

LOS MARCADORES DISCURSIVOS EN CHIEDE, UN CORPUS DE HABLA INFANTIL ESPONTÁNEA

ANA GONZÁLEZ-LEDESMA
MARTA GARROTE SALAZAR

UNIVERSIDAD AUTÓNOMA DE MADRID

Resumen: *Presentamos en este artículo un estudio sobre los marcadores discursivos utilizados en niños de tres, cuatro y cinco años de edad. CHIEDE ha sido el corpus de datos que hemos utilizado para el análisis, un corpus de habla oral infantil en el que los marcadores discursivos han sido etiquetados después de un proceso de desambiguación categorial realizado por medio de reglas contextuales. Hemos clasificado los marcadores discursivos siguiendo dos criterios de naturaleza cognitiva: las relaciones semánticas entre los enunciados —adición, oposición, concesión, etc.— y la interacción social, dentro de la cual, hemos diferenciado a su vez entre dos grupos de marcadores: los conativos, que son aquellos marcadores que marcan un contacto con el interlocutor; y los modalizadores epistémicos, a través de los cuales se marca la negociación sobre la verdad de los enunciados que implica toda interacción social. A través del estudio de estas palabras que codifican la relación entre lo social, lo mental y lo lingüístico, pretendemos acercarnos a la adquisición del lenguaje en relación con otras capacidades cognitivas del ser humano. Finalmente, contrastaremos los resultados obtenidos con el estado del arte en este campo de investigación.*

Palabras clave: *habla oral infantil, pragmática, marcadores discursivos y adquisición del lenguaje.*

Abstract: *We show in this paper a study about discourse markers used by children from three to five years old. CHIEDE is the data corpus we have employed in our analysis, a child speech corpus in which discourse markers have been tagged after a process of categorical disambiguation through contextual rules. Discourse markers have been classified following two cognitive criteria: the semantic relationships between utterances —addition, opposition, concession, etc.; and the social interaction, within which, two subgroups of markers are distinguished: connatives, that is, those discourse markers which mark contact with the interlocutor; and epistemic modalizers, through which the negotiation about utterances truth is marked. This papers aims to be an empirical approach to language acquisition in connection with other human cognitive capacities through the study of certain words, discourse markers, which precisely encode this relationship between language, cognition and social interaction. Finally, the results will be contrasted with the data presented so far in this research field.*

Keywords: *child speech, pragmatics, discourse markers and language acquisition.*

1. Introducción

El objetivo de esta comunicación¹ es presentar un estudio interdisciplinar en el que la Lingüística de Corpus y el tratamiento computacional de los textos se ponen al servicio de otra

¹ Esta investigación ha sido financiada parcialmente por el MEC (proyecto RILARIM TIN2004-07588-C03-02) y por la Comunidad de Madrid a través de una beca de Formación de Personal Investigador para

área de conocimiento, la Psicolingüística, con el fin de aportar luz, desde una perspectiva cuantitativa, sobre cuestiones que giran en torno a la adquisición del lenguaje y su relación con el desarrollo cognitivo del niño. En concreto, la inquietud que nos ha llevado a realizar esta investigación ha sido la de conocer cuál es el estado en el que se encuentran las relaciones tanto discursivas como interactivas de las conversaciones protagonizadas por los niños de tres, cuatro y cinco años de edad que conforman nuestro corpus. El objeto de estudio a través del cual hemos intentado responder a este interrogante ha sido el de los marcadores discursivos, palabras cuya semántica contribuye de una forma clave a la interpretación de los procesos comunicativos en las interacciones, y que han recibido una escasa atención —como en realidad todo el estudio del nivel pragmático de la lengua— en los análisis de adquisición del lenguaje en niños de estas edades, a juzgar por la escasa bibliografía encontrada.

Tal y como se especifica en el título de este trabajo, el banco de datos con el que hemos trabajado ha sido un corpus de habla infantil espontánea. No obstante, convertir el corpus en una fuente de conocimiento válida para hacer generalizaciones relevantes para el mundo de la psicolingüística no es una cuestión baladí. Antes de extraer los datos, los corpus digitalizados deben someterse a procesos computacionales de formalización, etiquetado y validación; y más aún, como es nuestro caso, cuando las partículas que son objeto de nuestro estudio presentan en algunos casos una elevada tasa de ambigüedad categorial.

