lenguaje alemán para obtener ejemplos significativos de ese lenguaje para propósitos de lingüística teórica.

En la siguiente sección se presenta el experimento Mago de Oz, con sus principales características y las fases que conforman una ejecución típica de este experimento.

### 3. EL EXPERIMENTO MAGO DE OZ

Ogden y Bernick (1996) describen ampliamente el uso de Interfaces de Lenguaje Natural en la Interacción humano-computador. En esta descripción se incluyen aspectos de evaluación y diseño de este tipo de interfaces, para las cuales una de las técnicas principales de diseño está constituida por los denominados “Experimentos Mago de Oz” (en adelante MDO).

Estos experimentos se realizan con la participación de un humano que se conoce como “Mago”o “Cómplice”, quien simula, sin conocimiento del participante en el experimento, el papel que desempeña el computador durante una interacción humano-computador (Fraser y Gilbert, 1991).

La primera referencia al nombre de este tipo de experimentos se remonta paradójicamente a una época en la cual no existían computadores, con la novela de Baum (1900) que lleva el mismo nombre de esta técnica y en la cual un Mago Gigante gobierna los destinos en la mágica ciudad de Oz, pero se revela posteriormente que es un hombre protegiendo su identidad a través de una máquina. Se presume que quien bautizó esta técnica fue Kelley (1983), por una serie de experimentos en los cuales literalmente se escondía de la vista de los participantes a través de una cortina. Esta técnica también se suele llamar PNAMBIC (Pay No Attention to the Man Behind the Curtain – no preste atención al hombre tras la cortina), que se suele atribuir a J. Berstein (Newell, 1987). Respecto del nombre “Mago de Oz” Kelley (2006), en su página Web, entrega una alusión al término “OZ” como un acrónimo de “Offline Zero”, para expresar el hecho de que el mago interpreta las solicitudes del usuario en un sistema en tiempo real.

La primera noción a una experimento de este tipo se remonta a Gould et al. (1983), aunque no tenía ese nombre aún. En su artículo se trataba de simular una máquina de escribir que respondiera a la voz humana, mostrando automáticamente los textos que fuera reconociendo; como en esa época la tecnología de reconocimiento del habla no estaba suficientemente avanzada, se recurrió a

simulaciones en las cuales los usuarios miraban a una pantalla y dictaban a través de un micrófono, el cual estaba conectado a un digitador humano en otra alcoba y a través de un teclado escribía en la pantalla del usuario.

Para poder emplear esta técnica se deben cumplir tres precondiciones (Fraser y Gilbert, 1991):

- Es posible simular el sistema futuro, dadas las limitaciones humanas del Mago, tales como la velocidad de respuesta ante la solicitud de un usuario y la posibilidad de que el Mago no sea capaz de procesar una solicitud.
- Es posible especificar el funcionamiento futuro del sistema.
- Es posible que la simulación sea convincente.

En Dahlbäck et al. (1993) se encuentra una descripción detallada de la técnica, al igual que en Fraser y Gilbert (1991). Estos últimos autores reportan las tres fases que se suelen tener en la metodología, en las cuales coinciden con Gibbon et al. (1997):

- Fase Pre-experimental: en la cual se definen las estructuras y restricciones del diálogo. En esta fase se analiza el dominio de la aplicación, los escenarios posibles (rol del sujeto y antecedentes), la locación y el software y hardware requerido para la realización del experimento.
- Fase Experimental Primaria: en la cual se realiza una primera evaluación del sistema. En esta fase se aplican algunas restricciones al Mago o al sujeto y se debe recopilar información básica sobre el vocabulario inicial, la gramática y el modelo de diálogo.
- Fase experimental secundaria o subsecuente: En la cual se realizan recolecciones de corpus y se realizan mejoras al sistema. Puede no ser solo una fase sino un conjunto de iteraciones en las que paulatinamente se vaya sustituyendo el papel del mago por funcionalidades de la aplicación definitiva; en este caso se suele hablar de un Mago “biónico” en el sentido de que algunas de sus funciones las desempeña el sistema definitivo. En cada iteración se añaden

restricciones identificadas a partir de la fase anterior.

