

Durham E-Theses

Children's knowledge of indefinite and definite reference

Kerry. M. Sims

How to cite:

Sims, Kerry. M. (1990) Children's knowledge of indefinite and definite reference. Doctoral thesis, Durham University.

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a <https://etheses.durham.ac.uk/id/eprint/6233/> is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

Children's knowledge of Indefinite and Definite reference

Kerry. M. Sims

1990

The copyright of this thesis rests with the author.
No quotation from it should be published without
his prior written consent and information derived
from it should be acknowledged.

A thesis submitted for the degree of Doctor of Philosophy
in the University of Durham

Volume 1

Department of Psychology
University of Durham



26 AUG 1992

to my son, Oliver

Also, I dedicate this thesis to Robert. W. Sims
(November 1912 - June 1984).

Table of contents

<u>Volume 1</u>	page
Table of contents	i
List of tables	vi
List of figures	xii
Declaration and copyright	xiii
Acknowledgements	xiii
Abstract	xv
<u>Chapter 1: Introduction</u>	1
1.1 Theoretical Literature	2
A. Definitions of in/definiteness	2
B. Representation and Processing of in/ definite reference in discourse	11
1.2 Developmental Literature	23
A. Cognitive Development and Language Acquisition research	
1.2.A1 General knowledge and pragmatic inferences	23
1.2.A2 Interpersonal abilities for language acquisition	28
1.2.A3 Knowledge of the relationship between social - pragmatic and linguistic expressions	32
B. Developmental Research and indefinite/ definite reference	42
1.3 Summary of the literature	47
<u>Chapter 2: Psychological studies of children's indefinite and definite reference.</u>	50
2.1 Knowledge of 'Familiarity' and indefinite/ definite reference	51
A. Sensitivity to the listener's knowledge	51
B. Sensitivity to pragmatic factors	56
2.2 Knowledge of 'Specificity' and indefinite/ definite reference	58
A. The exophoric context	58
B. The anaphoric context	62
2.3 Studies of Children's Pronominal reference	64
A. Linguistic factors	
2.3.A1 Syntactic knowledge	65
2.3.A2 Lexical knowledge	73
B. Non-linguistic factors	
2.3.B1 Thematic-discourse structures	74
2.3.B2 Pragmatic inferences	75
2.3.B3 Heuristic principles	80
2.3.B4 Knowledge of ambiguity	

2.4	Summary and overview of studies on children's knowledge of indefinite & definite reference	81
2.5	Summary of the current experiments	87
<u>Chapter 3: The effects of listener 'familiarity'.</u>		90
3.1	Abstract	90
3.2	General Introduction	91
3.3	Experiment 1 - Children's use of indefinite and definite expressions as a function of the knowledge of the listener.	96
3.4	Experiment 2 - The effects of using the experimenter as the listener and comparing story-telling with discussions about real-life experiences.	120
3.5	Experiment 3 - The effects of story context on children's sensitivity to the listener's perception of referents.	135
3.6	General Discussion	157
<u>Chapter 4: The effects of referent 'specificity'.</u>		164
4.1	Abstract	164
4.2	General Introduction	165
4.3	Experiment 4 - The effects of the status of referents in the immediate context.	171
4.4	Experiment 5 - The effects of increasing the salience of referents in the immediate context	184
4.5	General Discussion	197
<u>Chapter 5: Reference in discourse production.</u>		200
5.1	Abstract	200
5.2	General Introduction	201
5.3	Experiment 6 - Expressions used in a story-telling task as a function of social and comparative factors.	211
5.4	Experiment 7 - Using information for unambiguous reference I.	229

5.5	Experiment 8 - Using relative clauses for anaphoric reference.	272
5.6	Experiment 9 - Using information for unambiguous reference II.	283
5.7	Experiment 10 - Model village: Identification and anaphoric reference.	302
5.8	Experiment 11 - Pronominal reference as a function of lexical and pragmatic factors.	327
5.9	General Discussion	344
<u>Chapter 6: Interpreting anaphoric reference</u>		357
6.1	Abstract	357
6.2	General Introduction	358
6.3	Experiment 12 - Comprehension of definite reference as a function of specificity and pragmatic context.	363
6.4	Experiment 13 - Comprehension of pronominal reference as a function of lexical and pragmatic ambiguity.	387
6.5	General Discussion	405
<u>Chapter 7: Interactions between linguistic and non-linguistic factors.</u>		413
7.1	Abstract	413
7.2	General Introduction	414
7.3	Experiment 14 - Understanding definite reference as a function of Principle C of binding theory and pragmatic context.	420
7.4	Experiment 15 - Story completions I: as a function of contextual information about available and/or plausible referents.	451
7.5	Experiment 16 - Story completions task II	468
7.6	General Discussion	485
<u>Chapter 8: Discussion</u>		493
8.1	Summary of results	493

8.2	General conclusions: Children's use and understanding of indefinite and definite reference	502
8.3	Evaluation of theoretical perspectives regarding indefinite and definite reference	
	A. Definition	505
	B. Processing	506
	C. Development	510
	D. Summary of Conclusions	514
8.4	A proposed model of processing anaphoric definite reference in young children	514
8.5	Suggestions for further work.	519
	<u>Annexe: Comments on experimental results where the robustness of ANOVA is uncertain</u>	522

APPENDICES: Verbal materials, data and analyses
summary tables.

APPENDIX A	EXPERIMENT 1	Tables A.1 - A.15	1
APPENDIX B	EXPERIMENT 2	Tables B.1 - B.5	14
APPENDIX C	EXPERIMENT 3	Tables C.1 - C.18	17
APPENDIX D	EXPERIMENT 4	Tables D.1 - D.21	33
APPENDIX E	EXPERIMENT 5	Tables E.1 - E.16	48
APPENDIX F	EXPERIMENT 6	Tables F.1 - F.10	59
APPENDIX G	EXPERIMENT 7	Tables G.1 - G.27	70
APPENDIX H	EXPERIMENT 8	Tables H.1 - H.13	91
APPENDIX I	EXPERIMENT 9	Tables I.1 - I.22	101
APPENDIX J	EXPERIMENT 10	Tables J.1 - J.39B	119
APPENDIX K	EXPERIMENT 11	Tables K.1 - K.10	149
APPENDIX L	EXPERIMENT 12	Tables L.1 - L.25	158
APPENDIX M	EXPERIMENT 13	Tables M.1 - M.17	175
APPENDIX N	EXPERIMENT 14	Tables N.1 - N.70B	191
APPENDIX O	EXPERIMENT 15	Tables O.1 - O.15	259
APPENDIX P	EXPERIMENT 16	Tables P.1 - P.13	281
BIBLIOGRAPHY			301

<u>Tables Volume 1-Chapters 3 to 7</u>		page
<u>Experiment 1</u>		
1.A	Mean indefinite and definite expressions on first mention	111
1.B	Mean scores for definite articles and pronouns on second mention	113
1.C	Mean indefinite expressions for main and subsidiary referents on first mention	115
1.D	Percentage indefinite and definite scores in Emslie's experiments and in this experiment	117
<u>Experiment 2</u>		
2	Counterbalanced order of presentation of conditions and pre-experimental session	122
2.A	Percentage indefinite and definite scores for the story-telling condition of this experiment as well as for the listener conditions of experiment one	130
<u>Experiment 3</u>		
3.A	Percentage definites by listener condition for each age group to referent in picture 2	149
3.B	Percentage definites to referent in P2 across the four story types	151
3.C	Percentage definites to referent in P2 across all stories	152
3.D	Indefinite and definite scores in each condition for reference to picture 1	153
<u>Experiment 4</u>		
4.A	Adult indefinite and definite scores for singleton items in each listener condition	176
4.B	Adult indefinite and definite scores for singleton items in each listener condition	177
4.C	Indefinite and definite scores for first identicals in each listener condition and in each age group	178

4.D	Indefinite and definite scores for second identicals in each listener condition and in each age group	180
4.E	Indefinite and definite scores for singletons in each listener condition and in each age group	182
<u>Experiment 5</u>		
5.A	Percentage indefinite and definite scores for non-specific items by each type of task and in each age group	190
5.B	Percentage indefinite and definite scores for last identicals by type of task and in each age group	191
5.C	Percentage indefinite and definite scores for singletons by each type of task and in each age group	193
<u>Experiment 6</u>		
6.A	Percentage of expressions on first mention for animates/non-specific items	222
6.B	Percentage of expressions on first mention for inanimates/specific items	223
6.C	Percentage of expressions on second mention for animates/non-specific items	225
6.D	Percentage of expressions on second mention for inanimates/specific items	226
<u>Experiment 7</u>		
7.A	Discriminative ability on second mention in Physical group	257
7.B	Discriminative ability on second mention in Episodic group	258
7.C	Percentage indefinite/definite scores on first and second mention for each age group	260
7.D	Percentage selection scores in 'Physical' group	262
7.E	Percentage selection scores in 'Episodic' group	263
7.F	Percentage distractor selection scores	265
7.G	Percentage listener choices as a function of discriminative scores from the speaker in the Physical group	266

7.H	Percentage listener choices as a function of discriminative scores from the speaker in the Episodic group	267
7.I	Ascriptions of blame from the listener for picture mismatching	269
<u>Experiment 8</u>		
8.A	Expressions used to answer questions in the 'same' condition.	277
8.B	Expressions used to answer questions in the 'different' condition.	279
8.C	Total 'same' and 'different' scores in each category	280
<u>Experiment 9</u>		
9	Pretest scores	287
9.A	Discriminative scores in each condition	291
9.B	Discriminative character scores: + and - relatives	293
9.C	Time to make a response	294
9.D	Discriminative and irrelevant colour and character scores (out of 6)	295
9.E	Contrastive expressions (different -gender)	297
9.F	Listener selection of referents	298
<u>Experiment 10</u>		
10.A	Mean percentage discriminative scores	313
10.B	Mean percentage discriminative scores for physical, quantity and spatio-temporal discriminations	314
10.C	Mean redundancy scores as a percentage of discriminative scores	315
10.D	Total redundancy scores across redundancy type (physical, quantity and spatio-temporal)	316
10.E	Spatio-temporal references: discriminative and redundant	317
10.F	Total appropriate and inappropriate pronominal reference	318
10.G	Pronouns and definite articles for referents in high focus	319

10.H	Predicted and unpredicted indefinite expressions	320
10.I	Mean display errors and mean proportion of errors due to ambiguity	322
10.J	Mean listener intervention scores	323
10.K	Listener intervention and process questions as a function of speaker ambiguity	324
<u>Experiment 11</u>		
11	Results of pilot study on adult undergraduates	337
11.A	Mean pronoun and name scores in each age group	339
11.B	Mean pronoun scores in each condition: lexical constraint and pragmatic associations	340
<u>Experiment 12</u>		
12.A	Mean choice scores for +/- pragmatic associations and +/- referential specificity conditions	376
12.B	Mean choice scores for R OR P (conflict condition). Scores for doll either specifically referred to or most plausible	379
12.C	Justification scores for +R+P, when choices were to doll specifically referred to and most plausible	380
12.D	Justification scores for +R-P, when choices were to doll specifically referred to	381
12.E	Justification scores for -R+P, when choices were to most plausible doll	382
12.F	Justification scores for ambiguity condition (-R-P) when correct response (silly sentence).	383
12.G	Justification scores for R OR P, when choices were to doll specifically referred to	384
<u>Experiment 13</u>		
13	Pretest ratings across all sentences with each object	394

13.A	Choosing first or second mentioned referent	399
13.B	Discriminative choice scores for +-lexical and +- pragmatic conditions	400
13.C	Discriminative choice scores for conflict condition (L OR P)	402
<u>Experiment 14</u>		
14.A	Non-identity scores as a function of pragmatic context	432
14.B	Non-identity scores in each context position	434
14.C	Unrestricted backward anaphora as a function of pragmatic context	437
14.D	Unrestricted backward anaphora scores in each context position: outward scores	438
14.E	Unrestricted backward anaphora scores as a function of pragmatic context: outward scores	440
14.F	Unrestricted forward anaphora scores as a function of position of context: outward scores	441
14.G	Yes sores as a function of pragmatics and context position for non-identity examples	443
14.H	Responses to question 1	445
14.I	Responses to question 2	446
14.J	Responses to question 3	446
14.K	Responses to question 4	447
<u>Experiment 15</u>		
15.A	Relative and non-relative completions as a function of both semantic and pragmatic context	460
15.B	Relative and non-relative completions as a function of pragmatic context for 'with' and 'that' stories	462

Experiment 16

16.A	Relative and non-relative completions as a function of both the semantic and pragmatic context for 'with' constructions	475
16.B	Relative scores as a function of context (no relative, 'that', 'who' or 'which' relatives) for 'with' constructions	477
16.C	Relative and non-relative completions as a function of both the semantic and pragmatic context for 'that' constructions	479
16.D	Relative and non-relative completions as a function of both the semantic and pragmatic context for 'that' constructions, looking at questionnaire and verbal presentation separately: adult group	482

<u>Figures Volume 1-Chapters 3 to 7</u>		page
<u>Experiment 1</u>		
1.1-1.8	Picture stories	100
<u>Experiment 2</u>		
2.1	Histogram presenting reference scores in each condition on first mention	129
2.2	Histogram presenting reference scores in each condition on second mention	132
<u>Experiment 3</u>		
3.1A-3.4B	Picture stories	140
<u>Experiment 4</u>		
4	Experimental layout	172
<u>Experiment 6</u>		
6.1-6.4	Picture stories	215
<u>Experiment 7</u>		
7	Layout of experiment	250
7.1A-7.3B	Picture stories- Physical condition	234
7.4A-7.6B	Picture stories- Episodic condition	240
7.7A-7.8B	Picture stories- Distractors	246
<u>Experiment 10</u>		
10.1	Experimental materials and layout	309
<u>Experiment 11</u>		
11.1	Labelled picture of the four individuals serving as referents	332
11.2	Picture showing 'second reference'	333
11.3	Picture showing 'third reference' (neutral pragmatics)	334
<u>Experiment 12</u>		
12.1	Pictures on cards as experimental materials	366
<u>Experiment 13</u>		
13	Experimental layout	395
13.1	Pictures on cards as experimental materials	389

I declare that the work in this thesis is all my own and has not been submitted for any other degree.

The copyright of this thesis rests with the author. No quotation from it should be published without her prior written consent and information from it should be acknowledged.

Acknowledgements

I would like to thank Rosemary Stevenson for her helpful supervision throughout this research. Also, many thanks to David Kleinman, Marion England, Rosalind Crawley and Charles Crook of the Durham psychology department for their useful advice and encouragement. I am further grateful to Malcolm Rolling and Chris Mullaney, also of the Durham psychology department, for their technical support.

I particularly wish to thank those head teachers and staff of local schools, as well as parents and pupils, who were cooperative and supportive during my gathering of data.

Lastly, I thank Mustafa Chaidamli, Matthew Sims and Mandy Sims whose extra support helped towards the completion of this thesis.

I acknowledge support of the S.E.R.C and the University of Durham.

Children's Knowledge of Indefinite and Definite Reference

K. Sims

Abstract

After reviewing previous theoretical and empirical investigations, the decision is made that an understanding of both 'Familiarity' and 'Specificity' must be necessary for a full appreciation of indefiniteness/definiteness. Moreover, it is argued that knowledge of 'Specificity' requires an ability to co-ordinate and integrate a number of linguistic and non-linguistic factors appropriately and for this reason, it is suggested that young children's knowledge of indefinite/definite reference is incomplete.

16 experiments are reported, most of which were designed to investigate children's knowledge of 'Familiarity' and 'Specificity' for indefinite/definite reference. Other areas of investigation include children's ability to use both linguistic and non-linguistic factors in their understanding of definite reference and the way in which older children and adults use these factors for anaphoric reference.

The results showed that whilst children from 3½ years of age are sensitive to the perceptual information given to listeners and will use their general knowledge to define 'Familiarity', they do not appreciate the need to consider the relative status of a referent in their choice of expression. It seems that young children will only consider 'Specificity' when the relative status and/or the identifying features of a referent are salient to them.

The failure of the young child to appreciate the necessity of 'Specificity' for indefinite/definite reference may be explained as due to a failure to consider more than one possible interpretation of reference. By contrast, from about 8 years of age, children appear to be integrating contextual information about the number of available referents with their knowledge of linguistic constraints on reference. A model of the processing of definite anaphoric reference in children is proposed based on the 'weak' interactionist view of the effects of informational context on choice of definite construction.

CHAPTER 1

INTRODUCTION

The aim of the current research was to investigate children's understanding and use of indefinite and definite referring expressions. After looking at the different philosophical and linguistic definitions of indefiniteness/definiteness, (subsequent mentions of this term will be shown as in/definite(ness)), it was decided that the criterion for accepting that children have knowledge of indefinite and definite reference should be that they show an appreciation of the significance of two factors. These factors are FAMILIARITY and SPECIFICITY. With regards to recent linguistic and psychological approaches to anaphoric reference in discourse, it was decided that a model for unambiguous anaphoric reference must be one which considers an interaction between linguistic and non-linguistic factors. Furthermore, after considering evidence from cognitive and linguistic developmental theories, it was suggested that DEICTIC reference (referring to salient objects) precedes PHORIC reference (EXOPHORIC: referring to an object in the perceptual array or AN/CATAPHORIC: to a concept in the preceding (anaphoric) or succeeding (cataphoric) linguistic context). This is because the latter, unlike the former, requires an interaction between a number of linguistic and non-linguistic factors. The following section looks at THEORETICAL approaches to 1, the definition of in/definiteness and 2, in/definite representation and processing in discourse. The subsequent section looks at



DEVELOPMENTAL approaches from the perspective of 1, cognitive development and language acquisition in general and 2, indefinite and definite reference in particular.

1.1 THEORETICAL LITERATURE

A. Definitions of in/definiteness

Definitions of in/definiteness have been considered from both syntactic and semantic perspectives. However, in order to generalise a definition of indefinite and definite expressions to reference in discourse and so that current linguistic theories are compatible with traditional philosophical approaches, it was decided that 'semantic' theories should provide the general background for deciding on criteria for evidence of knowledge of indefinite and definite reference. Most of the literature has focused on the articles: A and THE. In fact, until recently most attention was given to the singular definite article (the).