Por todas estas razones, presentamos en este trabajo un experimento desde su inicio (la conformación del corpus) hasta su etapa final, la interpretación de los datos, pasando, como ya hemos explicado, por la fase de formalización y procesamiento de los mismos. Como consecuencia de ello, las partes de dicho artículo aludirán claramente a las diferentes etapas de esta investigación.

Así pues, en el primer apartado el lector podrá acceder a toda la información que respecta al diseño y características de CHIEDE, nuestro corpus de habla infantil. A continuación, en el segundo apartado, presentamos nuestra clasificación de marcadores discursivos y su relación con el objetivo final de nuestro estudio. En el apartado siguiente, nos centramos en la fase de procesamiento computacional del corpus. Después, presentamos los resultados de las frecuencias de uso de los marcadores discursivos seleccionados en función de la variable edad, y posteriormente ofrecemos algunas generalizaciones descriptivas que en el siguiente apartado se enmarcarán dentro de las teorías desarrolladas sobre la adquisición del lenguaje, el nivel pragmático y el desarrollo de la cognición en el individuo. Veremos cómo en contra de las pocas afirmaciones que se puedan relacionar con este nivel de la lengua, nuestros datos muestran que los niños en estas edades están aprendiendo a manejar la interacción con el adulto y que utilizan palabras que conectan y dan coherencia a su discurso, al tiempo que saben diferenciar entre sus estados de conocimiento y los del interlocutor. Finalmente, en el último apartado, presentamos un resumen de aportaciones más interesantes de este estudio.

2. CHIEDE, Corpus de Habla Infantil Espontánea del Español

CHIEDE es un corpus conformado por 60.000 palabras cuyo diseño está estructurado en torno a cuatro criterios: el tipo de situación comunicativa, la edad de los participantes, el sexo de los mismos, y finalmente, la simetría en cada uno de los grupos establecidos.

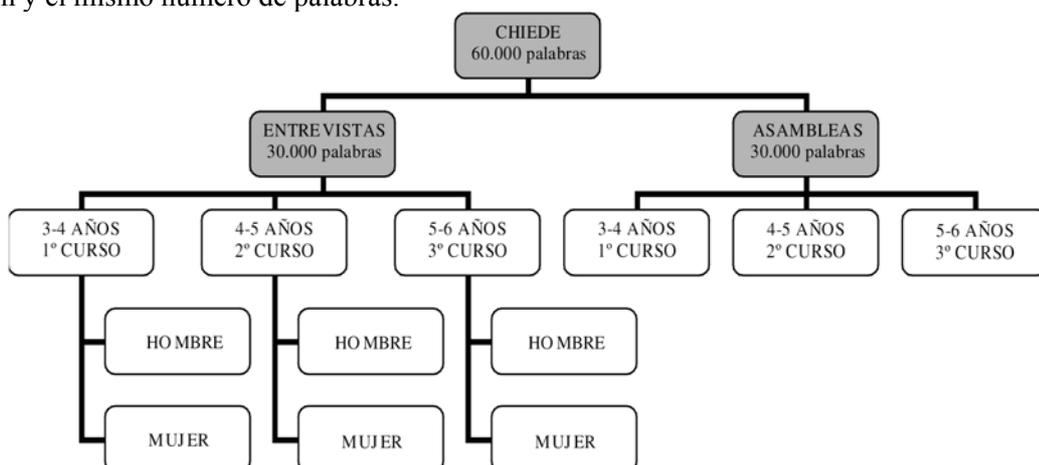
En lo que respecta al primer criterio, hemos diferenciado dos situaciones comunicativas: las entrevistas y las asambleas. Las entrevistas son interacciones conformadas por un niño y un adulto, en las que normalmente el adulto le hace preguntas al niño sobre su edad, su familia, su vida cotidiana o lo que está pasando en alguna foto. En cambio, las asambleas son situaciones de interacción colectiva en las que el profesor guía la conversación con el fin de estimular la actividad verbal en el niño.

la realización de la tesis “Un modelo de anotación pragmática para el corpus oral del español C-ORAL-ROM: estrategias de etiquetado y aplicaciones”.