En Meyer (2002) se establecen algunos usos identificados de la técnica MDO en relación con la elaboración de prototipos de sistemas:

- Probar la eficacia del sistema: Se pueden modificar las estrategias de diálogo de acuerdo con los resultados sucesivos que vaya arrojando el experimento.
- Analizar el comportamiento de las personas en la interacción con una máquina, con el fin de extraer la información lingüística que necesitará el sistema real cuando deba interpretar los mensajes del usuario. Se deberían comparar los corpus persona-persona que se tengan disponibles con los corpus persona-sistema informático que se recolecten con el experimento; para ello, se toman en consideración: la duración de la interacción, los tiempos de espera, el número de turnos utilizados desde que el usuario solicita la información hasta que se le facilita y el número de llamadas que se desvían porque el Mago no es capaz de contestarlas. Estos elementos suministran parámetros de diseño del sistema definitivo.
- Entrenar el módulo de reconocimiento, ya que en la mayoría de las aplicaciones se requiere el reconocimiento del habla espontánea.

Es de aclarar que en desarrollo del experimento se pueden usar modalidades de lenguaje natural (hablado o escrito) o usar otras modalidades, como por ejemplo manipulación de símbolos o tonos de toque, como se realiza por ejemplo en experimentos para máquinas contestadoras de teléfono, como se muestra en la siguiente sección (Fraser y Gilbert, 1991). Además, en dicha Sección se realiza un listado de las aplicaciones identificadas para el experimento MDO, constituyendo una visión retrospectiva del mismo.

#### **4. UNA MIRADA RETROSPECTIVA A LAS PRINCIPALES APLICACIONES DEL EXPERIMENTO MAGO DE OZ**

Entre los usos identificados para el Experimento MDO, uno de los principales es el diseño, construcción y evaluación de Sistemas de Diálogo. Otras aplicaciones que se incluyen en esta sección se relacionan con la recolección de corpus para propósitos generales y los entornos de desarrollo para la realización de experimentos MDO. Se encontraron pocas aplicaciones para apoyar el ciclo de vida del desarrollo de software, un área en la cual tendría mucho potencial este tipo de experimentos, en especial para la recolección de requisitos para el desarrollo de sistemas. Al final de la sección se muestran otro tipo de aplicaciones que por su especificidad no se pueden clasificar en ninguna de las categorías anteriores.

##### **4.1 Sistemas de Diálogo:**

Una de las principales aplicaciones del experimento MDO se relaciona con los sistemas de diálogo y especialmente en el tema de errores en la comunicación. A este respecto, algunos de los trabajos identificados son:

- Mayfield y Burger (1999) emplean dos técnicas (una de las cuales es MDO) para recolectar una base de datos de hablantes no nativos del inglés, con el propósito de analizar los errores más comunes en que incurrían.
- En Batliner et al. (2003) se trata la identificación de problemas en la comunicación de los usuarios de un sistema de call center y los turnos para pasar la atención del usuario a un operador humano; aquí se emplea la técnica MDO para simular el sistema y obtener los indicios de problemas en la comunicación.

Sin embargo, con el uso de esta técnica se pueden presentar problemas puesto que, cuando el usuario se percata de que está hablando con un computador inhibe el manejo involuntario de ciertos errores en la comunicación, restándole utilidad al manejo de la técnica (Skantze, 2005). En Karsently y Botherel (2004) se recurre a la denominada estrategia de “transparencia” (hacer visible el sistema de acuerdo con las necesidades del usuario) para ayudar a los usuarios a responder adecuadamente a los errores del sistema; en este caso se utilizó la técnica MDO para evaluar el sistema TRAVELS, con el cual se pretendía obtener itinerarios aéreos y férreos por vía telefónica.

Otro de los aspectos de los Sistemas de Diálogo en que es útil la información arrojada por los experimentos MDO es el habla espontánea. Al respecto, en Zhang et al. (2005b) se propone un método para determinar la prioridad pragmática y semántica del habla espontánea en inglés, utilizando los conceptos de contraste y enfoque; en este caso, el experimento MDO se empleó para recolectar un escenario de diálogo de enseñanza que posteriormente permite validar los cálculos del contraste. En esta misma línea, pero en idioma español, Bonafonte et al. (2000) reportan el desarrollo de un sistema de diálogo oral que utiliza tanto corpus de diálogos persona-persona como corpus recopilados con MDO para analizar los fenómenos de habla espontánea.