The traditional philosophical definitions of definiteness have arisen within the domain of logic: defining definiteness as a function of the truth and falsity of sentences. More recent philosophical work has leaned towards pragmatics: defining definiteness as a function of the illocutionary and social act of reference. Recent linguistic work has aimed to combine logical and pragmatic perspectives into an integrated theory of both indefiniteness and definiteness and such definitions tend to regard logical distinctions between indefinite and definite expressions as being a necessary consequence of their different pragmatic functions.

The brief summary of the research literature that is described below is an attempt to trace the foundations for the decision that the basic principles of in/definiteness are FAMILIARITY and SPECIFICITY.

An early 'logical' definition was proposed by Russell (1905), who regarded the definite article as essentially denoting UNIQUENESS. He argued that certain propositions needed to be true for a sentence such as 'The King of France is bald' to be true. These propositions were defined by Russell as being:

- a. Existence: There IS a King of France.
- b. Uniqueness: There is not more than ONE King of France.
- c. This individual is bald.

It was some years later that this view was challenged by Strawson (1950), who replaced Russell's Falsity theory with his own Presupposition theory. Basically, he argued that Existence and Uniqueness are presupposing principles and thus bear no relation to the question of truth and falsity. According to Strawson, only Russell's last proposition: that the King of France should be bald, was regarded as depending on the question of falsity, because this latter principle, unlike the others, was asserting rather than presupposing. Moreover, Strawson argued that EXISTENCE and UNIQUENESS were both illocutionary acts of asserting for reference, despite the fact that the propositions themselves were presupposing. The problem with Russell's theory was that he failed to appreciate that denotation was different from referential

uses of definite descriptions (Donnellan 1966).

One 'pragmatic' approach has led to the conclusion that the fundamental definition of the definite article is FAMILIARITY (Christopherson 1939, Jespersen 1949). Christopherson argued that 'the' is used when it can be assumed that the listener is knowledgeable of the referent. Christopherson also listed a number of 'usage types' of the definite article to demonstrate the significance of listener 'familiarity' for definiteness. Christopherson defined the indefinite article as UNITY, which applies only to the singular indefinite article, rather than to plural indefinites such as 'some'.

A strong 'pragmatic' view was proposed by Searle (1969), who attempted to define the propositions of Existence and Uniqueness in 'speech act' terms. Searle argued that there were 2 axioms of the singular definite article: EXISTENCE and IDENTIFICATION. Searle incorporated both existence and uniqueness within his axiom of 'Existence' by defining it as meaning that 'there is one and only one X'. With regards to the axiom of 'Identification', he implied that the hearer must be able to identify the referent from the speaker's expression. Moreover, Searle argued that identification must not be ambiguous for the hearer in that particular context if the hearer is to identify it successfully.

Thus, the main definitions of definite reference up to this point have been EXISTENCE/UNIQUENESS (basically defined by those studying logic) and FAMILIARITY/IDENTIFICATION (mainly from pragmatic

theories). Hawkins (1978) came up with a theory of indefinite and definite reference that served to unify this philosophical divide. He did this by proposing that definite descriptions depended on 'set identification' and 'location'. As an extension of Christopherson, he defined 8 'usage types' of the definite article to show that all usage possibilities imply that all the objects in the shared set (of the speaker and listener) satisfying the descriptive predicate are being referred to and none are being excluded. Thus, the only difference between each usage type was the type of situation which defined the shared set of the speaker and listener. The following presents Hawkins' main usage types together with the situations which define the shared sets of each:

<u>USAGE TYPE</u>	<u>SHARED SET</u>	<u>EXAMPLE</u>
Anaphoric	Entities from earlier discourse	'A car... The vehicle
Immediate situation	Entities salient to both speaker and listener in the current situation	'Pass the salt'
Larger situation	Entities known by both speaker and listener through either: a. Shared knowledge: (shared experience) b. General knowledge:(cultural)	'The dog': (our dog) 'The sun'
Associative anaphoric	Shared knowledge and/or general knowledge about entities from earlier discourse.	'A room.. The door' (ie. door of

Cruse (1980) argued that these usage types were also in order of precedence: He argued that lower ranking types would only be used if there were no competitors. His order of precedence was 1. ANAPHORIC/ASSOCIATIVE ANAPHORIC - 2. IMMEDIATE SITUATION - 3. LARGER SITUATION. According to this interpretation, whenever it was possible, the 'shared set' would be the discourse context of the discourse participants.

Nevertheless, for each of these usage types, the definite article refers to all those entities in the 'shared set' for which it describes. Hawkins defined the definite article as involving 3 speech acts:

A. Introducing a referent to the listener.

B. Instructing the listener to LOCATE the referent in some shared set.

(Lyons (1979) also included a 'looser' usage of Location: The definite description may also inform the listener that the referent is part of a particular set, even when the referent is not already part of a set).

C. Referring to the totality of the objects or mass within this set which satisfies the referring expression (INCLUSIVENESS).

Hawkins preferred the term LOCATION as opposed to 'familiarity', because he argued that the knowledge of the

listener was not necessarily one of knowing the object of reference per se, but rather was knowledge of a concept/object within either a situational, discourse or prototypical context. Hawkins preferred the term INCLUSIVE as opposed to 'uniqueness', because the totality of a shared set that satisfies the definite description is not necessarily one (unique) referent. For example, '2 men were waiting...THE MEN'. 'The men' refers to all the entities that satisfy the description 'men', which in this case is both of the men introduced in the previous sentence.

Hawkins argued that the difference between indefinite and definite articles was not a pragmatic one and he provided examples to show that both referred to 'shared sets'. He claimed instead, that the in/definite distinction was of a logical kind. Whilst definite articles were inclusive, indefinite articles were EXCLUSIVE: referred to a proper subset of the entities that satisfied the indefinite description. Thus, given the sentences below:

'There were cats and rabbits in the window. After giving it some thought Jenny decided to buy a cat'

Jenny's decision was to buy any one of those cats represented in the 'shared set' (in this case those introduced earlier in the discourse). Also, as with definite reference, indefinite references may be referring to more than one entity, in which case the indefinite

plural, 'some' would be used. However, according to Hawkins, both singular (a/an) and plural forms (some) are exclusive: neither refer to the totality of objects in the set which fit the indefinite expression.

Furthermore, Hawkins concluded that the reason why indefinite and definite descriptions were distinct logically rather than pragmatically was really a reflection of pragmatic requirements: that definite articles must be used for locating a referent, whereas indefinites need not involve any location of referents. Thus, the fact that the pragmatic 'location' requirement is necessary only for definite reference means that it is only the latter which requires the strict criterion of 'inclusiveness'.

However, Lyons (1979) argued that, whilst many definite references must be inclusive, the crucial aspect of definite reference that distinguishes it from indefinite reference is that the former is UNAMBIGUOUS. For example, the statement 'close the door' would be appropriate in a context where there are 3 doors and where only one of them is open, despite the fact that the definite reference is not inclusive in such a situation. Thus, Lyons claimed that definite reference must be unambiguous even when it is not inclusive. He argued that reference may be unambiguous because a referent is 'salient' in relation to other entities which satisfy the description, by some feature of the utterance or situation.

Given that FAMILIARITY does not necessarily mean that the object of reference itself has to be familiar to the

listener in the strong sense of the term, the use of 'familiarity' in the current investigation means that, when the listener can locate the referent in the particular context in which it is being used, a definite expression may be used. By contrast, when a referent is unfamiliar, it is argued that definite reference cannot be used and thus, indefinite reference is necessary. For that dimension of definite reference which has been labelled as either uniqueness, inclusiveness or non-ambiguity, the term preferred in the current investigation is SPECIFICITY. The reason for this definition is because 'specificity' seems to cover the idea that, if definite expressions refer to the totality of potential entities (inclusiveness), the listener knows exactly which particular referents are intended, whereas when only a subset of potential entities are referred to, the listener need not be aware of the particular referents, (the listener does not know which referents are intended. All they know is that some of them are being left out/excluded). Moreover, the question of ambiguity does not apply to indefinite reference because in a sense, indefinite references are always ambiguous owing to their non-specificity. However, ambiguity is relevant for definite reference, essentially because it is referring specifically: Ambiguous reference would imply that definite reference was non-specific and could be referring to other entities not intended by the speaker. Thus, knowledge of both 'familiarity' and 'specificity' has been accepted as the criterion for measuring indefiniteness and

definiteness in the experiments of the current investigation.

However, while a referent needs to be both familiar and specific for appropriate definite reference, an indefinite expression need not refer to an unfamiliar referent and it need not refer to a non-specific referent. For example:

1. Look at the 3 puppets on the stage. Will you get a puppet (one of the puppets) and bring it to me.
2. I gave the flowers to a lady in hospital. Her name was Betty Woods.

Example one, shows indefinite reference (a puppet) when the puppet is clearly familiar to the listener: The listener knows that it is on the stage. Example two, used indefinite reference to refer to a specific lady: Betty Woods rather than any of the other ladies in hospital. Nevertheless, indefinite reference must EITHER be referring to an unfamiliar referent OR a non-specific one. Thus, for example one, the indefinite article is used because reference is non-specific even though the referent is familiar to the listener (referring to any of the 3 puppets on the stage), whereas for example two, the indefinite article is used because the referent is unfamiliar to the listener, despite the fact that the speaker is referring specifically to Betty Woods.

However, BOTH 'familiarity' PLUS 'specificity' are

essential for appropriate definite reference. Taking the examples below:

1. Look at the 3 puppets on the stage. Will you get the puppet (one of the puppets) and bring it to me.

2. I gave the flowers to the lady in hospital. Her name was Betty Woods.

Neither satisfies the requirements for definiteness, because for the first, there is no specific referent and for the second, the listener does not know which lady is being talked about: Although the listener is subsequently told the recipient's name, the listener does not know who Betty Woods is.

Thus, the terms 'FAMILIARITY' and 'SPECIFICITY' have been defined as the essential factors necessary for determining knowledge of indefinite and definite reference, because the use of 'FAMILIARITY' here covers both Christopherson's 'familiarity' and Hawkins' 'shared sets' and the current use of the term 'SPECIFICITY' covers all the previous definitions, including 'uniqueness' (Russell, Strawson), 'existence' (Searle), 'inclusiveness' (Hawkins) and 'non-ambiguity' (Lyons).

B. Representation and processing of in/definite reference in discourse

As the linguistic constraints of definite reference are insufficient for determining indefinite and definite reference in discourse, more recent linguistic

investigations have centred on non-linguistic determinants of anaphoric reference, such as pragmatic inferences, speech acts and textual structure. Whilst these non-linguistic approaches can account for the FAMILIARITY factor of in/definiteness, it seems that the SPECIFICITY factor can only be accounted for by using a discourse model approach. Some recent processing theories are also described, which aim to account for the way that pragmatic factors interact with linguistic constraints (Marslen-Wilson and Tyler 1980, 1982) and the way that the particular model of the discourse interacts with linguistic factors (Crain and Steedman 1985, Altmann and Steedman 1988). A proposal for a more integrated processing theory is also discussed.

Chomsky's 'Government Binding Theory' (1981) lays down these three binding principles of definite reference.

- A. A reflexive (eg 'herself') is BOUND within its governing category.
- B. A pronominal (eg 'she/her') is FREE within its governing category.
- C. An r-expression (eg 'The girl, Susan') is FREE everywhere.

where bound means must be co-indexed with a noun phrase that is higher in the phrase structure tree (bound by a c-commanding noun phrase), free means must not be co-indexed with a c-commanding noun phrase and governing category means local domain or clause.

Therefore, a reflexive pronoun must have an antecedent which c-commands it in the same clause in which

it occurs. In the following example:

John told Peter to lock himself in.

the reflexive must refer to Peter. A pronoun cannot refer to a c-commanding antecedent in the same clause in which it occurs and therefore, in the following sentence:

John told Peter to lock him in.

the pronoun cannot refer to Peter but can refer to John. Finally, a definite noun phrase cannot have an antecedent that c-commands it. In the following example:

He wished that John would open the door.

John cannot be co-referential with 'he' because John is c-commanded by the pronoun. This is often referred to as the 'non-identity restriction.'

However, in the following example:

Jane told Peter that he had a letter.

the pronoun cannot be assigned through syntactic means because both Jane and Peter are outside the clause of the pronoun. However, 'he' cannot refer to Jane because an additional feature of pronouns is that they are marked lexically. In this case it is gender marking (he / she) which disambiguates reference, but pronouns are also marked

for number (he/ they) and case (he / him).

Nevertheless, if the sentence above was preceded by a different sentence, such as:

John was delivering his friend's mail.

John told Peter that he had a letter.

then, even when syntactic and lexical constraints are considered, the pronoun can refer to either John or Peter. However, given the context sentence that John was delivering his friend's mail, one can infer that the pronoun would be more likely to refer to Peter than to John: Peter is more plausible because John is delivering mail rather than receiving it.

Thus, many investigators have looked at discourse comprehension from a perspective which places primary emphasis on pragmatic contexts and on the way that anaphoric reference makes use of general knowledge inferences when finding an antecedent. Moreover, non-linguistic factors for co-reference have received increasing interest from both linguists and psychologists.

Referring again to the significance of pragmatic context, Moore (1986) has argued that the interpretation of anaphoric reference involves uncertainty: One has to consider the most probable, rather than the correct interpretation of the anaphor.

Clark and Haviland (1974, 1977) put forward the notion of 'Bridging Inferences': that finding an antecedent involves a 'memory search' of previous discourse for making

an inference. However, more recently, investigators have shown that context determines reference even before the need for a search arises in the discourse (Hirst and Brill 1980, Stevenson and Vitkovitch 1986) and that pragmatic inferences do not cost any more in processing time (Garrod and Sanford 1977, Marslen-Wilson and Tyler, 1980). Thus, recent studies have concentrated on the primary role of context in guiding the interpretation of discourse.

Therefore, the evidence suggests that, just as non-linguistic factors are significant for deictic reference, where there are cues enabling the listener to use both the situation and previous experiences of situations, anaphoric reference similarly requires a construction of a representation that makes it congruent with the hearer's view of the world.

There have been a number of theories of discourse representation which focus on the way that context is used in discourse. This general knowledge representation that is claimed to be relevant to discourse has been called 'schemas' (Bartlett 1932, Rummelhart and Ortony 1977) 'frames' (Minsky 1975) and 'scripts' (Shank and Abelson 1977, Sanford and Garrod 1981). These are names that have been used to describe the way that default prototypes are represented in discourse and are used when there is no other information available. Artificial Intelligence researchers have often modelled these 'scripts' etc. in computer programs. Moreover, some linguists have attempted to model anaphoric processing (particularly

pronominal anaphora) as part of the process of general discourse comprehension and these models have therefore, been attempts to account for anaphora in pragmatic and contextual terms (Charniak 1972, Reiger 1974, Hobbs 1976, 1978, Nash-Webber 1978, Sidner 1979, Carter 1983). For example, Hobbs used the principle of 'knitting' to account for inferences in discourse that depended on the redundancy of concepts which were not explicitly mentioned in the text. Carter used the term 'indirect co-reference' to account for anaphors co-referring with earlier evoked items, as well as for items evoked by the anaphor co-referring with earlier explicit items and recently Carter has represented context in a form of semantic network.

Another form of non-linguistic knowledge which has been of interest has been conventional knowledge about the way that participants of communication should behave. Austin (1962) emphasised the significance of the illocutionary force of utterances: the intentions of the speaker or 'speech acts'. Grice (1975) postulated 4 maxims for a contract of interaction: QUANTITY, QUALITY, RELATION (be relevant) and MANNER (avoid obscurity and ambiguity). Moreover, Clark and Marshall (1981) have concentrated on the way that MUTUAL knowledge is required to make dialogue felicitous. Mutual knowledge implies that people use these conventional heuristics in discourse and can assume that their partners will be aware of these conditions. For example, the speaker can assume that the listener is knowledgeable from the prior conversation, common community

and/or from some physical presence.

Yet another non-linguistic feature of discourse which relates to in/definite reference has been the way that the discourse is structured. H.H. Clark (1977) put forward the 'Given-new contract' which basically states that a sentence is syntactically divided into old (presupposed) elements and new elements and the former must have an antecedant: Anaphoric reference is referring to something introduced in the previous text. Others have divided sentences into topic: known information and comment : unknown information (Grieve 1973, Grieve and Wales 1973, Wright and Glucksberg 1976, Crawley 1982). Grieve showed experimentally that 'known' persons and objects were marked for definiteness and he related this to the notion of topic.

Thus, the predominant focus of discourse processing for anaphoric reference has been on non-linguistic social and pragmatic factors. This concern is consistent with that part of the above definition of in/definite reference which has been labelled FAMILIARITY. However, SPECIFICITY cannot be accounted for solely by these 'pragmatic' approaches. Instead, 'specificity' can be shown by using discourse models to represent discourse. Johnson-Laird (1978, 1980) and Johnson-Laird and Garnham (1980) put forward the notion of a 'mental model', which basically is a hypothetical representation of entities introduced in discourse, where the entities in the model are said to be structurally identical to the referents that are being represented.

True discourse requires that the speaker considers the listener's model. Definite reference is defined as finding a unique token in the model, whereas indefinite reference requires introducing into the model an arbitrary number of tokens or finding in the model any token that satisfies the description.

Murphy (1984) empirically tested the reality of discourse models with adults and found that indefinite expressions and references to more than one object in the model took longer because they required introducing something into the model (indefinite) or finding more than one entity within the model (plural references).

Stevenson and Crawley (unpublished manuscript) used a mental model framework to account for the way that pronouns referred to referents in explicit focus in the current discourse model. Emslie (1986) has actually incorporated Hawkins' location theory into a mental model representation of indefinite and definite reference by equating a mental model with the speaker's and listener's 'shared sets'. (Hawkins' anaphoric and associative anaphoric usage types).