El segundo criterio viene definido por la edad de los participantes. Tal y como ya señalamos en la introducción, los participantes de nuestro corpus tienen de tres a seis años de edad, y están agrupados por cursos escolares, que van del primer al tercer año de la primera etapa de escolarización, la Educación Infantil. Hemos elegido este margen de edad porque en él sitúan el principal período de aprendizaje los expertos en adquisición y desarrollo del lenguaje —con anterioridad el niño desarrolla diferentes formas de comunicación pero aún no un lenguaje similar al de los adultos (A. ANULA REBOLLO 1998). Como se podrá observar en el diagrama que presentamos más abajo, no todos los niños de un mismo curso tienen los mismos años. Así por ejemplo, en el primer curso la mayoría de los niños tiene tres años, pero también encontramos niños que ya han cumplido los cuatro. Este dato es importante para el análisis, porque dentro de cada grupo nos encontramos con niños en diferentes estados del proceso de aprendizaje.

A su vez, dentro de cada curso, las interacciones están divididas en dos grupos más, de acuerdo con el tercer criterio de nuestro diseño, el sexo de los hablantes.

En cuanto al cuarto y el último parámetro, el de la simetría, cada grupo y cada subgrupo del corpus está constituido por el mismo número de interacciones. Esto es, dentro de cada curso, podemos encontrar el mismo número de niños que de niñas —cuatro sujetos de cada subgrupo sumando un total de veinticuatro entrevistas—, y a su vez, en cada curso también podemos encontrar el mismo número de grabaciones, las cuales tienen aproximadamente la misma duración y el mismo número de palabras.



Es importante destacar en esta sección que, si bien no nos encontramos ante un corpus de grandes dimensiones, es sin embargo el primero en España que se ha elaborado respondiendo a los criterios ya señalados. Hasta el momento, lo que el estado de la cuestión en nuestro país nos ofrece en materia de recursos para analizar el habla infantil son en su mayoría corpus longitudinales, esto es, corpus conformados por un solo participante donde se sigue el desarrollo del lenguaje del niño a lo largo del tiempo (J. LLISTERRI 1996). En cuanto al panorama internacional, en lo que a corpus de habla infantil espontánea en lengua española se refiere, contamos con el proyecto CHILDES, un corpus multilingüe que incluye muestras en lengua española. Dicho proyecto fue concebido como un sistema de intercambio de datos que se nutre de aportaciones por parte de investigadores particulares que cuelgan sus propios trabajos en la página Web de CHILDES (<http://chilides.psy.cmu.edu/>).

3. Tipología de marcadores discursivos

Tal y como señala PORTOLÉS (1998) en una de las definiciones más citadas por los estudios de pragmática lingüística del español:

Los marcadores del discurso son unidades lingüísticas invariables, no ejercen una función sintáctica en el marco de la predicación oracional y poseen un cometido

coincidente en el discurso: el de guiar de acuerdo con sus distintas propiedades morfosintácticas, semánticas y pragmáticas las inferencias que se realizan en la comunicación.

No obstante, aunque en términos generales vamos a ceñirnos a su definición de marcador discursivo, no vamos a hacer lo mismo con su clasificación, ya que, a pesar de ser también una de las más consolidadas en los estudios pramalingüísticos del español, hacerlo significaría alejarnos de los fines de nuestra investigación, que es relacionar la cognición con el desarrollo del lenguaje del niño. Así pues, inspirándonos en este autor pero también en la tradición francesa sobre estudios de la conversación (ROULET 1985), hemos hecho una tipología más sencilla, en la que se diferencian dos grupos principalmente: los marcadores relacionales y los interactivos. El primero de ellos tiene que ver con las relaciones semánticas que tienen lugar entre los enunciados; mientras que en el segundo grupo, se atiende a la labor interactiva de algunas partículas dentro de la comunicación. Hemos dividido a su vez este segundo grupo en dos subtipos más en función de estos dos criterios: la preocupación por la eficacia en la transmisión del mensaje y la negociación sobre la verdad de los enunciados que tiene lugar entre hablante e interlocutor durante la interacción.

Antes de presentar cada uno de los grupos más detenidamente, queremos hacer dos aclaraciones. La primera de ellas tiene que ver con el inventario de marcadores discursivos. Como se podrá observar en las tablas que presentamos a continuación, la lista de marcadores seleccionados es un compromiso entre los que realmente aparecen en el corpus de habla infantil consultado y los que nos han parecido más relevantes en términos cognitivos para el objeto final de nuestro estudio.

