

DIÁLOGOS INTELIGENTES

EN ESPAÑOL

MULTIMODALES

SUIJ COPPA - RAFAELA GÓMEZ - CAROLINA MALDONADO
GABRIELA MUÑOZ - YASMARA VERDUGO

1- AUTOR: Dr. LUIS PINEDA CORTÉS



-Ingeniero Electrónico por la Universidad Anahuac (1976-1980), M. en C. en Ciencias Computacionales por el ITESM, Campus Morelos (1983-1986) y doctor (Ph. D.) por el Centre for Cognitive Science de la Universidad de Edimburgo, Escocia (1986-1989).

- Es investigador titular en el IIMAS-UNAM, miembro regular de la Academia Mexicana de Ciencias y de la Academia Mexicana de Informática (AMIAC). Actualmente es el Coordinador de la Red Mexicana de Investigación y Desarrollo en Computación (REMIDEC).



Líneas de investigación: -Sistemas multimodales inteligentes con lenguaje hablado en español.

-Razonamiento diagramático en inteligencia artificial y ciencia del conocimiento.

Proyectos: -Diálogos Inteligentes Multimodales en Español (DIME).

-Navegación en un robot móvil mediante información visual y de lenguaje natural (GOLEM), entre otros.

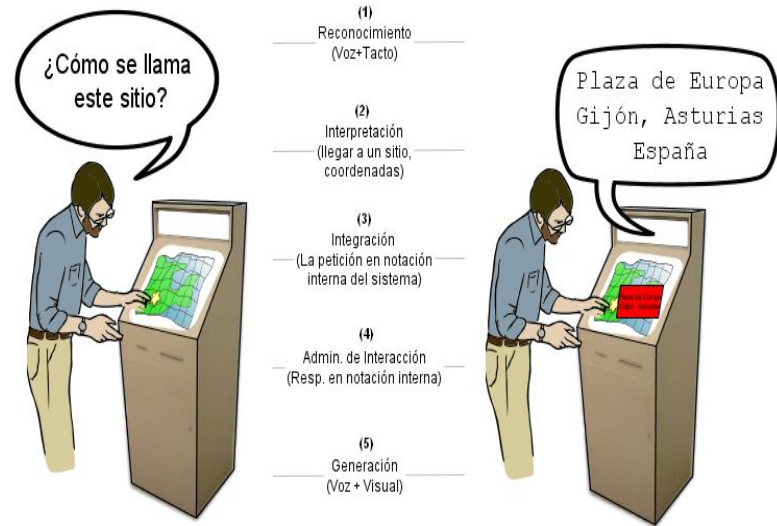
Publicaciones: -Inteligencia Artificial y Linguística Computacional, Investigación Aplicada, entre otros.

- Capítulos en libros, artículos y reportes no publicados, memorias, artículos de divulgación, etc.

Diálogos Inteligentes Multimodales en Español (DIME)

INTRODUCCIÓN:

De manera más general, nuestra investigación tiene por objetivo la creación de modelos de diálogos para contextos conversacionales en los que las intenciones que los participantes puedan expresar o interpretar tengan un carácter concreto, además de que su número sea finito y reducido y, por lo mismo, puedan elucidarse y especificarse a través del análisis. (Pineda, 2004, p.1)



Diálogos Inteligentes Multimodales en Español (DIME)

*Multimodalidad?

***Estructuración del texto:** -Como antecedente de esta investigación se describe un corpus constituido por un conjunto de diálogos orientados hacia el diseño de cocinas.

-Luego, se discuten algunas propiedades de estos diálogos y los problemas que enfrenta su modelación computacional.

-Posteriormente se presenta un modelo para un diálogo de carácter más sencillo en el que un robot da una visita guiada a las instalaciones del Departamento de Ciencias de la Computación IIMAS, UNAM, y se describe su implementación.