Thus, the most recent linguistic definition of in/definiteness (shared sets), which appears to require some way of representing the social and pragmatic context, plus the psychological approach which requires a representation of non-linguistic aspects of discourse (discourse models) are both consistent with the idea that definite reference must be 'specific'. It appears that discourse models can account for indefinite and definite distinctions because they incorporate one

other non-linguistic factor in addition to pragmatics: Discourse models emphasise the 'relative - comparative' features of referents in the discourse context, that is they focus on the 'specificity' of referents in relation to other potential referents in the model.

Therefore, a theory of discourse processing requires a consideration of the way that non-linguistic factors work together with linguistic factors in comprehending discourse. There is some empirical evidence that adults do show an interaction of linguistic and pragmatic factors when interpreting discourse: syntax and pragmatics (Hirst and Brill 1980), lexical gender marking and pragmatics (Caramazza, Grober, Garvey and Yates 1977, Marslen-Wilson 1980) and discourse structure and pragmatics (Marslen-Wilson). In fact, Marslen-Wilson 1980, Marslen-Wilson, Levy and Tyler 1982, Marslen-Wilson and Welsh 1978, and Tyler 1981, have proposed an 'on-line interactive' theory of discourse processing which accounts for the way that each word is comprehended in relation to preceding discourse context. This theory suggests that all linguistic and pragmatic processes occur concurrently and anaphoric mapping processes are not a special case but involve the same procedures as normal discourse processing.

Their main argument was that pragmatic checking plays the predominant role in discourse interpretation. In a number of experiments they showed that pragmatic inferences required no extra processing time compared to situations where lexical or thematic factors could disambiguate reference. For example, in one experiment,

the investigators presented adult subjects with short texts followed by an incomplete sentence and a probe word ('Him' or 'Her'). The subject's task was to name the word as quickly as possible after the presentation of the probe. It was found that reaction times were consistently slower when the probe word was inappropriate, even when the only information available was pragmatic. Similar conclusions were made regarding pronoun production, which they tested using a quasi-experimental picture-story task. However, opposite results have also been found : Shillcock (1982) used a task where subjects had to make a lexical decision at certain points throughout a text. It was found that comprehension of test words that were close associates of the referent of the pronoun were faster immediately after the pronoun was read as opposed to when the test words were presented later in the text, despite the fact that it was only some time after the presentation of the pronoun that pragmatic inferences were possible. This result suggests that pragmatic checking is not always necessary when processing discourse.

Moreover, Goodluck (1988) proposed that parsing discourse information required separate systems working sequentially (cf. Forster 1979), rather than a parallel system. She claimed that lexical and syntactic analyses occur before this information is integrated into a representation of the entire discourse.

However, recently Crain and Steedman (1985) and Altmann and Steedman (1988) have proposed an on-line interactive model of anaphoric reference which suggests

that context interacts with linguistic factors, despite their claim that these factors are logically modular. Moreover, this interaction is particular to anaphoric mapping and unrelated to discourse processing in general. They looked at sentence constructions which contained a linguistically ambiguous word. For example:

The psychologist told the woman that

The word 'that' is ambiguous with regards to whether it could be followed by a complement clause (he was having trouble with her husband) or a relative clause (he was having trouble with to visit him again). Frazier (1978) found evidence to show that the relative clause version engendered 'garden path effects' and she explained this in terms of a linguistic heuristic known as 'minimal attachment'. However, Crain et al proposed a processing theory which they called a 'weak' interactive model because the influence of context was to dispose among all possible syntactic analyses. Moreover, Crain et al argued that the specific context of interest must be within a discourse model. From their theory, they predicted that garden path effects would depend on the 'referential support' of the specific linguistic context: If the context contained only one potential antecedent, then they expected a garden path with a relative clause (The psychologist told the woman that he was having trouble with to visit him again), whereas if the context contained more than one potential antecedent, they predicted a garden

path with a complement clause (..... that he was having trouble with her husband). This is because restrictive relatives serve to disambiguate reference when there is more than one referent available. Using 'Grammatical judgement tasks', the investigators' hypotheses were confirmed: Expected linguistic analyses were a function of the number of referents available.

However, there is also conflicting evidence: Ferriera and Clifton (1987) found that context did not affect syntactic 'garden pathing'. Nevertheless, the latter investigations looked at pragmatic context rather than the specific conversational context.

Thus, there does seem to be a difference between context used for general discourse interpretation (coherence relations) and the type of context used for anaphoric reference (specific co-reference). Perhaps an ideal processing model of anaphoric processing should incorporate both 'comparative-relative' contextual factors and pragmatic factors. If pragmatics and the relationship between referents in the discourse context interact, then one might expect a complement clause when there are 2 potential referents, in those situations where pragmatic inferences disambiguate reference. Investigations of this proposal may help to reveal the type of linguistic/non-linguistic interactions that occur for anaphoric mapping.

1.2 DEVELOPMENTAL LITERATURE

A. Cognitive Development and Language Acquisition research

This section has been divided into subsections. Initially, evidence for children's use of general knowledge, particularly studies on children's use of pragmatic inferences is discussed. This is followed by studies of children's social development and evidence concerning the child's early interactive abilities. The third subsection looks at children's use of social-pragmatic knowledge in linguistic tasks and discusses a number of developmental theories which have attempted to account for the empirical findings.

1.2.A1 General knowledge and pragmatic inferences

Lyons (1981) defined reference as "the relation that holds between linguistic expressions and what they stand for in the world or universe of discourse". As this definition shows, reference is impossible without a knowledge of the objects and individuals that are being talked about.

The studies described in this section are conflicting with regards to the young child's use of pragmatic inferences and general knowledge in understanding indefinite and definite reference. However, it seems that in those tasks where the child can correctly interpret the instructions, young children will make use of pragmatic knowledge. In fact, there is evidence to show that they will overuse their general knowledge by making pragmatic inferences in situations which do not require them.

Piaget's theory stresses the primary role of the object world in early cognitive development. Another of Piaget's claims was that the young child's inferential abilities were not acquired until later, when the child had reached the 'concrete operational stage' (approximately 6 to 7 years).

However, there is empirical evidence to show that young children can make 'pragmatic inferences' (inferences using knowledge about the world) as opposed to 'logical inferences' (inferences about propositions). For example, Macnamara, Baker and Olson (1976) gave 4 year olds questions about stories and found that they were able to deduce implicit information from the story details. Harris (1974) found that young children could make inferences about nonsense syllables when they were paired with categories, such as: 'If a 'nib' is a man then a 'nib' can eat'. Other experiments that involved asking questions about inferences have produced similar results (Kail and Lèveillé 1977, Paris and Upton 1976).

Other investigations have shown children using 'probable event strategies' (Strohner and Nelson 1974) and 'referential event strategies' (Gowie and Powers 1978) in preference to syntactic 'actor-action-object' strategies and thus, they were more likely to interpret non-reversible passives as they did active constructions and to reject improbable constructions such as 'The bunny chased the dog'.

Moreover, those studies which have produced negative findings have been interpreted as due to difficulties

extraneous to inferential competence itself. For example, Paris and Lindauer (1976) found that, between 7 to 9 years of age, children were better at recalling explicitly mentioned items than implicitly mentioned items. Nevertheless, when the children were asked to read the sentences aloud and to act out the actions, they were just as able to recall implicit information. The investigators concluded that age differences in inferential abilities are due to "older subjects recognising the value of the inferential process and....the deliberate employment of such strategies". Thus, age differences appear to be due to strategy production deficiencies rather than due to inferential competence per se.

Furthermore, Gelman (1978) found that young children could make inferences about antecedent states, causes and outcomes when the causal relationships were well understood by the children. As Stevenson (1988) argued, any age differences may be due to the fact that young children are less likely to have had enough experience of the events to know what can be inferred.

Johnson and Smith (1981) and Ackerman (1984) found that, whilst young children have the basic inferential competence, they are limited as to the range of situations in which they can use this ability. For example, they have more difficulty when they have to integrate two premises which are embedded within a larger story or which are presented in separate paragraphs. Schmidt, Schmidt and Tomalis (1984) found that 4 year olds used inferences less efficiently and they were based on extra-story information.

They found that older children would use story cues appropriately and would reinterpret the story when the clues followed anomalous information, such that inferences were continually reorganised throughout the story setting. Similar results were found by Ackerman (1988).

Furthermore, Wykes (1981) claimed that, although young children could make inferences, they failed to do so in interpreting pronouns. She presented sentences such as 'I put the pencil in my pocket and because it had a hole it fell out'. Wykes found that 4 year olds were relatively poor at interpreting the sentence to mean that the pocket had a hole and the pencil fell out, because they were more concerned with the way that the sentence was structured. However, elsewhere in this thesis (chapter 6), it is argued that Wykes' experiments were possibly confounded by factors such as failing to test explicitly for the type of inferences that were actually required.

Piaget also claimed that the end of the sensorimotor stage of cognitive development marked the beginning of the 'semiotic' function: the ability to dissociate the concept of an object from the actual object itself. A similar point was obviously being made when Piaget used the term "word realism": that children tend to treat names as inherent properties of objects. Lyons used the label "protoreference" to account for a similar claim, that children fail to dissociate the linguistic terms used for reference from the referent itself. In fact many studies do show that, for young children, meaning is

more profoundly embedded in context than it is for adults (Bloom 1970, Halliday 1975). Bloom found that the speech of the young child can only be understood by reference to the context of the situation and concluded that it can be maintained that language development consists of progressively freeing the system from dependence on situational constraints. Halliday carried out a longitudinal study on his own son 'Nigel' and argued that initially children are "instrumental" because meaning and context are undifferentiated for them. Ackerman (1981), argued that, children of 5 and 6 years of age were poor at interpreting ambiguous communications in referential tasks because, they respond to the illocutionary 'performative force of their utterances as opposed to the locutionary 'propositional force' of the expressions they use, that is, they are more concerned with selecting a referent than with reflecting on the referential expressions used by the speaker.

Karmiloff-Smith (1980, 1985) claimed to have found evidence of children under 7 years of age using 'pragmatic procedures' in their interpretation and use of nominal determiners. Karmiloff-Smith found that, after telling young children a story about a garden, when she referred to 'the flower', the child usually interpreted this as referring to more than one flower, owing to the fact that one normally finds more than one flower in a typical garden. Therefore, children misinterpreted the 'specificity' of definite reference in favour of an interpretation which was more familiar with their

experience of everyday situations. She concluded from this that the reason why children under 7 years of age are influenced by the way the experiment is presented to them is because they are overimpressed by speech acts rather than using 'metaprocedures' for coming to terms with the language as a problem per se.

Donaldson and Lloyd (1974) found that, when young children were asked 'Are all the cars in the garages?', they answered 'no' when there was an empty garage. This result was interpreted as due to the child's bias to fixed objects, as a function of their pragmatic experiences of the canonical function of garages: that they should contain cars. This experiment is an example of young children misinterpreting the requirements of a task as a result of their concern to make 'human sense' (Donaldson, 1978) of the situation.

To summarise the evidence in this section, it seems that when a task is not confounded by factors unrelated to inferential knowledge, even 3 and 4 year olds will make pragmatic inferences. In fact, there is evidence to suggest that they will make pragmatic inferences even in tasks/situations which do not require them. Thus, there is strong evidence of the young child being overimpressed by general knowledge in interpreting linguistic information.

1.2.A2 Interpersonal abilities for language acquisition

Searle (1969) defined referring expressions as, any expressions which serve to identify any thing, process, event, action or any other kind of 'individual' or

'particular'. The crucial word in this definition is 'identify'. Identification is only functional if reference involves the participation of at least two individuals: a speaker and a listener.

Research into children's early interpersonal abilities is also conflicting. However, in contrast to those studies which have pointed to the 'egocentricity' of the young child, much recent work has focused on the strong social role of young children interacting with their parents. It seems that young children are not as 'egocentric' as was originally thought. However, their sensitivity to others is still not as developed as for older children and adults.

In Piaget's 'the child's construction of reality' (1958), he emphasised that mental development was the differentiation of 'mental states' from 'world states'. According to Piaget, in contrast to young children's strong interest in the objects around them, they fail to appreciate that others have different perspectives from their own. Piaget's writings continually refer to the young child as 'egocentric': unable to take another's view into account.

In sharp contrast to this negative view of the young child's social abilities, there has recently been a surge of studies that have provided evidence to the contrary of Piaget's claims. For example, observational studies have shown that pre-linguistic infants secure the joint attention of an object with their parent by pointing (Bruner 1975, 1976, Bates, 1976). Bates found that at 12

months of age, infants used pointing to ensure that they and their interlocutor were attending to the same object. Bruner found the beginnings of a locution-dependent reciprocal concept (I/You) emerging in action well before it was ever used in formal language. He argued that the highly ritualized play between parent and child was evidence of a "Language Acquisition Support System (LASS)" that was necessary to enable the Chomskian "Language Acquisition Device (LAD)" to function. He found that language occurred in the context of an 'action dialogue' in which joint action was being undertaken by infant and adult. Bruner also observed what he called the early 'inter-subjective functions' of children's play. One of his subjects, Carlotta, was observed to go through a series of steps as she focused her attention alternately between her mother and an object. First she pointed to the object, then to the adult and then to the object again. The child seemed to be putting together in a chained form, the components that eventually form the smooth deictic act of simultaneously pointing at an object while turning to the other for confirmation. This initial interest in the interpersonal domain may also account for the child's early mastery of the personal pronoun system and particularly the first and second person pronouns, which include reference to the shifting roles of speaker and hearer.

Trevarthen (1979) also found that the child first behaved in ways that elicited responses from the parent and that only subsequently did they show any evidence of beginning to understand the physical world.

Oshima-Takone (1988) found evidence of young children imitating adult's pointing gestures and using the terms 'Me/You' even when the children were not actually being addressed themselves. This evidence suggests that young children are interested in their social environment and will initiate communication even without elicited social interaction from the parent.

These findings conflict with Piaget's claim that young children are "egocentric". In the classic 'three mountains task' of Piaget, the child had to appreciate the doll's perspective of the scene. Piaget found that young children could only appreciate the doll's scene when it corresponded with their own view. However, Borke (1983) applied the same experimental design to a more age appropriate task (small toy figures were used instead of three mountains and an identical second display was used to select the doll's view instead of asking the child to select from pictures or to build another model). Performing on this task, 3 and 4 year olds were able to demonstrate perceptual role taking ability, suggesting that young children are not as egocentric as Piaget claimed and that one needs to be sure that the child's problem is not task specific. Shatz and Gelman (1973) found that, whilst 4 year olds failed on the Piagetian mountain task, in a toy task involving a more natural and familiar situation, children adapted their utterances according to the demands of the listener: The younger the listener, the greater the tendency to use short, simple utterances and to make efforts to attract and sustain the listener's attention.

Many other studies have shown pre-schoolers "talking down" to younger children (Brown and Fraser 1964, Jakobson 1968, Cazden 1970). Similar studies have shown that young children will adapt their utterances according to whether the listener can see the referents or not (Maratsos 1973, Menig-Peterson 1983, Perner and Leekam 1986).

Contrary to the evidence which suggests that children can appreciate the perspective of another from a very early age, there are many referential communication studies which show that pre-school children cannot label one of several objects so that the listener can unambiguously identify it (Glucksberg, Krauss and Welsberg 1966, Robinson and Robinson 1978, Flavell 1985). Flavell even found the same negative results when the task complexity was reduced so as to match the child's cognitive level.

Thus, it appears that young children are not egocentric in the way Piaget claimed. To the contrary, there is much evidence to suggest that early social interaction is crucial for early linguistic achievements. Yet, there are age differences, which suggest that young children's social competence is still not the same as that of adults. This issue is looked at more thoroughly in the next sub-section.

1.2.A3 Knowledge of the relationship between social-pragmatic knowledge and linguistic expressions

Whilst the previous sub-sections have provided evidence of young children showing an early mastery of pragmatic inferences and appreciating the knowledge of others, there

is evidence to show that young children do not use their social and pragmatic knowledge effectively when communicating to a listener. It appears that, as Olson claimed (1977), "the ability to assign meaning to the sentence per se, independent of its non-linguistic interpersonal context, is achieved only well into the school years".

Overall, the referential communication task literature shows that young children have difficulties due to their inability to appreciate the significance of the message for referential success. Perspective taking and 'appearance verses reality' tasks show that young children's problems stem from their inability to divorce mental states from world states and because their interpretation of a task depends on what is most salient or 'up front' for them, as opposed to what is objective and necessary.

Contrary to the evidence that young children can appreciate the perspective of another, many referential communication studies have shown that pre-school children produce ambiguous messages. Some investigators argued that this was a reflection of the child's problem in comparing referents with non-referents (Asher and Oden 1976), whilst others have shown that young children produce ambiguous instructions for building lego models even when they can make the relevant comparisons (Robinson and Robinson 1982). Moreover, Deutsch and Pechmann (1982) found that young children would not spontaneously produce unambiguous messages, but would do so only after 'adaptive specific

requests' from an adult. The fact that referential communication failure can reflect a number of different problems led Flavell (1974) to divide the task into a number of different skills: existence, need, prediction, maintenance and application. Lloyd and Beveridge (1981) used more natural referential communication tasks, arguing that the performance of discourse participants depended on the specific demands of each particular task. Whitehurst and Sonnenschein (1978) found that 5 year olds did not differentiate between pairs of triangles. However, as results depended on the type of instructions given, Robinson and Robinson suggested that the children were describing what was perceptually salient (values rather than dimensions) and thus, they concluded that young children do not spontaneously see the need for unambiguous messages.

Thus, the problem seems to be due to the young child failing to appreciate the difference between intended referents and the message itself: Robinson and Robinson (1983) found that the young speaker often claimed that s/he had produced the unambiguous (intended) message, when the message had actually been ambiguous.

Similarly, there is evidence that comprehension of referential ambiguity correlates with production ability (Robinson and Robinson 1982, Pratt and Bates 1982). Bearison and Levey (1977) found developmental effects when the task involved evaluating messages. In fact Piaget (1926) claimed that the pre-operational child judged message adequacy by idiosyncratic measures. Again, there

are investigators who have claimed that message evaluation problems are related to the performance difficulties of making the relevant comparisons (Asher and Oden 1976, Bearison and Levy 1977). However, others have found evidence to show that the problem is one of appreciating the need for comparison activity (Flavell, Speer, Green and August 1981, Robinson and Robinson 1982) and that difficulties are due to judging 'too general' messages as adequate as well as judging 'idiosyncratic' ones inappropriately. Moreover, Beal and Flavell (1982) found that young children were poor message evaluators even when message ambiguity was purposely increased in salience.

