



Universitat d'Alacant Universidad de Alicante

Esta tesis doctoral contiene un índice que enlaza a cada uno de los capítulos de la misma.

Existen asimismo botones de retorno al índice al principio y final de cada uno de los capítulos.

[Ir directamente al índice](#)

Para una correcta visualización del texto es necesaria la versión de [Adobe Acrobat Reader 7.0](#) o posteriores

Aquesta tesi doctoral conté un índex que enllaça a cadascun dels capítols. Existeixen així mateix botons de retorn a l'índex al principi i final de cadascun dels capítols .

[Anar directament a l'índex](#)

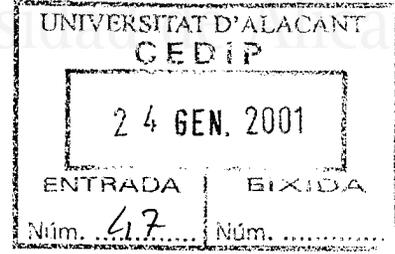
Per a una correcta visualització del text és necessària la versió d' [Adobe Acrobat Reader 7.0](#) o posteriors.

T/VA/I/2001/035



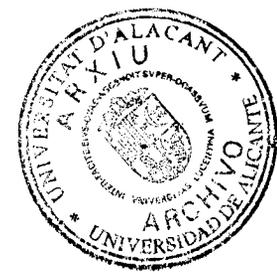
Resolución computacional de la anáfora en diálogos: estructura del discurso y conocimiento lingüístico

Tesis Doctoral



Presentada por
Patricio Manuel Martínez Barco

Dirigida por
Dr. Manuel Palomar Sanz



Depto. de Lenguajes y Sistemas Informáticos
Universidad de Alicante

Alicante, 15 de enero de 2001



Universitat d'Alacant
Universidad de Alicante



Universitat d'Alacant
Universidad de Alicante

Prefacio

El trabajo que aquí se presenta supone la culminación de mi trayectoria investigadora desde sus orígenes en el año 1994, fecha en la que comencé mis estudios doctorales hasta la actualidad. Quiero por tanto comenzar la presentación de este trabajo con una breve descripción de mi actividad investigadora durante este periodo.

Tras finalizar en 1994 mis estudios de la Ingeniería en Informática en la Universidad de Alicante me matriculé en el programa de doctorado “Sistema Industriales, computación y reconocimiento de formas” impartido por el Departamento de Tecnología Informática y Computación (DTIC) durante el curso 1994/1995 en la misma Universidad. Ese mismo año ingresé en el departamento como Colaborador Docente.

Durante el año 1995, el Departamento de Tecnología Informática y Computación se divide generando el nuevo Departamento de Lenguajes y Sistemas Informáticos (DLSI), en el que ingreso como Profesor Asociado y continúo mis estudios doctorales en el programa de doctorado que con el título “Reconocimiento Interpretación y Traducción Automática” ofrece el nuevo departamento.

En el año 1996 ingreso en el Grupo de Investigación de Procesamiento del Lenguaje y Sistemas de Información perteneciente al DLSI y liderado por el Dr. Manuel Palomar Sanz donde comienzo mi investigación en el campo del Procesamiento del Lenguaje Natural, en concreto en la temática del tratamiento y resolución de fenómenos lingüísticos tales como la extraposición de elementos, la anáfora y la elipsis, el desarrollo de técnicas de análisis sintácticos

II Prefacio

parciales para la resolución de estos fenómenos lingüísticos, y la aplicación de estas técnicas a los sistemas de diálogo.

Durante los años 1997 al 2000, participo en el proyecto subvencionado por la Comisión Interministerial de Ciencia y Tecnología (CICYT) *Construcción de analizadores híbridos de lenguajes naturales* (TIC97-0671-C02-01/02) que desarrolla mi grupo de investigación junto al Grup de Processament del Llenguatge Natural de la Universitat Politècnica de València dirigido por la Dra. Lidia Moreno. A consecuencia de este proyecto se producen diversas publicaciones en revistas y congresos de impacto internacional. En 1997 obtengo una plaza de Ayudante de Escuela Universitaria en el departamento y obtengo la suficiencia investigadora a finales de ese año.

Como consecuencia de estas actividades comienzo el desarrollo esta Tesis, mientras se producen en paralelo, y como consecuencia de él, diversas publicaciones y una Acción Integrada titulada *Development of a Corpus-based integrated anaphora resolution system for Spanish and English* (1999-2001) (HB1998-0068) que se desarrolla junto con el *Computational Linguistics Research Group* de la Universidad de Wolverhampton (United Kingdom) que dirige el Dr. Ruslan Mitkov. A consecuencia de esta acción se produce una estancia de investigación en dicha Universidad durante el mes de septiembre de 1999.

Se destaca además la reciente concesión de dos nuevos proyectos de investigación en los que participo: el proyecto subvencionado por la Comisión Interministerial de Ciencia y Tecnología (CICYT) *Desarrollo de un sistema de comprensión de textos aplicado a la recuperación de información (TUSIR)* (2001-2003) (TIC2000-0664-C02-02) desarrollado junto al Grupo de Procesamiento del Lenguaje Natural de la Universidad Politècnica de Valencia; y el proyecto *Extending the EuroWordNet with Public Sector Terminology (EUROTERM)* (2001) dentro del programa de acción preparatoria subvencionado por la Comisión Europea, a desarrollar junto al *Department of Computer Engineering and Informatics* de la Universidad de Patras (Greece) y el *Center for Applied Research* de Tilburg (The Netherlands).



Universitat d'Alacant
Universidad de Alicante

Agradecimientos

Al finalizar este trabajo, quiero agradecer sinceramente a todos aquellos que me han proporcionado colaboración y aliento en la consecución del mismo.

En primer lugar quiero destacar todo el apoyo recibido por parte de Manuel Palomar como director de esta Tesis. Su inestimable ayuda, ha quedado demostrada desde el principio de mi investigación tanto desde el rol de director, proporcionándome todo tipo de ideas, consejos, inspiración, así como un seguimiento pormenorizado de mi trayectoria, como desde el rol de compañero participando activamente en todas las tareas de investigación.

También estoy particularmente agradecido a Lidia Moreno, directora del Grup de Processament del Llenguatge Natural de la Universitat Politècnica de València por todas las revisiones y consejos que me ha proporcionado tanto en la elaboración de esta memoria como en el desarrollo de mi actividad investigadora.

Por otra parte, también quiero destacar especialmente la labor de Antonio Ferrández, quien ha actuado también como promotor y supervisor del trabajo, y que ha prestado una colaboración muy especial en el desarrollo de la herramienta, proporcionando su sistema de restricciones y preferencias como base para nuestro trabajo, y colaborando activamente en las tareas de implementación.

Tampoco puedo pasar por alto el apoyo de mis compañeros del Grupo de Procesamiento del Lenguaje y Sistemas de Información: Andrés Montoyo, Armando Suárez, Borja Navarro, Cristina Cachero, Fernando Llopis, Jaime Gómez, Jesús Peral, Jose Luis Vicedo, Juan Carlos Trujillo, Maxi Saiz, Rafa Muñoz, y Rafa Romero, además de los ya nombrados Manuel Palomar y Antonio

IV Agradecimientos

Ferrández, así como al resto de colaboradores del grupo que sería muy largo nombrar aquí. En muchas ocasiones han intervenido directamente en la investigación, en otras lo han hecho a través de la depuración formal de la memoria, y siempre han estado apoyándome desde el plano puramente personal.

A todos los compañeros del Departamento de Lenguajes y Sistemas Informáticos que me han respaldado en el proceso de investigación y que en ocasiones me han descargado de otras tareas para permitirme dedicar el tiempo necesario al desarrollo de este trabajo.

A los compañeros del Grup de Processament del Llenguatge Natural de la Universitat Politècnica de València, Antonio Molina, Emilio Sanchís, Encarna Segarra, Ferrán Pla y Natividad Prieto, con los que he intercambiado frecuentemente impresiones acerca del trabajo y que me han brindado las herramientas necesarias para realizar el entrenamiento y evaluación del mismo.

A todos los revisores anónimos de congresos y revistas a los que hemos sometido nuestra investigación, ya que con sus orientaciones y críticas nos han permitido centrar nuestros objetivos.

Y a David Calduch, cuya revisión formal de la memoria ha permitido depurar diversas erratas y cuestiones de estilo.

Finalmente quiero hacer un agradecimiento especial a mi familia: a Paloma, que nunca ha dejado de alentarme para que pudiera finalizar esta Tesis con éxito y cuya colaboración en el plano profesional siempre ha sido determinante; y a mis padres y hermanos, a quienes les debo no sólo su apoyo en este trabajo, sino también en todo mi curriculum tanto académico como profesional.



Índice General

Universitat d'Alacant
Universidad de Alicante

1. Introducción	1
1.1 Sistemas de diálogo	4
1.2 El problema de la anáfora	7
1.3 Objetivos	14
1.4 Organización de la Tesis	15
2. El diálogo y su estructura	19
2.1 Taxonomía de los diálogos	19
2.1.1 Modalidad: <i>Diálogos escritos (tecleados) y hablados</i>	20
2.1.2 Clases de agentes: <i>hombres y máquinas</i>	21
2.1.3 Contexto compartido: <i>espacial - temporal</i>	22
2.1.4 Distancia de la tarea del diálogo: <i>orientados a tareas y de recuperación de información</i>	22
2.1.5 Número y tipo de las tareas	23
2.1.6 Clases de conocimiento compartido: <i>perceptual, lingüístico y temporal</i>	23
2.2 Los sistemas de diálogo	24
2.2.1 Arquitectura de los sistemas de diálogo	25
2.2.2 Componentes de un sistema de diálogo genérico	27
2.2.3 La gestión del discurso en los sistemas de diálogo	30
2.2.4 Actualidad de los sistemas de diálogo	33
2.3 Segmentación y estructuración del discurso en diálogos	36
2.3.1 Segmentación basada en el discurso	36
2.3.2 Segmentación basada en el tópico del discurso	39
2.4 Anotación de la estructura del diálogo	44

VI Índice General

2.4.1	Esquemas de anotación	44
2.4.2	Fiabilidad de los esquemas de anotación manual	57
2.5	Conclusiones del capítulo	59
3.	La anáfora en diálogos	61
3.1	Taxonomía de las relaciones anafóricas en el diálogo	62
3.1.1	Relaciones anafóricas según la categoría gramatical del elemento anafórico	62
3.1.2	Relaciones anafóricas según la naturaleza del antecedente	69
3.2	Estrategias de resolución de la anáfora en diálogos	71
3.2.1	Enfoques estratégicos en la resolución de la anáfora	74
3.2.2	Métodos integrados basados en información lingüística	75
3.2.3	Métodos integrados basados en la estructura del discurso	81
3.2.4	Métodos integrados basados en restricciones y preferencias	90
3.2.5	Métodos alternativos basados en información estadística	92
3.3	Influencia de la estructura del diálogo en la resolución de la anáfora	96
3.3.1	Uso de anáforas dentro de un par adyacente	97
3.3.2	Uso de anáforas en el siguiente par adyacente	98
3.3.3	Uso de anáforas en pares adyacentes anidados	99
3.3.4	Uso de la anáfora en segmentos con más de un antecedente	101
3.3.5	El tópico del diálogo y la anáfora	103
3.4	Sistemas de diálogo, estructura del discurso y resolución de la anáfora	104
3.5	Conclusiones del capítulo	105
4.	El espacio de la solución	109
4.1	Propuesta de un esquema de anotación para la estructura del diálogo	110

4.2	Definición del espacio de accesibilidad anafórica estructural (EAAE)	114
4.3	Estudio sobre el espacio de accesibilidad anafórica estructural	116
4.3.1	Estudio empírico	116
4.3.2	Otras alternativas sobre el espacio de accesibilidad anafórica	119
4.4	Estandarización del esquema de anotación	121
4.4.1	Anotación de textos mediante el lenguaje de estándares SGML	122
4.4.2	Anotación de la estructura del diálogo con SGML	123
4.5	Conclusiones del capítulo	125
5.	ARIADNA: Sistema de resolución de la anáfora en diálogos	127
5.1	Orígenes de ARIADNA	128
5.1.1	Propuesta inicial de resolución de anáfora basada en la detección automática de la estructura del diálogo	129
5.1.2	Características del sistema ARIADNA	137
5.2	Módulos de preproceso en ARIADNA	138
5.2.1	Etiquetador léxico-morfológico	139
5.2.2	Segmentador de enunciados	140
5.2.3	Analizador sintáctico parcial SUPP	141
5.2.4	Etiquetado de la estructura del diálogo	152
5.3	Módulos de resolución anafórica en ARIADNA	153
5.3.1	Disparador anafórico	154
5.3.2	Generador del espacio de accesibilidad anafórico	155
5.3.3	Sistema de Restricciones	156
5.3.4	Sistema de Preferencias	162
5.3.5	Algoritmo ARDi	167
5.4	Adaptabilidad de ARIADNA a un sistema de diálogo	168
5.5	Conclusiones del capítulo	170

VIII Índice General

6. Experimentación	171
6.1 Corpus	172
6.1.1 Corpus Infotren: Persona	173
6.2 Anotación de la correferencia	174
6.3 Módulo de evaluación automática de resultados	177
6.3.1 Precisión	178
6.3.2 Cobertura	178
6.4 Entrenamiento del sistema de preferencias	178
6.4.1 Entrenamiento en ARIADNA v.1	179
6.4.2 Entrenamiento en ARIADNA v.2	185
6.5 Evaluación de ARIADNA v.2.	190
6.5.1 Discusión y consideraciones sobre los resultados	193
6.6 Conclusiones del capítulo	194
7. Conclusiones y trabajos futuros	197
7.1 Trabajos en progreso	200
7.2 Producción científica	202
Referencias	207



Índice de Tablas

Universitat d'Alacant
Universidad de Alicante

3.1	Tipos de transición en el <i>Centering</i>	83
3.2	Refinamiento de los tipos de transiciones según Brennan <i>et al.</i>	83
4.1	Estudio empírico sobre el espacio de accesibilidad anafórico	118
4.2	Comparativa entre espacio de accesibilidad anafórico estructural (EAAE), espacio basado en ventana de oraciones (EAAVO) y espacio completo (EAAC)	120
4.3	Estudio del espacio de accesibilidad anafórico basado en ventana de oraciones (EAAVO) adaptada a enunciados .	121
5.1	Preferencias de la anáfora pronominal en ARIADNA v.2.	167
5.2	Preferencias de la anáfora adjetiva en ARIADNA v.2. .	167
6.1	Ordenación de preferencias durante el entrenamiento de ARIADNA v.1. Anáfora Pronominal	185
6.2	Ordenación de preferencias durante el entrenamiento de ARIADNA v.1. Anáfora Adjetiva	186
6.3	Ponderación de preferencias durante el entrenamiento de ARIADNA v.2. Anáfora Pronominal	191
6.4	Ponderación de preferencias durante el entrenamiento de ARIADNA v.2. Anáfora Adjetiva	191
6.5	Evaluación de ARIADNA v.2.....	192
6.6	Análisis de errores en la evaluación de ARIADNA v.2...	193



Universitat d'Alacant
Universidad de Alicante



Universitat d'Alacant
Universidad de Alicante

Índice de Figuras

2.1	Arquitectura de un sistema genérico de diálogos por Allen <i>et al.</i> (2000)	27
3.1	Árbol probabilístico de la anáfora pronominal en inglés	95
4.1	Ejemplo de diálogo anotado según nuestra propuesta	113
4.2	Estudio empírico sobre el espacio de accesibilidad anafórico	117
4.3	Ejemplo de anotación SGML de la estructura del diálogo	125
5.1	Sistema de restricciones y preferencias de Ferrández <i>et al.</i> (1998)	128
5.2	Algoritmo de resolución de la anáfora con segmentación automática basada en el tópico	130
5.3	Visión global del sistema ARIADNA	137
5.4	Salida ejemplo del segmentador de enunciados	142
5.5	Ejemplo de un gramática SUG	144
5.6	Gramática SUG parcial simplificada	146
5.7	Ejemplo de análisis parcial con SUPP	148
5.8	Ejemplo de una estructura de huecos generada por SUPP (1/2)	150
5.9	Ejemplo de una estructura de huecos generada por SUPP (2/2)	151
5.10	Ejemplo de la salida del interfaz SGML-PROLOG para la estructura del diálogo	154
5.11	ARDi: Algoritmo de resolución de la anáfora	168
6.1	Proceso de evaluación de ARIADNA	172



Universitat d'Alacant
Universidad de Alicante



Universitat d'Alacant
Universidad de Alicante

1. Introducción

El lenguaje natural como medio de transmisión de conocimientos es una de las características fundamentales de la esencia humana. Sin lenguaje no es posible la comunicación y sin comunicación no es posible el intercambio de información.

Una de las formas de comunicación más habituales es el uso del diálogo, que se define en Casares (1994) como *el mecanismo de conversación entre dos o más personas que alternativamente manifiestan sus ideas o afectos*.

El mecanismo de conversación y en concreto el lenguaje conversacional se estudia desde diversas disciplinas: la lingüística, la psicolingüística, la filosofía y la lingüística computacional. En concreto, la *lingüística computacional* es una disciplina que intenta reproducir la transmisión natural de la información mediante el modelado de la producción del hablante (o escritor) y la interpretación del oyente (o lector) en un sistema de ordenador (Hausser, 1999).

En la lingüística computacional se pueden distinguir dos tipos de aplicaciones según el tipo de discurso para el que se desarrollan: aplicaciones basadas en el tratamiento masivo de información textual y aplicaciones basadas en diálogos que tratan la comunicación hombre-máquina tanto oral como escrita (Moreno *et al.*, 1999).

Entre las aplicaciones basadas en el tratamiento masivo de información destacaremos los trabajos realizados en las siguientes áreas:

- Extracción y Recuperación de Información: los Sistemas de Extracción de Información son sistemas donde, a partir de un texto, se extrae información relevante de forma estructurada. Los

Sistemas de Recuperación son sistemas que, a partir de una colección de textos con formato heterogéneo, extraen aquellos que se consideran relevantes de acuerdo a algún requerimiento. Estas tareas alcanzan mayor significación en la actualidad, donde gracias a la proliferación de textos publicados en Internet -se calcula que en 10 años se habrá duplicado la cantidad de información escrita durante toda la historia (Hausser, 1999)-, cada vez se hace más necesario poder extraer y recuperar información a través de los contenidos semánticos del texto.

- Traducción automática: sistemas capaces de realizar traducciones entre dos o más lenguajes mediante la interpretación semántica del texto y que van mas allá de los antiguos sistemas de traducción literal palabra a palabra. Especialmente, la Unión Europea, una comunidad que actualmente está compuesta por quince países que hablan en once lenguas nativas distintas (sin contar lenguas y dialectos oficiales en sus respectivas autonomías), requiere con urgencia sistemas de traducción automática capaces de generar textos simultáneamente en todas las lenguas de los países integrantes.
- Producción automática de textos: sistemas que permiten el mantenimiento automático de textos como páginas Web, manuales técnicos, generación automática de circulares informativas, etc. Estos sistemas toman generalmente la información principal de una base de datos y a partir de ésta construyen el texto explicativo en lenguaje natural.
- Corrección automática de textos: sistemas usados para la detección de errores ortográficos, errores en la ordenación de las palabras, concordancias, etc. Estos sistemas se implantan especialmente en los programas de tratamiento de textos.
- Producción automática de resúmenes: sistemas creados para proporcionar sinopsis automáticas de los textos que se pueden encontrar en las bases de datos textuales, en páginas publicadas en Web, etc.

Por otra parte, las aplicaciones basadas en diálogos se pueden clasificar en dos grandes áreas de trabajo:

- Sistemas de acceso a otros sistemas: como bases de datos donde el usuario solicita información almacenada en una base de datos mediante el lenguaje natural. Por su parte, el sistema informático tiene como objetivo la compilación de expresiones producidas en lenguaje natural a una forma interpretable por sistemas de gestión de bases de datos a través de su propio lenguaje de consulta. Estos sistemas tratan de evitar la necesidad actual que tienen los usuarios de bases de datos de conocer lenguajes artificiales o formales para poder acceder de forma libre a la información almacenada. También se incluyen en este grupo otros sistemas de acceso a otros dominios entre los que destacan interfaces para sistemas expertos y los interfaces con sistemas operativos. En general, se trata de adaptar el dominio de trabajo del sistema informático al entorno de comunicación humano en lugar de ser el usuario el que adapte al entorno del sistema informático.
- Sistemas inteligentes de diálogo: en los que se estudia el comportamiento del diálogo en su aspecto intelectual. En estos sistemas se produce una comunicación entre hombre y máquina donde el usuario conversa con el sistema mediante el lenguaje natural, y el sistema genera respuestas en lenguaje natural que se construyen tras la formalización de aspectos como las intenciones y los deseos de los usuarios, el conocimiento y las creencias de ese conocimiento y la relación entre conocimiento y acción. La investigación relativa a estos sistemas suele estar focalizada en dos grandes grupos:
 - sistemas inteligentes de diálogo orientados a tareas donde el fin principal es conseguir que la máquina dirija al usuario (o viceversa) a la consecución de una tarea concreta
 - sistemas inteligentes de diálogo orientados a la extracción y/o recuperación de información donde el fin principal es la obtención de una información (por parte del usuario o de la máquina).

En concreto, en esta Tesis nos centraremos en aplicaciones basadas en diálogos, y más concretamente, en los sistemas inteli-

gentes de diálogo (que a partir de ahora llamaremos sistemas de diálogo).

1.1 Sistemas de diálogo

Desde el punto de vista de la lingüística computacional, los sistemas de diálogo, se tratan desde tres perspectivas distintas:

- el procesamiento del lenguaje natural,
- la gestión del diálogo, y
- el procesamiento del habla

En los sistemas de diálogo, una de las tareas principales de la lingüística computacional tiene como objetivo principal el modelado del mecanismo de la comunicación en lenguaje natural mediante un análisis de oraciones a través de los distintos niveles del lenguaje: morfológico, léxico, sintáctico, semántico y pragmático. Este modelado del mecanismo de la comunicación humano se conoce con el nombre de *procesamiento del lenguaje natural (PLN)*. Así, el procesamiento del lenguaje natural se lleva a cabo mediante algoritmos y estructuras de datos computacionales.

Para centrar el concepto de procesamiento del lenguaje natural y sus tareas en los sistemas de diálogo, lo introduciremos desde los distintos niveles del lenguaje donde se produce.

De acuerdo con Moreno *et al.* (1999), el procesamiento del lenguaje natural se produce a través del análisis de oraciones desde las diferentes perspectivas que generan los distintos niveles de lenguaje:

- El análisis léxico realiza una transformación de las cadenas de caracteres de entrada en una secuencia de unidades significativas o unidades léxicas mediante el uso de un diccionario.
- El análisis morfológico complementa el análisis léxico mediante la introducción de información morfológica de género, número y persona obtenida mediante reglas morfológicas y el apoyo del diccionario.

- El análisis sintáctico, que se efectúa a partir del morfológico y léxico, analiza la secuencia de unidades léxicas y produce una representación de su estructura sintáctica. Esta estructura suele representarse por medio de árboles, redes o listas de constituyentes. El análisis sintáctico está generalmente dirigido por una gramática capaz de derivar constituyentes sintácticos a partir de unidades léxicas y de la concordancia morfológica de los constituyentes. Según las necesidades, este análisis puede realizarse completo, cuando intenta obtener todos y cada uno de los constituyentes de la oración (análisis sintáctico completo); o puede realizarse parcialmente, cuando sólo se contemplan algunos constituyentes (análisis sintáctico parcial). Para aplicaciones que trabajan en dominios no restringidos el análisis parcial se hace especialmente necesario, ya que un análisis completo que contemple todas y cada una de las circunstancias posibles sería inviable a efectos prácticos.
- En el análisis semántico, a partir de la estructura generada por el análisis sintáctico se genera otra estructura o forma lógica asociada (en este caso semántica) que representa el significado o sentido de la oración independiente del contexto.
- Por último, el análisis contextual o pragmático utiliza la forma lógica o estructura semántica anterior para obtener la interpretación final de la oración dependiente del contexto. Una tarea fundamental de este análisis es el tratamiento de los distintos fenómenos lingüísticos, de los que destacamos la resolución de la anáfora y la elipsis entre otros.

Además de las tareas de procesamiento del lenguaje descritas, la lingüística computacional tiene una tarea fundamental que es la propia gestión del diálogo. Esta tarea tiene una misión dual en el sistema de diálogo, por una parte, debe reconocer las intenciones del usuario en el diálogo concreto que se produce en un determinado momento, y por otra, saber interpretar cuáles son las obligaciones del sistema respecto a ese diálogo, y por tanto, ejecutar esas operaciones. Estas obligaciones pueden incluir la ejecución de operaciones de consulta sobre una base de datos, planificar tareas, y sobre todo, el control sobre la generación de

turnos de conversación de la máquina hacia el usuario para la transmisión de respuestas de preguntas aclaratorias.

Por último, desde el punto de vista del reconocimiento del habla, cuando el proceso anterior se aplica al lenguaje hablado, la cadena de entrada además debe ser preprocesada por un reconocedor de habla y la respuesta generada a través de un sintetizador de voz.

El trabajo que aquí se presenta se centra en la tarea de procesamiento del lenguaje natural, y en concreto, en la resolución de la anáfora en aplicaciones basadas en diálogos. De esta forma, una aplicación inmediata es su incorporación a los sistemas de diálogo como uno de los módulos que actúa en la tarea del procesamiento del lenguaje. De acuerdo con estos sistemas, el módulo de resolución de la anáfora se usará como una herramienta consultiva que permitirá una mejor interpretación del diálogo tratado. Así el sistema de resolución de la anáfora definido en esta memoria encaja con las características básicas de los sistemas de diálogo ya que proporciona la solución de la anáfora según las necesidades de estos sistemas y utiliza únicamente la información que éstos disponen. Además, este trabajo se basa en el uso de información independiente del dominio, por lo que puede ser adaptado a cualquier sistema de diálogo definido sobre cualquier dominio.

Destacamos, no obstante, que se estudia el problema de la resolución de la anáfora independientemente de la aplicación sobre la que finalmente sea aplicado. Para ello se realiza un estudio y descripción de las diferentes herramientas que proporcionan la información necesaria para la resolución de la anáfora. Por tanto, las aplicaciones posibles para este trabajo no se limitan exclusivamente a los sistemas de diálogo. El procesamiento masivo de corpus de texto con transcripciones de diálogos para generar una interpretación anafórica podría encontrarse entre otras aplicaciones posibles.

1.2 El problema de la anáfora

Tal y como se muestra en Moreno *et al.* (1999) la anáfora es uno de los problemas más difíciles de resolver en el procesamiento del lenguaje natural. Hirst (1981) define la anáfora como *el mecanismo que permite hacer en un discurso una referencia abreviada a alguna entidad o entidades, con la confianza de que el receptor del discurso sea capaz de interpretar la referencia y por consiguiente determinar la entidad a la que se alude*. De esta definición extraemos los componentes básicos del proceso anafórico: la referencia abreviada a la que se denomina *expresión* o *elemento anafórico* y la entidad referenciada que se denomina *referente* o *antecedente*.

En Brown y Yule (1983) se hace una diferenciación importante entre los conceptos de referente y antecedente. Se considera como referente la representación mental de los objetos que han sido evocados en el texto, mientras que el antecedente constituye la representación lingüística del referente.

En el ejemplo 1 se muestra un fragmento de diálogo entre dos hablantes A, y B donde se produce una anáfora generada por el pronombre *lo*. En este caso, el referente sería la representación mental del reloj de B, mientras que el antecedente sería la representación lingüística, *tu reloj*. En la notación empleada para éste y ejemplos sucesivos, el texto subrayado muestra el antecedente, el texto en negrita muestra la anáfora y la coincidencia de subíndices indica que ambos están relacionados (la no coincidencia indicaría lo contrario).

- (1) A: ¿Has encontrado ya tu reloj_i?
 B: No, creo que **lo**_i he perdido.

Siguiendo con la exposición de Moreno *et al.* (1999), el problema de la anáfora debe tratarse como dos procesos distintos: la resolución y la generación de la anáfora. La resolución de la anáfora busca la entidad a la que hace referencia, mientras que la generación crea referencias sobre una entidad del discurso. Los sistemas de diálogo necesitan ambos procesos. Por un lado, resolver la anáfora creada por el usuario y que debe ser interpretada por la

máquina para poder comprender lo que el usuario transmite. Pero, por otra parte, las anáforas se generan por los hablantes para enfatizar ciertas entidades (acto conocido como *topicalizar*), incluso para remarcar la propia estructura del discurso en el diálogo (la estructura del diálogo¹). Es, por tanto, necesario que una máquina que simula el comportamiento humano en el diálogo también sea capaz de generar anáforas.

Esta Tesis se centra en la resolución y no en la generación. Sin embargo, las conclusiones que se alcanzan relativas a la influencia de la estructura del diálogo en la anáfora son perfectamente aplicables al proceso de la generación anafórica.

Por otra parte, existe cierta confusión terminológica respecto a los términos anáfora y referencia. A continuación intentaremos clarificar ambos conceptos antes de profundizar más en el problema.

De acuerdo con el *Linguistic Glossary* (Summer Institute of Linguistics, 1997) la *referencia*, en su acepción más general, se define como la relación simbólica que una expresión lingüística tiene con el objeto concreto o abstracto al que representa. Así se deriva la definición de la *correferencia* como la referencia en una expresión al mismo referente en otra expresión.

La referencia, a su vez puede clasificarse en *endófora* y *exófora*.

Endófora. Se considera como endófora la correferencia de una expresión lingüística con otra que se encuentra antes o después de ella pero siempre dentro del discurso lingüístico. Una de las expresiones contendrá la información necesaria para interpretar a la otra. Cuando la expresión que contiene la información necesaria para interpretar a la otra se encuentra antes de ésta se habla de *anáfora* propiamente dicha. Sin embargo, cuando la expresión que

¹ El término *estructura del discurso* es utilizado ampliamente en la literatura para hablar de la forma en la que los enunciados del discurso se reúnen en secuencias para formar segmentos con un determinado valor intencional y atencional (Grosz & Sidner, 1986) (Fox, 1987). Sin embargo, cuando se trata de un concreto de un discurso de tipo diálogo, algunos autores se han referido a esta estructura como la *estructura del diálogo* (Carletta *et al.*, 1997). En este sentido, adoptaremos en general esta última denominación cuando hagamos referencia a la estructura del discurso en el diálogo, y mantendremos la denominación *estructura del discurso*, cuando hagamos referencia a ella en general.

lleva la carga semántica se encuentra después de la otra, entonces se habla de *catáfora*².

Exófora. Se entiende como exófora la referencia directa a un objeto extralingüístico. Las exóforas más comunes son la *deixis* y la *homófora*.

Se asignan a la categoría de *deixis* aquellos casos de anáfora en los que la identificación del antecedente depende de la información disponible en el entorno físico donde ocurre el diálogo.

(Lyons, 1977) proporciona la siguiente definición de la deixis: *La localización e identificación de personas, objetos, eventos, procesos y actividades de las que se habla, o a las que se refiere, en relación con el contexto espacio-temporal creado y sostenido por el acto del enunciado y la participación en él; típicamente, de un hablante simple y al menos un direccionamiento (alguna forma de apuntar a algo).*

Según (Huls *et al.*, 1995), la interpretación de las expresiones deícticas dependen de la identidad del hablante y de la audiencia, el tiempo verbal, la localización temporal del hablante y la audiencia, y actos comunicativos no lingüísticos como expresiones faciales y movimientos de ojos, manos y cuerpo. Así se distinguen tres tipos de deixis: personal, temporal y espacial. La deixis personal implica pronombres de primera y segunda persona, y su solución implica el conocimiento de quién está participando en el diálogo en cada momento, es decir, quién habla y quién escucha. La temporal relaciona el tiempo del habla con las relaciones expresadas por el enunciado. Su resolución implica el conocimiento de cuál es el tiempo actual. La deixis espacial implica el uso de demostrativos o de otras expresiones de referencia que se producen en combinación con un gesto de apuntar. En este caso hace falta tener un conocimiento visual del escenario donde se desarrolla el diálogo para poder resolver.

La *homófora*, por su parte, se define como una referencia que depende del conocimiento cultural o de otro conocimiento general

² Algunos autores como Ersan y Akman (Ersan & Akman, 1994) sólo consideran el término *anáfora* válido para nombrar al tipo de referencia endófora que ocurre después de nombrar el antecedente. Sin embargo en algunos trabajos de la literatura actual se ha venido considerando el término *anáfora* para nombrar también a la forma catafórica

más que de características específicas de un contexto particular. En este caso se encontraría el uso de *el Presidente* para referirse al presidente de un país o el uso de *el niño* entre padres para referirse a su propio hijo.

Nos centraremos, por tanto, en la anáfora como un tipo de correferencia endófora tal y como se ha definido anteriormente y no en la referencia en general.

Por otra parte, las estrategias para la resolución de la anáfora se han venido clasificando en dos grandes grupos (Ferrández, 1998): sistemas *integrados* y sistemas *alternativos*.

Los sistemas integrados agrupan una serie de conocimientos (fuentes de información) que se suponen necesarios para resolver la anáfora. Este tipo de estrategia se basa en la imitación del comportamiento humano, es decir, se pretende dotar al ordenador con las mismas fuentes de información que tiene el humano cuando resuelve la anáfora que ha generado su pareja en la comunicación. Estas fuentes de información se pueden fusionar por medio de una serie de mecanismos de inteligencia artificial para que todas ellas contribuyan a proporcionar finalmente una elección (estos sistemas se conocen como *democráticos*). O también pueden utilizar una única fuente de información (que se conocen entonces como sistemas *consultivos*).

Los sistemas alternativos, a diferencia de los anteriores, no intentan imitar las fuentes de conocimiento humano, sino que se basan en el estudio de corpus a través de herramientas estadísticas de forma que la información para la toma de decisiones se extrae mediante técnicas de aprendizaje de los textos usados para el entrenamiento.

En este trabajo se ha utilizado una estrategia integrada democrática desarrollada mediante un mecanismo de restricciones y preferencias. Este mecanismo define una serie de reglas de obligado cumplimiento por cualquier candidato a antecedente de una anáfora (restricciones), así como una serie de reglas cuyo cumplimiento no es necesario, pero que en caso de cumplirse destacan a ese candidato respecto a aquellos que no las cumplen (preferencias). De esta forma, las restricciones se usan para descartar candidatos imposibles y las preferencias para ordenar los candi-

datos. El candidato que alcance mejor posición en esa ordenación es elegido como antecedente.

Por último destacaremos que para la resolución de la anáfora en este sistema se ha considerado necesaria la confluencia de varias fuentes de información. Estas fuentes de información han sido clasificadas en dos grandes grupos: fuentes de información lingüística y estructural.

- Fuentes de información lingüística: Hemos agrupado en este apartado aquellas fuentes de información basadas en las reglas lingüísticas.
 - Fuentes de información léxica: proporcionan información sobre ciertas palabras o grupos de palabras tal como su valencia causal verbal implícita. Esta valencia causal es la que provoca que en la oración del ejemplo 2, sea *Ana* y no *María* el antecedente del pronombre. Esto cambiaría si el verbo *vencer* se sustituye por *perder*, o si el adjetivo *mejor* se sustituye por *peor*.

(2) Ana_i venció a María porque **ella_i** fue mejor

Este tipo de información se suele proporcionar a través de un analizador-etiquetador léxico que extrae la información de un diccionario y la adjunta como etiqueta a cada palabra del texto.

- Fuentes de información morfológica: proporcionan información sobre el número, género y persona que debe contener el antecedente para ser compatible con la anáfora. Esta información se genera gracias al analizador-etiquetador morfológico que la incluye como una etiqueta junto a las palabras del texto. Generalmente, las fuentes de información léxica y morfológica se generan conjuntamente en la misma herramienta proporcionando una única etiqueta para cada palabra.
- Fuentes de información sintáctica: proporcionan los distintos constituyentes sintácticos necesarios para el proceso de la resolución de la anáfora. Esta información se extrae de los analizadores sintácticos que pueden generar información completa o

parcial tal y como se vio anteriormente. Entre otras aplicaciones se puede destacar la definición de paralelismos sintácticos que permiten expresar la compatibilidad entre anáforas y antecedentes.

- Fuentes de información semántica: proporcionan ciertas informaciones como pueden ser los roles semánticos de los antecedentes dentro de una oración (agente o paciente), o las características semánticas que deben tener el sujeto, el objeto directo o el objeto indirecto de un determinado verbo. Esta información se puede obtener de herramientas basadas en recursos semánticos (Saiz-Noeda & Palomar, 2000) tales como WordNet Español, una base de datos léxica que incorpora relaciones semánticas (Miller *et al.*, 1993).
- Fuentes de información estructural: que no dependen tanto de las reglas lingüísticas que rigen un discurso, sino de la forma en la que se construye el discurso. Esta información se extrae generalmente de la anotación que se genera sobre un texto e indica cada una de las partes de dicha estructura. Esta anotación podría realizarse de forma automática mediante las herramientas pertinentes. Una de las mayores utilidades de las fuentes de información estructural es que permiten especialmente la definición de *un espacio de accesibilidad anafórica* entendido como el espacio en el que se pueden extraer los antecedentes para una determinada anáfora. Este espacio es fundamental para la resolución anafórica en diálogos tal y como se demostrará en los próximos capítulos.
- Otras fuentes de información: entre las que destacamos aquellas que proporcionan un conocimiento general del mundo, es decir, aquellas que indican por ejemplo, que existe una correferencia entre el *Presidente del Comité Olímpico Internacional* y la persona *Juan Antonio Samaranch*, y que además esto se produce en un determinado contexto temporal (no se podría determinar por ejemplo en un texto fechado en 1947).

En concreto, para el sistema definido se usarán las siguientes fuentes de información lingüística:

- Información léxica y morfológica extraída de un diccionario y aplicada por un analizador y etiquetador léxico-morfológico.
- Información sintáctica que proporciona un analizador sintáctico parcial. Este tipo de analizadores proporciona las estructuras sintácticas de aquellos constituyentes que se necesitan para la resolución de la anáfora.

Además, junto a estas fuentes se usará la siguiente fuente de información estructural:

- Información estructural generada por la propia estructura del diálogo que extrae de ella un espacio de accesibilidad anafórica donde las anáforas encuentran su antecedente. Además, una ordenación de preferencias sobre los posibles lugares donde puede encontrarse el antecedente dentro de esa estructura permitirá obtener preferencias válidas para la elección del mejor candidato. Cuando el sistema de resolución de anáfora está integrado en un sistema de diálogo, esta información se obtiene de forma automática mediante las herramientas definidas para la propia gestión del diálogo. Sin embargo, si se quiere evaluar el sistema de resolución de la anáfora de forma independiente, la información estructural se obtendrá tras un proceso manual de etiquetado de la estructura del diálogo. Este es el caso que se trata en el trabajo presentado.

El sistema propuesto se plantea como un sistema de resolución de la anáfora en diálogos independientemente del dominio en el que se aplique. Puesto que no existen en español herramientas independientes del dominio que proporcionen la información semántica necesaria para la resolución de la anáfora, esta fuente de información no ha sido contemplada en nuestro desarrollo. Se relegará por tanto al sistema de diálogo o a su aplicación final, ya dependiente del dominio, el uso de estas fuentes para completar la información que sobre la resolución de la anáfora proporcionará nuestro sistema.

1.3 Objetivos

El objetivo principal de este trabajo es la construcción de un sistema de resolución de la anáfora en diálogos independiente del dominio sobre el que se aplique. Para la consecución de este objetivo se plantean las siguientes líneas de actuación:

- Estudio de los sistemas de diálogo como aplicación del sistema de resolución de la anáfora.
- Estudio y desarrollo de estrategias de resolución de la anáfora basadas en el uso de restricciones y preferencias. Para ello se tomarán como hipótesis inicial los trabajos previos de resolución de la anáfora en español.
- Estudio de la influencia de la estructura del diálogo en la resolución de la anáfora, y propuesta de un espacio de accesibilidad anafórica basado en esta influencia.
- Desarrollo de mecanismos que permitan mantener la independencia del sistema respecto al dominio.
- Definición del ámbito de trabajo, en concreto para las anáforas pronominales (generadas por pronombres personales y demostrativos) y las anáforas adjetivas, como anáforas más frecuentes en los diálogos.
- Desarrollo de un proceso de entrenamiento del sistema que garantice la mejor estrategia de resolución anafórica al margen de su evaluación final.
- Desarrollo de un proceso de evaluación del sistema que deberá realizarse totalmente independiente de su aplicación final, con el fin de obtener resultados debidos exclusivamente al sistema de resolución de la anáfora estudiado y no a su integración en la aplicación de diálogos.

Junto con estas líneas de actuación, se tendrá en cuenta que el sistema debe quedar abierto para plantear estrategias de resolución sobre otros tipos de anáfora como la anáfora superficial numérica, la anáfora por omisión del sujeto pronominal, la anáfora de sintagma verbal o las descripciones definidas. Además, aunque el fin principal del sistema sea su integración en sistemas

de diálogo, su aplicación no debe quedar sólo limitada a este tipo de sistemas. Cualquier aplicación computacional que requiera la comprensión, interpretación y/o generación del lenguaje natural en diálogos puede ser receptora de un módulo de resolución anafórica.

1.4 Organización de la Tesis

Este trabajo se ha estructurado en nueve capítulos:

- Este primer capítulo es una introducción al problema del procesamiento del lenguaje natural, en concreto al problema del tratamiento de la anáfora en diálogos, en el que se presentan los objetivos principales así como la estructura de esta memoria.
- El segundo capítulo presenta el diálogo como el ámbito que abarca nuestro trabajo. En él se muestra una taxonomía de los diálogos según su capacidad para generar fenómenos lingüísticos, y se centra finalmente nuestra investigación en los diálogos manejados por los sistemas de diálogo. Además se presenta la estructura del discurso en estos diálogos como una fuente de información esencial para el estudio de dichos fenómenos lingüísticos.
- En el capítulo tercero se presenta la anáfora como el objetivo a investigar ya dentro del ámbito de los diálogos. Así se analiza en profundidad el problema de su resolución mediante la presentación de una taxonomía de los distintos tipos de anáfora que se pueden encontrar en los diálogos, y se matizan aquellos tipos en los que se centrará nuestro trabajo. Además se realiza una revisión de algunos de los métodos para la resolución de la anáfora más característicos con especial énfasis en su aplicabilidad a los diálogos. Por último se presenta la forma de integrar el módulo de resolución de la anáfora y sus fuentes de conocimiento en un sistema de diálogo.
- En el capítulo cuarto se presenta uno de los pilares fundamentales de esta investigación: nuestra propuesta de definición de un espacio de accesibilidad anafórica (el espacio de la solución) basado en la propia estructura del diálogo que se genera a partir de un sistema de diálogo. Sin embargo, y puesto que el sistema

de resolución de la anáfora presentado será evaluado independientemente de su aplicación final (el sistema de diálogo), se define también un esquema de anotación para la obtención de la estructura que genera el espacio de accesibilidad anafórica. Además, se realiza un estudio empírico que destaca los beneficios del uso de un espacio de accesibilidad anafórica basado en la estructura del diálogo frente a otros esquemas tradicionales.

- El capítulo quinto presenta las principales aportaciones de nuestra investigación: en concreto, el sistema de resolución de la anáfora en diálogos (ARIADNA). Este sistema, incluido en aquellos que utilizan una estrategia integrada democrática, está basado en restricciones y preferencias lingüísticas y estructurales. Este conjunto de restricciones y preferencias se aplican sobre el espacio de la solución, y mientras las restricciones se consideran absolutas y descartan candidatos, las preferencias seleccionan al mejor candidato tras una ordenación de todos los posibles.
- El capítulo sexto incluye la evaluación completa de este sistema independientemente de su aplicación final así como todo el proceso de entrenamiento que ha llevado el sistema hasta obtener la mejor configuración posible. Este entrenamiento se realiza en dos fases: en la primera fase se busca la definición del mejor conjunto de preferencias y se evalúa la utilidad de cada una de ellas; la segunda fase trata de definir la importancia óptima que se le asigna a cada preferencia para producir la mejor elección del candidato final. Para cada uno de estos entrenamientos se han usado definiciones diferentes del sistema de preferencias tal y como se mostrará. Así mismo se muestran los resultados de la evaluación realizada sobre un corpus de evaluación con independencia del proceso de entrenamiento, lo que garantiza la fiabilidad de los resultados obtenidos. Además se realiza una exposición y estimación de las posibles causas de fallos junto con propuestas para su mejora.
- El capítulo séptimo muestra las conclusiones y beneficios de este trabajo junto con una propuesta de trabajos futuros.

- Por último se presenta una relación bibliográfica de los trabajos que se han utilizado como referencia para el desarrollo de esta memoria.

Universitat d'Alacant
Universidad de Alicante



Universitat d'Alacant
Universidad de Alicante

2. El diálogo y su estructura

Universitat d'Alacant
Universidad de Alicante

En este capítulo se pretende centrar el ámbito global de la memoria: los diálogos. Para ello, en primer lugar se presenta una taxonomía de los distintos tipos de diálogo atendiendo a su capacidad para generar fenómenos lingüísticos. En segundo lugar, presentaremos los sistemas de diálogo como una de las principales tendencias actuales de investigación en el tratamiento automático de los diálogos, lo que permitirá centrar nuestro estudio en aquellos tipos de diálogo objeto de estos sistemas. Finalmente, de acuerdo con una de las propiedades fundamentales de los diálogos, su estructura del discurso, presentaremos un estudio sobre segmentación y estructuración del discurso que nos permitirá analizar y comprender la generación de fenómenos lingüísticos en los diálogos.

2.1 Taxonomía de los diálogos

Las necesidades de los investigadores del comportamiento del diálogo han generado diversas tendencias de estudio que atienden fundamentalmente a las características de los distintos tipos de diálogo a tratar. Sin embargo, cada tipo de diálogo presenta unas características propias en cuanto a la generación y resolución de fenómenos lingüísticos entre los que se incluye la anáfora. De esta forma, comenzaremos nuestro estudio sobre los diálogos presentando una taxonomía de los distintos tipos de diálogo según sus características en cuanto a la generación y resolución de anáforas propuesta por Dahlbäck (Dahlbäck, 1997), para posteriormente centrar la investigación en aquellos tipos de diálogo objeto de

nuestro sistema de resolución de la anáfora. Esta taxonomía clasifica los diálogos atendiendo a diversos aspectos.

2.1.1 Modalidad: *Diálogos escritos (tecleados) y hablados*

Actualmente, según Dahlbäck, existe una cierta tendencia investigadora que desplaza el interés por el estudio de los sistemas de diálogo teclado hacia un mayor interés por los sistemas de diálogo hablado. Y puesto que existe una gran cantidad de resultados empíricos y modelos para los sistemas de diálogo teclado, el autor se plantea si estos resultados y modelos podrían extenderse aplicándose a los sistemas de diálogo hablado.

Las tendencias de muchos investigadores en este campo parecen indicar que las teorías lingüísticas desarrolladas no son relevantes para los sistemas de diálogo hablado, y para ello argumentan que el lenguaje hablado es diferente al escrito, que las teorías lingüísticas actuales se desarrollan para el lenguaje escrito y como consecuencia habría que empezar más o menos desde cero. Sin embargo, Dahlbäck argumenta que muchas de las características que se le atribuyen al diálogo escrito no dependen tanto del medio (ser escrito) sino de ser un diálogo. Y concluye que si bien existen diferencias entre ambos, éstas son bien conocidas y por el contrario existen muchas similitudes independientemente del medio de comunicación usado.

Hay un alto número de estudios de interacción hombre-máquina que usan lenguaje natural teclado, y que muestran cómo estos diálogos contienen muchos de los atributos característicos del diálogo hablado. Sin embargo, también existen diferencias entre los diálogos teclados y hablados.

Cohen (1984) estudió los efectos del canal de comunicación (conversaciones telefónicas frente a conversaciones con teletipo) donde encontró que la interacción del teclado hace énfasis en el empaquetamiento óptimo de la información en un espacio lingüístico más pequeño, y esto afecta a la organización del discurso. Por ejemplo, se encontró que en un 90% la introducción de un nuevo segmento en lenguaje hablado se realiza por el uso de

marcadores del discurso (*cue-words*). Sin embargo, en los escritos es menor del 45%.

Otra de las diferencias concierne a la estructura del diálogo. Dahlbäck muestra cómo modelos del diálogo desarrollados para la recuperación de información en diálogos hablados y otros desarrollados con el mismo fin para diálogos tecleados llegan a conclusiones distintas en cuanto a la estructura. Los diálogos hablados estudiados por Bilange (1991) muestran una estructura de tres movimientos (Negociación, Reacción y Elaboración), mientras que los tecleados, estudiados por Dahlbäck muestran que una estructura con dos movimientos (Iniciativa y Respuesta) es suficiente.