2- EL CORPUS DIME

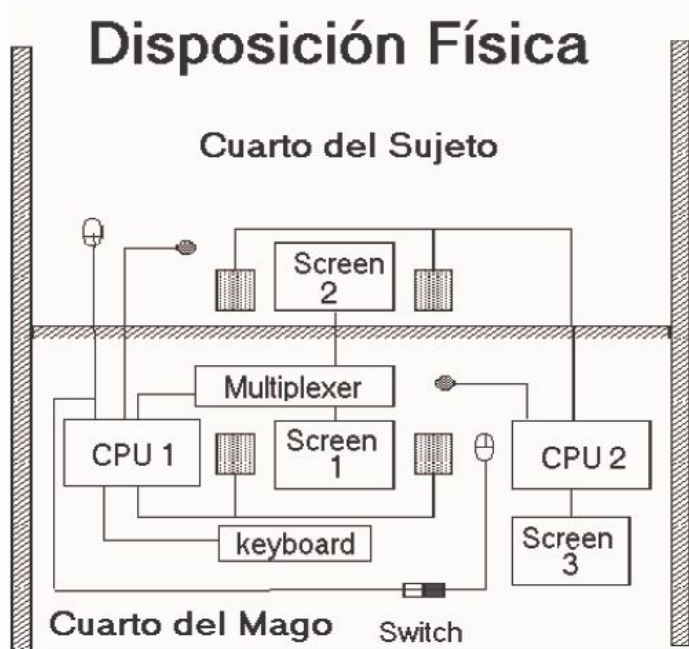


Figura 1. Escenario del Mago de Oz

Base empírica: recolección de un corpus de conversaciones orientadas hacia la solución de una tarea de diseño de cocinas. Recopilación automática.

Escenario: tipo Mago de Oz. Los sujetos saben que el mago era un ser humano.

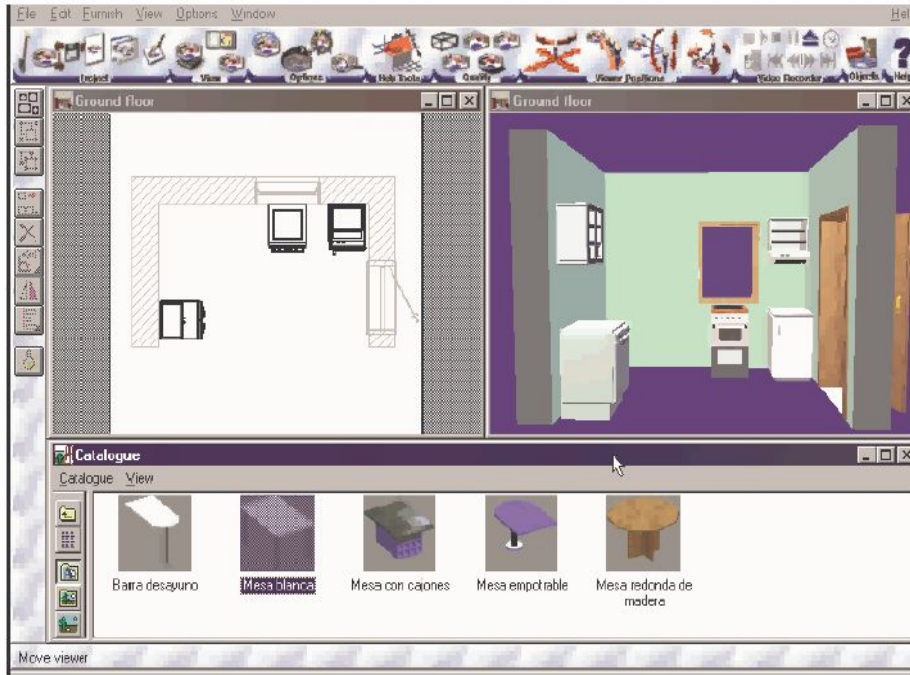


Figura 2. Escenario para el diseño de cocinas

Interfaz: Tres áreas de trabajo principales.