Un tercer aspecto de utilidad de los experimentos MDO en relación con los Sistemas de Diálogo es la evaluación de dichos sistemas. A este respecto, en Sparck y Galliers (1995) se provee una perspectiva histórica de los esfuerzos realizados para proveer evaluaciones de los sistemas de Procesamiento del Lenguaje Natural, describiendo las deficiencias de esos esfuerzos y proponiendo un direccionamiento de las mismas; uno de los esfuerzos listados es el experimento MDO. Un estudio similar se puede consultar en Whitakker y Stenton (1989). Paek (2001) emplea un protocolo para la conducción de un experimento MDO y un conjunto básico de estadísticas descriptivas de los datos recolectados en el experimento para realizar la evaluación de dichos sistemas.

Un último aspecto de los Sistemas de diálogo tiene que ver con los Modelos de diálogo. En este tema, Dybkjaer et al. (1993) se realiza un estudio basado en MDO de los Modelos de diálogo en lenguaje oral, haciendo especial énfasis en la negociación que se produce entre la naturalidad del sistema y las restricciones tecnológicas impuestas por el reconocedor de voz.

#### 4.2 Recolección de Corpus con fines generales.

En cuanto a recolección de corpus con fines generales, algunos de los trabajos que se pueden listar son:

- En Moore y Morris (1992) se emplea la técnica MDO para la recolección de un corpus que denominan “genuino”, porque no se refiere a un escenario previamente preparado; el corpus se refiere a un servicio de demanda telefónica en lenguaje hablado. Otra aplicación telefónica fue realizada por Gamm et al. (1997) para el proceso completo de desarrollo de un sistema de habla que controla una máquina contestadora; en este caso se usó, durante el análisis, un prototipo MDO, en el cual un mago simplemente escuchaba las órdenes del usuario y presionaba los botones correspondientes.
- Hindle y Rooth (1993) utilizan un corpus recolectado mediante un experimento MDO sobre información de viajes para resolver la ambigüedad de posición de modificadores en frases preposicionales.
- Jokinen (1996) emplea un corpus recolectado con la técnica MDO para encontrar información en compañías de alquiler de autos y restaurantes en un área particular, para presentar lo que denomina el “Modelo de diálogo constructivo” como un nuevo enfoque para la formulación de metas en Sistemas de Diálogo inteligentes.
- Villaseñor et al. (2002) recopilan, mediante experimentos MDO, una serie de corpus para el diseño de cocinas, denominados DIME (Diálogos Inteligentes Multimodales en Español) y los utilizan para compararlos con otros corpus recolectados en la Web para la construcción de modelos del lenguaje.

- Steininger et al. (2002) emplean la técnica MDO para recopilar corpus de datos multimodales. Este proyecto, denominado SmartKom, busca el desarrollo de una interfaz inteligente computador-usuario que permita una comunicación casi natural con una máquina adaptativa y autoexplicatoria. El sistema acepta lenguaje natural y gestos.
- Zhang et al. (2005a) usan un corpus recolectado con experimentos MDO en los cuales el tutor (un mago humano) intentaba ayudar niños a aprender física y matemáticas con fichas ensamblables. Con este corpus se intenta estudiar un modelo híbrido que permite la comprensión del habla y que combina connotación y significado literal de las palabras. En Benz Müller et al. (2003) se trabaja utiliza la misma técnica para la recolección de datos empíricos, escritos en diálogos tutoriales de matemáticas en alemán y se presenta adicionalmente un enfoque metodológico que permite optimizar las ganancias de los estudios empíricos empleando la técnica MDO.