La segunda aclaración tiene que ver con la naturaleza categorial de estas partículas. Los marcadores discursivos que hemos seleccionado tienen una variada naturaleza gramatical. Algunos incluso son lexicalizaciones de partes de una oración, como por ejemplo, *lo que pasa es que*, o incluso oraciones completas como *¿sabes qué?* o *claro que sí*. Detrás de esta elección hay una postura teórica sólidamente fundamentada. Los fenómenos pragmáticos se codifican en cada lengua a través de recursos lingüísticos que abarcan todos los niveles gramaticales descritos hasta el momento por la tradición lingüística. Así por ejemplo, la modalidad epistémica, y dentro de ella las operaciones de atenuación, son un buen ejemplo que ilustra la idea que acabamos de expresar, ya que el español ofrece diversos recursos para atenuar el discurso que abarcan desde la morfología apreciativa, *¿nos tomamos una cervecita?*, hasta la sintaxis, *Esta chica es un poco como ñoña, ¿no?*. En este caso, tanto para la modalidad como para las relaciones entre los enunciados, hemos seleccionado tanto palabras como también sintagmas ya lexicalizados (*es que*) a los que les hemos atribuido el estatus de marcador discursivo. Algunos de ellos, como el caso de *así que* y *es que*, ya están aceptados en la bibliografía como marcadores discursivos; otros como *claro que sí* o *sí que*, a pesar de que no los hemos encontrado registrados en la literatura consultada, han sido incluidos en nuestro inventario porque consideramos que también son codificaciones lingüísticas de un fenómeno pragmático.

A continuación, sin más preámbulos, presentamos más detalladamente cada uno de estos dos grupos.

3.1 Marcadores relacionales

Éstas que presentamos más abajo son las relaciones discursivas más frecuentemente utilizadas en nuestro corpus. La escuela francesa hace una distinción clara entre marcadores textuales y marcadores interactivos (ROULET 1985). Nosotros hemos preferido adoptar un criterio de designación más cognitivo. PIAGET defendía que el lenguaje de la lógica era el más apropiado para representar las operaciones mentales. Sin embargo, como sabemos, las lenguas naturales ofrecen también sus dispositivos para codificar dichas operaciones mentales. Buena parte de los marcadores discursivos se ocupan de esta tarea; pensemos, por ejemplo, en el condicional lógico *si*. A la tabla que exponemos a continuación, presentamos un número de

operaciones que consideramos de gran importancia en la relación entre lengua y pensamiento, y junto a ellas, un inventario de unidades de análisis que en español codificarían dicha relación.

MARCADORES RELACIONALES	
Relación semántica	Codificación lingüística en español
ADICIÓN	<i>y, luego</i>
OPOSICIÓN	<i>pero</i>
CONSECUCIÓN	<i>pues, así es que, porque, así que, entonces</i>
CONCESIÓN	<i>aunque</i>
JUSTIFICACIÓN	<i>es que, lo que pasa es que</i>
FINALIDAD	<i>para que</i>
HIPÓTESIS	<i>si</i>

3.1.2 Coordinación y subordinación: un estudio detallado del *que*

Una de las críticas que se le puede hacer a esta clasificación es que esta relación entre lengua y pensamiento a la que acabamos de aludir pasa en una lengua por la sintaxis. Algunos de los marcadores que hemos incluido en este inventario son claramente conectores, como por ejemplo *pues*, que funciona siempre a nivel extraoracional. Sin embargo, otros como la conjunción *pero* y la conjunción *y*, como ya sabemos, pueden relacionar tanto enunciados como sintagmas. Asimismo, también hemos incluido en el inventario otras partículas que la tradición gramatical define como subordinantes como *porque* y *aunque*. Sin embargo, aunque la diferencia entre el grado de subordinación con respecto a la oración principal sea más laxo en el caso de *porque* que en el caso de *aunque*, en la modalidad oral es francamente difícil determinar si su comportamiento es hipotáctico o paratáctico, ya que estas conjunciones pueden encabezar enunciados que son comienzo de turno o que se inician después de una pausa fuerte y descendente.

Sin embargo, a pesar de que en la modalidad oral estas fronteras teóricas se desdibujen, consideramos de gran relevancia someter a análisis si hay alguna diferencia en la frecuencia de uso de la subordinación frente al uso de la coordinación en los niños de estas edades. Para ello hemos utilizado como objeto de estudio la partícula *que*.

