

**UNIVERSIDAD NACIONAL AUTÓNOMA DE  
MÉXICO**  
FACULTAD DE FILOSOFÍA Y LETRAS  
COLEGIO DE LENGUA Y LITERATURAS HISPÁNICAS

*ANÁLISIS DE ACTOS DEL HABLA CON EL ESQUEMA DIME-DAMSL:  
MODELACIÓN DE DIÁLOGOS PRÁCTICOS EN TRANSACCIONES  
PARA LA INTERACCIÓN HUMANO-COMPUTADORA*

TESIS  
QUE PARA OBTENER EL TÍTULO DE:  
LICENCIADA EN LENGUA Y LITERATURAS HISPÁNICAS  
PRESENTA:  
**VARINIA MARGARITA ESTRADA GARCÍA**

ASESOR: Dr. Luis Alberto Pineda Cortés

México, D.F. 2013

1. Introducción	6
<i>Objetivo</i>	
2. Nociones preliminares	8
<i>Lingüística computacional</i>	
<i>Procesamiento del lenguaje natural</i>	
<i>Sistemas de Diálogo para la Interacción Humano-Computadora</i>	
<i>Arquitectura básica de un sistema de diálogo</i>	
<i>Diálogos prácticos y Sistemas de diálogo</i>	
<i>Sistemas de diálogo desde una perspectiva pragmática</i>	
<i>Actos del habla y su modelación computacional</i>	
<i>Diálogos prácticos como esquemas de conductas conversacionales</i>	
3. Antecedentes: el Proyecto DIME y el Grupo Golem	22
<i>Proyecto DIME: 1998-2008</i>	
<i>Robot conversacional Golem</i>	
<i>Grupo Golem 2009-2012</i>	
<i>Adivina la carta</i>	
<i>Robot de servicio Golem-II+</i>	
4. Corpus DIME	28
<i>Motivación empírica: lingüística de corpus</i>	
<i>La tarea</i>	
<i>El escenario físico</i>	
<i>Recomendaciones, reglas y convenciones</i>	
<i>Niveles de análisis del Corpus DIME</i>	
<i>Segmentación de turnos y enunciados: transcripción del corpus</i>	
<i>Análisis con 'Speech View'</i>	
<i>Análisis pragmático con DIME-DAMSL</i>	
5. DAMSL	36
<i>Notas preliminares</i>	
<i>Draft of Dialogue Act Markup in Several Layers</i>	
<i>Estado comunicativo [EDO-COM]</i>	
<i>Nivel de información [NIV-INF]</i>	
<i>Actos del Habla como 'Funciones hacia adelante' y 'Funciones hacia atrás'</i>	
<i>Funciones hacia adelante [FUNC-ADELANTE]</i>	
<i>Funciones hacia atrás</i>	
<i>Alcances y limitaciones de DAMSL</i>	
6. DIME-DAMSL: Modificación de niveles y adición de etiquetas	60
<i>Distinción entre tipos de solicitudes de información</i>	
<i>Inclusión de acciones gráficas</i>	
<i>Gestos</i>	
<i>Acciones</i>	
<i>Organización de diálogos prácticos en unidades mayores</i>	
<i>Transacciones</i>	
<i>Planos de la expresión: 'Obligaciones' y 'Acuerdo mutuo'</i>	
<i>Balanceo de transacciones: cargos y abonos en los planos de expresión</i>	

7. Análisis del Corpus DIME con el esquema DIME-DAMSL	78
<i>Datos generales</i>	
<i>Anotación DIME-DAMSL</i>	
<i>Estado comunicativo y Nivel de información</i>	
<i>Funciones hacia adelante y Funciones hacia atrás</i>	
<i>Distribución de las Funciones hacia adelante y hacia atrás en los enunciados</i>	
<i>Análisis de transacciones</i>	
<i>Transacciones de tres turnos conversacionales</i>	
<i>Modelo de tres turnos: directivas de acción</i>	
<i>Modelo de tres turnos: preguntas imperativas</i>	
<i>Otras transacciones de tres turnos</i>	
<i>Transacciones de más de tres turnos conversacionales</i>	
<i>Protocolos de interacción</i>	
<i>Inclusión de protocolos de interacción a los modelos de tres turnos</i>	
8. Conclusiones	114
<i>Sobre el Corpus DIME</i>	
<i>Sobre el esquema DIME-DAMSL</i>	
<i>Trabajo Futuro</i>	
9. Bibliografía	118

### *Agradecimientos*

Muchas gracias, mamá, por tu valioso ejemplo, por tu inagotable amor, por darle sentido a mis pasos. Muchas gracias, Aurora, Jorge y Yiréh, por llenar mis días de alegría. Muchas gracias a mi familia, la de sangre y la elegida, por el apoyo y el impulso incondicional.

Gracias, Grupo Golem, por la paciencia, por el apoyo constante, por su gran corazón: Ivan Meza, Caleb Rascón, Lisset Salinas, Esther Venegas, Gibrán Fuentes, Mauricio Reyes, Hernando Ortega, Mario Peña, Joel Durán, Arturo Rodríguez, Iván Sánchez, muchas gracias por integrarme al grupo y compartirme sus aventuras. Especialmente, muchas gracias, Dr. Luis A. Pineda, por su incansable motivación, sus valiosos comentarios y su fe en mi trabajo.

*Tal vez hubo un error en la grafía  
o en la articulación del Sacro Nombre;  
a pesar de tan alta hechicería,  
no aprendió a hablar el aprendiz de hombre.*  
Borges

## 1. Introducción

Dentro del desarrollo de las tecnologías del lenguaje se crean métodos para la representación y el procesamiento computacional de contenidos lingüísticos. Estos contenidos permiten que los sistemas computacionales “interpreten” las señales lingüísticas que reciben y que “produzcan” respuestas. Las representaciones lingüísticas requeridas por el sistema pueden ser fonético-fonológicas, morfológicas, sintácticas, léxicas, semánticas, pragmáticas, etc. Cualquier nivel de análisis lingüístico representado computacionalmente corresponderá a lo que el sistema “conoce” y potencialmente puede “reconocer”. Por ejemplo, para el reconocimiento de habla se requiere la definición de un conjunto de unidades fonéticas (o un alfabeto) que represente los sonidos relevantes para determinada lengua. Una vez definidas las unidades de análisis se recopilan grandes bases de datos con suficientes muestras para cada unidad. Con ayuda de lexicones, heurísticas y análisis estadísticos, estas muestras permiten definir una serie de estrategias fonético-fonológicas para la combinatoria de palabras en esa determinada lengua.

### *Objetivo*

Conversar implica aprender desde edad muy temprana los lineamientos sobre cuándo hablar y cómo. Estas conductas conversacionales se especializan de acuerdo a los usos de la conversación y de las circunstancias enunciativas. El principal objetivo de este trabajo se centra en la modelación esquemática de las conductas que surgen en el ámbito de la conversación colaborativa. La inclusión de estos protocolos vislumbra el desarrollo de sistemas conversacionales que integren hábitos de la conversación a la interacción entre humanos y computadoras.

¿Es pertinente modelar actos del habla computacionalmente? Este cuestionamiento se realiza sobre el supuesto de que los *actos del habla* son unidades de análisis que pueden ser sometidas a una sistematización. La noción de actos del habla se ha definido de diversas maneras, se habla de fuerzas, de intenciones, de funciones comunicativas, de efectos. Como punto de partida, los actos del habla son concebidos como las acciones que se realizan a través de la

palabra (prometer, preguntar, afirmar, amenazar, etc.). Por el simple hecho de decir algo realizamos alguna de estas acciones, su enunciación tiene un impacto directo en el curso que seguirá el intercambio lingüístico.

Las actividades que podemos realizar a través de la palabra son ilimitadas y un modelo que las integre todas es impensable. Para la modelación computacional de la conversación es posible y pertinente enmarcar como uso lingüístico la conversación colaborativa. Si bien este uso lingüístico limita las estrategias comunicativas disponibles, sabemos que incluso en este marco del lenguaje no es posible sistematizar los protocolos interactivos de forma absoluta.

Para el cumplimiento del objetivo principal de esta investigación se siguió la metodología propuesta por el Proyecto DIME en el esquema DIME-DAMSL (Pineda, *et al*, 2006). Este esquema permite analizar los enunciados en conversaciones asignando etiquetas que corresponden a los actos del habla expresados en ellos. Analizar actos del habla en los lineamientos de este esquema conlleva un ejercicio de abstracción en el que se conjuntan lo que el hablante “quiso decir” en comparación con lo que dijo literalmente y siempre considerando los efectos que se provocan por decirlo.

Dentro de los motivos que impulsan este tipo de esquemas se encuentra el desarrollo de sistemas conversacionales que ofrezcan servicios de manera más dinámica y flexible. La integración de sistemas conversacionales para la prestación de servicios telefónicos está ampliamente extendida en instituciones y empresas, públicas y privadas. Hasta ahora se ha resuelto con una grabación que ofrece un menú numérico, la interacción con la máquina se realiza apretando botones e el teléfono.

A través de la observación de la conversación colaborativa DIME-DAMSL propone un método para el desarrollo de sistemas que intenten una iniciativa conversacional compartida con el usuario. Cada acto del habla cumple una función comunicativa específica y es posible encontrar vínculos y correspondencias entre estas funciones que permiten comprender la coherencia conversacional.

DIME-DAMSL propone la unidad de análisis conversacional de la “transacción” para representar las correspondencias de los actos del habla en unidades mayores. Los enunciados expresan actos del habla, los actos van ensamblando transacciones y a través de las transacciones quedan capturados los mecanismos en que los participantes definen metas, organizan planes de trabajo y negocian acuerdos (Pineda, *et al*, 2004). Con el esquema DIME-DAMSL, se analizó una muestra particular de lenguaje hablado llamada el Corpus DIME. En el presente trabajo se describe a detalle el esquema DIME-DAMSL y se documenta el análisis que con este esquema se realizó al Corpus DIME. Como conclusión, se proponen una serie de estrategias conversacionales o modelos de interacción que surgen a partir de este análisis. Estas estrategias representan las conductas conversacionales más frecuentemente observadas en el corpus.

La organización textual de este trabajo pretende la buena adquisición de sus contenidos. Por esta razón, comenzamos el Capítulo 2 con una serie de nociones preliminares que ubican esta investigación en su dominio teórico: lingüística computacional, sistemas de diálogo, modelación de actos del habla. El Capítulo 3 adscribe este trabajo dentro de los intereses generales del Grupo Golem, para ello se cuenta la evolución teórica de lo que originalmente fue llamado el Proyecto DIME. El Capítulo 4 presenta las características del Corpus DIME que es la base empírica de este trabajo. Posteriormente, los Capítulos 5 y 6 presentan DIME-DAMSL de manera descriptiva y detallada. Este par de capítulos explican y ejemplifican minuciosamente qué metodología se siguió durante la anotación del corpus, por esta razón, pueden considerarse un manual de etiquetación del esquema DIME-DAMSL. En el Capítulo 7 se presenta el análisis cuantitativo y cualitativo del corpus, se especifican los fenómenos más frecuentes, se exponen los casos más relevantes y se presentan los resultados de la anotación, tanto al nivel de los enunciados (los actos del habla que expresan) como al nivel de las transacciones (cómo se siguen y organizan los actos para concretar objetivos). Asimismo, a partir de este análisis, se van definiendo un par de modelos de interacción que toman como punto de partida la consistente y frecuente repetición de un mismo grupo de

actos del habla. Este capítulo evoluciona hacia la definición de este par de modelos y su posterior refinamiento con la integración de protocolos de interacción que posibilitan un intercambio más flexible de recursos discursivos. Este par de modelos capturan esquemáticamente las conductas conversacionales más estables en los diálogos colaborativos del Corpus DIME. Por último, el Capítulo 8 resume los alcances y limitaciones, tanto del esquema DIME-DAMSL como de los modelos resultantes del análisis del Corpus DIME. Además, se plantea el trabajo futuro y los potenciales vínculos que esta investigación puede establecer con otras disciplinas.

Los esfuerzos detrás de esta propuesta pueden ayudar a futuras investigaciones a la definición de un modelo computacional robusto que abarque cada vez más fenómenos conversacionales.

## 2. Nociones preliminares

### *Lingüística computacional*

La lingüística computacional se ha definido como la disciplina que se dedica al desarrollo de teorías, métodos, herramientas, aplicaciones y productos que someten al lenguaje a procesos computacionales o informáticos (Moure y Listerri, 1996). Definiciones similares a ésta se encuentran en la literatura sobre el tema, sin embargo, es posible observar que se trata de una definición amplia que abarca al menos dos terrenos de competencia: en uno de ellos se distingue por sus métodos, recursos y herramientas, es una lingüística que realiza el procesamiento de grandes corpus a través de sistemas informáticos; en el otro se caracteriza por sus objetivos y aplicaciones, es una lingüística que busca desarrollar sistemas computacionales que reconozcan, procesen y produzcan contenido lingüístico de forma automática. La primera perspectiva se emparenta con la lingüística de corpus y los estudios fonéticos, sintácticos, semánticos, etc. que ha realizado tradicionalmente la lingüística, sólo que ahora cuenta con herramientas computacionales que agilizan el análisis de gran cantidad de datos. La segunda perspectiva establece vínculos estrechos con el procesamiento de lenguaje natural (PLN), disciplina perteneciente a las ciencias de la computación en el campo de la inteligencia artificial (IA).

Sea para su estudio o para su implementación en tecnologías del lenguaje, la lingüística computacional aporta las pautas para la representación y procesamiento computacionales de unidades lingüísticas; es decir, representa *conocimiento lingüístico* a través de modelos formales.

*“Speech and language technology relies on formal models, or representations, of knowledge of language at the levels of phonology and phonetics, morphology, syntax, semantics, pragmatics and discourse. A number of formal models including state machines, formal rule systems, logic, and probabilistic models are used to capture this knowledge”* (Jurafski, 2000:15).

### *Procesamiento del lenguaje natural*

El PLN estuvo inmerso en la revolución que desde los años 50 se dio en el desarrollo de sistemas computacionales y el nacimiento de la inteligencia artificial.

Lo que distingue al PLN de otros procesamientos computacionales, dice Jurafski, es que procesa *conocimiento lingüístico*; es decir, sus unidades representan valores lingüísticos relevantes para su procesamiento.

En sus orígenes, el PLN tenía como meta inicial la traducción automática de textos. Las gramáticas libres de contexto permitieron pensar que las lenguas podían ser modeladas tomando como base su estructura sintáctica profunda, y que traducir automáticamente una oración de una lengua a una oración de otra consistía en crear un sistema que, basándose en una serie de reglas, empatara dos estructuras sintácticas de dos sistemas lingüísticos distintos.

Actualmente, el PLN ha multiplicado sus metas y diversificado sus metodologías. Como disciplina colabora con otras áreas del conocimiento como la lingüística, las ciencias de la computación, las matemáticas, la ingeniería, entre otras. Los desarrollos creados en el marco multidisciplinario del PLN se traducen en aplicaciones o tecnologías del lenguaje.

*“The goal of this new field is to get computers to perform useful tasks involving human language, tasks like enabling human-machine communication, improving human-human communication, or simply doing useful processing of text or speech”*  
(Jurafski, 2001:1).

El estado del arte en el desarrollo de estas tecnologías es notable en el procesamiento de habla y texto. En el procesamiento de texto se producen traductores multilingües, sistemas automáticos de extracción y recuperación de información, resúmenes automáticos, correctores ortográficos, sistemas de pregunta-respuesta, etc. Las tecnologías del habla, por su lado, buscan la creación de aplicaciones que permitan la interacción humano-computadora a través del lenguaje natural. Aunque falte todavía mucho en los alcances de las tecnologías del lenguaje, actualmente existen útiles aplicaciones disponibles para traducir, editar y procesar textos; lo mismo que reconocedores y sintetizadores de voz o sistemas de diálogo hablados o escritos.

El imaginario colectivo asume que las tecnologías del lenguaje culminarán con sistemas con capacidades lingüísticas como las de HAL en *2001: Una Odisea en el Espacio*, C3PO en *La Guerra de las Galaxias* o incluso la emotiva Robotina de *Los Supersónicos*. Estas máquinas, capaces de conversar sobre cualquier tema fluida y coherentemente con seres humanos, están lejos de ser una realidad. Por lo pronto, el PLN y otras disciplinas se enfocan en el desarrollo de tecnologías que posibilitan la interacción humano-máquina a través del lenguaje natural hablado para la ejecución de comandos básicos.

### ***Sistemas de Diálogo para la Interacción Humano-Computadora***

La Interacción Humano-Computadora (IHC) es una disciplina que tiene como meta posibilitar y mejorar el intercambio de información entre humanos y computadoras a través de distintos medios. Para ello, diseña, implementa y evalúa dispositivos, interfaces y aplicaciones computacionales en términos de su eficiencia interactiva. Cada sistema (*software* y *hardware*) describe el tipo de interacción que el usuario tendrá con él. Esta interacción depende del tipo de tarea que realiza el sistema y de los medios y recursos con que cuenta para hacerlo. En la actualidad, la interacción con una computadora se realiza principalmente a través de plataformas gráficas con ayuda de dispositivos como la pantalla, el teclado, la impresora, o más recientemente el *touch pad*. Estas interfaces son exitosas; sin embargo, podemos imaginar escenarios en que los usuarios prefieran dar instrucciones a través del lenguaje natural sin utilizar sus manos para apretar botones o sus ojos para ver a la pantalla. Los sistemas de diálogo o sistemas conversacionales son aplicaciones que buscan la interacción entre humanos y máquinas a través del lenguaje natural.

Los sistemas conversacionales son vistos como una alternativa poderosa para el intercambio informativo. La implementación de estos sistemas en diversos ámbitos cotidianos será una realidad para la sociedad del futuro. Desde esta perspectiva, la modelación computacional del lenguaje natural plantea un enorme reto para la lingüística computacional, la ingeniería lingüística, el procesamiento del lenguaje natural, la inteligencia artificial y otras disciplinas.

### Arquitectura básica de un sistema de diálogo

Desde una perspectiva computacional, los sistemas de diálogo están conformados por distintos módulos que realizan procesos específicos en el tratamiento del lenguaje (Llisterri, 2006). En la Figura 2.1 se muestra una arquitectura de un sistema de diálogo monomodal<sup>1</sup>.

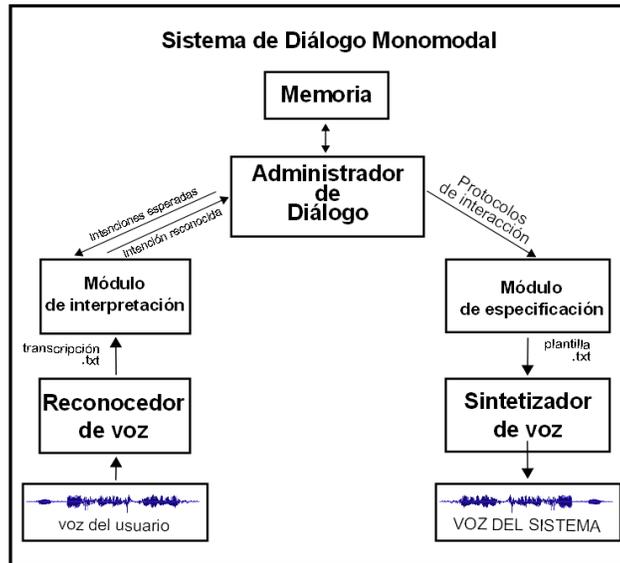


Figura 2.1 Arquitectura de un sistema de diálogo monomodal<sup>2</sup>

Al inicio del proceso, la señal acústica de la voz del usuario entra por el micrófono directamente al reconocedor de voz para ser convertida en una representación escrita. La transcripción que hace el reconocedor no es interpretada en términos de su contenido lingüístico. Hasta ahora, la voz se convirtió en una cadena de texto. El reconocedor envía el texto al módulo de interpretación que consulta al administrador de diálogo para “interpretar” el enunciado. El administrador consulta sus bases de datos o “memoria” para crear una respuesta coherente al enunciado interpretado, esta respuesta aun no tiene especificados ciertos parámetros del contexto conversacional. Para ser una respuesta determinada,

<sup>1</sup> La modalidad se refiere al modo de entrada y salida de información (*input-output*) que procesa un sistema. En un sistema monomodal el *input* y el *output* es lingüístico. Un sistema multimodal integrará otros modos de interacción (visuales, motrices, etc.). Los sistemas monomodales son primordialmente desarrollados para aplicaciones de asistencia telefónica. Los sistemas multimodales están pensados para sistemas que integran más funcionalidades de interacción, como sucede, por ejemplo, en un robot.

<sup>2</sup> La imagen es una reproducción de un sistema desarrollado por el Proyecto DIME.

el administrador envía su respuesta al módulo de especificación. Este módulo ordena la respuesta en una plantilla de texto y la envía al sintetizador de voz para que la “lea”. El usuario recibe una señal sonora que es la voz del sistema.

Potencialmente, los modelos que resultan de este trabajo pueden integrarse al módulo de administración del diálogo, de tal manera que permita una efectiva organización de los mensajes que entran y los que salen. Estas pautas de interacción conversacional fueron observadas en el corpus, los modelos resultantes de este trabajo tienen como objetivo restringir las opciones durante la interpretación de enunciados y fomentar la formulación de respuestas coherentes.

### ***Diálogos prácticos y Sistemas de diálogo***

La conversación es la manifestación lingüística favorita para la interacción humana cotidiana; es una actividad importante en el ámbito social que crea rutinas vinculadas a protocolos sociales. Cada comunidad desarrolla hábitos conversacionales que están ligados a contextos de uso.

*“It does seem clear that full natural language understanding by machine will not occur in the foreseeable future. Our belief, however, is that sufficient understanding of practical dialogues is feasible”* (Allen, et al, 2001:28).

Dentro de la amplia gama de los usos de la conversación, podemos hallar diálogos que son sumamente colaborativos y que son llamados diálogos prácticos. En este tipo de diálogo existe una tarea o meta común a resolver entre los interlocutores.

*“The goal-seeking nature of such conversations naturally creates a specific genre of conversation, which we call practical dialogues”* (Allen et al., 2000:2).

Lo interesante de la noción de diálogos prácticos es que se entiende como una modalidad conversacional que permite hacer cualquier cosa con alguien a través del lenguaje. Se trate de construir una casa o realizar un ambicioso proyecto, de comprar una hamburguesa o ayudar en una mudanza, los diálogos prácticos son entendidos como el “vehículo” lingüístico en que se realizan estas tareas conjuntas.

Son la abstracción de una estructura conversacional que subyace independiente al dominio o escenario en que se realizan los diálogos. Lo curioso es que si bien los diálogos prácticos son independientes del dominio, el dominio a su vez limita la conducta lingüística de los interlocutores durante un diálogo. Los temas conversacionales y el tipo de acciones que se realizan están limitados por la tarea.

En el ámbito de la modelación computacional, este par de características permiten ver una pauta para la creación de sistemas de diálogo eficientes. James Allen y su equipo lo postularon de la siguiente manera:

**- Hipótesis de los diálogos prácticos:**

*“The conversational competence required for practical dialogues, although still complex, is significantly simpler to achieve than general human conversation” (Allen et al., 2001:29).*

**- Hipótesis de la independencia del dominio:**

*“Within the genre of practical dialogue, the bulk of the complexity in the language interpretation and dialogue Management is independent of the task being performed” (Allen et al., 2001:29).*

Los diálogos prácticos comparten esencialmente las mismas estructuras de interacción subyacentes y restringen la conducta lingüística en comparación a la conversación general humana. Si las hipótesis son ciertas, dice Allen, es posible la creación de un modelo computacional genérico para diálogos prácticos. Un modelo de estas proporciones no será simple ni reducido y siempre estará sujeto a ciertos ajustes y especificaciones. El modelo deberá atender aspectos de la conversación que se dan en múltiples niveles comunicativos simultáneos. Deberá, además, integrarse a un sistema conversacional que incorpore el reconocimiento del habla, el procesamiento e interpretación del lenguaje, la administración del diálogo, la planeación de respuestas y acciones; y deberá tener, también, cierta flexibilidad para permitir su implementación en diversos dominios y tareas. (Allen et al., 1998; Pineda, 2008).

Si las computadoras sirven para hacer cosas y los diálogos prácticos son el tipo de conversación que surge cuando los humanos hacen cosas juntos, parece pertinente desear que en el terreno de la interacción entre humanos y computadoras esté disponible la modalidad lingüística.

### ***Sistemas de diálogo desde una perspectiva pragmática***

Originalmente, el PLN estuvo motivado por el desarrollo de las gramáticas libres de contexto de mediados del siglo pasado.

*“Chomsky (1956) first considered finite-state machines as a way to characterize a grammar and defined a finite-state language as a language generated by a finite-state grammar. These early models led to the field of formal language theory, which used algebra and set theory to define formal languages as sequences of symbols. This includes the context free of grammar” (Jurafski, 2000:9).*

Los lenguajes formales impulsaron el desarrollo de sistemas computacionales de forma explosiva. Sin embargo, en el marco del PLN interesado en los procesos de la significación lingüística, las gramáticas libres de contexto no permiten interpretar ciertos fenómenos comunes del habla, como son las metáforas o las ironías, por mencionar algunos deslices entre el sentido convencional y la forma en que se expresa. Esta situación tiene impacto directo en el campo de los sistemas conversacionales, pues no considerarla conlleva que el sistema interprete lo que está especificado en su gramática y no necesariamente lo que el hablante “quiso decir”. Con el paso del tiempo, el desarrollo de sistemas de diálogo fue integrando modelos con perspectivas de análisis un tanto más pragmáticas.

La pragmática es una rama de la lingüística que se interesa en los mecanismos, recursos y procesos de significación que se activan durante la formulación o interpretación de los enunciados en un contexto dado. Es decir, estudia los usos de la lengua. Como disciplina, surgió durante la segunda mitad del siglo XX a partir del trabajo de teóricos de la filosofía del lenguaje como Austin y Searle (1962). Originalmente, pusieron en el debate lingüístico y filosófico la

necesidad de atender ciertas manifestaciones del lenguaje que tradicionalmente se negaban a ser representadas de forma sistemática. Estas manifestaciones, defendían Austin y Searle, son tan lingüísticas como las proposicionales y están reguladas por contextos enunciativos, es decir, son convencionales.

Austin (1962) propuso que los enunciados están motivados por fuerzas *locucionarias*, *ilocucionarias* y *perlocucionarias*. Alternativamente a una gramática libre de contexto, desde esta perspectiva, “decir algo” siempre contiene un significado, una intención y un efecto que serán afortunados o no dependiendo del momento en que se enuncie. Estas fuerzas establecen vínculos y correspondencias discursivas ligadas a su momento de enunciación.

*“Intentamos un nuevo comienzo consistente en considerar los sentidos en los que decir algo es hacer algo. Distinguimos así el acto locucionario (y dentro de él los actos fonéticos, ‘fáticos’ y ‘réticos’) que posee significado; el acto ilocucionario, que posee una cierta fuerza al decir algo; y el acto perlocucionario, que consiste en lograr ciertos efectos por (el hecho de) decir algo” (Austin, 1962:168).*

Los niveles locutivo, ilocutivo y perlocutivo inspiraron la observación de múltiples aspectos presentes en la comunicación pues se distinguió entre lo que está dicho literalmente, lo que se quiso decir y los efectos que se causan por decirlo. De esta manera, la lingüística puede lidiar con el hecho de que en el debido momento las lenguas permiten decir cosas sin que estén dichas.

La propuesta de la pragmática propone un escenario de análisis que considera en paralelo distintas dimensiones comunicativas presentes durante la interacción lingüística. Desde la perspectiva de los sistemas de diálogo esta propuesta ofrece la integración de mayor número de elementos a considerarse relevantes durante la interacción. Entre mayor sea el conocimiento lingüístico del sistema, mayor será su capacidad para lidiar con las conductas lingüísticas que recibe y para generar la conducta lingüística que se espera.

### ***Actos del habla y su modelación computacional***

La teoría de los actos del habla sentó las nociones que inauguraron la pragmática como disciplina. Se postularon los actos del habla (*speech acts*) como unidades de análisis que representan las distintas acciones que se pueden realizar a través del lenguaje. Cuando hablamos, expresamos intenciones que son de diversa índole y que contribuyen de distinta manera a la conversación. Preguntamos, respondemos, ordenamos, pedimos, acordamos, ofrecemos, etc. Todas estas acciones corresponden a lo que tradicionalmente se conoce como “actos del habla”.

En el uso cotidiano de la lengua es común encontrar intenciones expresadas literalmente (“*Pásame la sal*”), sin embargo, también es común encontrar que no sucede así (“*¿Puedes pasarme la sal?*”). La distinción que la teoría de los actos del habla hace entre “actos del habla directo” y “actos del habla indirecto” permite representar las intenciones comunicativas de forma independiente a la forma en que se expresan. Esta perspectiva abre la posibilidad de analizar los enunciados como “órdenes” aunque no tengan la forma sintáctica de una imperativa.

Un modelo de actos del habla es una representación formal de las acciones lingüísticas que se pueden realizar y de los efectos que éstas pueden causar durante un contexto enunciativo controlado. El modelo asignará el valor de “orden” o algo similar a cualquiera de las formas sintácticas que haya aparecido mientras “quiera decir” o “intente decir” o “logre” que le pasen la sal, por ejemplo. Para especificar el modelo, se definen un conjunto de unidades pertinentes o etiquetas que corresponden a las acciones. Estas acciones están fuertemente vinculadas a las funciones que cumplen dentro de los enunciados. Podemos anticipar que las funciones comunicativas (explícitas o implícitas) pueden asociarse a representaciones sintácticas, de esta forma se pueden vincular significados convencionales marcados en el lenguaje.

Para su implementación en un sistema de diálogo, un modelo de actos del habla ayuda a establecer vínculos entre las funciones que cumplen los enunciados en los diálogos, posibilita la representación coherente de turnos conversacionales y la interpretación de los mensajes en el contexto enunciativo. Actualmente existen distintos modelos de anotación propuestos por diversos intereses teóricos, en cada

propuesta se define un conjunto de actos del habla (una serie de etiquetas o identificadores) y sus métodos de asignación. Los modelos se evalúan y validan analizando grandes corpus lingüísticos. Algunos de estos modelos utilizan métodos de anotación semi-automáticos. Otros realizan el análisis manualmente para crear bases de datos que sirvan de entrenamiento para un sistema computacional. Sin considerar las diferencias entre definiciones, terminologías, métodos o analizar la naturaleza del inventario de actos del habla listados en distintos modelos, todos ellos evalúan los enunciados buscando asignar significados a las “intenciones” que comunican, a las “funciones” que cumplen o a los “efectos” que provocan en el contexto del diálogo.