Finalmente, Dahlbäck destaca una diferencia importante entre los diálogos hablados y los tecleados que afecta al discurso, y es que en los diálogos tecleados, algunas partes del propio diálogo permanecen frente al usuario mientras se planea el siguiente movimiento. De esta forma, se encuentra que en este tipo de diálogo, incluso después de que haya pasado largo tiempo desde que se emitió el antecedente, los usuarios hacen uso de expresiones anafóricas incluso elipsis, ya que físicamente pueden ver el antecedente. Con la terminología de Grosz (1977), podríamos decir que lo tienen físicamente en su foco de atención.

2.1.2 Clases de agentes: *hombres y máquinas*

Dahlbäck propone que el hecho de tener como pareja de diálogo a una persona o a un ordenador influye en el número de aspectos lingüísticos, como por ejemplo el uso de pronombres. La interacción con un ordenador hace que la estructura del diálogo sea más simple: los enunciados son más cortos, la variación léxica es menor y se minimiza el uso de pronombres, aunque en muchos casos es imposible discernir si esto ocurre por influencia del canal (teclado *vs* hablado) o por las características del agente (humano *vs* ordenador).

Otra conclusión a la que llega el autor es que en tareas de recuperación de información, la estructura de los diálogos tecleados que tienen ordenadores como agentes es bastante simple, lo que

permite modelarla con una gramática de diálogo independiente del contexto. Esto es cierto también para los diálogos hablados con ordenadores. Sin embargo, la estructura de los diálogos entre personas es más compleja.

2.1.3 Contexto compartido: *espacial - temporal*

El espacio y el tiempo que comparten los participantes parecen influir en ciertos aspectos del discurso. Por ejemplo, en un corpus de diálogo sobre el funcionamiento de un determinado producto se encontró que los pronombres, o tienen su antecedente en el subdiálogo actual (un diálogo abierto para aclarar o matizar algo del diálogo general), o se refieren al producto que estaba físicamente presente en todo el diálogo y que apenas se menciona explícitamente. Sin embargo, en otras clases de discurso donde no hay un contexto físico compartido (de espacio y/o tiempo) y donde se minimiza la interacción, ocurre a veces que hay ciertas entidades privilegiadas, llamadas referentes primarios, que pueden ser referidas mediante el uso de un pronombre incluso si su antecedente no ha sido mencionado durante un largo tiempo: por ejemplo, los personajes principales en una novela.

2.1.4 Distancia de la tarea del diálogo: *orientados a tareas y de recuperación de información*

Si bien es conocido que la estructura de la tarea influye en la estructura del diálogo, también es cierto que existe un conjunto continuo de variaciones desde los diálogos de sobremesa totalmente desestructurados hasta los diálogos de tareas altamente estructurados. En los diálogos orientados a tareas, la estructura de la tarea influye sobre la estructura del diálogo y como consecuencia se ha incluido en el análisis del discurso. Así diferentes tareas generarán diferentes estructuras del discurso.

Además, Dahlbäck, basándose en la observación propone otro factor influyente en la estructura del discurso que llama la distancia de la tarea (*dialogue-task distance*). En diálogos orientados a tareas sabemos que es necesario entender los actos y tareas no

lingüísticas para responder correctamente a los actos de habla del compañero. Sin embargo, en recuperación de información esto no es necesario. Por ejemplo, si alguien pregunta por los trenes que salen hacia una ciudad a una hora, no hace falta saber para qué quiere ir a ese sitio para proporcionar una respuesta correcta. Conforme se hace menor la distancia entre el lenguaje y la tarea subyacente (*short dialogue-task distance*), más apropiado se hace tener un modelo basado en planes o intenciones. Conforme se hace mayor esta distancia (*long dialogue-task distance*), más difícil se hace la inferencia de intenciones subyacentes desde la estructura lingüística, y a la vez, más inútil se hace dicha información para proporcionar respuestas válidas.

2.1.5 Número y tipo de las tareas

Cuando crece el número de tareas de un diálogo se hace mucho más compleja la determinación del tópico del diálogo y por lo tanto la obtención de un modelo basado en planes o intenciones. En este caso, los diálogos hombre-máquina presentan como ventaja que pocas cosas pueden ser el tópico de ese diálogo, frente a la complejidad del diálogo entre humanos.

2.1.6 Clases de conocimiento compartido: *perceptual, lingüístico y temporal*

Para que una referencia se pueda usar con éxito es necesario que tanto el hablante como el oyente compartan conocimiento mutuo, creencias y suposiciones. Clark (1985) propone la existencia de tres clases de información básica que pueden ser usadas para inferir el conocimiento común entre hablante y oyente: el conocimiento compartido perceptual, lingüístico o cultural.

- el perceptual se produce cuando se comparte contexto visual o físico,
- el lingüístico se produce cuando comparten conocimiento de textos o diálogos anteriores,
- por último, el cultural, se usa en el sentido más amplio incluyendo formas de vestir, de hablar, etc.

En este sentido, las capacidades de inferencia de los ordenadores son mucho menores que las de la persona, no sólo porque tenga menor capacidad de deducción, sino porque que no son capaces de ver a sus interlocutores, ni los pueden recordar de encuentros anteriores; etc, es decir, pierden la potencia del conocimiento compartido.

Nuestro estudio se centrará especialmente en los diálogos teclados hombre-máquina puesto que el objetivo final es su aplicación a un sistema de diálogo.

2.2 Los sistemas de diálogo

Los sistemas de diálogo, constituyen una de las aplicaciones más inmediatas para el sistema de resolución de la anáfora en diálogos que se desarrolla en este trabajo. Su integración en estos sistemas de diálogo permite la comunicación hombre-máquina mediante una adecuada interpretación y generación del lenguaje.

La investigación para el diseño de sistemas de diálogo constituye una de las líneas más clásicas de la Lingüística Computacional y que sin embargo hasta hace unos pocos años no había generado resultados realmente aplicables. Esto ha sido debido, en parte, a que se ha necesitado del desarrollo de productos pertenecientes a diversas áreas tecnológicas para poder darle una utilidad real a toda la investigación que hasta el momento se había realizado sobre la comprensión y generación de conversaciones hombre-máquina.

En el trabajo realizado por Fernández (Fernández, 2000) se muestran las tres tendencias de investigación actuales en el campo de los sistemas de diálogo visto desde las tres grandes perspectivas de la Inteligencia Artificial. Desde el punto de vista *metodológico*, dicha investigación pretende una obtención de modelos de diálogo hombre-máquina así como su formalización teórica. Desde el punto de vista *cognitivo* se busca la motivación de los modelos propuestos. Y por último, desde el punto de vista *computacional* lo que se pretende es la implementación de los modelos de gestión de diálogo, y su integración en otros componentes de procesamiento del lenguaje natural como el reconocimiento y generación de voz o

la interpretación y generación del lenguaje natural. Estos componentes deben incluir la resolución de sus fenómenos lingüísticos.

En concreto, en este apartado nos centraremos en mostrar las necesidades de esta última línea de investigación, y haremos un especial énfasis en la necesidad de crear un módulo de procesamiento de lenguaje natural que resuelva e interprete los fenómenos lingüísticos.

2.2.1 Arquitectura de los sistemas de diálogo

Existe un criterio tradicional sobre el desarrollo de sistemas de diálogo por el cual éstos se construyen generalmente a partir de una arquitectura básica formada por tres niveles que se puede extraer del modelo básico conversacional definido por Bernsen *et al.* (1998). De acuerdo a este modelo la arquitectura básica de un sistema de diálogo debe contener una serie de módulos distribuidos en los siguientes niveles tecnológicos:

- Procesamiento del habla (reconocimiento y síntesis).
- Procesamiento del lenguaje natural (comprensión y generación).
- Gestión del diálogo (control y contexto).

El nivel de procesamiento de habla contiene los módulos de reconocimiento y síntesis del habla que permiten al sistema de diálogo realizar una conversión entre el lenguaje hablado introducido en el sistema por el usuario humano y el lenguaje escrito que será el analizado y procesado por la máquina. En este sentido existe una doble tarea a procesar. El sistema debe ser capaz de reconocer el lenguaje hablado para interpretar las entradas del usuario pero también deberá ser capaz de sintetizar el lenguaje hablado para obtener una salida para ese usuario.

El nivel de procesamiento del lenguaje natural describe el procesamiento léxico, gramatical y semántico que es necesario, por una parte, para obtener la información básica que se usará en la gestión del diálogo mediante la comprensión del lenguaje, y por otra parte, para tratar los aspectos lingüísticos de la generación del lenguaje.

Por último, el nivel de gestión de diálogo se encarga de controlar la estructura del diálogo desde dos aspectos distintos. Por un lado, debe modelar la estructura del diálogo desde el mantenimiento de la estructura intencional, el estado atencional y la estructura lingüística (Grosz & Sidner, 1986). Y por otro lado, debe encargarse del mantenimiento de un contexto en el que se incluya la historia del diálogo, un modelo del usuario, un modelo del dominio, o las características del diálogo. De esta forma, el nivel de gestión del diálogo permitirá llevar a cabo los fines de dicho diálogo, es decir, la ejecución de una tarea concreta si se trata de un sistema de diálogo orientado a tareas, o la obtención de una información si estamos ante un sistema de recuperación de información.

En el esquema básico descrito por Bernsen *et al.* (1998) la información fluye secuencialmente de un nivel a otro, de forma que la entrada del sistema se produce a través de un módulo *reconocedor de habla* que produce una salida escrita hacia un módulo de *comprensión del lenguaje natural* que analiza el lenguaje y lo interpreta, y produce la entrada del módulo *gestor de diálogo*. Este módulo finalmente genera la estructura controlada provocando que el módulo de *generación del lenguaje natural* construya una salida de lenguaje que debe ser sintetizado en forma de voz pseudo-humana gracias al módulo *sintetizador de voz*. De esta forma, se crea una cierta independencia entre los tres niveles tecnológicos, enlazados únicamente por el flujo de información que constituye las entradas y salidas entre los distintos niveles.

Sin embargo, trabajos recientes publicados por Allen *et al.* (2000), demuestran que el mantenimiento de una arquitectura con este flujo de información simple no es viable cuando se pretende construir un sistema capaz de gestionar diálogos de cierta complejidad. En este sentido, Allen *et al.* proporcionan una arquitectura genérica para los sistemas de diálogo por medio de la cual, existe un flujo principal de información desde la entrada de un enunciado hablado hasta la generación de la respuesta por parte de la máquina, pero además existe un flujo de información secundario entre los diferentes componentes que remarcan la necesidad

de una interacción entre los distintos niveles más allá del flujo principal.

La figura 2.1 muestra el flujo de información entre cada uno de los componentes de la arquitectura propuesta por Allen *et al.*

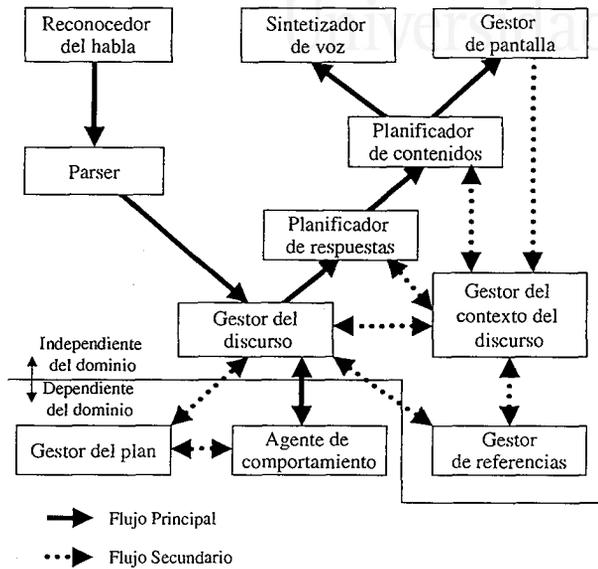


Figura 2.1. Arquitectura de un sistema genérico de diálogos por Allen *et al.* (2000)

En la figura mostrada, se distingue mediante líneas de trazo continuo el flujo de información principal que corresponde con el comportamiento descrito en el modelo básico de Bernsen *et al.* Sin embargo, existe además un flujo de información secundario representado mediante líneas de trazo discontinuo que permite una interacción constante entre los distintos niveles del sistema, y que rompe con el esquema cerrado basado en tres niveles independientes.

2.2.2 Componentes de un sistema de diálogo genérico

De acuerdo a la arquitectura planteada por Allen *et al.* se distinguen los siguientes componentes en un sistema de diálogo:

- Reconocedor del habla: Capaz de transformar la entrada hablada procedente de un humano en un secuencia o trama de palabras escritas. Actualmente, existe un amplio abanico de productos de propósito general capaces de interpretar el habla. La mayoría de estos productos se basan en el empleo de tácticas de concordancia de patrones y de modelos estocásticos alimentadas por diccionarios fonéticos y corpus hablados. Sin embargo la generalización de estos productos provoca salidas no exentas de errores considerables que impiden el correcto funcionamiento del resto del sistema. Esto hace que los sistemas de diálogo entrenen usualmente dichos sistemas para conseguir mejores prestaciones dentro de sus dominios específicos.
- *Parser*: Donde el término *parser* engloba el analizador léxico, morfológico, sintáctico y semántico que convierte la salida del Reconocedor de habla en interpretaciones, cada una de ellas formada por un conjunto de actos de habla convencionales gracias a técnicas robustas de análisis. El análisis semántico juega un papel fundamental en este módulo puesto que proporciona la información que servirá posteriormente para identificar las intenciones de los actos de habla, y además puede distinguir el tema, el agente y el instrumento en cada uno de ellos.
- Gestor del discurso: Tiene una doble misión: por una parte, reconocer las intenciones del acto del habla emitido por el usuario y por otra, determinar cuáles son las obligaciones que va adquiriendo el sistema conforme se desarrolla el diálogo. Es decir, qué información se debe buscar (en sistemas orientados a la recuperación de información) si es que se le ha preguntado por algo o qué tarea se debe realizar (en sistemas orientados a la tarea) si es que se le ha solicitado una tarea. Para poder efectuar su trabajo, el Gestor del discurso coordinará una serie de procesos que le proporcionarán la información necesaria. Por una parte, recibirá la entrada de los actos de habla y su información asociada generada por el *parser*. Por otra parte, una serie de reglas de interpretación actuarán sobre dichos actos de habla invocando a los módulos Gestor de referencias, Gestor del contexto, Gestor del plan, y el Agente de comportamiento para producir

y evaluar las interpretaciones posibles que se le den al acto de habla.

- Gestor del contexto del discurso: Se encarga de mantener el contexto del discurso a nivel global, mediante el desarrollo del tópico, y a nivel local, mediante el desarrollo de subtópicos que contribuyen al desarrollo del tópico.
- Gestor de las referencias: La interpretación de las referencias forma parte de la fase de interpretación semántica llevada a cabo por el Gestor del discurso. El Gestor de referencias tiene como misión la identificación de referencias y la resolución de dichas referencias usando información independiente del contexto. En este sentido, Allen *et al.* apuestan por módulos de resolución parcial de la referencia. Es decir que proporcionen una lista de posibles candidatos según la información independiente del dominio que manejan, más que proporcionar una única salida. Así dotará al Gestor del discurso de la capacidad para elegir la interpretación más adecuada para dicha referencia según la información que otros procesos dependientes del contexto le proporcionen. Además, el propio Gestor del discurso puede emitir actos de diálogo destinados a la clarificación de la anáfora si existiera ambigüedad en la resolución de la misma. Este tipo de actos pueden incluir preguntas directas al usuario para confirmar si la solución propuesta es realmente acertada o no.

Por otra parte, el Gestor de las referencias puede ser consultado también para la generación de las referencias en la salida del sistema. Esto contribuye a la producción de un lenguaje mucho más natural.

- Agente de comportamiento: Determina cuáles deben ser las acciones del sistema ante un determinado estímulo de entrada. El Agente de comportamiento actúa como consultor del Gestor del discurso independientemente de la conversación actual pero con dependencia del dominio en el que se encuentra. Este agente propone acciones, planes y objetivos al Gestor del discurso que será quien finalmente decida según la conversación concreta que se está desarrollando y del resto de procesos involucrados.

- Gestor del plan: El Gestor del plan ejecuta las interpretaciones que el Gestor del discurso ha determinado sobre las intenciones del diálogo. Es el órgano ejecutor de las tareas que se le han encomendado al sistema. Si se trata de una búsqueda de información deberá realizarla, si se trata de la ejecución de una tarea, deberá planificar cómo conseguirla.
- Planificador de respuestas: Determina cuáles son los actos de habla (y sus contenidos) que se deben emitir para llevar a cabo los objetivos y obligaciones adquiridas por el sistema.
- Planificador de contenidos: Construye los actos de habla que van a ser emitidos por el sistema.
- Sintetizador de voz: Módulo complementario al Reconocedor de habla, es capaz de generar pronunciaciones a partir del texto escrito. En muchas ocasiones los reconocedores y los sintetizadores de voz comparten las mismas herramientas (diccionarios fonéticos, modelos estocásticos de lenguaje hablado, etc.) por lo que puede ser habitual encontrarlos en un sistema conjunto. También, como ocurre con los reconocedores de habla, el sintetizador de voz mejora sus prestaciones cuando se entrena y adapta a dominios específicos.
- Gestor de pantalla: Gestiona las representaciones visuales de la salida del sistema (si es que existen).

2.2.3 La gestión del discurso en los sistemas de diálogo

De acuerdo con la arquitectura propuesta anteriormente, el módulo Gestor del discurso constituye la pieza esencial en los sistemas de diálogo. Este módulo tiene dos misiones fundamentales: interpretar la entrada proporcionada por el usuario y controlar el diálogo para generar la salida adecuada en cada caso. Para la realización de dichas tareas, el Gestor del discurso se apoya en una serie de procesos paralelos que le proporcionan información tanto dependiente como independiente del discurso, tal y como hemos visto en la sección anterior.

Sin embargo, la tarea de interpretación de la entrada que proporciona el usuario puede verse dificultada por la existencia de

una serie de fenómenos descritos por Fernández (2000) y que presentamos a continuación:

- Errores de reconocimiento de voz, producidos tanto por una falta de capacidad por parte del reconocedor de voz, como por fallos en la pronunciación por parte de los participantes.
- Errores en la interpretación semántica, producidos por una desambiguación incorrecta del sentido de algunas palabras.
- Expresiones deícticas y anafóricas, que no pueden ser interpretadas fuera de contexto.
- Elipsis, producida por razones de economía, estilo o énfasis y que en muchos casos deberá recuperarse del propio contexto del diálogo.
- Usos figurados del discurso gracias a un conocimiento del mundo real compartido por los participantes y que no debe ser interpretado de forma literal.
- Habla espontánea, que incluye el uso de dudas, falsos comienzos, elementos incompletos y rellenos que no constituyen actos de habla por sí mismos y que únicamente sirven para proporcionar cierta cohesión al diálogo y para reafirmar el papel de oyente.

La existencia de cierta dificultad para interpretar la entrada obliga al Gestor del discurso a modificar en ocasiones el comportamiento del diálogo, lo que genera actos de habla que permiten confirmar o clarificar los actos de habla emitidos por el usuario. Estas actuaciones que proporcionan cierta fiabilidad a las respuestas del sistema, cuando se producen en exceso pueden llevar a la construcción de diálogos antinaturales y poco flexibles, situación que no resulta deseable en un sistema de diálogo.

En definitiva, la interpretación de las entradas y la generación de las salidas es una tarea compleja que requiere entre otras cosas del mantenimiento de un equilibrio entre la falta de información para controlar el diálogo y la necesidad de proporcionar una respuesta natural y flexible. Para conseguirlo, los investigadores de los sistemas de diálogo emplean diversos métodos. Estos métodos se describen a continuación.

En el trabajo de Fernández (2000) se realiza una exposición de los distintos enfoques utilizados por los investigadores de sistemas

de diálogo para realizar la gestión del discurso. Estos enfoques se describen brevemente a continuación.

- Gramáticas de diálogo: Este enfoque usa modelos gramaticales que pueden estar basados tanto en gramáticas de estados finitos y de grafos, como en colecciones de reglas independientes del contexto. Dichos modelos se usan para representar y analizar la estructura del diálogo, y describen las secuencias más comunes del diálogo como los pares adyacentes tal y como se definen en Sacks *et al.* (1974). De esta forma, las gramáticas pueden predecir los elementos de una secuencia desde una secuencia anterior.
- Enfoques declarativos: Uno de los problemas de las gramáticas de diálogos radica en que cada transición posible debe ser especificada con anterioridad, y el número de transiciones crece exponencialmente a medida que crece el número de estados. Así cada posible pregunta puede ir seguida de cualquier otra, y en los sistemas complejos con un elevado número de preguntas posibles sería computacionalmente ineficiente mantener una red que albergue todas las posibilidades. Para evitarlo, los enfoques declarativos añaden a cada regla unas precondiciones que limitan el número de transiciones que pueden conducir a un determinado estado.
- Enfoques basados en marcos: Se basan la estructura temática o tópico. Los marcos se utilizan para modelar un conjunto jerárquico de entidades que constituyen la estructura temática del diálogo. Así los marcos se componen de una serie de casillas que constituyen los requisitos a conocer de cada una de estas entidades temáticas o tópicos. El gestor del diálogo se orienta a la producción de pares adyacentes destinados a rellenar las distintas casillas que permanecen vacías. Cuando el marco está completo se considera esa tarea finalizada, y por tanto alcanzado el tópico.
- Enfoques basados en planes: Tratan de modelar los objetivos de cada conversación incluyendo los sub-objetivos necesarios para conseguirlos, es decir, el plan. Una vez que se infiere cuál es el propósito del usuario, el sistema será capaz de reconocer incluso

usos figurados del discurso. Sin embargo, en este caso la complejidad se produce al intentar predecir cuál es el plan que un usuario tiene en cada momento. En muchos casos, no se consigue la predicción a tiempo y el diálogo puede quedar sin control.

- **Enfoques colaborativos:** Por último, los enfoques colaborativos surgen como una extensión de los modelos basados en planes. Sin embargo, a diferencia de éstos, no se basan en la identificación del plan del usuario como objetivo principal. Estos enfoques parten del hecho que en todo diálogo cada uno de los agentes participantes tiene sus propias intenciones y sus propias creencias sobre las intenciones del otro agente. Y los agentes emplean los actos de habla para representar esas intenciones y esas creencias. Así estos enfoques dan prioridad a las intenciones conjuntas sobre las particulares. El diálogo se modela por tanto como una actividad conjunta en la que los participantes se comprometen a entenderse uno a otro generando actos de habla para la confirmación, clarificación, etc.

Como conclusión, extraemos la idea de que existen diferentes enfoques para plantear un método que permita la gestión del diálogo. Cada uno de estos enfoques se define con una complejidad diferente y una serie de ventajas e inconvenientes que lo hacen más o menos adecuado según el tipo y la complejidad del sistema que se quiera modelar. Sin embargo, también extraemos la idea de que todos los enfoques se basan en el mantenimiento de una estructura del diálogo que se puede definir a nivel local, como en el caso de los enfoques basados en gramáticas, o a nivel global, como en los enfoques basados en planes o colaborativos. Pero, en cualquier caso, utilizan la producción de una secuencia de pares adyacentes como base para la construcción de las distintas estructuras.

2.2.4 Actualidad de los sistemas de diálogo

Existe actualmente una abundante proliferación de proyectos que tratan sobre el desarrollo de sistemas de diálogo. Gracias a estos proyectos se han estudiado las posibles tendencias investigadoras y se han evaluado las necesidades reales de dichos sistemas para

obtener las conclusiones mostradas en la sección anterior. Mostraremos a continuación algunos de los proyectos más significativos realizados en este campo, y que han servido como base para las hipótesis manejadas en esta Tesis.

Uno de los sistemas más significativos es el TRAINS (Allen *et al.*, 1995), construido en 1995 en la Universidad de Rochester. TRAINS es un sistema complejo y robusto orientado a tareas cuyo módulo gestor está basado en un enfoque colaborativo. La tarea principal de TRAINS es la búsqueda eficiente de rutas para trenes en el noreste de los Estados Unidos.

Basado en el TRAINS, el mismo grupo de la Universidad de Rochester desarrolla un nuevo sistema utilizando la arquitectura genérica definida en Allen *et al.* (2000) mostrada anteriormente. Como resultado se presenta TRIPS (*The Rochester Interactive Planning System*) (Ferguson & Allen, 1998), un sistema para la planificación de tareas portable a diferentes dominios. Una de las principales características de TRIPS es que contiene un módulo de resolución de la anáfora pronominal robusto independiente del dominio que proporciona al Gestor del Discurso una lista de las posibles soluciones de la anáfora. Finalmente, éste usará información de las herramientas dependientes del dominio para escoger la solución más adecuada.

Otro sistema de diálogo destacable es el Verbmobil (Alexandersson *et al.*, 1995), un sistema de traducción conversacional en diálogos de negocios que genera traducción simultánea al inglés, alemán y japonés. El sistema permite que varios agentes conversen sin que éste intervenga en el diálogo, reservándose la posibilidad de tomar el control sólo cuando se necesite aclarar algo.

Destacamos también el sistema de diálogo Philips (Aust *et al.*, 1995) como un sistema orientado a la recuperación de información basado en el uso de un enfoque declarativo para la gestión del diálogo. Este sistema proporciona información acerca del horario de trenes y posibilidades de conexión entre ciudades de Alemania.

Entre los trabajos españoles actuales destacamos el que se desarrolla en el proyecto BASURDE (*Sistema de diálogo para habla espontánea en un dominio semántico restringido*) (Proyecto BASURDE, 1998-2001), un consorcio formado por la Universitat

Politécnica de Catalunya, Euskal Herriko Unibersitatea, Universitat Jaume I, Universitat Politècnica de Valencia y Universidad de Zaragoza que estudia el comportamiento de un sistema de diálogo hombre-máquina orientado a proporcionar información sobre los servicios de una compañía de ferrocarriles. Estos trabajos son especialmente relevantes ya que trabajan sobre el mismo dominio del cual fueron extraídos los corpus empleados para evaluar nuestro sistema.

También destacamos el sistema TELEDELFO descrito en Fernández (2000), un sistema de diálogo para la gestión de un entorno de ayuda en operaciones telefónicas básicas. El sistema TELEDELFO utiliza la base del sistema DELFO, una herramienta definida para la implementación de diálogos inteligentes, robustos y eficientes en sistemas colaborativos de gestión del discurso dirigidos por voz (Fernández, 1999) (Quesada *et al.*, 2000). Tanto TELEDELFO como DELFO han sido desarrollados por los miembros del Grupo de Investigación Julietta del Departamento de Filología Inglesa de la Universidad de Sevilla y del Centro Informático Científico de Andalucía (CICA).

Por último, destacamos la existencia de una serie de proyectos definidos para el estudio, evaluación y propuesta de nuevas líneas de investigación en el campo de los sistemas de diálogo. Entre ellos destacamos el DISC (*Spoken Language Dialogue System and Components, Best practice in development and evaluation*) (Bernsen & Dybkjaer, 1995), un proyecto europeo en el que se evalúa la metodología empleada para la construcción de los sistemas de diálogo con el fin de mostrar las mejores tendencias para futuras líneas de investigación.

Siguiendo esta tendencia, el proyecto europeo TRINDI (*Task oriented Instructional Dialogue*) (Zaenen *et al.*, 2000) busca el desarrollo de una tecnología de diseño e implementación de sistemas de diálogo inteligentes y robustos, que pueda ser aplicada a diferentes dominios y lenguas.

En este sentido también destacamos, por último, el proyecto SIRIDUS en el que trabaja un consorcio con participación española formado por el SRI International de Inglaterra, la Universidad de Göteborg (Suecia), la Universidad de Saarland de Saarbrücken

(Alemania), la Universidad de Sevilla y el Grupo de Telefónica I+D. Este proyecto busca el desarrollo de herramientas computacionales para el diseño de sistemas de diálogo robustos, funcionales y con facilidad de uso.

2.3 Segmentación y estructuración del discurso en diálogos

Uno de los aspectos que diferencian fundamentalmente a los diálogos frente a los monólogos es la forma en la que se estructura su discurso. En este sentido, puesto que la estructura del discurso en los diálogos es pieza fundamental en el trabajo que estamos presentando, consideramos necesario hacer un estudio en profundidad de cómo se organiza y se segmenta esta estructura en los diálogos con los que trabajaremos.

Partiremos de la clasificación que hace Reynar (1999) de las distintas tendencias en investigación acerca de la estructuración del discurso en los diálogos. Reynar distingue dos categorías según su granularidad: las que se centran en la segmentación basada en el discurso (a nivel local) y las que se centran en la segmentación basada en el tópico del discurso (a nivel global).

2.3.1 Segmentación basada en el discurso

La segmentación basada en el discurso trabaja generalmente a nivel local de forma que centra su interés en la identificación de relaciones entre los enunciados (*utterances*). Definimos enunciado como la expresión oral o escrita de sintagmas y/o oraciones con sentido propio emitidas por el participante de un discurso.

Teoría de la estructura del discurso (Grosz & Sidner, 1986). Uno de los trabajos más representativos sobre segmentación del discurso es la teoría computacional sobre la estructura del discurso propuesta por Grosz y Sidner (Grosz & Sidner, 1986). De acuerdo a dicha teoría, la estructura del discurso, que en este caso se considera como una forma de lenguaje que puede implicar múltiples enunciados y múltiples participantes y que puede ser

tanto escrita como hablada, está compuesta por tres componentes distintos aunque relacionados entre sí: la estructura de la secuencia de enunciados (*estructura lingüística*), una estructura de propósitos (*estructura intencional*), y el estado del foco de atención (*estado atencional*). La estructura lingüística está formada por segmentos del discurso en los que los enunciados se enlazan de forma natural. La estructura intencional captura los propósitos del discurso expresados en cada segmento lingüístico y en las relaciones entre ellos. El estado atencional es una abstracción del foco de atención de los participantes conforme avanza el discurso. De acuerdo con las autoras, la distinción entre estos componentes es esencial para explicar las interrupciones, el uso de ciertos tipos de expresiones de referencia, y otras expresiones que afectan a la segmentación del discurso y a su estructura.

Esta teoría intenta modelar el comportamiento de los participantes del discurso en cada una de las estructuras:

1. Por un lado, cómo el participante conversacional que inicia un segmento (ICP) es capaz de indicar el principio y el fin de un segmento y cómo el otro participante no iniciador (OCP) es capaz de reconocerlo.
2. Cómo el OCP es capaz de reconocer las intenciones de un segmento del discurso.
3. Cómo opera el foco de los participantes en un segmento del discurso.

Una vez conocidos estos tres factores, se puede comprender cómo se segmenta el discurso en cada una de las tres estructuras: la lingüística, la intencional y la atencional respectivamente.

- *Segmentación en la estructura lingüística.* Según esta teoría, en la estructura lingüística se cumple la siguiente relación: los enunciados en un segmento, así como las palabras en una frase, cumplen roles particulares respecto al segmento. Así mismo, los segmentos como las frases cumplen ciertas funciones en el discurso completo.

Por otra parte, hay una interacción dual entre la estructura de segmentos del discurso y los enunciados que constituyen el discurso:

- En un sentido, las expresiones lingüísticas pueden ser usadas para definir la estructura del discurso. Cierta tipo de expresiones lingüísticas como palabras o frases clave, incluso ciertas entonaciones suelen ser usadas por los hablantes para introducir un cambio de segmento que el oyente será capaz de detectar. Este tipo de expresiones se conocen como *marcadores del discurso*.
- En el otro sentido, la estructura del discurso restringe la interpretación de las expresiones (y afecta a lo que el hablante dice y a lo que el oyente interpretará de esto). La estructura del discurso puede restringir el uso de ciertas expresiones lingüísticas como son las expresiones de referencia. Por ejemplo, los pronombres y los sintagmas nominales definidos no se usan en los límites entre segmentos mientras que sí suelen ser usados en el interior del segmento.
- *Segmentación en la estructura intencional*. Hay tres factores que influyen en la determinación de la intención o propósito de un segmento del discurso. Por un lado, los marcadores lingüísticos, además de marcar los límites de los segmentos pueden ayudar a reconocer el propósito del segmento de discurso en el que se encuentra. En segundo lugar, la intención de cada enunciado a nivel particular, ayuda a la detección de la intención del segmento. Y por último, cuando los marcadores lingüísticos, y las intenciones a nivel de enunciado no consiguen clarificar el propósito del discurso, sólo el conocimiento compartido del tópico puede determinarlo.
- *Segmentación en el estado atencional*. El estado del foco del discurso limita la información que debe considerarse para reconocer el propósito del segmento de discurso. Además, permite al OCP restringir la búsqueda de posibles referentes para sintagmas nominales definidos y pronombres. En este sentido, sólo aquellas entidades que se encontraran en el espacio del foco de atención podrían ser propuestas como referentes posibles. Basándose en esta idea, Grosz *et al.* (1983; 1995) plantean el *centering* como un mecanismo para la resolución del pronombre.

2.3.2 Segmentación basada en el tópico del discurso

Los trabajos definidos en la sección anterior estudian la segmentación del discurso desde el estudio del enunciado y sus relaciones con otros enunciados para la constitución de un segmento de mayor nivel. Sin embargo, los trabajos que se basan en la segmentación del tópico, identifican la coherencia respecto al tópico de bloques de texto constituidos por varias oraciones que se extienden a lo largo de diversos párrafos.

Existe cierta confusión sobre la definición y ámbito del término tópico. El término *tópico de discurso* (*topic of discourse*) o *tópico de conversación* (*topic of conversation*), entendido como tema general de un texto o discurso (sea del tipo que sea), fue planteado y desarrollado por Teun A. van Dijk (van Dijk, 1977) y traducido como tal en la versión al español (van Dijk, 1995). En este trabajo, el autor plantea que todo texto está organizado desde un punto de vista semántico (y en un plano muy abstracto) a partir de una unidad semántica (una proposición lógica) que contiene todo el desarrollo semántico del texto. Es decir, el *tópico del discurso* es lo que coloquialmente se denomina *tema de un texto*. A partir de este tópico, denominado más explícitamente *tópico general global*, un texto se estructura jerárquicamente en diferentes subtópicos y sub-subtópicos, en cada uno de los cuales se especifican los diferentes subtemas y sub-subtemas que componen un texto. Todo esto constituye la denominada *macroestructura* de un texto (o estructura profunda textual).

Este término, sin embargo, plantea un problema: al poder ser confundido con el término *tópico* aplicado a la oración, no al discurso. En un nivel oracional, el *tópico* es aquella parte de la oración que consta de información conocida por el receptor, a diferencia del *comento*: aquella parte de la oración con información desconocida por el receptor. Así se forma el par *tópico - comento*, también denominado *tema - rema*. En este sentido la definición coincidiría con el concepto de foco o del *center* usado en Grosz *et al.* (1983; 1995).

Para evitar confusiones, a nivel general, el término *tópico* suele ser usado en el sentido más amplio como el *tópico del discurso* o de

un segmento del discurso (coincidiendo en este caso con el concepto de propósito del discurso o del segmento de discurso estudiado por Grosz and Sidner (1986)). En nuestro caso utilizaremos esta última acepción para distinguir aquellos trabajos de segmentación centrados en un estudio a bajo nivel (relaciones entre constituyentes básicos) de aquellos que con el nombre segmentación del tópico estudian la segmentación a alto nivel.

En la línea de los trabajos sobre segmentación del tópico destacamos los siguientes:

Métodos basados en la localización de los primeros usos de palabras. Youmans (1991) se basa en la localización de los primeros usos de ciertos tipos de palabras. El autor postula que los núcleos de texto que contienen primeros usos de palabras generalmente introducen cambios de tópico ya que un cambio de tópico introduce generalmente una aportación de nuevo vocabulario.

Métodos basados en la cohesión léxica. Morris y Hirst (1991) se basan en relaciones de cohesión léxica mediante el uso de un *Thesaurus* capaz de identificar sinónimos.

Hearst (1994) desarrolló una técnica llamada *TextTiling*, algoritmo para la partición de textos en unidades de discurso multipárrafo que reflejan la estructura de subtópico del texto. El algoritmo usa información de frecuencia y distribución léxica para reconocer las múltiples interacciones de tópicos simultáneos. En este trabajo, Hearst plantea dos métodos distintos. El primero de ellos, para una ventana dada, compara la similitud léxica entre cada par de bloques de texto adyacentes. Así si los bloques son similares, es probable que el flujo del subtópico continúe. En cambio, si los bloques se hacen distintos, es bastante probable encontrar un cambio de subtópico. El segundo método le sigue la pista a cadenas de términos repetidos, determinando el flujo de subtópicos mediante marcas en el discurso que indican dónde termina el volumen de un conjunto de cadenas y dónde comienza un nuevo conjunto. En las pruebas efectuadas sobre 13 textos extraídos de una publicación popular sobre ciencia, el primer método alcanza una precisión del 64% con una cobertura (*recall*) del 58%. En cuanto al segundo método, la precisión alcanzada es de 66% junto con una cobertura del 61%.

Métodos basados en la repetición de las palabras. Reynar (1994) encuentra las fronteras que marcan los segmentos de tópicos basándose en la repetición de las palabras. Posteriormente, en Reynar (1999) se presenta una mejora del método anterior que incluye la frecuencia de repetición de pares de palabras mediante bigramas. De esta forma, si en una sección aparecen ocurrencias como *wild plant*, *native plant* y *woody plant* y en otra aparece *chemical plant*, *manufacturing plant* y *processing plant* el método detectará que las tres primeras no tratan el mismo concepto *plant* que la tres siguientes, y podría establecer una frontera de tópico entre ambas secciones. Además Reynar usa información sobre *domain cue words*¹, sinónimos, entidades relevantes, incluso el uso de los pronombres (entendiendo que la existencia de pronombres no es habitual al comienzo de un nuevo tópico). Con esta información Reynar evalúa su método sobre un fragmento del corpus HUB-4 en inglés, y alcanzó una precisión del 59% y una cobertura del 60% frente al 55% y 52% respectivamente alcanzado por el mismo método que usaba sólo la frecuencia de las palabras. Una versión del método de la frecuencia de las palabras alcanzó una precisión del 50% y una cobertura de 62% sobre la versión del corpus HUB-4 en español.

En Richmond *et al.* (1997) se diseñó un algoritmo para la segmentación del tópico que asigna peso a las palabras basándose en la frecuencia de aparición dentro de un documento y posteriormente se usan estos pesos en una fórmula basada en la distancia entre repeticiones de tipos de palabra.

Métodos basados en la búsqueda de candidatos a tópico. Rocha (1998) propone un método para la segmentación del diálogo basándose en la definición previa de candidatos a tópico. Según propone el autor, el diálogo puede dividirse en los siguientes componentes:

1. Fragmentos, originados por los tópicos globales del discurso.
2. Segmentos, originados por los tópicos locales del discurso.

¹ Se consideran como *domain cue words* ciertas palabras (marcadores del discurso) que típicamente, dependiendo del dominio, dan entrada a nuevas secciones en un texto.

3. Subsegmentos, originados por tópicos locales (llamados sublocales o subordinados) que complementan otro tópico local.

Rocha define un procedimiento previo para la búsqueda de candidatos a tópico global del discurso tomando como base que el tópico del discurso se define como el elemento de mayor importancia en un diálogo. Dicha importancia se mide en función de la frecuencia, la uniformidad de la distribución, la posición de la primera aparición y la adecuación semántica. Así el procedimiento completo para la búsqueda de candidatos a tópico global del discurso se define en los siguientes pasos:

1. Realizar un recuento de elementos aparecidos en el diálogo.
2. Seleccionar los cinco más frecuentes descartando pronombres, artículos, preposiciones y conjunciones, así como sintagmas nominales sin contenido semántico específico como *cosa*, *hecho*, etc.
3. Comprobar la distribución de dichos elementos a través del diálogo y seleccionar aquellos que tengan una distribución más uniforme como hipótesis de trabajo (aquellos que aparecen frecuentemente en un corto espacio del diálogo y ya no vuelven a aparecer, no son buenos candidatos).
4. Comprobar mediante un análisis contextual de corto alcance si el mismo elemento está enlazado a diferentes referentes y elegir aquellos que siempre tienen el mismo referente.
5. Comprobar si el referente de los elementos más frecuentes están referidos por varios elementos y analizar el efecto de esto en la frecuencia y distribución.
6. Considerar la frecuencia de los lemas de los elementos para el análisis de frecuencia y distribución.
7. Comprobar la posición de la primera aparición de los elementos y preferir aquellos más próximos al principio del diálogo.
8. Preferir elementos explícitos a sintagmas nominales genéricos que cubren varios elementos.
9. Preferir elementos que hacen referencia a objetos en lugar de elementos que hacen referencia a gente.
10. Comprobar manualmente la elección leyendo el diálogo completo.

Una vez efectuada la definición de la hipótesis de trabajo sobre los tópicos del discurso, se realiza un análisis que permite dividir el diálogo en fragmentos de acuerdo a estos posibles tópicos. En este sentido, la observación de la distribución y la posición de la primera aparición de ese posible tópico global determina el comienzo de un nuevo fragmento. El método a seguir para la división del diálogo en fragmentos es:

1. Analizar los elementos del diálogo con contenido léxico y estudiar la frecuencia, distribución y la posición de la primera aparición de los tópicos globales del discurso elegidos para detectar posibles fragmentos.
2. Los fragmentos cortos, con menos de 300 líneas, deben ser evitados, excepto si ocurren al principio o al final del diálogo.
3. Si un tópico global se retoma después de un paréntesis provocado por otro tópico global que no tiene nada que ver con el anterior, entonces se prefiere considerar todo como un mismo fragmento.

Además Rocha propone un método para la identificación de los tópicos locales y sub-locales:

1. Analizar los elementos con contenido léxico del diálogo y estudiar la frecuencia, distribución y concentración para detectar los tópicos potenciales para segmentos largos.
2. Analizar las frecuencias y distribuciones de los elementos mediante una ventana de 40 líneas con el fin de detectar tópicos locales, y añadir estos resultados a los del paso anterior.
3. Si fuera necesario, realizar conteos de frecuencias en ventanas más cortas.
4. Realizar un análisis manual de los intercambios para detectar movimientos o enunciados que puedan producir cambios de segmento, basándose en técnicas de coherencia como describe Sinclair (Sinclair, 1993).
5. Analizar dichos movimientos para determinar si introducen
 - a) un segmento, cuando el tópico local introducido está relacionado con el tópico global del discurso, pero no desarrolla un tópico de segmento anterior.

- b) un subsegmento, cuando el tópico local introducido desarrolla un segmento de tópico y está claramente subordinado a él.
- 6. Considerar que algunos tópicos locales no siempre aparecen como elementos definidos, sino que pueden ser *chunks* o trozos de discurso.
- 7. Preferir un nuevo segmento antes que un subsegmento de dudosa subordinación.

De esta forma Rocha propone un método semi-automático para la segmentación del discurso a nivel global con fragmentos, y a nivel local con segmentos y subsegmentos.

2.4 Anotación de la estructura del diálogo

El mantenimiento de una estructura del discurso propia es una de las características principales del discurso conversacional. Esta estructura puede alcanzar diversos niveles de complejidad con segmentos y subsegmentos dependientes de éstos tal y como se ha mostrado en la sección anterior. Así, los investigadores que trabajan en el estudio de los diálogos necesitan conocer dicha estructura hasta el nivel de complejidad que sea necesario en cada uno de los casos.

De esta forma, los investigadores que trabajan en la temática de los diálogos dedican un gran esfuerzo a la definición de esquemas de anotación de las estructuras del diálogo que les permitan justificar sus investigaciones. Así se muestran a continuación algunos de los trabajos más representativos en definición de esquemas de anotación, junto con una técnica para comprobar la fiabilidad de las anotaciones.

2.4.1 Esquemas de anotación

La definición de un esquema de anotación supone una formalización de la anotación del discurso, generalmente con la intención de ser procesada automáticamente por algún sistema informático.

Así, tradicionalmente, los autores interesados en el estudio de determinados sistemas de diálogo han propuesto esquemas de anotación que les sirven para sus propósitos concretos. Entre ellos, Alexandersson *et al.* (1995) definen un conjunto de tipos de actos de habla que pueden ocurrir en diálogos entre gente que intenta definir la fecha de una reunión de negocios, y los usan para hacer predicciones estadísticas sobre el acto que sigue al actual. En el caso de Ahrenberg *et al.* (Ahrenberg *et al.*, 1995), los autores dividen los movimientos de los diálogos de búsqueda de información *Mago de Oz* en comienzos y respuestas, y los clasifican posteriormente de acuerdo a la función que realizan en la transferencia de información, mostrando su relación con el foco del discurso.

A continuación se muestran algunos esquemas de anotación que se tomarán como referente para el desarrollo del esquema de anotación usado en este trabajo.

Anotación basada en las unidades de la estructura conversacional según Gallardo. Gallardo (Gallardo, 1996) estudia las distintas unidades que se manejan en la toma de turno conversacional según la siguiente escala de rango: la unidad básica es el acto, los actos se agrupan formando movimientos, los movimientos se agrupan para formar el turno y por último, éstos forman los intercambios o pares adyacentes. Esta definición, que se presenta a continuación, es una adaptación al español de los trabajos sobre análisis conversacional realizados por Sacks *et al.* (Sacks *et al.*, 1974).

Acto (act). Definido por Coulthard (1977), es la unidad mínima de la pragmática enunciativa que se encuentra tanto en el monólogo como en el diálogo. Sin embargo, no es una unidad interactiva.

Movimiento (move). Definido por Owen (1981), es la unidad dialogal mínima y hace referencia al movimiento que hace un jugador en el ajedrez. Son unidades en las que se fragmenta una intervención. Polanyi (1985) habla de él como *chunk*². Es el acto o conjun-

² En análisis sintáctico se considera el término inglés *chunk* para designar un fragmento de la oración que no llega a formar un constituyente sintáctico completo pero contiene al menos el núcleo de dicho constituyente y algunos de sus subconstituyentes. Esta acepción del término *chunk* se considera igual que la acepción sintáctica sólo que adaptado a la pragmática.

to de actos del habla dotado de valor interactivo. Un movimiento puede ser de tres tipos:

- Tipo 1 o movimiento de enlace retroactivo:
 - Prefacios: anticipan la orientación del movimiento de tipo 2 que le sigue (“por cierto que”, “es que”, “pero en fin”, “bueno pues”, “*well*”, “*ok*”, “*good*”, “*right*”, “*now*”, ...)
 - Preinicios (*pre-starts* o *appositional beginnings*): no proporcionan información sobre el contenido de la intervención (“pero”, “bueno”, “pues”, “sí”).
- Tipo 2 o movimiento constitutivo: Son actos ilocutivos.
- Tipo 3 o movimiento de enlace proyectivo:
 - Post-cierres (*post-completers*, *tag questions*): función directa de ceder el turno mediante la selección directa del hablante siguiente.
 - Prolongadores: no tratan de ceder el turno, sino mantenerlo.

Turno de palabra (turn). Según Gallardo (1996) los turnos son unidades de conversación dotadas de polaridad perceptiva natural, lo que significa que se identifican sin necesidad de reflexiones metalingüísticas: cada cambio de hablante supone un nuevo turno de habla. El turno de palabra tiene distintas características según el tipo de conversación que se trate. En un diálogo informal con personas bien conocidas entre sí y socialmente equivalentes es habitual el uso de interrupciones frecuentes. Sin embargo, en un diálogo formal se tomaría como una impertinencia. Según Gallardo hay dos tipos de turnos:

- *Intervención (IN)*: que forma el sistema primario de la conversación. Con la sucesión de interrupciones los hablantes favorecen el progreso informativo y temático de la conversación.
- *Aportación* o *Continuación (CT)*: son turnos vacíos, típicos del oyente, que responden a la ley del refuerzo formal y que ratifican el reparto de papeles conversacionales.

1. Las *aportaciones* o *continuaciones* pueden ser de dos tipos:
 - a) Con referencia ajena: hacen referencia a una intervención ajena y se pronuncian en solapamiento: “mm”, “ya”, “hum-hum”, “sí”, “claro”.

- b) Sin referencia ajena (o turnos de paso): no se produce en solapamiento, se usa para rechazar la posesión de la palabra, especialmente en secuencias de cierre: “y nada”, “y eso”, “pues nada”, “vale”...
2. La *intervención* ideal es una estructura tripartita con tres movimientos uno de cada tipo (Sacks *et al.*, 1974) (Stubbs, 1983), aunque esto no siempre se cumple. Más profundamente, una intervención puede ser clasificada en :

a) *Intervenciones iniciativas* IN_I :

- i. *Inicio* (I): (inicial y predictivo) se encuentra en la posición inicial y tiene carácter predictivo puesto que desencadena acciones inmediatas. Se trata de preguntas, invitaciones, ofrecimientos...

(3) IN_I (I) A: dame la maleta ³

IN_I (I) A: ¿Cuántos años tiene la tía?