Tarea: Escoger el inmobiliario y posicionarlo en su lugar en la cocina.

Dos tipos de problemáticas iniciales:

Con el cuarto de cocina parcialmente amueblado y con un cuarto vacío

Recolección: Se recolectaron 26 diálogos útiles, siendo cada uno la solución de un problema simple de diseño propuesto por el mago a través del mismo escenario y resuelto mediante la conversación.

Lenguaje: No se impusieron restricciones en el tipo de lenguaje a utilizar. El lenguaje presenta fenómenos característicos del habla espontánea (interjecciones, pausas, etc). Corpus con complejidad de diálogos.



Turno	Part.	Expresión de lenguaje natural	Acción gráfica
utt1	s:	quieres que desplace o traiga algún objeto a la cocina?	
utt2	u:	<ruido> no	
utt3		¿puedes mover la estufa hacia la izquierda?	
utt4	s:	<ruido> ¿hacia dónde ?	Apuntar objeto (estufa)
utt5	u:	<ruido> hacia <sil> hacia la derecha	
utt6	s:	<no-vocal> hacia la derecha	
utt7		<no-vocal> okey	
utt8		¿que tanto quieres que la desplace ?	
utt9	u:	a la mitad del espacio que hay entre la ventana y la pared	
utt10	s:	okey	
utt11		¿quieres que mueva este objeto <sil> hacia acá ?	Apuntar objeto e indicar. dir. mov. (derecha)
utt12	u:	no	
utt13		hacia el otro lado	
utt14	s:	okey	
utt15		¿este objeto hacia acá ?	Apuntar objeto e indicar. dir. mov. (izquierda)
utt16	u:	sí	
utt17	s:	okey	
utt18		<ruido>	Mueve estufa (izquierda)
utt19		¿ahí está bien ? <no-vocal>	
utt20	u:	un poco menos por favor	
utt21	s:	okey	
utt22		<ruido>	Mueve estufa (derecha)
utt23		¿ahí está bien ?	
utt24	u:	ahí está bien	

Tabla 1. Segmento de un diálogo


Se cuenta con videos con las imágenes y eventos de interacción que ocurren en todos los diálogos.

Frecuencia muy baja de expresiones gramaticalmente complejas y la gran mayoría de elocuciones son fragmentos y frases sueltas.

Proceso de determinación de los parámetros de acción: Como un proceso de resolución de referencia.

Proceso de planeación: Consiste en decidir la secuencia de movimientos necesarios para satisfacer la intención solicitada.

Modelación de los diálogos: Implica tomar en cuenta el lenguaje, el contexto de interpretación, las expectativas acerca de las intenciones, asociadas a estas están los actos motores que tienen como función confirmar la información necesaria para determinar las acciones(apuntar, indicar)



3- MODELOS Y ADMINISTRACIÓN DE DIÁLOGOS

3.1.- Conceptos y comprensión del modelo de diálogo

- Modelo basado en la **situación de conversacional**: cada situación está asociada a un conjunto de intenciones presentes en el dominio del diálogo; junto con una modalidad de interacción.

Situación: Objeto representacional que satisface las intenciones asociadas a una situación y es interpretado por un agente conversacional.

Modelo: conjunto o red de situaciones, en la que cada instancia está asociada a un **tipo definido**-> Dominio conversacional.

Con el modelo de diálogo, es posible establecer un **conjunto de actos retóricos** que un agente puede realizar durante una situación para obtener la interpretación de intenciones.