No obstante, para hacer un análisis de esta partícula en relación a un modelo de tipos de conexión, la tradicional división entre el *que* conjunción y el *que* pronombre relativo no nos ha parecido suficiente. Con lo cual, durante el etiquetado, hemos introducido dos tipos de información en el estudio de esta partícula. El primero es el que atiende a su naturaleza gramatical: pronombre relativo o conjunción; el segundo tipo en cambio se atiende a su posición discursiva, diferenciando dentro de este parámetro tres valores: principio de turno, principio de enunciado (después de una pausa fuerte y descendente que en nuestro formato de transcripción se marcó con tres barras ///) y, por último, interior de enunciado.

3.2 Marcadores de subjetividad

Presentamos en este grupo un conjunto de expresiones que tienen que ver, como ya hemos señalado más arriba, con la percepción que tiene el sujeto del conocimiento que está transmitiendo durante la interacción verbal. En el primer grupo, incluimos los marcadores que en PORTOLÉS (1993) se clasifican como conativos, y que tal y como este autor indica, aseguran el mantenimiento del contacto del canal de comunicación; además, tenemos que añadir también que, desde nuestra perspectiva, el uso de estos marcadores implica que los hablantes saben que

la comunicación puede modificar el mundo de conocimientos que cada individuo posee, de ahí que intenten asegurarse mediante dichas partículas de que este proceso de transmisión se esté llevando a cabo de forma eficaz.

En el segundo grupo presentamos aquellas palabras que marcan las operaciones pragmáticas de modalización epistémica. La modalización es un tipo de fenómeno pragmático que alude al grado de compromiso que el hablante muestra hacia la verdad del enunciado que está transmitiendo. Hemos decidido incluir las marcas de modalización dentro de este estudio porque consideramos que constituyen un parámetro de análisis clave para conocer el desarrollo cognitivo del niño, ya que el uso de un marcador de modalidad implicaría no solamente que el niño sabe diferenciar entre la realidad y la percepción o grado de conocimiento que tiene de ella, sino que además también indicaría que es consciente de que el interlocutor puede tener un grado de conocimiento diferente al suyo sobre la verdad de las cosas. Con lo cual, según esta última idea, la modalización tiene también un valor interactivo, en tanto que supone una negociación con el interlocutor sobre el conocimiento o la verdad de los enunciados. En este sentido, los atenuantes, como su propio nombre indica, servirían para aminorar o atenuar el grado de compromiso del sujeto con la verdad de lo que dice, bien porque no está seguro, bien porque percibe que su interlocutor pueda tener una posición diferente ante el mismo estado de cosas. Mientras, por su parte, la intensificación marcaría, por las mismas razones a las que acabamos de aludir, un aumento del grado de certeza sobre la verdad enunciativa que se está comunicando. A continuación presentamos el inventario de expresiones que hemos identificado en el corpus para expresar dichos fenómenos de naturaleza pragmático-cognitiva.

MARCADORES DE SUBJETIVIDAD		
CONATIVOS		<i>fíjate, sabes qué, ¿eh?, ¿vale?, ¿a que sí?, ¿sabes_X?, ¿a_que_no_sabes_X?, ¿o no?</i>
MODALIZADORES EPISTÉMICOS	Atenuantes	<i>me parece que, a lo mejor, bueno</i>
	Intensificadores	<i>seguro que, vale, claro, eso sí, sí que, claro_que_sí</i>

4. Anotación de los textos: procesamiento del corpus

Una vez que tenemos el corpus, conformado por los textos de la transcripción de las interacciones, etiquetamos esta información en formato XML, ya que esto nos permite un mayor control sobre los datos del corpus. A continuación, el siguiente paso es localizar y etiquetar las palabras con la información discursiva procedente de la clasificación que acabamos de presentar dentro del corpus. Para la anotación, hemos dividido la lista de palabras en dos tipos: ambiguas y no ambiguas. Las palabras no ambiguas, como por ejemplo *pero*, se localizan y etiquetan automáticamente en los textos. En cambio, el etiquetado de las palabras ambiguas, esto es, aquellas que tienen más de un análisis o significado en el texto – como por ejemplo, la palabra *bueno*, que puede ser marcador discursivo o adjetivo – requieren un coste de procesamiento mayor. Como nos enfrentamos a un corpus de unas dimensiones francamente manejables, nuestra metodología de etiquetado ha consistido, para las palabras ambiguas, en crear reglas contextuales que desambigüen los usos discursivos de otros usos categoriales. Posteriormente, la validación se ha llevado a cabo de forma manual, ya que como decimos, las dimensiones del corpus permiten la realización esta tarea. En la elaboración de las reglas contextuales de desambiguación nos hemos servido tanto de información categorial como de información prosódica y discursiva. El marcador discursivo que más coste de procesamiento ha implicado ha sido por su multifuncionalidad la partícula *que*. A pesar de que hemos elaborado reglas contextuales de desambiguación, finalmente hemos validado manualmente el etiquetado de esta partícula en todo el corpus.