- ii. *Relanzamiento* (RL): (predictivo) el inicio no triunfa y no logra lanzar su carácter predictivo, y el hablante lo reintroduce de nuevo.

(4) IN_I (I) A: ¿qué tal el examen?

IN_R (R) B: Uf, estoy cansadísima.

IN_I (RL) A: pero, ¿qué tal te ha salido?.

- iii. *Informe* (Inf): (inicial y final) se produce en la posición inicial pero no obliga a contestar al otro interlocutor. Las intervenciones de informe permiten al interlocutor *topicalizar*, es decir, convertir en tema o tópico de la conversación.

(5) IN_I (Inf) A: Estuve anteanoche con Eduardo.

³ Notación: en cada ejemplo expuesto se representa en primer lugar el tipo de turno, entre paréntesis el subtipo y por último el identificador del hablante.

48 2. El diálogo y su estructura

IN_{R+I} (R.Ev+I) B: Ah ¿sí?. ¿Y que tal está?.

b) *Intervenciones reactivas* IN_R:

- i. *Respuesta* (R): (predicha y puede ser final). Está desencadenada directamente por la intervención del hablante anterior y enlazada a él por relaciones de predicción.

(6) **IN_I (RL) A:** pero, ¿qué tal te ha salido?.

IN_R (R) B: creo que bien, ya veremos

- ii. *Respuesta/Inicio* (R/I): (predicha y predictiva). Reúne las dos cosas en un movimiento único.

(7) **IN_I (I):** ring.

IN_R (R/I) A: ¿diga?

IN_R (R/I) B: ¿Carmen?

IN_R (R/I) A: holaa

IN_R (R) B: buenaas

- iii. *Reacción Evaluativa* (R.Ev.): (final). Proporciona una valoración sobre el turno previo aunque éste no imponga restricciones de encadenamiento.

(8) **IN_I (Inf) A:** Estuve anteanoche con Eduardo.

IN_R (R.Ev) B: Pues bien. ¿Y a mí qué?.

- c) *Intervenciones mixtas* IN_{R+I}: Son aquellas que contienen movimientos con distinta orientación interactiva. Pueden darse R+I, R.Ev+Inf, R.Ev.+I, R+Inf,...

Intercambio (exchange) o Par Adyacente. Es una agrupación de turnos encabezados por un turno de intervención iniciativa y finalizado por un turno de intervención reactiva. Desde el punto de

vista estructural el intercambio prototipo constaría únicamente de dos intervenciones con orientación interactiva complementaria, es decir, una iniciativa y la otra reactiva. Los ejemplos estándar serían pregunta y respuesta, invitación y aceptación, oferta y aceptación, o requerimiento y cumplimiento entre otros.

1. Pregunta-respuesta:

- (9) **IN_I A:** ¿Vienes a cenar esta noche?
IN_R B: Hoy creo que no podré. Salgo muy tarde de trabajar.

2. Invitación-aceptación:

- (10) **IN_I A:** Pásate mañana a las tres y hablemos.
IN_R B: De acuerdo. Mañana nos vemos.

3. Oferta-aceptación:

- (11) **IN_I A:** Te puedo ceder mi carnet de socio si me invitas a cenar a cambio.
IN_R B: No gracias. Creo que me sale más barato pagar la entrada que invitarte a cenar.

4. Requerimiento-cumplimiento:

- (12) **IN_I A:** Por favor, muéstreme sus credenciales
IN_R B: Soy del servicio de seguridad. Aquí puede ver mi licencia

El concepto de par adyacente se usa puesto que se trata de un par de intervenciones la de iniciación y la de reacción que suelen encontrarse juntas. Sin embargo, esto no siempre ocurre así. Según Sacks *et al.* (1974) hay tres situaciones en las que el par adyacente

implica más de una intervención de iniciación y una intervención de reacción:

1. Expansión previa del par adyacente: Ocurre cuando se produce una expansión de la intervención iniciativa. Esto puede ocurrir cuando el hablante iniciador genera una nueva intervención iniciativa complementaria a la anterior antes de que el otro participante comience a responder:

(13) **IN_I A:** ¿Has visto el informe que te he dejado encima de la mesa?

IN_I A: Te lo quería haber preguntado en el almuerzo y se me ha olvidado.

IN_R B: Sí, ya le he echado un vistazo

o bien que el hablante continúe con una intervención iniciativa que ha sido interrumpida por un turno de aportación:

(14) **IN_I A:** Me gustaría conocer tu opinión sobre el caso de ayer

CT B: Sí

IN_I A: ¿Crees que le concederán una reducción de la pena?

IN_R B: Francamente, lo veo difícil

2. Expansión posterior del par adyacente: Ocurre cuando se produce una expansión de la intervención reactiva, bien porque el hablante que responde genera intervenciones reactivas aclaratorias a una previa, o bien porque la intervención reactiva ha quedado interrumpida por una aportación:

(15) **IN_I A:** ¿Te gustaría venir con nosotros a ver el museo?

IN_R B: La verdad es que no tengo nada mejor que hacer

IN_R B: Además, hace bastante tiempo que estoy queriendo ir y nunca encuentro el momento

3. Expansión intermedia del par adyacente: En algunos casos, entre la intervención iniciativa y la reactiva se puede encontrar un nuevo par adyacente generando una estructura de pares adyacentes anidados, como la que se puede apreciar a continuación:

(16) **IN_I A:** Tenemos que quedar todos para tratar un tema urgente

IN_I B: ¿A qué todos te refieres?

IN_R A: Me refiero a los que participan en el proyecto

IN_R B: Bien, pues podemos convocar una reunión para esta misma tarde

En el ejemplo, la segunda y tercera intervención forman un par adyacente aclaratorio que se inserta entre las intervenciones primera y cuarta, las cuales formarían un par adyacente de orden superior

Esquema de anotación de la estructura del diálogo de Carletta *et al.* Carletta *et al.*(1997) proponen un esquema de anotación para la estructura de los diálogos orientados a tareas con el propósito de estudiar y comprender este tipo de diálogo.

La anotación que presentan los autores plantea como ventajas principales respecto a otros esquemas de anotación que, por un lado, las categorías de los movimientos se definen por modelos computacionales del diálogo. Esto le proporciona una mayor independencia de la tarea que aquellas anotaciones que se basan en un tipo de diálogo particular. Por otra parte, el esquema de anotación intenta representar la estructura genérica del diálogo para usarse en conjunción con anotaciones de otros fenómenos del diálogo. Esta propiedad lo diferencia de otros esquemas de anotación que parecen estar diseñados para un propósito determinado.

Así, los autores distinguen tres niveles de anotación:

- En el nivel más alto los diálogos se dividen en transacciones. Estas transacciones son subdiálogos donde se lleva a cabo un gran paso en el plan de los participantes para alcanzar la tarea del diálogo. El tamaño y forma de la transacción depende de la tarea. En el mapa de tareas, dos participantes tienen diferentes versiones de un mapa simple con 50 marcas de referencia en él. Uno de los mapas tiene una ruta dibujada que no tiene el otro. La tarea del segundo participante es duplicarla. En este tipo de diálogo, una transacción típica sería el subdiálogo que diera con la ruta finalmente duplicada.
- Las transacciones se dividen a su vez en juegos conversacionales (también llamados juegos de diálogo, interacciones o intercambios). Todas las formas de los juegos conversacionales encarnan la observación de que las preguntas son seguidas de respuestas, las declaraciones por aceptaciones o negaciones, y así. De esta forma se distingue entre comienzos (que abren expectación en el discurso) y respuestas (que cierran dicha expectación). Los juegos conversacionales se inician con un comienzo y terminan cuando el propósito se ha completado.
- Los juegos conversacionales se componen de movimientos conversacionales, que son clases diferentes de comienzos y respuestas clasificadas de acuerdo a sus propósitos.

De esta forma, el esquema de anotación que proponen los autores es el siguiente:

1. Esquema de anotación de los movimientos.

Los movimientos se pueden anotar de acuerdo a la siguiente clasificación:

- Movimientos de iniciación:
 - a) INSTRUCT: Ordena a la pareja que lleve a cabo una acción.

(17) **A:** Gira a la derecha hasta que estés justo delante de ellos

- b) EXPLAIN: Expone una información que no ha sido provocada por la pareja.

(18) **A:** Tengo que cruzar el arroyo

- c) CHECK: Pide confirmación de una información a la pareja.

(19) **B:** Gira a la derecha

A: ¿A mi derecha? –CHECK

B: Sí

- d) ALIGN: Comprueba la atención, disponibilidad o acuerdo de la pareja para el siguiente movimiento.

(20) **B:** Gira a la derecha

B: ¿De acuerdo? –ALIGN

A: Sí

- e) QUERY-YN: Pregunta a la pareja cualquier cuestión que implique una respuesta si o no y no esté enclavada en un movimiento CHECK ni ALIGN.

(21) **A:** ¿Ves un rótulo rojo? –QUERY-YN

B: Sí

- f) QUERY-W: Cualquier pregunta no cubierta por las otras categorías.

(22) **B:** Voy hacia la cafetería

A: ¿Hacia dónde? –QUERY-W

B: Hacia la cafetería

• Movimientos de respuesta:

- a) ACKNOWLEDGE: Respuesta verbal que indica mínimamente que la pareja ha oído el movimiento al que se responde.

(23) **B:** ¿Ves un rótulo rojo?

A: No

B: No lo ves – ACKNOWLEDGE

b) **REPLY-Y:** Respuesta a cualquier pregunta con forma afirmativa.

(24) **B:** ¿Quieres que pase por allí otra vez?

A: Si puedes, sí –REPLY-Y

c) **REPLY-N:** Respuesta a cualquier pregunta con forma negativa.

(25) **B:** ¿Ves el lago oeste, abajo a tu izquierda?

A: No, de momento no – REPLY-N

d) **REPLY-W:** Cualquier respuesta que no sólo significa si o no.

(26) **B:** ¿Qué tienes delante?

A: Un bosque – REPLY-W

e) **CLARIFY:** Respuesta a cualquier clase de cuestión en la que el que habla responde a la pareja algo más sobre y de lo que le habían preguntado exactamente.

(27) **B:** ¿Cruzaste el arroyo?

A: Sí, a través del agua – CLARIFY

f) Otras respuestas posibles: Otras respuestas posibles aparecen muy raramente en corpus. En estos casos no es justificable asignarles una categoría determinada. Dependería pues del idioma en el que se trabaja.

- Movimiento de acuerdo:

a) **READY:** Cierra un juego de diálogo y prepara al sistema para un nuevo juego.

(28) **A:** De acuerdo, ahora seguiremos con el siguiente paso.

2. Esquema de anotación del juego.

Los movimientos son las piezas básicas de la estructura de un juego conversacional. Hay dos componentes principales en cualquier esquema de anotación del juego: en primer lugar, la identificación del propósito del juego (nombre del movimiento de iniciación) y en segundo lugar, una explicación de cómo los juegos se relacionan uno con otro. Para ello se anota en primer lugar el principio de cada nuevo juego nombrando el propósito del juego de acuerdo al movimiento de iniciación. En segundo lugar se anota el lugar donde los juegos terminan o son abandonados. Finalmente se marca si el juego ocurre en el nivel superior o si ocurre embebido dentro de otro juego sirviendo para el fin de éste. Estas definiciones pretenden dar suficiente información para estudiar las relaciones entre las estructuras de los juegos y otros aspectos del diálogo.

3. Esquema de anotación de la transacción.

La anotación de la transacción le proporciona al subdiálogo tareas completas de asistencia al diálogo. Cada transacción se construye con varios juegos de diálogo cada uno de los cuales corresponde a un paso de la tarea. El sistema de anotación tiene dos componentes: el primero, muestra cómo se divide una tarea en subtareas (en el corpus, el guía divide cada tarea en una serie de acciones a realizar), y qué partes del diálogo sirven a cada subtarea, y el segundo, qué acciones se toman y cuándo (acciones que toma el agente). Además, según la intención de la transacción se clasifica en NORMAL, REVISIÓN, VISIÓN GENERAL e IRRELEVANTE. Así la anotación implica marcar dónde transcribe el diálogo un principio de transacción y de qué tipo es, y para todos, exceptuando los tipos irrelevantes, se indica el principio y el fin de la tarea del corpus a la hace referencia esa transacción.

Esquema de anotación para la resolución de la anáfora de Rocha. Rocha (1999) propone un esquema de anotación de

la estructura del diálogo orientado a la extracción de reglas probabilísticas que permitan la resolución de la anáfora. El autor considera la importancia de la información estructural en la correcta resolución de algunas anáforas y propone crear un modelo probabilístico que determine para cada tipo de anáfora y para cada tipo de relación anafórica, la influencia que la información estructural tiene en su resolución. De esta forma, Rocha propone etiquetar los siguientes elementos estructurales:

1. Tópicos globales
2. Tópicos locales
3. Subtópicos locales
4. Fragmentos generados por los tópicos globales
5. Segmentos generados por los tópicos locales
6. Segmentos generados por los subtópicos globales

Estos elementos estructurales deberán etiquetarse tanto en los corpus de entrenamiento para el modelo probabilístico como en los corpus de test que se procesen en el sistema de resolución de anáfora basado en dicho modelo.

Para anotar el esquema de anotación, Rocha utiliza el siguiente criterio: toda la información introducida en el nivel de discurso se introducirá en el corpus a través de líneas marcadas con un asterisco siempre antes de comenzar el elemento a marcar. Así se utilizarán la marcas:

- * [FIRST|SECOND|...|SINGLE] FRAGMENT ‘tópico global’
Para indicar la ocurrencia del fragmento primero, segundo, etc o único, expresando entre comillas el tópico de dicho fragmento.
- * s_i ‘tópico local’ (donde $i = 1, 2, \dots, n$)
Para indicar la ocurrencia del segmento i -ésimo generado por el tópico local descrito.
- * ss_j/s_i ‘subtópico local’ (donde $j = 1, 2, \dots, m$ y $i = 1, 2, \dots, n$)
Para indicar la ocurrencia del subsegmento j -ésimo generado por el subtópico local descrito dentro del segmento i -ésimo indicado.

En el ejemplo 29 se muestra un caso que contemplaría dicha anotación.

- (29) * SINGLE FRAGMENT 'trenes que salen mañana de Alicante a Valencia'
- A:** Disculpe, podría informarme qué trenes salen mañana de Alicante a Valencia
- B:** Por supuesto. Mire, tiene uno a las siete y veinticinco de la mañana, otro a las nueve treinta y cinco, otro a las once quince,...
- * s1 'otro a las once quince'
- * ss1/s1 'hora'
- A:** Sí, este... ¿a que hora llega a Valencia?
- B:** A las doce cuarenta está en Valencia.
- * ss1/s2 'precio'
- A:** Ya... ¿Podría decirme el precio?
- B:** Pues mire le saldría por dos mil trescientas cincuenta.

Como conclusión de los trabajos expuestos se extrae la necesidad de conocer la estructura de discurso para poder comprender y generar el diálogo, y por tanto, la necesidad de establecer un esquema de anotación de dicha estructura adecuado a los propósitos que se quieren estudiar.

2.4.2 Fiabilidad de los esquemas de anotación manual

Las anotaciones basadas en un esquema pueden realizarse de forma automática mediante la definición de herramientas apropiadas (como, por ejemplo, los sistemas de diálogo) o pueden realizarse de forma manual si los investigadores usan dicha información como entrada para un determinado producto que se quiere evaluar de forma independiente. En este último caso es necesario asegurar que la información manual no altera los resultados del experimento, y por tanto, se debe ejecutar un test que mida la fiabilidad de la anotación.

De acuerdo con Carletta (Carletta *et al.*, 1997; Carletta, 1996), para asegurar una buena anotación, es necesario que gente externa a aquellos que están anotando efectúen una serie de pruebas para determinar si la anotación es creíble o no. Para ello plantea las siguientes pruebas de fiabilidad:

- Estabilidad. Mide la varianza entre las pruebas puesto que los juicios del anotador no deben cambiar con el tiempo. Para probarlo el anotador deberá anotar la misma información en distintas ocasiones y se medirá su varianza.
- Reproducibilidad. La información anotada por distintos anotadores realizando juicios sobre los mismos objetos debe ser idéntica. Para comprobarlo se realizan entrenamientos con distintos anotadores y se comprueban los resultados.
- Precisión. Los anotadores deberían anotar de la misma forma que algún estándar ya conocido. Esto se comprueba comparando los resultados producidos por los mismos anotadores respecto al estándar, si es que existe.

Además de estas mediciones directas, la fiabilidad debe medir en esencia la cantidad de ruido en la información. Si interfiere o no con los resultados dependerá del ruido y de la fuerza de la relación que se está midiendo. Los autores argumentan que puesto que la cantidad de acuerdo que uno puede esperar por casualidad depende del número y frecuencias relativas de las categorías bajo test, la fiabilidad para las clasificaciones de categorías debería medirse usando el coeficiente *kappa* definido en Siegel y Castellan (1988). El coeficiente *kappa* (*k*) mide la afinidad de acuerdo entre un conjunto de anotadores cuando hacen juicios de categorías y se calcula de la siguiente forma:

$$P_E = \left(\frac{C_1}{N \times C} \right)^2 + \left(\frac{C_2}{N \times C} \right)^2 + \dots + \left(\frac{C_n}{N \times C} \right)^2$$

$$k = \frac{P_A - P_E}{1 - P_E}$$

Donde *N* es el número total de objetos anotables, *C* es el número de anotadores, *C_n* es el número total de objetos asignados

a cada clase, P_A es el porcentaje de acuerdo entre los anotadores, $P_A = \frac{Z}{N}$, siendo Z el número total de acuerdos entre los anotadores y P_E el porcentaje de acuerdo esperado por casualidad.

Según el trabajo de Carletta, una medida de k tal que $0.68 < k < 0.8$ permitiría hacer conclusiones favorables, y si $k > 0.8$ significa entonces que existe fiabilidad total entre los resultados de ambos anotadores.

2.5 Conclusiones del capítulo

En este capítulo se ha realizado una exposición de los distintos tipos de diálogo existentes según su capacidad para generar fenómenos lingüísticos. Esta exposición se ha empleado para centrar el ámbito concreto de nuestro estudio que son los diálogos tecleados hombre-máquina ya que serán éstos los empleados por los sistemas de diálogo, los cuales son la aplicación del trabajo que se ha desarrollado.

Además, sobre los sistemas de diálogo se ha mostrado su arquitectura básica y las necesidades de fuentes de información que permiten la gestión eficiente del diálogo. En concreto, se ha demostrado la necesidad básica de resolución de las referencias existentes en estos sistemas y cuáles deberían ser las características que debe cumplir todo módulo de gestión de referencias: su adaptación al contexto y su independencia del dominio.

Los sistemas de diálogo plantean esquemas de anotación de su estructura que les ayuden a comprender y analizar las intenciones de dichos diálogos. Cualquiera de estas aplicaciones contiene mecanismos automáticos para la detección y generación de dichos esquemas.

Por último, destacamos que el estudio de técnicas para la segmentación y estructuración de los diálogos, así como las propuestas de anotación de éstos, nos ha permitido vislumbrar la importancia de estas técnicas y su considerable influencia en la resolución de la anáfora.



Universitat d'Alacant
Universidad de Alicante



3. La anáfora en diálogos

Universitat d'Alacant
Universidad de Alicante

En este capítulo se presenta un estudio del fenómeno lingüístico de la anáfora y de las distintas estrategias empleadas para su resolución orientado en concreto al caso de los diálogos. Así comenzaremos por realizar una taxonomía de los distintos tipos de relaciones anafóricas que pueden ser encontradas en los diálogos atendiendo a las clasificaciones de diversos autores.

En segundo lugar, se mostrará una revisión bibliográfica en la que se tratarán algunas de las estrategias para resolución de anáfora más características junto a su aplicabilidad a los diálogos.

En tercer lugar, se muestra un estudio de la influencia que tiene la estructura del diálogo sobre la generación, y por tanto sobre la resolución de la anáfora, basándose en los estudios de Fox (Fox, 1987) y de Rocha (Rocha, 1999). Dicho estudio lo usaremos como base para nuestra propuesta de resolución de la anáfora que se presentará en los capítulos posteriores.

Por último, se presentan las necesidades de resolución de la anáfora en los sistemas de diálogo como aplicación final del sistema que hemos desarrollado. Se enfatizará especialmente el modo en el que se produce la integración de los sistemas de diálogo con el módulo de procesamiento de lenguaje natural que resuelve e interpreta los fenómenos lingüísticos. En concreto, centraremos el estudio en cómo se puede integrar el sistema de resolución de la anáfora ARIADNA y sus fuentes de información en un sistema de diálogo completo.

3.1 Taxonomía de las relaciones anafóricas en el diálogo

En el diálogo, igual que en cualquier otro tipo de discurso, se pueden clasificar las relaciones anafóricas de acuerdo a varios parámetros, que afectan al elemento anafórico, al antecedente, y a la relación entre ambos. Para los objetivos de este trabajo es importante hacer una distinción de anáforas tanto por la categoría gramatical del elemento anafórico como por la naturaleza del antecedente. A continuación se exponen las clasificaciones asociadas a cada parámetro.

3.1.1 Relaciones anafóricas según la categoría gramatical del elemento anafórico

Según el trabajo de Rocha (1998), de acuerdo con la categoría gramatical del elemento anafórico podemos distinguir los siguientes tipos de relaciones anafóricas en el diálogo:

Anáfora pronominal. Una de las anáforas más comunes en diálogos es la provocada por pronombres personales, demostrativos y posesivos. Según los estudios empíricos realizados por Rocha sobre diálogos en inglés, se encontró que sobre 3085 anáforas detectadas, el 50,5% eran pronominales. El mismo estudio fue efectuado sobre diálogos en portugués encontrando un 17,4% de anáforas pronominales sobre un total de 3045. Este descenso es particularmente debido a la existencia de frecuentes anáforas cero (conocidas también como omisiones del sujeto pronominal) que afectan a ciertas lenguas como el portugués y el español, y que permite la omisión de ciertos pronombres que en otras lenguas como el inglés son necesarios. Por tanto, aunque Rocha no efectuó estudios sobre el número de anáforas pronominales en español, el paralelismo entre ambas lenguas indica que estará próximo al valor obtenido para el portugués.

Las anáforas pronominales se pueden clasificar atendiendo a los siguientes tipos:

- Anáfora pronominal de Sujeto: Ocurrencias de pronombres personales *yo, tú, él, ella, ello, nosotros, nosotras, vosotros, vosotras, ellos, ellas* con función sintáctica de sujeto.

(30) **B:** Me han dicho que ayer estuvo aquí la directora del nuevo centro_i.

A: Sí, ella_i estuvo esperando hasta las nueve y media, pero luego se tuvo que marchar.

- Anáfora pronominal de Objeto: Ocurrencias de pronombres personales *me, te, se, le, lo, la, mí, ti, él, ella, ello, nosotros, nosotras, vosotros, vosotras, ellos, ellas, les, las, los, conmigo, contigo* con función sintáctica de objeto (directo o indirecto).

(31) **A:** Voy a intentar hablar con Luis y Paco_i antes de que se haga demasiado tarde.

B_j: Por favor, ¿les_i puedes decir que hoy no me_j esperen para desayunar?

Este tipo de pronombre puede aparecer embebido en el verbo del enunciado (llamados pronombres enclíticos). El ejemplo 31 tiene la misma validez si modificamos la segunda intervención tal y como puede apreciarse en el ejemplo 32

(32) **B**_j: Por favor, ¿puedes decirles_i que hoy no me_j esperen para desayunar?

- Anáfora pronominal reflexiva: Ocurrencia de los pronombres personales *me, te, se, nos, os* actuando reflexivamente.

(33) **A:** ¿Crees que aprobará mucha gente?

B: No sé, pero el profesor_i se_i se está preparando para lo peor.

- Anáfora pronominal recíproca: Ocurrencia de los pronombres personales *me, te, se, nos, os* actuando recíprocamente.

(34) **A:** ¿Cuándo podré conocer al profesor alemán?

B: No te preocupes. Todos los conferenciantes_i os_i saludaréis antes de subir al estrado.

- Anáfora pronominal demostrativa: Ocurrencias de pronombres demostrativos *éste, ésta, esto, ése, ésa, eso, aquél, aquélla, aquello*.

(35) **A:** Dispone de trenes que salen de la estación de origen a las diez, a las doce y a la una y media. El último tren_i es a las tres.

B: ¿Y a qué hora llega éste_i?

- Anáfora pronominal posesiva pura: Ocurrencias de pronombres posesivos *mío, mía, míos, mías, tuyo, tuya, tuyos, tuyas, suyo, suya, suyos, suyas, nuestro, nuestra, nuestros, nuestras, vuestro, vuestra, vuestros, vuestras*.

(36) **A:** ¿Es esta la casa de Rafa_i?

B: No, ésta no es. La suya_i tiene un toldo de color verde y blanco.

- Anáfora pronominal posesiva como determinante: Ocurrencias de pronombres posesivos que actúan de determinante y hacen referencia al poseedor de la entidad a la que modifican *mi, tu, su, mis, tus, sus, nuestro, nuestra, nuestros, nuestras, vuestro, vuestra, vuestros, vuestras*.

(37) **A:** ¿Sabes qué hay que hacer para hablar con el encargado_i?

B: Muy fácil. Marca su_i número de teléfono y espera a que conteste.

Sin embargo, hay que indicar que no todos los pronombres son anafóricos. A este tipo de pronombres Hirst (Hirst, 1981) les denomina *pronombres no referenciales*. En español, este tipo de pronombres suelen ser frecuentes en la forma “se” impersonal, como ocurre en el caso *se vive bien en esta ciudad*.

Sintagmas nominales no pronominales. En este apartado se incluyen los sintagmas nominales que sin tener un pronombre como núcleo, hacen referencia a otro sintagma nominal enunciado previamente. El enlace entre el antecedente y la anáfora se hace a través de repeticiones léxicas, repeticiones con adición de modificadores, o mediante relaciones semánticas como la hiponimia e hiperonimia (relación es-un), meronimia (relación parte-todo), sinonimia y antonimia.

- (38) **A:** Podría decirme el horario de el trayecto entre Alicante y Benidorm_i
B: Lo lamento mucho, pero **ese recorrido_i** no lo gestionamos nosotros. Tendrá que llamar a la compañía responsable.

En el ejemplo 38 existe una relación de sinonimia entre *el trayecto entre Alicante y Benidorm* y *ese recorrido* que produce la relación anafórica.

Anáfora adjetiva. La anáfora adjetiva se produce por la ocurrencia de un adjetivo que requiere la búsqueda de un antecedente nominal para su interpretación semántica puesto que el núcleo del sintagma nominal en el que se encuentra ha quedado omitido. Este adjetivo puede venir expresado tanto en su forma normal como ocurre en el ejemplo 39 o en la forma superlativa como en el ejemplo 40

- (39) **A:** ¿Me puedes traer la chaqueta que hay en el armario_i?
B: ¿Cuál de todas? Hay varias.
A: **La verde_i**.
- (40) **A:** ¿Podría decirme los trenes que tengo para ir el día veinticinco por la mañana desde Murcia hasta Castellón_i?

B: Bueno, **el más rápido_i** es el de las once, que sólo tarda tres horas y media, pero también es **el más caro_i**.

Además la anáfora adjetiva puede aparecer como una cláusula de relativo con función de modificador adjetivo como ocurre en el ejemplo 41

(41) **A:** Tiene un tren_i a cada hora.

B: **El que sale a las diez_i** ¿a qué hora llega?

Y también puede aparecer como un sintagma preposicional con función de modificador adjetivo como en 42

(42) **A:** Tiene un tren_i a cada hora.

B: **El de las diez_i** ¿a qué hora llega?

La anáfora adjetiva equivale en la mayoría de los casos a la *one-anaphora* en inglés con la única variación que la *one-anaphora* inglesa incluye el pronombre *one* para sustituir al núcleo del sintagma nominal (cuando es un nombre contable) que queda omitido.

Destacamos por último, que la anáfora adjetiva puede ser entendida como un tipo de elipsis al contemplar una omisión de elementos. Sin embargo, esta omisión produce el efecto de la sustitución de un sintagma nominal completo por otro sintagma nominal, que aún estando incompleto, guarda la información léxica, morfológica y sintáctica, y que además correfiere con un elemento anterior del discurso (el sintagma nominal completo). Este fenómeno así definido corresponde con la definición de anáfora, matizando que el objetivo de la resolución de la anáfora adjetiva no es el de recuperar el elemento omitido, tarea que le corresponde a los sistemas de tratamiento de la elipsis, sino localizar el antecedente de la anáfora creada.

Anáfora adverbial. La anáfora adverbial se produce por la ocurrencia de adverbios o complementos circunstanciales que hacen referencia a lugares, modos, tiempos, etc.

(43) **A:** ¿A qué hora llega el tren que sale a las nueve para Barcelona_i?

B: Suele llegar **allí**_i sobre las doce menos cuarto.

Un caso particular de este tipo de anáfora son las expresiones temporales referenciales expresadas mediante sintagmas nominales con función adverbial como en el ejemplo 44. Este tipo de expresiones necesitan obtener el antecedente para poder ser calculadas.

(44) **A:** El avión no sale hasta las doce y media_i?

B: Sí, pero hay que estar en el aeropuerto **una hora antes**_i.

Anáfora superficial numérica. La anáfora superficial numérica es un tipo de anáfora bastante similar a la adjetiva. En este caso el componente adjetivo debe ser un numeral de tipo ordinal, cardinal, o en general, que exprese algún tipo de orden como ocurre en *el primero*, *el tercero*, *el último*, *el dos*, etc. La principal diferencia de este tipo de anáfora respecto a la anáfora adjetiva es que para ser resuelta no sólo es necesaria la localización del antecedente, sino que también hay que mantener la estructura superficial del enunciado donde aparece dicho antecedente con el orden de los constituyentes. En el ejemplo 45 se puede observar cómo la resolución del antecedente no sería posible sin mantener en la consciencia del hablante B el orden en el que el hablante A enunció los tres elementos.

(45) **A:** En el menú del día hoy puede elegir entre tres platos: arroz a banda_i, gazpacho de merro o ensalada especial.

B: Bueno, pues me quedo con **el primero**_i.

Anáfora verbal. La anáfora verbal se produce por la sustitución de un verbo, o incluso un sintagma verbal completo, por un verbo auxiliar para evitar la repetición del mismo. Este tipo de anáfora es muy común en lenguas como el inglés donde se utilizan los

auxiliares *do* y *have* para responder de forma abreviada como se presenta en 46

(46) **A:** Did you finish your thesis_i?

B: Yes, I **do**_i.

Sin embargo, la traducción de este tipo de anáfora al español suele ir acompañada por un pronombre de objeto que es quien realiza realmente la función referencial (ejemplo 47), por lo que puede ser tratada generalmente como una anáfora pronominal de objeto.

(47) **A:** Terminaste tu tesis_i?

B: Sí, **lo hice**_i.

La anáfora verbal puede ser confundida en ocasiones con una elipsis verbal. La principal diferencia entre la elipsis verbal y la anáfora verbal es que en la primera el verbo se elimina y no se sustituye por ningún otro componente, mientras que en la anáfora verbal existe una expresión anafórica que la sustituye.

Además, según estudia Franchini (Franchini, 1986), la anáfora verbal nunca puede contener un objeto que no sea el pronombre, mientras que un verbo elíptico admite objetos no referenciales como se muestra en el ejemplo 48, donde el verbo ha desaparecido sin ser sustituido por ninguna expresión anafórica, y además admite un objeto directo *la foto*.

(48) **A:** Debes entregar_i la ficha antes del lunes.

B: Sí ya lo sé, y \emptyset _i la foto también.

Omisión del sujeto pronominal (*zero-pronoun*). La omisión del sujeto pronominal es un fenómeno muy habitual del español que se produce cuando el pronombre de sujeto se omite recayendo todo el peso semántico sobre el verbo. Este fenómeno es especialmente propio de aquellas lenguas cuyos verbos conjugan la persona y el número del sujeto, por lo que al contar con infor-

mación morfológica suficiente en el verbo se permiten la omisión del pronombre.

La omisión del sujeto pronominal, aún considerándola un caso de elipsis, se trata como una anáfora puesto que la información morfológica sobre el pronombre omitido que contiene el verbo actúa como expresión anafórica que se relaciona con el antecedente. Al conocer la persona y el número, el pronombre omitido es fácilmente recompuesto como un pronombre de sujeto a falta de definir su género en el caso de la tercera persona. Una vez recompuesto el pronombre, la tarea de resolución de la anáfora se convierte en un caso más de anáfora pronominal de sujeto.

(49) **A:** ¿Qué te dijo Juan_i sobre el informe de contabilidad?.

B: Ø(él)_i Me comentó que habría que revisarlo de nuevo.

3.1.2 Relaciones anafóricas según la naturaleza del antecedente

Según la naturaleza del antecedente Asher (Asher, 1993) distingue entre anáforas individuales y anáforas abstractas.

Anáfora individual. La anáfora individual engloba todas aquellas anáforas cuyo antecedente es un objeto concreto (o individual). Esta concreción hace que el antecedente corresponda sintácticamente con un sintagma nominal.

(50) **A:** ¿Donde está la carta que me querías enseñar_i?

B: La_i tienes sobre la mesa.

Anáfora abstracta. La anáfora abstracta, en contraposición a la anáfora individual no tiene como referente un objeto concreto sino abstracto, como eventos, hechos o proposiciones. Sintácticamente, el referente no es por tanto un sintagma nominal sino un sintagma verbal, una oración, o incluso una cadena de oraciones.

Una de las anáforas abstractas más usadas en diálogos es la *deixis del discurso*. La deixis del discurso es definida por Levinson (Levinson, 1983) como una referencia deíctica a una parte del discurso relativa a la localización actual del hablante en el discurso. La deixis del discurso es por tanto una referencia a un elemento extralingüístico, con la salvedad que este elemento extralingüístico es una porción del discurso. En el ejemplo 51 el pronombre demostrativo *eso* hace referencia al fragmento de discurso que acaba de enunciar el hablante anterior.

- (51) **A:** La opinión de Fernando es que no tendremos otra oportunidad como ésta para conseguir el espacio que nos corresponde._i
B: Bueno, **eso**_i no es del todo cierto. Aún existe otra posibilidad que se podría evaluar.

En el discurso narrativo es también muy frecuente hacer referencia a segmentos del discurso como *en el próximo capítulo*.

En los estudios realizados por Eckert y Strube en (Eckert & Strube, 1999b) sobre el *Switchboard corpus of telephone conversations* se encontró que un 22,6% de las anáforas encontradas eran deixis del discurso.

En el mismo estudio, Eckert y Strube encontraron que otro tipo de anáfora abstracta bastante frecuente era la llamada *anáfora vaga*, (el 13,2%). Este tipo de anáfora hace referencia al tópico global del discurso pero, al contrario de lo que ocurre en la deixis de discurso, éstas no tienen como antecedente un fragmento específico del discurso. En el ejemplo 52 el pronombre demostrativo *esto* se usa para hacer referencia al tópico de la conversación pero no hay ningún fragmento determinado del discurso al que se haga referencia.

- (52) **A:** ¿Que tal te van las cosas? ¿Cómo va todo?
B: Pues mira, estoy bastante cansado la verdad. Me levanto todos los días a las cinco de la mañana para empezar a trabajar a las seis. Trabajo toda la mañana hasta que se hacen

las dos y media o las tres. Como en una hora y a las cuatro ya estoy allí de nuevo hasta las ocho que salgo.

A: Vaya, no tienes tiempo casi para descansar.

B: No, no lo tengo. **Esto** es bastante duro.

En concreto, nuestro trabajo se centrará en la resolución de las anáforas pronominales y adjetivas, en los que el antecedente es de naturaleza individual.

3.2 Estrategias de resolución de la anáfora en diálogos

A través de la literatura sobre resolución de anáfora se han planteado muchos métodos para llevar a cabo esta tarea diferenciándose básicamente en la fuente de información que usa cada uno de ellos. A continuación se mostrará una revisión bibliográfica en la que se tratarán algunas de las estrategias para resolución de anáfora más características junto a su aplicabilidad a los diálogos.

Los primeros métodos de resolución de la anáfora en diálogos no se desarrollaron como mecanismos independientes sino integrados en el sistema de diálogo en el que se encontraban.

Un ejemplo de estas primeras aproximaciones es el sistema ELIZA (Weizenbaum, 1966), uno de los primeros sistemas de simulación humana que representaba el comportamiento de un psiquiatra dialogando en inglés con su paciente. ELIZA, lejos de tratarse de un sistema experto, únicamente era capaz de construir nuevos enunciados enlazados a partir de los enunciados que emitía el usuario mediante un mecanismo de reconocimiento de patrones (*pattern-matching*). El sistema realizaba un tratamiento de anáfora pronominal personal de primera y segunda persona siendo capaz de intercambiar el *yo* y el *tú* según el hablante.

El sistema interactivo SHRDLU (Winograd, 1972; Winograd, 1986) permitía al usuario entablar un diálogo con el sistema obteniendo información sobre el mundo de las figuras geométricas, las posiciones, formas y colores que pueden adoptar, y además era

capaz de aceptar instrucciones para el movimiento de un brazo robot. El sistema podía tratar algunos pronombres personales y sintagmas nominales definidos y casos de anáfora generada por el pronombre *one* en inglés (*one anaphora*) en el que los adjetivos tienen sentidos opuestos, como el mostrado en el ejemplo 53 (en la segunda línea se puede apreciar el mismo caso traducido al español por medio de una anáfora adjetiva).

- (53) A big green pyramid_i and a little **one**_i
 Una gran pirámide verde_i y **una pequeña**_i

SHRDLU estaba formado básicamente por un analizador sintáctico, un analizador semántico y el módulo de razonamiento para la resolución de las cuestiones. Para la resolución de la anáfora, el sistema usaba una lista con todos los sintagmas nominales que aparecían en el diálogo a la que le aplicaba restricciones de concordancia morfológica en género, número y persona para elegir finalmente el candidato más próximo que cumplía dichas restricciones. El propio diálogo era usado por el sistema para que el usuario confirmara la elección del antecedente. Sin embargo, una de las principales desventajas de este sistema era que no manejaba restricciones sintácticas para evitar, por ejemplo, situaciones de no correferencialidad dentro de una misma oración, como en el caso del ejemplo 54.

- (54) Él puso la caja sobre la mesa. Puesto que ésta_i no estaba nivelada, **ésta**_j resbaló.¹

LUNAR (*Lunar Sciences Natural Language Information*) (Woods, 1977) fue un sistema de diálogo creado para actuar como interfaz en lenguaje natural sobre una base de datos de minerales extraídos de las piedras lunares que se adquirieron en el viaje del Apolo-XI a la luna. El sistema estaba formado por un analizador sintáctico, un módulo de interpretación semántica y un módulo gestor de la base de datos. Durante el diálogo, el sistema almacenaba la información sintáctica y semántica de cada enti-

¹ la no correferencialidad se indica por medio de subíndices distintos

dad aparecida. Con esta información, LUNAR resolvía dos tipos de anáforas: parciales y completas. La anáfora parcial emplea un pronombre que hace referencia a una parte de un sintagma nominal que ya ha aparecido anteriormente, tal y como se puede apreciar en la adaptación al español del problema efectuada en el ejemplo 55.

- (55) Dame los análisis de hidrógeno de la muestra 10046_i. Dame **los de oxígeno_i**.

Este es un caso especial de anáfora adjetiva que se detecta por la aparición de una oración de relativo o por un sintagma preposicional modificando al pronombre. La anáfora completa sin embargo se refiere al sintagma nominal completo. En cualquiera de los dos casos, la estrategia de resolución dependía del tipo de pronombre encontrado. Si el pronombre era demostrativo, el sistema buscaba un sintagma nominal cuyo núcleo fuera el mismo que acompañaba al demostrativo. En caso de ser un pronombre personal, se usaría la información semántica para establecer una concordancia semántica.

El sistema, sin embargo, tenía como principal limitación que dependía del paralelismo estructural entre anáfora y antecedente, por lo que era incapaz de hacer correferir entidades con estructuras diferentes. Por otra parte, igual que le ocurría al SHRDLU, al depender de una ontología semántica definida *ad-hoc* para un dominio semántico concreto, el sistema era difícilmente extensible a entornos de dominio no restringido.

En definitiva se podría concluir que el principal problema de los mecanismos de resolución de anáforas integrados en los sistemas clásicos de diálogo era que carecían de consistencia suficiente para ser extendidos a otros dominios. Esto obligaba a que cada diseñador de un sistema de diálogo tuviera que construir su propio mecanismo de resolución anafórico. Tras estos primeros sistemas, los investigadores del lenguaje natural empezaron a plantear mecanismos de resolución anafórica que fueran totalmente independientes del sistema de diálogo.

3.2.1 Enfoques estratégicos en la resolución de la anáfora

En Ferrández (1998) se clasifican las distintas estrategias para la resolución de la anáfora en dos tipos: estrategias *integradas* y estrategias *alternativas*. Las primeras agrupan una serie de conocimientos (fuentes de información) que se suponen necesarios para resolver la anáfora. Este tipo de estrategia se basa en la imitación del comportamiento humano, es decir, se pretende dotar al ordenador con las mismas fuentes de información que tiene el humano cuando resuelve la anáfora que ha generado su pareja en la comunicación. Estas fuentes de información se pueden fusionar por medio de una serie de mecanismos de inteligencia artificial para que todas ellas contribuyan a proporcionar finalmente una elección (conociéndose estos sistemas como *democráticos*). O también pueden utilizar una única fuente de información (conociéndose entonces como sistemas *consultivos*). Por su parte, los sistemas alternativos no intentan imitar las fuentes de conocimiento humanas, sino que se basan en el estudio de corpus a través de herramientas estadísticas de forma que la información para la toma de decisiones se extrae de los textos usados para el entrenamiento.

Por otro lado, en la línea anteriormente comentada, Dahlbäck (Dahlbäck, 1991) hace una clasificación de los sistemas de estrategias integradas según el tipo de fuente de información utilizada. Así reconoce, por un lado, los sistemas tradicionales basados en fuentes de información lingüística y por otro, los sistemas basados en la información de la estructura del discurso.

A continuación se proporciona un estudio de los métodos más característicos desde el enfoque de las estrategias integradas: métodos basados en información lingüística, métodos basados en la estructura del discurso, métodos basados en restricciones y preferencias (que podrían integrar las dos anteriores), así como los métodos alternativos basados en información estadística.

3.2.2 Métodos integrados basados en información lingüística

Los métodos de resolución de anáfora basados en información lingüística han sido ampliamente usados para el procesamiento de textos en general. Dichos métodos utilizan reglas lingüísticas que se apoyan en informaciones léxicas, morfológicas, sintácticas, semánticas y/o contextuales para explicar las relaciones entre las anáforas y sus antecedentes. Su aplicación al procesamiento de diálogos no es *a priori* una tarea compleja debido a que la información lingüística, a diferencia de la información estructural, no varía demasiado de un discurso con un único participante (monólogo) a un discurso con múltiples participantes (diálogo). Sin embargo, será necesario hacer algunas adaptaciones de las propuestas para su aplicación a diálogos. A continuación se describen algunos de los métodos más clásicos y significativos.

Algoritmo de Hobbs. En Hobbs (1978), reeditado en Hobbs (1986), se presenta uno de los algoritmos más citados en la literatura de resolución de la anáfora pronominal basándose únicamente en información lingüística. Este algoritmo trabaja con árboles de análisis sintáctico superficial de las oraciones y se basa en el uso de un algoritmo simple de búsqueda en árboles. El procedimiento de búsqueda selecciona el primer candidato encontrado mediante un recorrido izquierda-derecha y que satisfaga ciertas restricciones morfo-sintácticas. Como restricción morfológica se aplica la concordancia morfológica. Como restricción sintáctica se aplica la siguiente:

El algoritmo escoge como antecedente de un pronombre P el primer SN_i del árbol obtenido por recorrido de izquierda a derecha, primero en anchura, desde las ramas a la izquierda del camino T tal que:

- T es el camino desde el constituyente SN que domina a P hasta el primer constituyente SN u oración O que domina a este SN,
- T contiene un nodo N que representa a un constituyente SN u O que contiene el SN que domina a P,
- y, N no contiene a SN_i

Si el antecedente que satisface esta condición no se encuentra en la misma oración que contiene a P, entonces el algoritmo selecciona el primer sintagma nominal en recorrido de izquierda a derecha, primero en anchura.

Esta restricción sintáctica se basa en la *Teoría de rección y ligamiento (Government and Binding Theory)* analizada por Chomsky (1981) acerca de las relaciones anafóricas intraoracionales.

El método propuesto por Hobbs se basa en la aplicación de restricciones y entre todos aquellos que las cumplen se elige al candidato más próximo. El método fue evaluado sobre cien frases en inglés obteniendo un porcentaje de éxito del 81,8%. Así, este método básico se ha tomado como punto de partida para otros métodos que, empleando un sistema de restricciones similar, enriquecen la elección del mejor candidato con otras fuentes de información.

Algoritmo de Lappin y Leass. En Lappin y Leass (1994) se presenta un algoritmo para la identificación de los sintagmas nominales antecedentes para pronombres personales de tercera persona y anáforas léxicas (pronombres reflexivos y recíprocos) reportando una tasa de éxito del 85% cuando se aplica a textos en inglés. El algoritmo de resolución de la anáfora que se alimenta con información morfológica y sintáctica contiene los siguientes elementos:

- un filtro morfológico para descartar candidatos incompatibles,
- un filtro sintáctico para determinar la correferencialidad intraoracional entre el pronombre y el candidato,
- un procedimiento para identificar pronombres pleonásticos (no correferentes),
- un procedimiento para asignar valores a distintos parámetros que muestran la relevancia discursiva de cada candidato,
- un procedimiento para computar los diversos parámetros de cada candidato obteniendo un factor de relevancia global único para cada uno de ellos,

- y un procedimiento de decisión para seleccionar el candidato preferido de una lista de candidatos partiendo de sus factores de relevancia globales.

A continuación se muestra un detalle de los elementos fundamentales del algoritmo.

Filtro morfológico. El filtro morfológico propuesto se basa en la concordancia de género, número y persona entre el pronombre anafórico y el sintagma nominal candidato a antecedente. Cualquier candidato que no cumpla esta condición se descarta automáticamente.

Filtro sintáctico. Está formado por una serie de condiciones sobre correferencialidad y no correferencialidad de pronombres y sintagmas nominales que aparecen en la misma oración para descartar candidatos incompatibles. Dicho filtro se presentó con detalle en Lappin y McCord (1990) y necesita la información generada por un analizador sintáctico completo para poder ser aplicado.

Para presentar las condiciones de filtro es necesario exponer previamente la terminología usada por Lappin y Leass:

- Un sintagma P está en el *dominio argumento* de un sintagma N si y sólo si P y N son ambos argumentos del mismo núcleo. El sujeto, el objeto directo y el indirecto son argumentos del verbo. Así en el ejemplo 56, *the player* está en el dominio argumento de *the ball* y viceversa.

(56) **The player pitched the ball.**

El jugador lanzó la bola.

- Un sintagma P está en el *dominio adjunto* de N si y sólo si N es un argumento de un núcleo H, P es el objeto de una preposición PREP, y PREP es un adjunto de H. Por ejemplo, el SN de un SP que actúa como complemento circunstancial del verbo estará en el dominio adjunto del sujeto o de los objetos de ese verbo. Así en el ejemplo 57, *the bat* está en el dominio adjunto de *the ball* y de *the player*.

(57) **The player pitched the ball with the bat**
El jugador lanzó la bola con el bate

- Un sintagma P está en el *dominio SN* de N si y sólo si N es el determinante de un nombre Q y (i) P es un argumento de Q, o (ii) P es el objeto de una preposición PREP y PREP es un adjunto de Q. En 58 *Peter* pertenece al dominio SN de *John*.

(58) **John's portrait of Peter.**
El retrato de John de Peter.

- Un sintagma P *está contenido en* un sintagma Q si y sólo si (i) P es un argumento o un adjunto de Q, es decir, está *inmediatamente contenido en Q*, o (ii) P está inmediatamente contenido en algún sintagma R, y R está contenido en Q. En el ejemplo 58 *John* está contenido (inmediatamente) en *John's portrait*.

Así se definen las siguientes cinco reglas de filtrado sintáctico que determinan los casos en los que un pronombre P² (no reflexivo ni recíproco) no puede correferir con un sintagma nominal N en la misma oración:

1. P está en el dominio argumento de N.

(59) The player_i pitched **it_j**.
El jugador_i **la_j** lanzó.

2. P está en el dominio adjunto de N.

(60) He_i sat near **him_j**.
Él_i se sentó cerca de **él_j**.

3. P es un argumento de un núcleo H, N no es un pronombre y N está contenido en H.

(61) **He_i** believes that the man_j is wrong.

² En este caso P representa un sintagma formado por un pronombre

Él_i piensa que el hombre_j está equivocado.