3.1.- Conceptos y comprensión del modelo de diálogo

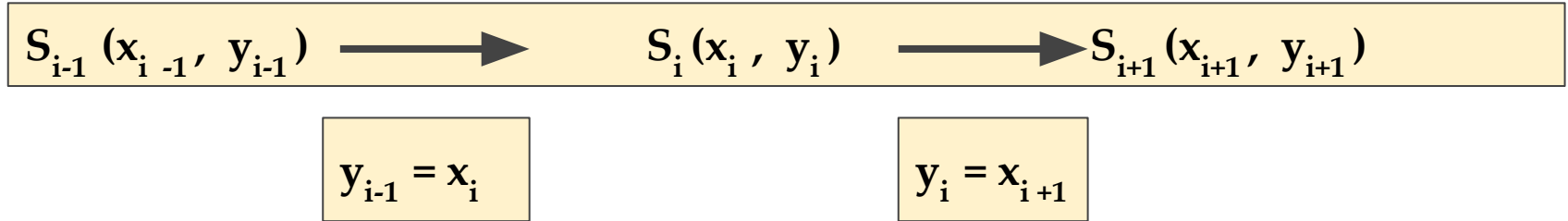
Tipos definidos de cada situación	
Inicial(initial)	Control
Final(final)	
Recursiva(recursive)	
Diciendo(telling)	Acción
Motora(motor)	
Escuchando(listening)	
Error	

- ❖ Todo diálogo tiene una situación inicial y final
- ❖ Mismo formato de situaciones y misma interpretación.
- ❖ Son de tipo **recursivo**: simplificación de un problema complejo en la división de este en subproblemas del mismo tipo
 - Múltiples situaciones recursivas.



3.2.- Transiciones entre situaciones

- ❖ Situación: un objeto abstracto “s” con un par de parámetros de entrada y salida




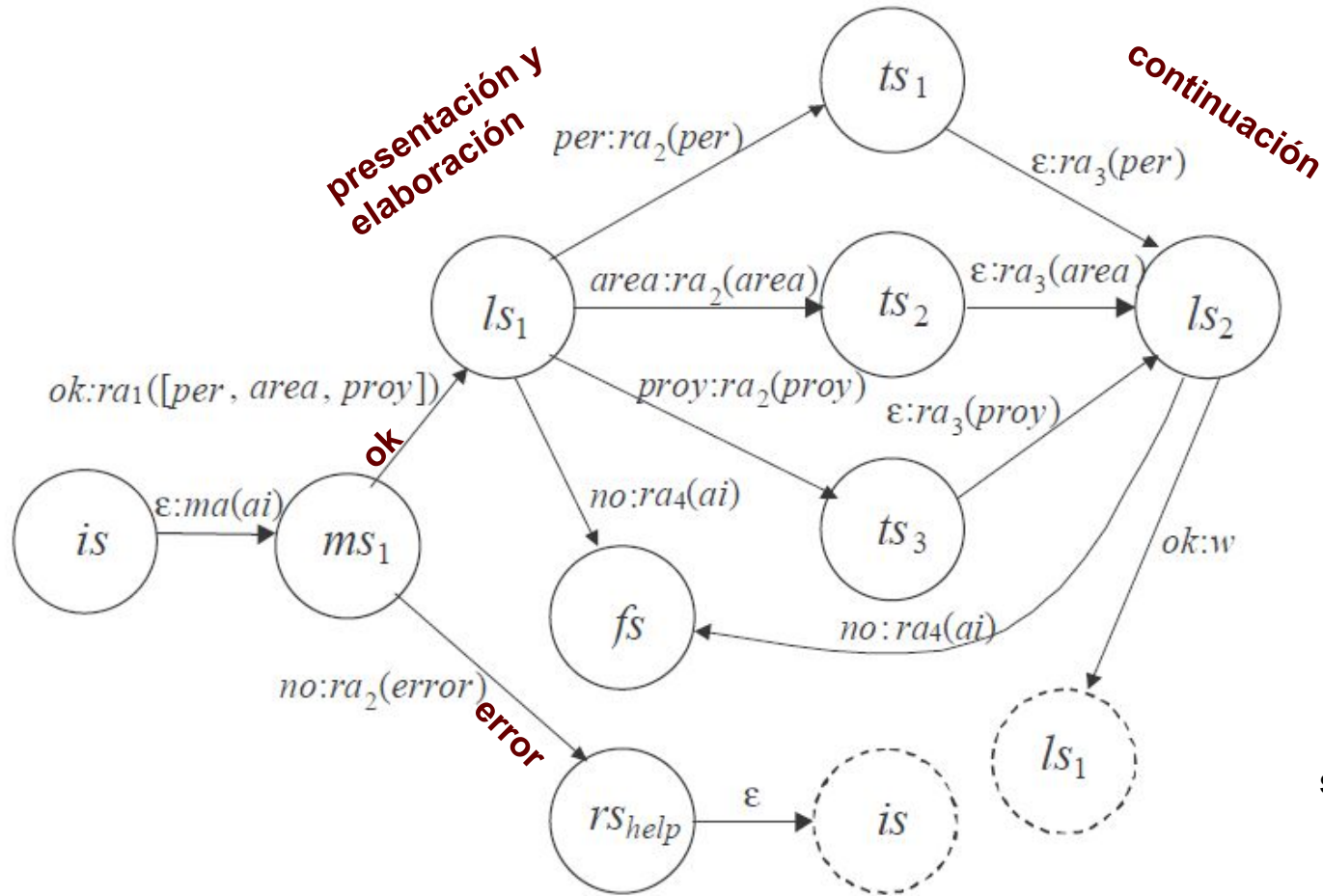
Par de entrada: es el nombre de la situación previa y una transición de la forma
i-sa : o-rha