There is also much experimental evidence to suggest that listeners in referential communication tasks do not respond appropriately to ambiguous messages (Roberts and Patterson 1983, Jackson and Jacobs 1982, Bredart 1984). Moreover, Lempers and Elrod (1983) found that young children failed to discriminate between different types of message inadequacies (situation, listener and language/message dependent problems). However, the success of young listeners appears to depend on whether they have access to information or not (Marvin, Greenberg and Mossler 1976, Wimmer, Hogrefe and Perner 1986): Markman (1977) found that 6 year olds were only aware of message inadequacy when attempting to carry out instructions or when they saw them demonstrated. Beal and Flavell (1984) and Flavell (1986) found that 6 year olds only recognised ambiguity when they were unable to tell which object a speaker intended. Robinson and Robinson

(1983) found that, if the young child was told what the intended referent was, they would subsequently identify an unambiguous message as being the message delivered by the speaker. Robinson and Whittaker (1985) argued that these effects were due to the fact that young children can only assess message adequacy when they can attend to their own feelings of uncertainty. Similarly, Sodian (1988) argued that problems arose when the child did not share the perspective of the listener.

However, Ackerman (1979) found that 6 year old children were more likely to return with an object requested which only approximated the speaker's description if the speaker's memory was said to be bad. This result suggests that children were actually performing as a function of the speaker. However, in a subsequent experimental replication, Mitchell and Russell (1989) found that when asked 'Is this the object that the speaker really meant?', they were more likely to say no, suggesting that they were probably less strict about whether the object was correct when the speaker's memory was bad, rather than regarding the object as the one that was really intended by the speaker.

Robinson and Robinson (1977) found that young children tended to blame the listener for ambiguous messages. Moreover, young listeners fail to spontaneously request clarification from the speaker, whereas older children tend to ask 'process questions' (which one do you mean?) when messages are ambiguous (Lloyd and Beveridge 1981, Cosgrove and Patterson 1977, Patterson 1978,

Robinson and Robinson 1983). Ackerman (1981) found evidence to show that listeners tended to make a response if they could (performative bias) rather than attending to the propositional content of utterances. This evidence is supported by Singer and Flavell (1981), who found that 6 year olds were more likely to rate closure (ambiguous and listener responds) instructions as less ambiguous than no closure (ambiguous and listener fails to respond) instructions.

There is evidence to suggest that the difficulty young children have in understanding ambiguity of reference stems from their less developed theory of mind: their inability to understand that sources of information are related to knowledge acquisition. For example, Olson and Astington (1987) presented pairs of animals, some of which were completely different and some of which were similar. They found that 6 year olds failed to appreciate that knowing what a referent was depended on the type of information provided: When the children were subsequently only given colour information, they failed to appreciate the ambiguity when the colour matched either one of two possible referents (similar condition).

Moreover, the evidence that young children do not appreciate informational sources of knowledge has also been applied to the social domain: There is evidence that young children can divorce mental states from world states (Johnson and Maratsos 1977, Johnson and Wellman 1980, Wimmer, Hogrefe and Perner 1984) in tasks which involve acting out a narrative where an individual sees or is told

about an object in place a. and then another individual moves the object secretly to place b. In these tasks, the subject is then asked, 'Does x know/think that object is in place a/b?'. Young children appear to be very good at understanding mental verbs, but their performance depends on whether mental events or world states are most salient. For example, Wimmer et al (1984) found that when the seeker unintentionally misinformed a third protagonist about the location of the object, the child said that the person lied. Thus, they were unable to divorce mental states from the perceived outcome. From these studies, it seems that what young children fail to do is to appreciate the division between intensional and extensional states (Russell 1987).

Flavell, Flavell and Green (1983) argued that the problem for young children is that they respond to whatever is most cognitively salient in a particular situation. Thus, it seems that whether young children interpret outcomes as due to reported mental states (intensions) or perceived events (extensions), depends on what is most salient to the child in that particular situation.

This explanation of developmental differences is in accordance with the evidence that, in some tasks, young children are overimpressed by the EXTENSION. For example, in experiments of children's drawings, it has been found that young children are 'intellectual realists' (Luquet 1927). For example Freeman and Janikoun (1972) found that young children would draw a mug with a handle,

when asked to draw what they could see, even through the handle was occluded by the mug and was thus, invisible to the child. These results are interpreted as the children drawing what they know of the 'world' as opposed to what they actually see (Barrett and Light 1976, Pillow and Flavell 1983). The Johnson et al tasks are also in accordance with the evidence that, for certain tasks/situations, young children are more concerned with the INTENSION than the extension. For example, Braine and Shanks (1965) presented visual illusions to 5 and 6 year olds and whilst they could correctly answer questions 'Does it look the same?' and 'Is it really the same?', they made phenomenistic errors in interpreting a neutral question, 'Is it the same?': They were more likely to say what the object looked like (appearance), rather than what it really was (reality).

These findings of young children confusing appearance with reality have become known as the appearance-reality problem. Flavell, Flavell and Green (1983) found evidence of phenomenistic and reality errors and argued that type of error depended on what was most salient to the young child. For example, they argued that when the focus was on the properties of an object rather than on the entirety of an object, 'phenomenistic errors' (interpreting as appearance when the task requires a reality interpretation) were more likely, whereas focusing on the wholistic aspect of objects was more likely to produce 'realist errors' (interpreting as reality when the task requires an appearance interpretation).

Appearance/Reality errors appear to be a reflection of perspective-taking abilities: Salates and Flavell (1976) and Flavell, J.H, Everett, Croft and Flavell, E.R (1981) argued that appearance-reality difficulties were a reflection of a lingering influence of level 1 thought: Flavell proposed levels of development for perspective-taking (visual percepts). Level 1 represents whether another person sees or does not see an object. Young children are at level 1 when they speak more clearly to a listener who has her eyes closed. Level 2 represents how an object looks from a particular spatial perspective and appreciating that there may be more than one perspective for the same object/situation. Roberts and Patterson (1983) and Flavell (1985) found evidence of referential ability correlating with level 2 perspective taking ability. Flavell argued that perspective taking ability, together with appearance-reality knowledge are a result of an "increased cognizance of the fact that people are sentient subjects who have mental representations of objects and events". This theory accounts for the development of 'intensionality': the ability to dissociate what is 'meant' from what is 'said'.

J.Russell (1983) argued that Appearance/Reality problems were due to the young child adopting a subjective 'propositional attitude' to the situation. A propositional attitude is a philosophical term for an attitude of mind towards linguistic propositions. Thus the meaning of a sentence is the object of a propositional attitude. Subjective and phenomenal propositional attitudes,

according to Russell, account for why children are impressed by what is salient to them, rather than what is objective and necessary.

The ability to approach a situation objectively can be related to metacognitive achievements (Markman 1977, 1979, Flavell 1979, Flavell, Speer, Green and August 1981). Flavell defined metacognition as any knowledge that takes any cognitive phenomenon as its object and this development is reported to occur mainly throughout middle childhood and adolescence. Donaldson (1978) referred to reflective thinking as 'disembedded' thought: thought not embedded in context. Olson (1977), Olson and Torrance (1983) have argued that meta-LINGUISTIC achievements occur as a consequence of schooling and literacy. In fact, there are a number of sub-cultural and cross-cultural studies which suggest that logical, syllogistic reasoning problems are related to culture and education and there is evidence that even adults are overimpressed by pragmatic knowledge, whereas 'intellectuals' are able to approach a situation 'objectively' because this is a reflection of their own experiences with academia (Cole and Scribner 1959, Thouless 1959, Evans-Pritchard 1963, Hymes 1974, Scribner 1975, Cole 1977, Luria 1971, 1977). Nevertheless, Donaldson (1978) argued that the difference between children and adults is the amount of weight that is given to sheer linguistic form and De Villiers and De Villiers (1978) wrote, "We do not wish to claim that language use ever becomes totally free from context, but the influences on what we say and how we might say it become increasingly

remote from the stimuli of the immediate physical and social environment".

To summarise the evidence presented in this sub-section, it seems that the problems for young children are due to their inability to appreciate that more than one factor is often important in a situation. Thus, whilst they can show a high level of sophistication in making pragmatic inferences and are sensitive to the knowledge of the listener, they are unable to use these perspectives in a situation which requires linguistic objectivity and they fail to dissociate linguistic expressions from the non-linguistic (social-pragmatic) features of the situation.

B. Developmental research and in/definite reference

Two areas of conflicting evidence are discussed in this section. The first of these is concerned with whether young children are only able to refer deictically when using definite reference and therefore whether anaphoric reference is a relatively late acquisition. The second area focuses again on the 'egocentricity' issue. In both areas there appear to be contradictory findings.

Returning to the first issue, Lyons (1975, 1977, 1979) argued that deixis was "a source of reference". He also argued that deixis is ontogenetically prior to anaphora. He supported this claim by citing evidence that primitive attention directing gestures allow reference and hence communication to occur at the outset of language use.

The term 'deixis' is used to describe gestural

reference and essentially is the same as identification by pointing. Linguistic expressions which are deictic serve to direct the hearer's attention to spatial or temporal aspects of the situation of utterance and the type of expression used depends on the immediate extralinguistic context (the speaker's here-and-now). Demonstratives (this and that), locatives (here and there), personal pronouns (I, you, he, she, it, they, them, his, her, him) and the definite and indefinite articles (A and the), are all typical instances of deictic expressions.

Ehrlich (1982) claims that "the deictic procedure is a linguistic instrument for achieving focusing of the hearer's attention towards a specific item, which is part of the respective deictic space."

In summary then, deixis makes use of spatial, temporal and interpersonal contextual features of situations as aids in the management of joint reference.

Ehrlich stated that there is a functional difference between deixis and anaphora. The anaphoric procedure is defined as "a linguistic instrument for having the hearer continue a previously established focus towards a specific item on which he had oriented his attention earlier." Basically anaphoric reference means referring back to some antecedent which has already appeared in the discourse, for the purpose of restatement or elaboration. Like deictic reference, anaphoric utterances can only be interpreted within an actual context of use, the difference in this case being the relevance of the linguistic context as opposed to the immediate perceptual environment.

There is evidence that children as young as 2 years of age do produce deictic terms and the most frequently used terms by parents and children when they are interacting together are mainly deictic (Brown 1973, Wales 1986, Tanz 1980). Brown found that the earliest utterances produced by his subjects, Adam, Eve and Sarah, were the first and second personal pronouns (I and YOU). This finding appears to support the above prelinguistic studies, because only the first and second personal pronouns are 'pure' deictics, that is to say, their use depends on changes in the immediate spatio-temporal location, or their role in the locutionary act. For example, if there were two participants engaged in a conversation called Sarah and John, Sarah might say "I am talking to you": 'I' being used to refer to herself, Sarah, and 'you' referring to John. John in the same context might also say "I am talking to you". Although the sentence is identical to Sarah's, the roles of 'I' and 'You' are reversed: 'I' now refers to John and 'you' now refers to Sarah. If Susan then appeared on the scene and John was no longer involved in the conversation, Sarah may repeat the sentence again. Now the word 'you' would be referring to Susan. Thus, the deictic pronouns 'I' and 'You' are directly dependent on the immediate context for use.

From a number of experiments, Karmiloff-Smith (1979) concluded that initially, children use 'spatial deixis' for those determiners that serve both deictic and anaphoric functions, such as definite articles and pronouns. However, she claimed that it was not until about 9 years of

age that a child could fully understand the anaphoric function of pronouns and definite articles. For example, the words 'the same' appearing before a noun phrase should refer anaphorically to a previously mentioned referent. Karmiloff-Smith found evidence of a child of 4 years of age interpreting 'same' to mean 'same kind' and she suggested that this was because young children are more concerned with similarities of attributes than identity. She concluded from this that the young child is simply making deictic reference to the extralinguistic context.

In fact, Piaget's conversation with his daughter Jacqueline ('Play, dreams and imitation in childhood', 1951, p226), that led him to conclude that an infant child cannot appreciate the differences between generality and particularity, was criticised by Karmiloff-Smith. According to Piaget, Jacqueline said 'encore la limace' (there's the slug again) after two separate appearances of a slug, and when she was asked 'is it the same slug?' and then 'is it a different slug?', she responded 'yes' to both questions. Karmiloff-Smith argued that, Jacqueline was actually interpreting 'same' to mean 'same kind' and was using 'la' (the) deictically rather than anaphorically: 'la' was merely a verbal gesture towards the slug and therefore, 'encore la limace' cannot be interpreted as a failure to differentiate generality from particularity.

There have been some criticisms and counter-evidence to Karmiloff-Smith's findings, many of which will be dealt with in the next chapter (Clibbens 1985, 1986, 1988, Emslie 1986). However, Karmiloff-Smith is about the only

investigator of children's indefinite and definite reference who actually looks at development over a broad age span: 4 to 11 years.

Jarvella and Klein (1982), looked at American Sign Language (ASL) and found that, whereas 'pointing' is used to sign for present entities, non-present entities are assigned unique spatial loci, which can then be used for anaphoric reference by subsequent pointing to the specific spatial loci. They found that it was these kinds of anaphoric distinctions that young deaf children often failed to make. This evidence suggests that the development from deictic to anaphoric reference is possibly a universal intellectual achievement.

Finally, Halliday and Hasan (1976) argued that, after the 'exophoric mode' (reference to items relative to a physical context) had been acquired, this knowledge was later transferred to the 'endophoric mode' (reference to another part of the verbal text). This claim is in accordance with Karmiloff-Smith's theory: that development of definite reference progresses from deictic to exophoric and finally to anaphoric reference. The exophoric to anaphoric change seems to require an ability to refer to hypothetical entities in discourse contexts as opposed to simply concrete visible items.

The second area of investigation originated from claims that, young children develop from using definite expressions 'egocentrically' to a non-egocentric form of indefinite and definite reference. Brown (1973) and Gallaway (1985), concluded from their observational

studies, that the young child was still unable to use the articles non-egocentrically, because they failed to use indefinite articles when the listener was unfamiliar with the referent. Similar findings have also been shown experimentally (Warden 1978). However, other empirical findings have found evidence to the contrary: that young children do use indefinite and definite expressions as a function of the knowledge of the listener (Maratsos 1973, Emslie and Stevenson 1981, Emslie 1986).

Thus, there appears to be no conclusive evidence regarding either the claim for a development from deictic to anaphoric definite uses or for the egocentricity debate. The next section briefly summarises the evidence so far from philosophical, linguistic, cognitive and developmental research, together with some speculative hypotheses with regards to the development of understanding and using indefinite and definite reference.

1.3 Summary of the literature

From a survey of the philosophical and linguistic literature, it was decided that the crucial definition of in/definite reference depended on the two factors of FAMILIARITY and SPECIFICITY. With regards to representing and processing discourse, evidence was cited that stressed the important role of non-linguistic pragmatic and social-conventional aspects of discourse. Whilst this could partly account for the 'familiarity' of in/definite reference, the 'specificity' factor

could only be accounted for by assuming that discourse required using and updating discourse models (a model of the current content of discourse). The latter was necessary because specific reference required knowing the relative status of a referent in comparison with other available referents. Moreover, knowing that 'specificity' within a model is necessary for definite reference, requires an ability to integrate linguistic with non-linguistic factors whilst processing discourse.

From the developmental literature, research suggests that young children can understand that 'mental states' are different from 'world states' and they can make pragmatic inferences about 'world states'. Their problem appears to be one of being able to dissociate 'mental states' from 'world states' and rather than reflecting on the significance of all factors, they appear to be interested in whatever is most salient to them at the time.

This evidence might lead one to speculate that the young child should have no trouble discriminating between indefinite and definite reference with regards to listener 'familiarity', because young children seem to be well advanced in their social-pragmatic knowledge. However, if as proposed by Flavell, J.H. Everett, Croft and Flavell, E.R (1981) they are still at level 1 of perspective taking and they cannot appreciate the appearance-reality distinction, together with the evidence that young children fail to appreciate ambiguity in referential communication tasks, one might argue that, with regards to definite reference,

they are not using mental models effectively, nor are they processing discourse in a way which allows the particular context to interact with linguistic constraints for definite reference. If this is the case, then the young child does not understand the 'specificity' requirement of definite expressions.

Referring back to the previous section, one might argue that young children are using definite articles deictically, because they are mentioning a salient object rather than referring to a specific object (in comparison to other available referents). Moreover, it seems that both 'exophoric' and 'ana/cataphoric' reference require a knowledge of 'familiarity' and 'specificity', the only difference being that the referential context for exophoric reference is the perceptual array, whereas it is the imaginary discourse model for anaphoric reference. One might predict that exophoric reference would be easier than anaphoric reference because the disambiguating information is more likely to be salient in an exophoric context, where physical features are important (the red clown), than in an anaphoric context, where it is previous verbal information which is significant (the clown that was skipping).

However, before the experiments of this research are individually discussed, the next chapter looks at past studies of children's understanding and use of the singular articles (a, the) and third person pronouns (he, she, him, her). The chapter then describes the experiments reported in the current investigation.

CHAPTER 2

PSYCHOLOGICAL STUDIES OF CHILDREN'S INDEFINITE AND DEFINITE REFERENCE

This chapter summarises previous psychological investigations of children's understanding and use of indefinite and definite expressions and the chapter closes with a brief introduction to the experiments of the current research.

The main body of this chapter describes and discusses previous investigations. As discussed in the previous chapter, it was decided that the required knowledge for indefinite and definite reference was FAMILIARITY and SPECIFICITY. Thus, each study is described under one of these two titles. Under 'familiarity' are reported those studies which have looked primarily at a) sensitivity to the listener's knowledge and b) sensitivity to pragmatic factors. Under 'specificity', are reported those studies which have been mainly interested in a) the exophoric context and b) the anaphoric context. This is followed by a separate section, which looks only at studies of children's pronominal reference. The reason why the latter were kept apart from studies of definite articles was because pronominal reference is often looked at from additional linguistic and structural perspectives.