Para extraer la información relevante del corpus hemos elaborado un programa en lenguaje PERL que extrae el número de ocurrencias de cada marcador discursivo por hablante y por curso académico. En el cuadro que viene a continuación exponemos los resultados.

5. Presentación de los resultados y análisis de los datos

A continuación presentamos los datos de frecuencia de uso de los marcadores seleccionados para el análisis de acuerdo con su clasificación semántica. Dichos datos expresan directamente en términos cuantitativos las recurrencias de cada grupo dentro del corpus en función del curso en el que los alumnos se encuentran.

MARCADOR DISCURSIVO		EDAD			
Tipo	Subtipo	3-4	4-5	5-6	
MARCADORES DE RELACIÓN	Adición	206	359	409	
	Consecución	22	46	95	
	Oposición	20	55	46	
	Justificación	14	17	36	
	Hipótesis	5	12	9	
	Finalidad	0	4	3	
	Concesión	0	0	4	
COORDINACIÓN VS. SUBORDINACIÓN	CATEGORÍA GRAMATICAL	Conjunción	44	105	135
		Relativo	12	49	65
	POSICIÓN DISCURSIVA	Ppio. turno	2	7	18
		Ppio. enunciado	17	34	38
		Interior enunciado	25	64	79
MARCADORES DE SUBJETIVIDAD	Conativos	3	7	15	
	Intensificadores	2	3	5	
	Atenuantes	2	11	15	

La primera generalización descriptiva que podemos apuntar después de observar los resultados obtenidos es el marcado carácter ascendente en la frecuencia de uso desde los tres hasta los seis años. En casi todos los tipos el uso de marcadores aumenta de forma clara del primer al segundo curso, y de forma un poco más relativa del segundo al tercer curso.

Dentro de los marcadores de relación, las relaciones discursivas más utilizadas por frecuencia de uso son la adición y la consecución, seguidas de las de oposición y justificación, y finalmente, en el último lugar, se encuentran la formulación de hipótesis y la expresión de la concesividad.

No podemos dejar de hacer hincapié en la aparición de la partícula *aunque* en el tercer curso, que anteriormente no ha sido empleada por ninguno de los hablantes de menor edad. Si añadimos este dato a las generalizaciones que podemos obtener sobre el *que*, esto es, aumento claro de la subordinación a medida que aumentan los años pero en ningún caso comparable a la frecuencia con la que se recurre a la coordinación como estrategia sintáctica para unir

enunciados, podemos concluir que efectivamente la adquisición de la subordinación como estrategia para relacionar enunciados parece que resultaría más tardía en el aprendizaje. Lo mismo podemos afirmar de la formulación de hipótesis y del uso de *aunque* con valor concesivo.

En segundo lugar, también podemos observar una diferencia importante en términos cuantitativos entre los marcadores de relación y los marcadores de interacción. Lo cual nos llevaría a pensar que aprendemos antes a relacionar enunciados que a ser conscientes de las consecuencias que conlleva la transmisión del conocimiento en términos de eficacia comunicativa y de modificación de estados mentales.

A continuación, veremos si desde el marco de la Psicolingüística se puede aportar más luz sobre las consecuencias cognitivas que tiene el empleo de dichas partículas por parte de los niños durante la comunicación.

6. Adquisición del lenguaje: conocimiento pragmático y desarrollo cognitivo

Existen diferentes teorías sobre la ontogénesis del lenguaje (innatismo, conductismo, etc.) y múltiples estudios al respecto. La gran mayoría de los experimentos en los que se han sustentado estas posturas teóricas han consistido en la realización de investigaciones de carácter longitudinal donde se observa el desarrollo del lenguaje de un niño (o de un número muy reducido de ellos) a lo largo de las diferentes etapas de su crecimiento.