Dialogar es coordinar actos lingüísticos individuales de forma conjunta (Clark, 1996). La secuencia conversacional no es una cadena de actos lingüísticos enunciados al azar sin orden alguno, conversar implica mantener la coherencia de temas y acuerdos mediante la coordinación de las conductas lingüísticas. En el contexto de los diálogos prácticos esta coordinación se bifurca en dos sentidos colaborativos, por un lado se administra el dialogo como actividad conjunta y por el otro se planea y desarrolla la tarea como objetivo común. Es decir, el diálogo y la tarea se organizan, administran y ejecutan de forma independiente pero utilizando el mismo canal de comunicación.

Un enunciado puede hacer una aportación a la conversación desde diferentes dimensiones. Conllevará distintas conductas decir “*No te entiendo*” porque no se escucha y se está administrando el canal de comunicación, que decir “*No te entiendo*” porque faltan referentes y se está administrando el conocimiento compartido del mundo. Un modelo de actos del habla permite capturar esta multidimensionalidad comunicativa pues asume que cada enunciado puede hacer aportes comunicativos en distintos niveles. La noción básica es que el tipo de tarea que se soluciona puede variar, pero las formas en que se administra el diálogo y en que se realiza la tarea son similares. La modelación de actos del habla en diálogos prácticos permite la representación de intenciones comunicativas relevantes.

Preguntar, responder, pedir, aceptar, rechazar, ofrecer, etc. son las acciones lingüísticas que componen los diálogos prácticos.

Para este trabajo se utiliza como base el esquema desarrollado por James Allen y su equipo hacia finales de los 90, *Dialog Act-Markup in Several Layers*. Este esquema permite la representación multidimensional de los actos del habla detallando el tipo de aporte comunicativo que expresan. DAMSL fue publicado junto a un manual de fácil aprehensión y con el consejo constante de hacer los ajustes necesarios para su evaluación en otros corpus conversacionales.

*“For any particular project, we would expect that the annotation scheme would be refined to provide further retail on phenomena of interest”* (Allen y Core, 1997:2).

Siguiendo el consejo de sus creadores, un grupo de investigación llamado Proyecto DIME ajustó el esquema original a sus intereses de estudio particular creando DIME-DAMSL. La historia de su aprovechamiento y el detalle de los ajustes que se realizaron se expondrán puntualmente en los Capítulos 5 y 6.

### ***Diálogos prácticos como esquemas de conductas conversacionales***

La vinculación entre los actos del habla durante los diálogos se puede observar desde dos perspectivas: por un lado, podemos apreciarla de forma local, los actos del habla se muestran como moneda de cambio entre los enunciados de los interlocutores, uno pregunta y el otro responde, uno afirma y el otro rechaza, uno pide y el otro cumple, hay una correspondencia entre un acto del habla que se expresa y el que le sigue; por otro lado, podemos verla de forma más global, los actos del habla forman grupos que secuencialmente van resolviendo la tarea por partes. Para representar esta doble coordinación entre los actos del habla, el Proyecto DIME propuso que de forma local los enunciados se cohesionan en pares adyacentes del tipo “cargo/abono”; y que de forma global, las “transacciones” plasman las conductas esquemáticas que completan la tarea.

DAMSL se centra en las relaciones que los actos del habla tienden hacia atrás y hacia adelante. La integración de las transacciones en el esquema DIME-DAMSL permite construir modelos de interacción que contemplen más de un par

de actos. Involucrar grupos de actos de habla en la planificación y solución de objetivos específicos permite observar y representar los mecanismos internos que se activan durante la organización y ejecución de la tarea. La propuesta del Proyecto DIME permite representar la solución de tareas en diálogos prácticos como la solución incremental de transacciones. Para validar su propuesta el grupo analizó el Corpus DIME. Los modelos resultantes de este análisis pretenden representar estas conductas esquemáticas de forma robusta para su futura implementación en un sistema de diálogo flexible.

### **3. Antecedentes: el Proyecto DIME y el Grupo Golem**

Este trabajo se adscribe a los intereses del Grupo Golem, equipo de investigación que se dedica al desarrollo de tecnologías en el área de la inteligencia artificial. Estas tecnologías son diversas y abarcan el procesamiento del lenguaje natural, la visión computacional, la navegación automática, la manipulación de objetos y la implementación de estas funcionalidades en robots de servicio. El grupo es coordinado por el Dr. Luis A. Pineda Cortés y trabaja en el Departamento de Ciencias de la Computación del Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas (DCC-IIMAS) de la UNAM. El antecedente del Grupo Golem es el Proyecto DIME. A continuación se cuenta brevemente su historia y su evolución metodológica.

#### ***Proyecto DIME: 1998-2008***<sup>3</sup>

En 1998 surgió el Proyecto DIME (Diálogos Inteligentes Multimodales en Español). El interés primordial de este proyecto se centró en el desarrollo de tecnologías para el español de la Ciudad de México, en particular, de sistemas conversacionales multimodales<sup>4</sup>. En un primer momento, el grupo pensó que era posible crear estos sistemas a partir del cumplimiento de tres objetivos: 1) la compilación, la transcripción y el análisis de un corpus conversacional multimodal; 2) el desarrollo de un reconocedor de habla para el español de la Ciudad de México; y 3) la definición de una gramática para el español y de un analizador sintáctico automático o *parser* que comprendiera el lexicón y la gramática observadas en el corpus recopilado. Esta perspectiva estaba fuertemente influenciada por las gramáticas generativas, que apuntan que el nivel de representación sintáctico es autónomo y principal, y que la interpretación del lenguaje puede obtenerse de la inducción sintáctica (Pineda, *et al*, 2002; Pineda, 2008).

---

<sup>3</sup> <http://leibniz.iimas.unam.mx/~luis/DIME/>

<sup>4</sup> 'Multimodal' se refiere al modo de interacción en términos de *input-output* informativo. Los sistemas multimodales administran información proveniente de distintos dispositivos de entrada y salida, como son bocinas, micrófonos, cámaras, motores, sensores, etc. En contraste, un sistema monomodal sería, por ejemplo, un sistema de diálogo para asistencia telefónica que utiliza de entrada y salida la señal acústica.

Los esfuerzos del grupo se dirigieron al cumplimiento de estos objetivos. El primero produjo el Corpus DIME<sup>5</sup>. El segundo objetivo, el reconocedor de voz, planteó tareas específicas que se desarrollaron de forma exitosa y paralela (Pérez-Pavón, 2006). El tercer objetivo, la gramática para el español, resultó en un modelo de los verbos auxiliares de las perífrasis verbales (Meza, 2002).

Conforme el análisis sintáctico fue avanzando, se fue haciendo evidente que la interpretación de los significados intentados por los hablantes no siempre resulta de la inducción sintáctica de las oraciones. Por un lado, los hablantes no formulan enunciados “gramaticalmente”. Todo el tiempo ocurren silencios, tropiezos y reformulaciones durante la enunciación. Por otro lado, los enunciadores en un diálogo difícilmente siguen las regulaciones de la gramática prescriptiva; la lengua en uso es vehículo de estrategias de significación complejas que dependen de las circunstancias enunciativas. La modelación conversacional resultó en algo más complejo que un análisis sintáctico. El grupo se encontró de frente con una vieja problemática para la lingüística tradicional: la lengua en uso.

Consideremos los siguientes ejemplos tomados del Corpus DIME:

- 1) bueno este el extractor con repisas de madera lo puedes colocar en sobre la co[cina] sobre la estufa ?
- 2) <sil> la estufa pegar vamos a quitar <sil> a <sil> a intercambiar vamos a poner este fregadero esto <sil> de este lado de acá y la estufa de este lado de acá

En ambos ejemplos ocurren difluencias del habla. El Ejemplo 1) muestra que los hablantes reformulan y organizan los enunciados conforme hablan. El Ejemplo 2) invita a cuestionar profundamente qué es “interpretar” un enunciado y a explorar representaciones alternativas a la sintáctica para los procesos de interpretación lingüística. Para el Proyecto DIME, encontrarse con que la predicción sintáctica no resultaba del todo precisa conllevó un cambio paradigmático en el desarrollo de sus sistemas de diálogo.

---

<sup>5</sup> Dado que es el recurso empírico utilizado para este trabajo, se describe a detalle en el Capítulo 4.

*“La dificultad de inferir la intención de un agente a partir de la estructura sintáctica de la oración, en un contexto conversacional dado, nos llevó a cambiar la orientación del proyecto y adoptar al nivel pragmático e intencional como el nivel de representación central para analizar la comunicación lingüística” (Pineda, 2008:2).*

Este cambio de paradigma asume que las conversaciones se mantienen a partir de la producción y comprensión de intenciones. “Desde este punto de vista, una elocución es el vehículo de una acción que se realiza por medio del lenguaje o, de acuerdo a la terminología tradicional, un acto del habla” (Pineda, 2008:27).

La noción de “acto del habla” creó hace más de 50 años una rama de estudio lingüístico. Para el Proyecto DIME, incorporar esta perspectiva de análisis creó un nuevo objetivo: la definición de un modelo de actos del habla que capture las intenciones comunicativas de los diálogos del Corpus DIME. Con este giro surgió el esquema de anotación de actos del habla DIME-DAMSL.

Con el paso del tiempo, el Proyecto DIME maduró lo suficiente sus recursos lingüísticos y tecnológicos para desarrollar sistemas conversacionales básicos. A continuación hacemos un breve resumen de ellos:

#### *Robot conversacional Golem*

En 2005, el Proyecto DIME implementó un sistema conversacional sencillo en un pequeño robot llamado Golem (Figura 3.1). Este robot daba visitas guiadas por los proyectos de investigación del cuarto piso del IIMAS. Las investigaciones estaban plasmadas en pósters que Golem explicaba a través de una conversación simple con el usuario. Golem podía moverse de póster a póster, hablar sobre sí mismo y contar uno que otro chiste. Fue presentado en diversos espacios académicos y universitarios del país. Se jubiló en 2008.



Figura 3.1 *Robot conversacional Golem*

### ***Grupo Golem 2009-2012<sup>6</sup>***

En 2009 el Proyecto DIME se renombró Grupo Golem en homenaje al primer robot. El cambio de nombre no alteró el objetivo de este grupo de investigación, mantiene el interés en la creación de sistemas de interacción multimodal.

### ***Adivina la carta***

En colaboración con el museo de las ciencias Universum, el grupo implementó un sistema conversacional multimodal en una aplicación fija que juega “Adivina la carta” (Figura 3.2). Esta plataforma se exhibe de forma permanente en el museo desde 2010 (Meza, *et al*, 2010). Los usuarios son primordialmente niños que hacen visitas escolares.



Figura 3.2 *Adivina la Carta en Universum*

Las cartas que motivan el juego son imágenes con el tema del “universo”: planetas, galaxias, estrellas, etc. El sistema elige al azar una de diez cartas a las que

---

<sup>6</sup> <http://golem.iimas.unam.mx>

el niño tiene acceso visual. El niño tiene que “adivinar” qué carta eligió el sistema. Las pistas las consigue preguntándole cosas como: “¿Tu carta tiene algo rojo?”, “¿Es azul?” o “¿Tiene vida?”



Figura 3.3 Sistema Adivina la Carta

El sistema reconoce palabras clave y contesta, permitiendo que el usuario realice operaciones de eliminación y deduzca qué carta es. La interacción es a través del lenguaje y el reconocimiento visual. La Figura 3.3 presenta las imágenes que el sistema muestra a sus usuarios en la pantalla. Se evaluó el sistema en términos del reconocimiento del habla, de si los niños lo consideraban interesante y se determinó que funciona adecuadamente (Venegas, 2010). Los esfuerzos en la implementación de esta plataforma están motivados por el interés de difundir los desarrollos en sistemas conversacionales en español en el museo de ciencia y tecnología de la UNAM.

#### *Robot de servicio Golem-II+*

El grupo se concentra actualmente en el desarrollo de Golem-II+ (Figura 3.4). Al igual que su hermano mayor, es un robot de servicio que interactúa por la vía lingüística y que navega a través de ambientes dinámicos. Con el paso del tiempo, el Grupo Golem ha mejorado las funcionalidades de sus sistemas. En el área de visión computacional ha integrado el reconocimiento visual de objetos, personas y rostros. Desde la perspectiva de la interacción humano-robot, Golem-II+ puede orientarse hacia el usuario a partir de la localización de fuentes acústicas; de esta manera, se atiende la necesidad de la comunicación “cara a cara” entre el usuario y el robot. Golem-II+ integra un par de brazos que le permiten tomar objetos e interactuar con el mundo.



Figura 3.4 *Golem-II+*

La propuesta detrás de Golem-II+ surge del mismo lugar que los anteriores desarrollos del grupo: a partir de estructuras de diálogo o modelos de diálogo es posible administrar la interacción con el usuario independientemente de si la información requerida es visual, lingüística, motriz, etc.

El sesgo original de crear sistemas multimodales sigue motivando la integración de funcionalidades y modos de interacción en un mismo orden de administración (lenguaje, visión computacional, navegación automática, manipulación de objetos). Para validar el trabajo realizado, el grupo ha decidido participar en la competencia RoboCup@Home<sup>7</sup>. Esta competencia ha permitido probar a Golem-II+ y sus capacidades en la solución de tareas de servicio del hogar. Cada prueba de RoboCup@Home involucra un tipo de interacción distinta; en una el robot sigue al usuario, en otra reconoce gente, en otra tiene que escuchar una orden del tipo “*trae la botella de jugo que está en la cocina*” y ejecutarla, etc.

De forma sintética, este ha sido el camino que, desde sus inicios, el Proyecto DIME y, posteriormente, el Grupo Golem ha seguido. A continuación se presenta el Corpus DIME y los distintos niveles de análisis en los que ha sido observado.

---

<sup>7</sup> [www.robocup.com](http://www.robocup.com)

#### **4. Corpus DIME**<sup>8</sup>

##### ***Motivación empírica: lingüística de corpus***

El Proyecto DIME diseñó un experimento con la intención de recopilar un corpus oral conversacional del español hablado de la Ciudad de México (Pineda, *et al*, 2002). El Corpus DIME es un recurso lingüístico característico conformado por un conjunto de 31 diálogos prácticos realizados por 14 informantes (7 hombres y 7 mujeres entre los 30 años de edad). Todos ellos con estudios universitarios. El dominio en que se desarrollan estos diálogos es el diseño de cocinas.

*“The DIME Corpus is a multimodal corpus that provides empiric information for studying the use and interaction between spoken language, deictic gestures and the graphical context during human-computer interaction” (Villaseñor, et al, 2001:1).*

Este recurso ha sido la fuente de estudio de diversos fenómenos lingüísticos observables en el habla. Desde una perspectiva fonética, el corpus inspiró la definición de *Mexbet*, un alfabeto fonético computacional para el español mexicano (Cuétara, 2004); también permitió el estudio y sistematización de ciertos fenómenos fonéticos, como la diptongación en el habla del español de la Ciudad de México (López, 2004), así como la silabificación (Espinoza Cruz, 2007) y pérdida fonética (Ceballos, 2007). Desde una perspectiva sintáctica, el corpus permitió la creación de un sistema para verbos auxiliares de las perífrasis del español (Meza, 2002). Asimismo, permitió el estudio de la formulación, corrección y reformulación de enunciados a partir del análisis de las reparaciones del habla (Moreno, 2006). Desde una perspectiva más pragmática, este corpus también sirvió de base empírica para el reconocimiento de patrones prosódicos asociados a intenciones comunicativas (Coria, 2008). Todas estas investigaciones fueron desarrolladas en el marco del Proyecto DIME. Cada investigación requirió la creación de metodologías y herramientas de etiquetación del corpus.

El Corpus DIME constituye un recurso empírico singular en su tipo, pues conjunta una muestra lingüística para el estudio de diversos fenómenos del español

---

<sup>8</sup> <http://leibniz.iimas.unam.mx/~luis/DIME/CORPUS-DIME.html>

de la Ciudad de México en uso conversacional. A continuación se describen la tarea y el escenario (virtual y real) en que se recopiló.

### **La tarea**

Para modelar diálogos orientados a la solución de tareas se diseñó un experimento: se le pidió a los usuarios que diseñaran o modificaran cocinas a través de un software CAD<sup>9</sup> llamado *Home Designer*.

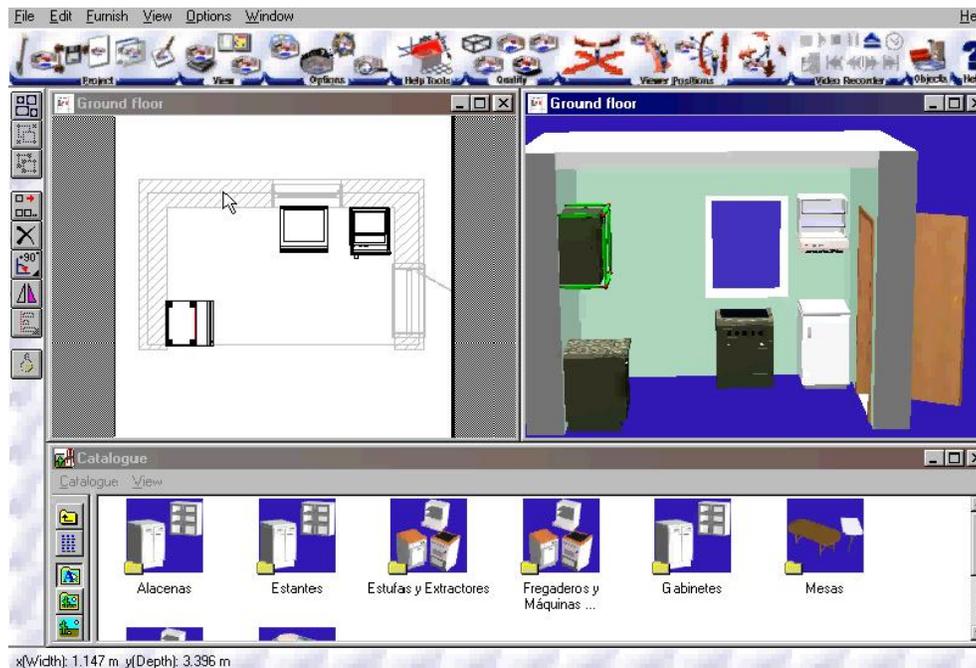


Figura 4.1 Interfaz CAD del Corpus DIME

En la Figura 4.1 podemos observar el escenario virtual del Corpus DIME. La imagen muestra tres ventanas principales. Dos en el centro arriba que dividen la pantalla por la mitad, estas ventanas muestran representaciones gráficas del espacio de diseño y tienen el mismo título: *Ground Floor*. La ventana de la izquierda es una vista 2D de la cocina desde arriba; la de la derecha es la vista 3D de la misma cocina vista de frente. Los objetos podían colocarse en una u otra ventana indistintamente. Justo debajo de estas dos ventanas aparece otra llamada *Catalogue*; en ella se encuentra el directorio o catálogo de objetos para colocar en la cocina (estufas, mesas, sillas, etc.).

<sup>9</sup> Por sus siglas en inglés: *computer-aided design*.

En cada diálogo, un “usuario” (*u*) diseña o modifica una cocina con la ayuda del “sistema” (*s*). Se definió esta tarea porque tiene un nivel de complejidad medio: es bastante sencilla para que el usuario pueda completarla, pero un tanto compleja para requerir la ayuda del sistema. Se trata de una tarea que no requiere un aprendizaje previo de terminología, pues el dominio de la cocina es un campo que engloba conocimiento compartido por gran parte de la sociedad. Tampoco era necesario que el usuario aprendiera a utilizar el programa CAD, pues el sistema es el que realiza las acciones que el usuario solicita (Villaseñor, *et al*, 2001).

### ***El escenario físico***

En la Figura 4.2 se muestra la disposición física del experimento. Se contaba con dos cuartos separados pero contiguos para que los participantes no se vieran el uno al otro: el cuarto del usuario y el cuarto del sistema respectivamente<sup>10</sup>.

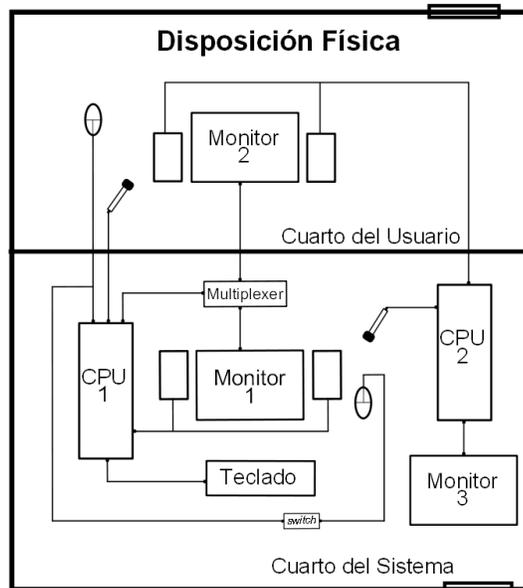


Figura 4.2 Disposición física para la recopilación del Corpus DIME

<sup>10</sup> La descripción del Corpus DIME menciona el “Protocolo del Mago de Oz” en la creación del experimento (Villaseñor, *et al*, 2001; Pineda *et al*, 2002). El citado protocolo tiene como característica principal hacerle creer a los informantes que interactúan con una máquina. Esta propuesta se basa en la idea de que las personas ajustan su conducta, sea lingüística o no, al interactuar con computadoras. El Corpus DIME siguió de cerca los lineamientos generales de este protocolo, sin embargo, los informantes siempre supieron que su interlocutor era un humano.

### ***Recomendaciones, reglas y convenciones***

La distinción entre los roles de “usuario” y “sistema” en los diálogos es de fundamental importancia. El tipo de acciones que puede realizar uno y otro intentaron ser limitadas por los objetivos del Proyecto DIME. El usuario necesita la ayuda del sistema para solucionar la tarea, no puede tomar el control obviando al sistema. El sistema, por su parte, es un experto en el manejo del programa de diseño y tiene información que el usuario desconoce. El sistema puede sugerir cambios, puede notificar errores y puede crear planes de trabajo; sin embargo, la última palabra sobre las acciones a realizarse la tiene siempre el usuario. El Proyecto DIME estableció una serie de indicaciones para restringir las conductas de los participantes<sup>11</sup> (Villaseñor, *et al*, 2001:4).

### ***Niveles de análisis del Corpus DIME***

#### ***Segmentación de turnos y enunciados: transcripción del corpus***

Los audios de los diálogos fueron segmentados en turnos y en enunciados, cada enunciado se transcribió manualmente en un archivo *.txt* que incluía un identificador correspondiente al número de diálogo y al número de enunciado distinguiendo si fue expresado por el usuario o por el sistema.

En diálogos prácticos es muy común encontrar una alternancia de turnos coordinada. Por lo general, los traslapes ocurren en contextos en que los interlocutores avisan que están prestando atención o que están de acuerdo (“ajá”, “sí”). Sin embargo, a veces los interlocutores hablan al mismo tiempo. La transcripción puede solucionar ese problema poniendo una marca de la que se hablará posteriormente.

La segmentación de enunciados es complicada. No queda siempre claro dónde empieza o dónde acaba un enunciado en un mismo turno. Algunos aspectos fonéticos y fonológicos permiten establecer ciertos límites a partir de la

---

<sup>11</sup> Los lineamientos de conducta son guías que se cumplieron en la medida de lo posible. El diseño del experimento buscaba controlar variables para tomar una muestra del lenguaje en uso. Lo cierto es que el lenguaje en uso se comporta de formas extraordinarias e impredecibles. Fue un reto para la persona que representó al “sistema” lidiar con las restricciones al uso convencional del lenguaje que se le fijaron, pues en las conversaciones ocurrieron fenómenos comunicativos no esperados del todo.

información prosódica (tono, volumen, tiempo, silencio). Sin embargo, la naturaleza de esta información puede ser muy variada, un turno y un enunciado pueden tener una carga entonativa delimitada por pausas, pero también puede ser un conjunto de sonidos vocálicos o no vocálicos que tienen aporte comunicativo (“*mmmm*”, “*eeh*”, “*shhhh*”), pueden ser monosílabos, palabras o incluso frases extensas. Los enunciados no siempre corresponden al orden gramatical de las oraciones, pero transmiten significados exitosamente.

Para transcribir el Corpus DIME se establecieron una serie de lineamientos<sup>12</sup>:

- un turno se compone al menos de un enunciado
- un enunciado puede estar delimitado por cambios en el tono de voz
- el inicio de una categoría gramatical mayor (oración o frase) ayuda a definir fronteras
- las pausas considerables pueden marcar límites de enunciados (sin embargo, puede haber enunciados largos con silencios entre sus frases)
- la transcripción de un enunciado debe reflejar los fenómenos del habla, incluidos chasquidos, interjecciones, pausas y sonidos vocálicos y no vocálicos

En el Ejemplo 4.1 se muestra la transcripción ortográfica de un fragmento del Corpus DIME. Al inicio se señala el número de diálogo (*D*). Para la identificación de los enunciados se estableció la anteposición del prefijo *utt* (del inglés *utterance*) seguido del número de enunciado en el diálogo. Después se indica con una 's' si fue expresado por el sistema o con una 'u' si fue expresado por el usuario.

#### Ejemplo 4.1

```
D06
utt260:s: <no-vocal> el sillón obstruye un poco la e[ntrada]
un poco la entrada
utt261:u: <no-vocal> <sil> podrías subirlo un poquito hacia
arriba ?
utt262:s: okey
utt263:s: + + así ? + +
utt264:u: + + no + +
utt265:u: <no-vocal> podrías quitarlo ?
```

La transcripción intentó reflejar “fielmente” la producción del hablante. Por ello, se apunta la expresión de ruidos, sean vocálicos <*vocal*> o no vocálicos <*no-*

---

<sup>12</sup> Basados en VERBMOBIL-2 (Alexandersson, *et al*, 1997) y el Proyecto MATE (*Multilevel Annotation Tools Engineering*).

*vocal*> como los chasquidos. También se anotan las pausas <*sil*>. Asimismo, si una palabra se abandona en un enunciado, se completa con los símbolos “[ ]” (como en *e[ntrada]*). Las convenciones de transcripción permiten señalar si los enunciados se traslapan con los símbolos “+ +”; en el ejemplo, el usuario y el sistema hablan al mismo tiempo en los enunciados 263 y 264, en la transcripción se muestra que el sistema pregunta “¿*así?*” al mismo tiempo que el usuario dice “*no*”.

A continuación se presenta la Tabla 4.1 que resume totales y promedios obtenidos en la transcripción del Corpus DIME (Villaseñor, *et al*, 2001:8):

Tabla 4.1 *Totales del Corpus DIME transcrito*

Total (31 diálogos)	Promedio (por diálogo)
27459 palabras	886 palabras
5779 enunciados	185 enunciados
3606 turnos	115 turnos
7:10 horas	14 minutos de duración

#### *Análisis con 'Speech View'*

*Speech View* es una herramienta de análisis lingüístico desarrollada por el Centro para el Entendimiento del Lenguaje Hablado, por sus siglas en inglés CSLU (*Center for Spoken Language Understanding*), integrada en el *CSLU Toolkit*<sup>13</sup>. Este programa permite estudiar diversos fenómenos lingüísticos y asignar distintas etiquetas a un mismo enunciado sincronizándolas con la señal sonora y el tiempo de enunciación. Con esta herramienta, es posible un análisis multidimensional de lo que ocurre en la expresión de un enunciado en el tiempo. Esta interfaz permite la representación gráfica de oscilogramas, espectrogramas, frecuencia fundamental e intensidad de la señal acústica de los archivos de audio que se analizan. Adicionalmente, permite la alineación de estas gráficas con la inclusión de líneas de trabajo o capas (*layers*) que permiten asignar valores o comentarios. No hay límite en el número de líneas de análisis. Con la información que se anota en estas capas se crean archivos *.txt* en los que se guardan representaciones lingüísticas (sean fonéticas, léxicas, morfológicas, etc.) con información temporal.

<sup>13</sup> <http://www.cslu.ogi.edu/toolkit/index.html>



diseñó una cocina desde el principio. En los once restantes, la tarea consistió en hacer modificaciones a una cocina previamente diseñada. Se analizaron poco más de 4 horas de grabaciones.

Para este análisis no se utilizó la herramienta *Speech View* dado que la etiquetación de actos del habla requiere la observación de los enunciados considerándolos en el contexto de la conversación y no como unidades aisladas; sin embargo, es posible crear análisis cruzados entre los valores asignados en distintas anotaciones. Como este análisis es el objeto principal de esta investigación se presenta completo el esquema DIME-DAMSL y su metodología en los capítulos siguientes.

## 5. DAMSL

### *Notas preliminares*

#### *Antecedentes*

Un grupo multidisciplinario de investigadores de diversas universidades denominado *Multiparty Discourse Group* se reunió tres veces entre 1996 y 1997 en lo que llamaron *Discourse Research Initiative*; se juntaron con la intención de sumar recursos y desarrollos provenientes de distintas investigaciones para crear un marco teórico y empírico de alto nivel para el análisis de actos del habla en corpus conversacionales. De estas reuniones surgió *Dialogue Act Markup in Several Layers* o DAMSL, un modelo de anotación de actos de diálogo o actos del habla que cubre distintos niveles de análisis pragmático. La traducción más cercana al español sería algo así como “Anotación de actos de diálogo en varios niveles”. La originalidad de DAMSL como esquema de anotación de actos del habla se concentra en su carácter multidimensional; esto significa que cada enunciado es valorado desde diversas perspectivas y se le asignan etiquetas (con atención en el plural) que identifican el nivel conversacional en que se expresa, las funciones comunicativas que cumple y los efectos que provoca en el diálogo.

Los corpus conversacionales anotados con DAMSL son un recurso empírico de utilidad en dos sentidos: desde una perspectiva lingüística, el análisis con este esquema permite establecer pautas de organización dialógica que aparecen en la solución de tareas; desde una perspectiva computacional, esta base empírica sirve para el entrenamiento de sistemas conversacionales.

James Allen y Mark Core publicaron en 1997 un “borrador” que describe la última versión de DAMSL. Se trata de un manual en el que se explica paso a paso cómo se valora un enunciado cualquiera con este esquema. Asimismo, incluyen una serie de árboles de decisión que auxilian al anotador en la elección de la etiqueta.

*“This manual describes a system for annotating dialogs. It marks important characteristics of utterances that indicate their role in the dialog and their relationship to each other”*  
(Allen y Core, 1997:2).

### *Etiquetas y la ejemplificación*

#### *Etiquetas*

La exposición de DAMSL en este trabajo requirió una serie de decisiones. Los niveles de anotación y sus unidades se van presentando junto con su etiqueta entre corchetes [ ]. Las mayúsculas representan niveles de análisis y las minúsculas representan las unidades o los actos del habla incluidos en esos niveles. Los nombres de los niveles de análisis, de los actos presentados y sus etiquetas fueron traducidos del inglés literalmente. Vale la pena resaltar que las etiquetas no tienen acentos para su óptimo manejo computacional; por ello, pueden aparecer [monologo], [opcion-abierta] o [dir-accion].