2.1 KNOWLEDGE OF 'FAMILIARITY' AND INDEFINITE/DEFINITE

REFERENCE

A. Sensitivity to the listener's knowledge

Studies which have focused on the child's sensitivity to social aspects of indefinite and definite reference have used a number of different tasks and materials. It seems that young children are rather good at considering the listener's knowledge when using the articles and any difficulties appear to be due to those aspects of a task that are too complex and/or confusing for the young child.

From a longitudinal study of 3 children (Adam, Eve and Sarah) by Brown (1973), it was found that the articles (a, the) were the 7th and 8th morphemes to be used and were acquired in children at approximately 3 years of age. Brown claimed to have found evidence of both deictic and anaphoric examples of definite reference, but he found that the infant would use 'the' when an object was specific for the speaker but not the listener. Thus, Brown concluded that young children still used the articles 'egocentrically': they failed to take the listener's perspective into account.

However, Hickmann (1980) argued that the problem with Brown's naturalistic study was that his anaphoric examples 'that a jeep.....the jeep', were stated in the absence of a referent/extralinguistic context and therefore, it is not valid to conclude that this represented true anaphoric reference. Moreover, Gallaway (1985) also carried out a longitudinal study and concluded from her findings that it is unlikely that the 2 year old

children who were producing articles were actually viewing the articles as words, but instead, they appeared to be part of a larger fixed chunk probably as a result of experiences with certain adjective co-occurrences, for example 'it's a', 'where's their', and 'the same'.

Warden (1974, 1976, 1981) from experimental studies of a wider age range of subjects, also concluded that younger children were using the articles 'egocentrically'. In one of his experiments (1976), he manipulated the knowledge of the listener by having subjects either looking at the pictures together with the listener (social condition) or having subjects looking at pictures alone (isolated condition). Warden found that both children and adults were poor at varying their expression with condition and he argued that maybe subjects were reluctant to treat a listener experimenter as a listener. Therefore, in a subsequent experiment, he used a peer communication procedure, where the speaker had to narrate a 3 picture cartoon story to a listener who was behind a screen. The results were that, from 3 to 8 years of age, subjects used mainly definite descriptions or non-referring indefinites (naming responses) on first mention of a referent. Warden also found evidence of an inconsistent use of indefinite referring expressions on first mention from children of intermediate ages (5 to 7 years), such as 'A cat jumped up a tree and the dog came along' and from these results, he concluded that the young child was 'egocentric' and that 'non-egocentric' references emerged gradually and were only fully understood at around 9 years of age. Warden

found the same results when he manipulated the presence/absence of both referents and listener and therefore, he concluded that 'egocentric' responses were not a function of physical presence. Emslie and Stevenson (1981) and Emslie (1986), criticised Warden for failing to include cases of first mention definites which would be appropriate given the pragmatic context, because they would allow the listener to infer the existence of a referent through pragmatic inferences. Emslie also criticised Warden's materials for being very vague and ambiguous and argued that if, in the Warden task, the young child was having difficulty representing the events taking place, then it is hardly surprising that they failed to appreciate the knowledge of the listener.

Hickmann (1980) used a task whereby 4 to 10 year olds and adults had to narrate films to an ignorant listener. Each film involved 2 inanimate characters talking about 2 non-participants. Hickmann found that 4 year olds never created referents in discourse with an indefinite noun phrase and appropriate first mentions increased with age. However, it is possible that developmental differences observed in this experiment were a reflection of the complexity of the task: reporting dialogue which required rephrasing linguistic material.

Another study which looks at the question of listener sensitivity is an experiment by Power and Martello (1986) who showed that whilst young children were more likely to use definite articles on second mention than on first mention when speaking to two different listeners in

sequence, the young child was more likely to use definites on first mention when talking to the second listener than when talking to the first listener. The investigators concluded from this finding that speakers may be using an 'egocentric rule': use indefinite expressions when a referent has just come into attention. Thus, they argued that young children are more likely to use a definite expression when they (themselves) are familiar with the referent, irrespective of the knowledge of the listener. However, overall second mentions were more definite than first mentions even for the second listener, which suggests that at least some social sensitivity was taking place.

Maratsos (1976) carried out a number of production, comprehension and imitations experiments on children between 3 and 4 years of age. In one experiment he presented stories, for example, he began a story about a man in a jungle and then said 'He saw 2/some animals. He saw a/some monkey/s and a /some pig/s', followed by the question 'Who came to the man?' (Answer: a/the monkey/pig). Maratsos found that from 4 years of age, subjects' choices of indefinite or definite articles in answering the questions depended on whether one particular referent had been introduced in the story or not. He also concluded that between 4 to 5 years of age, the young child is able to appreciate the perspective of the listener. However, there are a number of methodological criticisms of Maratsos' experiments (Haviland 1976, Hickmann 1980, Warden 1981). Some of these criticisms are that the

situation was unnatural because subjects were made to sit on their hands to prevent deictic pointing gestures, subjects were in a strange environment and were bored, the sample was too small, the sessions were too long and the type of questions used elicited the given types of referring forms. Haviland (1976) made many such criticisms when she commented, "Single sessions could run to an hour.....story completions were particularly disliked, which was not surprising as they involved responding to questions with no real answers. Such things can play havoc with any attempt to assess non-egocentric behaviour".

However, Emslie and Stevenson (1981) carried out 3 experiments on 2 to 4 year olds and adults using the 3 picture story task of Warden, with a peer listener situated behind a screen. They found that, contrary to Warden's results, from 3 years of age, indefinite articles were reserved predominantly for first mention.

Emslie (1986) found that when carrying out the same task, comparing Warden's pictures with the pictures used in the Emslie and Stevenson experiment, 4 year olds were more likely to use indefinites on first mention for the Emslie and Stevenson versions, but were more likely to use definites on first mention for the Warden versions. Thus, it seems that the criticism of Warden's materials above does account for his negative results. In a subsequent experiment, Emslie found that children of 4 years, who preferred an indefinite article on first mention to an ignorant listener, would preferentially use a definite article on first mention when the listener was

knowledgeable. Emslie concluded that when young children could clearly interpret the actions in the pictures, their choice of an article was a function of the knowledge of the listener.

Zehler and Brewer (1982) looked at 'shared narratives' and sentence completions in tasks involving talking about a set of objects. They found that, even 2 and 3 year olds produced a high percentage of introductory (identifying) uses of the indefinite article (84% - 100%). They argued that any overuse of definite expressions in younger subjects was related to the 'specificity' of referents (see next section).

Finally, Bennett-Kastor (1983) looked at 2 to 5 year olds' narration of a story and found no evidence of subjects' overusing the definite article to introduce referents.

B) Sensitivity to pragmatic factors

The results of the experiments below seem to suggest that young children do use their general knowledge and make pragmatic inferences when deciding upon indefinite or definite reference: There is evidence to show that they will use a definite article to introduce a referent to a listener when that referent can be inferred from the context. The only differences between findings seem to be due to the relative amount of significance that general knowledge has for the child in relation to the particular features of the experimental situation. The question still to be answered is whether young children's varying

sensitivity to pragmatic factors as a function of different tasks reflects their more limited conceptual experiences or is due to their weaker tendency to use the conceptual knowledge they already have to make pragmatic inferences compared with older children.

Recalling the 'usage types' of Hawkins (1978), he argued that anaphoric types could be anaphoric by using synonyms of entailment, for example, 'A car - the engine'. The engine is anaphoric and does not require previous introduction with an indefinite article, because one can infer the existence of an engine through prior mention of a car. Moreover, Hawkins used the term 'associative anaphora' to refer to first mention definites that could be inferred to exist through prior introduction of a concept that the anaphoric term was related to contextually, for example 'shop - the assistant': One can infer the existence of assistant from a prototypical representation of a shop. From these examples, one can see that determining whether a referent is familiar to a listener not only depends on the immediate physical situation, but also depends on background shared experiences and general knowledge of discourse participants. Thus, the use of pragmatic inferences is relevant to the question of familiarity for in/definite reference.

Brown (1973) claimed that he observed 3 year olds using definites which were specified by entailment, such as "the sticky of the bandage". Zehler and Brewer (1982) found that young children were very good on 'context uniques' (car - the steering wheel) and 'context

intermediates' (car - the door: defined as 'intermediates' because the context would predict 4 car doors, but only one of them is relevant in the story). Moreover, Bennett-Kastor (1983) found that 4 and 5 year olds used more definites on first mention as a function of "generic or archetypal noun phrases" (the wolf, the witch) as well as for class inclusion associations (a zoo - the zookeeper).

Finally, Emslie (1986) found evidence of contextual influences on article production. She found that items which normally appeared as one of a class, such as 'cup', tended to elicit indefinite expressions and objects which served a locative function were generally definite, for example 'on the table'. This is in accordance with Clark's (1978) claim that locative phrases are normally definite. Nevertheless, Emslie claimed that the listener's perception of the referents in the immediate situation was the strongest factor in determining young children's choice of expression, regardless of their general knowledge of contexts.

2.2 KNOWLEDGE OF 'SPECIFICITY' AND INDEFINITE / DEFINITE REFERENCE

A. The exophoric context

Studies of children's exophoric reference are again rather varied and findings are contradictory. From the experiments so far, it would appear that young children's performance is a function of the way that they perceive the task as well as due to the particular type

of task that is used.

The Bresson group (1970) carried out experiments on 4 and 5 year old children, who were required to refer to singletons or classes of visible farmyard animals. They found evidence of an overuse of definite articles for reference to sub-classes of items that were not specific. Zehler and Brewer (1982) found that when young children overused definite articles this was when there was a large number of non-specific items present, such as 'the girl opened a bag of blocks and took out the block' (expected expression - 'a block'). The investigators argued, it is "possible that young children do not have the linguistic means at their disposal for making reference to one of several identical objects which are visible to speaker and listener".

Karmiloff-Smith (1979) carried out a number of quasi-experimental investigations on French children between 3 and 11 years of age. The experiments mainly involved the experimenter manipulating toys and then asking questions. Karmiloff-Smith found evidence of children under 7 years of age using the definite article to refer to the experimenter's action on one of several identical objects: objects that were non-specific. For example, in the 'hide and seek' experiment, when the experimenter hid an object, in response to the question 'What did I do?', young children's responses were predominantly definite, regardless of whether the object was 'identical' (non-specific) or a 'singleton' (specific).

Some criticisms of Karmiloff-Smith's experiments

(Francis 1980, Emslie 1986) relate to the fact that no tight control was exerted over variables and that there may have been carry-over effects from the same subjects performing in more than one experiment. Another criticism (Emslie 1986) was that she failed to control for social features of the experiments: In many of her studies, the experimenter was the listener who pretended to be ignorant of the object and who could infer the correct object through the extralinguistic context. Emslie argued that it is possible that the subjects were not regarding the listener as ignorant, in which case it is not surprising that they overused the definite article. In fact, in a similar 'hide and seek' experiment, which involved a toy farmer moving amongst an array of animals, Garton (1983) manipulated whether the listener was 'blindfolded' or 'seeing'. Whilst, young children still failed to increase indefinites to non-specific items, there was a difference in the expressions used between these two conditions: The 'seeing' subjects mainly failed to use articles at all, whereas the 'blindfold' group were far less likely to omit articles. Nevertheless, Garton found that 60 per cent of second utterances were demonstratives especially when referents were 'identical' and therefore, she concluded that young children do not have the linguistic means to distinguish between 'identical' items.

Emslie (1986) partially replicated the 'hide and seek' task, but both participants had to close their eyes for the hiding and only the child could see the remaining, unhidden objects. In this experiment, except for one case,

children between 4 and 7 years of age were better at discriminating between 'singleton' (specific) and 'identical' (non-specific) items with regards to type of article used (indefinite or definite). However, Karmiloff-Smith herself claimed that tasks which required indirect reference such as in the 'hide and seek' task were better at eliciting indefinites from young children than were tasks which required direct reference to an item. For example, in Karmiloff-Smith's 'playroom' task, where the children had to ask a doll to lend them the toy touched by the experimenter, the youngest children were clearly non-discriminative in their choice of in/definite expression, as they preferred definite references regardless of the 'specificity' of the referent. Moreover, it was not until about 9 years of age that children would discriminate between non-specific and specific reference. Nevertheless, the latter experiment may be criticised for not actually requiring non-specific reference, because the experimenter always pointed to a visible object and thus the children may have interpreted the task as one which required only deictic reference (pointing to a salient object).

However, Emslie (1986) used other tasks which required direct reference and she found similar findings: Young children were influenced by referent 'specificity' in their choice of indefinite or definite expression. Nevertheless, there are large differences between the Emslie tasks and the Karmiloff-Smith tasks. One of Emslie's tasks (farmyards) showed that sensitivity to

'specificity' of items only occurred when the subject, rather than the experimenter, chose the referents. However, it is not clear as to the extent of choice subjects had in Karmiloff-Smith's experiments. Also, for Emslie's 'balancing' task, speakers could see the performance of the listener and therefore, visible matching between arrays may have lead subjects to focus on features that they would not otherwise have done had they been given a task which did not allow them to see the consequences of the speaker's failure to discriminate between specific and non-specific items.

Finally, the results of Karmiloff-Smith's experiments showed that the youngest children were better at understanding the differences between the articles as a function of 'specificity' (they were more likely to use indefinites for non-specific items), when the experiment involved a comprehension rather than a production task.

B. The anaphoric context

Whilst there is evidence to suggest that young children prefer definite articles when mentioning a referent for the second time, there is also evidence that this reference is often ambiguous, suggesting that young children have problems with anaphoric reference when more complicated noun phrases are required to make a referent specific.

Karmiloff-Smith (1979) tested anaphoric reference by looking at the way subjects made reference to the same object twice. She found that even the youngest subjects would prefer definites on second mention. However,

she argued that this was no indication of anaphoric reference, because her young subjects used definites non-discriminatively on first mention, suggesting that they were always using definites deictically.

However, Emslie (1986) found that young children would prefer definite reference for second mention even when indefinites were preferred on first mention. The fact that even the youngest subjects preferred the pattern, indefinite first and definite second when the listener was 'ignorant' led Emslie to conclude that even 4 year olds are able to use definite expressions for anaphoric conservation.

Nevertheless, definite reference which is used for second and subsequent mentions of an entity is not necessarily anaphoric: It is only anaphoric if it unambiguously refers to a specific entity that is mentioned or can be inferred from preceding discourse. Emslie found that young subjects used many ambiguous pronouns and failed to use explanatory modifiers or relative pronouns for anaphoric reference. Thus, many of their second definite references were not referring to a specific entity. Furthermore, Karmiloff-Smith observed that younger children were more likely to use demonstratives and pronouns for definite reference than older subjects. Emslie concluded that these age effects were due to differences in 'task perception'. However, Emslie did not go on to discuss why younger children should be more influenced by task variation, nor did she suggest the possibility of these age differences in approaching a task as possibly related to

their knowledge of 'specificity' and anaphoric reference. It seems that her failure to consider this possibility is a consequence of the way that she interpreted anaphoric reference in her experiments: She was concerned with the listeners' knowledge of referents per se (knowledge of referents in the 'world') as opposed to their knowledge of referents in relation to other available referents (knowledge of referents in a 'model').

2.3 STUDIES OF CHILDREN'S PRONOMINAL REFERENCE

This section has been divided into studies of LINGUISTIC factors of pronominal reference and studies of the NON-LINGUISTIC factors. Within each of these sections are further sub-divisions. The 'linguistic' section looks at syntactic and lexical knowledge separately and the 'non-linguistic' section is divided into thematic principles, pragmatic factors, heuristics and finally, knowledge of pronominal ambiguity. Whilst studies of children's pronominal reference are discussed in a separate section from studies of children's understanding of the articles, the overall finding here, that young children's difficulties with pronouns stem from their problems in effectively integrating linguistic and non-linguistic factors, are consistent with the findings that young children fail to show reliably any evidence of a knowledge of 'specificity' for definite reference. This issue is discussed in the final section of this chapter.

A. LINGUISTIC FACTORS

The experiments discussed below seem to indicate that children are aware of linguistic constraints on pronominalisation from an early age. Any difficulties seem to be a reflection of cultural and linguistic variations as well as due to the involvement of non-linguistic factors.

2.3.A1 Syntactic knowledge

There has been conflicting evidence in the literature concerning children's treatment of Backward Anaphora (where the pronoun precedes the noun phrase). The importance of backward anaphora is that it is often constrained by the NON-IDENTITY RESTRICTION. Put in its simplest form, this rule stated that, when the pronoun which precedes the noun phrase is in a main clause, the noun phrase and the pronominal are NOT co-referential: the pronoun is restricted to NON-IDENTITY. For example, 'She was pleased that Susan had passed the exam': 'she' cannot be 'Susan'. However, when the preceding pronoun is in the subordinate clause, this restriction does not apply. For example, 'As she passed the exam, Susan was pleased': 'she' may be 'Susan'.

However, a number of linguists have disagreed over the syntactic explanation of 'non-identity'. The disagreement has been over what it means for one node to be 'higher' than another in the tree.

Langacker (1969) claimed that, if a pronoun PRECEDES a noun phrase and the pronoun S-COMMANDS the noun phrase,

then anaphora is impossible. S-command is where the S-node that dominates node A also dominates node B, but neither node A nor node B dominate the other. However, in order to explain why sentences such as 'I told his wife that John should consume less cholesterol' can be co-referential, Lasnik (1976) and Wasnow (1972) used the term d-command to define the restriction. For d-command, rather than the S-node being the dominating node, it is the cyclic node that most immediately dominates node A that dominates node B.