Par de salida: es la transición de la situación actual a la siguiente y el nombre de la situación a la que se llega con dicha transición: **i-sa : o-rha**

Coherencia de diálogo se logra garantizando que el par de salida de una situación sea siempre el par de entrada de la que sigue.

3.3.- Modelo de diálogo principal y sus situaciones recursivas

- ❖ Visita guiada a las instalaciones de Dep. Ciencias de la Computación
 - ❖ **Variedad lingüística:** interpretación directa de los actos de habla con intenciones relevantes
 - ❖ **Tránsito de las situaciones recursivas:**
 - administrador del diálogo toma el parámetro del acto retórico de entrada
 - pasa el control a la situación inicial de dicho diálogo
 - Situación final se regresa el control a la situación recursiva del diálogo subordinante, el cual ejecuta su transición de salida de manera estándar.
- 



i	initial
f	final
m	motora
l	listening
r	recursive
t	telling

Todas las transiciones de una situación a la siguiente realizan un acto de habla

Figura 4. Modelo de diálogo con situaciones motora y diciendo

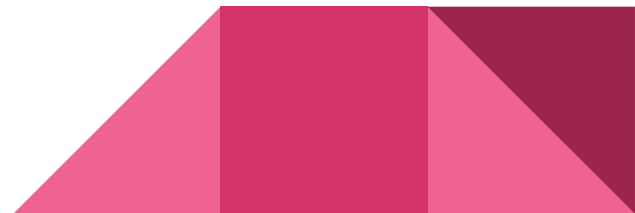
ACTOS RETÓRICOS

Cada acto retórico debe ser definido de manera explícita - **Teoría de los Actos Retóricos (RST)**.

Actos retóricos incluidos RST se definen en términos de **un conjunto de actos retóricos cuya forma está especificada de antemano para todos los modelos.**

Tabla 2 - Actos retóricos figura 1

Identificador, un tipo, una lista de parámetros y conjunto de actos básicos.