#### *Ejemplificación*

Conforme se van presentando los niveles y sus unidades se van mostrando ejemplos en tablas. Todos los ejemplos son tomados del Corpus DIME, para identificarlos, se introducen con una D#) seguida de una cifra que refiere al número de diálogo. Las etiquetas y los enunciados aparecen en columnas distintas. Los diálogos se transcriben siguiendo las especificaciones expuestas en el Capítulo 4 de este trabajo. Cada tabla de ejemplos se verá más o menos así:

D#)

<b>etiqueta tópico</b>	utt#: <enunciado>
etiqueta	utt#: <enunciado>

El principio básico de DAMSL es la anotación de varios niveles; esto quiere decir que los enunciados reciben más de una etiqueta al mismo tiempo. Por cuestiones de espacio y para una adquisición más sencilla de la metodología de anotación, las tablas irán mostrando incrementalmente el carácter multidimensional del análisis. Al principio, cada tabla mostrará sólo la etiqueta que se explica; conforme avance la descripción del modelo, los enunciados se presentarán con más etiquetas. En negritas se resalta la etiqueta o grupo de etiquetas que son tópico. Es necesario recalcar que aunque no se muestre en los ejemplos, los enunciados reciben más de una etiqueta siempre.

Por último, DAMSL está diseñado para anotar diálogos en que los participantes no comparten información visual, como cuando se habla por teléfono.

Dadas las características del Corpus DIME, se incluyeron etiquetas correspondientes a la ejecución de acciones no lingüísticas como mover, quitar o poner objetos en la cocina que se diseña. La ejecución de estas acciones tiene un impacto directo en el curso que sigue la conversación. Los ajustes del modelo para crear DIME-DAMSL se desarrollarán en el Capítulo 6 con mayor detalle; por lo pronto, adelantamos que los ejemplos incluyen etiquetas correspondientes a estas acciones no lingüísticas. Si bien DAMSL no las contempla, es un modelo flexible que no impide analizarlas.

*Draft of Dialogue Act Markup in Several Layers*<sup>14</sup>

DAMSL inicia con una definición de diálogo con la que delimita el tipo de interacción a la que se dirige: “A dialogue is a spoken, typed or written interaction in natural language between two or more agents” (Allen y Core, 1997:2). Los autores puntualizan que se trata de un modelo genérico para la anotación de conversaciones orientadas a la solución de tareas; con modelo genérico se refieren a que puede utilizarse para analizar diálogos prácticos provenientes de distintos dominios. Para utilizar DAMSL se requiere un corpus conversacional segmentado en turnos y enunciados. Para cada enunciado DAMSL propone una valoración de distintos aspectos comunicativos.

*“For each utterance, the annotation involves making choices along several dimensions, each one describing a different orthogonal aspect of each utterance unit.”* (Allen y Core, 1997:2).

*Estado comunicativo [EDO-COM]*

En este nivel se marcan los enunciados que no aportan contenido o que su estado comunicativo es muy bajo. Esto puede deberse a diferentes motivos, DAMSL considera principalmente tres tipos: abandonos, monólogos y enunciados ininteligibles.

- **Abandono [abandono]:** enunciados en los que el hablante se detiene y deja la formulación lingüística sin mayores consecuencias en el flujo conversacional.

---

<sup>14</sup> <http://www.cs.rochester.edu/research/speech/damsl/RevisedManual/>

D17)

abandono	utt11:u: + ahora vamos a poner +
sol-info	utt12:s: + <ruido> así ? +
resp	utt13:u: así

- **Monólogo [monologo]:** enunciados en los que el hablante parece “pensar en voz alta”. Se ha observado que en diálogos prácticos los monólogos ocurren poco antes de tomar una decisión, mientras el hablante repasa el plan o las alternativas durante la planeación de la tarea.

D02)

opcion-abierta	utt120:s: éstas son las mesas que tenemos
monologo	utt121:u: mesa uno mesa dos mm <sil> barra de desayuno

- **Ininteligible [ininteligible]:** enunciados incomprensibles, como cuando hay ruido en el entorno o el hablante articula de forma poco afortunada.

D18)

ininteligible	utt177:u: eh <sil> los dos muebles tienen que estar <sil> <no-vocal> no ?
---------------	--

Éstas son las tres etiquetas pertenecientes al nivel de análisis del estado comunicativo. Para asignarlas es posible observar los enunciados en su contexto conversacional y valorar su aportación a partir de las consecuencias que provocó su expresión. Los enunciados ininteligibles muchas veces son respondidos por una pregunta del tipo “¿Qué?” o “¿Perdón?”, pero también es muy común encontrar que la conversación sigue sin que los participantes se detengan a clarificar lo enunciado. Los monólogos y los abandonos por lo general son escuchados pero no son atendidos. Una recomendación que hace DAMSL cuando se pretende valorar un enunciado en este nivel es probar qué pasa si se “quita” el enunciado. Si la conversación no pierde coherencia es muy probable que el enunciado no haya tenido aporte comunicativo.

Una característica singular de este nivel de análisis es que si un enunciado recibe alguna de sus etiquetas ya no se analiza en los demás niveles de DAMSL; es decir, se asume que no hizo un aporte comunicativo a la conversación y se pasa al siguiente enunciado. En cambio, si el enunciado no recibió ninguna de estas etiquetas entonces el análisis continúa para especificar con nitidez exactamente qué tipo de aporte comunicativo hizo.

*Nivel de información* [NIV-INF]

En este nivel se distingue el contenido informativo de un enunciado en tres categorías: administración de la comunicación, administración de la tarea y tarea.

- **Administración de la comunicación [admin-com]:** se marcan con esta etiqueta los enunciados que monitorean el éxito comunicativo, sea para mantener el flujo de la conversación actualizado o para manifestar que no se entendió.

*“The utterances should be marked at the Communication-management level only when they make no direct contribution to solving the task”* (Allen y Core, 1997:8).

D02)

admin-com	utt4:s: me escuchas ?
admin-com	utt5:u: sí
admin-com	utt6:s: okey

- **Administración de la tarea [admin-tarea]:** se pone esta etiqueta a los enunciados que contienen planes y negociaciones sobre cómo se hará la tarea (es decir, que la administran); estos enunciados explícitamente hablan sobre los procesos de solución de la tarea, sobre la coordinación y negociación de las acciones y sobre la planificación de objetivos subordinados a la tarea principal.

D18)

admin-tarea	utt79:s: te voy a explicar los elementos que hay en la cocina ++
admin-com	utt80:u: + sí +
admin-tarea	utt81:s: éste es una campana de para la estufa <ruido> ésta es una estufa éste es un frigobar o frigorífico ésta es una alacena pequeña <ruido> y ésta es una alacena superior
admin-com	utt82:u: ajá ++
admin-tarea	utt83:s: ++ tu tarea consiste en acomodar estos muebles como los tienes en el dibujo que tienes en una + hoja +
admin-com	utt84:u: + ajá +
admin-tarea	utt85:s: y hay unos muebles que faltan esos me los tienes que + pedir +
admin-com	utt86:u: + sí +
ininteligible	utt87:s: <no vocal>
tarea	utt88:u: entonces me falta el lavatrastos

- **Tarea [tarea]:** se pone esta marca a los enunciados que expresan acciones que completan la tarea, es decir, que la hacen. La mayoría de enunciados en un diálogo práctico involucran la realización de la tarea.

*“Utterances whose content is at the Task level directly move ahead (or attempt to move ahead) the goals of the domain”* (Allen y Core, 1997:7).

D22)

tarea	utt120:s: <no-vocal> éstos son los refrigeradores con que contamos
tarea	utt121:u: bien <sil> eh <sil> necesito colocar un frigobar
tarea	utt122:s: <ruido> en dónde ?
tarea	utt123:u: eh <sil> aquí
tarea	utt124:s: okey

*Actos del Habla como ‘Funciones hacia adelante’ y ‘Funciones hacia atrás’*  
Hasta ahora, hemos presentado los niveles de *Estado comunicativo* y *Nivel de información*. Si estuviéramos analizando un enunciado en concreto, habríamos señalado si tuvo éxito comunicativo o no, y si sí lo tuvo, tendríamos también una representación abstracta sobre el tema del que habla. Los siguientes niveles de análisis del esquema son las *Funciones hacia adelante* y las *Funciones hacia atrás*. Las unidades mínimas de estos niveles corresponden a los *actos del habla*.

*“While there are many possible ways to define utterances, here we based the notion of utterance on an analysis of the intentions of the speaker (the speech act)”* (Allen y Core, 1997:2).

Entrar al mundo del “¿qué quiso decir cuando dijo?” es como andar en un terreno pantanoso y confuso. Un análisis basado en las intenciones del hablante es de suma dificultad. Sabiendo que es terreno difícil, tal vez por esta razón DAMSL nombró “funciones” y no “intenciones” a este par de niveles. En realidad, el esquema se basa en la certeza de que el anotador del diálogo tiene acceso al diálogo ya que terminó, por esta razón sabe qué consecuencias produjo determinado enunciado. El anotador debe hacer un balance entre lo que pareció intentar el hablante cuando dijo algo y lo que significó para su interlocutor. ¿Qué más contexto

que el conversacional para saber qué intención comunicativa tuvo y qué función cumplió un enunciado en determinado momento?

DAMSL distingue entre las funciones que se vinculan con enunciados que se dirán y las que se vinculan con enunciados que se dijeron. Con esta distinción se establecen relaciones hacia adelante y hacia atrás entre los enunciados, permitiendo capturar la alternancia de turnos de forma local y coordinada a partir de los actos del habla que expresan. Junto con las etiquetas correspondientes a los niveles de análisis previamente presentados es posible crear un mosaico amplio en que un enunciado puede estar aportando comunicativamente algo al diálogo. De esta manera, DAMSL representa el tejido conversacional, evaluando cada enunciado hacia adelante y hacia atrás en distintos niveles.

#### *Funciones hacia adelante [FUNC-ADELANTE]*

Las etiquetas pertenecientes a este nivel indican que el enunciado influye en lo que se dirá posteriormente durante el diálogo. La expresión de un enunciado puede provocar distintas consecuencias, incluso si no son intentadas por el hablante. Una afirmación puede cambiar el estado del conocimiento compartido entre los participantes, un compromiso obliga a quien lo realizó a ejecutar determinadas acciones, una pregunta motiva una respuesta en el interlocutor, etc. El conjunto de etiquetas propuestas para este nivel se presentan a continuación.

#### *Funciones hacia adelante [FUNC-ADELANTE]*

- *Convencional* [conv]
  - Abre [abre]
  - Cierra [cierra]
- *Declaración* [dec]
  - Afirmación [afirm]
  - Reafirmación [reafirm]
- *Influenciar acciones futuras del escucha* [influ-acc-fut-esc]
  - Opción abierta [opcion-abierta]
  - Directiva de acción [dir-accion]
- *Solicitud de información* [sol-inf]
- *Comprometer acciones futuras del hablante* [comp-acc- fut-hab]
  - Oferta [oferta]
  - Compromiso [compromiso]
- *Performativo* [performativo]
- *Exclamación* [exclama]
- *Convencional* [conv]

- **Abre diálogo [abre]:** con esta etiqueta se marcan los enunciados que inician los diálogos. Por supuesto, además de la etiqueta [abre] estos enunciados otras etiquetas. En el Corpus DIME las conversaciones se inician convencionalmente con una oferta por parte del sistema, sin embargo, el diálogo puede comenzar con una directiva de acción del usuario o algún otro acto.

D05)

abre oferta	utt1:s: quieres que desplace o traiga algún objeto a la cocina ?
rechaza	utt2:u: no

D17)

abre dir-accion	utt1:u: bueno em vamos a quitar esta mesa
compromiso	utt2:s: <no-vocal> okey <ruido>

- **Cierra diálogo [cierra]:** se asigna esta marca a los enunciados que establecen la conclusión de la tarea y, por lo tanto, del diálogo. Por lo general, son afirmaciones que indican que la tarea está terminada. Estos enunciados pueden ser emitidos por cualquiera de los interlocutores, los agradecimientos y despedidas posteriores se toman con otro valor comunicativo.

D11)

cierra afirm	utt282:u: es todo ya ya terminé
-----------------	---------------------------------

D01)

cierra afirm	utt115:s: la tarea está terminada
-----------------	-----------------------------------

*Declaración [dec]*

- **Afirmación [afirm]:** los enunciados que reciben esta etiqueta son siempre proposiciones, incluyendo hipótesis y sugerencias de que algo puede ser cierto o falso. Todas las afirmaciones deben establecer un cierto estado de cosas verificable. “We only code utterances that make an explicit claim about the world” (Allen y Core, 1997:10).

Las afirmaciones intentan cambiar o confirmar las creencias del otro. Por ello, motivan que el escucha acepte o rechace explícita o implícitamente el contenido expresado en ellas.

D05)

afirm	utt94:s: esa operación no la puedo realizar
acepta	utt95:u: bueno

D13)

afirm	utt137:u: eh <sil> me parece que [a]sí ya queda bien la cocina
rechaza	utt138:s: no
afirm	utt139:s: tu cocina no tiene un módulo de limpieza <sil> un lavatrastes o una máquina lavatrastes
acepta	utt140:u: ah sí cierto

Cuando se duda si un enunciado es una afirmación o no, DAMSL recomienda responder a lo dicho en él con un “*eso no es cierto*”. Si resulta coherente, se trata de una proposición, si no, es un enunciado que presenta una forma declarativa pero que expresa algo más. Por ejemplo, el enunciado “*Vamos a poner la estufa aquí*” parece una afirmación y tiene una forma declarativa pero expresa una orden.

- **Reafirmación [reafirm]:** se pone esta etiqueta cuando la afirmación ha sido hecha previamente en el diálogo. Por lo general, se reafirma cuando se cree que no se ha entendido anteriormente.

D19)

afirm	utt92:s: eh entre la estufa + <sil> + y la mesa no hay espacio suficiente para que se siente alguien
ack	utt93:u: + sí +
sol-info	utt94:u: entre la estufa y la mesa ?
reafirm	utt95:s: entre <ruido> aquí + <sil> + <ruido> y aquí <sil> <ruido> <sil> no hay espacio suficiente para que s[e] siente

La Figura 5.1 muestra el árbol de decisión que DAMSL propone para anotar un enunciado como afirmación o reafirmación.



Figura 5.1 *Árbol de decisión para 'Declaración'*

*Influenciar acciones futuras del escucha* [influ-acc-fut-esc]

En esta clasificación se anotan los enunciados en los que el hablante motiva acciones no comunicativas del escucha. Así sucede en las órdenes, las peticiones, las sugerencias, etc. Este nivel excluye las preguntas, pues en ellas se solicita información y DAMSL considera que dar información requiere una categoría particular. Este nivel se divide en directivas de acción y opciones abiertas. La Figura 5.2 presenta el árbol de decisión para esta clasificación.

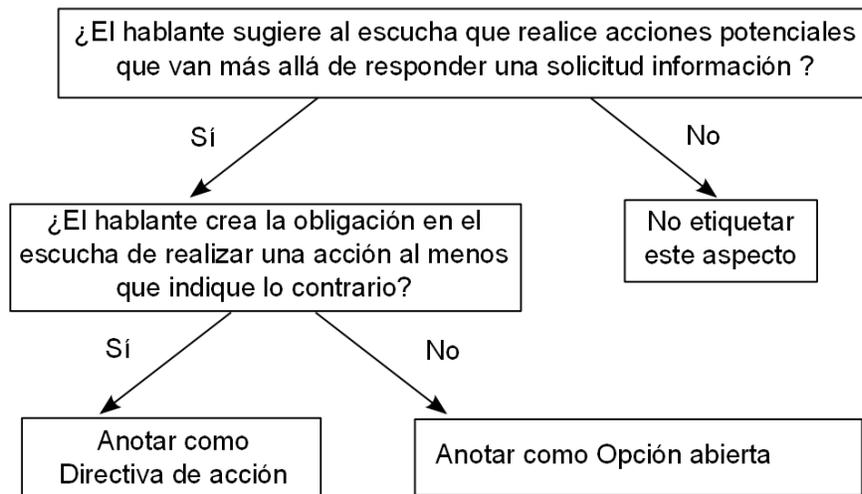


Figura 5.2 *Árbol de decisión para 'Influenciar acciones futuras del escucha'*

- **Directiva de acción [dir-accion]:** las directivas de acción son órdenes que expresa el hablante. La fuerza del mandato puede estar suavizada por convenciones de interacción social, como cuando se pide algo con una interrogativa del tipo “¿Puedes poner la mesa en la esquina?” o una declarativa como “Necesito poner el lavatrastos debajo de la ventana”, por mencionar algunas formas comunes de solicitar sin ordenar. Desatender una petición es considerado poco colaborativo en el contexto de un diálogo práctico. Por ello, cuando una directiva es expresada, se inicia un protocolo de interacción específico para atender la orden. De manera local, se observa que las órdenes se aceptan y se ejecutan, o si no, se rechazan.

D05)

dir-accion	utt91:u: podrías acomodarlos de tal manera <sil> que la parte superior <sil> estuviera a la misma altura <sil> de estos cuatro objetos ?
espera	utt92:s: <no-vocal> <sil> un segundo
rechaza	utt93:s: no

- **Opción abierta [opcion-abierta]:** en esta categoría se anotan los enunciados que ofrecen una serie de opciones para que el escucha elija. Estos enunciados no lo obligan a atender la sugerencia; sin embargo, por lo general lo motivan a aceptar o a rechazar las opciones que se le ofrecen.

D20)

opcion-abierta	utt45:s: contamos con <sil> estas estufas con hor[no]
rechaza	utt46:u: mm <ruido> no
sol-inf	utt47:u: mejor podemos ver las alacenas mejor?

*Solicitud de información [sol-info]:*

Las solicitudes de información son preguntas u otros tipos de requerimientos informativos que crean la obligación de responder o proveer algún tipo de información a través de cualquier medio.

*“Note, answers can be nonverbal actions providing information such as displaying a graph” (Allen y Core, 1997:11).*

Entre ellas se encuentran característicamente las preguntas cerradas y las preguntas pronominales.

D17)

sol-info	utt127:s: dónde quieres que lo ponga ? ++
respuesta	utt128:u: + + aquí

D19)

sol-info	utt101:u: otra sugerencia ?
respuesta	utt102:s: eh <sil> no

De la misma manera, se incluyen en esta categoría las solicitudes de información que parecen directivas de acción pero que solicitan que se despliegue cierto tipo de información en la pantalla, aunque no sea expresada lingüísticamente.

D20)

sol-info	utt2:u: por favor mm <sil> muéstrame las estufas
acepta	utt3:s: de acuerdo <ruido>
respuesta	utt4:s: éstas son <sil> las estufas con las que contam[os] <sil> <ruido>

02)

sol-info	utt81:u: eh <sil> bien eh quisiera volver a ver el catálogo de refrigeradores
acepta	utt82:s: okey
respuesta	utt83:s: <no-vocal> éstos son los refrigeradores

*Comprometer acciones futuras del hablante* [comp-acc-fut-hab]

Los enunciados anotados en esta categoría son aquellos que potencialmente comprometen al propio hablante a realizar algún tipo de acción futura. La distinción que se hace corresponde a si el compromiso lo toma el hablante por su propia cuenta o si lo adquiere cuando el escucha acepta su propuesta.

- **Oferta [oferta]:** los enunciados que reciben esta marca son aquellos en que el hablante propone realizar una acción. Las ofertas necesitan que el escucha esté de acuerdo con lo ofertado para que el hablante esté comprometido a realizarlo. Las ofertas pueden ser aceptadas, rechazadas o no atendidas sin mayores consecuencias.

D01)

oferta	utt192:s: quieres poner otro mueble en la cocina ?
rechaza	utt193:u: no creo que así está bien

- **Compromiso [comp]:** los enunciados en que el hablante se obliga a realizar una acción futura reciben la marca de compromisos. Éstos pueden variar en la fuerza, sin embargo, si el hablante deja claro que tiene la voluntad de realizar algo y

lo realiza, se anota en esta categoría. Los compromisos aparecen cuando se acepta una directiva de acción y van seguidos de la acción prometida.

D10)

dir-accion	utt92:u: ahora <sil> puedes colocar este fregadero dentro de del cuarto <sil> en el <sil> en esta <sil> en esta pared pegada a la pared mm <sil> del fondo del cuarto
compromiso	utt93:s: okey
agrega-obj	utt94:s: <acción gráfica>

D01)

dir-accion	utt7:u: <ruido> ahora <sil> vamos a poner <sil> la estufa <sil> aquí
compromiso	utt8:s: okey
mueve-obj	utt9:s: <acción gráfica>

A continuación se presenta la Figura 5.3, en ella se muestra el árbol de decisión propuesto por DAMSL para esta categoría.

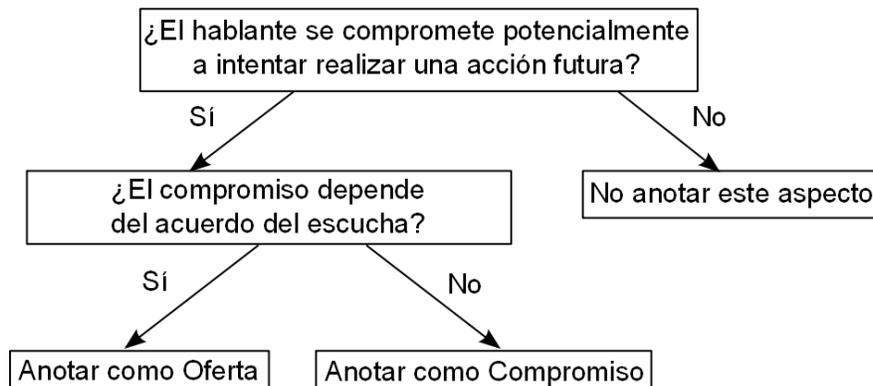


Figura 5.3 Árbol de decisión para 'Comprometer acciones futuras del hablante'

#### Otras Funciones hacia adelante

- **Performativo [performativo]:** En el análisis del Corpus DIME los agradecimientos fueron marcados como “performativos”.

*“Explicit-performative: Is the speaker performing an action by virtue of making the utterance (e.g., “Thank you”, “I apologize”)” (Allen y Core, 1997:14).*

Es muy probable que a este nivel le falte un desarrollo futuro. Digamos que tal vez debería llamarse “Agradecimientos”.

D15)

performativo	utt89:s: gracias
performativo	utt90:u: de nada

- **Exclamación [exclama]:** Se anotan con esta marca los enunciados que contienen exclamaciones.

D17)

exclama	utt180:u: mm <sil> ay ay ay qué es lo que estoy haciendo ?
---------	--

#### *Funciones hacia atrás*

Así como las *Funciones hacia adelante*, las *Funciones hacia atrás* tejen la continuidad del diálogo a través de los actos del habla.

*“Backward Looking Functions indicate how the current utterance relates to the previous discourse” (Allen y Core, 1997:17).*

La relación entre un enunciado y su antecedente por lo general es local e inmediata; sin embargo, podemos encontrar protocolos de interacción en que la correspondencia no es tan directa. Por ello, para asignar cualquiera de las siguientes etiquetas es necesario identificar a qué enunciado se vincula. Esto se hace escribiendo entre paréntesis el número que se asignó al enunciado durante la segmentación del corpus. Es decir, todas las etiquetas de las *Funciones hacia atrás* incluyen en su etiquetación el referente numérico del enunciado al que corresponde. La propuesta para las *Funciones hacia atrás* es la siguiente:

#### *Funciones hacia atrás [func-atras]*

- *Acuerdo* [acuerdo]
  - Acepta [acepta(#)]
  - Acepta parte [acepta-parte(#)]
  - Quizá [quiza(#)]
  - Rechaza parte [rechaza-parte(#)]
  - Rechaza [rechaza(#)]
  - Espera [espera(#)]
- *Entendimiento* [ent]
  - Señal de no entendimiento [SNE(#)]
  - Señal de entendimiento [SE]
    - Asentimiento [ack(#)]
    - Repetir/Refraseo [repeticion(#)]
    - Complemento [complemento(#)]
  - Corrección [correccion(#)]
- *Respuesta* [respuesta(#)]
- *Relación de información* [relacion-info]

*Acuerdo* [acuerdo]

Los acuerdos se realizan en múltiples niveles durante una conversación. Se evalúan en esta categoría los enunciados que establecen qué tanto coincide el hablante con algo dicho por su interlocutor anteriormente. Estas marcas junto con las de los otros niveles de DAMSL, permiten determinar si se negocia el acuerdo en el nivel de la “tarea” o en el de la “comunicación”, por ejemplo, o si el hablante está aceptando una orden previa o rechazando el contenido de una afirmación.

En el contexto de los diálogos prácticos, no sólo se verifica lo dicho, sino también lo hecho. Las acciones no verbales que se realizan se aprueban o se rechazan. En el caso del Corpus DIME, las acciones del sistema (quitar, poner, mover, etc.) siempre son verificadas por el usuario.

El acuerdo implícito está sobreentendido en los diálogos colaborativos con solo continuar la conversación; sin embargo, un enunciado que expresa el nivel de acuerdo explícitamente puede aparecer en cualquier momento durante un diálogo. Cuando aparece de forma explícita tiene el fin de evidenciar que se comparten o no las creencias sobre cierto estado de cosas. En esta dimensión de análisis se anotan, sobre todo, los posicionamientos explícitos. No obstante, en algunos casos ha sido necesario marcar negociaciones que se dan en un margen no tan explícito.

DAMSL propone en sus etiquetas de acuerdo una graduación que va desde aceptar algo por completo hasta rechazarlo categóricamente, pasando por aceptar o rechazar partes de lo enunciado o hecho. También cubre las posiciones que no son tan determinantes; como cuando se piensa que quizá una opción sea adecuada pero no se siguen las acciones que la completan, o como cuando se le pide o sugiere al interlocutor que espere un momento en lo que se reúne la información necesaria para tomar una decisión.

- **Acepta [acepta(#)]**: si el enunciado expresa que el hablante está absolutamente de acuerdo con el contenido del enunciado previo se pone esta marca. Las directivas de acción, las preguntas imperativas, las afirmaciones, las ofertas y muchos actos de las *Funciones hacia adelante* se aceptan explícita o implícitamente.

D05)

afirm	utt94:s: esa operación no la puedo realizar
<b>acepta (94)</b>	<b>utt95:u: bueno</b>

- **Acepta parte [acepta-parte(#)]**: esta marca se pone a los enunciados que sugieren o establecen estar de acuerdo con una parte de lo dicho o hecho.

D03

sol-info	utt44:s: ahí está bien ?
<b>acepta-parte (44,43)</b> <b>respuesta (44)</b>	<b>utt45:u: un poco más separadita</b>

- **Rechaza parte [rechaza-parte(#)]**: se pone esta marca a los enunciados que explícitamente rechazan una parte de lo propuesto por el interlocutor.

D19)

oferta	utt39:s: <ruido> <sil> quieres que ponga <sil> este frigobar <sil> a un lado de la estufa ?
<b>rechaza-parte (39)</b>	<b>utt40:u: mm no &lt;sil&gt; eh &lt;sil&gt; en la pared que estoy indicando</b>

Los creadores notaron que aceptar una parte de un enunciado es implícitamente rechazar la otra parte. Sin embargo, sólo se anota cuando es explícito que se está rechazando una parte de lo enunciado o lo dicho. DAMSL apunta que si explícita y simultáneamente se acepta una parte y se rechaza la otra parte, entonces se anoten las dos etiquetas como en los siguientes ejemplos.

D03)

oferta	utt1:s: quieres que desplace o traiga algún un objeto a la cocina ?
<b>rechaza-parte (1)</b> <b>acepta-parte (1)</b>	<b>utt2:u: mm que me traiga no que me m[uevas] que muevas el no sé cómo se le llama a las gabinete no sé a esto que está en la pared izquierda + + sí ah es lo único que está en la pared izquierda sobre el piso</b>

D25)

sol-info	utt8:s: ahí está bien ?
<b>acepta-parte (8)</b> <b>rechaza-parte (8)</b> <b>respuesta (8)</b>	<b>utt9:u: sí &lt;sil&gt; no</b>
dir-accion	utt10:u: más

- **Rechaza [rechaza(#)]**: se anotan con esta etiqueta los enunciados que rechazan por completo y explícitamente un enunciado previo.

D01)

sol-info	utt95:u: y especifica aquí de alguna manera la <sil> el tamaño del <sil> del fregadero ?
<b>rechaza (95)</b> <b>respuesta (95)</b>	<b>utt96:s: mm no</b>

- **Quizá [quiza(#)]:** se pone esta etiqueta a los enunciados que no comprometen al hablante a establecer el nivel de acuerdo. El “quizá” se sobreentiende como implícito; por lo general, estas negociaciones son abandonadas sin mayor impacto en el flujo de la conversación o de la tarea.

D01

sol-info	utt23:s: ahí está bien ?
<b>quiza (23)</b> <b>respuesta (23)</b>	<b>utt24:u: bueno par-parece que está</b>

- **Espera [espera(#)]:** en estos enunciados el hablante deja pendiente el establecimiento del acuerdo. Abre una especie de “paréntesis” en el que confirma, solicita información adicional o pide un poco de tiempo antes de decidir.

D22)

sol-info	utt159:u: eh hay algún dise[ño] algún problema de diseño ahora ?
<b>espera (159)</b>	<b>utt160:s: un segundo por favor</b>
acepta(160)	utt161:u: sí
rechaza(159) respuesta(59)	utt162:s: no no hay ningún problema de diseño

D20)

dir-accion	utt36:u: eh <sil> favor de elegir el fregadero con alacena
<b>espera (36)</b> <b>sol-info</b>	<b>utt37:s: éste ?</b>
acepta(37) respuesta(37)	utt38:u: ajá
dir-accion	utt39:u: y lo colocamos abajo de la ven[tana] de las ventanas
acepta(36,39)	utt40:s: de acuerdo

A continuación se presenta el árbol de decisión propuesto para este nivel de anotación por DAMSL.