Reinhart (1976) argued that both S-command and d-command were too restrictive and neither accounted for why the following sentence may be co-referential, 'Bill interviewed him when Woodward was in town'. Thus, she argued that c-command was necessary to account for all non-identity sentences. For c-command, the dominating node is the first 'branching node' which dominates node A that also dominates node B. Incorporating all these constraints, the Backward Anaphora Restriction (BAR) has been summarised as follows:

1. Rule out any Backward Anaphora.
2. Rule out anaphora between a pronoun and a noun phrase not in the same clause just in case the pronoun c-commands the noun phrase.
3. Rule out anaphora between a pronoun and noun phrase within the same clause (when pronoun and noun phrase are clausemates) just in case the pronoun d-commands the noun phrase.
4. Apply at levels of logical form and s-structure.

However, some investigators have found that children's initial hypotheses about co-reference are based on word order and there is evidence that young English children rule out anaphora in a sentence where the pronoun precedes the noun phrase (C.Chomsky 1969, Legum 1977, Tavokolian 1978 and Lust 1977, 1981). Tavokolian presented 3 to 5 year olds with sentences in a toy-moving task and Lust presented 3 to 7 year olds with an elicited-imitation task. Lust found that, even when the relative order of the pronoun and antecedent and the relative order of main clause and subordinate clause were varied, it was only the former (order of pronoun and antecedent) that affected children's imitations. Similar findings have also been shown by Legum.

C.Chomsky (1967) used a comprehension task and presented children with two cartoon figures of Mickey Mouse and Pluto Pup. They were given non-identity sentences such as 'He was five years old when Pluto broke his leg', followed by a question: 'Who was five years old?' Her results showed that a child chose non-identity referents 80 per cent of the time by the ages of 5 or 6 years, whereas, when a pronoun preceding a noun phrase was in the subordinate clause, the choice of a referent was random: There was no strong preference for co-reference or non-identity. Therefore, she concluded that this was the age when the non-identity syntactic rule had been acquired. However, she observed that 3 of her children (ages 6.5, 7.0 and 7.0) used purely directional restrictions on pronominal reference: failing to allow co-reference in any sentence

where the pronoun preceded the noun phrase.

It has been suggested that developmental effects are not a reflection of the child's lack of knowledge of principles such as c-command and local domain, but are rather a reflection of certain linguistic parameters which vary across languages. This is known as the 'parameter setting view' (Solan 1983). Solan (1981), from a series of 4 experiments, found that young children do over-apply BAR to all sentences where the pronoun precedes the noun phrase, but he found evidence that these children did show an awareness of c-command, for example, they were more likely to allow 'I saw him after Bill finished the trial' to be anaphoric than, 'He saw me after Bill finished the trial'. Also, children preferred pronominal reference when the pronoun and antecedent were in the same clause and Solan argued that this suggests that children initially conflate the restrictions on backward anaphora with those on the interpretation of reflexives. Goodluck (1988), argued that children use this word order hypothesis, because the child's initial constraint is too restrictive. Solan (1981, 1983) argued that those parameters which need to be fixed by experience, are those which are particular to a language and/or dialect, as opposed to being part of universal grammar. Support for this claim is that, whilst there is positive evidence that backward anaphora is restricted in all languages, there is also evidence that the application of c-command to this restriction is unnecessary in some languages: In Thai, for example, there are no sentences where identity can occur whenever a

pronoun precedes a noun phrase, and similar evidence is found for Japanese.

Nevertheless, other investigators have found evidence to show that young children do not interpret a pronoun preceding a noun phrase as referring to an antecedent outside the sentence (Goodluck 1986, Mathei 1981). However, BAR is consistent with principle C of Chomsky's binding theory: that an R-expression (eg. definite NP or name) is free everywhere. This rule can account for BAR because basically it is saying that when the pronoun c-commands the noun phrase/name, the latter cannot be co-referential with the former. Thus, one would expect even young children to obey the constraints of universal grammar.

Yet there is evidence to show that young children do not appreciate structural non-identity: Widowsky (1977) used a similar procedure to Chomsky, but instead of presenting two tangible figures, the referent was mentioned in previous sentences in a text. As the task was more complicated than Chomsky's, he only looked at older children. However, he found that, even after checking for memory of the texts, 9 year olds still tended to choose the referent in the same sentence as the pronoun when the pronoun was in the main clause. Thus, he concluded that Chomsky's proposed age for the acquisition of non-identity had something to do with a younger child's preference for tangible objects, and that only at a later stage, when a child is not bound by 'concrete reality', can they really appreciate non-identity in texts. Thus, the differences

between these two results seem to suggest that the child needs to develop from exophoric to anaphoric reference.

Berkowitz and Widowsky (1978), presented only single sentences as Chomsky had done, but they did not use tangible objects. They found results to suggest that only the oldest children (11 year olds) could appreciate non-identity; thus supporting Widowsky's earlier findings. However, a correct response in this experiment was to say 'One cannot know', because the referent was neither tangible nor introduced in the text.

Similar results were found by Kail and Léveillé (1977), who also used isolated sentences followed by a question. They found that only at 11 years of age did subjects appreciate non-identity and only when the lexical marking of the pronoun conflicted with the gender of the named noun phrase, did younger subjects tend to assign reference to 'somebody else' not mentioned in the sentence. However, it is questionable as to whether the non-identity constraint was relevant at all in the sentences used by Kail and Leveille, because they contained two main clauses joined by an 'and' connective, rather than having one main clause and one subordinate clause, as the following example shows: 'Elle/il a relevé le blesse et le pharmacien a appelé le medicine/ S/he became ill and the chemist called the doctor'.

Moreover, there is empirical evidence to suggest that young children are poorer at interpreting sentences using pronouns than they are at interpreting reflexives (Wexler and Chien 1985, Deutsch, Koster and Koster 1986).

However, there is also evidence to the contrary, that young children are relatively good at comprehending pronominals, as shown by Grammaticality Judgement tasks (Grimshaw and Rosen 1988) and Acting-out tasks (Deutsch, Koster and Koster 1986, Solan 1983). Grimshaw and Rosen (1988), argued that these differences in experimental findings can be accounted for by performance factors. One example they mentioned was that, in many of the studies where the performance of young children using pronouns was poor, no antecedent had been provided for the child. Yet, there is evidence to show that children will assume that a pronoun will have an antecedent within the discourse and they will rarely choose an antecedent that is not mentioned (Wexler and Chien, 1986). Solan (1983) also found that when using single sentences, if there was no antecedent mentioned in the sentence, young subjects tended to choose a referent exophorically 80 percent of the time: They chose a tangible object that was not mentioned as an antecedent, despite the evidence that young children do prefer a mentioned antecedent to an unmentioned one if both the object is present and the antecedent is mentioned. As Grimshaw and Rosen claimed, "Both pronouns and anaphors require antecedents, although for different reasons". In Widowsky's experiment of non-identity, where single sentences with no tangible objects were used, he found that younger children were poor at appreciating non-coreference. It is possible that when confronted with a situation where there was a conflict between pragmatic rules (be relevant to context) and linguistic constraints

(c-command or d-command), the children in Widowsky's experiment failed to abide by the latter as a consequence of their concern for the former.

Moreover, Grimshaw and Rosen argued that knowledge of the binding principle A was logically independent of knowledge of principles B and C, because principle A involves purely syntactic and lexical knowledge, whereas the other principles (B and C), involve knowledge of these linguistic constraints, plus an understanding of pragmatic and discourse based information. Thus, many investigators have argued that, as the understanding of pronouns involves both linguistic and non-linguistic factors, errors in sentences containing pronouns can be interpreted in terms of processing demands (Grimshaw and Rosen 1988, Goodluck 1988, Deutsch, Koster and Koster 1986, Stevenson and Pickering 1987). Bearing this in mind, it is possible that the results of previous studies are a result of the pragmatic context of the sentences that were used. Looking at C. Chomsky's sentences, a couple of her examples appear to pragmatically support a non-identity response. For example, 'He found out that Micky won the race', where it seems less plausible for Micky to find out that he (himself) had won the race than for somebody else (Pluto) to find out this information. Likewise, for the Widowsky and the Berkowitz and Widowsky sentences, the bias could just as easily go against non-identity. For example, 'He turned on the T.V before Roni finished preparing the lessons', seems to lead to identity interpretations: Roni turned on the T.V

even before he finished preparing the lessons'.

Although the current investigation focuses primarily on pronominal reference for simple forward and backward anaphoric sentences, there is also evidence that complicated structures, such as missing pronouns (PRO), seem to take time to acquire. Hsu, Cairns and Fiengo (1985) found that, for sentences such as 'the boy hits the girl after (PRO) jumping over the fence', only adults were fully aware of the rule that the nearest c-commanding noun phrase controls PRO, whereas the children showed a developmental trend of 4 different grammar types in interpreting PRO.

2.3.A2 Lexical knowledge

Unlike definite articles, pronouns mark gender (him/her), number (she, they) and case (she, her). Menyuk (1969) found that 3 to 7 year old children still had not mastered 'gender' or 'case' marking. However, Scholes (1981), in a relatively complicated sentence-verification task where the subject had to match sentences to line drawings of touching human figures, found that between 4 and 7 years of age, 'gender' was acquired first, followed by 'number' and 'case'. Nevertheless, opposite evidence is found which suggests that 'number' precedes 'gender' acquisition (Pinker 1984). Thus, the evidence so far is very controversial.

B. NON-LINGUISTIC FACTORS.

Overall in this section, it seems that whilst young

children do appear to be using heuristic strategies, thematic principles and pragmatic inferences, they do not seem to be able to effectively co-ordinate these non-linguistic factors with their linguistic knowledge, explaining why they fail to produce unambiguous pronominal reference in situations which require an integration of contextual and grammatical factors.

2.3.B1 Thematic-Discourse structures

Karmiloff-Smith (1985,1986) found evidence to show that children under 6 years failed to discriminate between main and subsidiary characters by using pronouns only for the main character and definite noun phrases only for subsidiary characters. However, her older subjects (approximately 6 to 9 years) did make this distinction, plus they tended to refer to the main referent in the subject position of the sentence ('Thematic subject strategy'). Beyond this age, subjects continued to use a pronoun only to refer to the main character, but they did not rigidly pre-empt the subject slot for this referent: The oldest subjects were more flexible about the positioning of referents and used contextual information to decide upon appropriate positioning of main and subsidiary referents.

This finding conflicts with Limber's claim (1973) that, as subject noun phrases are highly foregrounded, even infants use more pronouns in subject position.

However, Tyler (1983) found evidence of pronouns being used for the main character by children under 5

years of age and she found that this strategy in fact decreased with age. Clibbens (1986, 1988) found no evidence of the thematic subject strategy, but he did find that children below 6 years of age made the distinction between main and subsidiary characters: They were more likely to choose a pronoun for the main character and they clearly preferred a full noun phrase for the subsidiary character. However, beyond this age, he found no evidence of subjects discriminating between main and subsidiary characters and they preferred full noun phrases regardless of the characters' roles.

Finally, Bartlett (1984) found that both skilled and unskilled adult readers were sensitive to thematic principles, but the main factor determining their choice of definite expression was a function of the potential number of referents available as opposed to the thematic roles of available referents.

2.3.B2 Pragmatic Inferences

Some studies have focused on interactions between syntactic constraints and pragmatic inferences: Umstead and Leonard (1983) found that young children were better at interpreting pronouns when they could rely on within-sentence syntactic constraints, than when they had to rely on outside-sentence pragmatic contexts. Balfour (1983) carried out an experiment involving three conditions. In addition to a 'neutral' condition, where pragmatics did not bias reference to one referent over another, he also used sentences where the pragmatic content

was favourable to a non-identity response: the actions in the two clauses were incompatible, so that the same person could not be doing both simultaneously. For example 'He built sandpies while Jerry was paddling'. Other sentences were presented, where the pragmatic information suggested an identity response, in opposition to the syntax, for example 'He was very cold when Tom slept outside'. His results that 5 to 6 year olds would appreciate non-identity and use pragmatic information, plus his finding of a U-shaped developmental curve for allowing pragmatic factors to override syntactic ones when pragmatic and syntactic interpretations were incongruous, suggests that before 6 years of age, children's choice of a referent was guided predominantly by pragmatic factors. Beyond this age, they are able to understand the non-identity restriction, but pragmatic factors were still more salient to them. 9 year olds used strict linguistic criteria and failed to consider pragmatic inconsistencies (similar to Karmiloff-Smith's phase 2: rigidly guided by linguistic rules at the expense of relevant extra-linguistic factors). Finally, older children and adults reverted to plausible references when linguistic and pragmatic factors were conflicting. Thus, pragmatic factors do seem to have strong effects on the performance of both children and adults.

Moreover, the predominant role of pragmatic factors tends to support Reinhart's interpretation (1983). Reinhart argued that a referential interpretation of a pronoun is essentially a pragmatic rather than a sentence-level problem. She claimed that speakers and

hearers use the strategy: a syntactically constrained sentence should be used if it is possible to provide information about co-reference, but if a sentence structure does "not allow bound anaphora", then whether the noun phrases are interpreted as co-referential or not can only be determined on the basis of discourse information.

However, spontaneous comments from Balfour's subjects showed that, whilst adults commented on the unnaturalness of the sentences, younger children did not comment on the conflict. Balfour suggested that young children are probably using a 'purely semantically-based strategy.....while the older children use the semantic information in preference to what they know the syntax demands'. This finding also supports Chomsky's claim from her own experiment that older subjects were more motivated on the task, because they were aware of the differences between restricted and non-restricted sentences and thus, they produced comments such as 'It could be either one' for the non-restricted sentences, whereas the younger child appeared to view the task as arbitrary. Chomsky claimed that, for the younger children, "the skill itself was perhaps less clearly defined than in the case of the older children". This result suggests that allowing linguistic and pragmatic factors to interact in pronominal interpretation is probably a late acquisition.

Other experiments in this area have tended to compare pragmatic inferences with lexical knowledge, more specifically with that of 'gender' marking ('she' and 'he'). The experiments of Wykes (1981, 1983), Tyler (1983),

Tyler and Marslen-Wilson (1982, 1983) have made these comparisons, but the results so far have provided conflicting evidence, as the following discusses in more detail: Wykes used comprehension experiments, and she gave 5 year olds sentences such as:

1a. Jane found John's ball. 1b. She gave it to him.

2a. Jane found Susan's ball. 2b. She gave it to her.

The child's task was to act out the second sentence with puppets. Wykes found that when the antecedents were of different gender (1), the child could assign the pronoun to the referent which lexically matched the pronoun. However, performance was poorer when the antecedents were of the same-gender (2) and they could only assign the pronoun to an antecedent by making plausible inferences (that the person in possession of the ball would normally give it to the person who had lost it and therefore, the antecedent of 'she' would be 'Jane'). In another experiment, she also found that this difference was due to the complexity of making the inference rather than due to storage of premise information.

However, Tyler et al obtained results contrary to those of Wykes. They were interested in on-line interactive procedures and hence, chose to use a 'mispronunciation detection task', the rationale behind this being that, reaction time to detect a mispronounced word is facilitated when the actual word is more contextually predictable. They presented sentence pairs such as the following:

Mother saw the postman coming from a distance

He brought a LEFFER(letter) from Uncle Charles.... It was predicted that, if the child correctly assigned the pronoun in the second sentence to the postman, they should have quickly detected the mispronounced word (leffer), because of their expected association between 'letter' and 'postman'. Tyler found that 5 year old children were slower at detecting the mispronounced word when a pronoun was used, than when a complete noun phrase was presented. In three experiments of this type, they concluded that the problem for young children was not in making plausible inferences, but was due to their inability to fully exploit the lexical information carried by the pronoun.

Nevertheless, the results of both these studies are tentative; Firstly, because the type of tasks given were rather complicated for 5 year old children and it is hard to say whether problems were a reflection of more general difficulties with the actual tasks. Secondly, pragmatic inferences were never directly manipulated by varying the ease with which such inferences could be drawn. In fact, comparing lexically ambiguous and unambiguous pronouns may have obscured any genuine use of pragmatic inferences.

Stevenson and Pickering (1987) manipulated a number of factors and found that young children did not combine different sources of knowledge, syntactic (binding principles), lexical (gender marking) and pragmatics (plausibility), when interpreting pronouns.

Finally, Light and Capps (1986) found no differences between younger and older adults in their use of pragmatic inferences to assign pronouns.

2.3.B3 Heuristic principles

Interest in this area is concerned with the strategies that children use when comprehending pronouns in sentences. There is evidence to suggest that children make use of the way that the referents are positioned: Some studies have shown that young children tend to assign a pronoun to the 'subject' (Chai 1967, Wykes 1983, Crawley and Stevenson, (in press)). Others have claimed that children prefer to assign pronouns to the 'object' (Chomsky 1969: 'minimum distance principle'). Others have argued that children use a 'parallel function strategy' and choose a 'subject' antecedent when the pronoun is in 'subject' position, but choose the 'object' when the pronoun is in 'object' position (Grober, Beardsley and Caramazza 1978, Kail 1976, Sheldon 1974). Still, others have argued that a young child's choice of referent is dependent upon parallel semantic roles between the pronoun and antecedent (Ferreiro 1976, Pruner and Solan 1979). Finally, there is evidence that children take time to develop the rule that one should reverse the usual strategies whenever contrastive stress is given to the pronoun (Maratsos 1973, Solan 1983). Nevertheless, as Cromer (1976) pointed out, children's strategies are only useful 'tools' for understanding their knowledge of pronominal reference, when children use different strategies to adults. Yet there is evidence that children and adults often use similar strategies (Kail 1976, Karmiloff-Smith 1980, 1985).

Finally, Kail and Léveillé (1977) found that young children had problems using lexical information when they

could use the parallel-function strategy.

2.3.B4 Knowledge of ambiguity

Kail and Léveillé (1977) and Farioli (1979) found that adults, as opposed to children, would readily detect referential ambiguity of pronominal reference. Van Hekken, Vergeer and Harris (1980) carried out an observational study of pairs of children between 4 and 6 years and found that they tended to use pronouns ambiguously ('that one'), when non-verbal disambiguation was impossible (invisible items). Moreover, listeners of the same ages failed to respond any differently when ambiguous pronouns were used.

Finally, Bartlett (1981, 1984) found that a difference between good and poor readers and younger and older adults was that, when there were two potential referents and thus, when one referent had to be discriminated from the other to disambiguate reference, more nouns and appropriate modifiers were used by skilled readers and older adults. Bartlett argued that these differences could either be due to processing capacities or due to the fact that the latter were more developed in their knowledge of linguistic objectivity.