Al consultar la bibliografía sobre lenguaje infantil nos hemos encontrado, de cara a nuestro objeto de estudio, dos serios inconvenientes. En primer lugar, la edad límite en la que los estudios longitudinales ya mencionados se detienen no superan los treinta meses de edad, esto es, los dos años y medio. En segundo lugar, si analizamos la metodología de estos experimentos, todos los estudios consultados (S. GILI GAYA 1972, D. CRYSTAL 1981, M. SIGUÁN 1983 ó A. ANULA 1998) coinciden en realizar un análisis del lenguaje infantil por niveles lingüísticos, entre los que se incluye la fonología, la morfosintaxis y la semántica, pero en los que el nivel discursivo e interaccional de la lengua ni siquiera se contempla como parámetro. Así por ejemplo, el método para el análisis lingüístico del lenguaje infantil que propone SIGUÁN (1983:40-41) se divide en los siguientes apartados:

1. Plano léxico: anotación de categorías, clasificando sustantivos y verbos en campos semánticos.
2. Plano morfológico: estudio del uso y combinación de morfemas.
3. Plano sintáctico: catalogación de estructuras sintácticas.
4. Plano oracional: análisis de la coordinación y subordinación.

Con lo cual, nuestro margen de acción a la hora de contrastar nuestros resultados con otros estudios se ve seriamente reducido a unas cuantas referencias que se dan al respecto en algunas de las obras consultadas que pasamos a comentar a continuación.

En cuanto a la dimensión pragmática de la lengua en el lenguaje infantil, los trabajos que han tratado este tema se centran fundamentalmente en la fuerza ilocutiva de las emisiones de los niños, las cuales, según autores como K. ARONSSON y M. THORELL (1999) o H. MARCOS y J. BERNICOT (1997) tendrían la finalidad de captar la atención de sus padres con el fin de impulsarlos en la mayoría de las ocasiones a realizar una acción. Pero, como ya hemos mencionado, no se encuentra ninguna referencia al modo en que el niño maneja la interacción cuando ya ha alcanzado la edad de tres años y es capaz de mantener una conversación *pseudo-adulta*, a pesar de que en el plano teórico se haga mención a dicho aprendizaje, como manifiestan las observaciones de D. CRYSTAL al respecto: «existe una gran porción de gramática que el niño *no* ha aprendido aún antes de ingresar en la escuela, y el proceso de ir completando dicho conocimiento de su lengua se mantendrá hasta casi el umbral de la pubertad» (D. CRYSTAL 1981:52)

Esta última afirmación de CRYSTAL se ve confirmada por los datos de nuestros datos, ya que como hemos visto hay operaciones discursivas que se realizan con mayor facilidad que otras, como demuestra el hecho de que se usen más; al tiempo que hay formas de conexión,

tales como la coordinación, que se adquieren más tempranamente, frente a la subordinación, cuya baja frecuencia de uso demostraría mayor dificultad de procesamiento.

Sin embargo, en otras afirmaciones del autor, no tenemos más remedio que disentir, puesto que nuestros datos evidencian que sus afirmaciones son erróneas. Así por ejemplo, también según este autor (D. CRYSTAL 1981:64), durante el tercer año de edad no existen claves de respuesta que señalen continuidad de la interacción (*mm, ajá, claro, sí*). Mientras que los resultados de nuestro corpus nos demuestran que ya a la edad de tres años se utilizan marcadores de subjetividad conativos, y a pesar de que la frecuencia de uso no sea excesivamente elevada, este hecho indica que ya a esa edad el niño es consciente de la necesidad de mantener el contacto del canal de comunicación. Podríamos afirmar pues que a esta edad el niño ha adquirido este tipo de habilidad pragmática.

El caso de los marcadores de subjetividad es especialmente interesante, sobre todo teniendo en cuenta la teoría del lenguaje egocéntrico planteada por J. PIAGET (1923), la cual ha sido refrendada por numerosos estudiosos del lenguaje infantil. J. PIAGET, en su obra *El lenguaje y el pensamiento en el niño*, planteó que el tipo de lenguaje que utiliza un niño hasta la edad de siete años es egocéntrico, es decir, es un simple acompañamiento de la acción; esto implicaría que el niño no poseería otra perspectiva que no fuera la propia, con lo que se asumiría que los demás piensan y perciben la realidad de la misma forma que uno mismo. J. PIAGET entendía que el diálogo no se producía hasta la edad de siete años. Otros autores como L. S. VIGOSTKY (1964) o E. BATES (1976), sin embargo, sostienen la idea de la función comunicativa y social del lenguaje, pero siguen manteniendo el límite de edad alrededor de los siete años en cuanto a las capacidades pragmático-comunicativas del niño.