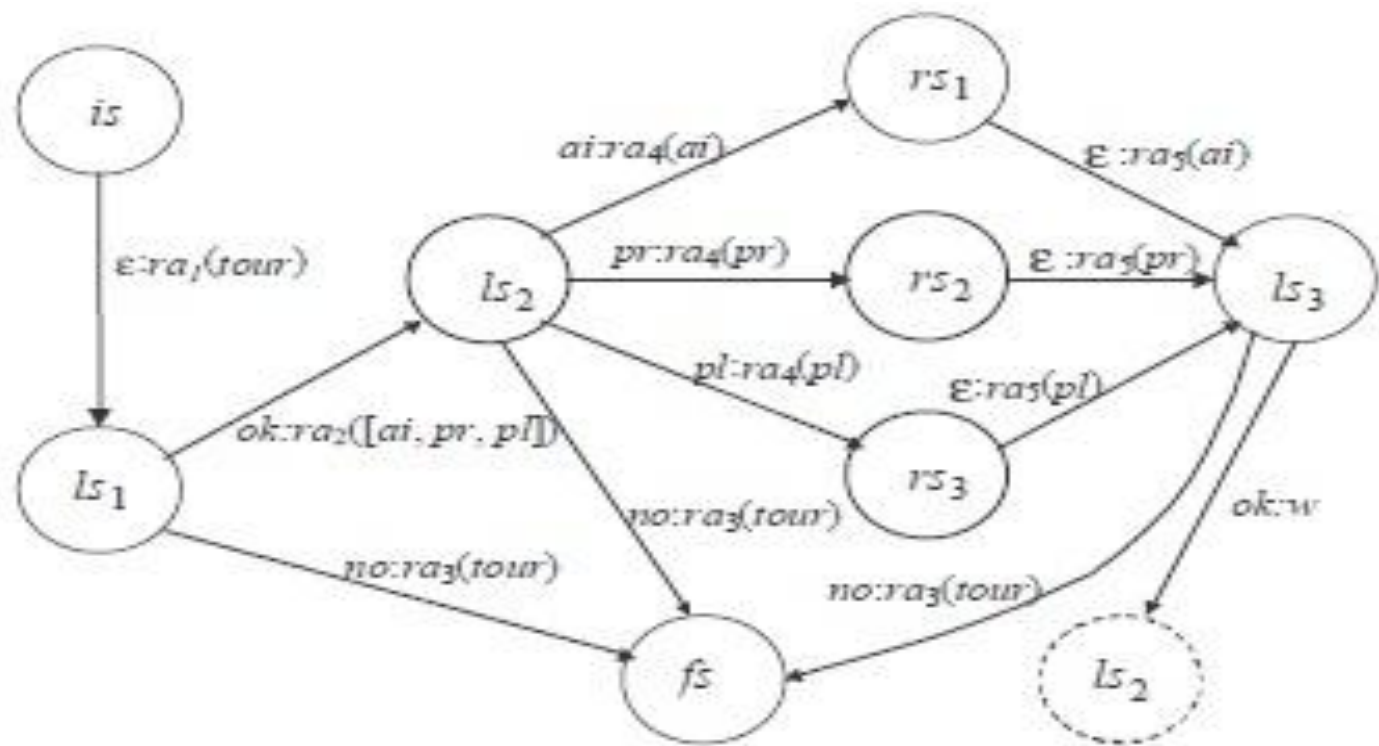


Figure 3. Modelo diálogo principal

Actos retóricos asociados al modelo de diálogo principal

Id	Tipo	Pars.	Actos Retóricos Básicos
<i>Ra₁</i>	<i>Invitación</i>	<i>Tour</i>	<p><i>Saludo</i></p> <p><i>Presentación(Tour)</i></p> <p><i>Opción-s/n(hacer, 'una visita guiada')</i></p>
<i>Ra₂</i>	<i>Oferta</i>	<i>Topicos</i>	<p><i>Introducción(Topicos, 'el departamento', 'las áreas')</i></p> <p><i>Opción-abierta-que('visitar', areas)</i></p>
<i>Ra₃</i>	<i>Despedida</i>		<p><i>Agradecimiento(visitar, departamento)</i></p> <p><i>Despedida</i></p>
<i>Ra₄</i>	<i>Confiración</i>	<i>Area</i>	<i>Confiración('vamos a', Area)</i>
<i>Ra₅</i>	<i>Continuation</i>	<i>Area</i>	<p><i>Agradecimiento(visitar, Area)</i></p> <p><i>Opción-s/n('seguir visitando', 'el departamento')</i></p>
<i>error</i>	<i>Error</i>	<i>Concepto</i>	<p><i>error(Concepto)</i></p> <p><i>solicitar-repetición</i></p>

Tabla 2. Actos Retóricos del modelo de diálogo principal

Actos retóricos asociados al modelo de diálogo subordinado.