2.4 Summary and overview of studies on children's knowledge of indefinite and definite reference.

Looking first at studies of children's knowledge of the articles, the evidence suggests that, when first using the articles, young children are aware of the 'familiarity'

factor. When the child's perception of the listener's knowledge was carefully manipulated and pragmatic bias was eliminated, the young child's choice of article depended on the listener's perception of the referents: When the listener was 'ignorant', an indefinite article was preferred, whereas when the listener was 'knowledgeable', a definite expression was preferred. Thus, it is necessary to test children's sensitivity to the listener by manipulating listener knowledge in a simple story telling task where pictures are unambiguous for young children. Furthermore, in order to find out whether using experimenters as listeners does truly affect the way that young children perceive a task, an experiment which uses the experimenter as listener should reveal whether the performance of younger subjects is deteriorated compared with when peers are listeners. Moreover, the child's appreciation of 'familiarity' seems to be related to the linguistic context in the sense that even young children have been shown to use 'novel definites' when the pragmatic context implies that the referent already exists. An experiment is needed to look at whether, when familiar contexts are used, young children will allow general knowledge factors to override immediate situational factors in determining their choice of indefinite or definite reference.

With regards to the 'specificity' factor, from the results of tasks using a concrete array (exophoric reference), it seems that the nature of the task used determines whether young children will discriminate between

specific and non-specific referents. For example, tasks requiring indirect reference and comprehension tasks seem to produce more positive results than do production tasks which look at direct reference. One possible explanation for these task differences may be that children perform better when the referential array is salient to them. What is required is an experiment which manipulates both listener 'familiarity' and referent 'specificity', to see whether young children are able to consider the significance of both these factors in determining form of reference. Moreover, a comparison of tasks is needed to see whether it is those factors which increase the salience of the referential array that bring about better results from young children.

From previous experiments of anaphoric reference, it appears that young children are not aware of the significance of 'specificity' because they have been shown to produce ambiguous definite reference on second mention and they fail to use relevant presuppositional information, such as modifiers and relatives. Requiring subjects to refer anaphorically in a referential communication task should reveal any evidence of the child's knowledge of unambiguous anaphoric reference. Moreover, if message evaluation follows a similar developmental pattern to message production, then both comprehension and production tasks are required.

The results of experiments on children's pronominal reference are conflicting. Overall, it seems that young children have little difficulty with linguistic

constraints and have greater difficulty when they are required to integrate non-linguistic and linguistic factors in assigning the pronoun to an antecedent. Referential communication requiring pronominal reference will therefore require careful control over the various possible linguistic and non-linguistic factors involved. Moreover, if 'familiarity' and 'specificity' apply to all forms of definite reference, then what is needed is a combination of tasks looking at both pronominal reference and reference with definite articles, in order to see if they both yield similar results.

As referential communication tasks have often been criticised as unnatural and creating an artificial situation for the child, it is also necessary to use a variety of experiments to see whether children's performance can be improved in more game-like naturalistic tasks.

Karmiloff-Smith (1979, 1980) has argued that developmental effects of indefinite and definite reference are independent of cognitive abilities, but instead are a result of a particular linguistic development: She argued that the reason why children develop from knowledge of deixis to a knowledge of anaphora is because they have to acquire the different plurifunctions of the indefinite and definite articles as opposed to regarding these determiners as each performing only a single function. In her account of young children's interpretation of 'same' as meaning 'same kind', she tried to explain these effects as due to the type of linguistic input that the child received and

that 'same' appears in a typical modifier position in normal linguistic input. One other explanation she put forward was that it may be that young children were more obsessed by similarities of attributes. Nevertheless, she failed to mention the possibility that the latter may suggest that knowledge of indefinite and definite reference is predominantly a reflection of general cognitive development. However, Karmiloff-Smith argued that there were linguistic universals of indefinite and definite reference, hence justifying her generalisation of results from French speaking children. Swinney and Prather (1981) stated in a review of Karmiloff-Smith (1979), "The reader will wish for more substantive detail on and explicit definition of the nature of the language-specific functional development that is posited to occur independently of logical-cognitive abilities".

Emslie (1986) carried out a number of experiments which led her to conclude the reverse of Karmiloff-Smith: that from 4 years of age, children could appreciate the identifiability of indefinite expressions and anaphoric conservation with definite expressions. She argued that any age differences were due to differences in task perception. Nevertheless, like Karmiloff-Smith she argued that knowledge of indefinite and definite reference was not related to cognitive development and thus she failed even to speculate on the possibility that task variations in performance were a reflection of the development of linguistic objectivity.

Some studies suggest that there may be cultural and

linguistic variations in in/definite development. For example, Farideh and Houssain (1986) found that Iranian children between 3 to 5 years of age were better at indefinite and definite reference than English children of the same age. Moreover, as already stated, studies of pronominal reference suggest that some linguistic acquisitions require 'parameter setting' (Solan 1983). Nevertheless, the developmental literature suggests that there are probably some universal linguistic achievements of indefinite and definite reference and it is possible that these depend on the development of the appearance/reality distinction and the emergence of an objective propositional attitude. Thus, perhaps it is more valuable to look at knowledge of indefinite and definite reference in terms of universal intellectual achievements, rather than focusing too much attention on those differences which do appear to be specific to culture and/or language.

Recent investigations of discourse processing have pointed to the way that certain interactions between linguistic and contextual factors take place when processing anaphoric expressions. Experiments concerned with the way that both children and adults use linguistic and non-linguistic factors in their interpretation and use of definite noun phrases are needed for looking more closely at the kind of processes required for anaphoric reference, as well as to see whether developmental effects are a reflection of the young child's failure to integrate several knowledge sources effectively.

2.5 Summary of the current experiments

The next four chapters report 16 experiments related to the question of children's knowledge of indefinite and definite reference. Chapter 3 looks at the effects of listener 'familiarity', chapter 4 is concerned with the effects of referent 'specificity', chapters 5 and 6 focus specifically on reference in discourse, with chapter 5 looking mainly at production and chapter 6 focusing on comprehension. Finally, chapter 7 is concerned with the interaction between linguistic and non-linguistic factors for definite reference.

In chapter 3 'the effects of listener 'familiarity'', 3 experiments are reported which look at whether children will use indefinite and definite expressions as a function of the knowledge of the listener. Experiment 1 looks at a wide age range of subjects and manipulates whether a peer listener is knowledgeable/ignorant in a picture-story task. Experiment 2 looks at the young child's sensitivity to an experimenter listener in both the picture-story task and in an elicited-discussion of real events. Finally, experiment 3 uses the picture-story procedure again but, regardless of the perceptual information provided for the peer listener, some referents can be assumed to exist by making pragmatic inferences using the contextual information provided in the pictures. This latter experiment is a test of the young child's use of general knowledge when defining 'familiarity' for indefinite and definite reference.

Chapter 4 'the effects of referent 'specificity'',

describes 2 experiments which involve exophoric reference (to a perceptual array of items). Experiment 4 manipulates listener 'familiarity' as well as the 'specificity' of referents and experiment 5 compares specific/non-specific reference to an array of objects which are made salient to the speaker with reference to an array of objects which are not made salient. The latter experiment is concerned with whether any evidence of a child's sensitivity to 'specificity' is task-specific and dependent on the salience of the referential context.

'Reference in discourse production' is the primary interest of chapter 5. Experiment 6 manipulates 'familiarity' and 'specificity' in a picture-story experiment and looks at how children identify and subsequently refer anaphorically. Experiments 7 to 9 look predominantly at anaphoric reference for referential communication and are interested in whether subjects will use relevant modifiers and relative clauses so that referents can be unambiguously differentiated. Experiment 10 uses a more naturalistic communication task and looks at pronominal reference as well as reference using definite articles. Finally, experiment 11 looks at pronominal reference more closely to see whether children use lexical factors, make pragmatic inferences and are sensitive to referential ambiguity when using pronouns.

Chapter 6, 'Interpreting anaphoric reference', consists of 2 comprehension tasks which are compared with the results of the production experiments. Experiment 12 looks at the comprehension of ambiguous reference using

definite articles and experiment 13 is concerned with the comprehension of ambiguous pronouns.

Finally, chapter 7 attempts to compare those aspects of linguistic knowledge which are constrained syntactically with the meta-linguistic knowledge required to dissociate linguistic from non-linguistic factors for definite reference. Experiment 14 looks at children's comprehension of one syntactic constraint known as the 'non-identity restriction' (Chomsky's binding principle C), to see whether the child's use of such constraints is dependent on pragmatic context. The last two experiments in this chapter (experiments 15 and 16) are story-completion tasks aimed to test Crain and Steedman / Altmann and Steedman's theory that adults process linguistic and non-linguistic factors interactively using the particular contextual information about the number of available referents to decide upon which choice of definite construction to use. Interest in this experiment is in whether adults and older children do appear to be performing according to this model of discourse processing and whether 'specificity' does require this particular interaction between linguistic and non-linguistic factors.

CHAPTER 3

THE EFFECTS OF LISTENER 'FAMILIARITY'

3.1 ABSTRACT

Experiment one was a partial replication of an experiment carried out by Emslie (1986). Children were required to produce narratives from a series of pictures, to a listener who was either unable to see the pictures (masked) or who could see the pictures (unmasked). The results showed that, for introducing referents, 3 to 5 year olds used predominantly indefinite articles when the listener was masked, whereas they used mainly definite descriptions when the listener was unmasked. Experiment two looked at the ability of young children to treat the experimenter as a listener and to see if their sensitivity to the experimenter listener varied according to the task. It was found that, for the story-telling task, 3 to 5 year olds used definite expressions on first mention, even when the experimenter pretended to hide from the pictures. However, when discussing real-life experiences, they used definite expressions on first mention when the experimenter had shared the experience with the child, but they used indefinite expressions on first mention when the experimenter had not shared the experience.

Finally, experiment three looked at whether children's choice of expression would depend on pragmatic inferences from the story context regardless of the listener masking conditions. Even young children were predominantly influenced by their general knowledge of

contexts and used definite expressions on first mention of a referent when the story implied the existence of that referent, irrespective of whether the listener had seen the pictures.

It seems that very young children are aware of the significance of social factors in determining whether to use indefinite or definite expressions. Moreover, their definition of 'familiarity' depends on their everyday experiences with different types of listeners and different types of contexts as well as the perceptual ability of listeners in the immediate situation. The only age differences were that with age, there appeared to be an increasing use of conventional story procedures for defining 'familiarity', a greater ability to co-ordinate the particular features of the ongoing experimental situation with general pragmatic knowledge and a decrease in ambiguous pronominal reference.

3.2 GENERAL INTRODUCTION

The three experiments in this chapter are concerned with one of the distinctions between in/definite reference: FAMILIARITY. Thus, interest is in the way that children of different ages choose between indefinite and definite reference as a function of their sensitivity of the listener's 'familiarity' with referents. Experiments 1 to 3 look at whether children use the social and pragmatic context of discourse to define 'familiarity' and hence, whether these factors affect their choice of indefinite or definite reference. Some investigators have found that

young children tend to overuse the definite article when the listener is unfamiliar with the referents (Brown, 1973, Maratsos, 1973, Karmiloff-Smith, 1979, Warden, 1974, 1976, 1981) and these results led to the conclusion that young children were egocentric and unable to take the listener's perspective into account when using indefinite and/or definite articles.

However, recently, not only has the notion of egocentricity been redefined (Donaldson 1978), but Emslie (1986) has found that by explicitly manipulating the social factors of the experiment, children as young as 3 years would use indefinite expressions when speaking to an ignorant listener and definite expressions for a knowledgeable listener. In experiment 1, using a wide age range of subjects, children were required to tell stories from pictures to a peer who either had or who had not seen the pictures before. If the children's choice of expression depends on whether the listener has been provided with the necessary perceptual information, then one would expect them to produce mainly indefinite expressions to a masked listener and mainly definite expressions to an unmasked listener.

Emslie argued that the reason why Karmiloff-Smith failed to elicit indefinite expressions from subjects is because she did not consider the social context in her experiments: Firstly, Karmiloff-Smith did not explicitly manipulate the knowledge of the listener and instead, often assumed that the listener was being treated as ignorant, despite the fact that sometimes the listener could infer

the existence of the referents on the basis of perceptual information. Secondly, Karmiloff-Smith herself usually served as both experimenter and listener and so often she had to pretend to be ignorant (because if she had really been ignorant she would not have been able to observe and record the performance of subjects). Therefore, it is possible that subjects realised that the experimenter was not really ignorant and thus, failed to treat her as such in the experiment. Furthermore, young children may be reluctant to treat an elder as ignorant because they are only experienced with situations where they learn from adults who are more knowledgeable than they are.

However, Karmiloff-Smith herself argued that, the way that subjects perceive the experimental task is a developmental phenomenon in itself. She found evidence of young children using "pragmatic procedures" at the expense of appreciating the events, objects and persons in the particular experimental context. Thus, young children apply their own everyday experiences to experimental tasks and hence, they often interpret the task in a different way to that which is intended by the experimenter.

If young children do appreciate the importance of listener 'familiarity' for in/definite reference, it is possible that they will also use pragmatic procedures to define the social context, in which case, pragmatic experiences may influence their decisions regarding in/definite reference. For experiment 2, which looked only at 3 to 5 year old children, again the story-telling procedure was used, but speakers had to

narrate a story to an experimenter who pretended not to have seen the pictures before. Also, an elicited-discussion procedure was used to get subjects to talk about real experiences to an experimenter who had either shared or who had not shared the experience with the subject. If young children's ability to use indefinite expressions depends on their ability to treat a listener as unfamiliar with referents, then their choice of expression will depend on whether they can accept the experimenter as an ignorant listener. More specifically, one would expect them to only treat the experimenter as unfamiliar with referents when talking about real experiences for which the experimenter was genuinely ignorant.

Emslie also questioned the results of Warden (1976), who showed evidence of young children being overdefinite, because it seems that some of his observed definite scores were due to referents being predictable from context, irrespective of the manipulation of the knowledge of the listener. Emslie argued that by using Hawkins' (1978) classification of appropriate uses, first mention definites are appropriate where the context presupposes the existence of that entity (for example, 'Associative Anaphoric Use': A shop -The assistant, where the context of a prototypical shop means that one can infer the presence of other items/individuals not explicitly mentioned, one of which would be a shop assistant).

However, whereas Emslie (1986) has concluded that children (and adults) are not influenced as much by pragmatic factors as they are with the status of the

listener in the current situation, others have found evidence to the contrary (McWhinney and Bates 1978, Zehler and Brewer 1982): that general knowledge of contexts has the greatest influence on children's choice of expressions.

Previous studies have shown that young children are very good at making pragmatic inferences (Johnson and Smith 1981, Nelson 1986), whilst other investigators have claimed the opposite (Paris and Lindauer 1976). In chapter 2, it was concluded that young children will use pragmatic inferences whenever it is possible as long as they have sufficient experience to know what can be inferred. Experiment 3 followed the procedure of experiment 1, with speaking partners telling picture-stories to either a masked or an unmasked peer, but interest was in the way that subjects referred to an entity which was predictable given the context of the story. If the strongest factor in determining the child's definition of the 'familiarity' of a referent depends on predictability from context, then one would expect children to use definite noun phrases on first mention of such predictable referents, even when the listener has had no opportunity to see the referents.

The two main questions underlying the above experiments are:

1. At what age do children begin to choose between indefinite and definite expressions as a function of listener 'familiarity'?
2. In what way do children of different ages use the context to define the 'familiarity' of the listener?

3.3 EXPERIMENT 1: Children's use of indefinite and definite expressions as a function of the knowledge of the listener.

INTRODUCTION

The following experiment was an attempt to replicate one of Emslie's experiments (1986). The knowledge of the listener was manipulated by either having a masked (ignorant) or an unmasked (knowledgeable) listener. Furthermore, the experiment looked at children between the ages of 3 to 11 years, as there remains the possibility that Emslie's selection of subjects may have obscured any real developmental differences that have been observed by Karmiloff-Smith (1985).

If Emslie's results are truly representative of the way that young children discriminate between indefinite and definite reference, then one would expect even the youngest children to use the indefinite article for introducing referents to a masked listener and to use definite articles only for introducing referents to an unmasked listener and for mentioning referents for the second time. However, if such results are to indicate that the development of the distinction between indefinite and definite expressions is virtually complete by the age of 5 years, then a similar pattern of reference should also appear for older subjects.

METHOD

Subjects

80 subjects took part in the experiment, with there being 20 subjects in each of the following age groups: 3 to 5 years (mean age 4.7), 6 to 7 years (6.6), 8 to 9 years (9.9) and 10 to 11 years (10.8).

Apart from the 6 subjects that were selected from the University nursery, most children were from a primary school in Gilesgate, Durham. Each child was picked on a random basis (each child in a class was assigned a random number and the first 20 children selected were used as subjects). There was approximately an equal number of males and females in each group. Each subject was randomly paired with another subject of approximately the same age. So there were 10 pairs of subjects (dyads) in each age group.

Materials

8 sets of picture stories and 2 practice stories were designed for the experiment, with 3 pictures making up each story.

The first picture of each story introduced the main character. The second picture depicted a scene which included the referent from the first picture plus a second subsidiary character. The third picture showed the outcome of the events of the first and second pictures and principally, this picture involved the main referent of the story. The introduction of a subsidiary character and the use of three pictures rather than two was to prevent

subjects from using one sentence for the entire episode, as opposed to using discourse.

For example, the following story is a hypothetical discourse representation of one of the stories used in the experiment:

P1: A man is fishing by a river. P2: A boy walks by and notices something pulling on the rod. P3: The man pulls up the fishing rod and there is a fish at the end of it.

For this story, if 'the boy' had not been introduced in the second picture, the child could easily have translated the events into one sentence, such as, 'A man is fishing and.....there is a fish at the end of the rod.' However, if there are at least two animate referents in the story, the child is probably more likely to refer to the main referent twice and the addition of P2 means that the speaker must shift reference back to the initial character in P3.