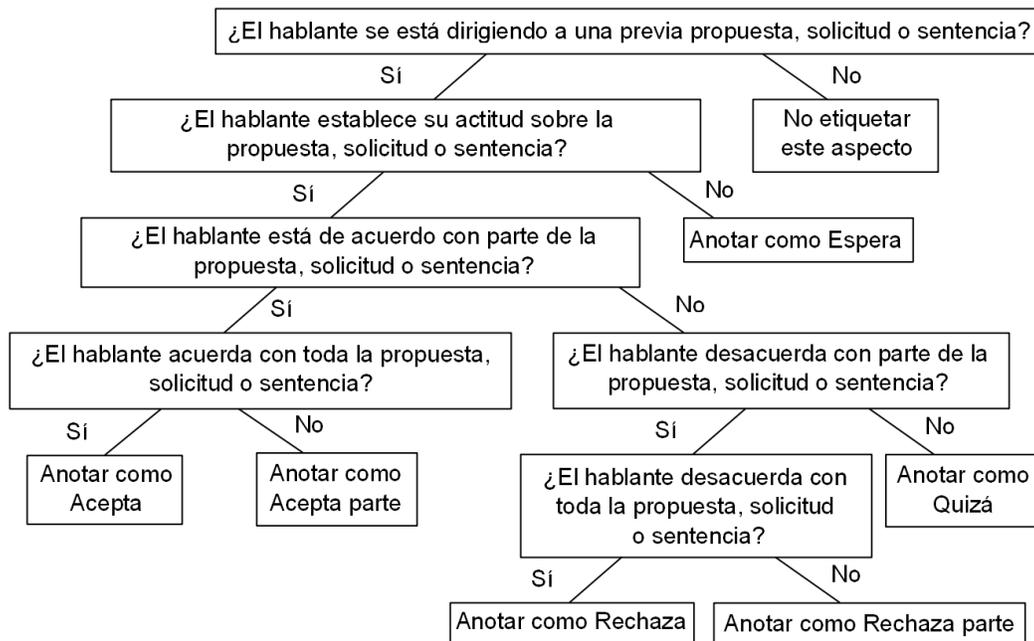


Figura 5.4 Árbol de decisión para 'Acuerdo'

#### Entendimiento [ent]

En esta dimensión de análisis se etiquetan los actos lingüísticos en que explícitamente se monitorea el canal de comunicación y la comprensión del enunciado previo. En turnos cortos y conversaciones fluidas, se mantiene implícito el buen entendimiento al seguir hablando. En turnos largos, en cambio, los hablantes van actualizando la comprensión del contenido informativo de manera constante y evidente.

*“There are many levels of “understanding”, ranging from merely hearing the words to fully identifying the speaker’s intention” (Allen y Core, 1997: 20).*

- **Señal de no entendimiento [SNE(#)]:** los enunciados que reciben esta marca son aquellos que manifiestan no haber comprendido lo dicho anteriormente por el interlocutor. Esto puede deberse a ruidos o problemas en el canal de comunicación, pero también a una mala articulación o a la falta de referentes compartidos en el conocimiento mutuo. Por lo general, estos enunciados son preguntas del tipo “¿Perdón?” o “¿Cómo?” o expresiones como “No entendi”.

D23)

dir-accion	utt18:u: quiero <ruido> ver eh éste <ruido> este fregadero <ruido> +aba[jo]+
<b>SNE (18)</b>	<b>utt19:s: + perdón ? +</b>

D18)

monologo	utt39:u: <no-vocal> pensé que estaba la alacena
<b>SNE (39)</b>	<b>utt40:s: perdón no t[e] no te entiendo puedes repetir ?</b>

- **Asentimiento [ack(#)]<sup>15</sup>**: los enunciados que tienen esta etiqueta son palabras o sonidos que son expresados para manifestar que el hablante “toma nota” o ha recibido lo que fue dicho sin que necesariamente se esté aceptando su contenido.

*“Acknowledgments are utterances consisting of short phrases such as ‘okay’, ‘yes’, and ‘uh-huh’, that signal that the previous utterance was understood without necessarily signaling acceptance” (Allen y Core, 1997:21).*

DAMSL dice que si un enunciado es al mismo tiempo un [acepta(#)] del nivel de *Acuerdo* y un [ack(#)] del nivel de *Entendimiento* se pongan las dos etiquetas. El Grupo DIME consideró pertinente restringir los *acknowledgments* a los enunciados que únicamente mantienen actualizado el flujo de la comunicación o que avisan al interlocutor que el mensaje se ha recibido. Si se trata de un enunciado que acepta explícitamente algo, se acordó etiquetarlo entonces en la dimensión de *Acuerdo* y no en la de *Entendimiento*.

D20)

opcion-abierta	utt116:s: éstas son las mesas con las que conta[mos]
<b>ack (116)</b>	<b>utt117:u: mm &lt;ruido&gt;</b>
sol-info	utt118:u: puedo ver las de abajo ?

- **Traslape [traslape(#)]**: Si bien en el contexto de diálogos colaborativos es poco frecuente que los interlocutores tomen la palabra al mismo tiempo “sin dejarse

<sup>15</sup> Esta etiqueta proviene del inglés “acknowledgment”. La traducción de “asentimiento” intenta reflejar de manera un poco más cercana el concepto detrás de este vocablo inglés. No debe confundirse con su sentido de “acordar” o “aceptar”, sino aprovecharse el sesgo de “asentar” algo en el conocimiento o la interpretación. La etiqueta [ack] conserva su origen anglosajón porque era más sencillo reconocerla visualmente.

hablar”, es común encontrar traslapes<sup>16</sup>. Pueden cumplir diferentes funciones, por lo general se trata de momentos en los que el hablante manifiesta prestar atención, sin embargo también encontramos reinicios en los que ambos toman la palabra. El corpus señala con los signos + + + cuando esto ocurre en los enunciados.

D03)

sol-info	utt111:u: cómo se le llama a la parte de arriba de la estufa la que absorbe ?
monologo	utt112:u: no me acuerdo
<b>respuesta (111)</b>	<b>utt113:s: + el extractor ? +</b>
<b>afirm traslape (113)</b>	<b>utt114:u: + campana +</b>
<b>repeticion (114) acepta (114)</b>	<b>utt115:s: + una campana +</b>
<b>acepta (113) traslape (115)</b>	<b>utt116:u: + + eso o extractor o ah sí</b>

- **Repetición o refraseo [repeticion(#)]:** Esta marca se pone a los enunciados que actualizan el entendimiento con frases cortas que repiten fragmentos del enunciado anterior o parafrasean su contenido.

D18)

afirm	utt103:s: tiene que coincidir con el que está ahí
<b>repeticion(103)</b>	<b>utt104:u: tiene que coincidir</b>

- **Complemento [complemento(#)]:** Los enunciados que reciben esta etiqueta son aquellos en que el hablante está completando el enunciado de su interlocutor con la intención de hacerle ver que ha entendido tanto que hasta puede terminar su frase. En el Corpus DIME no apareció ninguno de estos complementos. Un ejemplo imaginario sería:

sol-info	u: tienes un <sil> listado de mesas o una especie de <sil>
<b>complemento</b>	<b>s: catálogo ?</b>

<sup>16</sup> En la literatura sobre el estudio de intercambio de turnos es común encontrar que: “*overwhelmingly, one party speaks at a time*” (Sacks, Schlegoff y Jefferson: 1974). Evidentemente, los protocolos de intercambio de turno conversacional se adquieren a tan temprana edad que algunos autores olvidan que también eso se aprende. Se estima que sólo un 5% de la interacción se traslapa y que se puede encontrar cierta sistematicidad en el fenómeno: la mayor parte ocurre para aclarar que se escucha (“*ajá*”, “*sí*”, “*okey*”) y en los casos en que ambos interlocutores comiencen a hablar al mismo tiempo, alguno cede la palabra.

- **Corrección [correccion(#)]**: Se pone esta marca a los enunciados en que el hablante corrige un error en el enunciado previo de su interlocutor. Sólo se etiqueta cuando se está corrigiendo al otro, no cuando el hablante se corrige a sí mismo.

D17)

dir-accion	utt192:u: ahora vamos a buscar este <sil> un extractor para el refrigerador aquí + para éste +
<b>correccion(192)</b>	<b>utt193:s: + + no hay extractores para los refrigeradores + +</b>
dir-accion	utt194:u: + + digo un extractor para el la estufa

*Respuesta [respuesta(#)]*

Todos los enunciados que responden directamente a una solicitud de información son etiquetadas en esta dimensión, independientemente del medio en que la información sea vertida. En los paréntesis se anota el número de enunciado que expresa su pregunta, creando un par adyacente “pregunta-respuesta”. En el caso del Corpus DIME, ciertas preguntas se respondían mostrando información en la pantalla. Estas acciones recibieron la etiqueta de respuestas puesto que el sistema está dando la información solicitada.

D22)

sol-info	utt95:s: dónde quieres que lo coloque ?
<b>respuesta (95)</b> dir-accion	<b>utt96:u: aquí</b>

D18)

sol-info	utt5:u: me muestras el catálogo de las <sil> del mobiliario ?
acepta(5)	utt6:s: sí
<b>respuesta (5)</b> <b>opcion-abierta</b>	<b>utt7:s: &lt;despliega catálogo&gt;</b>

D11)

sol-info	utt274:s: quieres que gire el cuarto noventa grados ?
<b>respuesta (274)</b> <b>acepta (274)</b>	<b>utt275:u: sí</b>

*Relación de información [relacion-info]*

Los creadores de DAMSL presentan esta dimensión de análisis pero no la desarrollan en el texto. Apenas introducen que se vincula con el contenido informativo de los enunciados y que esperan proponer a futuro un nivel independiente de dominio en que se pueda considerar la continuidad informativa en

los enunciados durante una conversación. El Proyecto DIME no consideró esta dimensión de análisis para el Corpus DIME.

***Alcances y limitaciones de DAMSL***

La clara especificación de cada uno de los niveles de análisis de DAMSL permite la anotación de una gama amplia de posibilidades en que un enunciado puede comunicar algo. Por ejemplo, un par de preguntas como “¿Me escuchas?” y ¿Qué objeto es éste?” recibirán en las *Funciones hacia adelante* la marca de las solicitudes de información; sin embargo, en el *Nivel de Información*, la primera estará marcada como una pregunta que administra la comunicación y la segunda como una pregunta concerniente a la tarea.

*“This type of annotation is not typically seen in speech act theory but it is critical to interpreting dialogs since the utterances at these levels must be processed using different strategies.” (Core y Allen, 1997:7)*

DAMSL basa su anotación en la idea de que un enunciado puede aportar contenido en distintos niveles comunicativos simultáneamente durante una conversación. Asimismo, un enunciado puede expresar varios actos del habla y ser al mismo tiempo una respuesta, una afirmación y un compromiso, si se diera el caso. DAMSL privilegia la captura de estos niveles en los que un enunciado puede significar algo en la conversación, sin importar qué tan costoso resulte para un grupo de investigación aprender a utilizar el esquema y anotar rigurosamente cada enunciado de un corpus. A continuación presentamos un fragmento del Corpus DIME anotado con DAMSL.

D20)

NIV-INFO	FUNCIONES	ENUNCIADOS
tarea	dir-accion	utt82:u: mm <ruido> esta alacena <ruido> <sil> la colocamos en esta posición
tarea	acepta(82) compromiso	utt83:s: <ruido> [de] [a]cuer[do]
tarea	agrega-obj	utt84:s: <acción gráfica>
admin-tarea	sol-info	utt85:s: así está bien ?
admin-tarea	acepta(85) respuesta(85)	utt86:u: sí <ruido>

Al observar el ejemplo, DAMSL se muestra como un modelo minucioso que permite capturar la aportación de cada enunciado en distintos niveles conversacionales. Este análisis permite disociar la forma de expresión de un enunciado de su intención comunicativa; permite, también, distinguir los distintos niveles en que puede estar operando un acto del habla. A la par, este esquema permite establecer los vínculos que los enunciados tienden entre sí de forma local.

*“A corpus reliably annotated with DAMSL labels would provide a valuable resource in the study of discourse as well as a source of training and testing for a dialog system using DAMSL labels in its utterance representation” (Core y Allen, 1997:7).*

Se dijo anteriormente que los sistemas a los que está dirigido DAMSL están centrados en la interacción humano-computadora para la solución de tareas a través del lenguaje natural. El modelo está diseñado de tal forma que los vínculos hacia adelante y hacia atrás entre los enunciados representan estructuras conversacionales como pares adyacentes. Junto con una serie de reglas, esta metodología permite consolidar un recurso viable para el entrenamiento y las pruebas de sistemas conversacionales.

*“For example, a system might have a rule such as ‘a statement is something to add to the database’ [...] As the system adds utterances to its data structures, it will create a higher level forms such as hierarchical multi-agent plans and discourse structures analogous to paragraphs and chapters” (Core y Allen, 1997:7).*

Todos estos elementos hacen de este esquema una herramienta sólida, como punto de partida es pertinente considerar que los diálogos prácticos son secuencias de enunciados que se tejen hacia adelante y hacia atrás. Este análisis local abre un horizonte de posibilidades, si bien condiciona ver los diálogos como secuencias de enunciados que se relacionan en pares, no impide organizar esos pares en unidades mayores. No hace falta mucho tiempo estudiando diálogos prácticos para sentir que la cadena de enunciados se va organizando en unidades más complejas que una relación binaria “pregunta-respuesta”. El flujo del diálogo y, por ende, el desarrollo

de la tarea dependen de una organización más global. Esta organización puede verse como la planificación y cumplimiento de una serie de objetivos que completan la tarea. Un modelo que contemple la negociación de objetivos representará estructuras conversacionales mayores en el contexto de los diálogos prácticos.

El Proyecto DIME consideró conveniente expandir el potencial de análisis de DAMSL integrando una unidad conversacional mayor llamada “transacción” (Pineda *et al.*, 2007, 2006). Con esta unidad es posible entrenar sistemas conversacionales que vayan registrando los cambios que van ocurriendo en el mundo durante la conversación, sean cambios en el estado de las cosas o cambios en las creencias sobre el estado de las cosas. Este registro se va actualizando a partir de la satisfacción de objetivos subordinados a la tarea principal. DAMSL focaliza como unidades mínimas los actos del habla. DIME-DAMSL se basa en ellas para conformar unidades mayores. Mientras DAMSL dice que un diálogo práctico es una secuencia de enunciados coordinados, DIME-DAMSL dice que un diálogo práctico es una secuencia de transacciones balanceadas. En el siguiente capítulo se presentan puntualmente los ajustes y cambios que se le hicieron a DAMSL para crear DIME-DAMSL.

## 6. DIME-DAMSL: Modificación de niveles y adición de etiquetas

### *Distinción entre tipos de solicitudes de información*

DAMSL propone que todo tipo de preguntas que solicitan genuinamente información sean anotadas en la categoría de *Solicitud de información* sin importar la forma en que sea expresada. En el contexto de una conversación, preguntas como “¿Dónde quieres que lo ponga?”, “¿Me escuchas?” o “¿Puedes mostrarme el catálogo de muebles?” cumplen funciones conversacionales distintas. El Grupo DIME consideró pertinente distinguir tres tipos de solicitudes de información para corroborar si era posible rastrear funciones pragmáticas distintas vinculadas al tipo de información que se solicita y a la prosodia que la expresa (Coria, 2008).

- **Pregunta pronominal [preg-pro]:** se pone esta marca a las preguntas que buscan establecer referentes y que por ello incluyen adverbios y pronombres interrogativos como “dónde” y “cómo” o “qué” y “cuál”.

D17)

<b>preg-pro</b>	utt127:s: dónde quieres que lo ponga ? ++
respuesta (127)	utt128:u: + + aquí

- **Pregunta sí/no [preg-sn]:** se pone esta etiqueta a las solicitudes de información que buscan verificar si el interlocutor está de acuerdo o no con algo expresado. Se les llama también preguntas cerradas porque característicamente se responden con un “sí” o un “no”. Estas preguntas son enunciados que tienen la forma sintáctica de una afirmación pero no su tono; las afirmaciones son expresadas con un tono descendente, las preguntas cerradas con uno ascendente.

D06)

<b>preg-sn</b> <b>espera (35)</b> <b>apt-obj</b> <b>apt-zona</b>	utt37: quieres que ponga <sil> una estufa como éstas <sil> en esta esquina ?
respuesta (37) acepta (37)	utt38:u: sí

- **Pregunta imperativa [preg-imp]:** reciben esta etiqueta las preguntas que parecen directivas de acción (imperativas como “muéstrame”, interrogativas como “¿puedes enseñarme?” y declarativas como “quisiera ver”) pero que solicitan se despliegue información visual en la pantalla. En el contexto del Corpus DIME,

mostrar el catálogo de muebles es proveer información a través de un medio gráfico. Estas preguntas, por lo general, siguen el protocolo de las directivas de acción y es común encontrar que antes de dar la respuesta el interlocutor acepta proveer la información solicitada.

D20)

preg-imp	utt2:u: por favor mm <sil> muéstrame las estufas
acepta(2)	utt3:s: de acuerdo <ruido>
respuesta(2) opcion-abierta	utt4:s: éstas son <sil> las estufas con las que contam[os] <sil> <ruido>

### ***Inclusión de acciones gráficas***

DAMSL enmarca su análisis en las acciones lingüísticas porque fue creado para anotar corpus conversacionales provenientes de la asistencia telefónica. El Proyecto DIME integró la dimensión de *Acciones gráficas* al esquema de análisis porque el Corpus DIME no sólo se escucha, también se ve. Como se explicó en el Capítulo 4, este corpus está compuesto por archivos de video en los que se puede escuchar el diálogo y ver el desarrollo de la tarea. Esto quiere decir que el tipo de sistema conversacional que se busca pretende integrar la administración y vinculación de la información visual durante la ejecución de diálogos y tareas (Villaseñor *et al.*, 2001).

En la dimensión de *Acciones Gráficas* se anotan distintos aspectos relacionados con la información visual compartida que es relevante para el diálogo y el desarrollo de la tarea. La relevancia de la información visual puede ser valorada desde diferentes perspectivas. En este esquema se consideran dos: por un lado, se incluyen los gestos deícticos con que el hablante señala en la pantalla para establecer referentes; por el otro, se contemplan las acciones no lingüísticas que realizan la tarea, como mover o quitar objetos. Cualquiera de estas funciones conlleva la ejecución de una “acción gráfica” a la que ambos participantes tienen acceso visual.

Integrar este tipo de acciones a un esquema de anotación de actos del habla tiene dos vías de utilidad. Desde una perspectiva lingüística, los participantes en un diálogo práctico discuten, planifican y negocian acciones no lingüísticas a través del

lenguaje. Cuando estas acciones se realizan, cambia el estado de cosas en el mundo y el estado de las creencias de las cosas en el mundo; por ello, cambia también el curso que sigue la conversación. La anotación de estas acciones puede servir de base empírica a investigaciones que se interesan en los procesos en que la información visual y la lingüística se vinculan para establecer referentes y para actuar en el mundo. Desde una perspectiva computacional, la creación de sistemas capaces de interactuar con humanos administrando y organizando información proveniente de múltiples fuentes, además de la lingüística, ha sido de fundamental interés. Estos sistemas están motivados por la intención de interactuar con el mundo y no sólo con humanos. Por ello, integrar este tipo de acciones en un esquema de interacción permite abarcar dentro de una misma administración diversas fuentes de información y diferentes conductas vinculadas a la ejecución de estas acciones.

#### *Funciones hacia adelante*

##### *Acciones gráficas* [accion-grafica]

- *Gestos*
  - Apuntar objeto [apt-obj]
  - Apuntar zona [apt-zona]
  - Apuntar trayectoria [apt-tray]
  - Apuntar objetos coordinados [apt-objs-coor]
  
- *Acciones*
  - Agrega objeto [agrega]
  - Mueve objeto [mueve]
  - Quita objeto [quita]
  - Plan gráfico [plan-graf]
  - Rotar vista [rota-vista]
  - Despliega [despliega]

#### *Gestos*

Los gestos en el Corpus DIME se utilizan como un recurso común para establecer referentes o desambiguarlos cuando hay confusión. Se recordará que para la recopilación del corpus cada participante tenía acceso a una computadora que mostraba lo mismo en cada pantalla. Los participantes podían utilizar el puntero del ratón como dedo índice para señalar objetos y lugares. Los gestos se hicieron dando clic o pasando el ratón por encima haciendo círculos. Es común encontrar que los

gestos están acompañados por un déictico lingüístico, sin embargo no siempre sucede así.

- **Apuntar objeto [apt-obj]:** Se pone esta etiqueta a los enunciados en que el hablante utiliza el ratón para señalar un objeto en la pantalla. En la Figura 6.1 se señala con un círculo rojo el lugar en que se apunta un objeto.

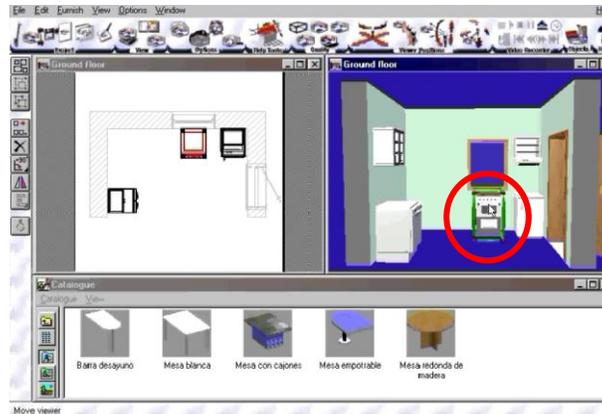


Figura 6.1 Apunta objeto

D18)

apt-obj preg-pro	utt189:u: me puedes indicar qué mueble es éste
apt-obj espera (189) preg-sn	utt190:s: <no-vocal> ésta ?
acepta (190) respuesta (190)	utt191:u: sí

- **Apuntar zona [apt-zona]:** reciben esta marca los gestos que refieren a lugares del espacio virtual que se muestra en la pantalla. La Figura 6.2 señala con un círculo rojo en la gráfica del lado izquierdo el puntero del ratón apuntando una zona en la pantalla.

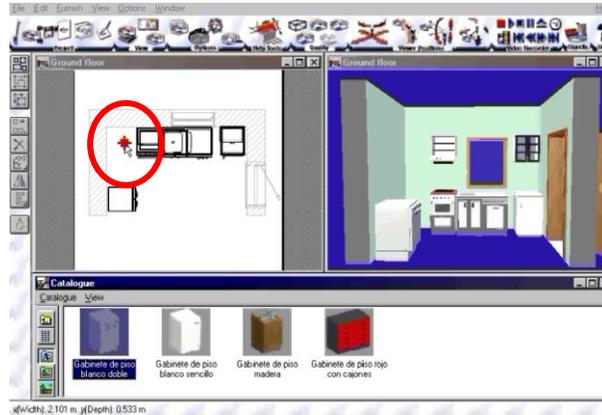


Figura 6.2 Apunta zona

D19)

dir-accion	utt9:u: eh por favor lo necesito <sil> en <sil> en la pared del fondo
espera(9) preg-pro	utt10:s: cuál es la pared del fondo ?
apt-zona respuesta(10) afirm	utt11:u: <ruido> a ver aquí

También puede encontrarse un gesto déctico sin referencia lingüística déctica.

D23)

apt-obj preg-pro	utt60:s: dónde quieres que la ponga ?
apt-zona respuesta(95) afirm	utt61:u: quiero que la pongas junto a la campana del extremo izquierdo

- **Apuntar trayectoria [apt-tray]:** se pone esta marca a los enunciados en los que el hablante manifiesta que desea mover un objeto indicando en la pantalla el punto de origen y el de término. En la Figura 6.3 se presentan dos imágenes, en la del lado izquierdo se apunta el punto de origen; en la del lado derecho se apunta el punto de término.

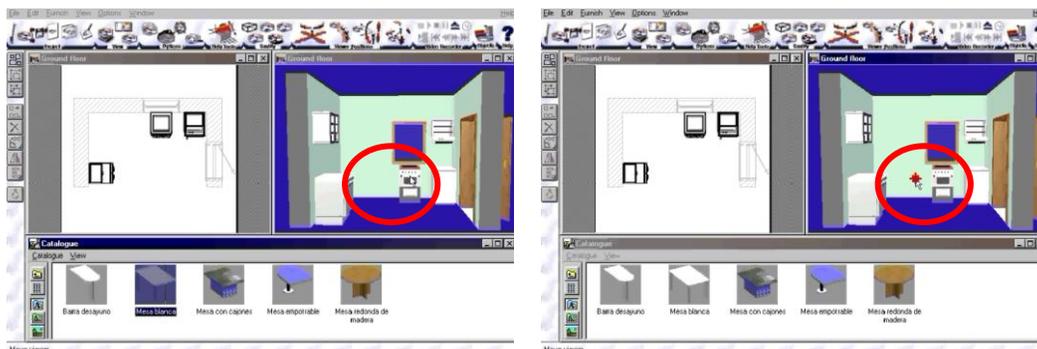


Figura 6.3 Apunta trayectoria

D14)

preg-pro	utt65:s: a dónde ?
apt-tray respuesta(261)	utt66:u: mm nada más recorrerla tantito <sil> has-hasta donde termina la ventana

- **Apuntar objetos coordinados [apt-objs-coor]:** reciben esta marca los enunciados en que el hablante señala más de un objeto.

D14)

apt-objs-coor preg-sn	utt102:s: quieres que <sil> sustituya <sil> este mueble por éste ?
acepta(102) respuesta(102)	utt103:u: mm mh
apt-objs-coor preg-sn	utt104:s: y <sil> éste por éste ?
acepta(104) respuesta(104)	utt105:u: sí

D05)

apt-objs-coor dir-accion	utt88:u: eh podrías alinear esos objetos y dejarlos a la misma altura
acepta(88) compromiso)	utt89:s: okey

Por supuesto, en un mismo enunciado el hablante puede hacer varios gestos y de cualquier tipo.

D18)

apt-obj apt-zona preg-sn	utt209:s: quieres que ponga este mueble <sil> aquí ?
acepta(209) respuesta(209)	utt210:u: sí

D17)

apt-obj apt-tray dir-accion	utt140:u: a ver haz el fregadero hasta la esquina acá entonces
acepta(140) compromiso	utt141:s: <ruido> okey

### *Acciones*

Cada tarea requiere la realización de acciones específicas. En el Corpus DIME, las acciones están vinculadas a mover, quitar y poner objetos en la cocina que se diseña, también a desplegar catálogos y a cambiar la configuración de la vista. Estas acciones son relevantes conversacionalmente porque, si bien se realizan sin enunciar una sola palabra, resuelven la tarea y son planificadas y refinadas a través del lenguaje. DIME-DAMSL representa su aportación al diálogo como una equivalencia. Las acciones cambian el estado de cosas en el mundo de forma

similar que las afirmaciones cambian las creencias sobre el estado de cosas en el mundo. Decir algo y que alguien lo crea verdadero se parece a hacer algo y que alguien lo considere correcto. Por esta razón, cuando se realiza una acción, se verifica que haya sido satisfactoria, en el contexto de diálogos colaborativos es común encontrar que este protocolo de acuerdo se hace explícito, sin embargo es posible que las acciones se acepten continuando con la tarea. En el caso del Corpus DIME, las acciones siempre son realizadas por el “sistema” y siempre es el “usuario” el que da el visto bueno a su ejecución.

- **Agregar objeto [agrega]:** esta etiqueta se pone cuando el sistema coloca desde el catálogo de muebles un objeto que no estaba previamente en la cocina.

D10)

dir-accion apt-obj apt-zona	utt3:u: este mira necesitamos otro <sil> estante creo como éste si lo puedes <sil> poner aquí en esta pared
acepta(3) compromiso	utt4:s: okey
<b>agrega</b>	<b>utt5:s: &lt;acción gráfica&gt;</b>
preg-sn	utt6:s: así está bien ?
acepta(5,6) respuesta(6)	utt7:u: <no-vocal> sí <ruido> así está bien <sil> <vocal>

- **Mover objeto [mueve-obj]:** se anota esta etiqueta cuando el sistema desplaza un objeto previamente colocado en el espacio virtual de la cocina.

D18)

dir-accion apt-obj	utt180:u: em <sil> este mueble que estoy señalando tiene que estar pegado hacia la pared izquierda
acepta(180) compromiso	utt181:s: okey
<b>mueve</b>	<b>utt182:s: &lt;acción gráfica&gt;</b>
preg-sn	utt183:s: ahí está bien ?
acepta(182,183) respuesta(183)	utt184:u: sí

- **Quitar objeto [quita-obj]:** se anota cuando el sistema elimina un objeto que estaba previamente en la cocina.

D06)

dir-accion	utt44:u: podrías quitarla ?
acepta(44) compromiso	utt45:s: okey
<b>quita</b>	<b>utt46:s: &lt;acción gráfica&gt;</b>
preg-sn	utt47:s: así está bien ?

acepta (46,47) respuesta (48)	utt48:u: sí
----------------------------------	-------------

- **Plan gráfico [plan-graf]:** se pone esta etiqueta a las acciones compuestas que requieren una planificación previa. Se trata de una serie de acciones que involucran mover, quitar o poner varios objetos.

D20)

dir-accion apt-obj apt-zona	utt165:u: podemos mover la mesa con junto con las sillas <sil> sólo un poco hacia la derecha
acepta (165) compromiso	utt166:s: sí
<b>plan-graf</b>	<b>utt167:s: &lt;acción gráfica&gt;</b>
acepta (167)	utt168:u: sí ahí est[á]

- **Rotar vista [rota-vista]:** esta marca corresponde a los movimientos que el sistema hace a la vista tridimensional de la cocina por solicitud del usuario. Es posible rotarla hacia arriba, hacia abajo y hacia los lados para ver la disposición de los muebles desde diversas perspectivas. Por lo general, el sistema gira la vista tridimensional como respuesta a la solicitud de ver la cocina desde otro ángulo por parte del usuario.

D26)

preg-imp apt-obj	utt21:u: mm mh <sil> a ver puedes mover éste eh girar la cocina de tal forma que vea de frente la estufa
acepta (21)	utt22:s: okey
<b>rota-vista respuesta (21)</b>	<b>utt23:s: &lt;acción gráfica&gt;</b>
preg-sn	utt24:s: ahí está bien ?
acepta (23,24) respuesta (24)	utt25:u: sí

- **Despliega [despliega]:** se pone esta marca a los enunciados en los que el sistema muestra en la pantalla el catálogo del mobiliario de la cocina. Al igual que en [rota-vista], esta acción conlleva proveer información en la pantalla; por ello, es común que respondan a solicitudes de información.

D06)

preg-imp	utt225:u: podría mostrarme las sillas ?
acepta (225)	utt226:s: <ruido> <sil> okey
<b>despliega opcion-abierta respuesta (225) apt-obj-coor</b>	<b>utt227:s: ése es el catálogo de bancos sillas y sillones las sillas son estos tres modelos que tienes aquí</b>

Hasta ahora hemos visto los ajustes menores que se hicieron al esquema de DAMSL en el análisis puntal de los enunciados. A continuación presentamos las modificaciones mayores que se hicieron al esquema para crear DIME-DAMSL.