#### 4.3 Entornos para el desarrollo de experimentos MDO

En este tipo de aplicaciones se busca desarrollar un sistema genérico que posibilite la realización de diferentes experimentos MDO de manera dinámica y no enfocarse en el diseño de un único experimento. A este respecto, se encontraron los siguientes trabajos en la literatura:

- En Dahlbäck et al. (1993) se describe ARNE, un entorno de simulación para experimentos MDO que es adaptable para cada nueva aplicación y que ha permitido la recolección de varios corpus usados en el diseño de sistemas de diálogos.
- Munteanu y Boldea (2000) presentan MDWOZ, un entorno de desarrollo de sistemas de diálogo hablado basados en la técnica MDO, que facilita la colección de datos y la interacción para la construcción de modelos. Por su manejo modular, se pueden reutilizar partes del mismo en el sistema de diálogo final.
- En Fiedler y Gabsdil (2005) se presenta DiaWoZ, un sistema que soporta el diseño y

ejecución de experimentos MDO para recolectar datos de diálogos tutoriales. En este sistema, la arquitectura es altamente modular y permite el refinamiento progresivo de los experimentos; esta característica lo puede hacer apropiado para examinar diálogos de tipo general.

#### 4.4 Aplicaciones orientadas al Ciclo de Vida del Software:

En este aspecto, se detectaron pocos trabajos en la literatura, pero puede ser un área muy promisoría para los experimentos MDO, especialmente en aspectos tales como la recolección de requisitos de los sistemas de información. Cabe aclarar que la funcionalidad de evaluación de sistemas de diálogo descrita en la sección 4.1 se podría extender a los sistemas de información.

Salber y Coutaz (1993) y Coutaz et al. (1994) describen una plataforma MDO genérica denominada NEIMO, que se emplea para la evaluación de la usabilidad de una pieza de software. Esta plataforma digitaliza y guarda datos comportamentales que varían desde acciones internas del software hasta comandos ejecutados por un usuario empleando una interfaz multimodal.

En Sommerville y Sawyer (1997) se discute una forma de elaboración de prototipos para cualquier tipo de sistema, empleando para ello la técnica MDO. En este caso sólo se desarrolla la interfaz de usuario, con la cual interactúan los usuarios de prueba; las peticiones correspondientes a esa interfaz se entregan a una persona que las interpreta y genera una respuesta adecuada. En esta línea de trabajo, Mørch et al. (2005) usan la técnica MDO para simular con usuarios humanos el comportamiento de agentes de software; este estudio lo realizan para elaborar prototipos de los agentes de software.

Fraikin y Leonhardt (2002) proponen un proceso de retroalimentación con el usuario, que puede ser considerado como una ampliación de la técnica MDO para el desarrollo de prototipos de piezas de software partiendo de los casos de uso. Para realizar la retroalimentación, se emplea una

herramienta de captura y repetición que posibilita que el prototipo desarrollado ejecute interacciones que cada vez más se aproximen al sistema definitivo.

#### 4.5 Otras aplicaciones:

Otras aplicaciones del experimento MDO con propósitos diferentes a los analizados en las secciones previas son:

- Petrelli et al. (1997) emplean una simulación de un sistema multimodal con la técnica MDO para determinar qué tan importante es la experiencia del usuario en este tipo de sistemas.
- Manstetten et al. (2001) describen un simulador de conducción que emplea parcialmente la técnica MDO para controlar un diálogo ficticio entre un conductor y un computador en su vehículo; el objetivo es determinar los efectos de las distracciones del conductor en su modo de conducción. Un sistema similar es presentado por Cheng et al. (2004) para recolectar diálogos humano-computador en los cuales el humano no tiene como prioridad el diálogo por estar realizando una actividad simultánea (por ejemplo conducir un vehículo).
- Mäkelä et al. (2001) emplean un experimento MDO para realizar la interpretación de los mensajes en lenguaje natural que caracterizan los diálogos durante el diseño de un sistema de computación denominado "Doorman", que se emplea para la localización de personas.
- Whittaker et al. (2002) describen una herramienta basada en la técnica MDO que aplica modelos de usuario basados en la teoría de Decisión Multiatributo al problema de la generación de respuestas de un sistema de diálogo hablado. El sistema se aplicó a un conjunto de restaurantes en la ciudad de Nueva York.
- Paiva et al. (2003) describen SenToy, una muñeca con sensores que capturan las emociones de su usuario en forma de juego. Se realizó para ello un experimento MDO con diferentes tipos de muñecos controlados a control remoto por un mago; así, se pudieron identificar las acciones más

comunes asociadas con ciertas emociones y emplearlas en el diseño de la muñeca.