4. P está en el dominio SN de N

(62) John's portrait of **him**_i is interesting.
El retrato de John de él_i es interesante.

5. P es el determinante de un nombre Q, y N está contenido en Q

(63) **His**_i portrait of John_j is interesting.
Su_i retrato de John_j es interesante.

Además, se definen las cinco condiciones para determinar cuándo un pronombre reflexivo o recíproco puede correferir con un sintagma nominal en la misma oración. Previamente se introduce un nuevo concepto en la terminología de los autores:

- Un sintagma P *ocupa un hueco de argumento mayor* que otro sintagma Q si $rol(P) > rol(Q)$ en la siguiente jerarquía de roles gramaticales:
 sujeto > agente (sujeto de la voz pasiva) > objeto directo > otros

Así, diremos que un sintagma nominal N es un antecedente posible para una anáfora léxica A (pronombre reflexivo o recíproco) si se cumple alguna de las siguientes condiciones:

1. A está en el dominio argumento de N y N ocupa un hueco de argumento mayor que A.

(64) They_i wanted to see **themselves**_i.
Ellos_i **se**_i querían ver.

2. A está en el dominio adjunto de N.

(65) John_i worked by **himself**_i.
John_i trabajó por **sí mismo**_i.

3. A está en el dominio SN de N

- (66) John likes Bill_i's portrait of **himself_i**.
 A John le gusta el retrato de Bill_i de **sí mismo_i**.

4. N es un argumento de un verbo V, hay un sintagma nominal Q en el dominio argumento o en el dominio adjunto de N, y Q no tiene determinante nominal, y (i) A es argumento de Q, o (ii) A es argumento de una preposición PREP y PREP es adjunto de Q.

- (67) They_i told stories about **themselves_i**.
Ellos_i contaron historias sobre **sí mismos_i**.

5. A es un determinante de un nombre Q, y (i) Q está en el dominio argumento de N y N ocupa un hueco de argumento mayor que Q, o (ii) Q está en el dominio adjunto de N.

- (68) John and Mary_i like each **other_i**'s portraits.
 A John y María_i les gustan los retratos de **sí mismos_i**.

Factores de relevancia. Una vez eliminados los candidatos incompatibles, el algoritmo considera algunos factores para destacar la relevancia discursiva de algunos candidatos frente a otros. Estos factores ayudarán a determinar el candidato más relevante, eligiendo finalmente a éste como antecedente. Los factores considerados son los siguientes por orden de importancia:

1. Oración actual: la valoración máxima se otorga a los candidatos que se encuentren en la misma oración que la anáfora.
2. Énfasis de sujeto: los candidatos con función de sujeto tienen mayor relevancia que cualquier otra función gramatical.
3. Énfasis de núcleo nominal: cuando el candidato no está incluido en otro sintagma nominal.

4. Énfasis existencial: los candidatos con función de predicado en una oración existencial (con verbo existir o haber).
5. Énfasis acusativo: los candidatos con función de sujeto.
6. Énfasis no adverbial: los candidatos que no están en un sintagma preposicional adverbial.
7. Énfasis de objeto indirecto: los candidatos que son objetos indirectos.

Con el orden expuesto, se van asignando pesos que se van acumulando por cada propiedad que cumple el candidato y se adquiere un valor global de relevancia. Además, cada vez que se produce un cambio de oración en el texto los valores de relevancia de los candidatos aparecidos en oraciones anteriores se degradan dividiéndose por factor 2. Así cuando el valor de relevancia de un candidato llega a ser 0, éste se elimina de la lista de candidatos. El candidato de mayor valor será el propuesto por el algoritmo como antecedente de la anáfora.

El principal problema de adaptación del algoritmo de Lappin y Leass a los diálogos estriba en el tipo de información que maneja. Tanto el filtro sintáctico, como los valores de relevancia utilizados se basan en la entrada de información de un analizador sintáctico completo. La complejidad sintáctica de los enunciados en diálogos, junto con las posibles incorrecciones, falta de signos de puntuación y enunciados interrumpidos hacen que el análisis sintáctico completo sea una tarea utópica en el procesamiento de diálogos reales.

3.2.3 Métodos integrados basados en la estructura del discurso

Los métodos de resolución de anáfora basados en la estructura del discurso se caracterizan por extraer la información de la propia forma de construir el discurso.

Marco global del *centering* (Grosz *et al.*). Muchos trabajos sobre resolución de anáfora que usan estructura del discurso se basan en el método del *centering* (Grosz *et al.*, 1983; Grosz *et al.*, 1995) como un marco global para modelar la coherencia local en el

discurso. El marco conceptual del *centering* explica la coherencia local que relaciona el foco local³ y la forma de las expresiones de referencia.

Este marco se basa en tres afirmaciones principales:

1. Dado un enunciado U_i , el modelo predice qué entidad del discurso será el foco de U_{i+1} .
2. Cuando el foco local se mantiene entre enunciados, el modelo predice que se expresará mediante un pronombre.
3. Cuando se encuentra un pronombre, el modelo proporciona un orden de preferencia sobre los antecedentes posibles del enunciado anterior.

Para ello, en cada U_i se crean las siguientes estructuras de datos:

- $Cf(U_i)$: La lista *forward-looking centers*, parcialmente ordenada, que incluye todas las entidades del discurso del enunciado U_i . Su primer elemento es el *center* preferido, $Cp(U_i)$, y será el candidato que se espera encontrar en $Cb(U_{i+1})$
- $Cb(U_{i+1})$: El *backward-looking center*, el elemento más alto de $Cf(U_i)$ que se realizará⁴ en el siguiente enunciado U_{i+1} .

El criterio de ordenación usado en Grosz *et al.* (1995) ordena los elementos de la lista C_f mediante roles gramaticales. De esta forma, las entidades con rol de sujeto se prefieren a aquellas que lo tienen de objeto y los objetos se prefieren a los otros (complementos circunstanciales, etc.).

El *centering* define un orden de preferencia basado en técnicas para efectuar un cambio de tópico como se muestra en la tabla 3.1

Al partir de las estructuras de datos previamente definidas y del criterio de ordenación anterior, el núcleo de la teoría se basa en dos reglas de *centering*:

³ El concepto de foco local representa a la entidad que tiene mayor interés en el enunciado actual y que se convierte en el principal candidato a ser referido mediante una anáfora según este marco.

⁴ Cuando se habla sobre coherencia del discurso se usa el término *realizar*, para indicar que un elemento es referido (realizado) por otro en un enunciado posterior.

	$Cb(U_i) = Cb(U_{i-1})$	$Cb(U_i) \neq Cb(U_{i-1})$
$Cb(U_i) = Cp(U_i)$	continuación	desplazamiento
$Cb(U_i) \neq Cp(U_i)$	retención	desplazamiento

Tabla 3.1. Tipos de transición en el *Centering*

Regla 1: Si cualquier miembro de $Cf(U_i)$ es realizado por un pronombre en U_{i+1} , entonces $Cb(U_{i+1})$ debe ser un pronombre.

Regla 2: las secuencias de continuaciones se prefieren a las secuencias de retenciones, y las secuencias de retenciones se prefieren sobre las secuencias de desplazamientos.

Algoritmo de *centering* (Brennan *et al.*). Tras la propuesta inicial del *centering* hay numerosos estudios de refinamiento y extensión a la teoría. Una de estas interpretaciones fue desarrollada por Brennan *et al.* (1987) con el fin de aplicarla a la resolución de la anáfora pronominal. Los autores basaron su algoritmo en las relaciones de transición entre enunciados que refinan la propuesta inicial de Grosz *et al.* (1983) tal y como se muestra en la tabla 3.2

	$Cb(U_i) = Cb(U_{i-1})$ o indefinido $Cb(U_{i-1})$	$Cb(U_i) \neq Cb(U_{i-1})$
$Cb(U_i) = Cp(U_i)$	continuación	desplazamiento-débil
$Cb(U_i) \neq Cp(U_i)$	retención	desplazamiento-fuerte

Tabla 3.2. Refinamiento de los tipos de transiciones según Brennan *et al.*

De esta forma, el algoritmo propuesto por los autores se construye con los siguientes pasos:

1. Cuando se encuentra un pronombre en U_i , los elementos en $Cf(U_{i-1})$ se prueban generando todas las asignaciones posibles de $Cb(U_i)$ a elementos en $Cf(U_{i-1})$ teniendo en cuenta que el antecedente del pronombre será $Cb(U_i)$, y la lista Cf se ordenará usando criterio gramatical (sujeto > objeto directo > objeto indirecto > otros > complementos circunstanciales).
2. Las asignaciones del pronombre a las entidades en Cf se filtran usando restricciones (morfosintácticas, semánticas, etc.).

3. Finalmente, la lista de asignaciones se ordena según los tipos de transición como preferencias. De esta forma se prefiere la *continuación* a la *retención*, ésta al *desplazamiento-débil* y ésta al *desplazamiento-fuerte*. Así la asignación de mayor orden se considerará la solución de la anáfora.

Adaptación del *centering* a diálogos (Byron y Stent). Para sistemas de diálogo, los beneficios de usar la teoría del *centering* incluyen una mejoría en la resolución de la anáfora y la generación de expresiones anafóricas más coherentes. Sin embargo, en la propuesta original no queda claro cómo adaptar la teoría para el discurso de diálogos. Por este motivo se han usado algunas variaciones de esta teoría para el tratamiento de diálogos. Destaca entre otras la propuesta de Byron y Stent (1998).

Para adaptar el *centering* a diálogos, Byron y Stent enumeran una serie de cuestiones a tener en cuenta. Entre otras:

1. ¿Cómo se definen los límites del enunciado? Es difícil precisar los límites de un enunciado en diálogo hablado, y su determinación afecta a las listas *Cf*. El hecho de definir cómo se rompen los turnos en enunciados tiene un gran impacto en el éxito del modelo.
2. ¿Se deben incluir en *Cf* los participantes del diálogo, referidos mediante pronombres de primera y segunda persona considerados como entidades del discurso?
3. ¿Qué enunciado debería ser considerado previo para localizar $Cf(U_{i-1})$, el enunciado previo del mismo hablante o el enunciado inmediatamente precedente sin importar su hablante?
4. ¿Qué debería hacerse con los enunciados abandonados o parciales y aquellos que no contienen entidades del discurso?

Sobre la base de las cuestiones anteriores, las autoras proponen tres modelos de *centering* distintos:

Modelo 1: *Cf* puede contener entidades referidas al uso de pronombres de primera y segunda persona, y $Cb(U_i)$ puede ser una entidad si es el elemento más alto de $Cf(U_{i-1})$ que aparece en $Cf(U_i)$. Sólo $Cf(U_{i-1})$, sin importar su hablante se busca en $Cb(U_i)$.

Modelo 2: No incluye pronombres de primera y segunda persona en Cf . $Cb(U_i)$ es el elemento de mayor rango en $Cf(U_i)$ que ocurre en el enunciado más reciente del hablante actual, o en el enunciado más reciente de otro hablante. La recencia (proximidad temporal) toma precedencia, así si un elemento de Cf del enunciado más reciente (sin importar el hablante) aparece en $Cf(U_i)$, se considera que es el $Cb(U_i)$, incluso si un elemento de mayor rango de la lista Cf de otro enunciado anterior está en $Cf(U_i)$.

Modelo 3: No incluye pronombres de primera y segunda persona en Cf . Sólo $Cf(U_{i-1})$, sin importar el hablante, se busca para $Cb(U_i)$.

Sobre la base de los modelos anteriores, Byron y Stent plantean las siguientes adaptaciones para responder a las cuestiones planteadas anteriormente. Así, para definir los límites de los enunciados se remiten al uso de un transcriptor. En este sentido, los transcriptores suelen separar los enunciados según los cambios de hablante o mediante la detección de pausas largas, o bien si la semántica o la sintaxis del lenguaje indica el fin de un enunciado. En el caso de las oraciones compuestas se separa cada cláusula no subordinada en un nuevo enunciado.

Por otra parte, en cuanto a la selección de elementos para Cf , se consideran todos los sintagmas nominales del enunciado. No se incluyen las entidades referidas mediante constituyentes nominales complejos tales como los infinitivos, asociaciones (parte/subparte) o las elipsis. La inclusión o no de las entidades referidas mediante primera y segunda persona dependerá del modelo usado. Como criterio de ordenación se usa el criterio gramatical así como la posición del componente en la oración para resolver el empate.

Por último, se propone no contar con los enunciados vacíos para la determinación del $Cf(U_{n-1})$. Puesto que la estructura del discurso en los diálogos se anota manualmente, U_{n-1} es el enunciado previo en el segmento del discurso, pero no necesariamente en orden lineal. En el modelo 2, para definir el Cb se usa el elemento más alto en el ranking de Cf del enunciado previo del actual ha-

blante o el del enunciado previo del otro hablante. Sin embargo, en los modelos 1 y 3 sólo se considera el enunciado inmediatamente precedente.

En los experimentos realizados por Byron y Stent, se usó un diálogo para entrenar a los anotadores que marcaron manualmente la solución de la anáfora. Tras la anotación del resto, los anotadores compararon sus resultados concertando una versión de reconciliación que se usó para producir los resultados. Así se calculó la precisión contando el número de enunciados que difieren de la información reconciliada.

Como conclusiones se obtuvo que sólo 10 de 664 enunciados violaron la regla 1 del *centering*, de tal forma que las suposiciones de la teoría que ligán el foco local a los pronombres parece mantenerse en el diálogo. El modelo 1, que incluye a los participantes como entidades del discurso, responde mejor en la evaluación. Sin embargo, Byron y Stent han detectado diversos problemas en la aplicación del método. Entre ellos:

- Cantidad excesiva de *Cbs* vacíos: Todos los modelos dejan al menos un 52% de enunciados no vacíos sin predicción de *Cb*, es decir, que $Cf(U_{i-1})$ y $Cf(U_i)$ están disjuntos.
- Poca coincidencia de *Cb* con el tópico real: Para aquellos enunciados donde ha podido seleccionarse un *Cb*, éste coincide con el tópico real sólo entre un 21% y 35% de las veces. Luego los modelos definidos son pobres en la predicción del foco local.
- Gran cantidad de transiciones caras frente a las baratas: Strube y Hahn (Strube & Hahn, 1999) proponen un método de evaluación de un modelo contra la regla 2 del *centering*, mediante la medición del coste de la carga de inferencia del oyente. Una transición barata tiene $Cb(U_i) = Cp(U_{i-1})$, en cualquier otro caso es cara. Los modelos con gran porcentaje de transiciones baratas reflejan mejor las nociones humanas de coherencia. Los tres modelos producen un porcentaje muy bajo de transiciones baratas en el experimento, especialmente cuando se comparan con el resultado de Strube y Hahn del 80%.

Como conclusión, las autoras creen que el comportamiento del *centering* en los diálogos es consistente con lo que se ha encontrado

en los monólogos. Sin embargo, la utilidad de los modelos definidos es bastante cuestionable.

Centering Funcional (Strube). Uno de los puntos fuertes del *centering* es la forma de ordenar los elementos de la lista *Cf*. En Grosz *et al.* (1995) se propone una ordenación mediante diferentes factores, pero en realidad únicamente se usa el criterio gramatical. Sin embargo, los trabajos de Strube (Strube, 1998) y Strube y Hahn (Strube & Hahn, 1999), demuestran la dificultad de la definición de roles gramaticales en lenguajes de orden libre de palabras como puede ser el alemán o el español. Por este motivo, proponen usar otro criterio de ordenación basado en información sobre la familiaridad de las entidades del discurso. Dicha información es crucial, al menos en los lenguajes de orden libre de palabras.

De acuerdo con el criterio de ordenación de Strube, hay dos conjuntos de expresiones distintos: las entidades del discurso conocidas para el oyente (*hearer-old*) (Old) y las nuevas para el oyente *hearer-new* (New). Así en el conjunto Old se incluyen las entidades del discurso *evocadas*, es decir, expresiones correferentes resueltas (anáfora pronominal y nominal, nombres propios que ya han sido mencionados previamente, pronombres relativos, aposiciones, etc.) y las conocidas pero no usadas que sean nombres propios y títulos. El resto de entidades se asignan al conjunto New.

De esta forma, el criterio básico para la resolución de la anáfora pronominal prefiere las entidades Old sobre las New⁵.

Así, Strube (Strube, 1998) propone la siguiente adaptación al modelo del *centering*:

- La lista *Cf* se sustituye por la lista de entidades de discurso relevantes (S-list) que contiene aquellas entidades del discurso que han sido referidas en el enunciado actual y en el previo.
- Los elementos de la lista S-list se ordenan de acuerdo al criterio básico definido anteriormente y la información sobre la posición:

⁵ Para resolver otras referencias como las descripciones definidas, los autores proponen incluir un conjunto Med con la información que se puede inferir, aunque esto no es necesario para la anáfora pronominal (Strube & Hahn, 1999)

Si $x \in Old$ y $y \in New$, entonces x precede a y .

Si $x, y \in Old$ o $x, y \in New$,

entonces si *enunciado*(y) precede a *enunciado*(x),

entonces x precede a y ,

si *enunciado*(y) = *enunciado*(x) y $pos(x) < pos(y)$,

entonces x precede a y .

- Puesto que no hay una definición clara de los que se considera como enunciado, los autores adoptan el siguiente criterio: las cláusulas verbales se definen como enunciados por sí mismo, mientras que las cláusulas no verbales se procesan con la principal constituyendo un único enunciado.

De esta forma, Strube propone el algoritmo siguiente:

1. Si se encuentra una expresión de referencia,
 - a) si es un pronombre, comprobar los elementos de la lista S-list por orden hasta que alguno sea válido
 - b) actualizar la S-list con la información de la expresión de referencia
2. Si se termina el análisis del enunciado U, eliminar todas las entidades del discurso de la lista S-list, que no hayan sido referidas en U.

La evaluación del algoritmo obtuvo una precisión del 85,4% mejorando los resultados del algoritmo de *centering* propuesto por Brennan *et al.* (1987), que sólo alcanzó el 72,9% cuando fue aplicado al mismo corpus.

Adaptación del *centering* funcional a diálogos (Eckert y Strube). Eckert y Strube (1999a) proponen un método para clasificar y resolver la anáfora individual (referente a objetos reales) y la anáfora abstracta (referente a objetos abstractos) de los pronombres personales y demostrativos. Para la segmentación se basan en el esquema propuesto por Carletta *et al.* (1997) y para la resolución de la anáfora se basan en el modelo alternativo al *centering* propuesto por Strube (1998).

La hipótesis planteada por Eckert y Strube es que el estado atencional de los participantes del discurso tiene relación con los

actos de diálogo, lo que se usa por los hablantes para indicar que se alcanza el *common ground* o conocimiento compartido del tópico (expresión que indica que todos participantes coinciden en su conocimiento sobre el tópico al que se está haciendo referencia, el objetivo del diálogo). Esta suposición la obtienen de la teoría sobre contribuciones de Clark y Schaefer (1989).

Para resolver las cuestiones básicas que plantearon Byron y Stent (1998) sobre el tratamiento de los diálogos, se adoptaron los siguientes criterios:

- Para definir los límites de los enunciados, los anotadores siguen un criterio independiente de la puntuación. Así evitan el problema de fiabilidad en cuanto a puntuación que genera el lenguaje.
- Los enunciados sin entidades del discurso son detectados y olvidados por el modelo.
- Los actos de diálogo actúan como marcadores que indican qué entidades han sido introducidas en el modelo de discurso y cuáles no.

La anotación efectuada divide los actos de diálogo en: Iniciaciones (*Initiations*, I) y Reconocimientos (*Acknowledgments*, A) según la jerarquía de Carletta *et al.* (Carletta *et al.*, 1997). Además, los enunciados que funcionan como A pero que tienen contenido semántico se etiquetan como *Acknowledgment/Initiation* (A/I). Los pares I y A adyacentes forman una *Unidad de Sincronización* (SU). Los SU's tienen doble función: por un lado, después de cada SU, las entidades del discurso no evocadas en él se eliminan de la lista S-list. Por otra parte, un acto I no replicado por el otro participante se queda fuera del SU. Así, las entidades de discurso que se extraen de actos I fuera de un SU se eliminan de la S-list y no se toman como antecedentes (no contribuyen al conocimiento compartido del tópico).

Para asegurar la fiabilidad de los resultados, se evalúan los criterios de los anotadores de la siguiente forma: la segmentación de turnos en actos del diálogo se convierte en una tarea de clasificación considerando que ser límite de un acto de diálogo es una clase, y no serlo, otra. A esta tarea se le aplica un test de

fiabilidad mediante *kappa statistic* (Carletta, 1996) y se comprueba su fiabilidad. Por otra parte, la tarea de clasificar los actos de diálogo en *Initiations* (I), *Acknowledgments* (A), *Acknowledgment/Initiation* (A/I) o *No actos del diálogo* (No) también se evalúa con *kappa*. Lo mismo respecto a la clasificación de los pronombres. Y finalmente, se evalúa la coindexación de la anáfora de la siguiente forma: los anotadores marcan los antecedentes y los coindexan con la anáfora para comprobar finalmente la salida del algoritmo. Entonces se comparan los resultados de ambos anotadores y se llega a una versión de reconciliación, proporcionando la medida de precisión de cada anotador respecto a la versión reconciliada.

Finalmente se evalúa el algoritmo para anáfora individual proporcionando una precisión del 66,2% y una cobertura del 68,2% y para la anáfora abstracta obteniendo una precisión del 63,6% y una cobertura del 70%. Según los autores, las causas de error fueron debidas a errores en la clasificación de anáforas entre individuales o abstractas, y que el dominio del antecedente esté limitado a una unidad de sincronización única. Sin embargo, esta restricción hace que muchas anáforas puedan ser resueltas correctamente.

3.2.4 Métodos integrados basados en restricciones y preferencias

Una línea de investigación paralela a la vista hasta ahora considera que la anáfora, como un fenómeno del discurso, debe ser tratada integrando el conocimiento que proporcionan varias fuentes de información (métodos democráticos).

En estos sistemas, los conocimientos se dividen en dos grupos: restricciones y preferencias. Las restricciones contienen una serie de reglas de obligado cumplimiento por cualquier candidato a antecedente de una anáfora. De esta forma, el candidato que incumple una de ellas será automáticamente rechazado por el sistema. Por otra parte, las preferencias agrupan una serie de reglas cuyo cumplimiento no es necesario, pero que en caso de cumplirse destacan a ese candidato respecto a aquellos que no las cumplen.

De esta forma, las restricciones se usan para descartar candidatos imposibles y las preferencias para ordenar los candidatos. El candidato que alcance mejor posición en esa ordenación es elegido como antecedente.

Entre las aproximaciones basadas en restricciones y preferencias destacamos los trabajos de Mitkov (1998), Baldwin (1997) y Ferrández *et al.* (1999) que se basan en una combinación de conocimiento lingüístico (léxico, morfológico, sintáctico y/o semántico) para la resolución de la anáfora. Estas aproximaciones aplican el conocimiento lingüístico en forma de restricciones y preferencias siguiendo la propuesta de Carbonell y Brown (1988) y de Rich y LuperFoy (1988). En estos trabajos, los sistemas de restricciones y preferencias se proponen como una técnica para la combinación de diferentes fuentes de información.

Las aproximaciones basadas en restricciones y preferencias aplican, intuitivamente, los siguientes tres pasos: a) obtención de posibles candidatos, b) aplicación de restricciones, y c) aplicación de preferencias.

El sistema de restricciones y preferencias debe obtener una lista con todos los candidatos posibles para ser el antecedente de la anáfora. Este paso es fundamental en el proceso restante ya que un fallo en la obtención de la lista por defecto causaría la exclusión de antecedentes válidos, mientras que una definición por exceso causa grandes listas con un incremento considerable de la posibilidad de fracaso. Generalmente, los sistemas de resolución de la anáfora con restricciones y preferencias definen la lista mediante una ventana de n oraciones anteriores a la anáfora, donde n es variable dependiendo de la clase de anáfora.

Una vez que ha sido definida la lista de los candidatos posibles, el conjunto de las restricciones se aplican con el fin de eliminar los candidatos incompatibles. El sistema de restricciones usa tradicionalmente información léxica, morfológica, sintáctica y semántica para la definición de las reglas.

Por último, una vez que se han eliminado los candidatos incompatibles, y cuando la lista mantiene más de un candidato, se aplican las preferencias con el fin de conseguir un único antecedente. En este caso, a diferencia de las restricciones, las preferen-

cias tienen una probabilidad de cumplimiento asociada menor al 100%. Los candidatos que cumplen una preferencia tienen mayor probabilidad de ser el antecedente que aquellos que no lo cumplen. El sistema de preferencias debe ser diseñado teniendo en cuenta que sólo debe quedar un candidato al final. Este candidato final se propondrá como antecedente para la anáfora. La información léxica, morfológica, sintáctica y semántica se usa habitualmente para definir el sistema de preferencias.

Los trabajos de Mitkov (1998) y Ferrández *et al.* (1999) muestran que los sistemas de resolución de anáfora basados en restricciones y preferencias tienen resultados satisfactorios cuando se aplican a discurso monólogo. Sin embargo estos trabajos carecen de una propuesta adecuada para el espacio de accesibilidad anafórica (tal y como se demostrará más adelante), es decir, el espacio de donde se puede extraer la lista de candidatos. Este hecho provoca una falta de consistencia para el tratamiento de otras clases de discurso en los que este espacio queda marcado por su estructura de discurso como ocurre por ejemplo, los diálogos.

3.2.5 Métodos alternativos basados en información estadística

Los métodos de resolución de la anáfora basados en información estadística intentan proporcionar un modelo probabilístico desde el estudio de corpus que previamente han sido anotados con aquella información que se cree influyente en la resolución de la anáfora.

Entre estos trabajos destacamos el propuesto por Dagan e Itai (1990) quienes desarrollaron un método de desambiguación para el pronombre *it* basado en patrones que asociaban la ocurrencia de una anáfora con la posición sintáctica (sujeto o objeto) que ocupaba su antecedente. Estos patrones se construían previo aprendizaje automático que era realizado sobre un corpus de entrenamiento. Los autores reportaron una tasa de éxito del 87%.

También en este sentido, la propuesta de Connolly *et al.* (1994) es una aproximación de resolución de la anáfora mediante aprendizaje computacional. En este caso, se basa en el uso de corpus de

entrenamiento en los que los antecedentes se anotan manualmente con una serie de características que los clasifican. Así, una vez realizado el aprendizaje automático se extraen unas preferencias de ciertas anáforas por ciertas clases de antecedente.

Siguiendo esta misma idea, pero aplicado a los diálogos, uno de los métodos estadísticos más destacables es la Teoría de probabilidad de antecedentes (*Antecedent-likelihood theory*) de Rocha (Rocha, 1999) que describiremos a continuación.

Teoría de probabilidad de antecedentes (Rocha). La teoría de la probabilidad de los antecedentes propuesta en Rocha (1999) se basa en la construcción de un modelo probabilístico que pueda especificar, dado un tipo de anáfora, la probabilidad que tiene una determinada entidad de ser su antecedente.

Para ello se basa en un esquema de anotación en el que se anotan cuatro propiedades para cada caso de anáfora. La primera es el tipo de anáfora que se clasifica según su categoría gramatical. La segunda es el tipo de antecedente concentrándose en la dicotomía implícito o explícito y alguna información adicional como si es referencial o no. La tercera propiedad define el rol tópico del antecedente e indica si pertenece a alguno de los elementos de la jerarquía de topicalidad del discurso: tópico global del discurso, tópico local del segmento, y tópico local del subsegmento. También se indica si se trata de un elemento temático relacionado de alguna forma con el tópico, pero sin serlo, o si se trata de un mecanismo de enfoque (marca el foco del hablante). Por último, la cuarta propiedad incluye la estrategia de procesamiento. Esta propiedad pretende ser una guía para la resolución de la anáfora en el sistema de procesamiento automático. Así, Rocha define la siguiente clasificación de estrategias de procesamiento:

1. Procesos léxicos:
 - a) Conocimiento compartido: la resolución se basa en información compartida por los participantes.
 - b) Conocimiento del mundo: se trata de información conocida por todo el mundo independientemente del lugar o el tiempo donde transcurre el diálogo.

- c) Indicación léxica: la anáfora se resuelve gracias a la información que se extrae de un diccionario.
 - d) Repetición léxica: la repetición de elementos léxicos establece el enlace entre anáfora y antecedente como, por ejemplo, el caso de una referencia a una entidad en singular a través de su plural.
2. Procesos sintácticos:
- a) Búsqueda del primer candidato: es el caso típico de resolución pronominal descrito en el algoritmo básico de Hobbs (Hobbs, 1986). Se busca hacia atrás el primer candidato que cumpla las restricciones sintácticas y morfológicas.
 - b) Búsqueda del primer candidato en cadena: variación del anterior en el que el primer candidato coincide con otra anáfora que enlaza con un antecedente común.
 - c) Memoria literal: especialmente en el caso de las anáforas abstractas, puede ser necesario recordar exactamente un enunciado o fragmento del diálogo para resolver, como en el pronombre *lo* en *yo también lo pienso*.
 - d) Paralelismo sintáctico: la identificación del antecedente puede depender de la formación de un paralelismo sintáctico con un enunciado previo. Así algunos casos de doble anáfora en el sujeto y el objeto de un enunciado se pueden resolver tomando como antecedentes el sujeto y el objeto respectivamente de un enunciado anterior.
3. Procesos del discurso:
- a) Modificación de antecedente: el antecedente y la anáfora únicamente difieren en el determinante usado. Este es el caso de un antecedente que se usa por primera vez con un artículo indeterminado y posteriormente como determinado.
 - b) Conocimiento del discurso: la identificación del antecedente sólo puede ser resuelta mediante el procesamiento del discurso completo.
 - c) Selección de un miembro de conjunto: la anáfora hace referencia a un elemento contenido en un conjunto previamente mencionado. Es el caso típico de la resolución de la anáfora superficial numérica o de algunos casos de one-anáfora.

- d) Creación de conjunto: la referencia se hace mediante la creación de un conjunto que engloba a una serie de antecedentes previamente mencionados.
 - e) Referencias secundarias: cuando se trata de pronombres de primera o segunda persona.
 - f) Deixis: cuando la resolución de la anáfora depende del entorno físico en donde se desarrolla el diálogo.
4. Coloquialismos: las anáforas pronominales que aparecen en una expresión coloquial pueden seguir estrategias particulares de resolución.

La recolección estadística de estas cuatro propiedades en cada caso de anáfora proporciona un modelo estadístico que combina cuatro variables entre sí mediante árboles de probabilidad. En la figura 3.1 se muestra un ejemplo del árbol probabilístico presentado en Rocha (1998) para la resolución de la anáfora pronominal en inglés. En este caso, la categoría gramatical de la anáfora es la raíz, la estrategia de procesamiento ocupa el primer nivel, a continuación los roles tópicos del antecedente y finalmente los tipos de antecedente.

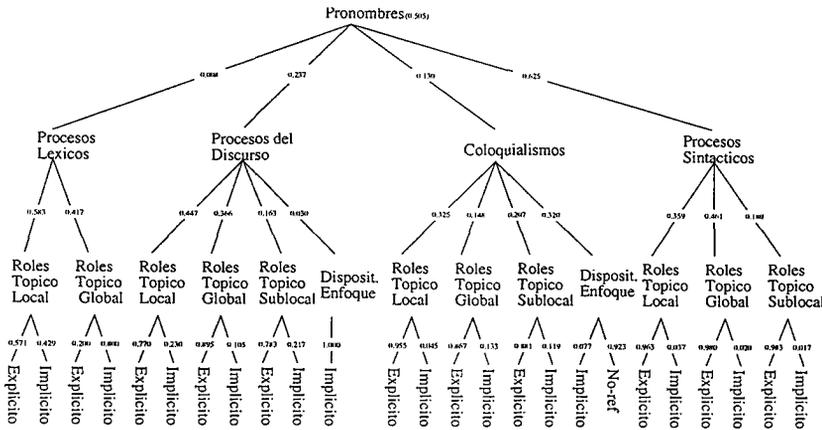


Figura 3.1. Árbol probabilístico de la anáfora pronominal en inglés

Del árbol de probabilidad se pueden extraer las siguientes conclusiones: dado un pronombre como anáfora, existe una probabi-

alidad de 0,625 de ser tratado mediante procesos sintácticos siendo por tanto la estrategia más probable. Además, el rol tópico más probable para ese antecedente es el tópico global con una probabilidad de 0,461, y con una probabilidad de 0,980 el antecedente será explícito.

Los árboles de probabilidad se generan automáticamente mediante un programa computacional. Dicho programa es capaz de generar árboles en todas las ordenaciones posibles, por lo que la información se puede examinar desde otros puntos de vista.

Sin embargo, la principal crítica que se atribuye en general a todos los modelos probabilísticos es su gran dependencia del dominio en el que trabaja, de los textos utilizados para el aprendizaje, y especialmente en el caso de los diálogos, de los propios participantes. Esto se debe a la falta de uso de reglas lingüísticas que tengan validez universal y hace que los métodos probabilísticos tengan que ser entrenados en cada dominio donde se aplique. Además, los resultados obtenidos sólo sirven como una orientación para ser aplicados en un módulo de resolución posterior, ya que en muchos casos los valores de probabilidad obtenidos no indican que una de las ramas del árbol sea claramente superior al resto, y por tanto nunca debería ser descartado el resto de posibilidades.

3.3 Influencia de la estructura del diálogo en la resolución de la anáfora

Fox (1987) investiga la anáfora basándose en sus relaciones con la estructura de discurso. En este sentido, usa el análisis convencional para definir una estructura jerárquica del discurso. Según este esquema, la unidad básica usada para definir la estructura de la organización de un discurso es el par de adyacencia.

Fox basa su estudio en la creencia de que existe una fuerte relación entre el flujo de información en un texto, la estructura del texto y el uso de la anáfora. De hecho es sabido que aquellos antecedentes que se encuentran en el “foco” o en “la consciencia del oyente” pueden ser pronominalizados. Así la autora destaca que cualquier tratamiento de la anáfora debe ser entendido desde

la estructura jerárquica así como desde el tipo de discurso usado como fuente de entrada.

En este sentido, se distinguen básicamente dos tipos de discurso, los conversacionales producidos por más de una persona y los no conversacionales escritos, y justifica el uso de modelos de representación de discurso distintos basándose en las siguientes diferencias:

- Los discursos conversacionales son muy interactivos y los modelos que lo representan deben capturar las relaciones sociales que se mantienen entre cada una de las piezas de la estructura (básicamente relaciones entre los enunciados).
- Por otra parte, el discurso no conversacional escrito es puramente informativo, y por tanto los modelos de discurso que lo representen deberán recoger las relaciones informativas que existen entre las piezas de la estructura (en este caso, relaciones entre proposiciones).

De acuerdo a los estudios de Fox, la forma básica de distribución de herramientas anafóricas en conversación se rige por la siguiente norma:

“La primera mención de un referente en una secuencia se realiza mediante un sintagma nominal completo. Después de esto, mediante el uso de un pronombre, el hablante subraya un conocimiento de que la secuencia anterior aún no ha quedado cerrada”.

A continuación analizaremos cada una de las situaciones en las que la anáfora se usa para indicar que la secuencia no ha sido finalizada aún.

3.3.1 Uso de anáforas dentro de un par adyacente

Según Fox (1987), la situación más obvia de una secuencia no cerrada se produce dentro del mismo par adyacente. Los hablantes generalmente usan anáforas en el turno de reacción para hacer referencia a un antecedente mencionado en el turno de iniciación. Así ocurre, por ejemplo, en el siguiente par adyacente:

(69) IN_I A: ¿Has encontrado ya tu reloj _i?

IN_R B: No, creo que **lo_i** he perdido

En este caso, el pronombre *lo* es usado por el hablante (B) en el turno de reacción para hacer referencia a *tu reloj*, para dar a conocer al otro participante (A) que se encuentran en la misma secuencia.

En otra situación, el hablante puede introducir una anáfora para añadir a un primer turno de intervención iniciativa otro que lo complementa (caso de una expansión previa del par adyacente):

(70) **IN_I A:** ¿Has encontrado ya tu reloj *i*?

IN_I A: **Ese_i** que buscabas la semana pasada

IN_R B: No, creo que **lo_i** he perdido

En el ejemplo, el hablante (A) introduce el pronombre *ese* para hacer referencia a un antecedente introducido por él mismo en el turno de intervención iniciativa anterior remarcando que debe ser completado con la nueva intervención.

La misma situación puede suceder cuando se producen expansiones posteriores del par adyacente, como ocurre en el siguiente ejemplo:

(71) **IN_I A:** ¿Quién vendrá a cenar con nosotros?

IN_R B: Vendrán Juan Antonio y Macu, su mujer_i

IN_R B: Creo que **la_i** conociste en la boda de Luis.

Encontrar un turno de intervención reactiva no indica necesariamente que haya finalizado el par adyacente, y en muchos casos se usa la anáfora para justificar esa expansión.

3.3.2 Uso de anáforas en el siguiente par adyacente

Por otra parte, la finalización de un par adyacente tampoco indica necesariamente que la secuencia haya finalizado. En muchas ocasiones, un segundo par adyacente puede vincularse al anterior de forma que la secuencia continúa a través de más de un

3.3 Influencia de la estructura del diálogo en la resolución de la anáfora 99

par adyacente. La anáfora es una herramienta que siempre indica continuidad de algo previamente comenzado, por ello se usa en estos casos. En el siguiente ejemplo se muestra cómo un primer par adyacente puede seguirse por otro de forma que el turno de intervención iniciativa del segundo incluye un pronombre haciendo referencia a un elemento del primer par:

(72) **IN_I A:** Perdona, ¿sabría decirme a qué hora sale el tren para Madrid_i?

IN_R B: Creo que a las tres.

IN_I A: ¿Pero ese_i es el talgo?

IN_R B: No le puedo decir seguro. Pregunte mejor en la ventanilla.

Hay dos causas que pueden provocar que dos o más pares adyacentes queden vinculados:

- La serie. En una serie, un par adyacente puede ser del mismo tipo que el anterior. Es decir, sería el siguiente en una serie de objetos similares.
- La post-elaboración. Uno de los pares proporciona una serie de noticias, o informa de algo que pares adyacentes posteriores matizan.

3.3.3 Uso de anáforas en pares adyacentes anidados

No siempre las dos partes de un par adyacente deben estar físicamente contiguas. En algunos casos pueden quedar separadas por pares adyacentes que tras aclarar alguna cuestión permiten volver a la secuencia iniciada (situación de expansión intermedia del par adyacente). En estas circunstancias, el uso de la anáfora a larga distancia puede ser necesario para conseguir los siguientes objetivos:

1. Cuando se trata de una situación en la que hay que cumplir un objetivo global a través del cumplimiento de pequeños sub-objetivos (caso típico de los diálogos orientados a tareas), los

hablantes suelen referirse mediante una anáfora al objetivo global mientras dialogan sobre uno de los subobjetivos:

(73) **IN_I A:** Mire, quería saber el horario del tren para Madrid_i?

IN_I B: Si. ¿Qué día viajará usted?

IN_R A: Pues a ser posible, el martes día quince.

IN_I B: Ya, y ¿**lo_i** quiere por la mañana o por la tarde?

IN_R A: Por la tarde.

IN_R B: Pues tiene un talgo a las diecisiete treinta y cinco y un intercity a las veinte cero cinco.

En el ejemplo anterior, el hablante A inicia una secuencia requiriendo del hablante B un objetivo global, conseguir un horario. Sin embargo para obtenerlo es necesario que el hablante B ponga en marcha algunos subobjetivos intermedios con el fin de obtener la información necesaria para proporcionar una respuesta. En el segundo de los pares adyacentes iniciados por B, el hablante usa el pronombre *lo* para referirse al objetivo global.

2. Cuando la secuencia iniciada por el hablante es interrumpida (por él mismo o por otro hablante) al introducir una nueva secuencia, entonces, la secuencia interrumpida podrá retomarse cuando las internas finalicen. Para indicar que vuelve a retomar la secuencia iniciada anteriormente, el hablante suele hacer uso de la anáfora:

(74) **IN_I A:** ¿Puedo usar tu teléfono_i?

IN_I B: Siempre estás igual. ¿Has probado a usar tu teléfono móvil_j?

IN_R A: Sí, lo siento, pero es que **lo_j** he perdido. Tengo que comprarme uno nuevo.

3.3 Influencia de la estructura del diálogo en la resolución de la anáfora 101

IN_I A: Bueno, ¿me lo_i dejas o no?

IN_R B: Sí, toma.

En este ejemplo, la secuencia iniciada por el hablante A en el primer turno, se ve interrumpida por un comentario del hablante B que genera un nuevo par adyacente. Para relanzar la secuencia, el hablante A hace uso del pronombre *lo*.

3.3.4 Uso de la anáfora en segmentos con más de un antecedente

Finalmente, Fox realiza un estudio sobre el uso de la anáfora en segmentos del diálogo en los que puede existir más de un antecedente activo, desmintiendo la creencia clásica de que la aparición de un segundo antecedente en el discurso implica inmediatamente el uso de sintagmas nominales completos para hacer referencia de nuevo al primero. Fox demuestra que hay dos situaciones en las que esto no ocurre así:

En primer lugar, cuando los antecedentes son de distinto género se suelen usar simultáneamente expresiones anafóricas de distinto género.

En segundo lugar, cuando hay dos antecedentes del mismo género presentes en la conversación, el hablante hace uso de la anáfora para indicar al oyente que el antecedente al que se hace referencia se debe encontrar en un par adyacente activo (perteneciente a una secuencia no cerrada, porque sucedan cualquiera de las tres situaciones aludidas anteriormente). Puede ocurrir, por tanto, que un primer antecedente esté aún en un par adyacente activo, mientras que un segundo antecedente, aunque sea nombrado con posterioridad, se encuentre en un par adyacente no activo.

(75) **IN_I A:** Perdone. ¿Ha terminado ya mi pedido_i?

IN_I B: A ver. ¿Puede esperar un momento que consulte el ordenador?

IN_R A: Sí, no tengo prisa.

IN_R B: Pues sí, ya **lo_i** puede recoger cuando quiera.

En este ejemplo, el pronombre *lo* se usa para hacer referencia a *mi pedido* y no a *el ordenador* o a *un momento* que son antecedentes que han aparecido posteriormente. El motivo es que *mi pedido* aparece en un par adyacente que aún está activo, y los otros dos antecedentes, están en una secuencia que ya está finalizada.

Sin embargo, si los antecedentes estuvieran en pares adyacentes que pudieran estar activos simultáneamente, entonces, el hablante necesitaría el uso de otra información anexa (por ejemplo información semántica o conocimiento del mundo) para poder utilizar una anáfora o debería referirse a ellos mediante sintagmas nominales completos.

(76) **IN_I A:** ¿Me puede decir qué vale un billete de Alicante a Madrid?

IN_I B: Sí, vamos a ver. ¿Desea asiento de primera o segunda clase?

IN_R A: De primera.

IN_I B: ¿**Lo_i** quiere de ida y vuelta?

IN_R A: Sólo ida.

IN_R B: Pues son cuatro mil quinientas.

En este caso, los dos antecedentes *i* y *j* pertenecen a pares adyacentes activos. El primero porque pertenece a un par adyacente que contiene al par donde aparece la anáfora, y el segundo porque pertenece al par adyacente anterior y forman una serie. Sólo el conocimiento del mundo que tiene el oyente acerca de billetes y asientos hace que sepa que un billete puede ser de ida y vuelta pero raramente el asiento.

Si cambiáramos el enunciado por:

(77) **IN_I B:** ¿Lo_i quiere cerrado?

puede crear mayor confusión ya que el oyente podría interpretar que es *el asiento* y no *el billete* el que está cerrado. En este caso el hablante habría formulado la pregunta como:

(78) **IN_I B:** ¿Quiere el billete cerrado?

3.3.5 El tópico del diálogo y la anáfora

La mayoría de los diálogos, tanto aquellos orientados a tareas como los de extracción de información, tratan sobre algún tema que los participantes han reconocido como el tópico del diálogo.

Exceptuando el caso de los diálogos cortos, cada tópico global de un diálogo puede descomponerse en tópicos locales, y así generar segmentos de menor nivel. Así, trabajos como el de Grosz y Sidner (Grosz & Sidner, 1986) y el de Rocha (Rocha, 1998; Rocha, 1999) muestran que muchas de las anáforas generadas en un segmento del discurso hacen referencia al tópico local de ese segmento e incluso al tópico global del discurso. Esto hace que la identificación de los tópicos globales y locales del discurso por parte de los participantes sea una de las herramientas usadas para determinar el antecedente adecuado de posibles anáforas.

Sin embargo, hay ciertos factores que no hacen sencilla esta identificación. Generalmente, el tópico se puede identificar bien porque haya sido introducido explícitamente en el diálogo como tal, por ejemplo con frases como *Vamos a hablar de ...*, o puede haber sido definido implícitamente como tópico puesto que ha sido referido en varias situaciones. Sin embargo, ciertos tópicos abstractos pueden estar en la consciencia de los hablantes aunque no hayan sido introducidos explícitamente. Por ejemplo, dos hablantes podrían estar hablando de *política* como tópico general sin que ninguna intervención haya nombrado explícitamente el tópico. En este caso la identificación del tópico forma parte más de la semántica e incluso del conocimiento del mundo.

Por otra parte, la segmentación del tópico en subtópicos no siempre obedece a criterios semánticos del tipo hiperonimia-

hiponimia. En muchos casos puede obedecer más al contexto del propio diálogo que a una relación semántica entre el tópico y los subtópicos.

Por último hay que tener en cuenta que muchos diálogos, especialmente aquellos de larga duración, pueden tener bruscos cambios en el tópico global y pasar de un tema a otro. En ese caso, el participante deberá también cambiar bruscamente el antecedente para anáforas posteriores.

En cuanto a los tipos de anáfora que permiten tratar este tipo de referencias, podemos destacar que, puesto que se resuelve a larga distancia, en pocos casos se usan las referencias pronominales, sino más bien otros tipos de anáfora que incluyen mayor información para la identificación del antecedente.

3.4 Sistemas de diálogo, estructura del discurso y resolución de la anáfora

Como se mostró en la sección 2.2, los sistemas de diálogo, y en concreto su módulo de Gestión del discurso, generan y hacen uso de una estructura básica formada por pares adyacentes e incluso, en algunos casos, generan estructuras temáticas superiores como la definición de tópicos.

Por otra parte, tal y como se mostró en la arquitectura genérica presentada en Allen *et al.* (2000), el módulo de resolución de la referencia es utilizado consultivamente por el módulo de Gestión del discurso para permitir la interpretación y la generación de los actos de habla que constituyen los pares adyacentes, y consecuentemente, la estructura temática del diálogo.

Además, según dicha arquitectura, el módulo de resolución de la anáfora⁶ debería proporcionar salidas independientes del dominio, de tal forma, que fuera el propio Gestor del discurso quien, tras consultar con otras fuentes dependientes del dominio, pro-

⁶ Aunque el sistema de Allen trata el problema de la resolución de referencias en general, puesto que nuestro estudio se centra en la resolución de la anáfora, enfocaremos la discusión hacia el tratamiento concreto de este tipo de referencia

porcionara la solución final para la interpretación de la expresión de referencia.

Por lo tanto, lo que se pretende del módulo de resolución de la anáfora en un sistema de diálogo es que sea capaz de resolver las anáforas dentro del contexto definido por la propia estructura del diálogo y utilizando toda la información independiente del dominio de que dispone el módulo Gestor del discurso.

Así, el sistema de resolución de la anáfora que queremos plantear deberá contemplar ambas características. Por una parte, debe ser capaz de adaptarse al contexto definido por la estructura del diálogo ya que hace uso de dicha estructura como una de las fuentes de información que toma como entrada para la resolución de la anáfora. Por otra parte, debe ser un sistema independiente del dominio ya que toda la información de entrada empleada es absolutamente independiente del dominio.

La integración del módulo de resolución de la anáfora en un sistema de diálogo podría relegar al Gestor del diálogo el uso de dichas herramientas dependientes del dominio para proporcionar la interpretación final de la anáfora, siendo el objetivo principal del módulo de resolución anafórico actuar como filtro previo para generar una lista reducida de candidatos a antecedente de la anáfora.

Sin embargo, puesto que el objetivo principal de esta Tesis es el desarrollo y evaluación de un sistema para la resolución de la anáfora en los diálogos, consideraremos a partir de ahora que nuestro sistema debe resolver la anáfora proporcionando una única respuesta y como tal será evaluado, independientemente de su posible integración en un sistema de diálogo.

3.5 Conclusiones del capítulo

En este capítulo hemos realizado un estudio de los distintos tipos de relaciones anafóricas existentes en el diálogo junto con una exposición de los diferentes enfoques para su resolución.

Respecto a los primeros destacaremos que, de todos los tipos de relaciones anafóricas presentados, nos centraremos en concreto

en la resolución de las anáforas pronominales y de tipo adjetivas, en aquellos casos en los que el antecedente es de naturaleza individual.