Los datos obtenidos en nuestro estudio contradicen estas afirmaciones: el hecho de que niños de edades comprendidas entre los tres y cinco años hagan uso de marcadores discursivos, no sólo relacionales, sino también interaccionales, demuestra que son conscientes del papel que desarrollan dentro de la interacción. Esto confirma la idea de que no sólo son capaces de utilizar el lenguaje con una función extralingüística y referencial, sino que además poseen la capacidad de discriminar entre el conocimiento que él tiene de la realidad y el conocimiento de la realidad que pueda tener su interlocutor.

7. Conclusiones

Las conclusiones que se pueden extraer de nuestro estudio son las siguientes:

1. Por una parte, se puede afirmar, en contra de lo que muchos autores piensan, que el niño ya posee habilidades pragmático-comunicativas a la edad de tres años, a pesar de que el dominio de las mismas sea una tarea que se extienda hasta la pubertad. Así pues, mantener la teoría sobre el lenguaje egocéntrico infantil resulta inadecuada, ya que la función comunicativa del lenguaje en los niños se puede observar desde edades muy tempranas.
2. En segundo lugar, creemos que ha quedado patente que es necesario incrementar los estudios sobre lenguaje infantil en la pragmática de la lengua, al igual que la extensión cronológica de los estudios longitudinales, que como hemos comentado se reducen a la primera etapa del aprendizaje lingüístico, dejando el periodo que va desde los tres o cuatro años hasta la edad adulta sin estudiar.
3. Finalmente, abogamos por la interdisciplinariedad de terrenos de conocimiento que nos puedan conducir a una explicación de la realidad más global y al mismo tiempo más rigurosa y científica.

Referencias bibliográficas

- ANULA REBOLLO, ALBERTO, *El abecé de la psicolingüística*, Madrid, Arco-Libros, D. L., 1998.
- ARONSSON, KARIN & THORELL, MIA, «Family politics in children’s play directives», *Journal of Pragmatics* 31, 1999, 25-47.
- BATES, ELIZABETH, *Language and context: the acquisition of pragmatics*, New York, Academic Press, 1976.
- CRESTI, EMANUELA Y MONEGLIA, MÁXIMO, *C-ORAL-ROM Integrated Reference Corpora for Spoken Romance Languages*, Amsterdam, John Benjamins, 2005.
- CRYSTAL, DAVID, *Lenguaje infantil. Aprendizaje y lingüística*, Barcelona, Editorial Médica y Técnica, 1981.
- FRANCESCATO, GIUSEPPE, *El lenguaje infantil. Estructuración y aprendizaje*, Barcelona, Península, 1987.
- GILI GAYA, SAMUEL, *Estudios de lenguaje infantil*, Barcelona, Bibliograf, 1972.
- LLISTERRI, JOAQUIM, *Informe sobre recursos lingüísticos para el español (II). Corpus escritos y orales disponibles y en desarrollo en España*, Alcalá de Henares, Instituto Cervantes, 1996.
- MACWHINNEY, BRIAN, *The CHILDES Project (3rd Edition). Volume II: Tools for Analyzing Talk: The Database*, Mahwah, NJ, Lawrence Erlbaum Associates, 2000.
- MARCOS, HAYDÉE & BERNICOT JOSIE, «How do young children reformulate assertions? A comparison with requests», *Journal of Pragmatics* 27, 1997, 781-798.
- MENDOZA LARA, ELVIRA y LÓPEZ HERRERO, PAZ, «Consideraciones sobre el desarrollo de la teoría de la mente (ToM) y del lenguaje», *Revista de Psicología General y Aplicada*, Vol. 57, Nº 1, 2004, 49-68.
- PIAGET, JEAN, *El lenguaje y el pensamiento en el niño*, Buenos Aires, Paidós, 1965.
- PORTOLÉS, JOSÉ, *Los marcadores del discurso*, Barcelona, Ariel, 1998.
- ROULET, EDDY [et al.], *L’articulation du discours en français contemporain*, Berna, Peter Lang, 1985.
- SIGUÁN, MIGUEL, *Metodología para el estudio del lenguaje en la infancia*, Barcelona, Universidad de Barcelona, 1983.
- SIGUÁN, MIGUEL, *Estudios sobre psicología del lenguaje infantil*, Madrid, Pirámide, 1984.
- VYGOTSKY, LIEV, *Pensamiento y Lenguaje*, Buenos Aires, Lantaró, 1964.