- Höysniemi y Read (2005) discuten los aspectos éticos y organizacionales de los experimentos Mago de Oz que involucran la participación de niños.
- En De Ruyter et al. (2005) se emplea el experimento MDO para tratar de comprender el comportamiento humano en relación con la tecnología en el hogar, con el fin de incorporar inteligencia social en este tipo de aplicaciones; esto puede contribuir al mejoramiento de la aceptación de la tecnología en el hogar y el inicio de comportamientos sociales del usuario en relación con los sistemas de diálogo en el hogar.

#### 5. CONCLUSIONES

Como principales conclusiones del experimento MDO, a partir de la mirada retrospectiva que se ha realizado en este artículo, se pueden mencionar las siguientes:

- La Lingüística Computacional ha empleado los corpus como uno de los recursos para estudiar ciertos fenómenos. Para la recolección de corpus hombre-máquina una de las técnicas que se ha utilizado en la literatura es Mago de Oz.
- En esta técnica un operador humano se suele esconder tras una interfaz de un sistema informático para simular su comportamiento, sin que el usuario humano se entere de ello.
- Como técnica se ha utilizado especialmente en el diseño, desarrollo y evaluación de sistemas de diálogo, la recolección de corpus con propósitos generales, los entornos para el desarrollo mismo de experimentos MDO, el apoyo al ciclo de vida del Software y otras aplicaciones más específicas que varían desde simuladores de conducción hasta diseño de juguetes.

#### 6. TRABAJOS FUTUROS

Dentro de las áreas examinadas para la aplicación de los experimentos MDO, algunas de las oportunidades identificadas de trabajos futuros se centraron en el Ciclo de Vida del

Software y la Ingeniería de requisitos, y son las siguientes:

- En lo que respecta al Ciclo de vida del Software, conjuntamente con la definición de entornos para la realización de experimentos Mago de Oz, se puede explorar la posibilidad de generar un sistema que emplee esta técnica para la recolección de requisitos de cualquier tipo de sistema informático. En este caso el Mago sería sustituido por un analista humano que se oculte a través de una interfaz, para lograr que los interesados en una pieza de software expresen sus requisitos de manera más concreta a través de la interacción con este sistema.
- Otra oportunidad de trabajo en el área de requisitos sería el análisis de los corpus existentes persona-persona para el levantamiento de requisitos de diferentes sistemas, con el propósito de identificar las características que permitirían diseñar el sistema de levantamiento de requisitos planteado.
- Continuando con el área de Ingeniería de requisitos, y tomando en consideración que mucha de la información que se recolecta en esta área se encuentra en forma de modelos, sería conveniente la realización de algún tipo de sistema que recopilara corpus de diagramas en lugar de diálogos y sobre el cual se pudieran realizar análisis y consultas.

## AGRADECIMIENTOS

Este artículo se realizó en el marco de los siguientes proyectos de Investigación: “Construcción Automática de Esquemas Conceptuales a partir de Lenguaje Natural”, financiado por la DIME y “Definición de un Esquema Preconceptual para la Obtención Automática de Esquemas Conceptuales de UML”, financiado por DINAIN y administrado por la DIME.

## REFERENCIAS

[1]. Aijmer, K. y Altenberg, B. (Eds). *Advances in Corpus Linguistics*. Rodopi, Amsterdam, 2004.

[2]. Batliner, A., Fischer, K., Huber, R., Spliker, J. Y Nöth, E. *How to Find Trouble in Communication*. En: *Speech Communication*, No. 40, 2003. pp. 117-143.

[3]. Baum, F. *The Wizard of Oz*. Collins, London, 1900.

[4]. Benzmüller, C., Fiedler, A., Gabsdil, M., Horacek, H., Kruijff-Korbayova, I., Pinkal, M., Siekmann, J., Tsovaltzi, D., Vo, B. Q. y Wolska, M. *A Wizard of Oz Experiment for Tutorial Dialogues in Mathematics*. En: *Proceedings of AI in Education (AIED 2003) Workshop on Advanced Technologies for Mathematics Education*, Sydney, 2003.