Id	Tipo	Pars.	Actos Retóricos Básicos
<i>ra₁</i>	<i>Oferta</i>	<i>Temas</i>	<i>Introducción(Topics, ai)</i> <i>Opción-abierta-como(visitar, ai)</i>
<i>ra₂</i>	<i>Descripción</i>	<i>Concepto</i>	<i>Presentación(Concepto, ai)</i> <i>Elaboración(Concepto, ai)</i>
<i>ra₃</i>	<i>Continuación</i>	<i>Concepto</i>	<i>Notificación(terminamos, Concepto, ai)</i> <i>Opción-s/n('seguir visitando', ai)</i>
<i>ra₄</i>	<i>Conclusión</i>	<i>Concepto</i>	<i>Notificación(terminamos, Concepto)</i>
<i>error</i>	<i>Error</i>	<i>Concepto</i>	<i>error(Concepto)</i> <i>solicitar-repetición</i>

Tabla 3. Actos Retóricos del modelo de diálogo de inteligencia artificial

Interpretación de expresiones lingüísticas del sujeto.

Intención	Expresiones
<i>ok</i>	<i>"si", "okey", "ok", "si, por favor", "por favor", "encantado", "me encantaría", "si, gracias", "aja", "hum, si"</i>
<i>no</i>	<i>"no", "no, gracias", "hum, no", "no tengo tiempo", "ahorita no, gracias", "nel", "nel pastel"</i>
<i>ai</i>	<i>"ai", "ia", "inteligencia artificial", "a inteligencia artificial", "inteligencia", "a inteligencia", "la de inteligencia", "la de inteligencia artificial"</i>
<i>pr</i>	<i>"lenguajes de programación", "lenguajes", "el area de lenguajes", "a lenguajes", "la de lenguajes"</i>
<i>lp</i>	<i>"por reconocimiento de patrones", "reconocimiento de patrones", "reconocimiento", "patrones"</i>
<i>per</i>	<i>"personal", "por personal", "el personal", "los investigadores", "por investigadores", "a personal"</i>
<i>proy</i>	<i>"proyectos", "por proyectos", "los proyectos", "quiero ver los proyectos", "los investigadores"</i>
<i>area</i>	<i>"áreas", "por áreas", "por área", "las áreas", "por secciones", "secciones", "las secciones"</i>

Tabla 4. Modelo del Lenguaje

Diálogo actual producido por versión actual del manejador.

Turno	Part.	Actos Retóricos	Texto
1	s:	Saludo Presentación Opción-s/n	Bienvenido Este es el departamento de ciencias de la computación quieres una visita guiada?
2	u:		Si, por favor
3	s:	Introducción Opción-abierta-que	El departamento tiene las áreas de inteligencia artificial, reconocimiento de patrones y lenguajes de programación Qué área quieres visitar?
4	u:		Inteligencia
5	s:	Confirmación Acto motor Introducción Opción-abierta-como	Okay, vamos a inteligencia artificial El robot se mueve a IA Esta es el área de inteligencia artificial; podemos hacer la visita por personal académico, área o proyectos Cómo quieres hacer la visita?
6	u:		Por proys
7	s:	Error solicitar repetición	Perdón no te entendí Podrías repetirlo?
8	u:		Por proyectos
9	s:	Presentación Elaboración Notificación	En inteligencia artificial se desarrollan los proyectos DIME y GOLEM Las tareas principales son la creación de prototipos con manejadores de diálogos inteligentes, con entrada y salida hablada en español, y visión por computadora Terminamos con proyectos de inteligencia artificial