Respecto a las diversas técnicas para la resolución de la anáfora, destacamos que algunas de ellas son aplicables directamente a los diálogos, y otras de ellas adaptables tras la realización de algunos cambios. Sin embargo, en pocos casos se plantean como una estrategia de resolución específica para los diálogos. Como ya se introdujo en el capítulo 1, en este trabajo proponemos un sistema de resolución de la anáfora en diálogos que se basa en una estrategia integrada democrática de restricciones y preferencias en la que se usarán fuentes de información tanto lingüística (léxica, morfológica y sintáctica) como estructural. Además, el sistema que aquí se presenta plantea la resolución de la anáfora en un sistema de diálogo en español, que pretende ser independiente del dominio y del sistema de diálogo en el que se aplica.

Por otra parte, tal y como se ha demostrado en los párrafos anteriores, existe una fuerte relación entre el flujo de información en un diálogo, la estructura del diálogo y el uso de la anáfora. Esta relación se demuestra con el uso de la anáfora como recurso empleado por los hablantes de un diálogo para hacer partícipe al oyente de la propia estructura del diálogo. Así se producen las siguientes situaciones de secuencias no finalizadas en las que se utiliza la anáfora para indicar esta circunstancia:

- en el par adyacente,
- en pares adyacentes consecutivos, y
- en el par adyacente que contiene a su vez otros pares adyacentes

En cualquiera de estos casos se considera que el par adyacente donde aparece el antecedente está activo respecto del par adyacente donde aparece la anáfora y por tanto, está justificado el uso de la anáfora.

Además de estas secuencias, el tópico del diálogo genera una secuencia de diálogo global, en la que está ampliamente justificado el uso de anáforas para hacer referencia precisamente a la entidad que conforma el tópico. Lo mismo ocurre con los subtópicos

que puedan generarse como descomposición del tópico global del diálogo, especialmente en largos diálogos.

Destacaremos, por último, que el sistema de resolución de la anáfora que proponemos, puesto que tiene como finalidad su integración en un sistema de diálogo, debe ser capaz de adaptarse a la estructura que genera dicho sistema como una de las fuentes de información que toma como entrada para la resolución de la anáfora.



Universitat d'Alacant
Universidad de Alicante



4. El espacio de la solución

Universitat d'Alacant
Universidad de Alicante

Una de las principales aportaciones de nuestro trabajo es que plantea la necesidad de definir un espacio de accesibilidad anafórico adecuado de donde se puedan extraer los candidatos a antecedentes para una determinada anáfora. Así la definición de este espacio se presenta como una de las tareas fundamentales para un sistema de resolución de la anáfora.

En este capítulo se desarrolla un estudio sobre el espacio de accesibilidad anafórica que genera la propia estructura del diálogo tomando como base las conclusiones extraídas en el capítulo 3 sobre las relaciones entre anáfora y estructura del diálogo. Esta estructura del diálogo se genera automáticamente en los sistemas de diálogo tal y como se mostró. Sin embargo, puesto que nuestro objetivo básico es el desarrollo y evaluación del sistema de resolución de la anáfora en diálogos independientemente del sistema de diálogo en el que se integre, se propone un esquema de anotación de la estructura del diálogo que servirá como base para la definición del espacio de accesibilidad anafórica estructural. A partir de esta propuesta se desarrolla un estudio empírico sobre 40 diálogos seleccionados al azar que demuestra los beneficios de usar esta propuesta de espacio de accesibilidad anafórica estructural frente a otras aproximaciones tradicionales. Por último se propone un método de estandarización del esquema de anotación utilizando el lenguaje de definición de estándares SGML para facilitar la construcción de herramientas de anotación.

4.1 Propuesta de un esquema de anotación para la estructura del diálogo

Como se mostró en la sección 3.3, la estructura del diálogo en la consciencia del oyente, construida mediante pares adyacentes y sus relaciones, y mediante la definición de los tópicos del discurso, influye decisivamente en la toma de decisiones para resolver la anáfora (Martínez-Barco & Palomar, 2000b). Por este motivo proponemos un esquema de anotación que sustituya el conocimiento que sobre la estructura del diálogo dispondría un oyente, para que pueda ser aplicado por un sistema automático de resolución de anáfora.

Según estas conclusiones se propone a continuación un esquema de anotación basado en la estructuración de diálogos en español propuesta por Gallardo (1996) quien se basó a su vez en las teorías desarrolladas por Sacks *et al.* (1974) tal y como se mostró en la sección 2.4.1. De acuerdo con estas teorías, la unidad básica de conocimiento en el diálogo es el *movimiento* que informa al oyente sobre una acción, requerimiento o pregunta. Los movimientos se llevan a cabo a través de los *enunciados* que finalmente se agrupan en los *turnos*. Puesto que se trabaja con diálogos que han sido previamente transcritos, el turno se convierte en la primera unidad identificable puesto que aparece anotado en las transcripciones con la marca del hablante que actúa en ese turno. Además, los enunciados son fácilmente identificables por la lectura de signos de puntuación (., ?, !, ...) que los delimitan.

Destacaremos que se ha elegido este modelo de discurso y no otro fundamentalmente por tres razones:

- Puesto que es un mecanismo genérico de generación de diálogos, es aplicable a cualquier tipo de diálogo, tanto orientado a la tarea como orientado a la recuperación y/o extracción de información. Por tanto, el uso de un modelo de discurso genérico en un sistema de resolución de la anáfora nos permitirá su aplicación directa a cualquier tipo de aplicación y dominio. Con ello se adquiere una considerable ventaja sobre aquellos sistemas de resolución de anáfora dependientes de un modelo de discurso específico para un determinado dominio y/o tipo de aplicación.

4.1 Propuesta de un esquema de anotación para la estructura del diálogo 111

- El esquema de anotación definido a partir de este modelo de discurso puede aplicarse fácilmente por medio de procesos automáticos sin necesidad de consideraciones metalingüísticas. En este sentido, aunque en este trabajo el esquema de anotación ha sido aplicado manualmente, cualquier sistema de diálogos podría proporcionar dicha información automáticamente.
- Por último, puesto que nuestra hipótesis parte de los estudios de Fox (1987) sobre la influencia de la estructura del discurso en la anáfora, y dichos estudios se basan en este modelo, su aplicación resulta inmediata.

Además, para definir el tópico del diálogo apropiadamente se utilizarán los criterios definidos por Rocha (1999) sobre la definición del tópico del diálogo.

De esta forma, a partir de las marcas de estructura conocidas, los enunciados y los turnos, y mediante un método apropiado para la definición del tópico se plantea un esquema de anotación basado en la anotación de los siguientes elementos estructurales:

- **Turno:** Un turno es identificado por el cambio de hablante en el diálogo. Cada vez que cambia el hablante supone un nuevo turno. El turno, en diálogos transcritos, es fácilmente identificable puesto que las convenciones sobre transcripción de diálogos suelen incluir una marca especial para cambio de hablante. Por ejemplo, mediante la marca “<H1>...”, o “H1:”. Sin embargo, será necesario realizar una clasificación de los turnos para posteriormente construir los posibles pares adyacentes. Así será necesario distinguir entre dos clases primarias de turnos:

1. **Intervención (IT):** es aquel turno que añade información al diálogo. Los hablantes usan los turnos de intervención para proporcionar cierta información que facilita el progreso del tópico de conversación.

Las intervenciones, a su vez se deben clasificar como:

- a) **Intervención Iniciativa (IT_I)** cuando formulan invitaciones, requerimientos, ofertas, etc.
- b) **Intervención Reactiva (IT_R)** cuando contestan o evalúan la intervención anterior del otro hablante.

- c) **Intervención Mixta ($IT_{R/I}$)**, que son intervenciones reactivas que finalizan introduciendo una intervención iniciativa. Una intervención mixta puede tratarse en cualquier caso como dos intervenciones, una iniciativa y otra reactiva, como si fueran dos turnos diferentes del mismo hablante.
2. **Continuación o aportación (CT)** representa un turno vacío típico de un oyente cuyo propósito es reforzar su papel de oyente. El oyente utiliza las continuaciones para indicar al hablante que sigue el turno que este último está emitiendo en ese instante, especialmente cuando el turno se alarga. La continuación carece de información para el progreso del tópico de la conversación.
- **Par Adyacente o Intercambio (AP)**: es una secuencia de turnos encabezados por una intervención iniciativa (IT_I) y finalizados por la intervención reactiva a ésta (IT_R). Internamente, el par adyacente puede contener otros pares adyacentes, incluso algunos turnos de continuación que interrumpen la secuencia.
 - **Tópico del discurso (TOPIC)**: es la entidad principal del diálogo. En Rocha (1998) se proponen cuatro características que se pueden tener en cuenta para seleccionar el mejor candidato para tópico del discurso: su frecuencia, su distribución en el texto y su adecuación semántica. El tópico debe ser un elemento léxico al que se hace referencia frecuentemente.

De acuerdo a la estructura mencionada, se considera necesario incluir las siguientes etiquetas para una anotación apropiada de la estructura del diálogo: IT_I , IT_R , CT, AP y TOPIC. Las etiquetas AP y TOPIC se usarán en el sistema de resolución de anáforas en el diálogo para definir el espacio de accesibilidad anafórica, y el resto serán usadas para poder construir los pares adyacentes. La etiqueta $IT_{R/I}$ para las intervenciones mixtas no se incluye aquí puesto que las intervenciones mixtas se propone que sean divididas en dos intervenciones: la reactiva IT_R y la iniciativa IT_I . Esta tarea se realiza en la fase de anotación.

Un ejemplo del esquema de anotación del diálogo aplicado a un texto concreto se presenta en la figura 4.1. Las etiquetas (OP)

4.1 Propuesta de un esquema de anotación para la estructura del diálogo 113

y (US) proporcionadas por el transcriptor de los diálogos indican los turnos de cada hablante.

TOPIC		tren que salga mañana por la tarde para Monzón
AP1	<i>IT_I</i> (OP)	información de Renfe, buenos días
	<i>IT_R</i> (US)	hola, buenos días
	CT (OP)	hola
AP2	<i>IT_I</i> (US)	me podéis decir algún tren que salga mañana por la tarde para Monzón
	<i>IT_R</i> (OP)	sí, vamos, mira hay un talgo a las tres y media de la tarde
AP3	<i>IT_I</i> (US)	sí, tiene que ser más tarde
	<i>IT_R</i> (OP)	más tarde. Hay un intercity a las cinco y media, un expreso a las seis y media
AP4	<i>IT_I</i> (US)	el de las seis y media ¿llega a Monzón?
AP5	<i>IT_I</i> (OP)	a ver. El de las seis y media me ha preguntado ¿verdad?
	<i>IT_R</i> (US)	sí
	<i>IT_R</i> (OP)	a las nueve y veinticinco
AP6	<i>IT_I</i> (US)	a las nueve y veinticinco está en Monzón
	<i>IT_R</i> (OP)	sí
	CT (US)	vale, pues ya está. Esto ya es suficiente.
AP7	<i>IT_I</i> (US)	gracias, ¿eh?
	<i>IT_R</i> (OP)	muy bien a usted. Hasta luego

Figura 4.1. Ejemplo de diálogo anotado según nuestra propuesta

Destacamos especialmente la compatibilidad de este esquema con la información proporcionada por los sistemas de gestión del diálogo. Así, la definición de los pares adyacentes se corresponde con la tarea básica de detección de preguntas y respuestas en un sistema de diálogo. Por otra parte, la definición del tópico corresponde a la tarea básica del reconocimiento de las intenciones del usuario por parte del gestor del diálogo. De hecho, el tópico tal y como lo consideramos en este esquema es una entidad del discurso, y por tanto, los sistemas de diálogo aplicados a un dominio determinado pueden reconocerla mediante el empleo de marcadores del discurso.

Una de las principales ventajas de este esquema de anotación es por tanto su compatibilidad con cualquier sistema de diálogo aplicado a cualquier dominio. Ello nos permitirá su inclusión en cualquiera de estos sistemas sin necesidad de modificaciones.

4.2 Definición del espacio de accesibilidad anafórica estructural (EAAE)

Los estudios presentados en la sección 3.3 muestran que la estructura de pares adyacentes, así como la definición del tópico del diálogo generan por sí mismos una gran parte de las anáforas provocadas en los diálogos. Esto significa que el oyente que recibe una de estas anáforas será capaz de identificar dichas estructuras con el fin de definir un espacio donde buscar los antecedentes que le permiten la comprensión de las anáforas generadas. Este espacio definido de esta forma es lo que hemos llamado el **espacio de accesibilidad anafórica estructural, EAAE** (Martínez-Barco & Palomar, 2000a).

De esta forma, y con el esquema de anotación definido anteriormente como base, se propone un *espacio de accesibilidad anafórica* útil para la resolución de las anáforas creadas por los pronombres, tanto personales como demostrativos, y las anáforas adjetivas. Consideraremos por tanto que los antecedentes de dichas anáforas pueden encontrarse principalmente en el conjunto de sintagmas nominales que se encuentran en alguno de los espacios que se detallan a continuación:

- El mismo par adyacente donde se encuentra la anáfora, como ocurre en el ejemplo 79.

(79) **AP₁ IT_I A:** ¿Me podría decir a qué hora sale
el último tren a Burgos_i?
IT_R B: Sí, **lo_i** tiene a las veintitrés treinta.

- El par adyacente anterior al de la anáfora en una secuencia de pares adyacentes consecutivos, como ocurre en la anáfora adjetiva del ejemplo 80.

4.2 Definición del espacio de accesibilidad anafórica estructural (EAAE) 115

(80) **AP₁ IT_I A:** ¿Tiene información sobre el talgo
que va a Barcelona_i?

IT_R B: Sí, tenemos horarios y precios.

AP₂ IT_I A: ¿Cuál es **el próximo_i** que sale?

IT_R B: Hay uno a las doce y media.

- En el par adyacente que contiene a su vez al par adyacente de la anáfora (en el ejemplo 81, AP₂ está contenido en AP₁ puesto que AP₂ desarrolla para aclarar algo de AP₁).

(81) **AP₁ IT_I A:** ¿Tiene información sobre trenes
a Barcelona_i?

AP₂ IT_I B: Sí, ¿qué desea saber de **ellos_i**?

IT_R A: Quiero horarios.

IT_R B: Sí, dígame entonces.

- En el sintagma nominal que representa al tópico principal del diálogo, como ocurre en el ejemplo 82 donde el tópico es el único elemento que sin aparecer recientemente, aún está en la consciencia de los hablantes.

(82) **TOPIC** el talgo que va a Barcelona_i

AP₁ IT_I A: ¿Tiene información sobre el talgo
que va a Barcelona_i?

IT_R B: Sí, tenemos horarios y precios.

AP₂ IT_I A: Estupendo ya creía que me habían
dado un número equivocado. Llevo un rato
llamando y no me cogían el teléfono.

IT_R B: Sí es que llevamos una mañana muy
ocupada.

AP₃ IT_I A: Bueno, no importa, ¿cuál es
el próximo_i que sale?

IT_R B: Hay uno a las doce y media.

Para la evaluación del espacio que se propone se ha llevado a cabo un estudio sobre la localización de los antecedentes tal y como se muestra en la siguiente sección.

4.3 Estudio sobre el espacio de accesibilidad anafórica estructural

Con el fin de conocer la validez de la definición anterior para el espacio de accesibilidad anafórica estructural se ha desarrollado un estudio empírico que nos permite conocer sobre un determinado corpus de diálogos la cantidad de anáforas que han sido generadas dentro de esta estructura. Este estudio (Martínez-Barco & Palomar, 2000c) se muestra en detalle a continuación.

4.3.1 Estudio empírico

Para la realización de este estudio se han tomado al azar 40 diálogos de los 204 pertenecientes al corpus Infotren: Persona proporcionados por el proyecto BASURDE (Proyecto BASURDE, 1998-2001). Estos diálogos muestran conversaciones telefónicas entre operadores de una compañía de ferrocarriles y usuarios de esta compañía que demandaban información sobre algunos de los servicios ofrecidos por ésta. El corpus Infotren: Persona fue transcrito manualmente por el proyecto BASURDE incorporando información ortográfica y marcas de los turnos y el hablante de cada turno.

A partir del corpus transcrito, tal como muestra la figura 4.2, se ha realizado una anotación manual del texto en dos fases: en una primera fase se han anotado los pares adyacentes así como el tópico de cada diálogo. En una segunda fase, se han anotado las anáforas pronominales y adjetivas encontradas, así como la solución adecuada para cada una de ellas. Para garantizar la validez de los resultados obtenidos, la anotación se realizó en paralelo por dos anotadores a los que se les realizó un estudio de fiabilidad por medio del coeficiente *kappa* tal y como se expone en la sección 2.4.2. El proceso de anotación de estos 40 diálogos se realizó junto

con el resto de los 204 diálogos que conforman el corpus Infotren: Persona. En el capítulo 6 se muestran más detalles acerca de esta anotación y su estudio de fiabilidad.

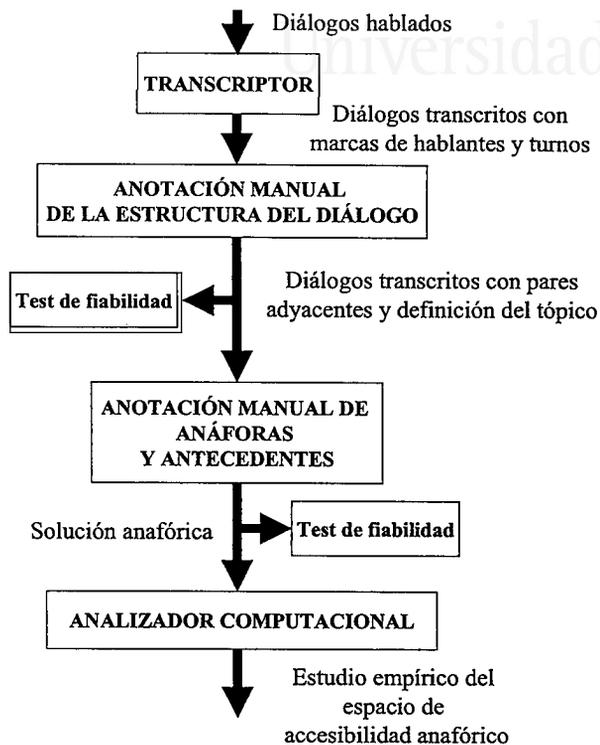


Figura 4.2. Estudio empírico sobre el espacio de accesibilidad anafórica

Tras los procesos de anotación manual, un analizador computacional implementado en LPA Prolog estudia la estructura del diálogo, las anáforas marcadas (pronominales y adjetivas) y las soluciones propuestas manualmente. Como resultado se obtiene el informe estadístico proporcionado en la tabla 4.1.

En dicha tabla, *Mismo AP* muestra un porcentaje de veces en las que el antecedente se encontró en el mismo par adyacente donde se encontraba la anáfora. En este caso se encontró que el antecedente de la anáfora pronominal estaba en este espacio un 60,6% de las veces y un 44,7% para la anáfora adjetiva. Este re-

	Mismo AP	Previo AP	AP anidado	Tópico	Otros
Pronominal	60,6%	24,6%	8,2%	4,9%	1,7%
Adjetiva	44,7%	28,9%	5,2%	13,4%	7,8%
Resultados totales	EAAE: 95,9%				4,1%

Tabla 4.1. Estudio empírico sobre el espacio de accesibilidad anafórico

sultado muestra cómo el pronombre hace uso de espacios un poco más cortos que la anáfora adjetiva debido principalmente a que contiene menor información para la localización del antecedente.

Previo AP muestra los casos en los que el antecedente se encontró en el par adyacente inmediatamente anterior al par donde se encuentra la anáfora. En este caso, ambos resultados son similares, 24,6% para la pronominal y 28,9% para la adjetiva.

AP anidado muestra casos en los que el antecedente se encontró en un par adyacente que contenía al par adyacente de la anáfora. Hay que destacar que, aunque el resultado obtenido en este espacio es relativamente bajo (8,2% y 5,2% para pronominal y adjetiva respectivamente), esta estructura anidada tampoco es demasiado habitual. Por ello, si ocurre que dicha estructura aparece, entonces existe una alta probabilidad de encontrar allí el antecedente.

Tópico muestra los casos en los que el antecedente se encontró en el tópico del diálogo. En este caso la anáfora adjetiva (13,8%) es más frecuente en este espacio que la anáfora pronominal (4,9%). Este hecho demuestra que los hablantes tienen cierta preferencia por escoger la anáfora adjetiva cuando quieren referirse al tópico.

Finalmente, *Otros* muestra aquellos casos en los que el antecedente no se encontró en ninguno de los anteriores (1,7% y 7,8% para pronominal y adjetiva respectivamente).

Destacaremos en este punto que sólo se muestra información de la ocurrencia más cercana a la anáfora, esto quiere decir que si un antecedente ha sido localizado en el mismo AP y además en el Tópico, sólo se muestra información sobre el primero siguiendo un criterio de búsqueda más cercana.

Según este estudio, el 95,9% de los antecedentes encontrados se localizaron en el espacio de accesibilidad anafórica estructural

propuesto. El resto de antecedentes no localizados (4,1%) se estimó que podrían localizarse en los subtópicos del tópico principal del diálogo, no definidos para este estudio. Esto demuestra que la gran mayoría de los antecedentes para la resolución de la anáfora podrían encontrarse en la estructura formada por los pares adyacentes, el tópico del diálogo y los subtópicos de éste.

4.3.2 Otras alternativas sobre el espacio de accesibilidad anafórica

Se han estudiado y comparado otras alternativas al espacio de accesibilidad anafórica estructural experimentado en la sección anterior. Sin embargo ninguno de ellos ha sido considerado eficiente para el tratamiento de los diálogos. A continuación se muestran las conclusiones alcanzadas.

El espacio completo (EAAC). La estrategia básica podría ser la consideración del *espacio completo* (EAAC), es decir, tomar todos los sintagmas nominales desde el principio del diálogo hasta la anáfora. Este es el caso de la propuesta de algunos métodos clásicos de resolución de anáfora como el de Carbonell y Brown (1988).

Mediante el espacio completo se llega forzosamente a recoger el 100% de los antecedentes, en lugar del 95,9% recogido por el EAAC en nuestro experimento. Sin embargo, tal y como se muestra en la tabla 4.2, el EAAC genera una media de 10,74 candidatos por anáfora (antes de aplicar ningún tipo de restricción) en lugar de los 34,14 candidatos por anáfora que se ha calculado en el experimento anterior que debería haberse procesado considerando el espacio completo. Esto significa que el uso del EAAC produciría un incremento de candidatos hasta el 318% respecto al EAAC. Así, el uso del espacio completo como estrategia para la definición del espacio de accesibilidad anafórica implicaría a) un mayor coste computacional y b) triplicar las posibilidades de obtener una respuesta incorrecta en el algoritmo de resolución de la anáfora que lo use. Esta diferencia se puede incrementar considerablemente en diálogos largos.

Candidatos considerados	EAAE	EAAC	EAAVO
Total	1063	3380	1292
Por anáfora	10,74	34,14	13,05
Proporción	100%	318%	122%

Tabla 4.2. Comparativa entre espacio de accesibilidad anáforico estructural (EAAE), espacio basado en ventana de oraciones (EAAVO) y espacio completo (EAAC)

Ventanas de oraciones de tamaño fijo (EAAVO). Otra alternativa podría ser la propuesta por métodos de resolución de anáfora en discurso monólogo que definen una ventana de accesibilidad fija según un número de oraciones. Por ejemplo en el trabajo de Ferrández *et al.* (1998b) se propone el uso de una ventana de 3 oraciones anteriores más la actual para el caso de la anáfora pronominal y de 4 oraciones anteriores más la actual para el caso de la anáfora adjetiva. Por otra parte, para el inglés Kameyama (1997) propone la misma ventana para la anáfora pronominal y una ventana de 10 oraciones para el caso de las descripciones definidas. No disponemos de propuestas específicas para el caso de la resolución de la anáfora en los diálogos en español por lo que no se puede comparar con ninguna de características similares a la nuestra.

Sin embargo, esta propuesta no tiene un fundamento estructural. Simplemente, los autores han realizado estudios empíricos que les han proporcionado estos valores. Por tanto, la clase de texto que se procese puede hacer variar considerablemente estas longitudes.

En nuestro estudio empírico con los 40 diálogos del corpus Infotren: Persona hemos realizado un estudio sobre la cantidad de oraciones que deberían haberse considerado para conseguir una lista de candidatos completa. Sin embargo, debido a la dificultad existente para considerar oraciones en los diálogos, se ha considerado su equivalente pragmático: el enunciado. Los resultados se muestran en la tabla 4.3.

De acuerdo a estos estudios, siguiendo un espacio EAAVO, para abarcar el mismo número de antecedentes que proporciona la estructura EAAE, deberían ser considerados 11 enunciados en

Ventana de Enunciados: "Desde el enunciado de la anáfora a:"	% antecedentes de anáforas pronominales	% antecedentes de anáforas adjetivas
Enunciado de la anáfora	37.7	18.4
-1	54.1	44.7
-2	70.5	52.6
-3	77.0	55.3
-4	80.3	57.9
-5	83.6	71.0
-6	88.5	73.7
-7	91.8	76.3
-8	91.8	81.6
-9	95.1	81.6
-10	96.7	92.1
-11	98.4	94.7
-12	98.4	97.4
-13	98.4	97.4
-14	100.0	100.0

Tabla 4.3. Estudio del espacio de accesibilidad anáforico basado en ventana de oraciones (EAAVO) adaptada a enunciados

el caso de la anáfora pronominal y 10 enunciados en el caso de la adjetiva. Dichos resultados incrementan bastante la ventana de oraciones definida por Ferrández *et al.* y Kameyama para el mismo tipo de anáfora. Esto es debido a que los enunciados del diálogo son mucho más cortos y por tanto más frecuentes que las oraciones de los monólogos.

De esta forma, tal y como muestra la tabla 4.2, el número de candidatos a considerar por anáfora pasa a ser 13,05, es decir el 122% del considerado en EAAE. Si a esto le añadimos la desventaja evidente de que no hay un criterio estructural más o menos estable que defina el tamaño óptimo de estas ventanas, sino que depende únicamente del tipo de textos, consideramos que la definición del espacio anáforico estructural EAAE es la mejor solución para el caso de los diálogos.

4.4 Estandarización del esquema de anotación

La necesidad de usar un esquema de anotación para la ejecución de cualquier proyecto de lenguaje natural conlleva implícito la utilización de una serie de herramientas de etiquetado que deben

estar al alcance de cualquier usuario del proyecto. En este sentido, una estandarización del esquema de anotación que permita el etiquetado de textos al margen de cualquier implementación es una tarea fundamental para permitir su conexión con otras herramientas. A continuación se presenta el lenguaje de estándares SGML como una solución a dicho requerimiento.

4.4.1 Anotación de textos mediante el lenguaje de estándares SGML

SGML (*Standard Generalized Markup Language*) es un lenguaje de generación de etiquetados estandarizado por la norma ISO 8879:1986 (Goldfarb, 1990). De acuerdo a este lenguaje, mediante la definición del rol de cada componente del texto en un modelo formal, los usuarios de programas basados en SGML serán capaces de comprobar que cada elemento del texto se ha etiquetado de forma correcta. Así, el uso de SGML permitirá por ejemplo comprobar que los anotadores de un texto no han introducido un tercer nivel de etiquetado sin haber introducido el segundo nivel, etc. De esta forma, una vez que se define el modelo formal del etiquetado para un tipo de documento en particular, los programas de análisis de SGML (conocidos normalmente como *SGML parsers*) podrán comprobar si un texto etiquetado según ese tipo de documento cumple o no las reglas establecidas por el modelo. Además, otro tipo de programas llamados editores de SGML podrán generar gran parte del etiquetado de forma automática y permitir al anotador ahorrar gran parte del trabajo de etiquetado.

SGML se basa en el concepto de un documento compuesto por varios objetos. Cada objeto contiene uno o más elementos lógicos con atributos que describen la forma en la que van a ser procesados. De esta forma, SGML proporciona una manera de describir las relaciones entre entidades, elementos lógicos y atributos para que sean reconocidas por el programa informático que lo procesa.

Una de las principales ventajas del uso de SGML para etiquetado de textos es que SGML permite que los usuarios definan su propio modelo de documento a través de un fichero llamado DTD (*Document Type Definition*). Este DTD, construido según

una sintaxis concreta, permite que desde el programa informático pueda ser interpretado el tipo de documento que se introducirá y cómo se van a encontrar sus diferentes partes relacionadas entre sí. El documento etiquetado en SGML y el DTD que lo interpreta, estandarizan el etiquetado permitiendo que cualquier herramienta posterior analice fácilmente el documento sin crear dependencias entre el etiquetador y la herramienta.

En nuestro esquema de anotación usaremos el lenguaje SGML como herramienta de etiquetado aprovechando la flexibilidad que nos brinda.

4.4.2 Anotación de la estructura del diálogo con SGML

Puesto que la anotación será procesada por algún sistema de resolución de anáfora automático, proponemos un formato de etiquetado basado en SGML que será independiente de la implementación final del sistema (Martínez-Barco & Palomar, 2000c). La utilización de SGML tiene algunas ventajas añadidas entre las que se encuentra que permite que la anotación manual se efectúe desde cualquier editor de SGML definiendo el fichero DTD correspondiente. Por otra parte, cualquier sistema automático de definición de tópicos y construcción de pares adyacentes podría hacer uso de este mismo DTD.

El etiquetado SGML tendrá la siguiente forma:

```
<NOMBRE-ELEMENTO NOMBRE-ATRIBUTO="VALOR" ...>
  texto
</NOMBRE-ELEMENTO>
```

De esta forma se proporcionará la siguiente notación para cada caso particular:

- Tópico. Del tópico nos interesa conocer la entidad que lo forma y su ámbito. Por este motivo se usa la etiqueta TOPIC con el parámetro ENTITY que muestra la entidad que forma el tópico, y la etiqueta /TOPIC que marca el fin del ámbito del tópico. En los diálogos con un único tópico, este ámbito alcanzará todo el

diálogo, sin embargo, para diálogos largos, el tópico podría cambiar introduciendo una etiqueta de fin del tópico y otra etiqueta de tópico nueva.

```
<TOPIC ENTITY="Entidad-tópico">
  Diálogo ámbito del tópico
</TOPIC>
```

- Pares adyacentes. En cuanto la información necesaria del par adyacente, una etiqueta AP indica el comienzo del par junto con el parámetro ID que contiene un número de identificador usado para identificar de forma única los diferentes pares adyacentes. La etiqueta /AP se usa para indicar el fin del ámbito del par adyacente

```
<AP ID="número">
  Texto-Par-Adyacente
</AP>
```

- Intervenciones. De las intervenciones es necesario conocer su ámbito, que quedará definido entre las etiquetas IT y /IT. Además el parámetro TYPE indicará el tipo de intervención adoptando los valores posibles R o I (Reactivo o Iniciativo), y SPEAKER que definirá la identidad del hablante de ese turno según el transcriptor.

```
<IT TYPE="R|I" SPEAKER="hablante">
  Texto-Intervención
</IT>
```

- Continuaciones. Por último, los turnos de continuación (o continuaciones) se definirán a través de las etiquetas CT y /CT que marcarán el comienzo y el fin del ámbito de la continuación, y el parámetro SPEAKER incluido en la primera etiqueta para marcar la identidad del hablante.

```
<CT SPEAKER="hablante">
  Texto-Continuación
</CT>
```

En la figura 4.3 se muestra el etiquetado mediante SGML del diálogo mostrado en la figura 4.1

```

<TOPIC ENTITY="tren que salga mañana por la tarde para Monzón">

<AP ID="1"> <IT TYPE="I" SPEAKER="OP"> información de Renfe,
buenos días </IT> <IT TYPE="R" SPEAKER="US"> hola, buenos días
</IT> <CT SPEAKER="OP"> hola </CT> </AP>
<AP ID="2"> <IT TYPE="I" SPEAKER="US"> me podéis decir algún
tren que salga mañana por la tarde para Monzón </IT> <IT TYPE="R"
SPEAKER="OP"> sí, vamos, mira hay un talgo a las tres y media de
la tarde </IT> </AP>
<AP ID="3"> <IT TYPE="I" SPEAKER="US"> sí, tiene que ser más
tarde </IT> <IT TYPE="R" SPEAKER="OP"> más tarde. Hay un
intercity a las cinco y media, un expreso a las seis y media </IT>
</AP>
<AP ID="4"> <IT TYPE="I" SPEAKER="US"> el de las seis y media
>llega a Monzón? </IT>
<AP ID="5"> <IT TYPE="I" SPEAKER="OP"> a ver. El de las seis
y media me ha preguntado >verdad? </IT> <IT TYPE="R" SPEAKER="US">
sí </IT> </AP>
<IT TYPE="R" SPEAKER="OP"> a las nueve y veinticinco </IT> </AP>
<AP ID="6"> <IT TYPE="I" SPEAKER="US"> a las nueve y veinticinco
está en Monzón </IT> <IT TYPE="R" SPEAKER="OP"> sí </IT> <CT
SPEAKER="US"> vale, pues ya está. Esto ya es suficiente. </CT>
</AP>
<AP ID="7"> <IT TYPE="I" SPEAKER="US"> gracias, >eh? </IT> <IT
TYPE="R" SPEAKER="OP"> muy bien a usted. Hasta luego </IT> </AP>
</TOPIC>

```

Figura 4.3. Ejemplo de anotación SGML de la estructura del diálogo

4.5 Conclusiones del capítulo

En la sección 3.3 se mostró que la estructura del diálogo en la consciencia del oyente, construida mediante pares adyacentes y sus relaciones, y mediante la definición de los tópicos del discurso, influye decisivamente en la toma de decisiones para resolver la anáfora. Desde estas conclusiones, en este capítulo se ha demostrado que la definición de un espacio de accesibilidad anafórica basado en la estructura del diálogo es una tarea fundamental en un sistema de resolución de anáfora específico para diálogos.

Este espacio, desde el cual (dada una anáfora) se extrae la lista de candidatos a antecedente, ha sido evaluado y comparado con otros métodos para la obtención de dicha lista encontrándose que éste se ajusta más a las intenciones del hablante y a las interpretaciones que de ellas hace el oyente.

Nuestra propuesta de anotación permitirá obtener la información estructural necesaria y suficiente para la generación del espacio de accesibilidad anafórica (espacio de la solución).

Sin embargo, puesto que el sistema que proponemos deberá integrarse en una aplicación basada en diálogos, deberemos plantear un esquema que sea compatible con los esquemas que generalmente usan estas aplicaciones. Por otra parte, puesto que la estructura ya vendrá generada desde la aplicación, para evaluar nuestro sistema independientemente de su aplicación final, deberemos simular la entrada de dicha estructura usando una anotación manual.

Además, se ha proporcionado un mecanismo para estandarizar la anotación del diálogo y facilitar así la construcción de herramientas que permitan la anotación tanto manual como automática de dichos diálogos. Cabe destacar en este sentido, que no es objetivo de este trabajo la construcción de generadores automáticos de esta estructura, ya que son los propios sistemas de diálogo los que interpretan y generan dicha estructura en dominios restringidos sin intervención del usuario.



5. ARIADNA: Sistema de resolución de la anáfora en diálogos

Universitat d'Alacant
Universidad de Alicante

Tal y como expresan Carbonell y Brown (Carbonell & Brown, 1988) la anáfora es un fenómeno lingüístico en el que intervienen diversas fuentes de información. Esto genera un sistema de reglas que deben ser aplicadas, unas para descartar candidatos no válidos y otras para ordenar los candidatos y seleccionar como solución aquel que, atendiendo a las propiedades que cumple, es el mejor antecedente.

En este capítulo presentamos el sistema completo para la resolución de la anáfora en diálogos ARIADNA (*Anaphora Resolution In Automatic Dialogue systems with Natural language Analysis*), que se plasma mediante el algoritmo ARDi (*Anaphora Resolution in Dialogues*), tomando como base un sistema de restricciones y preferencias que combina dos clases de fuentes de información: el conocimiento lingüístico de tipo léxico, morfológico y sintáctico, así como el conocimiento de la estructura del diálogo basado en la anotación manual de pares adyacentes y del tópico del diálogo. Las bases de este sistema han sido presentadas en diversas publicaciones y congresos (Martínez-Barco, 2000) (Palomar & Martínez-Barco, 2000) (Martínez-Barco & Palomar, 2000d) (Palomar & Martínez-Barco, 2001).

El sistema ARIADNA está actualmente definido para tratar las anáforas generadas por los pronombres personales de tercera persona (incluyendo los reflexivos y recíprocos), los pronombres demostrativos y las anáforas adjetivas, al ser considerados éstos los tipos de anáfora más comunes encontrados en los diálogos. Sin embargo, el sistema queda abierto al tratamiento de otros tipos de anáfora tal y como se muestra en este capítulo.

5.1 Orígenes de ARIADNA

Los orígenes del sistema ARIADNA toman como punto de partida las conclusiones extraídas por Ferrández *et al.* (1998b) sobre sistemas de restricciones y preferencias en la resolución de la anáfora. El sistema de restricciones y preferencias original de Ferrández (1998) estaba basado en el uso de información léxica, morfológica y sintáctica, y se aplicaba a la resolución de la anáfora pronominal en discursos monólogos, tal y como se puede apreciar en la figura 5.1

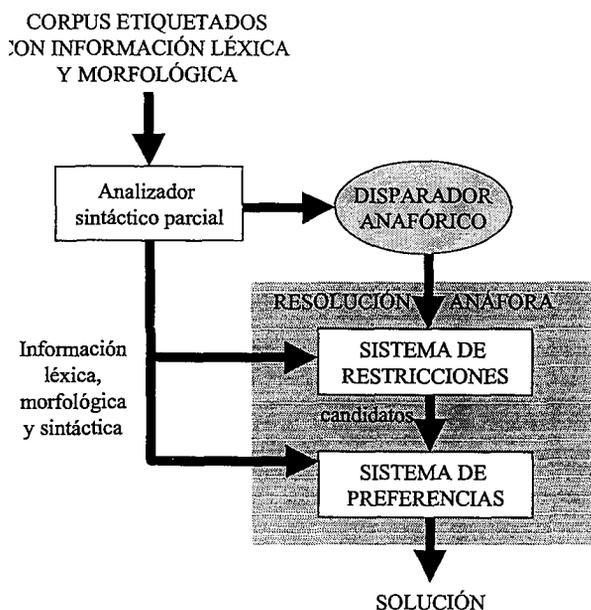


Figura 5.1. Sistema de restricciones y preferencias de Ferrández *et al.* (1998)

Por otro lado, según los estudios presentados sobre segmentación y estructuración del diálogo, se llega a la conclusión de que existe la necesidad de definir una adecuada estructura del diálogo que proporcione información para la resolución de la anáfora en diálogos. Es por este motivo que nuestra propuesta inicial para la resolución de la anáfora en diálogos se basaba en la obtención de una estructura del diálogo a través de la detección automática de

los tópicos, turnos de intervención y turnos de continuación. Esta propuesta inicial, embrión de la que se utiliza en ARIADNA, se muestra a continuación.

5.1.1 Propuesta inicial de resolución de anáfora basada en la detección automática de la estructura del diálogo

En Martínez (1999) y Martínez *et al.* (1999) presentamos una propuesta de un sistema de resolución de anáfora basado en una detección automática de la estructura del diálogo. En esta detección automática se propone un tópico para cada fragmento de diálogo y se realiza una clasificación de los turnos entre intervenciones y continuaciones. La estructura así definida se aplica a la resolución de la anáfora pronominal en diálogos. El sistema cuyo algoritmo principal se muestra en el pseudocódigo de la figura 5.2, permite obtener la información del tópico y de las intervenciones necesaria para proporcionar los antecedentes a un sistema de restricciones y preferencias que escogerá finalmente el antecedente apropiado.

El sistema de detección de tópicos define tres espacios para la accesibilidad anafórica:

1. El espacio de la intervención actual.
2. El espacio de la intervención anterior.
3. El tópico relevante en el segmento de diálogo en el que se encuentra la anáfora.

Este sistema se aplica sobre un corpus¹ de diálogos grabados en cinta y posteriormente transcritos incluyendo el etiquetado de los turnos de habla y el hablante de cada uno de ellos. Además el corpus se etiquetó incluyendo la información léxica y morfológica de cada palabra² y posteriormente fue procesado por un analizador sintáctico parcial del que se obtienen los sintagmas nominales y las estructuras anafóricas.

Con este etiquetado, los espacios primero y segundo se determinan directamente gracias al etiquetado de turnos que contiene

¹ Corpus oral proporcionado por el Laboratorio de Lingüística Computacional de la Universidad Autónoma de Madrid

² Etiquetado por el etiquetador léxico-morfológico XEROX

130 5. ARIADNA: Sistema de resolución de la anáfora en diálogos

```

Principio
Fa :=[]; Fn :=[]; Fa' :=[]; F :=[];
Turn :=1;
Mientras hay turnos
  Oración :=1;
  Mientras hay oraciones
    Fn :=relevante(np(Oración));
    Fa := Fa+Fn;
    Fn :=[];
    Si Buscar_anáfora
      Buscar_antecedente(Fa);
      Buscar_antecedente(Fa');
      Buscar_antecedente(F) por peso;
    FinSi
    Si Fa<>[]
      Clase(Turno):=INTERVENCION;
      Si Clase(Turno-1)=INTERVENCION o
        Participante(Anterior_intervención)
        <>Participante(Turno)
        Fa' :=[];
      FinSi
    FinSi
    Oración := Oración+1;
  FinMientras
  Si Fa<>[] /* intervención */
    F :=incorporar(Fa,F);
    Fa' := Fa+Fa';
    Fa :=[];
    Anterior_intervención :=turno;
  Sino
    Clase(turno) :=CONTINUACION;
  FinSi
  Turno :=Turno+1;
FinMientras
Fin

```

Figura 5.2. Algoritmo de resolución de la anáfora con segmentación automática basada en el tópico

el propio diálogo transcrito. El propio sistema hace una clasificación automática de los turnos en turnos de intervención o turnos de continuación (de acuerdo a la definición realizada por Gallardo y que se presentó en la sección 2.4.1) dependiendo de la detección o no de entidades (sintagmas nominales) en el turno.

El tercer espacio se identifica como se muestra a continuación. Se definen dos coeficientes: el primero (*C_{iu}*, coeficiente de incremento por uso) indica la importancia que adquiere una entidad

que aparece en la intervención actual; el segundo (Cdd , coeficiente de decremento por desuso) indica la pérdida de importancia de una entidad que aparece en intervenciones previas pero no en la actual. Ambos coeficientes tomarán los siguientes valores definidos de forma empírica:

- Ciu (Coeficiente de incremento por uso) = 10 unidades.
- Cdd (Coeficiente de decremento por desuso) = 1 unidad.

Además, el algoritmo usa tres listas diferentes para guardar las entidades que pueden ser un antecedente (de acuerdo con los espacios anteriormente definidos):

- La lista de entidades locales actuales $-Fa-$ contiene una lista de entidades que aparecen en el ámbito de la intervención actual.
- La lista de entidades locales anteriores $-Fa'-$ contiene una lista de entidades aparecidas en el ámbito de la última intervención.
- Y finalmente, la lista de tópicos generales $-F-$ contiene una lista de entidades relevantes en el diálogo que será evaluada de acuerdo con criterios de frecuencia y distribución.

El algoritmo analiza cada sentencia de cada turno de la siguiente forma:

1. Se produce una detección de las entidades con relevancia temática en el turno y se guardan en la lista Fa mediante el mecanismo *relevante*. Este mecanismo selecciona básicamente todos los sintagmas nominales encontrados excepto aquellos que concuerden con una lista de sintagmas nominales recogidos como no relevantes en un pequeño diccionario, como por ejemplo, “buenos días”, etc.
2. Cuando el turno cambia, las entidades almacenadas en la lista Fa se mueven a la lista Fa' puesto que ya forman parte del turno anterior, y además se incorporan a la lista de tópicos generales F mediante el mecanismo *incorporar*. Este mecanismo se comporta de la siguiente forma: por una parte, añade todas las entidades de Fa a F incrementando su peso con el coeficiente Ciu ; por otra parte el valor de Cdd se usa para decrementar el peso de aquellas entidades que están en F y no

en Fa , es decir, aquellas que habían aparecido anteriormente pero no lo han hecho en el último turno. Finalmente, Fa queda vacía para almacenar las nuevas entidades que se detecten.

3. Cuando un turno finaliza podremos determinar si se ha encontrado algún elemento temático (una entidad). Si esto ocurre, el turno de identificará como una intervención. Sin embargo, si no se encuentra ninguno, entonces el turno se identifica como un turno continuación. Si se tratara de una intervención, la primera oración que introduzca nuevos elementos temáticos provocará que la lista Fa' con las entidades relevantes de la intervención anterior se vacíe totalmente. Esto es debido a que consideramos que las entidades que aparecen en la intervención anterior sólo tienen importancia local mientras la intervención actual aún no ha introducido ninguna entidad. A partir de este instante ya sólo serán consideradas como tópicos generales de acuerdo con el valor de peso que tienen en la lista F . Sin embargo, si se trata de un turno continuación, Fa' no se vacía puesto que sus entidades siguen teniendo importancia local (ya que la intervención anterior sigue vigente).
4. Conforme se analiza cada oración, el mecanismo *buscar-anáfora* se pone en marcha intentando detectar la existencia de una anáfora pronominal. Si el resultado es afirmativo, se pone en marcha el mecanismo *buscar-antecedente* que usa el sistema de restricciones y preferencias. Este mecanismo comenzará con una búsqueda del antecedente en la lista Fa . Si no se encuentra ninguna entidad, el mecanismo lo buscará en la lista de entidades locales anteriores Fa' , y como último recurso lo buscará en F .

Funcionamiento del sistema. El ejemplo 83 es un fragmento de un diálogo en español donde se puede comprobar el funcionamiento del algoritmo mediante una traza (entre paréntesis se marca el número de turno).

- (83) (T01) H1: Buenos días.
 (T02) H2: Buenos días. ¿Qué deseas?
 (T03) H1: Quiero manzanas.

- (T04) H2: ¿De qué clase las quieres?
- (T05) H1: No importa si son buenas.
- (T06) H2: Éstas las he recibido esta mañana.
Son muy buenas.
- (T07) H1: Entonces dame de esas.
- (T08) H2: ¿Cuántas quieres?
- (T09) H1: Media docena.
- (T10) H2: Muy bien, ¿qué más quieres?
- (T11) H1: ¿Tienes limones?
- (T12) H2: Sí, los tengo en una caja por aquí...
Aquí están. ¿Los quieres muy verdes?
- (T13) H1: No... Son para hacer limonada. A
mis hijos les encanta.
- (T14) H2: ¿Cuánto quieres?
- (T15) H1: Un kilo.
- (T16) H2: Éstos te los dejo a buen precio. Muy
bien, ¿algo más?
- (T17) H1: Nada más, gracias. ¿Cuánto es to-
do?
- (T18) H2: Trescientas cincuenta.
- (T19) H1: Ahí tienes. Hasta luego.
- (T20) H2: Hasta luego.

El algoritmo comienza considerando los turnos T01 y T02 como continuaciones debido a que no contienen ninguna entidad. De esta forma, las tres listas permanecen vacías hasta el turno T03. Al procesar T03, *manzanas* es considerado como una entidad, y por lo tanto se introduce en la lista de entidades locales actuales Fa :

$Fa = [\text{manzanas}]$

T03 se considera por tanto una intervención. Cuando esta intervención finaliza, las entidades de Fa se incorporan a la lista de

tópicos generales F con su peso correspondiente a la vez que se guardan en la lista de entidades locales anteriores Fa' . Tras esta operación las listas quedan de la siguiente forma:

$Fa = []$
 $Fa' = [\text{manzanas}]$
 $F = [(\text{manzanas}, 10)]$

En T04, tras la incorporación de *clase* como entidad aparecida, se detecta la primera anáfora con la aparición del pronombre *las*. En este punto el estado de las listas es:

$Fa = [\text{clase}]$
 $Fa' = [\text{manzanas}]$
 $F = [(\text{manzanas}, 10)]$

Para resolver la anáfora, el algoritmo busca, en primer lugar, un antecedente en la lista de entidades locales actuales (Fa). En este caso se encuentra la entidad *clase*, sin embargo, tras la intervención del sistema de restricciones se rechaza puesto que no cumple las características morfológicas requeridas (se busca un antecedente en plural). De esta forma, al no contener más candidatos la lista Fa , el algoritmo pasará a buscar en la lista de entidades locales anteriores Fa' encontrando en este caso *manzanas*. Tras la aplicación de las restricciones, *manzanas* se propone como antecedente para el pronombre y se incluye como entidad en Fa :

$Fa = [\text{clase}, \text{manzanas}]$

Tras procesar la primera oración de T04, Fa' se vacía:

$Fa' = []$

Y al terminar T04 las listas quedan de la siguiente forma:

$Fa = []$
 $Fa' = [\text{clase}, \text{manzanas}]$
 $F = [(\text{manzanas}, 20), (\text{clase}, 10)]$

Puesto que T05 es un turno de continuación las listas no se modifican. Tras T06, el peso de *manzanas* se incrementa gracias a nuevas ocurrencias de esta instancia (introducidas por los pronombres), mientras que la entidad *clase* pierde peso por lo contrario:

$Fa = []$
 $Fa' = [\text{manzanas, esta mañana}]$
 $F = [(\text{manzanas}, 40), (\text{esta mañana}, 10), (\text{clase}, 9)]$

Destacamos que las anáforas producidas por los pronombres *éstos* y *los* en el turno T16 sólo pueden ser resueltas mediante las entidades de la lista de tópicos generales F puesto que las listas locales Fa y Fa' no contienen el antecedente correcto. El contenido de las listas en este momento es:

$Fa = []$
 $Fa' = [\text{un kilo}]$
 $F = [(\text{manzanas}, 45), (\text{limones}, 39), (\text{un kilo}, 10),$
 $(\text{caja}, 8), (\text{media docena}, 6), (\text{esta mañana}, 4),$
 $(\text{clase}, 3)]$

Así se cumple que, por un lado, la lista Fa está vacía, el sistema de restricciones rechaza el posible antecedente de la lista Fa' , y finalmente, tanto el sistema de restricciones como el de preferencias actúan sobre la lista F devolviendo la entidad *limones* como antecedente válido.