### ***Organización de diálogos prácticos en unidades mayores***

#### ***Transacciones***

DIME-DAMSL propone la noción de “transacción” como unidad de análisis conversacional (Pineda *et al.*, 2007 y 2006). Con ella se pretende modelar diálogos prácticos a partir de la solución de objetivos subordinados. La propuesta se fundamenta en la idea de que en el contexto de los diálogos prácticos la tarea principal se va resolviendo por partes. El objetivo de cada parte se establece y satisface a través de los actos del habla.

Mientras DAMSL dice que los diálogos prácticos son secuencias de actos del habla relacionados hacia adelante y hacia atrás; DIME-DAMSL propone que los diálogos prácticos son secuencias de transacciones vinculadas por un propósito común que es la tarea principal. Esta noción está fuertemente inspirada en los trabajos de Groz y Sidner sobre las contribuciones que los enunciados hacen al discurso.

*“A discourse is coherent only when its discourse purpose is shared by all the participants and when each utterance of the discourse contributes to achieving this purpose, either directly or indirectly, by contributing to the satisfaction of a discourse segment purpose” (Groz y Sidner, 1986:202).*

La propuesta en DIME-DAMSL es que los actos del habla hacen las contribuciones conformando transacciones. Las transacciones tienen una coherencia interna motivada por un propósito al que los dos participantes se abocan. Cuando el propósito se cumple pasan al siguiente hasta dar por concluida su interacción. La transacción es una unidad flexible que contempla la inclusión de un variado repertorio de recursos y posibilidades discursivas; la intención primordial radica en el interés por los procesos de planificación, desarrollo y validación de objetivos a través del lenguaje. Con la intención de crear un terreno abierto al estudio de estas

estrategias, el Grupo DIME presenta la transacción con una estructura interna muy básica de dos fases (Pineda *et al.*, 2007 y 2006):

- **Especificación de la intención:** en esta etapa se manifiesta la intención de solucionar un objetivo; se refinan los detalles y se establecen acuerdos sobre cómo se hará.
- **Satisfacción de la intención:** en esta etapa se intenta cumplir con el objetivo planteado; para ello, se ejecutan acciones y se verifica que hayan sido correctas y que el interlocutor esté de acuerdo con el resultado.

Esta estructura simple permite que dentro de cada fase pueda entrar cualquier protocolo seguido por los hablantes durante las transacciones para lograr sus objetivos. La segmentación de transacciones se realizó durante la etiquetación del corpus. Los enunciados limítrofes reciben una [i] o una [f] si inician o finalizan una transacción respectivamente. Consideremos el siguiente ejemplo:

**D06) Tran69)**

Fase de la Transacción	i/f	ENUNCIADO	FUNCIONES
Especificación de la intención	i	utt278:u: podrías colocar <sil> el comedor con sus sillas <sil> acá ?	dir-accion apt-objs-coor apt-zona
		utt279:s: okey	acepta (278) compromiso
Satisfacción de la intención		utt280:s: <acción gráfica>	plan-graf
		utt281:s:<no-vocal> así está bien ? + +	preg-sn
	f	utt282:u: + + sí	acepta (279,280) respuesta (280)

Esta transacción sólo requirió tres turnos conversacionales para desarrollarse con éxito. En la primera fase, el usuario expresa una directiva de acción señalando los referentes en la pantalla y el sistema acepta realizar lo solicitado comprometiéndose con un “okey”. En la segunda fase, el sistema realiza la acción solicitada (mover un comedor con sus sillas) y verifica que su acción se haya realizado con éxito preguntando “¿Así está bien?”. La transacción termina cuando el usuario acepta el resultado.

Para delimitar el inicio y el fin de las transacciones existen ciertas pistas conversacionales. Por ejemplo, el inicio de transacciones está caracterizado por la aparición de marcadores del discurso que revelan un ordenamiento secuencial en el desarrollo de la tarea (“*después* vamos a poner...”, “*ahora* quiero...”). Por otro lado, la confirmación de acciones realizadas con éxito con preguntas del tipo “¿*Así está bien?*” y su respectiva respuesta permite encontrar los enunciados que terminan transacciones. Consideremos otro ejemplo:

**D01) Tran04)**

Fase de la Transacción	i/f	ENUNCIADO	FUNCIONES
Especificación de la intención	i	utt65:u: ahora necesito mover este <sil> este mueble	dir-accion apt-obj
		utt66:u: ya no recuerdo qué es un gabinete ?	preg-pro
		utt67:s: <no-vocal> <sil> + es un gabinete superior de pared +	respuesta (66) afirm apt-obj
		utt68:u: + <vocal> + bueno el gabinete tengo que moverlo a <sil> al al lado de lo que es el extractor <au> <sil> y encima de lo que es la al[acena] lo que es la alacena	acepta (67) dir-accion apt-objs-coor
		utt69:s: quieres que mueva <sil> este gabinete <sil> acá ?	espera (68) preg-sn apt-obj apt-zona
		utt70:u: sí	acepta (69) respuesta (69)
		utt71:s: okey	acepta (68)
Satisfacción de la intención		utt72:s: <acción gráfica>	mueve-obj
		utt73:s: ahí está bien ? <ruido>	preg-sn
	f	utt74:u: sí	acepta (72, 73) respuesta (73)

La especificación de un objetivo puede requerir múltiples turnos conversacionales que permitan determinar referentes, realizar planes y confirmar información. La primera fase de esta transacción se completa en 6 turnos necesarios para que el usuario establezca un referente y el sistema confirme el plan antes de ejecutar la acción solicitada. La siguiente fase, la acción y su verificación, se resuelven de manera sencilla en dos turnos. Estos ejemplos sólo pretenden mostrar el carácter flexible de la transacción. Consideremos un último ejemplo que presenta otro camino:

**D15) Tran155)**

Fase de la Transacción	i/f	ENUNCIADO	FUNCIONES
Especificación de la intención	i	utt11:u: ahora vamos a es vamos a desplazar la estufa <sil> hacia la e izquierda <sil> para que quede <sil> junto a este mueble pero con una separación de unos diez centímetros + +	dir-accion apt-obj apt-zona apt-tray
		utt12:s: + + okey	acepta(11) compromiso
Satisfacción de la intención		utt13:s: <acción gráfica>	mueve-obj
		utt14:s: así está bien ?	preg-sn
Especificación de la intención		utt15:u: un poquito más hacia la derecha	acepta- parte (13,1 4) respuesta(14) dir-accion
		utt16:s: okey	acepta(15) compromiso
Satisfacción de la intención		utt17:s: <acción gráfica>	mueve-obj
		utt18:s: así está bien ?	preg-sn
	f	utt19:u: [es]tá bien	acepta(17,18) respuesta(18)

En esta transacción, la especificación de la intención se da en dos turnos, sin embargo, su satisfacción requiere una serie de reajustes. Si la acción realizada en una transacción no es del todo exitosa, los hablantes volverán a especificar el objetivo cuantas veces sea necesario e intentarán satisfacerlo hasta lograrlo. En este sentido, DIME-DAMSL propone que puede haber transacciones embebidas dentro de una transacción principal; ésta termina cuando los participantes finalmente acuerdan satisfactorio el resultado.

Segmentar un corpus en transacciones no es una tarea sencilla. Por un lado, los hablantes tienen formas diferentes de proceder al resolver la tarea; algunos son ordenados y planifican objetivos de forma certera, otros son indecisos y cambian de parecer constantemente. Por otro lado, los anotadores no siempre están de acuerdo sobre dónde empieza y dónde acaba una transacción, algunos consideran que los reajustes son transacciones en sí mismas, otros que existen grandes transacciones con muchas transacciones embebidas. Si bien es una noción que requiere más desarrollo e investigación, también es un punto de partida muy útil para organizar el discurso en unidades mayores que los actos del habla.

Recapitulando, hasta ahora hemos dicho que existe una organización: los diálogos prácticos son secuencias de transacciones; hemos dicho que la estructura interna de las transacciones es sencilla de dos fases y que los actos del habla expresan diferentes protocolos conversacionales que completan ambas fases. Hemos dicho que lo que cohesiona a una transacción es el cumplimiento de un objetivo subordinado a la tarea principal. De primera impresión, la transacción resulta una unidad demasiado abstracta y general. Si estuviéramos etiquetando un corpus, solamente tendríamos secuencias de actos del habla que son separadas por etiquetas de inicio y fin de transacción. Todavía no sabemos nada sobre cómo es que los actos del habla se organizan en el interior para conformar transacciones.

*Planos de la expresión: 'Obligaciones' y 'Acuerdo mutuo'*

¿Cómo es que los actos del habla conforman transacciones? DIME-DAMSL propone que las conversaciones se desarrollan sobre dos “ejes” o “planos de expresión” conversacional. Por un lado, define el plano de las *Obligaciones* que incluye los actos que crean y cumplen deberes y compromisos conversacionales (si uno pregunta el otro responde, si uno pide el otro facilita); por otro lado, propone el plano del *Acuerdo mutuo* que contempla los actos del habla que monitorean el entendimiento y la verificación de que las creencias sobre el mundo son compartidas (acuerdos, desacuerdos, malentendidos, etc.). Estos planos se establecen en un nivel abstracto de análisis conversacional y responden a la observación de consensos muy arraigados en la conversación humana.

*“We define the obligations and common ground structures as the relation between speech acts in a conversational transaction, in addition to a number of constraints on such relation” (Pineda et al., 2007:10).*

El esquema propone que la contribución de los actos del habla o que las relaciones que los actos del habla establecen entre sí pueden representarse como “cargos” o “abonos” en las *Obligaciones* y el *Acuerdo mutuo*. Cada que se enuncia un acto se contribuye de alguna u otra forma en esta negociación. Asignar a los actos del habla está “valencia” permite crear pares adyacentes. Del emparejamiento

de cargos y abonos depende el éxito de la transacción. En este sentido DIME-DAMSL postula que las transacciones se “balancean”.

*“Complete transactions are balanced in these two planes of expression; this is for every ‘charge’ in each of these planes there must be a ‘credit’; otherwise, the transaction cannot be completed successfully” (Pineda et al., 2006:1).*

Analizar los actos del habla como relaciones de “cargo-abono” permite la creación de restricciones o reglas que colaboran en el entrenamiento de sistemas computacionales que buscan predecir actos del habla en contextos conversacionales específicos.

### *Obligaciones*

En el plano de las *Obligaciones* se crean los deberes conversacionales. Es tal vez una noción muy abstracta y difícil de explicar detalladamente. Aprendemos desde pequeños cuándo debemos responder cuando nos hablan. Es lo que sucede, por ejemplo, cuando se pregunta algo directamente a alguien. Lo común es que se responda o que se dé una negativa a hacerlo; aunque puede suceder, quedarse callado después de una pregunta se considera poco amable. Es decir, el plano de las *Obligaciones* está asociado a convenciones sociales sobre el intercambio conversacional. Los actos del habla que característicamente crean cargos son las preguntas, las directivas de acción y los compromisos. Los actos del habla que abonan las obligaciones adquiridas son las respuestas y las acciones que responden tanto a directivas enunciadas por otros como a los compromisos hechos por uno mismo.

A continuación presentamos una transacción del Corpus DIME anotada en este plano de expresión. Las dos últimas columnas corresponden a los cargos (C) y abonos (A). El número del enunciado que crea el cargo se anota en la columna de los cargos, su abono se anota en la siguiente columna y lleva el mismo número para establecer la correspondencia con el cargo.

**D06) Tran48)**

i/f	ENUNCIADO	FUNCIONES	OBLIGACIONES	
			C	A
I	utt56:u: podrías colocar esa estufa aquí ?	dir-accion apt-obj apt-zona	56	
	utt57:s: <no-vocal> okey	compromiso acepta (56)	57	
	utt58:s: <acción gráfica>	agrega-obj		57, 56
	utt59:s: <no-vocal> ahí está bien ?	preg-sn	59	
F	utt60:u: sí	respuesta (60) acepta (60)		59

El primer enunciado es una directiva de acción, influye en lo que hará el interlocutor por lo que crea un cargo (56). En el siguiente enunciado el sistema dice “okey” y se compromete a realizar la acción, este compromiso también crea un cargo (57), pero en esta ocasión se crea en el propio hablante. La acción gráfica del siguiente enunciado abona ambos cargos, puesto que la directiva pedía se agregara un mueble que se agregó y el compromiso implicaba la voluntad de agregar un mueble que se agregó. El siguiente enunciado es una pregunta cerrada que crea un cargo (59) que se abona en el siguiente enunciado con la respuesta del usuario.

En la siguiente tabla se presentan las reglas de cargos y abonos en el plano de las *Obligaciones*. Estas reglas corresponden a relaciones frecuentes y estables encontradas en el Corpus DIME. En la columna de la izquierda se especifican los actos que crean cargos, en la columna central aparecen los actos que los abonan. En la columna derecha se especifica en qué participante se crea la obligación de abonar, si en el otro hablante como pasa con una directiva o en el mismo como sucede con los compromisos (Pineda *et al.*, 2006).

**REGLAS DE LAS OBLIGACIONES**

CARGO	ABONO	PARTICIPANTE
sol-info	respuesta	En el otro
dir-accion	accion-grafica	En el otro
compromiso	accion-grafica	En el mismo
oferta aceptada	accion-grafica	En el mismo

### *Acuerdo Mutuo*

El plano del conocimiento mutuo o el acuerdo mutuo se vincula a los actos del habla que establecen qué tanto concuerdan los interlocutores o qué tanto se entienden. Por esta razón, se distinguen dos niveles en este plano de expresión: el acuerdo y el entendimiento. Se podría decir que cada enunciado crea un cargo en ambos niveles de este plano y que se abonan implícitamente continuando la conversación; sin embargo, existen casos en que aparecen de forma explícita, sea para reafirmar que hay acuerdo o para comunicar que no se ha entendido o que se está en desacuerdo.

### *Acuerdo*

Los actos del habla hacen aportaciones al nivel del acuerdo. Las afirmaciones, las preguntas cerradas, las directivas de acción y las acciones gráficas comúnmente son seguidas de un acto de acuerdo explícito en el que se aceptan o se rechazan. Cuando hay acuerdo explícito se marcan cargos y abonos en los enunciados pertinentes. En este sentido, existe una correspondencia directa entre los actos del habla que funcionan como abonos en este plano de expresión y las etiquetas del nivel de *Acuerdo* de las *Funciones hacia atrás* previamente anotadas.

Consideremos los cargos y abonos del nivel de *Acuerdo* en el mismo ejemplo utilizado para las *Obligaciones*.

**D06) Tran48)**

i/f	ENUNCIADO	FUNCIONES	ACUERDO MUTUO	
			CARGOS	ABONOS
I	utt56:u: podrías colocar esa estufa aquí ?	dir-accion apt-obj apt-zona	56	
	utt57:s: <no-vocal> okey	compromiso acepta (57)		56
	utt59:s: <acción gráfica>	agrega-obj	58	
	utt60:s: <no-vocal> ahí está bien ?	preg-sn	60	
F	utt61:u: sí	respuesta (60) acepta (60)		58, 60

En este caso, los cargos se van abonando de forma local. A diferencia del plano de las *Obligaciones*, que el enunciado 56 crea un cargo que se abona después cuando se realiza la acción, en el nivel del *Acuerdo*, la directiva de acción se abona inmediatamente después, cuando el interlocutor acepta realizarla en el enunciado

siguiente. Otra diferencia con el plano de las obligaciones es que el compromiso del enunciado 57 no crea ningún cargo en el acuerdo puesto que comprometerse no requiere que el otro esté de acuerdo necesariamente.

La siguiente tabla presenta las reglas para el nivel de acuerdo mutuo obtenidas del Corpus DIME (Pineda *et al.*, 2006).

REGLAS DEL ACUERDO

CARGO	ABONO
preg-sn	acepta, rechaza, acepta-parte, rechaza-parte.
dir-accion	acepta, rechaza, acepta-parte, rechaza-parte.
oferta	acepta, rechaza, acepta-parte, rechaza-parte.
opcion-abierta	acepta, rechaza, acepta-parte, rechaza-parte.
afirm/reafirm	acepta, rechaza, acepta-parte, rechaza-parte.

*Balanceo de transacciones: cargos y abonos en los planos de expresión*

Como se puede observar, la definición de cargos y abonos en estos planos de expresión permite construir dos estructuras conversacionales que son independientes entre sí pero que de forma conjunta sostienen la cohesión de las transacciones. Los vínculos se van estableciendo a partir de los actos del habla.

*“The notion of conversational obligations and grounding have been applied to the definition of dialogue managers; however, these planes of expression are not reflected directly in annotation schemes” (Pineda et al., 2006:332).*

A continuación presentamos una transacción del Corpus DIME anotada con el esquema DIME-DAMSL expuesto hasta ahora. En el ejemplo consideramos los niveles de DAMSL y la anotación de cargos y abonos en ambos planos de la expresión incorporados por DIME.

D01) Tran06)

	ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DAMSL	OBLIGACIONES		ACUERDO MUTUO			
						ACUERDO		ENTENDIMIENTO	
				C	A	C	A	C	A
i	utt86:u: pue[des] eh puedes eliminar de <vocal> sacar del cuarto la <sil> esta <sil> este fregadero ?	tarea	dir-accion apt-obj	86		86			
	utt87:s: quieres que borre este fregadero ?	admin-com	espera(86) preg-sn apt-obj	87				87	
	utt88:u: sí	admin-com	respuesta(87) acepta(87)		87				87
	utt89:s: okey	tarea	acepta(86) compromiso	89		86			
	utt90:s: <acción gráfica>	tarea	quita-obj		89 86	90			
	utt91:s: así está bien ?	admin-tarea	preg-sn	91		91			
f	utt92:u: así está bien	admin-tarea	respuesta(91) acepta(90,91)		91		91 90		

Desde una perspectiva computacional, la definición de este tipo de estructuras conversacionales permite crear modelos de entrenamiento basados en las reglas obtenidas de la anotación de un corpus. De esta manera, se prevé un escenario en que las conductas conversacionales sean simuladas por sistemas que aprovechen esta información, puesto que a cada estado en que se encuentre el sistema le corresponderá una expectativa vinculada a las reglas de cada plano de expresión previamente presentadas.

El siguiente capítulo muestra el análisis del Corpus DIME con el esquema DIME-DAMSL especificado hasta ahora. El objetivo será determinar cuáles son las transacciones que se presentan de forma esquemática y constante en el corpus a partir de los tipos de situaciones que se crean internamente. Para ello será necesario analizar qué actos del habla son los más frecuentes y cómo es que se vinculan unos con otros.

## 7. Análisis del Corpus DIME con el esquema DIME-DAMSL

### *Datos generales*

Los 20 diálogos del Corpus DIME<sup>17</sup> seleccionados para el análisis abarcan poco más de 4 horas de grabación; fueron transcritos y segmentados en 3,092 enunciados que conforman 276 transacciones. En promedio, cada diálogo tiene 154 enunciados, dura 12 minutos y se desarrolla en 14 transacciones; cada transacción se conforma por una media de 11 enunciados. La Tabla 7.1 presenta la información general de los 20 diálogos que se utilizaron para el análisis.

Tabla 7.1 Información de cada diálogo

Diálogo	1	2	3	5	6	10	11	12	13	14	15	17	18	19	20	21	22	23	25	26	20
# enun	105	188	163	112	370	84	235	113	179	131	87	233	187	96	160	70	180	80	111	208	3092
# tran	6	14	9	14	36	9	30	9	16	10	8	22	11	11	18	7	16	9	9	12	276
Duración mm:ss	08:16	18:01	06:53	07:56	27:36	07:38	20:12	07:19	14:00	08:46	05:45	10:56	20:35	10:10	15:29	04:24	13:32	05:54	08:10	20:05	04:01:37

Hasta arriba, de izquierda a derecha se listan los diálogos. Se notará que la numeración es discontinua, esto se debe a que el rótulo corresponde al identificador dentro del corpus, en total son 20. De arriba abajo aparecen el número de enunciados, el número de transacciones y la duración en minutos y segundos que tuvo cada diálogo.

### *Anotación DIME-DAMSL*

Durante el análisis se asignaron más de 15,500 marcas al corpus, cada enunciado recibió en promedio 5 etiquetas del esquema DIME-DAMSL. Fue un arduo trabajo que requirió la colaboración de distintas personas y la revisión constante de la metodología de anotación.

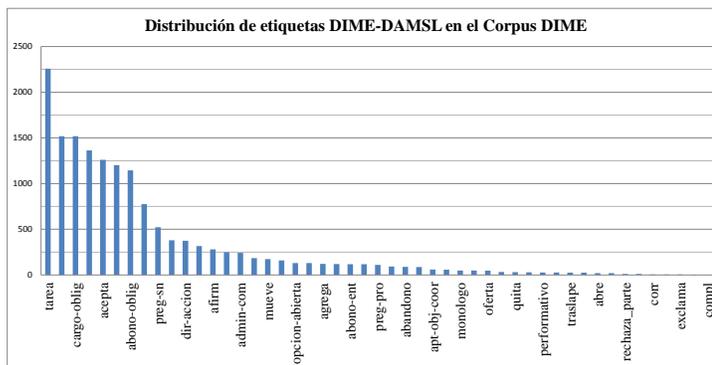
La Tabla 7.2 presenta a detalle la distribución de las etiquetas DIME-DAMSL en cada diálogo. Por limitaciones de espacio aparece un poco ajustada; sin embargo, es posible apreciar que exceptuando algunos casos especiales, las etiquetas del esquema se distribuyen con cierta uniformidad por el corpus.

<sup>17</sup> Sobre las características del corpus y su recopilación ver Capítulo 4.

Tabla 7.2 Distribución de las etiquetas DIME-DAMSL en los diálogos del Corpus DIME

D	CARGOS Y ABONOS					EDO-COM		NIV-INFO		FUNCIONES HACIA ADELANTE													ACCIONES GRÁFICAS					FUNCIONES HACIA ATRÁS					GESTOS																		
	CO	AD	CA	AA	CE	AE	inin	mond	aban	taea	ad-tar	ad-com	otro	afirm	resfir	p-sin	p-pro	dir-ac	lo-ab	oferta	compr	abre	abierta	perf	excl	despl	mueve	preg	quita	plap	el	rota	respl	acep	resp	quita	espera	rech	rech	SNE	ack	ack-cl	rep	compl	corr	apl-ob	apl-so	rep	trasp	obs	cor
01	46	35	49	40	6	6	3	4	35	40	12		19	16	6	1	11	3	2	10	1	1	1	1	2	7	3	1				23	37		1	6	3	3	2	5	4	1	1	19	10	1	2	488			
02	73	53	96	79	5	5	7	8	110	80	10		29	1	21	2	14	14	11	6	10	1	1	1	1	9	2	8	1	3	8	36	77		1	9	6	1	6	8	4	1	24	22	11	873					
03	47	41	74	59	7	7	8	6	76	67	20		26	2	35	7	11	2	2	10	1	1	5	1	2	7	4	2	1	1	29	56	2		15	4	1	1	13	7	7	13	5	1	675						
05	68	48	56	54	2	2	1	94	49	8	1	6	22	4	2	22	2	1	18	1	1	2			2	17	2		1	28	45	5		4	7			2	1	3	18	15	1	617							
06	202	153	192	185	5	5	1	3	322	142	16		22	1	68	9	29	47	16	6	44	1	1	2		16	9	24	11	2	17	105	158	6		14	24	1	12	2	4	1	38	38	4	1964					
10	45	32	37	35	5	5	1	1	61	38	9		6	17	2		12	1	3	12	1	1	1	1	1	7	3	1	2	19	35	1		6	3		1	5	1	3	15	12	2	443							
11	137	102	108	105	11	11	1	5	183	113	11		6	1	52	4	17	31	8	3	29	1	1	1	1	1	9	20	9	1	8	72	105	3		18	5	2	7	9	1	40	37	4	1293						
12	58	47	53	45	6	6			93	34	8		9	20	7	9	11	9	1	11	1	1	1		9	7	3	1		36	45	1		9	6			3	7	7	12	4	2	575							
13	96	71	91	86	7	7	3		138	71	11		14	1	30	5	11	26	11	2	24	1	1	1		11	15	6	1	2	45	75	5		1	7	9	3	2	5		18	13	1	930						
14	57	45	64	59	5	5	2	5	90	46	10		20	2	27	11	1	11	9	1	11	1	1	1		9	2	8		1	34	62			11	2			4	5		13	4	1	639						
15	44	33	40	36	6	6	1	3	56	32	12		7	3	18	3	2	10	2	1	10	1	1	2		2	8	2		23	38	1		8	1		2	3	5		13	6	1	444							
17	124	93	123	114	3	3	2	12	193	73	9	1	12	1	42	3	12	40	8		27	1	1	1		8	23	7	4	4	3	55	96	9		9	10	1	6	2	1	36	13	14	7	1206					
18	79	58	87	68	14	14	9	13	100	76	28	1	23	3	26	12	2	22	8	6	14	1	1	1		8	11	3	1		39	63	1		14	9	1	7	2	3	9		17	4	1	859					
19	39	31	49	45	3	3		1	76	33	2	1	19	1	14	6	4	7	6	1	6	1	1	2		5		6		4	24	41	2		5	2	2		2	2		8	8	2	464						
20	80	63	82	73	3	3	1	5	132	61	5	1	9	29	2	19	15	10	2	13	1	1	1		11	3	10		2	6	48	70			12	5			1	2	5		18	17		821					
21	35	26	33	30	5	5			54	24	13	1	7	3	11	6		11	3	2	7	1	1	1		3	3	3	1		17	31			4	1	3	4	2		8	4	2	365							
22	94	69	94	89	5	6		1	136	80	15	1	15	1	35	2	15	19	10	3	20	1	1	1		10	5	11	3		5	50	85	2		1	10	6		7	4	1	17	11	3	943					
23	46	33	35	34	5	5		3	61	29	12	1	4	1	14	4	2	14	2	1	12	1	1	1		2	9	3	1		20	32	3		3	2		2	4	3		17	9	2	433						
25	54	41	56	47	8	8	2	3	5	77	42	17	1	15	4	13	9	2	18	2	1	10	1	1	1		2	13	1		23	38	5		2	5	1	6	2	1	1	16	5	1	3	563					
26	94	70	99	81	6	6	11	12	152	72	14	1	13	3	28	7	17	22	7	5	18	1	1	1		9	6	6	1	7	50	72	2		18	10		3	4	2	1	19	12	3	3	974					
total	1518	1144	1519	1364	117	118	6	49	88	2259	1202	242	10	281	27	523	111	159	374	130	47	316	20	20	28	5	130	174	122	30	25	58	776	1261	48	3	184	120	11	32	92	26	86	1	6	379	249	29	60	15579	

La gráfica que se presenta a continuación corresponde a la distribución de estas 15,579 etiquetas organizadas de mayor a menor frecuencia de aparición. El eje horizontal lista las etiquetas; el eje vertical representa el número de veces que se asignó cada una de ellas.



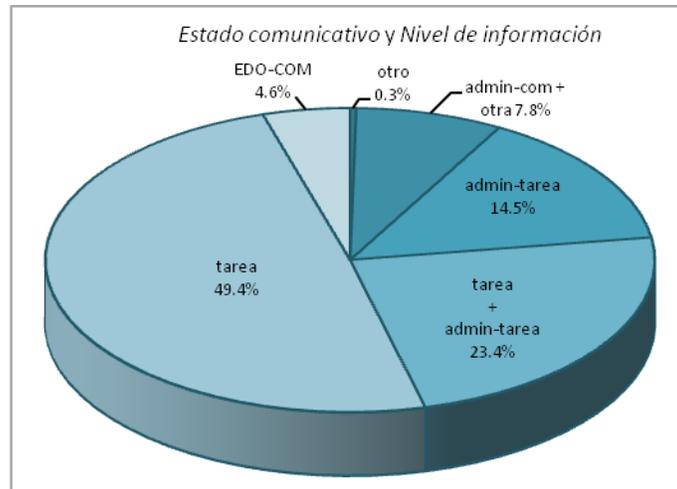
Gráfica 7.1 Distribución de etiquetas DIME-DAMSL en el Corpus DIME

Los datos que se han presentado hasta ahora corresponden al número de veces que apareció cada etiqueta; a continuación presentaremos cómo se distribuyeron estas etiquetas en los enunciados revisando cada nivel del esquema DIME-DAMSL.

### *Estado comunicativo y Nivel de información*

La Gráfica 7.2 muestra la distribución de etiquetas en estos dos niveles de análisis.

El total corresponde a los 3,092 enunciados del corpus.



Gráfica 7.2 Distribución de enunciados - 'Estado comunicativo' y 'Nivel de información'

Observando la gráfica a contrarreloj encontramos primero la sección más clara correspondiente al *Estado comunicativo*. Se recordará que las etiquetas de este nivel manifiestan que el enunciado no tuvo aportación comunicativa porque se abandonó, porque se trataba de un monólogo o porque fue ininteligible sin que el interlocutor se interesara en aclarar lo expresado. El total de enunciados sin aporte comunicativo [EDO-COM] conforman apenas el 4.6% del corpus.

La dimensión de *Nivel de información* incluye tres etiquetas sobre el contenido informativo del enunciado: [tarea], [admin-tarea] y [admin-com]. Estas etiquetas no son excluyentes entre sí, por lo que un enunciado podía recibir más de una. La gráfica permite ver que la etiqueta de [tarea], sin otra etiqueta en este nivel, fue asignada a casi el 50% del corpus.

El siguiente bloque corresponde a los 723 enunciados que fueron etiquetados como [tarea] + [admin-tarea]. Todos estos enunciados pertenecen a protocolos de verificación: “¿Así está bien?” y su respuesta afirmativa o negativa. Parece ser que los anotadores consideraron confuso distinguir si se habla de la tarea o de su administración durante estos protocolos. Arriba del 23% de los enunciados del corpus fueron anotados con este par de etiquetas.

El siguiente bloque corresponde a los enunciados que recibieron únicamente la etiqueta de [admin-tar] y suman 448; es decir, arriba del 14%. La mayoría son afirmaciones en que se explica de qué trata la tarea o cómo debe desarrollarse.