		<i>Opción-s/n</i>	<i>Quieres seguir visitando inteligencia artificial?</i>
10	u:		<i>Okey</i>
11	s:	<i>Introducción</i> <i>Opción_abierta</i>	<i>inteligencia artificial tiene personal académico y área</i> <i>Cómo quieres visitar inteligencia artificial?</i>
12	u:		<i>Por áreas</i>
13	s:	<i>Presentación</i> <i>Elaboración</i> <i>Notificación</i> <i>Opción-s/n</i>	<i>En inteligencia artificial se investiga en sistemas multimodales inteligentes, gramáticas computacionales del español, razonamiento gráfico y creatividad</i> <i>Y también en visión computacional</i> <i>Terminamos con área de inteligencia artificial</i> <i>Quieres seguir visitando inteligencia artificial</i>
14	u:		<i>No</i>
15	s:	<i>Notificación</i> <i>Agradecimiento</i> <i>Opción-s/n</i>	<i>Terminamos con inteligencia artificial</i> <i>Gracias por visitar inteligencia artificial</i> <i>Quieres seguir visitando el departamento?</i>
16	u:		<i>No, gracias</i>
17	s:	<i>Agradecimiento</i> <i>Despedida</i>	<i>Gracias por visitar el departamento de ciencias de la computación</i> <i>Hasta luego!</i>

4- ENRIQUECIENDO EL MODELO.

Complejidad y ampliación:

1. Revisión en la definición de situación conversacional
2. Caracterización de los actos del habla o intenciones de la situación
3. El modelo de interpretación directa de los actos del habla

-Información prosódica: parte de la fonología que estudia los rasgos sonoros que puede afectar a unidades mayores.

-Limitación del modelo: “escuchar”



CONSECUENCIA DE INTENCIONES SUB-ESPECIFICADAS

1. Necesidad de definir también transiciones sub-especificadas para determinar argumentos de la acción.

Ejemplo: “hacia la izquierda”

2. Para satisfacer las intenciones asociadas al dominio de diseño es necesario realizar acciones concreta.



NATURALEZA DE LOS ACTOS DE HABLA

- El sistema reconoce instancias concretas y no actos de habla.

Ejemplo: “si, por favor”: okey/ aceptación de una oferta

- Se enriquece el sistema no solo codificando la intención total sino su tipo de acto de habla en el esquema **Corpus DIME** (prosódica)



CONCLUSIÓN

Construir un sistema conversacional capaz de llevar a cabo el tipo de diálogos que aparecen en el corpus DIME.

Tareas:

- Terminar la etiquetación del corpus para una base empírica.
- Tener un inventario de los actos del habla y las estructuras retóricas que ocurren en el dominio, como la caracterización de su estructura entonativa.
- Tercer algoritmo: para interpretación directa de actos de habla para identificar el que emite el sujeto en base a información prosódica, como la información léxica que contribuye a la intención de dicho sujeto.

[DIÁLOGOS INTELIGENTES MULTIMODALES EN ESPAÑOL]

```
language_attributes)
charset=<?php bloginfo 'charset' );?>
name="viewport" content="width=device-width
rel="profile" href="http://gmpg.org/xfn/11"
rel="pingback" href="
fruitful_get_favicon();?>
<?php wp_head()
<?php wp_body_class();?>
<div id="page" class="hfeed site"
<?php
$theme_pos = $theme->get_theme_mod('
$log_pos = esc_attr(
if (isset($theme_pos) && $theme_pos)
    $logo_pos_class = fruitful_get_theme_mod('
    $menu_pos_class = fruitful_get_theme_mod('
    $responsive_menu_type =
    $responsive
```

Gracias!