Realimentación entre el sistema de tópicos y la propia resolución de la anáfora. El algoritmo propuesto incluye un mecanismo de detección de la anáfora (*buscar-anáfora*) y otro para su resolución (*buscar-antecedente*). Este último usa la información del sistema de detección de tópicos para determinar el antecedente más relevante en la lista de tópicos generales. La información de los tópicos se une a la información que se obtiene de otras fuentes para obtener el mejor antecedente. De la misma forma, una vez que la anáfora queda resuelta, el antecedente ocupa el espacio que ha dejado la expresión anafórica. Así, este antecedente verá incrementado su peso en la lista de tópicos generales. Podemos

afirmar, por tanto, que la estructura del diálogo influye sobre la resolución de la anáfora de la misma forma que la resolución de la anáfora influye sobre la estructura del diálogo.

Importancia del sistema. El sistema mostrado en esta sección presenta como innovación principal la extracción automática tanto de la información lingüística como de la información de la estructura del diálogo a través de una técnica para la detección automática del tópico. De esta forma no es necesario hacer una anotación previa.

Sobre este sistema se evaluó la resolución de la anáfora pronominal utilizando un corpus formado por diez diálogos. Cinco de estos diálogos fueron transcripciones directas de grabaciones proporcionadas por el Laboratorio de Lingüística Computacional de la Universidad Autónoma de Madrid. Los otros cinco fueron diálogos ficticios creados por nosotros mismos. Como resultado se obtuvo un prometedor valor del 77% de precisión. Sin embargo, cuando el sistema se aplicó sobre otro corpus de diálogo en los que los turnos se encontraban más difusos (entrecortados y mezclados entre sí) la precisión descendía. Además el sistema de resolución anafórico era totalmente dependiente de la detección del tópico, de tal forma que si el sistema hacía una propuesta incorrecta del tópico, la precisión caía irremediablemente.

Este sistema inicial, aunque constituyó un primer embrión del sistema ARIADNA que se mostrará a continuación, presentaba algunos inconvenientes como el hecho de considerar el espacio de accesibilidad anafórica desde el ámbito de los turnos, en lugar de hacerlo desde los pares adyacentes tal y como propone Fox (Fox, 1987). Además, puesto que la estructura del diálogo es generada automáticamente por los sistemas de diálogo, no consideramos necesario profundizar más en la obtención de dicha estructura.

Estos inconvenientes fueron subsanados en ARIADNA mediante una definición adecuada del esquema de anotación de la estructura del diálogo.

5.1.2 Características del sistema ARIADNA

A partir de las conclusiones extraídas de la propuesta inicial presentada anteriormente se desarrolla el sistema ARIADNA (ver figura 5.3) que incluye como características relevantes las siguientes:



Figura 5.3. Visión global del sistema ARIADNA

- Un etiquetador léxico-morfológico que genera las categorías gramaticales de cada palabra junto con su información morfológica, y que proporciona al sistema independencia en la entrada de documentos.
- Un segmentador de enunciados capaz de obtener los enunciados del diálogo mediante la detección de marcas de comienzo y fin de enunciados.
- Un analizador sintáctico parcial que genera la información sintáctica únicamente de los constituyentes que son relevantes para el proceso de resolución de la anáfora.
- Un disparador anafórico que detecta los casos de anáfora a tratar y que pone en marcha el mecanismo correspondiente a cada uno de ellos.
- Un interfaz de SGML a PROLOG que permite el empleo de un lenguaje de etiquetado estándar para la estructura del diálogo. Este interfaz abre el sistema al empleo de cualquier etiquetador estándar para anotaciones manuales, incluso, al empleo de anotadores automáticos.
- Un generador del espacio de accesibilidad anafórica estructural que define el espacio basado en la estructura del diálogo donde se buscará el antecedente adecuado.
- Un sistema de restricciones basado en información lingüística obtenida como salida del analizador sintáctico parcial que sirve para descartar los candidatos incompatibles con la anáfora.
- Un sistema de preferencias estructurales que utiliza la estructura del diálogo como fuente de información que sirve para ordenar y seleccionar los mejores candidatos.
- Un sistema de preferencias lingüísticas alimentado con información lingüística que contribuye a la ordenación final de los candidatos y selección del mejor como antecedente.

5.2 Módulos de preproceso en ARIADNA

A continuación se describen los módulos usados por ARIADNA en el preproceso de los diálogos para la generación de la información que posteriormente usarán los módulos de resolución de anáfora.

5.2.1 Etiquetador léxico-morfológico

El etiquetador léxico-morfológico usado en ARIADNA analiza y desambigua las categorías léxicas de las palabras para textos no restringidos en español basándose en modelos estocásticos ECGI Extendidos (Pla, 1999; Pla & Molina, 1999; Pla *et al.*, 2000) desarrollado por el Grup de Processament del Llenguatge Natural de la Universitat Politècnica de València. El etiquetador, como todo método basado en probabilidades, necesita una fase de entrenamiento previa a la fase de etiquetado. Para la fase de entrenamiento sus autores partieron de un corpus etiquetado que fue revisado por un experto. En concreto, se empleó una porción del corpus LexEsp³ (unas 60.000 palabras) etiquetado léxicamente con categorías gramaticales. A partir de las secuencias de categorías, se aprenden dos modelos de lenguaje probabilísticos: un modelo de bigramas y un modelo gramatical extendido basado en un algoritmo ECGI. Ambos modelos forman un modelo de lenguaje combinado que se usa para describir las secuencias posibles de categorías léxicas. Este modelo ha sido suavizado para poder incluir relaciones que no han sido encontradas en el entrenamiento.

Por otra parte se construye un modelo léxico que se estima a partir del texto de entrenamiento etiquetado en el que se tienen en cuenta los pares formados por cada palabra y las categorías léxicas almacenadas. Así se estiman las frecuencias de las palabras, de las categorías y las frecuencias de cada palabra en cada categoría. Estas probabilidades léxicas también han sido suavizadas.

Una vez que se ha finalizado la fase de entrenamiento, el etiquetador léxico-morfológico ya se encuentra disponible para realizar etiquetados (fase de etiquetado). La fase de etiquetado o desambiguación léxica del texto usa el analizador morfológico MACO+ (Carmona *et al.*, 1998) desarrollado en el Grup de Processament de la Llengua de la Universitat Politècnica de Catalunya. MA-

³ LEXESP es un corpus en español que contiene 5 millones de palabras etiquetadas léxicamente perteneciente al proyecto del mismo nombre desarrollado por el Departamento de Psicología de la Universidad de Oviedo y desarrollado por el Grupo de Lingüística Computacional de la Universidad de Barcelona, con la colaboración del Grupo de Tratamiento del Lenguaje de la Universidad Politècnica de Cataluña.

CO+ segmenta el texto en tokens y proporciona todas las categorías gramaticales posibles para cada token junto con su información morfológica según el juego de etiquetas PAROLE (Martí *et al.*, 1998). La salida de MACO+ (a la que se le ha eliminado la información morfológica por simplicidad) y el modelo léxico entrenado constituyen las probabilidades léxicas. Dichas probabilidades léxicas junto con el modelo de lenguaje ECGI extendido constituyen la entrada para el etiquetador. Finalmente, para cada frase de entrada, se buscará la secuencia de estados de mayor probabilidad en el modelo ECGI extendido mediante el algoritmo de Viterbi. Esta secuencia óptima tendrá una única categoría léxica que es la etiqueta léxica a devolver como salida. Por último se añade la información morfológica que ha sido extraída en la entrada y se reescribe la oración añadiendo a cada palabra su lema y la etiqueta léxico-morfológica PAROLE ya desambiguada.

Entradas y salidas. El analizador léxico-morfológico toma como entrada un texto plano y obtiene como salida la relación de palabras con sus lemas y una etiqueta que indica la categoría gramatical de la palabra junto con su información morfológica en el formato: PALABRA₁ LEMA₁ ETIQUETA₁ PALABRA₂ LEMA₂ ETIQUETA₂... El conjunto de etiquetas léxico morfológicas contiene unas 230 etiquetas estructuradas en categoría y subcategoría gramatical, y contemplan aspectos morfológicos de género, número, persona y tiempos verbales.

5.2.2 Segmentador de enunciados

El segmentador de enunciados es un módulo de preproceso para el analizador sintáctico parcial SUPP (Martínez-Barco *et al.*, 1998). Puesto que SUPP está preparado para trabajar con corpus en los que las oraciones están segmentadas y numeradas, el segmentador de enunciados reconoce caracteres de fin de enunciado (., ?, !, ...) y etiquetas de inicio de enunciado (la etiqueta léxico-morfológica del hablante) y segmenta cada enunciado introduciéndolo en un predicado PROLOG que se numera automáticamente y que incluye predicados para las palabras y etiquetas que lo forman.

Entradas y Salidas. El segmentador de enunciados toma como entrada un texto con el formato PALABRA₁ LEMA₁ ETIQUETA₁ PALABRA₂ LEMA₂ ETIQUETA₂... y obtiene como salida un fichero con una serie de predicados PROLOG $o()$, cada uno de ellos con tres argumentos:

Número (arg_1): Es un número único para cada enunciado.

Descripción (arg_2): Breve descripción del tipo de corpus a tratar.

Enunciado (arg_3): Lista de predicados $w()$ de las palabras que forman el enunciado. Cada uno de estos predicados contiene a su vez tres argumentos correspondientes a la palabra, el lema y la etiqueta.

La figura 5.4 muestra un ejemplo de la salida proporcionada por el segmentador de enunciados.

5.2.3 Analizador sintáctico parcial SUPP

El analizador sintáctico parcial SUPP (Slot Unification Partial Parser) (Martínez-Barco *et al.*, 1998) utiliza el formalismo gramatical SUG (Slot Unification Grammar) (Ferrández *et al.*, 1997) implementado en PROLOG (LPA Prolog) en un sistema de análisis sintáctico denominado SUP. Sobre este sistema de análisis sintáctico completo SUP, se desarrolla el analizador parcial SUPP. Las SUG fueron desarrolladas como una extensión de las DCG (Definite Clause Grammars) (Pereira & Warren, 1980) con el objetivo de ampliar las capacidades de éstas para facilitar la resolución de manera modular de diversos problemas lingüísticos. Las SUG se denominan así debido a las estructuras de huecos o *slot structures* generadas automáticamente por el analizador donde se incluyen de forma automática toda la información morfológica, sintáctica y semántica necesaria para resolver diversos problemas lingüísticos. Las SUG se definen como una quintupla (NT, T, H, P, S) , donde NT es un conjunto finito de símbolos no terminales. T es un conjunto finito de símbolos terminales disjunto con NT . H son hechos SUG, pudiendo ser *coordination*, *juxtaposition*, *fusion*, *basicWord* o *isWord*. P son las reglas de producción de la gramática, un conjunto finito de reglas de la forma: $\alpha ++ > \beta$ ó β_1 , donde $\alpha \in NT$, $\beta \in (T \cup NT \cup \{\text{llamadas a$

142 5. ARIADNA: Sistema de resolución de la anáfora en diálogos

Enunciado:

OP: información de Renfe, buenos días
 US: buenos días, quería saber los horarios de trenes hacia Madrid
 ...

Entrada:

OP OP X
 información información NCFS000
 de de SPS00
 Renfe renfe NP00000
 , , Fc
 buenos buen AQOMP00
 días día NCMPO00
 . . Fp
 US US X
 buenos buen AQOMP00
 días día NCMPO00
 , , Fc
 quería querer VMII1S0
 saber saber VMN0000
 los el TDMP0
 horarios horario NCMPO00
 de de SPS00
 trenes tren NCMPO00
 hacia hacia SPS00
 Madrid madrid NP00000
 . . Fp
 ...

Salida:

```
o(1,dialogo,[w('OP','OP','SPEAKER'),w('información','información',
'NCFS000'),w('de','de','SPS00'),w('Renfe','renfe','NP00000'),w(
',',' ','Fc'),w('buenos','buen','AQOMP00'),w('días','día','NCMP000
'),w('.',' ','Fp')])
```

```
o(2,dialogo,[w('US','US','SPEAKER'),w('buenos','buen','AQOMP00'),
w('días','día','NCMP000'),w(',',' ','Fc'),w('quería','querer','VM
II1S0'),w('saber','saber','VMN0000'),w('los','el','TDMP0'),w('hor
arios','horario','NCMP000'),w('de','de','SPS00'),w('trenes','tren
','NCMP000'),w('hacia','hacia','SPS00'),w('Madrid','madrid','NP00
000'),w('.',' ','Fp')])
```

...

Figura 5.4. Salida ejemplo del segmentador de enunciados

procedimientos})* y $\beta_1 \in H$. S es el conjunto de símbolos iniciales de la gramática.

Por ser las SUG una extensión de las DCG, heredan muchas de sus características. La principal diferencia es que en las reglas de producción de la forma $\alpha \rightarrow \beta$ cada subconstituyente de β puede omitirse en la oración si se escribe entre el operador opcional: $\langle\langle$ constituyente $\rangle\rangle$. Este operador opcional tiene la posibilidad de recordar si el constituyente opcional ha sido analizado en la oración o no. Esta información es muy útil para la resolución de problemas lingüísticos tales como la elipsis o extraposición. Para comprobar la presencia o ausencia de un constituyente opcional se le añade una etiqueta, por ejemplo $\langle\langle SN : sn \rangle\rangle$ para el caso de un sintagma nominal. Esta etiqueta será una variable Prolog sin instanciar si no existe el constituyente sn , es decir, el predicado $Prolog \text{ var}(SN)$ tendrá éxito.

Analizador sintáctico SUP. El analizador sintáctico SUP usa como motor básico una gramática SUG que se traduce a Prolog. Hay traductores que convierten las reglas DCG en cláusulas Prolog añadiendo dos argumentos extra a cada símbolo que no está entre llaves (llamadas a procedimientos Prolog) o corchetes (símbolos terminales). Cada regla DCG, por ejemplo $o \rightarrow sn, sv$ es traducida a Prolog del modo siguiente: $o(L1, L3):- sn(L1, L2), sv(L2, L3)$, donde $L1$, $L2$ y $L3$ son las listas de palabras de la frase a analizar. De la misma forma, en (Ferrández *et al.*, 1997) se desarrolló un traductor similar de reglas SUG a cláusulas Prolog que incluyen la llamada *estructura de huecos (EH)* o *slot structure*. Esta *EH* almacena información sintáctica, morfológica y semántica de cada constituyente. En el ejemplo 84 se puede comprobar la estructura de huecos simplificada para la oración “Los equipos auxiliares incluyen un dispositivo indicador de llamada exterior”.

- (84) **EH:** oracion(sn(determinante(los), -, sust(equipos), adjetivo(auxiliares), -), sv(verbo(incluyen), -, -, sn(determinante(un), -, sust(dispositivo), adjetivo(indicador), sp(

```

preposicion( de), sn( _, _, sust( llamada), ad-
jetivo( exterior), _ )))).

```

Aquellos huecos correspondientes a constituyentes opcionales que no han aparecido en la frase de entrada quedan como variables sin instanciar. La gramática SUG usada para construir la estructura de huecos anterior se muestra en la figura 5.5.

```

oracion ++> <<sn>>, sv.
sn ++> <<determinante>>, <<adj>>, sust, <<adj>>, <<sp>>.
sv ++> verbo, <<cc>>, <<sp>>, <<sn>>.
cc ++> adverbio.
sp ++>preposicion, sn.

```

Figura 5.5. Ejemplo de un gramática SUG

El analizador SUP accede una sola vez al diccionario durante todo el proceso de análisis para evitar repetidos accesos a la misma palabra del diccionario. Así, almacena la información de cada palabra en una lista (*lista SUG*) antes de comenzar el análisis y trabajará con esta estructura en lugar de la lista de palabras de un analizador DCG en Prolog, como se muestra en el siguiente ejemplo:

(85)

```

DCG: [este, libro, es, mío],
SUG: [word (este, [ adjSimple(sing, masc, dem),
pron(sing, masc, dem) ] ), word (libro, [
sust(...) ] ), ...].

```

Cada elemento de esta lista es una estructura con nombre *word* y con dos argumentos. El primero corresponde a la misma palabra de la oración como átomo Prolog. El segundo corresponde a una lista que almacena todas las entradas léxicas de la palabra. Por tanto, cada vez que el analizador tenga que acceder a la entrada léxica de una palabra la buscará en esta lista y no accederá al diccionario de nuevo. SUP está implementado con un algoritmo descendente.

Analizador Parcial SUPP. El Analizador SUPP (Martínez-Barco *et al.*, 1998) toma como entrada la salida de un analizador léxico-morfológico donde cada palabra contiene su categoría gramatical e información morfológica. Así las oraciones introducidas se transforman en formato de lista, donde cada etiqueta del analizador léxico-morfológico se sustituye por la etiqueta apropiada en la gramática SUG; por ejemplo, la etiqueta léxico-morfológica *trenes tren NCMP000* se transforma en la entrada: *word (tren, [sust (pl, masc, comun)])*. Esta transformación se realiza mediante un interfaz apropiado. Actualmente SUPP dispone de interfaces capaz de convertir dos juegos de etiquetas léxico-morfológicas: el juego de etiquetas del Xerox POS Tagger adaptado al español en el Proyecto CRATER (Proyecto CRATER, 1994-1995), y el juego de etiquetas PAROLE (Martí *et al.*, 1998) definido en el Proyecto ITEM (Proyecto ITEM, 1996-1999). Finalmente, esta lista de palabras será la entrada para la gramática descrita a continuación. Esta gramática realizará el análisis parcial del texto, y el sistema SUPP producirá la estructura de huecos final de la oración. Este simple interfaz entre el etiquetador y la gramática SUG es una de las ventajas de la modularidad que presentan las SUG: se puede trabajar con diferentes diccionarios o etiquetadores y usar la misma gramática SUG.

La gramática definida en SUPP analiza básicamente tres tipos de constituyentes principales:

- sintagmas nominales (que pueden estar coordinados) (sn)
- sintagmas preposicionales (que pueden estar coordinados) (sp)
- núcleos verbales (v)

Estos tres constituyentes pueden aparecer en cualquier orden, desechando cualquier otra palabra que encuentre. Las palabras desechadas se corresponden con constituyentes que no están cubiertos por la gramática (por ejemplo los adverbios) o palabras incorrectamente etiquetadas o erróneas. La inclusión de nuevos tipos de constituyentes a tratar (como por ejemplo, sintagmas adverbiales) se podría realizar fácilmente en la primera regla de la gramática junto a los ya existentes. En la figura 5.6 se muestra una versión simplificada de la gramática SUG definida en SUPP. La

gramática mostrada es un subconjunto extraído de la gramática original aplicada en ARIADNA que se ha simplificado aquí por claridad expositiva.

```

oracTag ++>
  << SP:sp >>, << SN:sn >>, <<V:nucleoVerbal>>,
  <# [ ],
  oracionRestante(SP,SN,V)
#> .
oracionRestante(SP,SN,V) ++>
  <## ( {( var(SP), var(SN), var(V))}, [W]),
  ( _ , _ )
##>,
oracTag.
%--- Reglas gramaticales para cada constituyente
coordinated( sp, simpleSP).
simpleSP ++> preposicion, sn.
coordinated( sn, simpleSN(_) ).
simpleSN(tipoSUST) ++>
  <<determinante>>, <<adjetivo>>, sust, <<adjetivo>>, <<sp>>.
simpleSN(tipoADJ) ++> <<determinante>>, adjetivo, <<sp>>.
simpleSN(tipoPRON) ++> pron.
nucleoVerbal ++>
  <# verboComp,
  perifVerbal,
  verbo(_,_)
#> .
verboComp ++>
  <<verbo(_,_)>>,
  verbo(_,_),
  <# verbo(_,_),
  verbo(participio,_)
#> .
perifVerbal ++>
  verbo(_,_),
  <# verbo(gerundio,noCopul),
  verbo(participio,noCopul)
#> .
adjetivo ++> <<adv>>, adjSimple.
adjetivo ++> <<adv>>, verbo.
determinante ++> art.
determinante ++> adjSimple.

```

Figura 5.6. Gramática SUG parcial simplificada

Esta gramática utiliza el operador opcional de la siguiente manera: << SS: constituyente >>. De este modo permite la comprobación, en el caso de no aparecer el constituyente,

mediante el predicado predefinido de Prolog `var(SS)`. También utiliza el operador `<# a, b #>`, que permite analizar el constituyente *a*, y si no se encuentra en el texto, entonces intenta analizar *b*; en caso de no encontrarse fracasará. El operador `<## a, b ##>` es similar al anterior salvo que tendrá éxito aunque no se encuentren ni *a* ni *b*. Esta gramática permite la coordinación de *sn* y *sp* mediante el hecho `SUG coordinated`. Por ejemplo, `coordinated(sn, simpleSN)` sería equivalente a las siguientes reglas DCG: `sn --> simpleSN, conjuncion, sn` y `sn --> simpleSN`. Este tipo de reglas `SUG` permitirá coordinar constituyentes de diferentes formas, por ejemplo, sintagmas nominales formados por un sustantivo, pronombre o verbo. El símbolo inicial de la gramática será *oracTag* que comenzará el proceso del análisis parcial aplicando las reglas de la gramática anterior. Como salida, el analizador parcial `SUPP` generará una estructura de huecos que nos mostrará los constituyentes analizados, y el orden en que los encontró.

En la figura 5.7 se muestra un ejemplo de la aplicación del análisis parcial sobre una frase del corpus. En ella se ha representado la lista de palabras que genera el analizador con las etiquetas originales y a continuación la salida del programa.

Especialmente importante en la resolución de la anáfora es la información que `SUPP` genera sobre los sintagmas nominales. El analizador sintáctico genera cada uno de los sintagmas nominales e incluye en su estructura de huecos la información sintáctica, morfológica y léxica (categorías gramaticales) asociada. Además, proporciona una clasificación del sintagma nominal en los siguientes tipos:

- Sintagma Nominal Sustantivo de tipo común (`SNcomun`): que contiene un nombre común como núcleo del sintagma.
- Sintagma Nominal Sustantivo de tipo propio (`SNpropio`): que contiene un nombre propio como núcleo del sintagma.
- Sintagma Nominal Adjetivo (`SNadj`): La presencia de un adjetivo en el sintagma y ausencia de un nombre genera la clasificación del sintagma nominal como adjetivo.

148 5. ARIADNA: Sistema de resolución de la anáfora en diálogos

Enunciado:

<US> buenos días , quería saber los horarios de trenes hacia Madrid

Corpus:

```
[w('US', 'US', 'SPEAKER'), w('buenos', 'buen', 'AQOMPO0'),
w('días', 'día', 'NCMPO00'), w(' ,', ' ,', 'Fc'), w('quería',
'querer', 'VMII1SO'), w('saber', 'saber', 'VMNO000'), w('l
os', 'el', 'TDMPO'), w('horarios', 'horario', 'NCMPO00'), w
('de', 'de', 'SPSO0'), w('trenes', 'tren', 'NCMPO00'), w('h
acia', 'hacia', 'SPSO0'), w('Madrid', 'madrid', 'NPO0000'),
w(' .', ' .', 'Fp')]
```

Resultado del análisis parcial:

```
** PALABRA: US
** SINT.NOMINAL:
  ** SINT.NOMINAL SIMPLE:
    ** ADYACENTE ADJETIVO:
      ** ADJETIVO SIMPLE: buen
    ** SUSTANTIVO: día
** CONJUNCION:
  ,
** NUCLEO VERBAL:
  ** VERBO COMPUESTO:
    ** VERBO: querer
    ** VERBO: saber
** SINT.NOMINAL:
  ** SINT.NOMINAL SIMPLE:
    ** DETERMINANTE 1:
      ** ARTICULO: el
    ** SUSTANTIVO: horario
    ** SINT.PREPOSICIONAL:
      ** SINT.PREPOSICIONAL SIMPLE:
        ** PREPOSICION:
          ** PREPOSICION SIMPLE: de
      ** SINT.NOMINAL:
        ** SINT.NOMINAL SIMPLE:
          ** SUSTANTIVO: tren
        ** SINT.PREPOSICIONAL:
          ** SINT.PREPOSICIONAL SIMPLE:
            ** PREPOSICION:
              ** PREPOSICION SIMPLE: hacia
          ** SINT.NOMINAL:
            ** SINT.NOMINAL SIMPLE:
              ** SUSTANTIVO: madrid
```

Figura 5.7. Ejemplo de análisis parcial con SUPP

- Sintagma Nominal de Pronombre en un Sintagma Preposicional (SNpronSP): se trata de un sintagma nominal cuyo núcleo es un pronombre y además este sintagma se encuentra incluido en un sintagma preposicional.
- Sintagma Nominal de Pronombre con función de Sujeto (SNpronSuj): cuyo núcleo del sintagma nominal es un pronombre y no se encuentra incluido en la categoría anterior.
- Sintagma Nominal Verbo en Infinitivo (SNverboInf): el núcleo del sintagma es un verbo en infinitivo.

Estos tipos de sintagma nominal serán usados por ARIADNA para detectar las anáforas a tratar así como sus candidatos a antecedente.

Entradas y salidas. El analizador parcial SUPP toma como entrada la salida de un etiquetador léxico-morfológico que ha sido procesada por el segmentador de enunciados y obtiene como salida una estructura de huecos por cada constituyente encontrado en la que se almacena la estructura sintáctica del componente junto con una estructura llamada *paral* que almacena información sintáctica adicional.

La estructura de huecos es un predicado PROLOG con los siguientes argumentos:

Functor (arg_0): es el nombre del predicado PROLOG. Corresponde al constituyente que representa (sn, sp, etc.).

Conc (arg_1): incluye la información de concordancia del constituyente (información morfológica de número, género y persona).

X (arg_2): variable usada para definir la fórmula lógica cuando se hace análisis semántico.

$\text{arg}_3 \dots \text{arg}_n$: el resto de los argumentos contiene las estructuras de huecos de sus subconstituyentes.

En las figuras 5.8 y 5.9 se muestra la estructura de huecos generada por SUPP para el enunciado cuyo análisis aparece en la figura 5.7. El predicado *oracTag* inicial contiene todos los constituyentes analizados en ese enunciado.

La estructura *paral* contiene información adicional del análisis sintáctico en los siguientes argumentos:

150 5. ARIADNA: Sistema de resolución de la anáfora en diálogos

```

oracTag(
  conc(,_,
  w(US),
  sn(
    conc(pl,masc,terc,sustAdj,_)_,_
    snCoordinado(
      conc(pl,masc,terc,sustAdj,_)_,_
      snSimple(
        conc(pl,masc,terc,sustAdj,_)_,_
        snSimplePodado(
          conc(pl,masc,terc,sustAdj,_)_,_,_,_
          adyAdj(
            conc(pl,masc)_,_,_
            adj(
              conc(pl,masc,cal)_,_
              [adjSimple(
                conc(pl,masc,cal)_,_
                buen)],_,_)_,_)
          sust(
            conc(pl,masc,comun)_,_
            día),_,_,_,_,_,_)_,_)
        conj(,_,_(,)),
        nucleoVerbal(
          conc(,_,_,_)_,_
          verboComp(
            conc(,_,_,_)_,_,_
            verbo(
              conc(sing,prim,imperfecto,noCopul)_,_
              querer),
            verbo(
              conc(sing,_,infinitivo,noCopul)_,_
              saber))),
        sn(
          conc(pl,masc,terc,sustAdj,_)_,_
          snCoordinado(
            conc(pl,masc,terc,sustAdj,_)_,_
            snSimple(
              conc(pl,masc,terc,sustAdj,_)_,_
              snSimplePodado(
                conc(pl,masc,terc,sustAdj,_)_,_
                det(
                  conc(pl,masc)_,_
                  art(
                    conc(pl,masc,det)_,_
                    el)),_,_,_

```

Figura 5.8. Ejemplo de una estructura de huecos generada por SUPP (1/2)

Enunciado (arg_1): Contiene el número de enunciado en el corpus donde aparece el constituyente. El número de enunciado es único en el corpus y se asigna automáticamente en el módulo segmentador de enunciados.

Cláusula (arg_2): Contiene el número de cláusula dentro del enunciado donde aparece el constituyente. El número de cláusula es único dentro de cada enunciado. Las cláusulas se reconocen en el analizador parcial SUPP por la aparición de conjunciones y pronombres relativos que separan distintos núcleos verbales. Así cada núcleo verbal pasa a ser el núcleo de la cláusula, y la conjunción o el pronombre relativo marcan sus límites.

Posición respecto al núcleo verbal (arg_3): Este argumento únicamente puede tomar dos valores, *av* si el constituyente se encuentra delante del núcleo verbal en la cláusula, o *dv* si se encuentra detrás.

Número de constituyente (arg_4): Indica la posición del constituyente dentro de la cláusula.

Número de subconstituyente en coordinación (arg_5): Indica la posición de un subconstituyente dentro de un constituyente superior coordinado.

5.2.4 Etiquetado de la estructura del diálogo

ARIADNA necesita un etiquetado de la estructura del diálogo que siga el esquema de anotación propuesto en la sección 4.1. Actualmente ARIADNA no dispone de un módulo de etiquetado automático para la estructura del diálogo. El principal motivo es que no existen etiquetadores automáticos de pares adyacentes que funcionen en dominios no restringidos en español, y tampoco es nuestro objetivo desarrollarlo ya que los sistemas de diálogo ya contienen esta información. Sin embargo, hemos propuesto el empleo de SGML como lenguaje de anotación de esta estructura para permitir que cualquier editor de SGML pueda realizar de los etiquetados manuales, y además define un estándar para la comunicación entre ARIADNA y la aplicación donde se instale.

Interfaz SGML - PROLOG. Un interfaz apropiado definido en ARIADNA trasladará posteriormente el etiquetado de la estruc-

tura del diálogo desde el formato SGML a un formato PROLOG para que sea procesado finalmente por el *Generador del espacio de accesibilidad anafórica* y por el *Sistema de preferencias estructurales*. Este interfaz utiliza también la información generada previamente por el módulo *Segmentador de enunciados*.

Entradas y salidas. El módulo de etiquetado de la estructura del diálogo toma como entrada un texto plano que tras ser etiquetado en formato SGML se traduce a formato PROLOG por medio de un interfaz apropiado. Este interfaz genera un fichero con los predicados PROLOG *topic* y *par*. El predicado *topic* tiene tres argumentos que se describen a continuación:

Enunciado inicial (arg_1): Número de enunciado dentro del corpus desde el cual el tópico empieza a tener ámbito.

Enunciado final (arg_2): Número de enunciado dentro del corpus hasta el cual el tópico tiene ámbito.

Entidad (arg_3): Lista de palabras que forman la entidad tópico.

Por otra parte, el predicado *par* tiene los siguientes tres argumentos:

Número (arg_1): Número de par adyacente dentro del corpus.

Enunciado inicial (arg_2): Número de enunciado dentro del corpus desde el cual el par adyacente empieza a tener ámbito.

Enunciado final (arg_3): Número de enunciado dentro del corpus hasta el cual el par adyacente tiene ámbito.

La figura 5.10 muestra un ejemplo de etiquetado traducido a formato PROLOG.

5.3 Módulos de resolución anafórica en ARIADNA

A continuación se analiza en detalle cada uno de los módulos específicos para resolución anafórica del sistema.

Además, el disparador anafórico revisa la información léxica almacenada del núcleo de los sintagmas nominales *SNpronSuj* buscando pronombres que sólo actúan como complemento. De esta forma, el disparador anafórico pone en marcha la resolución de la anáfora distinguiendo los siguientes tipos de anáfora:

- Anáfora pronominal de Complemento (*compl*): Se trata de un sintagma nominal de tipo *SNpronSuj* cuyo núcleo es un pronombre personal de complemento.
- Anáfora pronominal de Sujeto (*pronSuj*): cualquier otro caso de sintagma nominal de tipo *SNpronSuj*
- Anáfora pronominal de Sintagma Preposicional (*pronSP*): generada por un sintagma nominal de tipo *SNpronSP*
- Anáfora Adjetiva (*adj*): generada por un sintagma nominal de tipo *SNadj*.

Además, el disparador distingue los pronombres demostrativos (éste, ése, aquel, ésta, esa, aquella, esto, eso, aquello, éstos, éstos, aquéllos, éstas, éstas, aquéllas) de los pronombres personales. Dicha información se genera como parámetro de entrada al módulo de resolución de la anáfora.

5.3.2 Generador del espacio de accesibilidad anafórico

El módulo disparador anafórico activa el proceso de resolución de anáfora que se inicia con el Generador del espacio de accesibilidad anafórica. Este módulo ha sido construido tomando como referencia las conclusiones generadas a raíz de los estudios desarrollados en la sección 4.3 sobre la influencia de la estructura del diálogo en la resolución de la anáfora, y en concreto en la definición de un espacio de accesibilidad anafórica basado en la estructura del diálogo. En su funcionamiento básico, el módulo extrae del análisis sintáctico la lista de sintagmas nominales que se encuentran en el espacio de accesibilidad anafórica estructural formado por:

1. los SNs que estén incluidos en el mismo par adyacente que la anáfora,
2. los SNs que estén incluidos en el par adyacente anterior al de la anáfora,

3. los SNs que están incluidos en algún par adyacente que incluya al de la anáfora, y
4. el tópicico definido para el segmento donde se encuentra la anáfora.

Esta lista de sintagmas nominales se convierte en la lista inicial de candidatos de la anáfora que deberá ser procesada por el sistema de restricciones.

5.3.3 Sistema de Restricciones

El Sistema de Restricciones está formado por un conjunto de reglas cuyo cumplimiento debe ser obligatorio en cualquier antecedente anafórico (Palomar & Martínez-Barco, 2001). Este conjunto de reglas utiliza información lingüística para su definición (información léxica, morfológica y sintáctica). El conjunto de reglas restrictivas definido se detalla a continuación.

Restricciones morfológicas de los pronombres (R1p). El conjunto de las restricciones morfológicas aplicables a los pronombres (que etiquetaremos como R1p para referencias posteriores) incluye aquellas restricciones que pueden ser aplicadas sólo mediante el conocimiento morfológico de las palabras y constituyentes tratados.

- *Concordancia morfológica en género.* Cualquier pronombre en cuya información morfológica se especifique el género deberá concordar en género con su antecedente. En el ejemplo 86 el antecedente de *ella* cuyo género es femenino siempre debe hacer referencia a un sintagma nominal con género femenino y nunca al masculino. Los pronombres con género neutro pueden concordar tanto con antecedentes de género masculino como con el femenino.

(86) **A:** Ayer estuve comiendo con Luis y María_i.

B: ¿Hablaste con **ella**_i sobre el proyecto nuevo?

- *Concordancia morfológica en persona.* La concordancia morfológica en persona se aplica a las anáforas de pronombre personal.

En concreto, ARIADNA se centra en la resolución de pronombres de tercera persona únicamente puesto que el caso de los pronombres de primera y segunda persona en un diálogo con dos participantes es una tarea trivial, la primera persona siempre hace referencia al hablante, y la segunda al oyente, tal y como se muestra en el ejemplo 87. Este fenómeno se complica cuando se trata de diálogos multiparticipados con más de dos hablantes. No obstante, en estos casos la anáfora se convierte en deixis y no es objetivo de este trabajo.

(87) **A_i**: Ayer estuve comiendo con Luis y María.

B: ¿**Te_i** contaron su nuevo proyecto?

- *Concordancia morfológica en número.* La concordancia morfológica en número es una de las reglas restrictivas características de la anáfora pronominal. Un pronombre en singular siempre hace referencia a un antecedente en singular. El caso de los pronombres en plural debe ser estudiado con mayor detalle. Algunos pronombres en plural pueden hacer referencia a dos o más sintagmas nominales en singular que se convierten en plural al ser coordinados como ocurre en el ejemplo 88. Por este motivo, es importante que el analizador sintáctico proporcione la nueva información morfológica de los componentes que genera por coordinación.

(88) **A**: Ayer estuve comiendo con Luis y María_i.

B: ¿Hablaste con **ellos_i** sobre el proyecto nuevo?

Restricción morfológica de las anáforas adjetivas (R1a).

En el caso de la anáfora adjetiva se aplica únicamente la siguiente restricción:

- *Concordancia morfológica en género.* La concordancia morfológica en género se aplica también a la anáfora adjetiva tratada en ARIADNA. Cualquier anáfora adjetiva deberá concordar en género con su antecedente (ejemplo 89).

(89) **A:** Para viajar a Barcelona puedes tomar un tren_i
a las cinco que llega a la estación de Sants a
las nueve.

B: Ya, ¿y el siguiente_i a qué hora sale?

No se incluye restricción de concordancia morfológica en persona por carecer de ella, ni en número puesto que en ocasiones, la anáfora adjetiva hace referencias a antecedentes de distinto número, como en el ejemplo 90).

(90) **A:** Quisiera conocer el horario de los trenes a Valencia_i.

B: Pues mire, **el primero**_i sale a las siete treinta de la mañana.

Restricciones sintácticas de los pronombres (R2p). La anáfora intraoracional (en diálogos se debería hablar de anáfora intraenunciado) tiene unas condiciones sintácticas particulares que han sido ampliamente recogidas por diferentes teorías lingüísticas entre las que destacamos la *Teoría de las Condiciones Anafóricas C-Dominio (C-Command)* de Reinhart (Reinhart, 1983).

De acuerdo con esta teoría, en un árbol de análisis sintáctico un nodo A domina a un nodo B, si A se encuentra por encima de B en el árbol de análisis.

Además, se dice que un nodo A *c-domina* a un nodo B, si el nodo α_1 que domina inmediatamente a A, o bien domina a B, o es inmediatamente dominado por un nodo α_2 que domina a B, y α_2 es de la misma categoría que α_1 . Así, se define el dominio de un nodo A como el conjunto de todos los nodos *c-dominados* por A.

A partir de las condiciones de *c-dominio* se han propuesto las siguientes condiciones sintácticas para la anáfora:

1. Un SN no pronominal debe ser interpretado como no correferencial con cualquier SN completo que lo *c-domine*. En el ejemplo 91, los sintagmas nominales *la madre* y *la deportista*_j,

aunque sean compatibles, no pueden correferir puesto que *la madre c-domina a la deportista* al encontrarse el primero inmediatamente dominado por un nodo (el sintagma nominal *la madre de la deportista*) que domina al segundo⁴.

(91) **A:** La madre_i de **la deportista**_j llamó para felicitarla.

B: Supongo que estaría bien orgullosa.

- Un pronombre debe ser interpretado como no correferencial con cualquier SN completo al que *c-domine*. En el ejemplo 92 el pronombre *él* c-domina a *Juan* al estar inmediatamente dominados ambos por el mismo sintagma nominal coordinado, por tanto, serán no correferentes.

(92) **A:** Juan_i y él_j estuvieron muy amables.

B: Sí, siempre lo han sido.

- Un pronombre reflexivo o recíproco debe ser interpretado como correferencial con (y sólo con) un SN que lo *c-domine* dentro de un dominio sintáctico específico: su categoría mínima de gobierno. Se considera *categoría mínima de gobierno* de un nodo A, el menor nodo que determine la oración, o un sintagma nominal que c-domine a A y a su nodo principal.

(93) **A:** Juan_i se_i miró en el espejo.

B: Espero que le gustara el nuevo corte de pelo.

- Un pronombre no reflexivo ni recíproco debe ser interpretado como no correferencial con cualquier SN que lo *c-domine* en su categoría mínima de gobierno. En el ejemplo 94 el sintagma nominal *el primo* c-domina al pronombre *él*, y además se

⁴ En realidad, un análisis sintáctico correcto en español nunca extraería como sintagma nominal completo *la madre* puesto que siempre incluiría el sintagma preposicional *de la deportista* como modificador del núcleo *madre*. Sin embargo, puesto que esta estructura se permite en inglés por medio del genitivo sajón, hemos hecho una adaptación de cómo funcionaría en español.

encuentra dentro de su categoría mínima de gobierno que es *el primo*, un sintagma nominal que *c-domina* al pronombre *él* y a su nodo superior, el sintagma preposicional *de él*.

- (94) **A:** El primo_i de él_j estuvo aquí el otro día.
B: Ah!, ¿y qué se contaba?.

En ARIADNA se ha tomado como punto de partida la restricción de Reinhart, aunque se adaptan las condiciones a la información de la que dispone nuestro sistema. Esto significa básicamente redefinir las condiciones con la utilización de información sintáctica parcial en lugar de información sintáctica completa, adaptarlas a los diálogos y al español.

El filtro sintáctico pronominal se aplica en ARIADNA tomando como salida la información generada por el analizador sintáctico parcial SUPP. Éste proporciona información sobre sintagmas nominales, sintagmas preposicionales, verbos y *chunks* verbales, así como las relaciones entre ellos. Por otra parte, aunque no se analizan oraciones y no se definen las funciones gramaticales de los constituyentes, el analizador parcial utiliza un segmentador capaz de segmentar los turnos en enunciados, y además, detecta cláusulas subordinadas⁵. De esta forma, de cada constituyente analizado se conoce un número de enunciado al que pertenece y un número de cláusula dentro de éste si hubiera subordinadas.

Con esta información se definen las siguientes condiciones sintácticas para ARIADNA:

- *Condiciones sintácticas de pronombres no reflexivos ni recíprocos.* Un pronombre P no reflexivo ni recíproco no puede referir con un sintagma nominal SN_i si:
 - P y SN_i están en el mismo enunciado y en la misma cláusula y P modifica al núcleo de un SN, y SN_i modifica al mismo SN que P, como ocurre en la coordinación del SN_i *Juan* y el P *él* que modifican al SN *Juan y él* en el ejemplo 95.

⁵ El analizador detecta una nueva cláusula siempre que encuentre una conjunción que no se usa para coordinar constituyentes

(95) **A:** Juan_i y él_j estuvieron aquí el otro día.

B: Ah!, ¿y qué se contaban?.

– P y SN_i están en el mismo enunciado y en la misma cláusula y P no modifica al núcleo de ningún SN

(96) **A:** Juan_i habló seriamente con él_j.

B: Espero que le haga caso.

A diferencia de la propuesta de Reinhart, consideraremos que la condición de no correferencialidad de los pronombres reflexivos o recíprocos con sintagmas nominales que se encuentren fuera de su categoría mínima de gobierno no es válida para los diálogos en español. Esto es debido a que en español los pronombres reflexivos y recíprocos pueden hacer referencia a sujetos omitidos dentro de la categoría mínima de gobierno, con lo cual, a menos que se resuelva dicha omisión su antecedente debe ser buscado fuera de ella, como ocurre en el ejemplo 97.

(97) **A:** ¿Cómo va Juan_i?

B: Ya está casi preparado.

A: ¿Y \emptyset_i se_i ha afeitado ya?

Restricciones sintácticas de la anáfora adjetiva (R2a). En el caso de la anáfora adjetiva se aplica la siguiente condición:

- *Condición sintáctica de la anáfora adjetiva.* La anáfora adjetiva nunca puede correferir con un antecedente cuyo núcleo sea un nombre propio. La información sobre los distintos tipos de sintagma nominal proporcionada por el analizador sintáctico parcial SUPP permite a ARIADNA aplicar fácilmente esta condición eliminando los candidatos de tipo *Sintagma Nominal Sustantivo de tipo Propio*. En el ejemplo 98 se muestra un caso de no correferencia según esta condición.

(98) **A:** La primera parada es en Zaragoza_i.

B: ¿Y la siguiente_j?

Destacamos que en cualquiera de los casos tratados, tanto para la anáfora pronominal como para la adjetiva, puesto que la información sintáctica recibida es parcial, no se puede garantizar que siempre se disponga de la información necesaria para aplicar el filtro. Por este motivo se ha usado un filtro negativo en lugar de positivo, es decir, en lugar de aceptar aquellos candidatos que cumplen las condiciones, se eliminan los que no cumplen. De esta forma, si la información sintáctica disponible no es suficiente para determinar si se debe o no eliminar un candidato, éste permanecerá para que sean filtros posteriores los que lo eliminen.

5.3.4 Sistema de Preferencias

El sistema de preferencias se define como un conjunto de reglas que se aplican para ordenar los candidatos que han superado los filtros del sistema de restricciones. Tras la aplicación de las preferencias, de acuerdo con las condiciones que haya cumplido, cada uno de los candidatos obtiene un valor de preferencia. Aquel que obtenga el máximo valor será el propuesto como antecedente en la resolución de la anáfora.

Las reglas del sistema de preferencias deben representar una serie de condiciones cuyo cumplimiento no sea obligatorio, aunque sí aconsejable para el candidato que se proponga finalmente. El sistema de preferencias de ARIADNA se presenta en dos módulos diferentes atendiendo a la fuente de información que los alimenta. Por un lado, *el sistema de preferencias estructurales* define aquellas condiciones que hacen más probable a un candidato respecto a los demás según el segmento de la estructura del diálogo donde aparece con respecto a la anáfora. Por otra parte, *el sistema de preferencias lingüísticas* define las condiciones lingüísticas que hacen más probable a un candidato que las cumple respecto a otro que no las cumple.

A continuación se describe con detalle cada una de las preferencias aplicadas en ARIADNA.

Sistema de preferencias estructurales. El estudio empírico sobre el corpus de entrenamiento para nuestro sistema ha determi-

nado que es más probable encontrar el antecedente de una anáfora en ciertos segmentos que en otros. De esta forma el sistema de preferencias estructurales considera las siguientes prioridades en la ordenación de candidatos:

- E1. *Coincidencia del par adyacente*. Preferencia por los candidatos que se encuentren en el mismo par adyacente que la anáfora,
- E2. *Vecindad del par adyacente*. Preferencia por los candidatos que sin estar en el mismo par adyacente se encuentren en el par adyacente anterior a la anáfora,
- E3. *Anidamiento del par adyacente*. Preferencia por los candidatos que no estén en ninguno de los dos anteriores pero se encuentren en algún par adyacente que contenga al par adyacente de la anáfora,
- E4. *Topicalidad*. Preferencia por el candidato que ha aparecido en el discurso y representa al tópico del diálogo o segmento.

Sistema de preferencias lingüísticas. La definición del sistema de preferencias lingüísticas utilizado en ARIADNA se parte de los estudios sobre preferencias lingüísticas aplicadas por Ferrández (Ferrández, 1998) a la resolución de la anáfora en textos monólogos para el español. A partir de este conjunto de preferencias se realizaron los experimentos de entrenamiento para ARIADNA sobre la idoneidad de las mismas tal y como se muestra en el capítulo 6. Tras este estudio se ha determinado que el conjunto de preferencias lingüísticas idóneo en ARIADNA para cada tipo de anáfora es el siguiente:

Anáfora pronominal.

- L5p⁶. *Repetición*. Se ha comprobado que la preferencia por los candidatos que se han repetido alguna vez en el texto es útil en el tipo de anáfora que hemos considerado como anáfora pronominal de sintagma preposicional. Este tipo de anáfora, que ocupa un rol secundario en el análisis sintáctico del enunciado,

⁶ Para mayor claridad hemos mantenido la numeración de las preferencias tal y como aparece en los experimentos del capítulo 6, aunque debido al descarte de algunas de ellas, en el sistema final se pierde la correlatividad

busca generalmente antecedentes de importancia secundaria, y por tanto no aparecen en aquellos seleccionados como tópico. Por ello, el uso de esta preferencia adicional a la estructural del tópico se ha considerado necesaria. En otros casos de anáfora pronominal no se aplicará ya que la información estructural proporcionada por el tópico se considera suficiente.