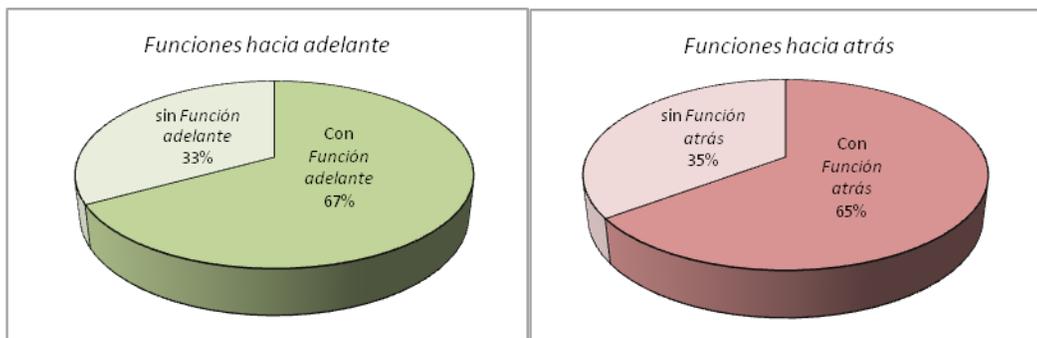
Los enunciados que administran la comunicación suman 242 (7.8%) y manifiestan el grado de entendimiento del enunciado previo, sea para decirle al interlocutor que se entendió lo que dijo, como para expresarle que no se entendió. En 201 casos la etiqueta [admin-com] fue asignada sin otra marca de este nivel; en 41 casos apareció junto a [tarea] o [admin-tarea].

Finalmente, “otro” corresponde a 10 enunciados (0.3%) en que los anotadores no asignaron alguna de las etiquetas previas puesto que se trata de exclamaciones o agradecimientos (que están listadas en las *Funciones hacia adelante* pero no refieren ni a la tarea, ni a su administración, ni al canal de comunicación).

Se recordará que los niveles de análisis representados en la Gráfica 7.2 corresponden a una primera revisión, posteriormente se anotan las *Funciones hacia adelante* o *hacia atrás* que contemplan los efectos comunicativos que tuvo un enunciado en su contexto de enunciación.

#### *Funciones hacia adelante y Funciones hacia atrás*

Consideremos el siguiente par de gráficas que presentan la distribución de enunciados con *Funciones hacia adelante* a la izquierda y *Funciones hacia atrás* a la derecha. Se recordará que estos niveles integran etiquetas como [oferta], [direccion], [acepta], [rechaza], etc.

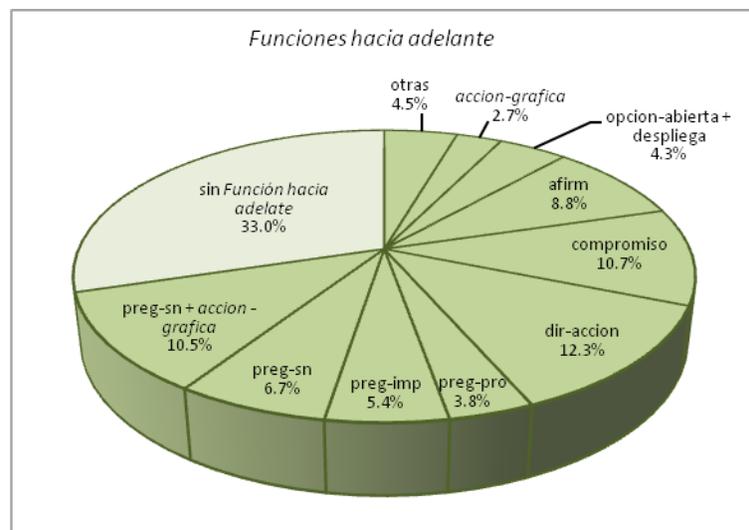


Gráfica 7.3 Comparativa de la distribución - 'Funciones hacia adelante' y 'Funciones hacia Atrás'

La distribución de enunciados con etiquetas en ambos niveles es similar porque es parecido el número de enunciados que recibieron al menos una etiqueta de las *Funciones hacia adelante* (2,073) y el número de enunciados que recibieron al menos una etiqueta de las *Funciones hacia atrás* (2,005)<sup>18</sup>. Estas gráficas permiten ver que los enunciados se tejen hacia adelante y hacia atrás en una proporción similar en el corpus. Ahora detallemos qué funciones se anotaron en cada nivel.

### *Funciones hacia adelante*

El esquema DIME-DAMSL define las *Funciones hacia adelante* a partir de la influencia que su expresión tiene en el flujo futuro del diálogo. La gráfica que se presenta a continuación muestra la distribución de las etiquetas de este nivel en los 3,092 enunciados del corpus. El segmento más claro del lado izquierdo de la gráfica corresponde a los enunciados que no recibieron etiqueta en este nivel. Posteriormente, de mayor a menor frecuencia de aparición, los segmentos verdes van presentando la distribución de las etiquetas de las *Funciones hacia adelante*.



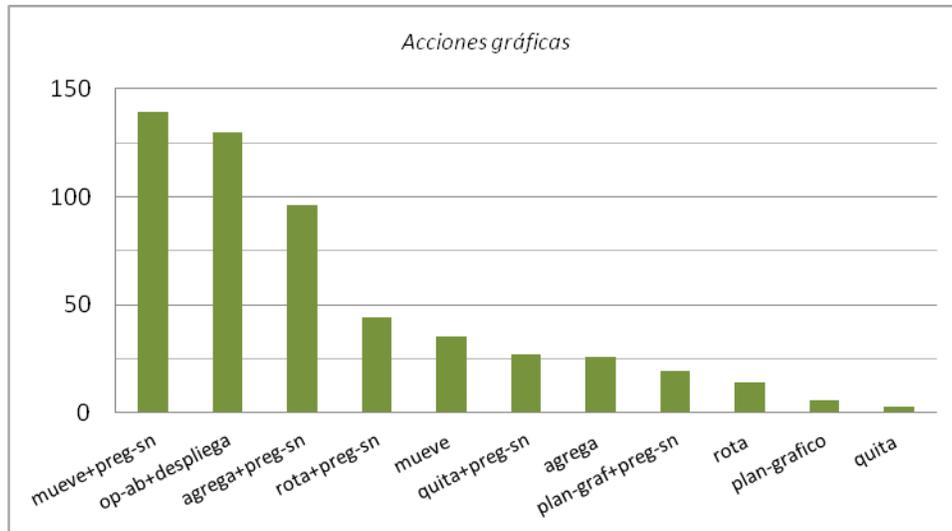
Gráfica 7.4 Distribución de enunciados - 'Funciones hacia adelante'

<sup>18</sup> Para realizar las gráficas tipo “pastel” se omitieron los datos de los gestos deícticos. Esto se debe a que los gestos aparecen siempre auxiliando a otras funciones comunicativas; es decir, no tienen un valor equivalente a una afirmación o a una pregunta. El hecho de que nunca aparezcan de forma independiente corrobora que funcionan como herramientas referenciales integradas a directivas, ofertas, afirmaciones, respuestas, etc. El análisis de los gestos se presenta más adelante.

Si bien el esquema es multidimensional, por lo general, dentro del nivel de las *Funciones hacia adelante*, las directivas son directivas, las ofertas son ofertas, las preguntas son preguntas, etc. De los 2,073 enunciados etiquetados con marca *hacia adelante*, 1,566 recibieron una sola etiqueta en este nivel y 507 recibieron dos. A continuación veremos su distribución.

- *Solicitudes de información*: Las protagonistas de las *Funciones hacia adelante* son las preguntas, con un total de 793 enunciados que corresponden a poco más del 25% del corpus. De mayor a menor frecuencia, las preguntas más encontrada son cerradas [preg-sn] que aparecieron 523 veces, después las preguntas imperativas [preg-imp] con 159 apariciones y, por último, las pronominales [preg-pro] con 111 enunciados. Del total de preguntas cerradas, 325 veces se anotó también alguna etiqueta de las *Acciones gráficas*: [mueve], [quita], [agrega], etc. En estos casos, el sistema pregunta “¿Así está bien?” justo después de realizar la acción. Es indudable que esto resultó de las metodologías de segmentación y anotación seguidas; no significa que la acción y la pregunta ocurrieran simultáneamente. Parece ser que resultó confuso entre la segmentación de enunciados y la anotación de DIME-DAMSL determinar qué metodología se seguiría cuando en un mismo turno hay un silencio mientras se ejecuta una acción gráfica.

- *Acciones gráficas*: En total se anotaron 539 acciones gráficas en todo el corpus. A continuación se presenta la gráfica de su distribución.



Gráfica 7.5 Distribución de etiquetas de 'Acciones gráficas'

Cabe destacar que estas acciones tienen una influencia directa en la solución de la tarea y aparecen en 270 transacciones. Esto significa que casi el 98% de las transacciones están enfocadas a la ejecución de acciones gráficas. Las 6 transacciones restantes en que no aparece ninguna acción gráfica son pequeños protocolos en que se termina el diálogo o se rechaza la realización de alguna acción por algún motivo que lo imposibilita (“*el mueble no cabe ahí*”).

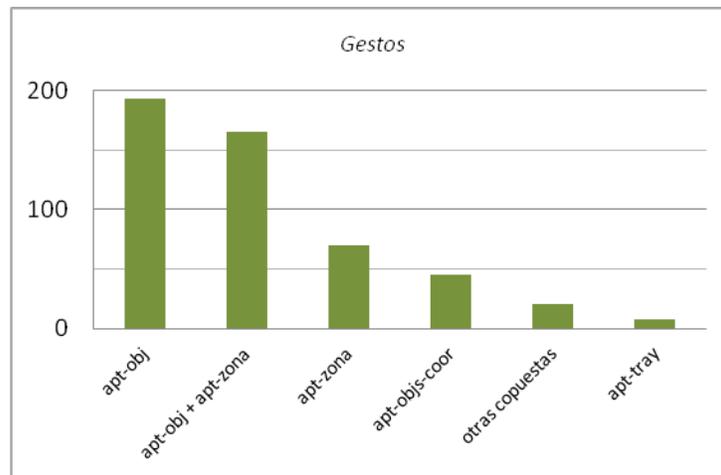
- *Directivas de acción:* El siguiente grupo de enunciados frecuentes corresponde a las directivas de acción con 374 apariciones distribuidas en 242 transacciones.

- Después, están los *Compromisos*, que aparecieron en 316 enunciados. De las 242 transacciones en que aparecen directivas de acción, en 231 aparece al menos un compromiso.

- *Otras Funciones hacia adelante:* Con menor frecuencia aparece la etiqueta de [afirm], asignada a 281 enunciados, 8.4% del corpus. También, un total de 126 enunciados (4.1%) se etiquetaron como [opcion-abierta] + [despliega]. Se recordará que este par de marcas manifiestan que el enunciado muestra en la pantalla un menú de opciones. El segmento titulado “otras” en la Gráfica 7.4 representa el 4.3% del corpus y corresponde a 134 enunciados, se trata de un cajón de sastre con etiquetas

simples y mixtas con una frecuencia muy baja. Entre ellas se encuentran 48 ofertas, 27 reafirmaciones, 25 agradecimientos, 5 exclamaciones, etc.

- *Gestos*: Las etiquetas de gestos se asignaron a los enunciados en que se apuntan objetos, zonas o trayectorias con la intención de establecer referentes. Un total de 513 enunciados distribuidos en 233 transacciones recibieron 717 marcas correspondientes a gestos. Tienen una alta frecuencia de aparición y siempre acompañan a otras etiquetas de las *Funciones hacia adelante o hacia atrás*. Más del 84% de las transacciones utilizan al menos un deíctico. La Gráfica 7.6 presenta la distribución de etiquetas en este nivel.



Gráfica 7.6 Distribución de etiquetas de 'Gestos'

Los gestos aparecen casi siempre acompañando directivas de acción. Un total de 221 enunciados recibieron la etiqueta [dir-accion] + [*Gesto*]; como en “quiero que pongas *esto aquí*”, “hazlo un poquito más para *allá*”). Lo más frecuente fue asignar el par de etiquetas [apt-obj]+[apt-zona] (“mueve *esto* para *acá*”) que apareció junto a 111 directivas.

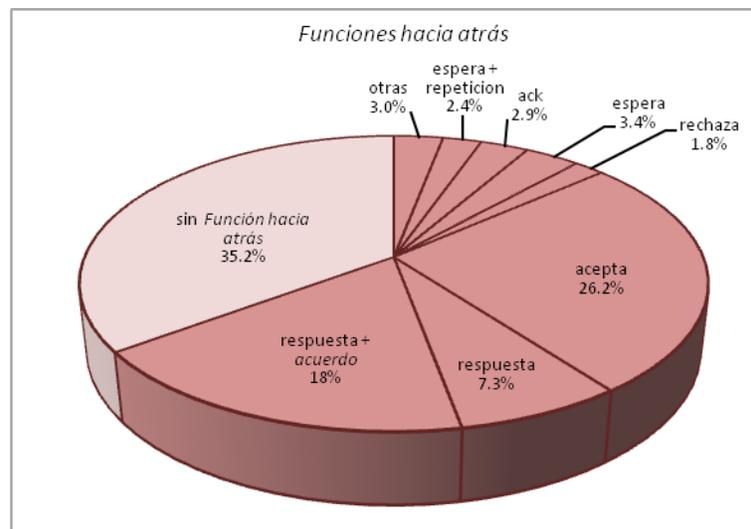
También fue común encontrar que los gestos acompañan a algunas preguntas cerradas. De 109 que tuvieron algún gesto, en 105 también se anotó [espera] en el nivel de *Acuerdo* y [admin-com] en el *Nivel de información*. Todas estas preguntas cerradas están vinculadas a la confirmación de que se ha entendido correctamente el enunciado previo. “¿Quieres que ponga *esto acá*?”, “¿*Éste*?”, “¿Quieres saber qué diferencia hay entre *este* y *éste*”, etc.

Hubo 106 enunciados que recibieron la etiqueta [afirm] más algún gesto (“*ésta* es una alacena”, “*éste* es más grande que *éste*”, etc.). Los 76 gestos restantes están distribuidos en otras *Funciones hacia adelante* con una frecuencia de aparición menor a 30 enunciados.

### *Funciones hacia atrás*

En la Gráfica 7.7 se presenta la distribución de enunciados con marcas de las *Funciones hacia atrás*. Se recordará que este nivel incluye etiquetas correspondientes a tres vínculos distintos que los enunciados establecen con enunciados previos:

- *Acuerdo*: [acepta], [rechaza], [rechaza-parte], [espera], [quizá] etc.
- *Entendimiento*: [ack], [sne], [repetición], [corrección], etc.
- *Respuestas*



Gráfica 7.7 Distribución de enunciados - 'Funciones hacia atrás'

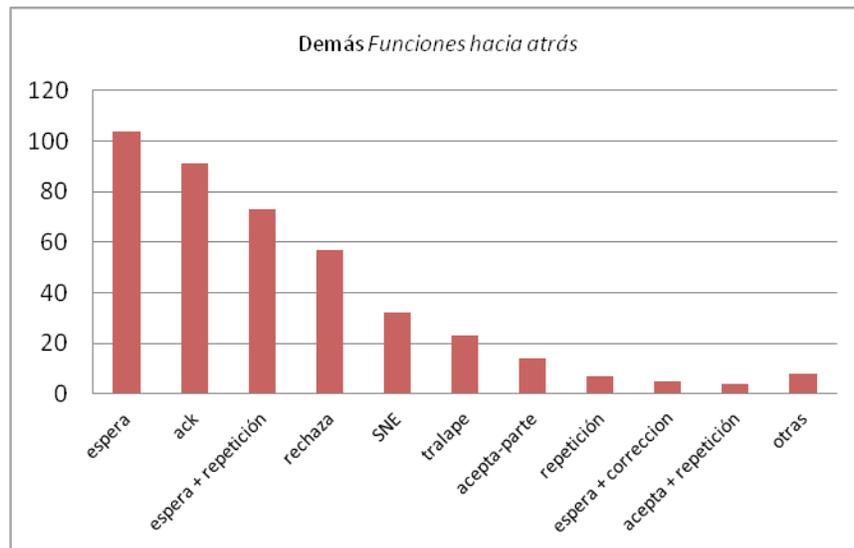
El 65% de los enunciados recibieron al menos una etiqueta de las *Funciones hacia atrás*. Las respuestas suman un total de 776 enunciados que representan el 25% del corpus. Se vincula al 25% de las solicitudes de información revisadas previamente. Del total de respuestas, 549 veces aparece con otro acto de *Acuerdo* o *Entendimiento* y en 227 no aparece con ninguna otra etiqueta de las *Funciones hacia atrás*. A continuación se presenta la Tabla 7.3 con la distribución de etiquetas en los enunciados que recibieron [respuesta] + otra *Función hacia atrás*.

Tabla 7.3 Distribución de las etiquetas compuestas - [respuesta] + otro acto

resp +	Acuerdo					Entendimiento	Total
	acepta	rechaza	acepta-parte	rechaza-parte	quizá		
	440	62	31	9	2	5	549

La etiqueta más anotada en las *Funciones hacia atrás* fue [acepta] con 1,261 apariciones, 811 individuales y 450 con otro acto; de estos, 440 fueron respuestas que aceptaban algo y 10 enunciados muy particulares que recibieron [acepta] + otra etiqueta de *Acuerdo* o *Entendimiento*.

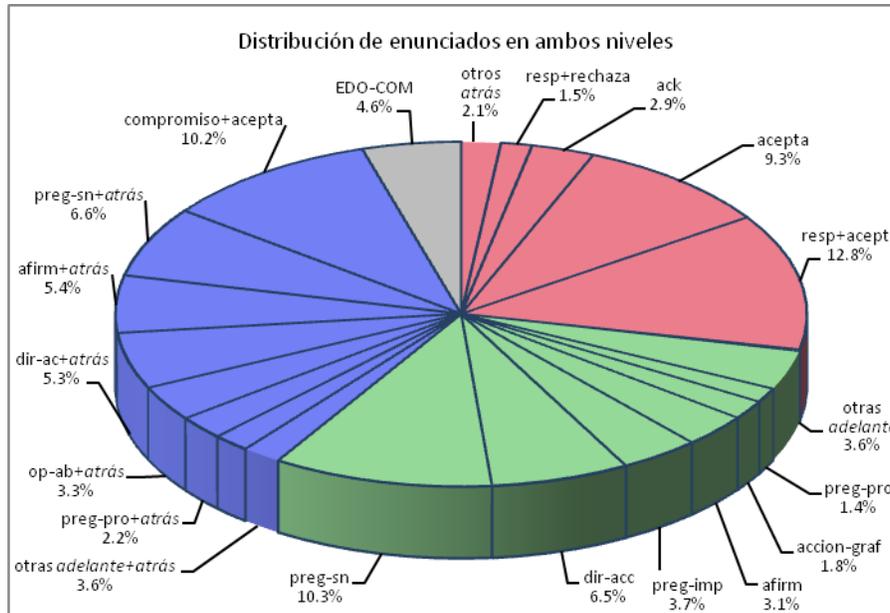
Aceptar y responder son acciones que corresponden a poco más del 50% del total del corpus. Las demás etiquetas de *Acuerdo* ([rechaza], [espera], [acepta-parte], etc.) y de *Entendimiento* ([ack], [sne], [repetición], etc.) suman 418 enunciados que corresponden al 13% del corpus. Todos ellos tuvieron una frecuencia de aparición menor a 150 enunciados. A continuación se presenta la gráfica con su distribución.



Gráfica 7.8 Distribución de las demás 'Funciones hacia adelante'

*Distribución de las Funciones hacia adelante y hacia atrás en los enunciados*  
 Ambos niveles distribuyeron etiquetas en más del 60% del corpus respectivamente. La gráfica que presentamos a continuación muestra la distribución de enunciados que tuvieron etiquetas sólo en las *Funciones hacia atrás* (segmento rosa) o que tuvieron etiquetas sólo en las *Funciones hacia adelante* (segmento verde) o que tuvieron etiquetas en los dos niveles (segmento azul) o que no recibieron etiqueta

en ninguno de los dos niveles (segmento gris que corresponde al 4.6% del corpus con etiquetas del *Estado comunicativo*).



Gráfica 7.9 Distribución - 'Funciones hacia adelante' y 'Funciones hacia Atrás' en el Corpus DIME

De forma general, habíamos revisado la distribución de las etiquetas sin considerar las etiquetas compuestas por *Funciones hacia atrás* y *hacia adelante*. Esta gráfica pretende poner en relieve la flexibilidad y detalle que adquiere un esquema multidimensional cuando se anota un corpus. Cada segmento de la gráfica representa un conjunto de enunciados que cumplen una función comunicativa específica. Los enunciados que reciben un mismo conjunto de etiquetas resultan similares en los contenidos que expresan y las formas que toman. Considérense los siguientes enunciados tomados de distintas transacciones que reciben el mismo par de etiquetas.

[respuesta] + [afirm]

D	T	ENUNCIADO	FUNCIONES ADELANTE	FUNCIONES ATRÁS
14	150	utt76:u: <ruido> arriba de la alacena inferior	afirm	respuesta
17	172	utt128:u: + + aquí	afirm	respuesta
19	196	utt26:u: a ver en este <sil> en este lugar	afirm	respuesta
18	184	utt31:s: tenemos dos refrigeradores ahora uno encima de otro	afirm	respuesta
25	260	utt68:s: es una alacena inferior	afirm	respuesta

Los enunciados que reciben [respuesta] + [afirm] tienen vínculo directo con las preguntas pronominales del tipo “¿Dónde quieres que lo ponga?” o “¿Qué es esto?” Ahora consideremos los siguientes ejemplos, todos corresponden a respuestas que simultáneamente son directivas de acción:

[acepta-parte] + [respuesta] + [dir-accion]

D	T	ENUNCIADO	FUNCIONES ADELANTE	FUNCIONES ATRÁS
5	32	utt26:u: eh <ruido> un poquito <sil> un poquito más despegado	dir-accion	acepta-parte respuesta
12	119	utt20:u: un poco menos por favor	dir-accion	acepta-parte respuesta
13	137	utt92:u: lo puedes poner más separado	dir-accion	acepta-parte respuesta
22	233	utt28:u: podrías ponerla un poco más alta ?	dir-accion	acepta-parte respuesta
25	258	utt49:u: un poco el extractor un poco más hacia la derecha	dir-accion	acepta-parte respuesta

Los enunciados anotados como [respuesta] + [dir-accion] responden a la pregunta “¿Así está bien?” y forman parte de los reajustes que se hacen a una acción previa. También van acompañados de una etiqueta [acepta-parte] o [rechaza-parte], según sea el caso y de las etiquetas [admin-tarea] o [tarea].

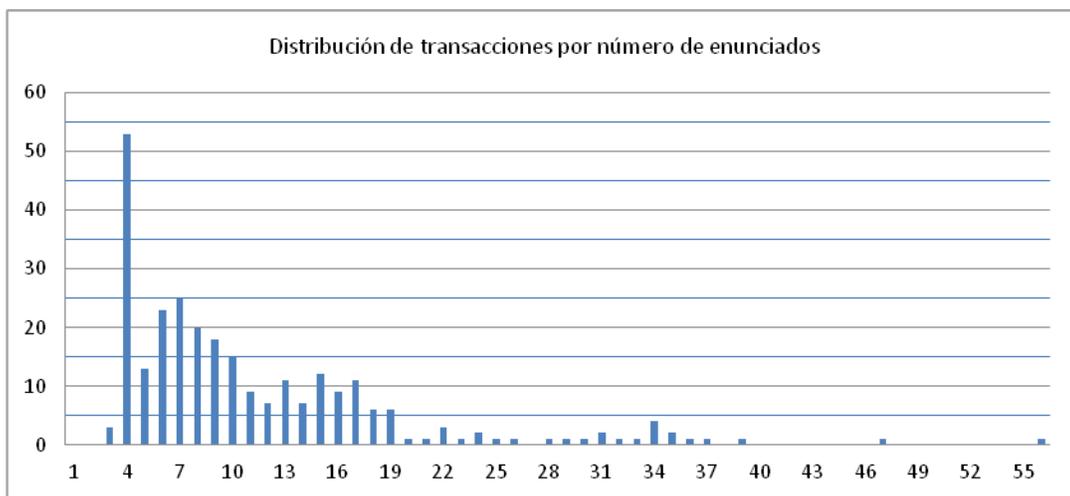
Con la presentación de estos ejemplos sólo pretendemos enfatizar la productividad de un esquema multidimensional como DIME-DAMSL. Al etiquetar diversos aspectos comunicativos que operan en un mismo enunciado es posible localizar enunciados que cumplen funciones similares y que pertenecen a protocolos conversacionales similares. Sería un trabajo extenuante hablar de todas las combinatorias del esquema de forma individual. Es justo aquí donde la noción de transacción resulta una oportuna organización que agrupa estas funciones comunicativas. A continuación presentaremos el análisis de las transacciones en el Corpus DIME.

### ***Análisis de transacciones***

Los enunciados que se agrupan en una transacción tienen un objetivo que se resuelve por mutuo acuerdo entre los hablantes. Los casos en que se abandona un objetivo también encontramos que por mutuo acuerdo se termina una transacción

con la intención de empezar otra. El cumplimiento de estos objetivos encamina la solución de la tarea de forma incremental. Dentro de una transacción, el intercambio de turnos y enunciados se mantiene cohesionado gracias a este propósito común que se resuelve; transacción tras transacción, el diálogo se dirige hacia la solución de la tarea.

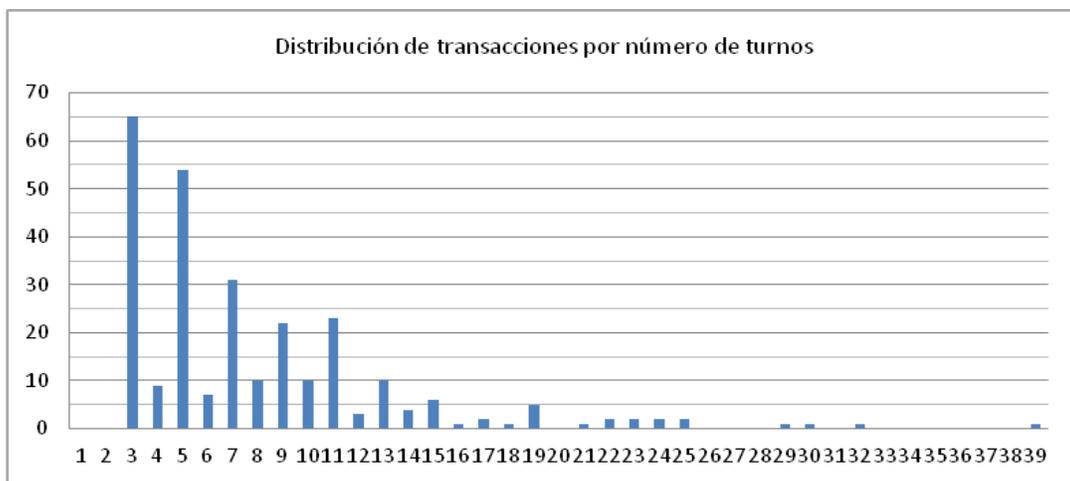
¿Cuántos enunciados se requieren para resolver una parte de la tarea? La longitud de las 276 transacciones del corpus es variable; va desde los 3 hasta los 56 enunciados. La Gráfica 7.10 presenta la distribución de la longitud de transacciones con relación al número de enunciados que la componen. El eje horizontal corresponde al número de enunciados, el eje vertical corresponde al número de transacciones con esa cantidad de enunciados.



Gráfica 7.10 *Distribución de transacciones por número de enunciados*

Como puede observarse, las transacciones de cuatro enunciados fueron las más frecuentes con 53 apariciones. Tras una revisión detallada, resultó que estas 53 transacciones se resuelven en tres turnos conversacionales y que se observa un mismo grupo de etiquetas que se repite de forma constante. Encontrar este pico en la distribución de enunciados estableció ciertas pautas básicas de interacción y resaltó los actos del habla preferidos por los hablantes para el cumplimiento de objetivos de manera rápida y eficiente. Se nos ocurrió, entonces, que lo que pasa en las transacciones de cuatro enunciados tal vez también pasa en las de cinco o en las de tres; es decir, que podía haber más transacciones de tres turnos en las que se

repitieran estos modelos de interacción básicos, incluso cuando hubiera un enunciado más o uno menos. A diferencia de la segmentación de enunciados, el intercambio de turnos no está sujeto a la interpretación del segmentador y, comúnmente, la alternancia de turnos está vinculada a la expresión de contenido informativo relevante para la tarea y para el diálogo. Con la intención de revisar la distribución de turnos en las transacciones presentamos la siguiente gráfica, el eje horizontal corresponde al número de turnos que componen la transacción; y el eje vertical corresponde al número de transacciones.



Gráfica 7.11 *Distribución de transacciones por número de turnos*

Llama la atención que la distribución de transacciones presente más de un pico cuando se considera con relación a los turnos conversacionales. También llama la atención que estos picos estén vinculados a las transacciones que se realizan en un número impar de turnos (3, 5, 7, etc.). Como puede observarse, las transacciones realizadas en número par tienen una frecuencia mucho menor.

Sea que se tome como base del análisis los enunciados o los turnos, la atención se dirige hacia un mismo punto de partida: las transacciones cortas. Para este análisis se consideró más relevante la perspectiva del intercambio de turnos que la de número de enunciados. Decidimos revisar primero las transacciones con tres turnos conversacionales que corresponden al 23% del corpus. Este primer acercamiento permitirá la definición de ciertos modelos de interacción básica en el contexto de los diálogos prácticos. Estos modelos serán un punto de partida para la

especificación incremental de esquemas que cubran con mayor amplitud los fenómenos conversacionales más relevantes que se observaron en las transacciones.

*Transacciones de tres turnos conversacionales*

Las transacciones más cortas del corpus se realizan en tres turnos. En total son 65 e incluyen las 53 transacciones de cuatro enunciados previamente mencionadas. Adicionalmente, se integran 3 transacciones de tres enunciados, 7 transacciones de cinco enunciados y 2 transacciones que tienen seis enunciados.

La Figura 7.1 presenta las etiquetas que encontramos una y otra vez en los enunciados de estas 65 transacciones. Como son las *Funciones hacia adelante* las que dirigen la conversación, presentamos cada estrategia conversacional con el acto que encabeza el primer turno. De izquierda a derecha se representa la secuencia de tres turnos, con las flechas se simboliza la alternancia entre hablantes.



Figura 7.1 *Secuencias de etiquetas más frecuentes en transacciones de tres turnos*

Vale la pena apuntar que las 65 transacciones de tres turnos están representadas en la figura sólo con seis modelos de interacción que surgen a partir de sólo cuatro *Funciones hacia adelante*. Se trata de un grupo limitado de etiquetas conformado por directivas de acción, preguntas, acciones gráficas, respuestas y acuerdos. Estas secuencias de etiquetas se muestran como esquemas estables de interacción en las transacciones cortas. Son 65 transacciones de tres turnos, 48 de ellas siguen el primer modelo, [dir-accion]. El segundo, [preg-imp], corresponde a 14 casos. Las etiquetas [oferta] y [preg-sn] comenzaron 3 transacciones. Si bien este último par de modelos tiene una frecuencia de aparición mucho menor, con relación

a los primeros dos, representan otras posibles y válidas rutas en el contexto de los diálogos prácticos. A continuación detallamos y ejemplificamos con algunas transacciones de tres turnos los modelos presentados en la Figura 7.1

*Modelo de tres turnos: directivas de acción*

Lo más frecuente con 48 casos es que las transacciones de tres turnos tengan una directiva de acción en el primero, que en el segundo el interlocutor acepte la directiva, ejecute la acción y pregunte si estuvo bien, y que en el tercero se responda aceptando la acción realizada. De estas 48 transacciones, 40 tienen cuatro enunciados y las demás tienen 3, 5 ó 6. Todas ellas siguen la secuencia de etiquetas que se presenta en la Figura 7.2.