[5]. Bonafonte, A., Aibar, P., Castell, N. Lleida, E., Mariño, J.B., Sanchis, E. y Torres, I. *Desarrollo de un sistema de diálogo oral en dominios restringidos*. En: *Memorias de las Primeras Jornadas en Tecnología del Habla*, Sevilla, 2000.

[6]. Cheng, H., Bratt, H., Mishra, R., Shriberg, E., Upson, S., Chen, J., Weng, F., Peters, S., Cavedon, L. y Niekrasz, J. *A Wizard of Oz framework for collecting spoken human-computer dialogs*. En: *Proceedings of INTERSPEECH - 8th International Conference on Spoken Language Processing*, Jeju Island, 2004.

[7]. Coutaz, J., Salber, D., Balbo, S. y Jambon, F. *Supporting Usability Evaluation through Software Engineering Tools*. En: *Proceedings of the Workshop on Models for Developing High Impact Formative Evaluation Methods at the ACM Conference on Human Factors in Computing Systems CHI'94*, Boston, 1994.

[8]. Dahlbäck, N., Jönsson, A. y Ahrenberg, L. *Wizard of Oz Studies – Why and How*. En: Gray, W. D., Hefley, W. E. y Murray, D. Editores, *Proceedings of the 1993 International Workshop on Intelligent User Interfaces (IUI 1993)*, Orlando, 1993. pp. 193-200.

[9]. De Ruyter, B., Saini, P., Markopoulos, P. y Van Bremen, A. *Assessing the effects of building social intelligence in a robotic interface*

- for the home*. En: *Interacting with computers*, No. 17, 2005. pp. 522-541.
- [10]. Dybkjaer, H., Bernsen, N. y Dybkjaer, L. *Wizard-of-oz and the trade-off between naturalness and recogniser constraints*. En: *Proceedings of EUROSPEECH'93, Third European Conference on Speech Communication and Technology*, Berlin, 1993. pp. 947-950
- [11]. Fiedler, A. y Gabsdil, M. *Supporting Progressive Refinement of Wizard-of-Oz Experiments*. En: Carolyn Penstein Rose and Vincent Aleven (eds.), *Proceedings of the ITS 2002 - Workshop on Empirical Methods for Tutorial Dialogue*, San Sebastian, 2002. pp. 62-69.
- [12]. Fraikin, F. y Leonhardt, T. *From Requirements to Analysis with Capture and Replay Tools*. PI-R 1/02, Software Engineering Group, Department of Computer Science, Darmstadt University of Technology, 2002. Disponible en: <http://www.pi.informatik.tu-darmstadt.de/publikationen/technische%20Berichte/2002/pi2002-1.pdf> [Citado 18 de Enero de 2006]
- [13]. Fraser, N. y Gilbert, G. *Simulating Speech Systems*. En: *Computer Speech and Language*, No. 5, 1991. pp. 81-99.
- [14]. Freddi, M. *Arguing Linguistics: corpus investigation of one functional variety of academic discourse*. En: *Journal of English for Academic Purposes*, No. 4, 2005. pp. 5-26.
- [15]. Gamm, S., Haeb-Umbach, R. y Langmann, D. *The development of a command-based speech interface for a telephone answering machine*. En: *Speech Communications*, No. 23, 1997. pp. 161-171.
- [16]. Gibbon, D., Moore, R. y Winski, R. *Handbook of Standards and Resources for Spoken Language Systems*. Mouton de Gruyter, 1997.
- [17]. Gould, J. D., Conti, J. y Hovanyecz, T. *Composing letters with a simulated listening typewriter*. En: *Communications of the ACM*, Volume 26, Issue 4, 1983. pp. 295 - 308
- [18]. Hindle, D. y Rooth, M. *Structural Ambiguity and Lexical Relations*. En: *Computational Linguistics* Vol. 19, No. 1, 1993. pp. 103-120.
- [19]. Höysniemi, J. y Read, J.C. *Wizard of Oz Studies with Children*. En: *Proceedings of Interact 2005 Workshop on Child Computer Interaction: Methodological Research*, Rome 2005.
- [20]. Hyland, K. y Tse, P. *Hooking the reader: a corpus study of evaluative That in abstracts*. En: *English For Specific Purposes*, No. 24, 2005. pp. 123-139.
- [21]. Jokinen, K. *Goal Formulation based on Communicative Principles*. En: *Proceedings of the 16<sup>th</sup> Conference of Computational Linguistics*, Vol. 2, Copenhagen, 1996. pp. 598-603.
- [22]. Jönsson, A. y Dahlbäck, N. *Distilling Dialogues – a method using natural dialogue corpora for dialogue systems development*. En: *Proceedings of the 6<sup>th</sup> Applied Natural Language Conference*, Seattle, 2000.
- [23]. Karsently, L. y Botherel, V. *Transparency strategies to help users handle system errors*. En: *Speech Communication*, No. 45, 2005. pp. 305-324.
- [24]. Kelley, J. F. *Natural Language and Computers: Six empirical steps for Writing an Easy-to-Use Computer Application*. Tesis Doctoral, The Johns Hopkins University, 1983.
- [25]. Kelley, J. F. *Where did the usability term Wizard of Oz come from?*. Disponible en: <http://www.musicman.net/oz.html> [Citado 18 de Enero de 2006].
- [26]. Llisterri, J., Carbó, C., Machuca, M. J., De la Mota, C., Riera, M. y Ríos, A. *El papel de la fonética en el desarrollo de las tecnologías del habla*. En: *Memorias de las VII Jornadas de Lingüística*. Cádiz, Servicio de Publicaciones de la Universidad de Cádiz, 2003.