L6p. *Paralelismo verbal.* La coincidencia de verbo entre la anáfora y su antecedente es una de las preferencias que ha sido valorada en el caso de las anáforas demostrativas. Si bien para el resto de las anáforas no se ha demostrado su utilidad, durante la fase de entrenamiento de ARIADNA se encontró que las anáforas demostrativas tenían una cierta preferencia por aquellos candidatos a antecedente que habían aparecido en el diálogo compartiendo el mismo verbo. Su aplicación se demostró que mejoraba la resolución de este tipo de anáforas durante el entrenamiento.

L7p. *Paralelismo posicional respecto al verbo.* Los estudios empíricos realizados sobre el corpus de entrenamiento nos han mostrado una cierta preferencia de la anáfora pronominal de complemento por aquellos candidatos que ocupan ciertos roles sintácticos según el rol que cumple el propio pronombre. Sin embargo, puesto que en la información sintáctica parcial no disponemos de esta información, se realizaron experimentos en el entrenamiento orientados a demostrar la influencia de la posición de anáfora y antecedente respecto al verbo del enunciado, es decir, si aparecen ambos antes o después del verbo. Tras los experimentos se demostró que los pronombres que aparecían antes del verbo tenían preferencia por antecedentes que aparecían antes del verbo, y los pronombres que aparecían tras el verbo, la tenían sobre los que aparecían tras él. La aplicación de esta preferencia mejoraba la resolución de las anáforas de complemento y por tanto se aplicó únicamente a este caso pronominal.

L8p. *Paralelismo posicional respecto al enunciado.* Se valoran los candidatos que estén en la misma posición sintáctica que la anáfora (posición de sujeto, posición de objeto, de complemento). De nuevo, y puesto que el analizador sintáctico parcial

no proporciona roles gramaticales, la preferencia aquí aplicada se basa en la posición de aparición del constituyente respecto al inicio del enunciado, esto es, si es el primer constituyente, segundo, etc. Esta preferencia y la anterior sustituyen a los criterios de preferencia por paralelismo de categorías gramaticales propuesto por otros sistemas que usan análisis sintáctico completo. Esta preferencia se ha considerado útil únicamente en el caso de la anáfora pronominal de complemento.

L12p. *Proximidad.* La selección del candidato más cercano a la anáfora es un criterio de preferencia a valorar especialmente en aquellos casos de empate en los que ninguna otra preferencia sea capaz de decidir un candidato final, tal y como ocurre en el procesamiento anafórico humano.

Anáfora adjetiva.

L3a. *Paralelismo de tipo de modificador.* Se prefieren los candidatos que compartan la misma clase de modificador (por ejemplo, un sintagma preposicional, un adjetivo, etc.) con la anáfora.

L4a. *Paralelismo de modificador.* Se prefieren después aquellos candidatos que compartan exactamente el mismo modificador (por ejemplo, el mismo adjetivo 'rojo').

L5a. *Concordancia morfológica en número.* Esta concordancia que no fue considerada como restricción, ha sido valorada, sin embargo, como una preferencia.

L8a. *Proximidad.* Por último, igual que ocurre en la anáfora pronominal, la proximidad se valorará en caso de empate.

Manejo de preferencias. Hay dos tendencias de tratamiento para las preferencias tal y como expone Ferrández (Ferrández, 1998), un tratamiento ordenado y un tratamiento ponderado:

- El tratamiento ordenado de las preferencias asigna cada una de las preferencias a un nivel distinto asignando mayor importancia a la preferencia de primer nivel y restando la importancia en niveles subsecuentes. De esta forma, las preferencias se empiezan a comprobar desde el primer nivel en adelante distinguiendo en cada nivel las siguientes posibilidades:

1. que ninguno de los candidatos de la lista cumpla con la preferencia de ese nivel: en este caso se procede a comprobar la siguiente preferencia
2. que sólo un candidato la cumpla: en este caso éste es la solución elegida
3. que varios candidatos la cumplan: se comprueba la siguiente preferencia pero ya sólo con aquellos elementos que cumplen ésta

Este tipo de tratamiento es computacionalmente bastante eficiente puesto que no siempre es necesario comprobar todas las preferencias. En el mejor de los casos, al comprobar la primera preferencia el sistema podría ya responder. Como desventaja, presenta una falta de flexibilidad al asignar importancias. Por ejemplo, si hay dos candidatos A y B, donde A cumple únicamente la preferencia de primer orden, y B cumple todas las demás preferencias excepto la primera, es siempre A el candidato propuesto y no B.

- El tratamiento ponderado de las preferencias consiste en asignar un peso a cada una de las preferencias. De esta forma, cada candidato tendrá asignado un valor que será el sumatorio de los pesos de aquellas preferencias que cumple. El candidato que obtenga el mayor valor será el propuesto finalmente. Como ventaja respecto al anterior, este modo de tratamiento permite mucha mayor flexibilidad gracias a que cada preferencias tiene un valor de importancia definible de forma independiente al resto. Como desventaja, para cada candidato se deben comprobar todas las preferencias.

En ARIADNA se han implementado ambos tipos de tratamiento siendo asignados a diferentes versiones del sistema. La primera versión utilizó un sistema de preferencias ordenado que sirvió para entrenar manualmente al sistema seleccionando las preferencias útiles. La segunda versión incorporó un sistema de preferencias ponderado que se usó para el entrenamiento del sistema al medir la importancia de cada una de las preferencias en la obtención de un resultado final. Los diferentes pesos asignados muestran la definición de tales importancias. La ordenación de las preferencias

y la asignación de pesos se muestra en detalle en la sección 6.4. Los mejores resultados obtenidos en el proceso de entrenamiento se consiguieron de acuerdo al sistema ponderado que se muestra en las tablas 5.1 y 5.2.

Anáfora Pronominal					
Preferencias		Pesos			
		Personal	Demostr	Compl	SP
Tipo	Pref.	Suj	Suj	Compl	SP
Estructural	E1	35	35	35	35
	E2	20	20	0	25
	E3	30	30	0	20
	E4	15	15	15	15
Lingüística	L5p	0	0	0	5
	L6p	0	5	0	0
	L7p	0	0	5	0
	L8p	0	0	5	0
	L12p	SI	SI	SI	SI

Tabla 5.1. Preferencias de la anáfora pronominal en ARIADNA v.2.

Anáfora Adjetiva		
Preferencias		
Tipo	Pref.	Pesos
Estructural	E1	35
	E2	10
	E3	10
	E4	35
Lingüística	L3a	5
	L4a	5
	L5a	5
	L8a	SI

Tabla 5.2. Preferencias de la anáfora adjetiva en ARIADNA v.2.

5.3.5 Algoritmo ARDi

Una vez definidas las fases principales del procedimiento de resolución de la anáfora se define el algoritmo ARDi (Anaphora

Resolution in Dialogues) capaz de manejar la información proporcionada por dichas fases. El pseudocódigo de dicho algoritmo se muestra en la figura 5.11.

```

Procedimiento RESOLUCION (A,L(AAS))
Sea ANAFORA = la anáfora A
Sea EAA = Espacio de Accesibilidad Anafórica de A
Sea LISTA = una lista L(EAA) de todos los SNs candidatos a
                antecedente en EAA

Para cada SN de LISTA
    aplicar restricciones para obtener la lista LISTA1
FinPara
Para cada SN de LISTA1
    aplicar preferencias estructurales para obtener LISTA2
    ordenada
FinPara
Para cada SN de LISTA2
    aplicar preferencias lingüísticas para obtener LISTA3
    ordenada
FinPara
Devolver la cabeza de LISTA3
FinProcedimiento

```

Figura 5.11. ARDi: Algoritmo de resolución de la anáfora

5.4 Adaptabilidad de ARIADNA a un sistema de diálogo

El sistema ARIADNA ha sido concebido de forma totalmente independiente a la aplicación a la que finalmente se destina. Por este motivo, para su definición se ha contado con las herramientas lingüísticas y estructurales que se han considerado necesarias para la obtención de toda la información usada en el proceso de resolución de la anáfora. Las primeras, por medio de procesos de análisis automático y las segundas a través de la anotación manual de acuerdo con el esquema definido para marcar la estructura del diálogo. De esta forma, ARIADNA nace como un sistema totalmente independiente del dominio y de la aplicación, y este hecho lo hace precisamente aplicable a cualquier aplicación de diálogo

trabajando sobre cualquier dominio. Sin embargo, puesto que la aplicación principal del sistema es su integración en un sistema de diálogo, a continuación se exponen algunos aspectos relativos al acoplamiento de ambos sistemas.

Tal y como se ha expuesto anteriormente, el núcleo de ARIADNA (el conjunto formado por lo que hemos llamado módulos de resolución anafórica en la sección 5.3: el módulo disparador anafórico, el generador del espacio de accesibilidad anafórica, el sistema de restricciones, el sistema de preferencias estructurales y el sistema de preferencias lingüísticas) utiliza directamente la información generada por el etiquetado de la estructura del diálogo y por el analizador sintáctico parcial SUPP.

Por otra parte, un sistema de diálogo genérico concebido según la arquitectura de Allen *et al.* (2000) dispondría de un parser que incorporaría el análisis léxico, morfológico, sintáctico y semántico de los actos de habla reconocidos. Esta información la tomaría el módulo gestor de discurso que reconocería y construiría la estructura del diálogo en cuestión. El módulo gestor del discurso en el sistema de diálogo contendría por tanto toda la información que necesita el núcleo de ARIADNA para resolver la anáfora.

Así, la tarea de adaptación de ARIADNA a un sistema de diálogo genérico pasaría por construir el interfaz correspondiente entre el sistema gestor del discurso y el núcleo de ARIADNA. Este interfaz debería expresar la información lingüística contenida en el gestor del discurso de acuerdo a la salida expresada por el analizador sintáctico parcial SUPP (ver sección 5.2.3). Además, la estructura del diálogo podría ser generada utilizando el esquema de anotación mediante SGML definido en la sección 4.4.2. Por último quedaría por convenir cómo realizar la salida de ARIADNA hacia el gestor del discurso en el sistema de diálogo. Esta salida dependería ya de las necesidades del propio gestor del discurso.

Cabría destacar, por último, que la definición de sus propias herramientas de preproceso lingüístico permitiría a ARIADNA completar la información proporcionada por el sistema de diálogo si es que dicha información fuera insuficiente para la resolución anafórica.

5.5 Conclusiones del capítulo

En este capítulo se ha presentado ARIADNA, un sistema completo para la resolución de la anáfora en diálogos basado en la estructura del diálogo. Este sistema de restricciones y preferencias utiliza, por una parte, la información lingüística (léxica, morfológica y sintáctica) proporcionada por un etiquetador léxico-morfológico junto con un analizador parcial para definir las restricciones y las preferencias usadas, y por otra parte, la información que sobre la estructura del diálogo le proporciona el etiquetado de dicha estructura.

Actualmente, el sistema trata las anáforas pronominales generadas por los pronombres personales de tercera persona⁷ y los pronombres demostrativos, además de las anáforas de tipo adjetiva.

Además debido a la estructura modular del sistema, éste queda abierto al tratamiento de otros tipos de anáfora sin la necesidad de invertir un mayor esfuerzo que la definición de las restricciones y preferencias adecuadas previa realización de un estudio correspondiente. El sistema ARIADNA ha sido implementado en PROLOG con el mostrado ARDi como algoritmo principal.

Destacamos, por último, la versatilidad del sistema ARIADNA que dada su independencia de dominio y de aplicación permite su adaptación a diferentes aplicaciones finales entre las que destacamos los sistemas de diálogo. Dicha adaptación se produciría con solo definir el interfaz apropiado.

⁷ El tratamiento de los pronombres de primera y segunda persona en el diálogo se considera una tarea trivial aunque no se puede tratar como una anáfora propiamente dicha puesto que los antecedentes no pueden encontrarse dentro del contexto lingüístico. Pertenece por tanto a una clase de exófora: la deixis personal



Universitat d'Alacant
Universidad de Alicante

6. Experimentación

La definición de un sistema para la resolución de la anáfora basado en restricciones y preferencias como ARIADNA requiere una definición del sistema que permita obtener la mejor configuración posible para sus metas. Esta configuración incluye tres objetivos diferentes:

- la definición de un espacio de accesibilidad anafórica óptimo,
- la definición de un conjunto de restricciones válido
- y la definición del conjunto de preferencias junto con una valoración de la importancia de cada una de ellas.

En ARIADNA, la definición de un espacio de accesibilidad anafórica óptimo se ha llevado a cabo tras la ejecución de los estudios empíricos mostrados en la sección 4.3. Por otra parte, la definición de un conjunto de restricciones válido se efectúa tras un estudio de las características propias de la lengua. Sin embargo, la definición de un conjunto de preferencias adecuado, así como la valoración de la importancia de cada una de ellas necesita un proceso de entrenamiento en el que se van ajustando los parámetros que regulan dicha importancia hasta obtener la configuración válida.

Para garantizar una experimentación independiente se ha de seleccionar un corpus, y de éste se toma una parte para el entrenamiento y se reserva el resto para la evaluación.

Esta fase de entrenamiento se deberá realizar mediante el corpus de entrenamiento elegido al azar sobre el cual se irán realizando las variaciones oportunas destinadas a conseguir cada uno de los objetivos citados. Para comprobar la efectividad del cambio efectuado se debe evaluar el sistema sobre el corpus de entrena-

miento después de la modificación y comprobar si los resultados han mejorado o por el contrario han empeorado.

Una vez que se ha ajustado el sistema con la mejor configuración, se tomará el corpus de evaluación y se realizará la evaluación final del sistema.

En la figura 6.1 se muestra el proceso completo de evaluación desarrollado sobre el sistema ARIADNA, tanto para el proceso de entrenamiento como para la evaluación final del sistema.

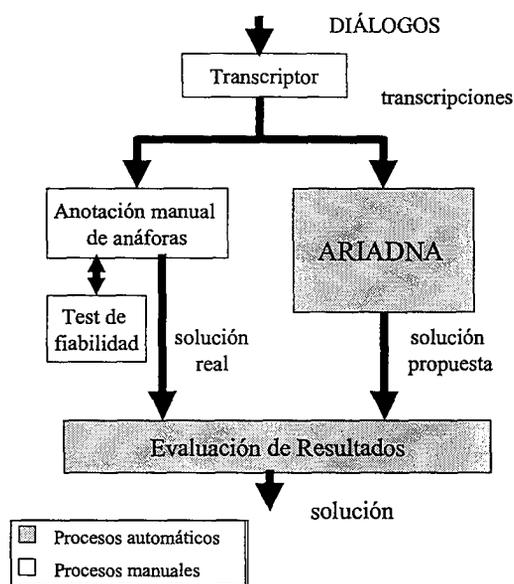


Figura 6.1. Proceso de evaluación de ARIADNA

En este capítulo se describe, en primer lugar, el proceso de entrenamiento desarrollado para el ajuste del sistema ARIADNA y a continuación el proceso de evaluación final del sistema.

6.1 Corpus

Para realizar la experimentación del sistema ARIADNA se dispone del corpus Infotren recopilado por el proyecto *Sistema de*

diálogo para habla espontánea en un dominio semántico restringido (BASURDE) (Proyecto BASURDE, 1998-2001) con dos subcorpus de diálogos definidos sobre el mismo dominio, diálogos en los que un usuario solicita información sobre trenes de viajeros, aunque con agentes de características diferentes. El primero de ellos es un corpus de diálogo real entre humanos y el segundo es un corpus de diálogo simulado hombre-máquina extraído de un experimento *Mago de Oz*¹. El corpus es propiedad del consorcio BASURDE formado por la Universitat Politècnica de Catalunya, Euskal Herriko Unibersitatea (Universidad del País Vasco), Universitat Jaume I, Universitat Politècnica de Valencia y Universidad de Zaragoza. Para el entrenamiento y evaluación de ARIADNA se ha elegido en concreto el subcorpus Infotren: Persona, decisión motivada principalmente por su mayor contenido anafórico.

6.1.1 Corpus Infotren: Persona

El corpus Infotren: Persona contiene llamadas reales en español realizadas al servicio de información de RENFE (Barcelona) durante noviembre de 1997. Estas llamadas fueron grabadas por RENFE, quien cedió las grabaciones al consorcio BASURDE para tareas de investigación. La transferencia de los diálogos se realizó mediante soporte DAT con una duración aproximada de 7 horas. La mayoría de los diálogos se refieren a conversaciones sobre trenes de largo recorrido lo que proporciona mayor riqueza informativa que aquellos referidos únicamente a cercanías.

El corpus fue transcrito de la siguiente forma: cada diálogo se almacenó en un fichero diferente, y cada línea del fichero representa un turno. Las líneas están precedidas por una cabecera de tres caracteres donde “US:” indica que el hablante del turno es el usuario del servicio y “OP:” indica que el hablante es el operador de la compañía. En algún caso, en el que interviene más de un usuario, se utilizan las marcas “US1:” y “US2:”. La transcripción usa el juego de caracteres ISO-8859 para la transcripción de

¹ Experimento en el que dos humanos conversan pero uno de ellos simula el comportamiento de una máquina en un sistema de diálogos.

caracteres no ascii (acentos, diéresis, interrogación, admiración, etc.). Además se han usado palabras como “uh” y “hum” para indicar sonidos de relleno, y códigos como <PAUSA> que indica pausas de discontinuidad del discurso, <RISA> que indica risas en la conversación, <SIMULTANEO> para indicar un solapamiento del turno con el anterior o <ININTELIGIBLE> que indica una palabra o fragmento que no se ha entendido.

En los 204 diálogos de los que consta el corpus hay aproximadamente 60.000 palabras con un léxico de 2.200 palabras distintas. Hay una media de 10 palabras por turno de operador y 6 por turno de usuario.

Para la evaluación del sistema, además ha sido necesaria realizar una anotación manual de la estructura del diálogo en el corpus de acuerdo a lo expuesto en el apartado 5.2.4, puesto que no hemos utilizado ningún mecanismo de detección automática para esta estructura.

Para garantizar la validez de los resultados obtenidos, la anotación se realizó en paralelo por dos anotadores a los que se les realizó un estudio de fiabilidad por medio del coeficiente *kappa* tal y como se expone en la sección 2.4.2.

La fiabilidad de la anotación de la estructura de pares adyacentes y definición del tópico se midió como una tarea de clasificación de turnos obteniendo un valor de *kappa* de $k = 0,91$, garantizando la utilidad del resultado.

Una versión de la anotación consensuada por ambos anotadores fue usada finalmente.

6.2 Anotación de la correferencia

Para comprobar el rendimiento de ARIADNA es necesario realizar una anotación manual de la correferencia en los casos de anáfora que se tratan para una posterior comparación de resultados. Para la anotación manual de correferencia seguiremos la propuesta de Gaizauskas y Humphreys en (Gaizauskas & Humphreys, 1996) sobre un esquema de anotación de la correferencia mediante SGML (Goldfarb, 1990). De acuerdo a esta anotación,

dado un antecedente A y una anáfora B, donde tanto A como B son cadenas del texto, y además, A es antecedente de B, la forma básica de anotación de dicha correferencia tendría la forma:

```
<COREF ID="valorID1" MIN="valorMIN"> A </COREF>
.
.
.
<COREF ID="valorID2" TYPE="valorTYPE"
REF="valorID1" STATUS="valorSTATUS"> B </COREF>
```

Donde ID tomará generalmente un valor numérico, mientras que MIN, TYPE y STATUS son cadenas alfanuméricas.

El atributo ID tiene como finalidad la identificación de cada una de las cadenas que toman parte de la relación de correferencia. Dicho atributo se podrá definir de forma arbitraria, sin necesidad de comenzar por un número o carácter en particular, ni existe necesidad de mantener una anotación consecutiva. Sin embargo será condición necesaria la unicidad. El atributo REF lo contiene la cadena anafórica indicando el número de identificador de la cadena antecedente. El atributo TYPE indica el tipo de relación existente entre la anáfora y su antecedente. En el etiquetado aplicable a ARIADNA proponemos los siguientes valores: "PRONSUJ", "PRONSP", "COMP" y "ADJ" para los casos de anáfora de sujeto, de sintagma preposicional, de complemento y anáfora adjetiva.

Además se proponen dos atributos MIN y STATUS que pueden ser opcionalmente usados por el anotador para indicar ciertas condiciones de la anotación. El atributo MIN contenido en la anotación del antecedente podría opcionalmente expresar el valor mínimo que sería aceptado para hacer referencia a este antecedente en la evaluación de la correferencia. Un ejemplo del uso de este atributo puede comprobarse en el ejemplo 99.

```
(99) <COREF ID="100" MIN="El talgo de
      las doce"> El talgo de las doce que
      llega a Alicante a las cinco
      </COREF>
```



<COREF ID="101" TYPE="PRONSUJ"
REF="100"> éste </COREF>

En este ejemplo se ha marcado como antecedente el mayor sintagma nominal que representa a la entidad referida (*El talgo de las doce que llega a Alicante a las cinco*). Sin embargo, el sintagma nominal interno *El talgo de las doce* representa igualmente de forma inequívoca a la entidad referencia. Por ello, el anotador puede considerar que cualquier solución al pronombre demostrativo *éste* que incluya al menos la cadena MIN puede ser considerada como acierto.

Por último el atributo STATUS puede ser incluido en la etiqueta de la expresión anafórica con el valor OPT (opcional) para indicar que existe cierta incertidumbre sobre el antecedente real de dicha anáfora aunque la creencia del anotador indica que el apuntado por REF es el más probable.

El proceso de anotación de la correferencia en los 204 diálogos del corpus se realizó paralelamente a la anotación de la estructura del diálogo, y se adoptó el mismo criterio. Dos anotadores se encargaron de realizar toda la anotación, que se redujo de nuevo a una tarea de clasificación consistente en definir la solución adecuada entre las propuestas en la lista de candidatos. En este sentido se estimó que cada anotador tuvo que decidir entre una media de 6,5 candidatos posibles por anáfora después de aplicar las restricciones que descartaban al resto. Posteriormente se llegó a una versión de anotación consensuada por ambos anotadores que fue la que se utilizó finalmente para la experimentación.

El test de fiabilidad mostró en este caso un valor de *kappa* de $k = 0,87$, lo cual garantizó la fiabilidad de la anotación.

6.3 Módulo de evaluación automática de resultados

Para comprobar el rendimiento de ARIADNA se ha empleado un módulo de evaluación automática de resultados. Este módulo recibe la salida de ARIADNA y la compara con el etiquetado manual de la correferencia para proporcionar una estadística de aciertos y fallos. El módulo de evaluación usa un interfaz para convertir el etiquetado anafórico SGML a formato PROLOG. Para ello se define un predicado PROLOG *oSol* que contiene dos parámetros:

Enunciado (arg_1): Contiene el número de enunciado que contiene la anáfora.

Lista de anáforas (arg_2): Contiene una lista PROLOG con un predicado *anaf* por cada anáfora detectada en ese enunciado. El predicado *anaf* contiene 4 parámetros:

Anáfora (arg_1): Es la lista de palabras que conforman la expresión anafórica. Se extrae de la etiqueta COREF de SGML, en concreto del texto encerrado entre las marcas `<COREF . . . >` y `</COREF>`.

Antecedente (arg_2): Es la lista de palabras que conforman la expresión antecedente. Se extrae de la etiqueta COREF de SGML relacionada con la anáfora. Si hay un valor en el argumento MIN se incluye éste, si no se toma el texto encerrado entre las marcas `<COREF . . . >` y `</COREF>`.

Tipo (arg_3): Correspondiente al tipo definido en el argumento TYPE de SGML.

Dirección (arg_4): Que puede adoptar dos valores “<” para indicar que el antecedente se encuentra antes de la anáfora o “>” indicando que el antecedente se encuentra después de la anáfora (catáfora). ARIADNA sólo está preparado para tratar la anáfora, no la catáfora, por tanto se utilizará siempre el primer símbolo.

En el ejemplo 100 se muestra la conversión a PROLOG del etiquetado SGML mostrado en el ejemplo 99

(100) oSol(11, [anaf([éste], [el, talgo, de, las, doce], pronSuj, <)]).

Una vez efectuada la conversión de SGML a PROLOG del etiquetado anafórico, el módulo de evaluación automática compara la salida de ARIADNA con la entrada anotada y proporciona datos de precisión y cobertura del sistema.

6.3.1 Precisión

$$P = \frac{N}{T}$$

La precisión (P) se calcula como el cociente entre el número de anáforas correctamente resueltas por el sistema (N) y el número total de anáforas tratadas (T). La precisión proporciona una medida de la capacidad del sistema para resolver correctamente las anáforas que ha detectado.

6.3.2 Cobertura

$$C = \frac{T}{R}$$

La cobertura (C) se calcula como el cociente entre el número de anáforas detectadas por el sistema (T) y el número real de anáforas existentes (R). La cobertura proporciona una medida de la capacidad del sistema para detectar los casos de anáfora que se van a tratar.

6.4 Entrenamiento del sistema de preferencias

El conjunto de preferencias mostrado en la sección ha sido generado como consecuencia del estudio que se muestra a continuación, en el que tomando como punto de partida el sistema de restricciones y preferencias original planteado por Ferrández *et al.* (1999) se han realizado diversas adaptaciones para su aplicación en diálogos hasta obtener la configuración óptima. Estas adaptaciones se han realizado en dos fases distintas correspondientes a dos versiones

distintas de ARIADNA. La diferencia básica entre ambas versiones ha estado dirigida por diferentes métodos para el tratamiento de las preferencias como se mostró en la sección 5.3.4.

Así, mientras la primera versión de ARIADNA se entrenó usando un tratamiento ordenado de preferencias, en la segunda se define un tratamiento ponderado que admite mayor flexibilidad en el entrenamiento. En concreto, la primera versión de ARIADNA se usó para definir el conjunto de preferencias a usar, ya que el tratamiento ordenado permitía distinguir fácilmente aquellas que eran útiles y las que no. La segunda se empleó para definir el grado de importancia de cada una de las preferencias.

Para el entrenamiento de ambas versiones se reservó una parte del corpus formada por 40 diálogos. Estos 40 diálogos fueron los mismos empleados para el estudio del espacio de accesibilidad que se mostró en la sección 4.3, de los cuales 5 fueron empleados para el entrenamiento de los anotadores, y 35 para obtener el entrenamiento del sistema.

A continuación se muestran los resultados de ambos entrenamientos sobre los 35 diálogos.

6.4.1 Entrenamiento en ARIADNA v.1

El sistema de restricciones y preferencias original de Ferrández *et al.* (1999) se define como un sistema que integra el conocimiento lingüístico proveniente de varias fuentes de información de acuerdo a la estructura definida en la figura 5.1.

Para este entrenamiento se han definido varios experimentos que implican cambios en el conjunto de preferencias hasta obtener la precisión óptima.

Experimento 0 (*baseline*): Uso de información lingüística únicamente. Tomamos como punto de partida el sistema de preferencias usado por el algoritmo descrito por Ferrández *et al.* (1999). Este algoritmo se basa únicamente en información lingüística, y sus resultados fueron evaluados sobre corpus de discurso monólogo obteniendo como resultado una precisión del 82% para la resolución de la anáfora pronominal. Puesto que esta primera aproximación carecía de información estructural, se adapta

el espacio de accesibilidad anafórica Ferrández *et al.* basado en ventanas de oraciones a los diálogos. El sistema de restricciones usado es el original de Ferrández *et al.* El sistema completo del experimento se define a continuación:

Espacio de accesibilidad anafórica. Para la resolución de la anáfora pronominal, el espacio considerado consistía en los tres enunciados anteriores a la anáfora. Para la anáfora adjetiva, el espacio considerado fue de cuatro enunciados anteriores. Por tanto, se tomarán como candidatos todos los sintagmas nominales encontrados en este espacio.

Restricciones. Se definen las siguientes restricciones según lo mostrado en la sección 5.3.3:

- En el caso de la anáfora pronominal las restricciones incluían:
 - R1p.** *Restricciones morfológicas:* concordancia morfológica (en género, número y persona)
 - R2p.** *Restricciones sintácticas:* condiciones sintácticas de los pronombres
- En el caso de la anáfora adjetiva, las restricciones incluían:
 - R1a.** *Restricciones morfológicas:* concordancia morfológica (en género)
 - R2a.** *Restricciones sintácticas:* concordancia sintáctica de la anáfora adjetiva

Preferencias.

- En el caso de la anáfora pronominal, las preferencias (originales de Ferrández *et al.*) fueron en orden de mayor a menor nivel, es decir, de mayor a menor importancia:
 - L1p** (*1ª*): candidatos en el mismo enunciado que la anáfora,
 - L2p** (*2ª*): candidatos en el enunciado anterior a la anáfora,
 - L3p** (*3ª*): candidatos que son nombres propios o sintagmas nominales indefinidos,
 - L4p** (*4ª*): candidatos que son nombres propios (para pronombres personales),
 - L5p** (*5ª*): candidatos que se han repetido más de una vez,

- L6p** (6^a): candidatos que han aparecido con el mismo verbo de la anáfora más de una vez,
- L7p** (7^a): candidatos que tienen la misma posición respecto al verbo de la anáfora (antes o después),
- L8p** (8^a): candidatos que están en el mismo constituyente sintáctico (tienen el mismo número de constituyente analizado que la anáfora),
- L9p** (9^a): candidatos que no están en complementos circunstanciales,
- L10p** (10^a): candidatos que se repiten más en el texto,
- L11p** (11^a): candidatos que aparecen con mayor frecuencia con el mismo verbo de la anáfora,
- L12p** (12^a): el candidato más cercano a la anáfora.
- En el caso de la anáfora adjetiva, las preferencias fueron para:

L1a (1^a): candidatos en el mismo enunciado que la anáfora,

L2a (2^a): candidatos en el enunciado anterior a la anáfora,

L3a (3^a): candidatos que comparten la misma clase de modificador que la anáfora (por ejemplo, un sintagma preposicional),

L4a (4^a): candidatos que comparten el mismo modificador que la anáfora (por ejemplo, el mismo adjetivo 'rápido'),

L5a (5^a): candidatos que concuerdan en número,

L6a (6^a): candidatos que se repiten más en el texto,

L7a (7^a): candidatos que aparecen más con el mismo verbo de la anáfora,

L8a (8^a): el candidato más cercano a la anáfora.

Discusión. Tras la primera configuración, se efectuó una primera evaluación en la que se obtuvo una precisión del 59,0% para la resolución de la anáfora pronominal y del 23,7% para la anáfora adjetiva. Como se puede apreciar, estos resultados fueron bajos para el caso pronominal y extremadamente bajos para la anáfora adjetiva. Una vez que se analizaron los fallos, se alcanzaron las siguientes conclusiones: el espacio de accesibilidad anafórica definido estaba demasiado restringido, y además había sido definido de forma arbitraria sin tener en cuenta la relación existente entre la anáfora y la estructura del diálogo. Consecuentemente, se proponen los cambios del experimento 1.

Experimento 1: Uso de información de la estructura del diálogo únicamente. Para este segundo experimento se modifica la definición del espacio de accesibilidad anafórica basado en una ventana de tamaño fijo por el espacio de accesibilidad anafórica estructural definido en la sección 4.2. Además se modifican las preferencias afectadas por esta definición.

Espacio de accesibilidad anafórica estructural. Para definir este espacio se considera el par adyacente y el tópico del diálogo. Concretamente, se tomarán todos los sintagmas nominales que se encuentren en el mismo par adyacente de la anáfora, en el par anterior, en todos los pares que contengan al de la anáfora y el tópico principal del diálogo.

Preferencias. Para este experimento se eliminaron las preferencias de la anáfora pronominal L3p a L11p descritas en el experimento anterior así como las preferencias de la anáfora adjetiva L3a a L7a (todas las preferencias lingüísticas). Así mismo, las preferencias L1p y L2p y L1a y L2a relacionadas con el espacio de accesibilidad se sustituyeron por cuatro nuevas preferencias estructurales E1 a E4. El sistema de preferencias para la anáfora pronominal queda con el siguiente orden:

- E1** (1ª): candidatos que se encuentran en el mismo par adyacente que la anáfora,
- E2** (2ª): candidatos que se encuentran en el par adyacente anterior a la anáfora,
- E3** (3ª): candidatos que se encuentran en algún par adyacente que contiene al de la anáfora,
- E4** (4ª): candidatos que están en el tópico
- L12p** (5ª): candidato más cercano

En el caso de la anáfora adjetiva se obtiene una ordenación paralela:

- E1** (1ª): candidatos que se encuentran en el mismo par adyacente que la anáfora,
- E2** (2ª): candidatos que se encuentran en el par adyacente anterior a la anáfora,

E3 (3ª): candidatos que se encuentran en algún par adyacente que contiene al de la anáfora,

E4 (4ª): candidatos que están en el tópico

L8a (5ª): candidato más cercano

Este cambio se efectuó para evaluar el rendimiento del sistema cuando la información lingüística se eliminaba y únicamente se incluye información estructural, aunque se mantienen las preferencias L12p para la anáfora pronominal y L8a para la adjetiva (el candidato más cercano) para garantizar una única solución al final del proceso.

Discusión. Después de incluir la información sobre la estructura del diálogo y eliminar las preferencias lingüísticas, las tasas de precisión fueron del 62,3% para la resolución de la anáfora pronominal y 65,8% para la anáfora adjetiva. En este caso se obtuvo un importante crecimiento en la resolución de la anáfora adjetiva con un único cambio en el espacio de accesibilidad anafórica. Evidentemente, esto demuestra que la anáfora adjetiva necesita un espacio de accesibilidad mucho mayor del definido mediante la ventana de oraciones en el experimento 0.

Pero estos resultados continúan siendo bajos, y demuestran que usar únicamente la información sobre la estructura del diálogo no es suficiente. De esta forma, un tercer experimento se lleva a cabo usando ambos tipos de información, la estructura del diálogo y la información lingüística. Este experimento resume varios experimentos que se llevaron a cabo hasta obtener el resultado óptimo mostrado en él.

Experimento 2: La información lingüística junto a la información sobre la estructura del diálogo.

Preferencias. Para este experimento se partió de un conjunto de preferencias formado por todas las preferencias lingüísticas y estructurales definidas anteriormente. Así, tras varias configuraciones, se obtuvo la óptima, la cual se presenta a continuación:

- Para la anáfora pronominal se preferirán en este orden:
 - preferencias estructurales (experimento 1): E1 a E4

- preferencias lingüísticas (experimento 0): L6p, L7p, L8p y L12p
- Del mismo modo, para la anáfora adjetiva:
 - preferencias estructurales (experimento 1): E1 a E4
 - preferencias lingüísticas (experimento 0): L3a, L4a, L5a y L8a

Estos cambios finales proporcionaron la mejor configuración para el sistema de preferencias lingüísticas y el sistema de preferencias estructurales de ARIADNA v.1.

Discusión y justificación. Generalmente, la información sobre los candidatos más repetidos se usa en los sistemas de preferencias con el fin de obtener un conocimiento sobre las entidades fundamentales del diálogo (que serán las más fácilmente referenciables). Sin embargo, en este experimento, se ha incluido ya la información sobre el tópico que sustituye este conocimiento, por este motivo, las preferencias L5p, L10p y L11p de la anáfora pronominal y las L6a y L7a de la anáfora adjetiva no son necesarias y su eliminación evita resultados erróneos.

Otra evaluación que eliminó la preferencia pronominal L3p (candidatos que son nombres propios o sintagmas nominales indefinidos) obtuvo mejores resultados. En este caso, no hay una clara justificación para su uso, y por tanto, la preferencia se eliminó definitivamente.

La preferencia L4p para la anáfora pronominal (preferencia por los nombres propios, si es un pronombre personal) también se eliminó mejorando los resultados para la anáfora. Después de estudiar el caso, encontramos que la preferencia por los nombres propios causa errores debido a que en el uso de nombres de lugar (tratados como propios) esta preferencia se incumple.

Tras eliminar la preferencia número L9p (candidatos que no están en complemento circunstancial), la precisión para la anáfora pronominal permaneció en el mismo valor. De nuevo se decide eliminar por su falta de justificación.

De esta forma, tras agotar todas las posibilidades, el mostrado anteriormente se consideró el conjunto mínimo de preferencias que obtenía la mayor precisión, y por tanto se consideró la configuración óptima para ARIADNA v.1. Al finalizar se obtuvo una

precisión del 73,8% para la anáfora pronominal y 78,9% para la adjetiva, finalizando el entrenamiento de ARIADNA v.1.

Las distintas ordenaciones estudiadas así como las precisiones obtenidas se resumen en las tablas 6.1 y 6.2².

Anáfora Pronominal				
Preferencias		Experimento		
Tipo	Pref.	0	1	2
Estructurales	E1	-	1 ^a	1 ^a
	E2	-	2 ^a	2 ^a
	E3	-	3 ^a	3 ^a
	E4	-	4 ^a	4 ^a
Lingüísticas	L1p	1 ^a	-	-
	L2p	2 ^a	-	-
	L3p	3 ^a	-	-
	L4p	4 ^a	-	-
	L5p	5 ^a	-	-
	L6p	6 ^a	-	5 ^a
	L7p	7 ^a	-	6 ^a
	L8p	8 ^a	-	7 ^a
	L9p	9 ^a	-	-
	L10p	10 ^a	-	-
	L11p	11 ^a	-	-
	L12p	12 ^a	5 ^a	8 ^a
Precisión (%)		59,0	62,3	73,8

Tabla 6.1. Ordenación de preferencias durante el entrenamiento de ARIADNA v.1. Anáfora Pronominal

6.4.2 Entrenamiento en ARIADNA v.2

Una vez definido el conjunto de preferencias considerado idóneo para la resolución de la anáfora en los diálogos, se incorpora a ARIADNA el sistema de preferencias modificado para permitir un tratamiento ponderado de las preferencias obteniendo la versión v.2. Esto supone la asignación de pesos a cada una de las preferencias consideradas, de tal forma que un entrenamiento basado en la asignación de diferentes pesos en cada una de las preferencias, permitirá un ajuste fino del sistema. Como punto de

² No se ha proporcionado información acerca de la cobertura obtenida puesto que este parámetro no ha sido objetivo del proceso de entrenamiento.

Anáfora Adjetiva				
Preferencias		Experimento		
Tipo	Pref.	0	1	2
Estructurales	E1	-	1 ^a	1 ^a
	E2	-	2 ^a	2 ^a
	E3	-	3 ^a	3 ^a
	E4	-	4 ^a	4 ^a
Lingüísticas	L1a	1 ^a	-	-
	L2a	2 ^a	-	-
	L3a	3 ^a	-	5 ^a
	L4a	4 ^a	-	6 ^a
	L5a	5 ^a	-	7 ^a
	L6a	6 ^a	-	-
	L7a	7 ^a	-	-
	L8a	8 ^a	5 ^a	8 ^a
Precisión (%)		23,7	65,8	78,9

Tabla 6.2. Ordenación de preferencias durante el entrenamiento de ARIADNA v.1. Anáfora Adjetiva

partida se toma por tanto un conjunto de pesos basado en un estudio empírico de los casos tratados en el sistema ARIADNA v.1. De acuerdo a este estudio se realiza el primer experimento con preferencias ponderadas que hemos numerado como experimento 3.

Experimento 3: Preferencias ponderadas.

Preferencias. En este experimento, se toman las mismas preferencias obtenidas en el experimento 2 añadiéndoles un peso a cada una de ellas, excepto a la última, la preferencia por el candidato más cercano que se mantiene como característica de desambiguación para el caso en el que los candidatos empaten en peso.

Para la determinación de los pesos se ha realizado un estudio sobre los resultados del experimento 2 estimando las frecuencias de intervención de cada preferencia en la obtención del resultado correcto. Así se han determinado los siguientes pesos según la importancia estimada para cada uno de ellos:

- Anáfora pronominal:

E1 (*peso=35*): candidatos que se encuentran en el mismo par adyacente que la anáfora,

E2 (*peso=20*): candidatos que se encuentran en el par adyacente anterior a la anáfora,

E3 (*peso=30*): candidatos que se encuentran en algún par adyacente que contiene al de la anáfora,

E4 (*peso=15*): candidatos que están en el tópico

L6p (*peso=5*): candidatos que han aparecido con el mismo verbo de la anáfora más de una vez,

L7p (*peso=5*): candidatos que tienen la misma posición respecto al verbo de la anáfora (antes o después),

L8p (*peso=5*): candidatos que están en el mismo constituyente sintáctico (tienen el mismo número de constituyente analizado que la anáfora),

L12p (*condición de desempate*): el candidato más cercano a la anáfora.

- Anáfora adjetiva:

E1 (*peso=35*): candidatos que se encuentran en el mismo par adyacente que la anáfora,

E2 (*peso=20*): candidatos que se encuentran en el par adyacente anterior a la anáfora,

E3 (*peso=30*): candidatos que se encuentran en algún par adyacente que contiene al de la anáfora,

E4 (*peso=15*): candidatos que están en el tópico

L3a (*peso=5*): candidatos que comparten la misma clase de modificador que la anáfora (por ejemplo, un sintagma preposicional),

L4a (*peso=5*): candidatos que comparten el mismo modificador que la anáfora (por ejemplo, el mismo adjetivo 'rápido'),

L5a (*peso=5*): candidatos que concuerdan en número,

L8a (*condición de desempate*): el candidato más cercano a la anáfora.

Discusión y justificación. Las preferencias E1 a E4 en ambos casos de anáfora se han determinado fundamentales para la elección del candidato correcto. El estudio sobre el espacio de accesibilidad anafórica presentado en la sección 4.3 muestra una distribución de anáforas encontradas en cada uno de estos espacios. Este estudio

se toma como base para la estimación de los pesos de las preferencias estructurales en el experimento 3. Así, puesto que la anáfora más probable ocurre dentro del par adyacente, se le asigna un peso de 35 a la preferencia E1. La siguiente más probable encuentra su antecedente en el par adyacente anterior, asignándose un peso de 20. El caso de las anáforas que encuentran su antecedente en algún par adyacente que contiene al de la anáfora no es demasiado habitual. Sin embargo, esto es debido a que la estructura anidada tampoco está presente en todos los casos. Así se ha estimado que si E1 no resuelve correctamente y hay estructura anidada, existe una gran probabilidad de que resuelva correctamente E3. Así se le asigna un peso intermedio entre E1 y E2 con valor 30. Para el antecedente que se encuentra en el tópico se le asigna un valor 15. Además se asigna un valor igual a 5 a cada una de las preferencias lingüísticas, ya que se estima que tienen valor especialmente en caso de producirse un empate con el sumatorio de los pesos estructurales. Un valor pequeño idéntico para cada preferencia lingüística se estima adecuado para desempatar.

Con estos pesos se evalúa la precisión del sistema obteniendo un 61,62% de anáforas pronominales correctamente resueltas y 76,32% en el caso de anáforas adjetivas. Estos valores disminuyen considerablemente los resultados obtenidos en el experimento 2 con preferencias ordenadas. Sin embargo, el valor de importancia asignado a cada preferencia fue considerado en conjunto y sin tener en cuenta las características de cada tipo de anáfora. Así se estimó que los resultados podrían incrementarse de forma notable al realizar un estudio de pesos adecuado para cada tipo de anáfora devuelta por el módulo Disparador Anafórico.

Experimento 4: Preferencias ponderadas con separación por tipos. El experimento 4 (definitivo) se realiza mediante el estudio independiente de los pesos asignados a cada tipo de anáfora distinguiendo entre adjetivas, pronominales personales con función de sujeto, pronominales demostrativas con función de sujeto,

pronominales con función de complemento, y pronominales en sintagmas preposicionales³.

Preferencias. Tras este experimento se determinan los siguientes cambios en la ponderación de las preferencias.

- Anáfora pronominal:

E1 : se mantiene el valor de peso igual a 35 en todos los casos pronominales,

E2 : se mantiene el valor de peso igual a 20 en la anáfora con función de sujeto, aunque se le asigna un valor 0 para la que tiene función de complemento y 35 para la de sintagma preposicional,

E3 : se mantiene el valor de peso igual a 30 en la anáfora con función de sujeto, aunque se le asigna un valor 0 para la que tiene función de complemento y 20 para la de sintagma preposicional,

E4 : se mantiene el valor de peso igual a 15 en todos los casos pronominales,

L5p : se recupera la preferencia por los candidatos que se han repetido alguna vez únicamente para el caso de la anáfora pronominal de sintagma preposicional y se le asigna valor igual a 5,

L6p : sólo se valora y se hace con peso igual a 5 en el caso de los pronombres demostrativos de sujeto,

L7p : sólo se valora y se hace con peso igual a 5 en el caso de los pronombres de complemento,

L8p : sólo se valora y se hace con peso igual a 5 en el caso de los pronombres de complemento,

L12p: se mantiene la condición de desempate.

- Anáfora adjetiva:

E1 : se mantiene con el mismo peso igual a 35,

E2 : se disminuye el valor del peso a 10,

E3 : se disminuye el valor del peso a 10,

E4 : se incrementa el valor del peso a 35,

³ Aunque se estudió la influencia de los pronombres personales y demostrativos por separado en los sintagmas preposicionales se determinó que el comportamiento en ambos casos era idéntico

L3a : se mantiene con el mismo peso igual a 5,

L4a : se mantiene con el mismo peso igual a 5,

L5a : se mantiene con el mismo peso igual a 5,

L8a : se mantiene la condición de desempate.

Discusión y justificación. En el caso de la anáfora de complemento, las preferencias E2 y E3 dejan de tener valor puesto que este tipo de anáforas utilizan un espacio de accesibilidad anafórica mucho más corto que el resto de pronominales. Además la preferencia E2 gana importancia en el caso de la anáfora pronominal de sintagma preposicional al incrementar su espacio de accesibilidad y lo pierde la E3 por el mismo motivo. La preferencia L5p que dejó de usarse en el experimento 2, ha sido recuperada ya que al separar los casos de anáfora se puede considerar en el caso de la anáfora de SP sin perjudicar los resultados del resto de pronominales. Igualmente, la L6p, L7p y L8p se aplica sólo en los casos donde resultan efectivas.

Finalmente, en el caso de la anáfora adjetiva que se ha demostrado ser mucho más sensible a la influencia del tópico que la secuencia de pares adyacentes se han reajustado los valores de los pesos para asignarle más peso por tanto al tópico que a otros pares adyacentes.

Como resultado de este entrenamiento se produce el mejor resultado de precisión obtenido con el corpus de entrenamiento. El 80,33% de la anáfora pronominal fue correctamente resuelta, alcanzando una precisión del 92,11% para la anáfora adjetiva.

Este resultado da por finalizado el proceso de entrenamiento, y por tanto, el ajuste de pesos y preferencias para el sistema ARIADNA v.2.

Las tablas 6.3 y 6.4 resumen los experimentos 3 y 4.

6.5 Evaluación de ARIADNA v.2.

Para la evaluación del sistema ARIADNA se tomaron los 164 diálogos del corpus Infotren: Persona que no fueron utilizados en la fase de entrenamiento. El proceso de evaluación se realizó al margen del entrenamiento y en tiempo posterior a éste. De

Anáfora Pronominal						
Preferencias		Experimento				
		3		4		
Tipo	Pref.		Personal Suj	Demostr Suj	Compl	SP
Estructural	E1	35	35	35	35	35
	E2	20	20	20	0	25
	E3	30	30	30	0	20
	E4	15	15	15	15	15
Lingüística	L5p	0	0	0	0	5
	L6p	0	0	5	0	0
	L7p	5	0	0	5	0
	L8p	5	0	0	5	0
	L12p	SI	SI	SI	SI	SI
Precisión (%)			74,07		86,67	75,00
		61,62	80,33			

Tabla 6.3. Ponderación de preferencias durante el entrenamiento de ARIADNA v.2. Anáfora Pronominal

Anáfora Adjetiva			
Preferencias		Experimento	
Tipo	Pref.	3	4
Estructural	E1	35	35
	E2	20	10
	E3	30	10
	E4	15	35
Lingüística	L3a	5	5
	L4a	5	5
	L5a	5	5
	L8a	SI	SI
Precisión (%)		76,32	92,11

Tabla 6.4. Ponderación de preferencias durante el entrenamiento de ARIADNA v.2. Anáfora Adjetiva

esta forma se garantiza que el entrenamiento no ha influido en el resultado final de la evaluación.

Tras el proceso de evaluación se obtienen los resultados mostrados en la tabla 6.5.