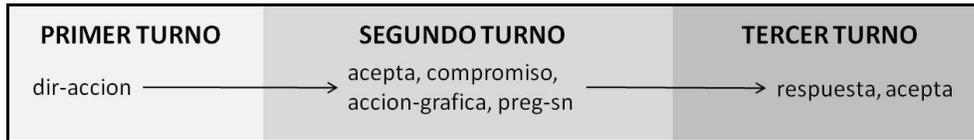


Figura 7.2 Modelo de tres turnos con directiva de acción

Consideremos los siguientes ejemplos de transacciones de tres turnos completadas en 3, 4, 5 y 6 enunciados respectivamente. Para facilitar la delimitación de turnos, se marca con gris la alternancia entre los interlocutores:

**3 enunciados - D17) Tran168) - dir-accion**

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO	
			C	A	C	A
utt67:u: este [re]frigerador lo quitamos	tarea	dir-accion apt-obj	67		67	
utt68:s: <no-vocal> así ?	tarea admin-tarea	acepta[67] accion-graf preg-sn	68	67	68	67
utt69:u: sí	tarea admin-tarea	acepta[68] respuesta[68]		68		68

Dentro del corpus, este ejemplo corresponde a la única transacción de tres turnos que se realiza en tres enunciados. En ella observamos presente el modelo con una pequeña variación: en ella se realiza un acuerdo implícito después de la directiva de acción. El “sistema” directamente ejecuta la acción sin proferir palabra alguna. La metodología de anotación de DIME-DAMSL especifica que los

acuerdos implícitos se marcan igual que los explícitos<sup>19</sup>. Sin embargo, los compromisos no. Esto implica que con relación al modelo presentado falta la etiqueta [compromiso] mientras están presentes todas las demás. Ahora veamos el ejemplo de tres turnos en cuatro enunciados:

**4 enunciados - D12) Tran125) - dir-accion**

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO	
			C	A	C	A
utt74:u: y ahora un la estufa un poco hacia el fregadero	tarea	dir-accion	74		74	
utt75:s: okey	tarea	acepta[74] compromiso	75			74
utt76:s: así está bien ?	tarea admin-tarea	accion-graf preg-sn	76	75 74	76	
utt77:u: así está bien	tarea admin-tarea	acepta[76] respuesta[76]		76		76

Este ejemplo presenta la secuencia de etiquetas más estable en las transacciones del Corpus DIME. Aparece en 40 transacciones que no presentan una sola variación. Por esta razón, el modelo de interacción se definió a partir de estos 40 casos. Las etiquetas que pueden o no aparecer acompañando el conjunto del modelo corresponde a los gestos deícticos, que funcionan como estrategia referencial. Ahora observemos una transacción de tres turnos completada en cinco enunciados:

**5 enunciados - D23) Tran253) - dir-accion**

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO	
			C	A	C	A
utt46:u: igual la campana un poco hacia la derecha	tarea	dir-accion apt-obj apt-tray	46		46	
utt47:s: <no-vocal> okey	tarea	acepta[46] compromiso	47			46
utt48:u: <ruido> ahora quiero ver	Abandonada					
utt49:s: ahí está bien ?	tarea admin-tarea	accion-graf preg-sn	49	46 47	49	
utt50:u: <ruido> sí <ruido>	tarea admin tarea	acepta[49] respuesta[49]		49		49

<sup>19</sup> La anotación de acuerdos implícitos está motivada por la noción de “cargos” y “abonos” del esquema DIME-DAMSL (Capítulo 6). Este modelo representa con las nociones de *Obligaciones* y *Acuerdo mutuo* las relaciones conversacionales entre los enunciados como pares cargo-abono. Comúnmente los cargos en el acuerdo son explícitos (una orden, una oferta, etc.); sin embargo, su abono no necesariamente se expresa de esta manera, sobre todo cuando se aceptan. El esquema representa en estos pares adyacentes (cargo-abono) la coherencia conversacional, permitiendo anotar actos del habla indirectos.

Las seis transacciones de cinco enunciados tuvieron algún abandono o un monólogo. Como se recordará, estas etiquetas indican que los enunciados no hicieron un aporte comunicativo relevante a la conversación. Es decir, como se ve en el ejemplo, un abandono puede no considerarse un turno, puesto que el hablante que abandona lo que está diciendo no está tomando la palabra. Los enunciados anotados en el nivel del *Estado comunicativo* son prescindibles. Por esta razón, podemos afirmar que se mantiene la secuencia de etiquetas presentadas en el modelo de las directivas de acción de tres turnos. Ahora revisemos la única transacción de seis enunciados y tres turnos que se encontró en el corpus:

6 enunciados - D11) Tran100) - dir-accion

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO	
			C	A	C	A
utt105:u: hÍjole	admin-tarea	exclama				
utt106:u: no cabe	admin-tarea	afirm				
utt107:u: <no-vocal> puedes mover el estante lo puedes <sil> bajar puedes poner en el piso pegado al <sil> en esta esquina ?	tarea	dir-accion apt-obj apt-zona	107		107	
utt110:s: okey	tarea	acepta[107] compromiso	108			107
utt111:s: <no-vocal> así está bien ?	tarea admin-tarea	accion-graf preg-sn	111	108 107	111	
utt112:u: sí así está bien	tarea admin-tarea	respuesta[111] acepta[111]		111		111

En un mismo turno, el hablante exclama “¡HÍjole!”, afirma “No cabe” y ordena “¿puedes mover el estante...?”. En este ejemplo se muestra que se sigue manteniendo estable el modelo de interacción de las directivas aunque aparezcan otros actos de las *Funciones hacia adelante* en el mismo turno.

#### *Modelo de tres turnos: preguntas imperativas*

También encontramos 14 casos en que las transacciones de tres turnos comienzan con una pregunta imperativa, siguen con una respuesta y terminan con un acto de acuerdo. La Figura 7.3 presenta la secuencia de etiquetas que se repiten.

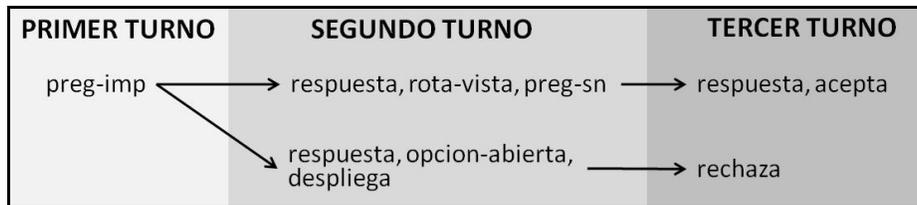


Figura 7.3 Modelo de tres turnos con pregunta imperativa

De estas 14 transacciones, 11 son del tipo “¿Puedes girar el cuarto 90 grados?” y se responden con una acción gráfica [rota-vista] que se acepta en el siguiente turno. Consideremos el siguiente par de ejemplos correspondientes a transacciones de tres turnos de cuatro y cinco enunciados respectivamente.

**4 enunciados - D11) Tran111)**

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO	
			C	A	C	A
utt222:u: puedes girar noventa grados el el el cuarto o la cocina ?	tarea	preg-imp	222		222	
utt223:s: okey	tarea	acepta[222]				222
utt224:s: ahí está bien ?	tarea admin-tarea	rota-vista respuesta[222] preg-sn	224	222	224	
utt225:u: sí está bien	tarea admin-tarea	acepta[224] respuesta[224]		224		224

**5 enunciados - D20) Tran217)**

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO	
			C	A	C	A
utt124:u: <ruido>		Ininteligible				
utt125:u: podemos rotar la vista hacia la izquierda ?	tarea	preg-imp	125		125	
utt126:s: de acuerdo <ruido> <sil> <ruido>	tarea	acepta[125]				125
utt127:s: así está bien ?	tarea admin-tarea	rota-vista respuesta[125] preg-sn	127	125	127	
utt128:u: <ruido> mm sí <ruido>	tarea admin-tarea	acepta[127] respuesta[127]		127		127

Observando estas transacciones parece que los enunciados anotados como [preg-imp] bien pudieron anotarse como [dir-accion]. El resultado sería que estas transacciones seguirían la misma secuencia de etiquetas vista anteriormente en el modelo de las directivas de acción<sup>20</sup>. Así, la diferencia entre el modelo de las

<sup>20</sup> La categoría de *pregunta imperativa* es extraña dentro del conjunto de valores del esquema DIME-DAMSL. En el nombre mismo se refleja un límite confuso y borroso entre si se trata de una pregunta o de una orden. Es decir, por un lado, pretende distinguirse de las

preguntas imperativas y el modelo de las directivas de acción podría afirmarse que es de interpretación.

En cuanto a las preguntas imperativas que siguen otros caminos, encontramos que hubo tres casos en que se rechaza la información proporcionada. En estos tres casos aparecen preguntas imperativas del tipo “*Muéstrame el catálogo de muebles*”, en las tres el usuario rechaza la [opcion-abierta] o el catálogo que se le muestra como respuesta. Con el rechazo la transacción se termina de inmediato. Estos casos reflejan que hubo un cambio de planes en el orden de las acciones a realizarse para completar la tarea, como cuando el “usuario” nota que antes de poner la estufa necesita mover el refrigerador. Consideremos los siguientes ejemplos de tres y cuatro enunciados.

**3 enunciados - D20) Tran210)**

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO	
			C	A	C	A
utt44:u: ahora <ruido> <sil> <sil> podemos ver los hornos ?	tarea	preg-imp	44		44	
utt45:s: contamos con <sil> estas estufas con hor[no]	tarea	acepta[44] respuesta[44] opcion-abierta despliega		44	45	44
utt46:u: mm <ruido> no	tarea	rechaza[45]				45

**4 enunciados - D13) Tran135)**

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO	
			C	A	C	A
utt75:u: me puedes mostrar los extractores ?	tarea	preg-imp	75		75	
utt76:s: okey	tarea	acepta[75]				75
utt77:s: éste y estos otros tres son los modelos de extractores que tenemos	tarea	respuesta[75] opcion-abierta despliega		75	77	
utt78:u: no	tarea	rechaza[77]				77

*Otras transacciones de tres turnos*

Las tres transacciones restantes muestran otros caminos que se pueden tomar durante un diálogo práctico. A pesar de tener una frecuencia mucho menor que lo encontrado en las previas transacciones, también existen estas otras posibilidades para el desarrollo de tareas a través del diálogo en general. Las presentamos a

---

directivas de acción puesto que solicitan información; por el otro, no puede ocultar su similitud con los protocolos imperativos.

continuación porque dejan una idea clara de lo que puede ocurrir cuando no se siguen los modelos previamente revisados.

Por un lado, encontramos que las ofertas también inician transacciones y que igual que las directivas refieren a la realización de acciones gráficas. Si una oferta se acepta se ejecuta una acción, si una oferta se rechaza no pasa nada. Consideremos los dos casos de ofertas encontrados en las transacciones de tres turnos:

**3 enunciados - D02) Tran18) utts: 170-172**

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO	
			C	A	C	A
utt170:s: quieres que lo borre ?	tarea	oferta			170	
utt171:u: por favor	tarea	acepta[170]	171			170
utt172:s: ahí está bien ?	tarea	Quita preg-sn		171	172	

En este caso, el primer turno hace una oferta que se acepta en el segundo, en el tercero se quita el objeto y se hace una pregunta de verificación que no obtiene la respuesta esperada. En este particular caso, a la pregunta “¿Así está bien?” se le contesta “y una opción <sil> un gabinete más pequeño sería”<sup>21</sup>. Como ya se ha comentado, desde la perspectiva de análisis del esquema DIME-DAMSL, en este tipo de casos el acuerdo es implícito; el hablante no manifestó desacuerdo con la acción realizada y demuestra más interés en continuar con la tarea, por esta razón la transacción se da por terminada con sólo tres turnos. Ahora consideremos una oferta de tres turnos que se rechaza:

**4 enunciados - D02) Tran20)**

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO	
			C	A	C	A
utt192:s: quieres poner otro mueble en la cocina ?	admin_tarea	oferta			192	
utt193:u: no creo que así está bien	admin_tarea	rechaza[192] afirm			193	192
utt195:s: okey la tarea está terminada	admin_tarea	acepta[193] afirm cierra				193
utt196:s: gracias	admin_tarea	performativo				

<sup>21</sup> Decimos que es un caso particular porque las preguntas de confirmación suelen contestarse explícitamente en el contexto de los diálogos prácticos. Sin embargo, como puede verse en este caso, no siempre sucede así.

En este caso vemos una transacción que termina un diálogo. De hecho, las ofertas que son rechazadas se dan comúnmente en el contexto del término de la tarea. Es interesante observar que cuando las transacciones no ejecutan una acción gráfica por lo general están dirigidas a expresar y acordar que la tarea está terminada.

Por último, consideremos la única transacción de tres turnos que comienza con una pregunta cerrada, en esta transacción tampoco se ejecuta una acción gráfica y el diálogo se dirige a su término.

**4 enunciados - D19) Tran205)**

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO	
			C	A	C	A
utt101:u: otra sugerencia ?	admin_tarea	preg-sn	101		101	
utt102:s: eh no está bien	admin_tarea	rechaza[101] respuesta[101]		101		101
utt104:s: gracias	admin_tarea	cierra performativo			104	
utt105:u: de nada	admin_tarea	performativo				104

Hasta ahora, hemos visto que con tres turnos se resolvieron 65 transacciones que completan el 23% del corpus. El análisis de estas transacciones cortas ha resultado en la definición de dos estables modelos de interacción que se parecen mucho entre sí (el de las directivas y el de las preguntas imperativas). Consideramos que la definición de este par de modelos de tres turnos representa de forma sintética la esquematización de conductas conversacionales relevantes en el dominio de los diálogos prácticos. La alta frecuencia de estos modelos pone en relieve que en este tipo de conversaciones se privilegian los intercambios cortos para resolver partes de la tarea de manera rápida y eficiente. Adicionalmente, de este análisis nos queda también una idea de lo que puede ocurrir en las transacciones en que no se repiten estos modelos. A continuación revisaremos qué ocurre en las transacciones que necesitan más turnos conversacionales para completarse.

*Transacciones de más de tres turnos conversacionales*

Si bien hasta ahora el análisis de las transacciones ha intentado ser exhaustivo, lo cierto es que sólo hemos cubierto el 23% del corpus. ¿Qué pasa en ese otro 67% de las transacciones en que hay más turnos conversacionales? Lo que hemos

observado es que ocurren primordialmente dos fenómenos: por un lado, encontramos que con cierta regularidad aparecen las estrategias de los modelos anteriormente revisados a los que se van sumando protocolos conversacionales que cumplen funciones comunicativas específicas. Por otro lado, encontramos caminos distintos, otros tipos de cosas que también pasan en los diálogos colaborativos. Este ha sido un arduo ejercicio en el que se ha pretendido la búsqueda de conductas esquemáticas, lo cierto es que el estudio de la lengua hablada siempre sorprende por la variedad de estrategias que tiene disponibles para lograr objetivos similares.

#### *Protocolos de interacción*

Para cerrar el análisis que se hace en este trabajo presentaremos los protocolos conversacionales más relevantes que se encontraron en el primer caso. En todos ellos hay al menos una alternancia de turno, es decir, en cada protocolo habla por lo menos una vez cada uno de los participantes del diálogo. Un par de turnos es suficiente para preguntar y responder, para solicitar y corregir algún detalle de la acción realizada o para recuperar la conversación de un malentendido. Esto quiere decir que, tomando como base los modelos de tres turnos, es posible integrar ciertos protocolos de “ida y vuelta” que conformen modelos más flexibles. La flexibilidad del modelo permite contemplar un repertorio más amplio de estrategias conversacionales comunes en el desarrollo de diálogos prácticos. A continuación presentamos los protocolos más frecuentes y relevantes con ejemplos de cada uno. Los ejemplos muestran en “negritas” los enunciados que componen el protocolo que se explica en cada caso. Distinguimos dos tipos principales:

*Protocolos de entendimiento:* los enunciados que componen este par de turnos están vinculados al monitoreo del canal de comunicación, del acuerdo o de la comprensión del diálogo.

- Señal de no entendimiento: durante una conversación pueden aparecer enunciados que manifiestan no haber comprendido el enunciado previo en cualquier momento (“¿Perdón?”, “¿Lo puedes repetir?”, “Disculpa, no te entendí”, etc.). Cuando aparece una de estas señales se echa a andar un

protocolo de recuperación que conlleva al menos un par de turnos. En el primero el hablante manifiesta no haber comprendido el enunciado anterior; en el siguiente turno el interlocutor atiende la solicitud repitiendo, parafraseando o algunas veces abandonando lo dicho previamente. La anotación DIME-DAMSL indica que en estos casos los enunciados que pertenecen a este protocolo reciben las etiquetas de [admin-com] en el *Nivel de información* y se anotan cargos y abonos en el plano del *Entendimiento*.

**SEÑAL DE NO ENTENDIMIENTO**  
**D13) Tran130) - fragmento**

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO MUTUO			
			C	A	ACUERDO		ENTENDIMIENTO	
					C	A	C	A
utt32:u: la mesa blanca <sil> la puedes poner aquí	tarea	dir-accion apt-ob apt-zona	32		32			
utt33:s: perdón	admin-com	sne [32] preg-pron	33				33	
utt34:u: me puedes poner esta mesa blanca aquí	admin-com	respuesta [33] reafirm apt-obj apt-zona		33	34			33
utt35:s: okey	tarea	acepta [34;32] compromiso	35			34 32		

- Confirmación de entendimiento: en el contexto de los diálogos colaborativos es común encontrar señales que confirman que se ha comprendido el mensaje del interlocutor cabalmente o que están de acuerdo con la información vertida previamente. Encontramos que se agregan turnos de dos maneras distintas:
  - Traslapes, Acuerdos y Asentimientos: Para indicar que se entiende el mensaje anterior los participantes de un diálogo comúnmente dicen “Ajá”, “Sí”, “Okey”, etc. Estos enunciados se expresan sin tomar propiamente el turno. La intención es dejar hablar al interlocutor dejándole saber que se le está prestando atención. La etiquetación DIME-DAMSL anota estos enunciados como [acepta] cuando se trata de información que explícitamente se añade al bagaje informativo; los anota como [traslape] cuando se expresan mientras el interlocutor habla; y los anota como asentimiento, [ack], cuando sin aceptar propiamente lo dicho por el interlocutor, manifiestan haberlo escuchado. Cualquiera de los tres implica

un protocolo similar en el que se agrega un par de turnos; en el primero, el hablante expresa a través de alguna de estas posibilidades que entendió o escuchó el mensaje, en el segundo turno, su interlocutor continua hablando. El primer turno contiene un enunciado con cualquiera de estas etiquetas ([accepta], [traslape] o [ack]), integradas en los niveles de *Acuerdo* o de *Entendimiento* de las *Funciones hacia atrás*.

**TRASLAPES, ACUERDOS y ASENTIMIENTOS**  
D02) Tran007) - fragmento

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO MUTUO			
					ACUERDO		ENTENDIMIENTO	
			C	A	C	A	C	A
utt15:s: te voy a explicar	admin-tarea	afirm			15			
utt16:s: ésta es la ventana	admin-tarea	afirm apt-zona			16			
utt17:u: sí	admin-tarea	accepta[16;15]				16 15		
utt18:s: ésta que tienes aquí	admin-tarea	reafirm apt-objs-coord			18			
utt19:u: okey	admin-tarea	accepta[18]				18		
utt20:s: + + o sea esta esquina corresponde a esta esquina	admin-tarea	afirm apt-objs-coord					20	
utt21:u: + ajá +	admin-com	traslape[20]						20
utt22:s: y ésta es la puerta que tienes acá	admin-tarea	afirm apt-objs-coord			22			

- Preguntas de confirmación: Por otro lado, encontramos que se hacen preguntas que tienen como función verificar que se ha entendido correctamente. Asegurar que se ha comprendido cabalmente toma al menos dos turnos; en el primero se hace una pregunta cerrada que, sin aceptar lo dicho anteriormente por el interlocutor, lo parafrasea o lo repite (“¿Me estás pidiendo que quite el refrigerador?”, “¿Quieres que ponga este objeto en este lugar?”, “¿Este?”, etc.), en el segundo turno se responde. Si la respuesta es afirmativa la conversación sigue su curso como si no se hubiera preguntado; si es negativa, en el mismo turno en que se responde, se corrige lo que no se entendió bien. Siguiendo la metodología del esquema DIME-DAMSL, los enunciados de este protocolo conforman una relación de cargo

y abono en el nivel del *Entendimiento*. La pregunta de confirmación recibe también las etiquetas de [espera] y [repeticion].

**CONFIRMACIÓN DE ENTENDIMIENTO  
D01) Tran004) - fragmento**

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO MUTUO			
					ACUERDO		ENTENDIMIENTO	
			C	A	C	A	C	A
utt68:u: + <vocal> + bueno el gabinete tengo que moverlo a <sil> al al lado de lo que es el extractor <au> <sil> y encima de lo que es la al[acena] lo que es la alacena	tarea	dir-accion apt-objs-coor	68		68			
utt69:s: quieres que mueva <sil> este gabinete <sil> acá ?	admin-tarea	espera[68] preg-sn repeticion[68] apt-obj apt-zona	69				69	
utt70:u: sí	admin-tarea	respuesta[69] acepta[69]		69				69
utt71:s: okey	tarea	acepta[68] compromiso	71			68		

*Protocolos de la tarea:* están dirigidos a la planificación y ejecución de acciones vinculadas a la solución de la tarea. Con relación a la estructura de dos fases que se propone en la definición de “transacción”, encontramos que ciertos protocolos aparecen en algunos momentos de la transacción y no en otros. Se recordará que la transacción es propuesta por DIME-DAMSL como una unidad de análisis conversacional compuesta por dos fases: *Especificación de la intención* y *Satisfacción de la intención* (Pineda *et al.*, 2007 y 2006). En la primera fase se manifiesta la intención de realizar una acción y se especifican los parámetros que la componen (qué hacer con qué). En la segunda fase se ejecuta la acción, si fue satisfactoria la transacción termina, si no, se realizan los ajustes pertinentes. Por esta razón dividimos los protocolos primordialmente en dos tipos:

*Protocolos de especificación de la intención:*

- Información: es común encontrar que se solicita información sobre la tarea o los elementos que la conforman antes de completar una acción. Estos protocolos informativos toman al menos dos turnos, los necesarios para preguntar y responder. A diferencia del caso de las preguntas de confirmación

que están dirigidas a verificar el entendimiento, estas preguntas solicitan que se provea información que se desconoce genuinamente y que se necesita para continuar con la tarea. Dentro de estos protocolos informativos se encuentran las preguntas pronominales “¿Qué es esto?”, “¿Qué diferencia hay entre éste y éste?” o “¿Dónde quieres que lo ponga?”; también se encuentran las preguntas cerradas “¿Éste es más grande que éste?”, “¿Tienes más sillas?”, “¿Alguna otra sugerencia?” y también las preguntas imperativas del tipo “¿Puedes mostrarme los muebles?”, “Muéstrame las estufas”, “Ahora quiero ver los refrigeradores”. Es claro que cada tipo de pregunta busca cierto tipo de información y obtendrá una respuesta distinta. Esto quiere decir que las preguntas pronominales tienen como respuesta una afirmación, las preguntas cerradas un acuerdo y las preguntas imperativas una acción gráfica en la que se despliega el menú de opciones. Estos protocolos pueden integrarse a los modelos de tres turnos permitiendo que se contemplen las necesidades informativas que surgen antes de realizar una acción.

**INFORMACIÓN - Pregunta pronominal**  
**D25) Tran262) - fragmento**

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO	
			C	A	C	A
utt82:u: éste <sil> este mueble qué es ?	tarea	preg-pro apt-obj	82			
utt83:s: ése es una alacena inferior	tarea	respuesta[82] afirm		82	83	
utt84:u: ah la alacena inferior ponlo debajo del gabinete	tarea	acepta[83] dir-accion apt-objs-coor	84			

**INFORMACIÓN - Pregunta cerrada**  
**D11) Tran089) - fragmento**

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO	
			C	A	C	A
utt14:s: <no-vocal> éstas son las estufas con que contamos	tarea	respuesta[7] opcion-abierta despliega		7	14	
utt15:u: tienes alguna que sea <sil> no sé no no tan moderna rústica <vocal>	tarea	espera[14] preg-sn	15		15	
utt16:s: no	tarea	respuesta[15] rechaza[15]		15		15

utt17:s: solamente tengo estas <sil> estos cinco modelos de estufas	tarea	afirm despliega			17	
utt18:u: <no-vocal> okey está bien	tarea	acepta[17]				17
utt20:u: puedes llevar <sil> mh <sil> esta <sil> estufa al <sil> al al centro de del cuarto ?	tarea	acepta[14] dir-accion apt-obj apt-zona	20		20	14

**INFORMACIÓN - Imperativa**  
**D05) Tran035) - fragmento**

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO	
			C	A	C	A
utt55:u: podría mostrarme el catálogo <sil> de este tipo de muebles ?	tarea	preg-imp apt-obj	55		55	
utt56:s: <no-vocal> <sil> okey	tarea	acepta[55]				55
utt57:s: éste es el catálogo de alacenas <sil> ah <sil> que- que tenemos hay de diferentes colores y tamaños tenemos blancas <sil> de madera <sil> rojas con negro <sil> y verdes <no-vocal> hay superiores e inferiores	tarea	respuesta[55] opcion-abiera despliega		55	57	
utt58:u: podrías <sil> tomar ésta y colocarla aquí ?	tarea	acepta[57] dir-accion apt-obj apt-zona	58		58	57

- Ofertas: la aparición de ofertas conlleva que se agregue un par de turnos conversacionales a las transacciones. En el primero el hablante se ofrece a realizar alguna acción (“¿Quieres que desplace o traiga algún objeto a la cocina?” o “¿Quieres que borre este objeto?”) y en el segundo el interlocutor lo acepta. Por convención, los diálogos del corpus DIME empiezan con una oferta que comúnmente se acepta junto con una directiva de acción; sin embargo, como puede verse en el siguiente ejemplo, la aparición de ofertas no es exclusiva al momento en que inicia el diálogo.

**OFERTAS**

**D01) Tran006) - fragmento**

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO	
			C	A	C	A
utt85:u: pero ahora me hace falta un <sil> me parece que es un lavatrastos <sil> o un	admin-tarea	afirm			85	
<b>utt86:s: quieres que te muestre los lavatrastes ?</b>	admin-tarea	acepta[85] oferta			86	85
<b>utt87:u: eh por favor</b>	admin-tarea	acepta[86]	87			86
utt88:s: éstos son lo fregaderos que tenemos	tarea	opcion-abierta despliega		87	88	

*Protocolos de satisfacción de la intención:*

- Reajuste de acciones: cuando se realiza una acción gráfica el resultado está sujeto a la aprobación de interlocutor de un diálogo práctico. Si el resultado es afortunado la transacción termina; si no lo es, los hablantes ajustan los detalles. Por lo general, la pregunta de verificación “¿Así está bien?” que suele concluir transacciones en este caso no recibe un [acepta] sino una orden “¿Puedes ponerlo más arriba?”, “Un poco más abajo”, “Un poco más a la izquierda, por favor”. Como podemos observar en el siguiente ejemplo, el protocolo de reajuste puede repetirse una y otra vez hasta que el resultado de la acción se considere satisfactorio. La metodología de anotación del esquema considera equivalente a “tener” el turno al momento en que se realizan acciones gráficas. Aunque no se pronuncien palabras mientras se hacen cosas, la atención de la conversación está enfocada a las acciones y el curso del diálogo está sujeto a los resultados que se obtengan de ellas.

**REFINAMIENTO DE ACCIONES**

**8 enunciados - D17) Tran169)**

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO	
			C	A	C	A
utt70:u: vamos a hacer la barra de desayuno más para acá aquí más o menos	tarea	dir-accion apt-obj apt-tray	70		70	
utt71:s: okey	tarea	acepta[70] compromiso	71			70
utt72:s: acción gráfica	tarea	mueve objeto		71 70	72	
<b>utt73:u: más arriba</b>	<b>tarea</b>	<b>acepta-parte[72] dir-accion</b>	<b>73</b>		<b>73</b>	<b>72</b>
<b>utt74:s: acción gráfica</b>	<b>tarea</b>	<b>acepta[72] mueve objeto</b>		<b>73</b>	<b>74</b>	<b>73</b>
<b>utt75:u: más arriba</b>	<b>tarea</b>	<b>acepta-parte[74] dir-accion</b>	<b>75</b>		<b>75</b>	<b>74</b>

utt76:s: + así está bien ? +	tarea admin-tarea	acepta[75] mueve objeto preg-sn	76	75	76	75
utt77:u: + okey +	tarea admin-tarea	acepta[74] respuesta[74]		76		76

- Fin de diálogo: cuando alguno de los hablantes considera que la tarea está terminada el diálogo se perfila a su término. Por lo general, terminar el diálogo implica al menos dos turnos, en el primero el hablante manifiesta que se acabó mediante una [afirmación] (“*Ya terminé*” o “*La tarea está terminada*”) y en el segundo su interlocutor acuerda el término. Es común encontrar que posteriormente vienen agradecimientos, el esquema DIME-DAMSL indica que se anotan como [performativo] en las *Funciones hacia adelante* y como [otro] en el *Nivel de información*.