- [27]. Llisterri, J. *Las tecnologías del habla: Entre la Ingeniería y la Lingüística*. En: Actas del Congreso Internacional “La Ciencia ante el Público. Cultura humanística y desarrollo científico y tecnológico”, Universidad de Salamanca, Instituto Universitario de Estudios de la Ciencia y Tecnología, 2003. pp. 44-67.
- [28]. Mäkelä, K., Salonen, E., Turunen, M., Hakulinen, J. y Raisamo, R. *Conducting a Wizard of Oz Experiment on a Ubiquitous Computing System Doorman*. En: Proceedings of the International Workshop on Information Presentation and Natural Multimodal Dialogue, Verona, 2001. pp. 115-119.
- [29]. Manstetten, D., Krautter, W., Grothkopp, B., Steffens, F. y Geutner, P. *Using A Driving Simulator To Perform A Wizard-of-Oz Experiment On Speech-Controlled Driver Information Systems*. En: Proceedings of the Human Centered Transportation Simulation Conference (HCTSC 2001), Iowa City, USA, 2001.
- [30]. Mayfield, L. y Burger, S. *Eliciting natural speech from non-native users: collecting speech data for LVCSR*. En: Proceedings of the ACL-IALL Joint Workshop on Computer Mediated Language Assessment and Evaluation in NLP, Maryland, 1999.
- [31]. McEnery, T. *Corpus Linguistics*. En: Mitkov, R. Editor, *The Oxford Handbook of Computational Linguistics*, Oxford University Press, 2003. pp. 448-463.
- [32]. Meurers, W. D. *On the use of electronic corpora for theoretical linguistic case studies from the syntax of German*. En: *Lingua*, No. 115, 2005. pp. 1619-1639.
- [33]. Meyer, Ch. *English Corpus Linguistics*. Cambridge University Press, Cambridge, 2002.
- [34]. Moore, R. y Morris, A. *Experiences collecting genuine spoken enquiries using WoZ Techniques*. En: *Speech and Natural Language – Proceedings of a Workshop held at Harriman*, New York, 1992. pp. 61-63.
- [35]. Mørch, A., Dolonen, J. y Nævdal, J. *An Evolutionary Approach to Prototyping Pedagogical Agents: from simulation to integrated systems*. En: *Journal of Network and Computer Applications*, No. 29, 2005. pp. 177-199.
- [36]. Munteanu, C. y Boldea, M. *MDWOZ: A Wizard of Oz Environment for Dialog Systems Development*. En: Proceedings of the 2<sup>nd</sup> International Conference on Language Resources and Evaluation – LREC, Athens, 2000.
- [37]. Newell, A. F. *Speech simulation studies – performance and dialogue specification*. En: Unicorn Seminar, London, 1987.
- [38]. Ogden, W. y Bernick, Ph. *Using Natural Language Interfaces*. En: M. Helander (ed.), *Handbook of Human-Computer Interaction*, Elsevier Science Publishers B. V., North Holland, 1996.
- [39]. Paiva, A., Costa, M., Chaves, R., Piedade, M., Mourão, D., Sobral, D., Höök, K., Andersson, G. y Bullock, A. *Sentoy: an affective sympathetic interface*. En: *International Journal of Human-Computer Studies*, No. 59, 2003. pp. 227-235.
- [40]. Paek, T. *Empirical methods for evaluating dialog systems*. En: Proceedings of the ACL 2001 Workshop on Evaluation Methodologies for Language and Dialogue Systems, 2001. pp. 3-10
- [41]. Petrelli, D., De Angeli, A., Gerbino, W., Cassano, G. *Referring in Multimodal Systems: the importance of user expertise and system features*. En: Proceedings of the Workshop on Referring Phenomena in a Multimedia Context and their Computational Treatment, ACL-EACL, 1997. pp. 14-19.
- [42]. Salber, D. y Coutaz, J. *Applying the Wizard of Oz technique to the Study of Multimodal Systems*. En: Len J. Bass, Juri Gornostaev and Claus Unger (eds), *Human-Computer Interaction Selected Papers*, Berlin, Springer-Verlag, Lecture Notes in Computer Science, No. 753, 1993. pp. 219-230.