Este proceso de evaluación se realizó en dos fases. En una primera fase se realiza un proceso de evaluación automática en la

	Existentes	Tratadas	Aciertos	Fallos	P(%)	C(%)
Pron. Suj.	124	101	78	23	77,23	62,90
Pron. Compl.	106	104	88	16	84,62	83,02
Pron. SP.	16	14	12	2	85,71	75,00
Pronominal	246	219	178	41	81,28	72,36
Adjetiva	146	146	119	27	81,51	81,51
TOTAL	392	365	297	68	81,37	75,77

Tabla 6.5. Evaluación de ARIADNA v.2.

que el sistema obtiene resultados sin intervención alguna del analista. En una segunda fase de depuración, el analista elimina manualmente los casos de anáfora que no son objetivo de nuestro estudio y que el sistema no puede descartar automáticamente, es decir, las exóforas y catáforas, además de las anáforas de segunda persona que se encuentran camufladas en terceras personas debido al tratamiento de cortesía “usted”. Además se realiza una monitorización de cada resultado obtenido con el fin de evitar interferencias provocadas por fallos en el etiquetado manual de la estructura del diálogo y de las soluciones, así como los fallos provocados por el etiquetado léxico del corpus. Por último se realiza una revisión manual de las anáforas que deberían haber sido tratadas por el sistema pero que por diferentes causas no han podido ser analizadas. Estas últimas se sumarán a las anáforas tratadas para obtener el total de anáforas existentes con el fin de calcular la cobertura del sistema.

Finalmente, tras la fase de revisión manual, se encontró un total de 392 anáforas existentes del tipo que constituye el objeto de nuestro estudio, es decir, anáforas individuales de tipo pronominal y adjetiva. Estas anáforas se distribuyen de la siguiente forma: 365 fueron tratadas automáticamente por el sistema, de las cuales 219 correspondían a anáforas pronominales y 146 a anáforas adjetivas. De las pronominales, 101 correspondían a anáforas de sujeto, 104 a anáforas de complemento y 14 a anáforas de sintagma preposicional. Como resultado se obtuvo una precisión de 77,23% para las anáforas pronominales de sujeto, una precisión de 84,62% para las anáforas pronominales de complemento y un 85,71% para

las de sintagma preposicional. Ello conforma una precisión total de 81,28% para la anáfora pronominal. En la anáfora adjetiva se obtuvo una precisión del 81,51%. Además se obtuvo una cobertura del 72,36% en el caso de la anáfora pronominal y un 81,51% para la anáfora adjetiva. Todo ello conforma una precisión total del sistema de 81,37%, y una cobertura total de 75,77%.

6.5.1 Discusión y consideraciones sobre los resultados

Algunas de las anáforas existentes no pudieron ser tratadas por el sistema. En la mayoría de los casos este hecho es debido a que sus antecedentes también tienen origen anafórico o contienen algún tipo de elipsis y por tanto deben ser tratados previamente a ser considerados. Estas formas no son consideradas por el analizador como sintagmas nominales y, por tanto, no las considera como antecedente. Aunque en algunos casos el sistema devolvía una solución para estas anáforas, el resultado fue eliminado manualmente para evitar interferir con los resultados del sistema.

Se ha realizado un estudio de casos para los fallos proporcionados por el sistema con estimaciones de sus causas que se resume en la tabla 6.6.

CAUSA DE FALLO					
	Semántica	Acumulación	Fallo EAA	Ambigüedad	TOTAL
Casos	37	10	8	12	67
%	55,22	14,93	11,94	17,91	100

Tabla 6.6. Análisis de errores en la evaluación de ARIADNA v.2.

De acuerdo a este estudio, los errores se estiman provocados por las siguientes causas:

- Falta de información semántica dependiente del dominio: se ha estimado que el 55,22% de la anáfora mal resuelta ha sido debido a la falta de información semántica dependiente del dominio que permita discernir entre candidatos semánticamente compatibles con la expresión anafórica y aquellos que no lo son.

Consideramos en este caso aquellos fallos en el que la información semántica descarta los antecedentes no válidos, y entre los válidos, el antecedente correcto es el que alcanza mejor valoración, y por tanto, sería el propuesto.

- Acumulación de errores: el 14,93% de los fallos se han estimado debidos a errores acumulados, es decir, anáforas mal resueltas que se toman como antecedentes para otras anáforas. Si las primeras hubieran sido detectadas correctamente, se estima que las últimas hubieran encontrado su antecedente correcto.
- Espacio de accesibilidad anafórica inadecuado: se ha estimado que el 11,94% de las anáforas mal resueltas han sido debidas a la consideración de un espacio de accesibilidad incorrecto, lo que supone que el antecedente correcto había sido descartado antes de aplicar restricciones y preferencias. Destacamos que en algunos de los casos podría ser debido a un criterio erróneo por parte del anotador de la estructura.
- Fallos de ambigüedad: incluimos en este último grupo aquellos fallos provocados por la propia ambigüedad del diálogo y de la forma de expresarse los hablantes, que impide determinar exactamente, incluso para un analista manual, cuál es la solución correcta.

Estimamos por tanto, que la inclusión de herramientas semánticas adecuadas (que podrían ser proporcionadas por el sistema de diálogo en el que se implante ARIADNA) proporcionaría resultados de resolución elevados, ya que aproximadamente un 55% de los errores podrían ser evitados. Esto supondría alcanzar una precisión próxima al 92%.

6.6 Conclusiones del capítulo

El entrenamiento y evaluación de ARIADNA han supuesto dos de las tareas fundamentales realizadas durante el proceso de investigación. Ambas han sido realizadas aprovechando los 204 diálogos proporcionados por el corpus Infotren: Persona.

Para el entrenamiento han sido necesarias dos fases distintas. La primera fase sirvió para detectar el conjunto de preferencias

útil para la obtención de la información que permita resolver la anáfora en los diálogos. Esta primera fase se aplicó a la primera versión del sistema ARIADNA que incorporaba un manejo ordenado de preferencias. Este primer entrenamiento proporcionó una precisión del 73,8% para la anáfora pronominal y del 78,9% para la anáfora adjetiva. La segunda fase del entrenamiento tuvo como objetivo principal la determinación de un nivel de importancia adecuado para cada una de las preferencias en cada tipo de anáfora. De esta forma se implementa una segunda versión de ARIADNA que incluye un manejo ponderado de preferencias. Tras este segundo entrenamiento se obtiene un conjunto de pesos adecuados con una precisión del 80,33% para la anáfora pronominal y del 92,11% para la anáfora adjetiva. Para la consecución de ambas fases de entrenamiento se usaron 40 diálogos tomados al azar entre los 204 del corpus.

Finalmente se realizó la evaluación del sistema mediante los 164 diálogos restantes. Como consecuencia se obtuvo un resultado de precisión del 81,28% para la anáfora pronominal y 81,51% para la anáfora adjetiva. Además se obtuvo una cobertura del 72,36% en el caso de la anáfora pronominal y un 81,51% para la anáfora adjetiva. Todo ello conforma una precisión global del sistema de 81,37%, y una cobertura total de 75,77%. Destacamos además que una precisión cercana al 92% podría llegar a alcanzarse mediante la inclusión de información semántica dependiente del dominio.



Universitat d'Alacant
Universidad de Alicante



7. Conclusiones y trabajos futuros

Universitat d'Alacant
Universidad de Alicante

En este trabajo se ha desarrollado un sistema de resolución de la anáfora generada por pronombres (personales y demostrativos) de tercera persona y anáforas adjetivas en los diálogos. El método de resolución de la anáfora propuesto se basa en el uso de conocimiento lingüístico (información léxica, morfológica y sintáctica) y de conocimiento estructural (la estructura del diálogo y el espacio de accesibilidad anafórica estructural que genera), ambos independientes del dominio.

Gracias al uso de información no dependiente del dominio, este sistema de resolución de la anáfora para diálogos al que hemos denominado ARIADNA (*Anaphora Resolution In Automatic Dialogue systems with Natural language Analysis*) está preparado para ser aplicado a sistemas que trabajen sobre cualquier dominio, y de una forma muy especial para los sistemas de diálogo que constituyen su aplicación más inmediata.

Las principales aportaciones de la Tesis han sido:

- La definición de un espacio de accesibilidad anafórica estructural. Para ello, mostramos mediante un estudio empírico, que en los diálogos en español, los antecedentes de la anáfora pronominal y adjetiva se encuentran en un conjunto de sintagmas nominales tomados de este espacio de accesibilidad anafórica. Este espacio está basado en los estudios de Fox (Fox, 1987) sobre las relaciones entre la anáfora y la estructura del diálogo, y en concreto, se define de acuerdo a una estructura basada en los pares adyacentes y el tópico del diálogo. Este espacio ha sido evaluado y se ha obtenido que el 96,4% de los antecedentes se encuentran en él, optimizando considerablemente el proceso de

resolución de la anáfora al reducir el conjunto de los candidatos a antecedente.

- La propuesta y definición de un esquema de anotación para diálogos en español con el objetivo de obtener el espacio de accesibilidad anafórica. Dicho esquema de anotación sigue las directrices indicadas por los sistemas de diálogo estudiados, lo que permite la correcta integración de la propuesta en un sistema de diálogo. Además se ha presentado una propuesta de anotación estándar del esquema mediante el *Standard Generalized Markup Language* (SGML) que permitirá su realización tanto de una manera semi-automática mediante herramientas de edición de textos SGML, como automática mediante herramientas definidas para la detección de la estructura del diálogo como, por ejemplo, las usadas por los sistemas de diálogo. En este último caso, SGML estandariza el interfaz con el sistema de diálogos.
- La definición de un conjunto de restricciones morfológicas y sintácticas que definen la no correferencialidad entre la anáfora y los candidatos a antecedente. Estas restricciones actúan como filtro lingüístico descartando los candidatos a antecedente.
- La definición de un conjunto de preferencias lingüísticas basadas en la estructura sintáctica parcial generada por el analizador sintáctico, además de un conjunto de preferencias estructurales basados en la propia estructura del diálogo.
- El desarrollo de un algoritmo de resolución de la anáfora basado en los conjuntos de restricciones y preferencias lingüísticas y estructurales definidos. Este algoritmo aplica el conjunto de restricciones como un filtro lingüístico que descarta los posibles antecedentes y propone el conjunto de preferencias desde dos puntos de vista: el primero de ellos como un conjunto de preferencias ordenadas, en el que los antecedentes se van seleccionando conforme van cumpliendo una serie de preferencias según un orden definido establecido previamente por criterios empíricos, y por otra parte, como un conjunto de preferencias ponderadas, en el que cada preferencia tiene un peso que define su importancia en la decisión final. En este último caso el mejor

antecedente se selecciona al calcular el sumatorio de todos los pesos obtenidos por las preferencias cumplidas.

- La construcción del sistema de resolución de la anáfora al que hemos denominado ARIADNA que integra el algoritmo de resolución de la anáfora y el espacio de accesibilidad anafórica estructural. La implementación del sistema se ha realizado en el lenguaje de programación lógico Prolog (LPA Prolog).
- La definición de un conjunto de herramientas interfaz que integran la información que proviene de las distintas fuentes de conocimiento usadas por ARIADNA con el módulo de resolución anafórica, es decir, el núcleo de ARIADNA. La existencia de estas herramientas proporciona una total independencia y facilidad de integración con cualquier sistema de diálogo o aplicación final a la que se destine ARIADNA.
- El entrenamiento del sistema de resolución de la anáfora con el objeto de obtener de un modo empírico y cuantitativo el mejor conjunto de preferencias a través de un corpus de 40 diálogos extraídos del corpus Infortren:Persona proporcionados por el proyecto BASURDE (Proyecto BASURDE, 1998-2001). Para ello se utilizó la doble implementación de los conjuntos de preferencias lingüísticas y estructurales: preferencias ordenadas y preferencias ponderadas. La primera implementación facilitaba la elección del conjunto de preferencias a usar, y la segunda facilitaba el ajuste de los valores de importancias asignados a cada una de las preferencias. Dicho entrenamiento se llevó a cabo utilizando un corpus independiente del corpus de evaluación final. Sobre los dos posibles sistemas de preferencias se realizó un estudio de comparación en el que se obtuvieron mejores resultados en el segundo caso: la aplicación de preferencias ponderadas.
- La evaluación del sistema sobre un corpus independiente formado por 164 diálogos extraídos del corpus Infortren:Persona proporcionados por el proyecto BASURDE (Proyecto BASURDE, 1998-2001). Para realizar la evaluación de un modo independiente y fiable, previamente se realizó un etiquetado de la estructura del diálogo y de la anáfora por medio de dos anotadores. La primera anotación se usó para obtener la información lingüística necesaria, mientras que la segunda fue usada únicamente para

comparar la solución propuesta por el sistema con la solución real interpretada por el anotador.

Sobre la anotación efectuada se aplicó el test de fiabilidad propuesto en Carletta (1996) y Carletta *et al.* (1997). Dicho test proporcionó unos valores *kappa* de fiabilidad del 0,91% para la anotación de la estructura del diálogo y un 0,87% para la anotación de la anáfora. Estos valores garantizaron que los resultados obtenidos no estaban influidos por dicha anotación.

Como resultado se obtuvo que el 81,28% de las anáforas pronominales y el 81,51% de las anáforas adjetivas tratadas fueron resueltas correctamente, proporcionando un valor global del 81,37%.

Como conclusión final indicaremos que en este trabajo se ha presentado un sistema de resolución de la anáfora en diálogos independiente del dominio y para textos no restringidos. Este sistema proporciona unas tasas de éxito del 81,37% y se ha estimado que se podrían alcanzar resultados del 92% si tenemos en cuenta que no se ha utilizado información dependiente del dominio ni información del mundo.

7.1 Trabajos en progreso

Como resultado del trabajo realizado y continuación de éste se muestran diferentes líneas de investigación que se están llevando a cabo o que se prevén para un futuro próximo:

- Estudio y propuesta de la incorporación de información semántica como restricción que permita descartar candidatos a antecedente. Los trabajos realizados hasta la fecha muestran que la incorporación de información semántica en la resolución de la anáfora es exitosa cuando se aplica a dominios concretos. Lo que se propone por tanto es el uso del WordNet español como herramienta léxica por una parte, y por otra, el aprendizaje de patrones semánticos entre sujeto y verbo, y verbo y objeto con el objetivo de extraer conocimiento que permita restringir semánticamente el uso de ciertos candidatos como objeto o sujeto de los verbos al ser substituidos en el lugar de la anáfora

(Palomar *et al.*, 2000). De esta forma, si el candidato no cumple las condiciones impuestas por el verbo, no podrá ser el antecedente de esa anáfora. Esta técnica ya probada en monólogos (Saiz-Noeda & Palomar, 2000) se pretende aplicar a los diálogos.

- Tratamiento y resolución de otros tipos de referencias en el diálogo. Se propone por una parte la resolución de las descripciones definidas. En este sentido, se sabe que las descripciones definidas a diferencia de los pronombres y anáforas adjetivas, contienen mayor información léxica que permiten definir mejor la compatibilidad entre la referencia y el antecedente. Sin embargo, presenta algunas dificultades añadidas como que los pronombres y anáforas adjetivas siempre encuentran su antecedente en español, mientras que las descripciones definidas pueden ser no anafóricas. Además, el espacio de accesibilidad de las descripciones definidas es bastante más largo y complejo que el derivado de un pronombre o de una anáfora adjetiva. Por tanto, a partir de los trabajos presentados por Muñoz *et al.* (Muñoz *et al.*, 1999; Muñoz *et al.*, 2000) se pretende continuar en esta línea aplicándose en la resolución de las descripciones definidas en los diálogos. El empleo de WordNet español nos permitirá el tratamiento de la mayoría de las descripciones definidas. Únicamente quedarán sin resolverse aquellas que necesiten información pragmática (conocimiento del mundo) para su resolución, información que no contemplamos actualmente (o referencias exóforas).

Por otra parte, en cuanto al tratamiento de las expresiones temporales, se trabaja actualmente en un método para el reconocimiento y resolución de las referencias temporales que se producen en el discurso. Este método propone la construcción de un modelo del lenguaje basado en una gramática de reconocimiento y el uso de un diccionario para la conversión de las expresiones en fechas (Saquete & Martínez-Barco, 2000). Actualmente se está trabajando en la adaptación de este método al discurso conversacional.

Además, como trabajo futuro se pretende el tratamiento de las anáforas abstractas, cuyos antecedentes no son sintagmas nominales sino sintagmas verbales, enunciados, intervenciones, e

incluso una secuencia de pares adyacentes. Este tipo de anáforas se produce con bastante frecuencia en los diálogos, bastante más que en los monólogos (Allen *et al.*, 2000) y por tanto deben ser contemplados por los sistemas de resolución de anáfora en diálogos. Este tipo de anáforas se han tratado en otros trabajos por medio de un filtro que de acuerdo con el tipo de verbo que acompaña a la anáfora determina si puede o no tratarse de una anáfora abstracta (Eckert & Strube, 1999b). En caso de resultar un caso de abstracta la lista de candidatos no contendrá sintagmas nominales sino sintagmas verbales, enunciados y pares adyacentes. La definición de un adecuado espacio de accesibilidad anafórica en este caso puede ser decisivo en la obtención de la solución correcta.

- Por último, se está trabajando en el desarrollo de un sistema completo para la resolución de la anáfora que incorpora las técnicas de resolución de anáfora aquí estudiadas, junto con las técnicas basadas en el uso de información semántica restringida, que resuelve anáforas pronominales, adjetivas y referencias provocadas por descripciones definidas, y que es capaz de procesar tanto diálogos como monólogos. Este sistema, al que hemos denominado PHORA, se describe en detalle en Palomar *et al.* (2000).

7.2 Producción científica

Como consecuencia de la labor de investigación llevada a cabo para el desarrollo de esta Tesis, se han producido las siguientes publicaciones relacionadas directamente con la resolución de la anáfora en diálogos.

- Revistas:
 - Palomar, M., Ferrández, A., Moreno, L., Martínez-Barco, P., Peral, J., Saiz-Noeda, M., y Muñoz, M. An Algorithm for Anaphora Resolution in Spanish Texts. *Computational Linguistics, Special Issue on Anaphora Resolution*. 2000. Pendiente de aceptación.

- Martínez-Barco, P., y Palomar, M. Resolución de la anáfora: estructura del diálogo y conocimiento lingüístico. *Procesamiento del Lenguaje Natural*, **26**. 2000. pp 47–52. ISSN: 1135-5948
- Ferrández, A., Palomar, M., Moreno, L., Martínez-Barco, P., Peral, J., Muñoz, R., y Saiz-Noeda, M. Sistema de procesamiento del lenguaje natural orientado a la resolución de la correferencia lingüística. *Procesamiento del Lenguaje Natural*, **25**. 1999. pp 217–218. ISSN: 1135-5948
- Martínez-Barco, P. Algoritmo de resolución de la anáfora pronominal en diálogos. *Procesamiento del Lenguaje Natural*, **24**. 1999. pp 75–82. ISSN: 1135-5948
- Peral, J., Martínez-Barco, P., Ferrández, A., y Navarro, B. Sistema de adquisición automática de reglas gramaticales. *Procesamiento del Lenguaje Natural*, **23**. 1998. pp 110–117. ISSN: 1135-5948
- Palomar, M., y Martínez-Barco, P. Computational Approach to Anaphora Resolution in Spanish Dialogues. *Journal of Artificial Intelligence Research*. 2000. Pendiente de aceptación.
- Proceedings (Congresos Internacionales):
 - Palomar, M., Saiz-Noeda, M., Muñoz, R., Suárez, A., Martínez-Barco, P., y Montoyo, A. PHORA: A NLP system for Spanish. In: Gelbukh, A. Ed. *Proceedings of the 2nd International conference on Intelligent Text Processing and Computational Linguistics, CICLing-2001*. Lecture Notes in Artificial Intelligence. Mexico City (Mexico). Febrero, 2001. Springer-Verlag. Pendiente de publicación.
 - Palomar, M., Saiz-Noeda, M., Muñoz, R., Suárez, A., y Martínez-Barco, P. PHORA: A system to solve the Anaphora in Spanish. In: *Proceedings of the Third International Conference on Discourse Anaphora and Anaphor Resolution, DAARC2000*. Lancaster (UK). Noviembre, 2000.
 - Martínez-Barco, P.; Palomar, M. An annotation scheme for dialogues applied to anaphora resolution algorithms. In: P. Sojka, I. Kopecek, K. Pala Eds. *TSD 2000: Third International Workshop on TEXT, SPEECH and DIALOGUE*. Lecture

- Notes in Artificial Intelligence, vol. 1902. Brno (Czech Republic). Septiembre, 2000. Springer-Verlag. pp 410–414. ISBN: 3-540-41042-2
- Martínez-Barco, P., y Palomar, M. Empirical study of the anaphoric accessibility space in Spanish dialogues. *In: Proceedings of the Fourth Workshop on the Semantics and Pragmatics of Dialogue, Götaolog 2000*. Göteborg (Sweden). Junio, 2000. pp 123–128
 - Palomar, M., y Martínez-Barco, P. Anaphora Resolution through Dialogue Adjacency Pairs and Topics. *In: D.N. Christodoulakis Ed. Natural Language Processing - NLP 2000*. Lecture Notes in Artificial Intelligence, vol. 1835. Patras (Greece). Junio, 2000. Springer-Verlag. pp 196–203. ISBN: 3-540-67605-8
 - Martínez-Barco, P., y Palomar, M. Dialogue structure influence over anaphora resolution. *In: O. Cairo, L.E. Sucar and F.J. Cantú Eds. MICAI 2000: Advances in Artificial Intelligence*. Lecture Notes in Artificial Intelligence, vol. 1793. Acapulco (Mexico). Abril, 2000. Springer-Verlag. pp 515–525. ISBN: 3-540-67354-7
 - Martínez-Barco, P. Dialogue structure as a preference in anaphora resolution systems. *In: Proceedings of the International Conference on Artificial and Computational Intelligence For Decision, Control and Automation In Engineering and Industrial Applications, ACIDCA '2000*. Monastir (Tunizia). Marzo, 2000. pp 37–42
 - Martínez-Barco, P., Muñoz, R., Azzam, S., Palomar, M., y Ferrández, A. Evaluation of pronoun resolution algorithm for Spanish dialogues. *In Proceedings of Venezia per il Trattamento Automatico delle Lingue, VEXTAL*. Venice (Italy). Noviembre, 1999. Unipress. pp 325–332. ISBN: 88-8098-112-9
 - Palomar, M., Ferrández, A., Moreno, L., Saiz-Noeda, M., Muñoz, R., Martínez-Barco, P., Peral, J., y Navarro, B. A Robust Partial Parsing Strategy based on the Slot Unification Grammars. *In: Proceedings of the Sixth Conference on Natural Language Processing, TALN'99*. Corsica (France). Julio, 1999. pp 263–272.

- Peral, J., Martínez-Barco, P., Muñoz, R., Ferrández, A., Moreno, L., Palomar, M. Una técnica de análisis parcial sobre textos no restringidos (SUPP) aplicada a un Sistema de Extracción de Información (EXIT). *In: Actas del VI Simposio Internacional de Comunicación Social Santiago de Cuba (Cuba)*. Enero, 1999. Oriente. pp 662–669. ISBN: 959-11-0250-X
- Martínez-Barco, P., Peral, J., Ferrández, A., Moreno, L., y Palomar, M. Analizador Parcial SUPP. *In: Coelho, H. Ed. Progreso en Inteligencia Artificial. Actas do 6º Congresso Iberoamericano de Inteligencia Artificial, IBERAMIA '98*. Lisboa (Portugal). Octubre, 1998. Colibri. pp 329–342. ISBN: 972-772-033-1

Además de las anteriores, se citan a continuación algunas publicaciones de temas relacionados que corresponden a desarrollos iniciales de la Tesis, así como a trabajos futuros.

- Revistas:

- Muñoz, R., Martínez-Barco, P., y Ferrández, A. Método para la resolución de correferencias de sintagmas nominales definidos incluyendo alias y acrónimos en el sistema de extracción de información EXIT. *Procesamiento del Lenguaje Natural*, 25. 1999. pp 143–149. ISSN: 1135-5948
- Ferrández, A., Peral, J., Martínez-Barco, P., Saiz, M., Romero, R. Resolución de la extraposición a izquierdas con las gramáticas de unificación de huecos. *Procesamiento del Lenguaje Natural*, 21. 1997. pp 167–183. ISSN: 1135-5948

- Proceedings (Congresos Internacionales):

- Saquete, E., y Martínez-Barco, P. Grammar Specification for the Recognition of Temporal Expressions. *In: Proceedings of Machine Translation and multilingual applications in the new millennium, MT2000*. Exeter (UK). Noviembre, 2000. pp 21.1–21.7

- Proceedings (Congresos Nacionales):

- Saiz, M., Martínez-Barco, P., y Palomar, M. Paralelismo sintáctico-semántico para el tratamiento de elementos extrapuestos en textos no restringidos. *In: Actas del CAEPIA'97-TTIA'97*. Málaga. Noviembre, 1997. pp 797–803

• Informes Técnicos:

- Llopis, F., Muñoz, R., Suárez, A., Montoyo, A., Palomar, M., Ferrández, A., Peral, J., Martínez-Barco, P., Romero, R., y Saiz, M. Sistema EXIT. *Report Interno - DLSI. Universidad de Alicante*. 1998.
- Moreno, L., Segarra, E., Prieto, N., Palomar, M., Sanchís, E., Pla, F., Molina, A., Ferrández, A., Gómez, J., Suárez, A., Martínez-Barco, P. Construcción de Analizadores Híbridos de Lenguajes Naturales Definidos sobre un Dominio Semántico Restringido. *Report Interno - DLSI. Universidad de Alicante*. 1997.



Universitat d'Alacant
Universidad de Alicante

Referencias

- Ahrenberg, L., Dahlbäck, N., & Jönsson, A. 1995. Coding schemes for studies of natural language dialogue. *Pages 8–13 of: Proceedings of the AAAI Spring Symposium on Empirical Methods in Discourse Interpretation and Generation.*
- Alexandersson, J., Maier, E., & Reithinger, N. 1995. A robust and efficient three-layered dialogue component for a speech-to-speech translation system. *Pages 188–193 of: Proceedings of 7th Conference of the European Chapter of the Association for Computational Linguistics (EACL'95).*
- Allen, J., Schubert, L.K., Ferguson, G., Heeman, P., Hwang, C.H., Kato, T., Light, M., Martin, N.G., Miller, B.W., Poesio, M., & Traum, D.R. 1995. The TRAINS project: A case study in defining a conversational planning agent. *Journal of Experimental and Theoretical AI*, 7, 7–48.
- Allen, J., Byron, D., Dzikovska, M., Ferguson, G., Galescu, L., & Stent, A. 2000. An Architecture for a Generic Dialogue Shell. *Natural Language Engineering*. To appear.
- Asher, N. 1993. *Reference to Abstract Objects*. Dordrecht: Kluwer.
- Aust, H., Oerder, M., Seide, F., & Steinbiss, V. 1995. The Philips automatic train timetable information system. *Speech Communication*, 17, 249–262.
- Baldwin, B. 1997. CogNIAC: high precision coreference with limited knowledge and linguistic resources. *In: Proceedings of ACL/EACL workshop on Operational factors in practical, robust anaphora resolution.*
- Bernsen, N.O., & Dybkjaer, L. 1995. The DISC Concerted Action. *Pages 35–42 of: Gaizauskas, R. (ed), Proceedings of the*

- SALT Club Workshop on Evaluation in Speech and Language Technology.*
- Bernsen, N.O., Dybkjaer, H., & Dybkjaer, L. 1998. *Designing interactive Speech Systems: From first Ideas to User Testing.* Berlin/New York: Springer.
- Bilange, E. 1991. A task independent oral dialogue model. *In: Proceedings of Fifth Conference of the European Chapter of the Association for Computational Linguistics (EACL'91).*
- Brennan, S.E., Friedman, M.W., & Pollard, C.J. 1987 (July). A centering approach to pronouns. *Pages 155-162 of: Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics.*
- Brown, G., & Yule, G. 1983. *Discourse Analysis.* Cambridge: Cambridge University Press.
- Byron, D.K., & Stent, A. 1998. A Preliminary Model of Centering in Dialog. *Pages 1475-1477 of: Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics (COLING-ACL'98).*
- Carbonell, J.G., & Brown, R.D. 1988. Anaphora resolution: a multi-strategy approach. *Pages 96-101 of: Proceedings of 12th International Conference on Computational Linguistics (COLING'88).*
- Carletta, J. 1996. Assessing agreement on classification task: the kappa statistic. *Computational Linguistics*, **22**(2), 249-254.
- Carletta, J., Isard, A., Isard, S., Kowtko, J.C., Doherty-Sneddon, G., & Anderson, A.H. 1997. The Reliability of a Dialogue Structure Coding Scheme. *Computational Linguistics*, **23**(1), 13-32.
- Carmona, J., Cervell, S., Márquez, L., Martí, M.A., Padró, L., nad H. Rodríguez, R. Placer, Taulé, M., & Turmo, J. 1998. An Environment for Morphosyntactic Processing of Unrestricted Spanish text. *In: Proceedings of First International Conference on Language Resources and Evaluation (LREC'98).*
- Casares, J. 1994. *Diccionario ideológico de la lengua española.* Barcelona: Gustavo Gili.

- Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris Publications.
- Clark, H.H. 1985. Language use and language users. In: Lindzey, G., & Aronson, E. (eds), *The Handbook of Social Psychology*, 3rd edn. Erlbaum.
- Clark, H.H., & Schaefer, E.F. 1989. Contributing to discourse. *Cognitive Science*, **13**, 259–294.
- Cohen, P.R. 1984. The pragmatics of referring and the modality of communication. *Computational Linguistics*, **10**, 97–146.
- Connolly, D., Burger, J., & Day, D. 1994. A Machine learning approach to anaphoric reference. *Pages 255–261 of: Proceedings of the International Conference on New Methods in Language Processing, NEMLAP*.
- Coulthard, M. 1977. *An introduction to discourse analysis*. London: Longman.
- Dagan, I., & Itai, A. 1990. Automatic processing of large corpora for the resolution of anaphora references. In: *Proceedings of 13th International Conference on Computational Linguistics (COLING'90)*.
- Dahlbäck, N. 1991. *Representations of Discourse-Cognitive and Computational Aspects*. Ph.D. thesis, Department of Computer and Information Science, Linköping University, Linköping, Sweden.
- Dahlbäck, N. 1997. Towards a Dialogue Taxonomy. In: Maier, Elisabeth, Mast, Marion, & LuperFoy, Susann (eds), *Dialogue Processing in Spoken Language Systems*. Lecture Notes in Artificial Intelligence, vol. 1236. Springer Verlag.
- Eckert, M., & Strube, M. 1999a. Dialogue Acts, Synchronising Units and Anaphora Resolution. In: *Proceedings of Amsterdam Workshop on the Semantics and Pragmatics of Dialogue (AMSTELOGUE'99)*.
- Eckert, M., & Strube, M. 1999b. Resolving Discourse Deictic Anaphora in Dialogues. *Pages 37–44 of: Proceedings of 9th Conference of the European Chapter of the Association for Computational Linguistics (EACL'99)*.
- Ersan, E., & Akman, V. 1994. *Focusing for pronoun resolution in english discourse*. Tech. rept. BU-CEIS-94-29. Bilkent Uni-

- versity. Department of Computer Engineering and Information Science.
- Ferguson, G., & Allen, J.F. 1998. TRIPS: An integrated intelligent problem-solving assistant. *Pages 567-573 of: Proceedings of the Fifteenth National Conference on AI (AAAI'98)*.
- Fernández, M.G. 1999. Delfos: Un modelo basado en unificación para la representación y el razonamiento en sistemas de gestión de diálogo. *Procesamiento del Lenguaje Natural*, **25**, 67-74.
- Fernández, M.G. 2000. *Un modelo para la especificación lingüística y la gestión computacional en diálogos hombre-máquina mediante instrucciones expresadas en lenguaje natural*. Ph.D. thesis, Universidad de Sevilla, Departamento de Filología Inglesa. Facultad de Filología, Sevilla.
- Ferrández, A. 1998. *Aproximación computacional al tratamiento de la anáfora pronominal y de tipo adjetivo mediante gramáticas de unificación de huecos*. Ph.D. thesis, Universidad de Alicante, Departamento de Lenguajes y Sistemas Informáticos.
- Ferrández, A., Palomar, M., & Moreno, L. 1997. Slot Unification Grammar. *Pages 523-532 of: Proceedings of the Joint Conference on Declarative Programming. APPIA-GULP-PRODE*.
- Ferrández, A., Palomar, M., & Moreno, L. 1998a. A computational approach to pronominal anaphora, one-anaphora and surface count anaphora. *Pages 117-128 of: Botley, Simon, & McEnery, Tony (eds), Proceedings of Second Colloquium on Discourse Anaphora and Anaphor Resolution (DAARC2)*.
- Ferrández, A., Palomar, M., & Moreno, L. 1998b. Anaphora resolution in unrestricted texts with partial parsing. *Pages 385-391 of: Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics (COLING-ACL'98)*.
- Ferrández, A., Palomar, M., & Moreno, L. 1999. An empirical approach to Spanish anaphora resolution. *Machine Translation*, **14**(2-3).

- Fox, B. 1987. *Discourse Structure and Anaphora*. Written and conversational English. Cambridge Studies in Linguistics. Cambridge: Cambridge University Press.
- Franchini, E. 1986. *Las condiciones gramaticales de la coordinación copulativa en español*. Romanica Helvetica, vol. 102. Bern: Francke Verlag.
- Gaizauskas, R., & Humphreys, K. 1996. Quantitative Evaluation of Coreference Algorithms in an Information Extraction System. In: Botley, Simon, & McEnery, Tony (eds), *Proceedings of the Discourse Anaphora and Anaphor Resolution Colloquium (DAARC)*.
- Gallardo, B. 1996. *Análisis conversacional y pragmática del receptor*. Colección Sinapsis. Valencia: Ediciones Episteme, S.L.
- Goldfarb, C.F. 1990. *The SGML Handbook*. Oxford University Press.
- Grosz, B. 1977. The representation and use of focus in a system for understanding dialogs. *Pages 67-76 of: Proceedings of Fifth International Joint Conference on Artificial Intelligence (IJCAI'77)*.
- Grosz, B., & Sidner, C. 1986. Attentions, intentions and the structure of discourse. *Computational Linguistics*, 12(March), 175-204.
- Grosz, B., Joshi, A., & Weinstein, S. 1983. Providing a unified account of definite noun phrases in discourse. *Pages 44-50 of: Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*.
- Grosz, B., Joshi, A., & Weinstein, S. 1995. Centering: a framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2), 203-225.
- Hausser, R. 1999. *Foundations of Computational Linguistics*. Berlin Heidelberg New York: Springer-Verlag.
- Hearst, M.A. 1994. Multi-paragraph segmentation of expository text. *Pages 9-16 of: Proceedings of 32nd Annual Meeting of the Association for Computational Linguistics*.
- Hirst, G. 1981. *Anaphora in Natural Language Understanding*. Berlin: Springer-Verlag.

- Hobbs, J. 1978. Resolving pronoun references. *Lingua*, 44, 311–338. Reprinted in (Hobbs, 1986).
- Hobbs, J. 1986. Resolving pronoun references. In: B.L. Webber, B. Grosz, & Spark-Jones, K. (eds), *Readings in Natural Language Processing*. Palo Alto, California: Morgan Kaufmann.
- Huls, C., Bos, E., & Claasen, W. 1995. Automatic Referent Resolution of Deictic and Anaphoric Expressions. *Computational Linguistics*, 21(1), 60–79.
- Kameyama, M. 1997. Recognizing Referential Links: An Information Extraction Perspective. *Pages 46–53 of: Proceedings of ACL/EACL workshop on Operational factors in practical, robust anaphora resolution*.
- Lappin, S., & Leass, H.J. 1994. An algorithm for pronominal anaphora resolution. *Computational Linguistics*, 20(4), 535–561.
- Lappin, S., & McCord, M. 1990. A syntactic filter on pronominal anaphora in slot grammar. *Pages 135–142 of: Proceedings of 28th Annual Meeting of the Association for Computational Linguistics*.
- Levinson, S.C. 1983. *Pragmatics*. England: Cambridge University Press.
- Lyons, J. 1977. *Semantics*. Vol. 2. Cambridge University Press.
- Martí, M.A., Rodríguez, H., & Serrano, J. 1998. *Declaración de categorías morfosintácticas*. Proyecto ITEM. Doc. nº2. <http://sensei.ieec.uned.es/item>.
- Martínez-Barco, P. 1999. Algoritmo de resolución de la anáfora pronominal en diálogos. *Procesamiento del Lenguaje Natural*, 24, 75–82.
- Martínez-Barco, P. 2000. Dialogue structure as a preference in anaphora resolution systems. *Pages 37–42 of: Proceedings of International Conference on Artificial and Computational Intelligence For Decision, Control and Automation In Engineering and Industrial Applications (ACIDCA '2000)*.
- Martínez-Barco, P., & Palomar, M. 2000a. An annotation scheme for dialogues applied to anaphora resolution algorithms. *Pages 410–414 of: Sojka, P., Kopecek, I., & Pala, K. (eds), Proceedings of the Third International Workshop on TEXT*,

- SPEECH and DIALOGUE (TSD'2000)*. Lecture Notes in Artificial Intelligence, vol. 1902. Brno, Czech Republic: Springer-Verlag.
- Martínez-Barco, P., & Palomar, M. 2000b. Dialogue structure influence over anaphora resolution. *Pages 515–525 of: Cairo, O., Sucar, L.E., & Cantu, F.J. (eds), MICAI 2000: Advances in Artificial Intelligence*. Lecture Notes in Artificial Intelligence, vol. 1793. Acapulco, México: Springer-Verlag.
- Martínez-Barco, P., & Palomar, M. 2000c. Empirical study of the anaphoric accessibility space in Spanish dialogues. *Pages 123–128 of: Proceedings of the Fourth Workshop on the Semantics and Pragmatics of Dialogue (GÖTALOG'2000)*.
- Martínez-Barco, P., & Palomar, M. 2000d. Resolución de la anáfora: estructura del diálogo y conocimiento lingüístico. *Procesamiento del Lenguaje Natural*, **26**, 47–52.
- Martínez-Barco, P., Peral, J., Ferrández, A., Moreno, L., & Palomar, M. 1998. Analizador Parcial SUPP. *Pages 329–341 of: Coelho, Helder (ed), Progreso em Inteligência Artificial*.
- Martínez-Barco, P., Muñoz, R., Azzam, S., Palomar, M., & Ferrández, A. 1999. Evaluation of pronoun resolution algorithm for Spanish dialogues. *Pages 325–332 of: Proceedings of the Venezia per il Trattamento Automatico delle Lingue (VEXTAL'99)*.
- Miller, G., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. 1993. Five Papers on WordNet. *International Journal of Lexicography*, **3**.
- Mitkov, R. 1998. Robust pronoun resolution with limited knowledge. *In: Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics (COLING-ACL'98)*.
- Moreno, L., Palomar, M., Molina, A., & Ferrández, A. 1999. *Introducción al procesamiento del lenguaje natural*. Alicante: Universidad de Alicante.
- Morris, J., & Hirst, G. 1991. Lexical cohesion computed by thesaural relations as an indicator of the structure of text. *Computational Linguistics*, **17**(1), 21–42.

- Muñoz, R., Montoyo, A., Llopis, F., & Suárez, A. 1998. Reconocimiento de entidades en el sistema EXIT. *Procesamiento del Lenguaje Natural*, **23**, 47–53.
- Muñoz, R., Martínez-Barco, P., & Ferrández, A. 1999. Método para la resolución de correferencias de sintagmas nominales definidos incluyendo alias y acrónimos en el sistema de extracción de información EXIT. *Procesamiento del Lenguaje Natural*, **25**, 143–149.
- Muñoz, R., Palomar, M., & Ferrández, A. 2000. Processing of Spanish Definite Descriptions. *Pages 526–537 of: Cairo, O., Sucar, L.E., & Cantu, F.J. (eds), MICAI 2000: Advances in Artificial Intelligence. Lecture Notes in Artificial Intelligence*, vol. 1793. Acapulco, México: Springer-Verlag.
- Owen, M. 1981. Conversational Units and the use of 'well...'. *Pages 99–116 of: P.Werth (ed), Conversation and Discourse*. New York: St Martin's Press.
- Palomar, M., & Martínez-Barco, P. 2000. Anaphora Resolution through Dialogue Adjacency Pairs and Topics. *Pages 196–203 of: Christodoulakis, D.N. (ed), Natural Language Processing - NLP 2000. Lecture Notes in Artificial Intelligence*, vol. 1835. Patras, Grece: Springer-Verlag.
- Palomar, M., & Martínez-Barco, P. 2001. Computational approach to anaphora resolution in Spanish dialogues. *Journal of Research in Artificial Intelligence*. Submitted.
- Palomar, M., Ferrández, A., Moreno, L., Saiz-Noeda, M., Muñoz, R., Martínez-Barco, P., Peral, J., & Navarro, B. 1999. A Robust Partial Parsing Strategy based on the Slot Unification Grammars. *Pages 263–272 of: Proceeding of 6e Conférence annuelle sur le Traitement Automatique des Langues Naturelles. TALN'99*.
- Palomar, M., Saiz-Noeda, M., Muñoz, R., Suárez, A., & Martínez-Barco, P. 2000. PHORA: A system to solve the Anaphora in Spanish. *Pages 206–211 of: Proceedings of Third International Conference on Discourse Anaphora and Anaphor Resolution (DAARC'2000)*.
- Palomar, M., Saiz-Noeda, M., Muñoz, R., Suárez, A., Martínez-Barco, P., & Montoyo, A. 2001. PHORA: A NLP aystem

- for Spanish. In: Gelbukh, Alexander (ed), *Proceedings of 2nd International conference on Intelligent Text Processing and Computational Linguistics (CICLing-2001)*. Lecture Notes in Artificial Intelligence. Mexico City: Springer-Verlag. To appear.
- Peral, J., Palomar, M., & Ferrández, A. 1999a. Coreference-oriented Interlingual Slot Structure and Machine Translation. *Pages 69-76 of: Proceedings of ACL workshop Coreference and its Applications*.
- Peral, J., Martínez-Barco, P., Muñoz, R., Ferrández, A., Moreno, L., & Palomar, M. 1999b. Una técnica de análisis parcial sobre textos no restringidos (SUPP) aplicada a un Sistema de Extracción de Información (EXIT). *Pages 662-669 of: Proceedings of VI Simposio Internacional de Comunicación Social*.
- Pereira, F., & Warren, D. 1980. Definite Clause Grammars for Language Analysis - A survey of the Formalism and a comparison with augmented transition networks. *Artificial Intelligence*, **13**.
- Pla, F. 1999. *Aproximación probabilística al etiquetado automático léxico de textos*. Tech. rept. DSIC-II/14/99. Departament de Sistemes Informàtics i Computació. Universitat Politècnica de València.
- Pla, F., & Molina, A. 1999. Etiquetador Léxico basado en Modelos ECGI Extendidos y su aplicación al Corpus BDGEO. *Pages 36-43 of: Actas de la Conferencia de la Asociación Española para la Inteligencia Artificial CAEPIA-TTIA '99*.
- Pla, F., Molina, A., & Prieto, N. 2000. Tagging and Chunking with Bigrams. In: *Proceedings of 18th International Conference on Computational Linguistics (COLING'2000)*.
- Polanyi, L. 1985. Conversational storytelling. *Pages 183-202 of: Dijk, Van (ed), Handbook of Discourse Analysis*, vol. 3. London: Academic Press.
- Proyecto BASURDE. 1998-2001. *Spontaneous-Speech Dialogue System in Limited Domains*. Comisión Interministerial de Ciencia y Tecnología TIC98-423-C06. <http://gps-tsc.upc.es/veu/basurde/Home.htm>.

- Proyecto CRATER. 1994-1995. *Corpus Resources And Terminology ExtRaction*. MLAP-93/20. <http://www.lllf.uam.es/proyectos/crater.html>.
- Proyecto ITEM. 1996-1999. *Recuperación de Información Textual en un Entorno Multilíngüe con Técnicas de Lenguaje Natural*. Comisión Interministerial de Ciencia y Tecnología TIC96-1243-C03. <http://sensei.ieec.uned.es/item>.
- Quesada, J.F., Amores, J.G., Fernández, M.G., Bernal, J.A., & López, M.T. 2000. Design Constraints and Representation for Dialogue Management in the Automatic Telephone Operator Scenario. *Pages 137-141 of: Proceedings of the Fourth Workshop on the Semantics and Pragmatics of Dialogue (GÖTALOG'2000)*.
- Reinhart, T. 1983. *Anaphora and Semantic Interpretation*. London and Sydney: Croom Helm.
- Reynar, Jeffrey C. 1994. An automatic method of finding topic boundaries. *Pages 331-333 of: Proceedings of 32nd Annual Meeting of the Association for Computational Linguistics*.
- Reynar, Jeffrey C. 1999. Statistical Models for Topic Segmentation. *Pages 357-364 of: Proceedings of 37th Annual Meeting of the Association for Computational Linguistics*.
- Rich, E., & LuperFoy, S. 1988. An architecture for anaphora resolution. *Pages 18-24 of: Proceedings of the Second Conference on Applied Natural Language Processing (ANLP'88)*.
- Richmond, K., Smith, A., & Amitay, E. 1997. Detecting subject boundaries within text: A language independent statistical approach. *Pages 47-54 of: Exploratory Methods in Natural Language Processing*.
- Rocha, M. 1998. *A corpus-based study of anaphora in dialogues in English and Portuguese*. Ph.D. thesis, University of Sussex, Sussex. UK.
- Rocha, M. 1999. Coreference resolution in dialogues in English and Portuguese. *Pages 53-60 of: Proceedings of ACL workshop Coreference and its Applications*.
- Sacks, H., Schegloff, E., & Jefferson, G. 1974. A simplest systematics for the organization of turn taking for conversation. *Language*, 50(4), 696-735.

- Saiz-Noeda, M., & Palomar, M. 2000. Semantic Knowledge-driven Method to Solve Pronominal Anaphora in Spanish. *Pages 204-211 of: Christodoulakis, D.N. (ed), Natural Language Processing - NLP 2000. Lecture Notes in Artificial Intelligence*, vol. 1835. Patras, Grece: Springer-Verlag.
- Saquete, E., & Martínez-Barco, P. 2000. Grammar Specification for the Recognition of Temporal Expressions. *Pages 21.1-21.7 of: Proceedings of the Machine Translation and multilingual applications in the new millennium, MT2000.*
- Siegel, S., & Castellan, J. 1988. *Nonparametric Statistics for the Behavioral Sciences*. 2nd edn. McGraw-Hill.
- Sinclair, J. 1993. Written discourse structure. *In: Sinclair, J., Hoey, M., & Fox, G. (eds), Techniques of description: spoken and written discourse: a Festschrift for Malcolm Coulthard*. London: Routledge.
- Strube, M. 1998. Never Look Back: An Alternative to Centering. *Pages 1251-1257 of: Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics (COLING-ACL'98).*
- Strube, M., & Hahn, U. 1999. Functional Centering - Grounding Referential Coherence in Information Structure. *Computational Linguistics*, 25(5), 309-344.
- Stubbs, M. 1983. *Discourse Analysis*. Oxford: Blackwell.
- Summer Institute of Linguistics. 1997. *Linguistic Glossary*. <http://www.sil.org/linguistics/glossary/>.
- van Dijk, T. 1977. *Text and context: explorations in the semantics and pragmatics of discourse*. London: Longman.
- van Dijk, T. 1995. *Texto y contexto: semántica y pragmática del discurso*. Madrid: Cátedra.
- Weizenbaum, J. 1966. ELIZA: a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9, 36-45.
- Winograd, T. 1972. *Understanding Natural Language*. New York: Academic Press.
- Winograd, T. 1986. A procedural model of language understanding. *In: B. J. Grosz, K. Sparck-Jones, B. L. Webber (ed),*

- Readings in Natural Language*. Los Altos, California: Morgan Kaufman.
- Woods, W. 1977. Lunar rocks in natural English: Explorations in natural language question answering. In: Zampolli, A. (ed), *Linguistics Structures Processing*. New York: Elsevier.
- Youmans, G. 1991. A new tool for discourse analysis: The vocabulary management profile. *Language*, 67(4), 763-789.
- Zaenen, A., Ericsson, S., Larsson, S., Mikheev, A., Milward, D., Pinkal, M., Poessio, M., Rupp, C.J., & Worm, K. 2000. *Robust interpretation and dialogue dynamics*. Tech. rept. LE4-8314. TRINDI. Task Oriented Instructional Dialogue. Deliberable D5.2.

UNIVERSIDAD DE ALICANTE

Comisión de Doctorado

Reunido el Tribunal que suscribe en el día de la fecha acordó otorgar, por UNANIM. a la Tesis Doctoral de Don/Dña. PATRICIO MANUEL MARTÍNEZ BARCO la calificación de SOBRESALIENTE CUM LAUDE

Alicante 3 de marzo de 2001

El Secretario,

El Presidente,


Fdo. LIDIA MORENO


Fdo. ANTONIO FERNANDEZ

UNIVERSIDAD DE ALICANTE

Comisión de Doctorado

La presente Tesis de D. Patricio Manuel Martínez Barco ha sido registrada al Folio con el n.º 750-21P del registro de entrada correspondiente.

Alicante 12 de marzo de 2001

El Encargado del Registro,