**FIN DE DIÁLOGO**  
**D06) Tran079) - fragmento**

ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO	
			C	A	C	A
utt364:s: <no-vocal> así está bien ? + +	tarea admin-tarea	mueve objeto preg-sn	364	363 362	364	
utt365:u: + sí +	admin-com	traslape[364]				
utt366:u: sí está bien	tarea admin-tarea	acepta[364] respuesta[364]		364		364
utt368:s: okey	admin-com	ack[366]				
utt369:s: la tarea está terminada	admin-tarea	cierra afirm			369	
utt370:s: gracias	otro	performativo				
utt371:u: de nada	otro	performativo				369

*Inclusión de protocolos de interacción a los modelos de tres turnos*

Los protocolos que hemos presentado corresponden a pares de turnos que pueden o no aparecer en las transacciones. Su presencia dependerá de las necesidades comunicativas que vayan surgiendo para el cumplimiento de los objetivos parciales. En una misma transacción puede aparecer uno o más de estos protocolos, no son excluyentes entre sí e, incluso, pueden aparecer protocolos de entendimiento embebidos entre los turnos de un protocolo sobre la tarea. Exceptuando los protocolos de entendimiento que pueden aparecer indistintamente en cualquier parte del diálogo, los protocolos sobre la tarea están sujetos al momento de la transacción en que se encuentren los interlocutores. Es decir, a pesar de que existe bastante

flexibilidad en su aparición dentro de una transacción, lo cierto es que si aparecen no es de forma azarosa

A continuación presentamos dos transacciones del Corpus DIME; creemos que son ideales para ejemplificar la forma en que estos protocolos operan en el contexto de los diálogos prácticos. Es nuestro último par de ejemplos y con él pretendemos concluir el análisis de las transacciones con el esquema DIME-DAMSL. La intención es recalcar la flexibilidad con que estos protocolos se pueden insertar en los modelos. Del lado izquierdo de las tablas indicamos con corchetes los protocolos y señalamos con “negritas” los modelos a los que se integran. El primer ejemplo corresponde a una transacción en el que está presente el modelo de las directivas de acción y el segundo ejemplo presenta el modelo de las preguntas imperativas.

17 enunciados - 12 turnos- D11) Tran089)

	ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO MUTUO			
				C	A	ACUERDO		ENTENDIMIENTO	
						C	A	C	A
oferta	utt1:s: quieres que traiga algún mueble a la cocina ?	tarea	abre oferta			1			
	utt2:u: <no-vocal> sí	tarea	acepta[1]	1		1			
informativo (preg-imp)	utt3:u: de hecho me gustaría algo así medio rústico no sé qué tipo de cocinas <sil> haya en el <sil> menú aquí	tarea	preg-imp	3		3			
	utt4:s: perdón no entiendo	admin-com	sne[3] preg-pron	4				4	
	utt5:u: sí qué tipo de cocinas tienen ? me puedes mostrar las co[cinas] el tipo de cocinas que hay + la estuf[a] la estufa me refiero a +	tarea	respuesta[4] preg-imp	5	4	5			4
	utt6:s: + eh <sil> + quieres que te muestre el catálogo de estufas ? + +	admin-tarea	espera[5] repeticion[5] preg-sn	6					6
	utt7:u: + sí +	admin-tarea	acepta[6] resp[6]		6				6
	utt8:s: okey	tarea	acepta[3;5]				5 3		
	utt9:s: <no-vocal> éstas son las estufas con que contamos	tarea	respuesta[3;5] opcion-abierta despliega		5 3	9			
	utt10:u: tienes alguna que sea <sil> no sé no no tan moderna rústica <vocal>	tarea	espera[9] preg-sn	10		10			
	utt11:s: no	tarea	rechaza[15] resp[15]		10	10			
	utt12:s: solamente tengo estas <sil> estos cinco modelos de estufas	tarea	reafirm despliega			12			
utt13:u: <no-vocal> okey está bien	tarea	acepta[17]				12			
utt14:u: puedes llevar <sil> mh <sil> esta <sil> estufa al <sil> al al centro de del cuarto ?	tarea	acepta[14] dir-accion apt-obj apt-zona	14		14	9			
utt15:s: okey	tarea	acepta[20] compromiso	15			14			
utt16:s: ahí está bien ?	tarea admin-tarea	agrega objeto preg-sn	16	15 14 1	16				
utt17:u: mm sí está bien	tarea admin-tarea	acepta[22] resp[22]		16		16			

En este ejemplo podemos ver que la fase de *Especificación de la intención* requirió la inclusión de cinco protocolos, dos de ellos referentes al entendimiento y tres vinculados al desarrollo de la tarea. El primero es una oferta que convencionalmente inicia el diálogo. El segundo es un protocolo informativo en el que se solicita ver el catálogo. Dentro de este protocolo informativo aparecen los dos protocolos de entendimiento, el primero señala que no se comprendió el mensaje, en el segundo se busca confirmar que se entendió el mensaje. El último protocolo es informativo, el hablante necesita saber si la [opcion-abierta] que se le hizo incluye todas las opciones antes de decidirse por una. Como puede observarse, los protocolos se integraron antes de la serie de etiquetas presentadas en el modelo de las directivas de acción. Por último, consideremos el ejemplo de las preguntas imperativas.

**11 enunciados - 8 turnos - D11) Tran118)**

	ENUNCIADO	NIVEL DE INFORMACIÓN	FUNCIONES DIME-DAMSL	OBLIGACIONES		ACUERDO MUTUO			
				C	A	ACUERDO		ENTENDIMIENTO	
						C	A	C	A
	utt273:u: puedes girar el cuarto ? <sil> 1[a] la cocina ?	tarea	preg-imp	273		273			
confirmación	utt274:s: quieres que gire el cuarto noventa grados ?	admin-tarea	espera[273] preg-sn					274	
	utt275:u: sí	admin-tarea	acepta[274] respuesta[274]						274
	utt276:s: así está bien ?	tarea admin-tarea	acepta[273] rota vista respuesta[273] preg-sn	276	273	276	273		
reajuste	utt277:u: hacia en el otro sentido	tarea admin-tarea	rechaza[276] respuesta[276] preg-imp	277	276	277	276		
	utt278:s: okey	tarea	acepta[277]				277		
	utt279:s: así está bien ?	tarea admin-tarea	rota vista respuesta[277] preg-sn	279	277	279			
	utt280:u: sí así está bien	tarea admin-tarea	acepta[279] respuesta[279]		279		279		
fin de diálogo	utt281:u: es todo ya ya terminé	admin-tarea	cierra afirm			281			
	utt282:s: okey	admin-tarea	acepta[82]				281		
	utt283:s: gracias	admin-tarea	performativo						

En este caso podemos observar que se integran tres protocolos, dos de ellos están embebidos en el modelo de las preguntas imperativas; el primero es una pregunta de confirmación que se hace en la fase de *Especificación de la intención*. El segundo es un reajuste que se realiza en la fase de *Satisfacción de la intención*. Después de que se acepta el resultado, se integra un protocolo de fin de diálogo con el que se da por terminada la tarea.

Para poder integrar estos protocolos de interacción a los modelos de las directivas de acción y las preguntas imperativas es necesario que se especifiquen en los mismos términos en que originalmente fueron definidos aquellos modelos. Al igual que los modelos a los que se integran, la definición de estos protocolos se balancea entre descriptiva y prescriptiva. Por un lado, describe las secuencias conversacionales más observadas en el Corpus DIME; por el otro, intenta prescribir una serie de expectativas que surgen a partir de que se enuncian ciertas intenciones comunicativas. Se recordará que la descripción de estos modelos pretende en un futuro implementarse en el desarrollo de sistemas conversacionales.

A continuación presentamos un par de imágenes, Figura 7.4 y Figura 7.5 que representan de forma sintética y abstracta los *Protocolos de entendimiento* y de *la tarea* previamente ejemplificados. Los protocolos se listan de lado derecho seguidos de las etiquetas que caracterizan cada caso. Las flechas que dirigen a unos puntos suspensivos indican que la transacción continúa su curso en alguna parte del modelo.

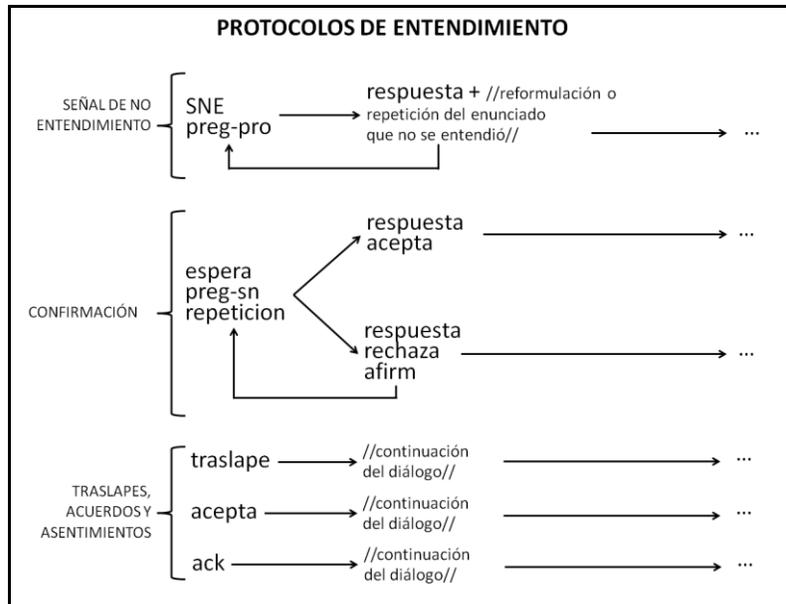


Figura 7.4 Especificación de 'Protocolos de entendimiento'

Como ya se dijo, los *Protocolos de entendimiento* tienen como primordial motivo mantener en buen estado el canal de comunicación y la buena comprensión

de los mensajes y sus intenciones. Esto quiere decir que la definición de los modelos de interacción, tanto para las directivas como para las preguntas imperativas, considera que es probable que se activen los protocolos de entendimiento en cualquier momento del diálogo.

De manera distinta, los *Protocolos de la tarea* están fuertemente vinculados al momento de la transacción en que se encuentran los interlocutores para que aparezcan (es decir, no es lo mismo si ya se realizó una acción o si está por realizarse). La representación gráfica de los modelos y los protocolos de forma integral requiere un nivel de detalle en la abstracción y generalización de las conductas que en este trabajo no podemos alcanzar. Creemos que el trabajo completado hasta ahora traza ciertas líneas de observación que a futuro pueden llevar a la implementación experimental de modelos conversacionales (basados en el análisis de actos de habla) en sistemas computacionales.

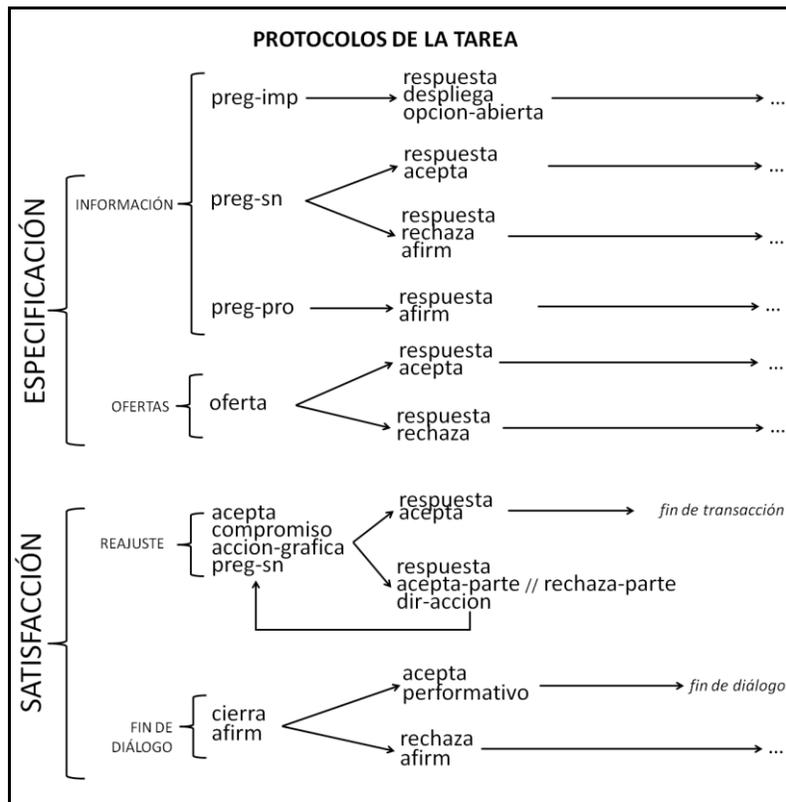


Figura 7.5 Especificación de 'Protocolos de la tarea'

Comenzamos el capítulo presentando los resultados obtenidos en la anotación de DIME-DAMSL en cada uno de los enunciados. Este punto de partida es un nivel de análisis conversacional local y con él pudimos darnos una idea de qué funciones o intenciones comunicativas fueron más frecuentes y relevantes. Posteriormente, revisamos cómo es que los enunciados conforman unidades conversacionales mayores llamadas transacciones. Asimismo, a partir de la anotación y revisando el tipo de interacción encontrada en las transacciones pudimos encontrar y definir un par de modelos de interacción y un conjunto de protocolos que permiten especificar esquemas conversacionales que cubren diversos fenómenos observados en los diálogos prácticos del Corpus DIME.



## 8. Conclusiones

### *Sobre el Corpus DIME*

El grupo de investigación al que pertenece este trabajo considera que el análisis lingüístico debe hacerse directamente sobre una muestra de lenguaje; es decir, su metodología está emparentada con la lingüística de corpus. La muestra de lenguaje que se elija para una investigación está directamente vinculada a los resultados que se obtengan al estudiarla. El control de las variables es responsabilidad del grupo de investigación. En el caso del Corpus DIME, se diseñó un experimento especial que tuvo impacto directo en la forma que tomaron los diálogos y las estrategias que surgieron durante su ejecución. Se recordará que ninguno de los participantes se ve el uno al otro pero los dos ven lo mismo en la pantalla. En este sentido podemos encontrar pros y contras en la utilización de un recurso lingüístico con las características del Corpus DIME para la definición de modelos de interacción en un sistema de diálogo. Los resumimos en la siguiente tabla:

Tabla 8.1 *Pros y contras del Corpus DIME*

PROS	CONTRAS
<ul style="list-style-type: none"><li>• corresponde a una muestra única del español de la Ciudad de México en uso conversacional colaborativo</li><li>• el recurso cuenta con archivos de texto, audio y video de cada enunciado y de cada diálogo posibilitando el análisis desde diferentes perspectivas</li><li>• el tipo de interacción que se desarrolla permite la investigación de las correspondencias entre la información visual y la información lingüística</li><li>• está disponible para la comunidad tanto el corpus como los análisis que se han realizado con él: fonéticos, fonológicos, morfológicos, sintácticos, prosódicos, pragmáticos, etc.</li></ul>	<ul style="list-style-type: none"><li>• los audios son de baja calidad</li><li>• la interacción de los diálogos no se pudo controlar con la rigurosidad especificada en su definición</li><li>• el control que sí se logró tiene impacto directo en el tipo de interacción que surgió</li><li>• la metodología de segmentación de enunciados no fue sistemática, algunos marcaron dos enunciados donde otros veían uno y viceversa</li></ul>

Creemos que la evaluación del corpus resulta positiva. Si bien puede reconocerse que para futuras investigaciones es probable que el Grupo Golem requiera el diseño y recopilación de un nuevo corpus que se adecue de manera más cercana a los intereses de investigación y que cumpla sistemáticamente con las especificaciones

de control y diseño; lo cierto es que el recurso ha sido un terreno válido para la investigación de diversos fenómenos lingüísticos.

#### *Sobre el esquema DIME-DAMSL*

De la misma manera que el recurso lingüístico utilizado para una investigación tiene impacto directo en los resultados, lo mismo ocurre con la metodología de análisis. El esquema DIME-DAMSL surgió de la necesidad de capturar las funciones comunicativas de los enunciados y su coherencia conversacional independientemente de la forma sintáctica que manifiestan. El esquema se fue refinando y ajustando al mismo tiempo que se realizaba el análisis del Corpus DIME, de tal suerte que los nuevos fenómenos que se iban encontrando pudieran ser cubiertos a cabalidad. El análisis con el esquema se realizó en un momento posterior a la transcripción y segmentación del corpus en enunciados; esto conllevó notar que las especificaciones de segmentación no correspondían del todo a los intereses del esquema DIME-DAMSL (por ejemplo, la anotación de acciones gráficas tuvo que hacerse casi siempre en el mismo enunciado en que se realizaba la pregunta de confirmación (“¿Así está bien?”) puesto que los segmentadores no marcaron una diferencia enunciativa). En la medida de lo posible, la definición del esquema intentó socavar estas y otras diferencias surgidas entre las metodologías de segmentación, transcripción y anotación del corpus.

La anotación fue realizada a través de varios ejercicios colectivos. Conforme iba avanzando se evaluaba el nivel de acuerdo entre los anotadores con el coeficiente Kappa. El resultado del acuerdo entre los anotadores en los diferentes niveles del esquema fue superior a 0.8, que es considerado afortunado muy por encima del azar (Pineda *et al.*, 2006). Esto quiere decir que la definición de la metodología quedó hasta cierto punto clara para los anotadores y que la asignación de etiquetas mantiene cierta sistematicidad.

Las unidades comprendidas en el esquema, si bien son limitadas, presentan un mosaico que pretende cubrir diversas funciones comunicativas observadas en los diálogos colaborativos. Podemos afirmar que exceptuando algunos casos muy particulares, las etiquetas del esquema fueron suficientes para dar cuenta de los

actos del habla expresados en el corpus. Es indudable que hace falta refinar los rasgos propios de algunas categorías; por ejemplo, los llamados “performativos” son en realidad agradecimientos, las exclamaciones se consideran dentro de las *Funciones hacia adelante* pero en sí mismas no incitan “hacia adelante” ninguna acción directa de su interlocutor, la etiqueta de [traslape] marca más una característica de enunciación que una función comunicativa en sí misma, las preguntas imperativas se evidencian como una categoría postulada por el Proyecto DIME para anotar los enunciados que inician protocolos similares a los de las directivas pero que buscan información, etc.

Existen múltiples aspectos de cada acto del habla contemplado y las relaciones que establecen entre sí que pueden ser refinados en un futuro trabajo con mayor detalle. Si bien los modelos encontrados en este experimento son estables, reconocemos que se deben en parte a las restricciones del propio experimento. La utilización de este método de análisis crea más preguntas que respuestas sobre el funcionamiento de la conversación general. Pero creemos que eso motiva la curiosidad lingüística y el quehacer del investigador. Como punto de partida el esquema DIME-DAMSL establece un camino propicio para la investigación de las conductas conversacionales en dominios específicos. De forma paralela, se presenta como un modelo de interacción humano-máquina basado en una estructura dialógica.

#### *Trabajo futuro*

Los modelos resultantes de este trabajo intentan capturar y recalcar la naturaleza ambivalente de la lengua, que por un lado es dinámica y cambia y se adapta a las necesidades de cada situación comunicativa, y que por otro es rutinaria y crea hábitos y consolida patrones de conducta que están asociados a situaciones comunicativas específicas. Estos patrones o regularidades pueden encontrarse en cualquier nivel del lenguaje y emergen del mismo lugar: la lengua en uso. Cada lengua tiene su norma fonética, morfológica, sintáctica, semántica, etc. y cada lengua rompe su normas para significar algo concreto de acuerdo al momento de enunciación.

En este sentido, el uso de un esquema como DIME-DAMSL para la definición de modelos de interacción que se integren en un sistema conversacional busca representar esta característica doble del lenguaje. Por un lado, los modelos y los protocolos marcan pautas fijas de interacción asociadas a momentos enunciativos concretos (estructura de la transacción). Por el otro, son descritos con cierta flexibilidad, de tal manera que el sistema pueda administrar los modelos integrando los protocolos necesarios cuando las necesidades del diálogo así lo indiquen. En el marco de su aplicación en sistemas conversacionales, este tipo de modelos permitirán al sistema saber qué hacer cuando se le presenten estas conductas al interactuar con humanos (como cuando falta un referente o se avisa de una falla en el entendimiento o se pide realizar una acción). Entre mayor sea el repertorio de estrategias conversacionales que incluya un sistema conversacional, mayor será la capacidad del sistema de lidiar con situaciones conversacionales distintas. Idealmente se espera que un sistema conversacional tenga acceso a un amplio repertorio de recursos conversacionales y la facilidad para administrar unos y otros; esto es, que tenga la capacidad de adaptar su conducta conversacional de acuerdo al caso. Esta interacción flexible y con capacidades dinámicas debe basarse en una definición puntual de los recursos de interacción que tiene disponibles.

Dentro del trabajo futuro que se vislumbra siguiendo estas líneas de investigación está la validación de los modelos obtenidos en un sistema conversacional. Para ello, será necesario un trabajo junto a expertos en computación que integren los modelos presentados a los procesos de administración del diálogo por parte del sistema.

## 9. Bibliografía

- ALEXANDERSSON, J., Buschbeck-Wolf, B., Fujinami, T., *et al.* (1997) *Dialogue acts in VERBMOBIL-2*. Verbmobil Report 204. Alemania: DFKI University of Saarbruecken.
- ALLEN, James, Donna Byron, Myroslava Dzikovska, George Ferguson, Lucian Galescu and Amanda Stent. (2001) "Toward Conversational Human-Computer Interaction". En *AI Magazine*. Vol. 22, No. 4. pp 27-37. Disponible en línea, mayo de 2013:  
<http://www.aaai.org/ojs/index.php/aimagazine/article/view/1590/1489>
- \_\_\_\_\_. (1998) "An Architecture for a Generic Dialogue Shell". En *Natural Language Engineering* 1 (1). pp. 1-16. Inglaterra: Cambridge Journals, Cambridge University Press. Disponible en línea, mayo de 2013:  
<http://www.cs.rochester.edu/research/cisd/pubs/2000/allen-et-al-nle2000.pdf>
- ALLEN, James y Mark Core. (1997) *Draft of DAMSL: Dialogue Act Markup in Several Layers*. Estados Unidos: Rochester University. Disponible en línea, mayo de 2013:  
<http://www.cs.rochester.edu/research/speech/damsl/RevisedManual/>
- AUSTIN, John. (1962) *¿Cómo hacer cosas con palabras?* Compilación: J. O. Urmson. Traducción: Genaro R. Carrió y Eduardo A. Rabossi. Barcelona: Paidós.
- CEBALLOS, Ana. (2007) *Fenómenos de pérdida en el Corpus DIME para su inclusión en un reconocedor de habla*. Tesis de licenciatura. México: Facultad de Filosofía y Letras, UNAM.
- CLARK, Herbert. (1996) "Language use". En *Using Language*. Inglaterra: Cambridge University Press. pp. 3-25.
- CLARK, Herbert y Edward Schaefer. (1989) "Contributing to Discourse". En *Cognitive Science*. No. 13. pp. 259-294. Disponible en línea, mayo de 2013:  
[http://onlinelibrary.wiley.com/doi/10.1207/s15516709cog1302\\_7/pdf](http://onlinelibrary.wiley.com/doi/10.1207/s15516709cog1302_7/pdf)
- CORE, Mark y James Allen. (1997) "Coding Dialogs with the DAMSL Annotation Scheme". En *Working Notes of AAI Fall Symposium on Communicative Action in Humans and Machines*. pp. 1-8. Estados Unidos: Massachusetts Institute of Technology. Disponible en línea, mayo de 2013:  
[http://www.fsktm.upm.edu.my/~aida/sak4801/notes/wk9\\_Core-Allen.AAI.pdf](http://www.fsktm.upm.edu.my/~aida/sak4801/notes/wk9_Core-Allen.AAI.pdf)

- CORE, Mark (1997) "Addendum to 'Coding Dialogs with the DAMSL Annotation Scheme'". Estados Unidos: Rochester University. Disponible en línea, mayo de 2013:  
<http://www.cs.rochester.edu/research/cisd/pubs/1997/core-allen-fs97add.pdf>
- CORIA, Sergio. (2008) *Reconocimiento de actos del habla en base a información prosódica por medio de árboles de decisión*. Tesis de Doctorado. Posgrado en Ciencia e Ingeniería de la UNAM. México: Departamento de Ciencias de la Computación, Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, UNAM.
- CUÉTARA, Javier. (2004) *Fonética de la Ciudad de México. Aportaciones desde las tecnologías del habla*. Tesis de maestría. Posgrado en Lingüística Hispánica. México: Facultad de Filosofía y Letras, UNAM.
- ESPINOZA CRUZ, Alejandra. (2007) *Sistematización del fenómeno de silabificación en el Corpus DIME para su aplicación en tecnologías del habla*. Tesis de licenciatura. México: Facultad de Filosofía y Letras, UNAM.
- GROZ, Barbara J. y Candance L. Sidner. (1986) "Attention, Intentions, and the Structure of Discourse". En *Computational Linguistics*, Vol. 12 No. 3. pp. 175-204. Disponible en línea, mayo de 2013:  
<http://acl.ldc.upenn.edu/J/J86/J86-3001.pdf>
- JURAFSKY, Dan. (2000) "Introduction". En *Speech and language processing, an Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. 2ª ed. New Jersey: Pearson-Prentice Hall. pp. 1-16. Disponible en línea, mayo de 2013:  
<http://www.cs.colorado.edu/~martin/SLP/Updates/1.pdf>
- LLISTERRI, Joaquim. (2006) "Introducción a los sistemas de diálogo". En *Los sistemas de diálogo*. Bellaterra-Soria: Manuals de la Universitat Autònoma de Barcelona, Servei de Publicacions - Fundació Duques de Soria. pp. 11-21. Disponible en línea, mayo de 2013:  
[http://liceu.uab.es/~joaquim/publicacions/Llisterri\\_06\\_Sistemas\\_Dialogo.pdf](http://liceu.uab.es/~joaquim/publicacions/Llisterri_06_Sistemas_Dialogo.pdf)
- \_\_\_\_\_. (2003) "Lingüística y tecnologías del lenguaje". En *Lynx. Panorámica de Estudios Lingüísticos 2*. Valencia: Universitat de València, Departament de Teoria dels Llenguatges. pp. 9-71. Disponible en línea, mayo de 2013:  
[http://liceu.uab.es/~joaquim/publicacions/Llisterri\\_03\\_Linguistica\\_Tecnologies\\_Lenguaje.pdf](http://liceu.uab.es/~joaquim/publicacions/Llisterri_03_Linguistica_Tecnologies_Lenguaje.pdf)

- LÓPEZ, Fernanda. (2004) *El estudio de los diptongos del español de México para su aplicación en un reconocedor del habla*. Tesis de licenciatura. México: Facultad de Filosofía y Letras, UNAM.
- MOURE, Teresa y Joaquim Llisterri. (1996) “Lenguaje y nuevas tecnologías. El campo de la lingüística computacional”. En *Avances en lingüística aplicada*. Coordinadora: Milagros Fernández Pérez. Santiago de Compostela: Universidade de Santiago de Compostela, Servicio de Publicacións e Intercambio Científico. pp. 147-228. Disponible en línea, mayo de 2013: [http://liceu.uab.es/~joaquim/publicacions/listerri\\_moure\\_96.html](http://liceu.uab.es/~joaquim/publicacions/listerri_moure_96.html)
- MEZA, Ivan. (2002) *Sistema de verbos auxiliares del español en HPSG*. Tesis de licenciatura. México: Facultad de Ingeniería, UNAM.
- MEZA, Ivan, Elia Pérez, Lisset Salinas, Héctor Avilés y Luis Pineda. (2010) “A Multimodal Dialogue System for Playing the Game ‘Guess the card’”. En *Procesamiento de Lenguaje Natural*. Vol 44. pp. 131-138. Disponible en línea, mayo de 2013: [http://turing.iimas.unam.mx/~ivanvladimir/static/pubs/2010/meza\\_perez\\_salinas\\_aviles\\_pineda-pnl-2010.pdf](http://turing.iimas.unam.mx/~ivanvladimir/static/pubs/2010/meza_perez_salinas_aviles_pineda-pnl-2010.pdf)
- MORENO, Antonio. (1998) *Lingüística computacional*. Madrid: Síntesis.
- MORENO, Iván. (2006) *Detección y reparación de difluencias del habla en el Corpus DIME*. Tesis de licenciatura. México: Facultad de Ingeniería, UNAM.
- PÉREZ-PAVÓN, Elia P. (2006) *Construcción de un reconocedor de voz utilizando Sphinx y el corpus DIMEx100*. Tesis de licenciatura. México: Facultad de Ingeniería, UNAM.
- PINEDA, Luis. (2008) *El proyecto DIME y el robot conversacional Golem: Una experiencia multidisciplinaria entre la computación y la lingüística*. México: DCC-IIMAS, UNAM. Disponible en línea, mayo de 2013: <http://turing.iimas.unam.mx/~luis/golem/papers/DIME-Golem.pdf>
- \_\_\_\_\_. (2004) “Diálogos Inteligentes Multimodales en Español”. *Curso de Tecnologías Lingüísticas “Preguntar al ordenador: las aplicaciones de los sistemas de diálogos”*. Soria: Fundación Duques de Soria, 12 al 16 de junio. Disponible en línea, mayo de 2013: <http://turing.iimas.unam.mx/~luis/golem/papers/pineda-soria-2004.pdf>

- PINEDA, Luis, Varinia Estrada, Sergio Coria y James Allen. (2007) “The obligations and common ground structures of task oriented conversations”. En *Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial* Vol. 11. No. 36. pp. 9-17. Disponible en línea, mayo de 2013:  
<http://polar.lsi.uned.es/revista/index.php/ia/article/viewFile/549/533>
- PINEDA, Luis, Haydé Castellanos, Sergio Coria, Varinia Estrada, Fernanda López, Isabel López, Ivan Meza, Iván Moreno, Patricia Pérez y Carlos Rodríguez. (2006) “Balancing transactions in practical dialogues”. En *Lecture Notes in Computer Science – CICLing 2006*. México: Springer-Verlag, Berlin Heidelberg. pp. 331-342. Disponible en línea, mayo de 2013:  
<http://turing.iimas.unam.mx/~luis/DIME/publicaciones/papers/DIME-DAMSL-38780331.pdf>
- PINEDA, Luis, Antonio Massé, Ivan Meza, M. Salas, E. Schwartz, E. Uranga y L. Villaseñor. (2002) “The DIME project”. En *Lectures Notes in Computer Science. MICAI 2002*. Edición: Carlos Coello, Álvaro de Albornoz, Luis Sucar, Osvaldo Cairó. Vol. 2313. Yucatán, México: Springer-Verlag. pp. 166–175. Disponible en línea, mayo de 2013:  
<http://turing.iimas.unam.mx/~luis/DIME/publicaciones/papers/dime-micai.pdf>
- SEARLE, John. (1962) “¿Qué es un acto del habla?”. En *Cuadernos Teorema*. Versión en español de Luis Valdez. Vol. 15. Valencia: 1977, pp. 1-17.
- VENEGAS, Esther. (2010) *Evaluación del sistema conversacional con español hablado “Golem - Adivina la Carta” en el museo Universum*. Tesis de licenciatura. México: Facultad de Filosofía y Letras, UNAM.
- VILLASEÑOR, Luis, Antonio Massé y Luis Pineda. (2001) “The DIME Corpus”. En *ENC01: 3er. Encuentro Internacional de Ciencias de la Computación*. Edición: C. Zozaya, M. Mejía, P. Noriega y A. Sánchez. Tomo II. Aguascalientes: SMCCINEGI.