- [43]. Segarra, E., Sanchis, E., Galiano, I., García, F., Hurtado, L. *Extracting semantic information through automatic learning*. En: International Journal of Pattern Recognition and Artificial Intelligence, No. 16, Vol. 3, 2002. pp. 301–307
- [44]. Skantze, G. *Exploring human error recovery strategies: Implications for spoken dialogue systems*. En: Speech Communication, No. 45, 2005. pp. 325-341.
- [45]. Sommerville, I. y Sawyer, P. *Requirements Engineering: A Good Practice Guide*. John Wiley and Sons, Chicester, 1997.
- [46]. Sparck, K. y Galliers, J. *Evaluating Natural Language Processing Systems: an Analysis and Review*. En: Goos, G., Harmanis, J. y Van Leeuwen, J. (Eds), Springer Verlag, Lecture Notes in Computer Science, 1083, Berlin, 1995.
- [47]. Steininger, S., Rabold, S., Dioubina, O. y Schiel, F. *Development of the User-State Conventions for the Multimodal Corpus in SmartKom*. En: Proceedings of the 3<sup>rd</sup> International conference on Language Resources and Evaluation, Workshop On Multimodal Resources and Multimodal Systems Evaluation, Las Palmas, 2002.
- [48]. Tognini-Bonelli, E. *Corpus-Linguistics at work*. John Benjamins Publishing Company, Amsterdam y Philadelphia, 2001.
- [49]. Villaseñor, L., Montes, M., Pérez, M. y Vaufreydaz, D. *Comparación léxica de corpus para generación de modelos de lenguaje*. En: Proceedings of IBERAMIA workshop on Multilingual Information Access and Natural Language Processing, Sevilla, 2002.
- [50]. Wang, S. *Corpus-based approaches and discourse analysis in relation to reduplication and repetition*. En: Journal of Pragmatics, No. 37, 2005. pp. 505-540.
- [51]. Whittaker, S. y Stenton, P. *User Studies and the Design of Natural Language Systems*. En: Proceedings of the Fourth Conference of the European Chapter of the ACL, proceedings, 1989. pp. 116-123.
- [52]. Whittaker, S., Walker, M., y Moore, J. *Fish or Fowl: A Wizard of Oz Evaluation of Dialogue Strategies in the Restaurant Domain*. En: Proceedings of Language Resources and Evaluation Conference, Spain, 2002.
- [53]. Zhang, T., Hasegawa-Johnson, M. y Levinson, S. *A Hybrid Model for Spontaneous Speech Understanding*. En: Proceedings of the AAAI Workshop on Spoken Language Understanding, Pittsburgh, 2005a.
- [54]. Zhang, T., Hasegawa-Johnson, M. y Levinson, S. *Extraction of Pragmatic and semantic salience from spontaneous spoken English*. Aceptado para publicación en Speech Communication, 2005.