When constraints are imposed upon children, there is always the danger of creating an artificial situation that is more representative of the way children approach a particular experimental task, rather than being typical of the way children choose to refer in more naturalistic situations. In order to minimise this possibility, the stories were constructed by considering those events which children normally experience in their everyday lives. Moreover, as emotional responses can affect performance, no events were used which could have been regarded as

frightening or upsetting.

(See figures 1.1 - 1.8 for picture stories).

A 'Listener Mask' was also used for the experiment. This was a funny face with big ears. It was hoped that the comical aspect of the mask should make the experiment seem more like a game and the emphasis on the ears was intended to encourage the children to focus on the listening role of the masked subject.

Projecter apparatus for enlarging the pictures onto a screen was also used in the experiment. This was to keep the intervention from the experimenter at a minimum and to keep the attention of the speaker on the listening peer, rather than on the experimenter.

Tape recording equipment was also used for recording the children's narratives.

Design and Procedure

In each age group, half the subject dyads were assigned to the 'Listener Ignorant' condition and the other half were assigned to the 'Listener Knowledgeable' condition.

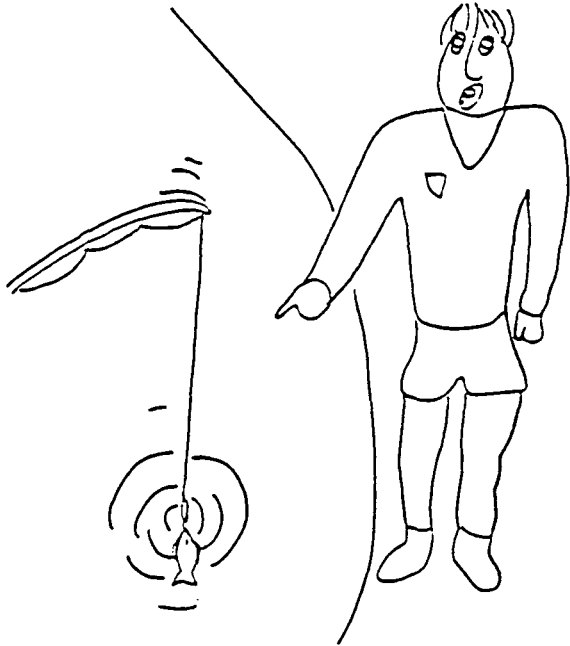
The 8 stories were presented to each pair, with each subject acting as speaker for 4 stories and as listener for 4 stories (alternating between speaker and listener for each trial). The slides were removed from the projecter before each new pair of subjects took part in the experiment. Cards numbered 1 to 8 were put into a hat and the order in which they were selected (by a member of

FIGURE 1.1 PICTURE STORY 1

First picture



Second picture



Third picture

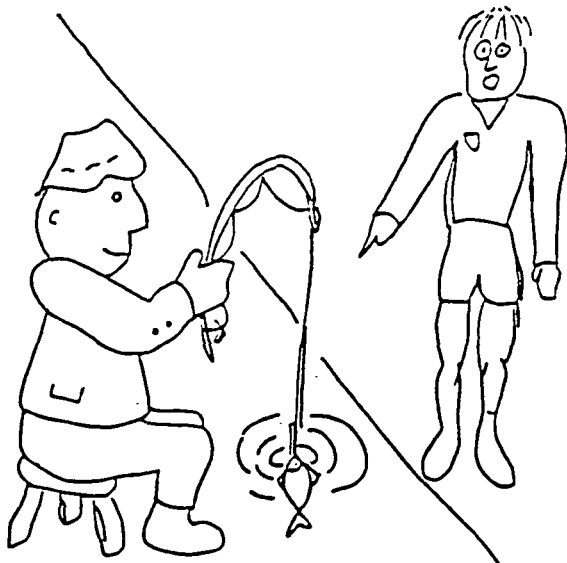
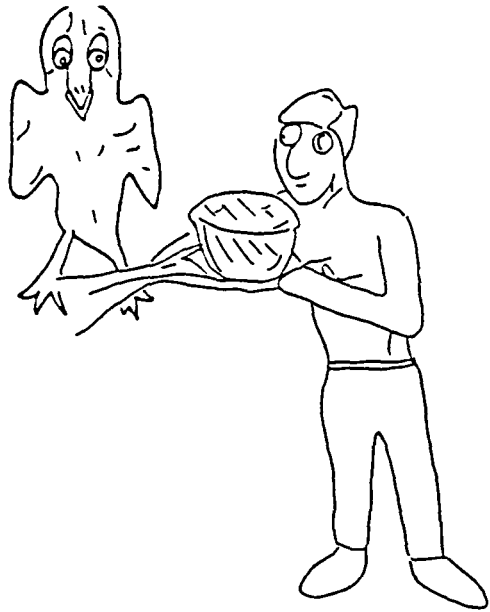


FIGURE 1.2 PICTURE STORY 2

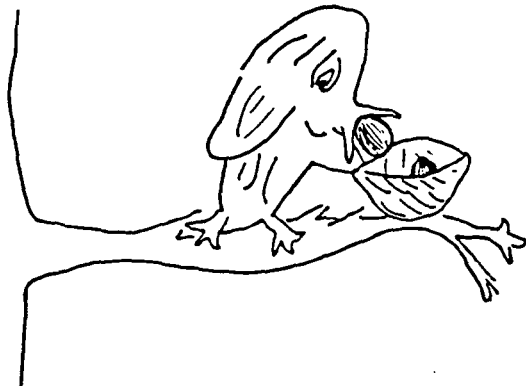
First picture



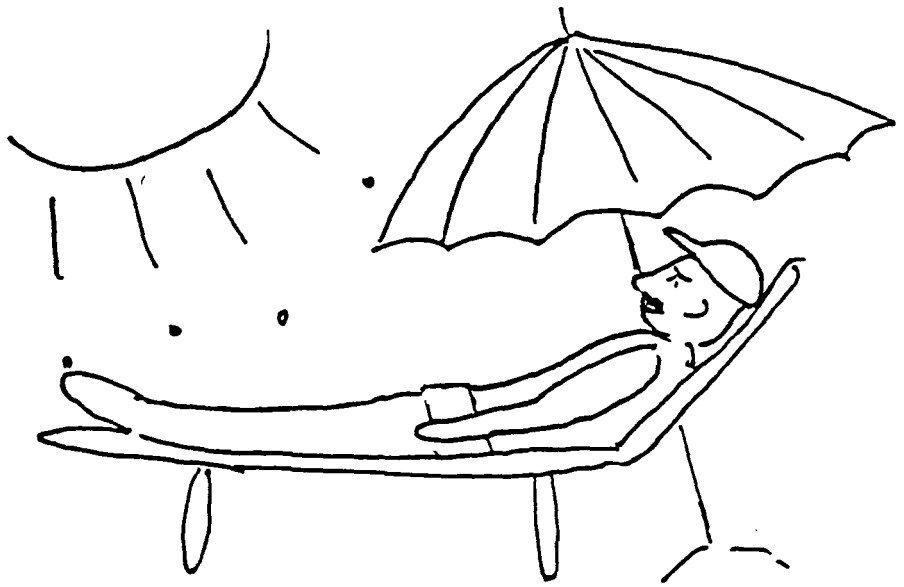
Second picture



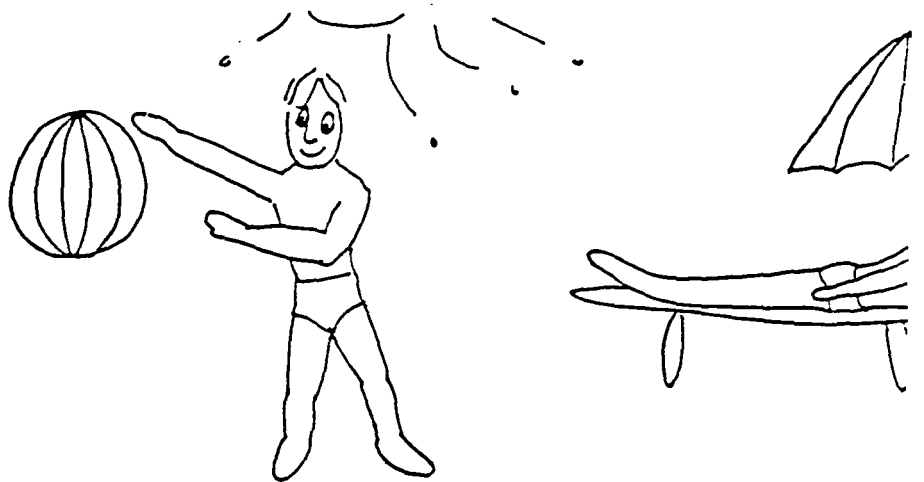
Third picture



First picture



Second picture



Third picture

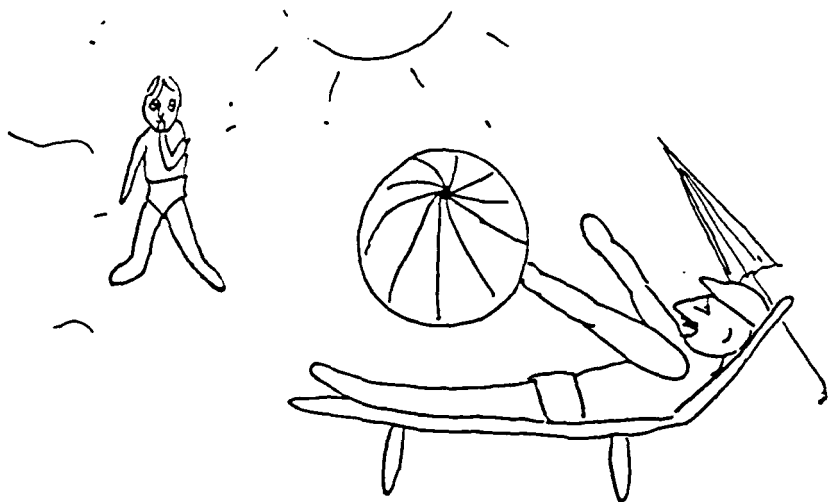
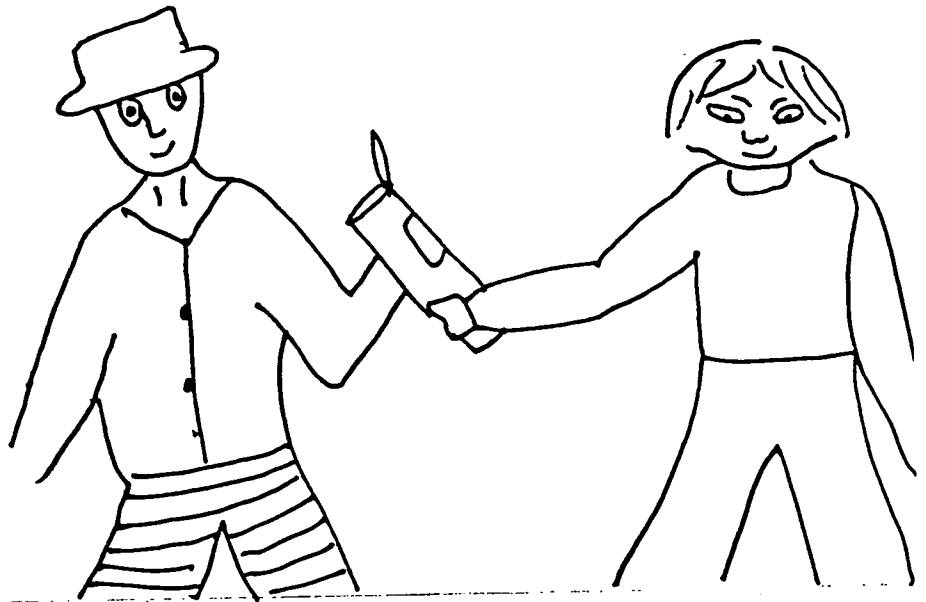


FIGURE 1.4 PICTURE STORY 4

First picture



Second picture



Third picture

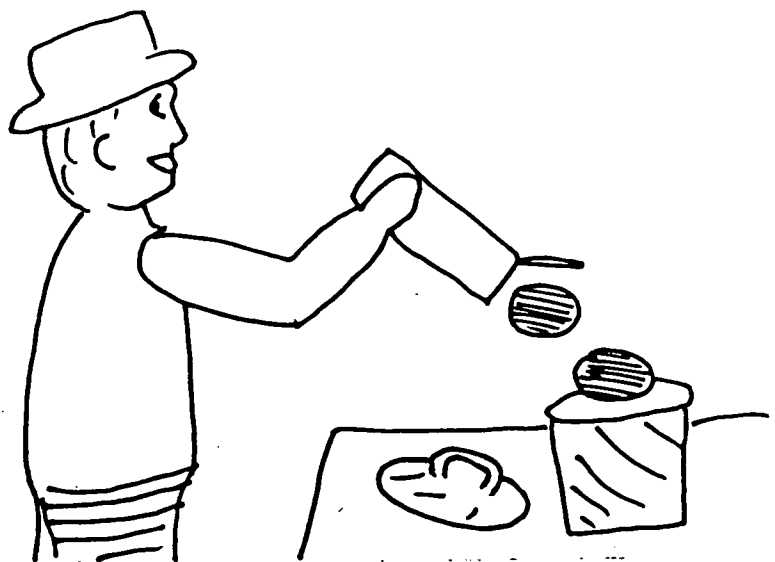
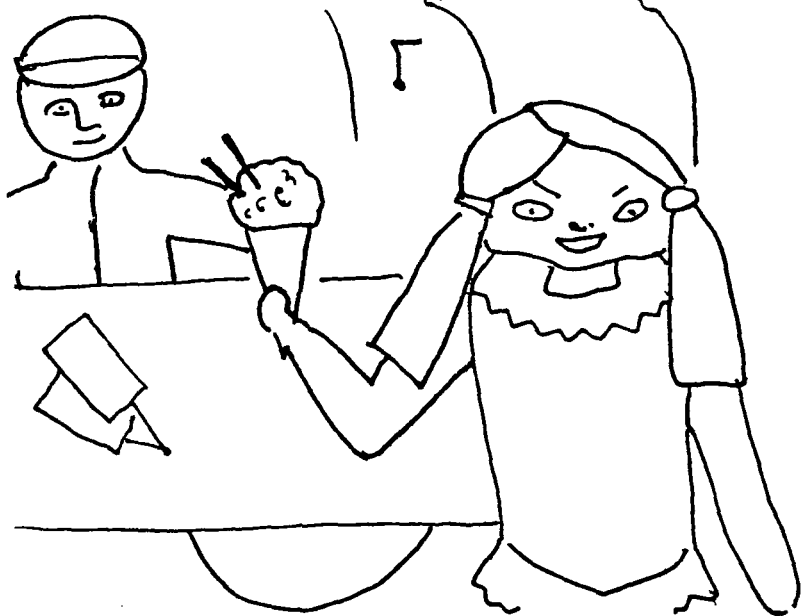


FIGURE 1.5 PICTURE STORY 5

First picture



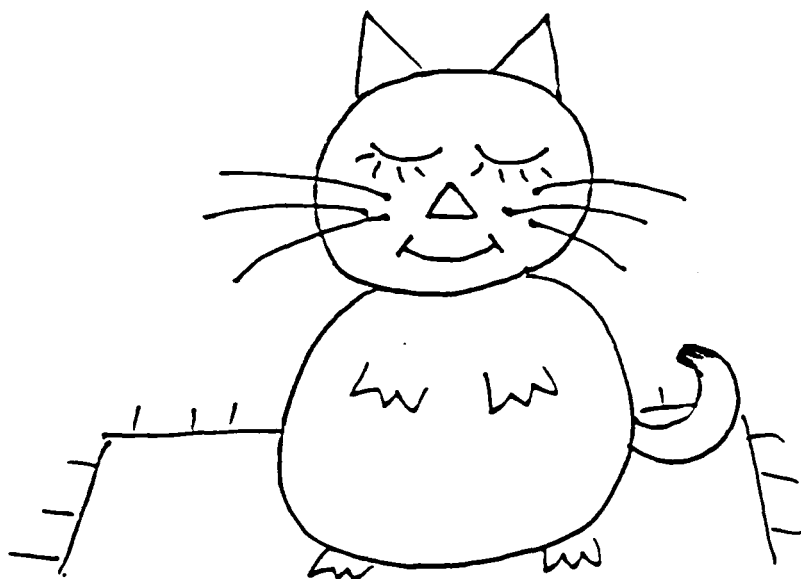
Second picture



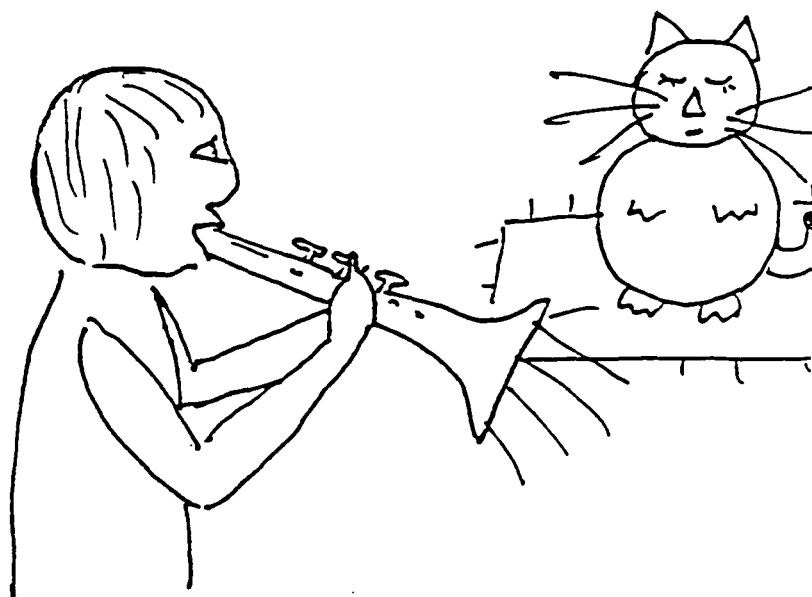
Third picture



First picture



Second picture



Third picture

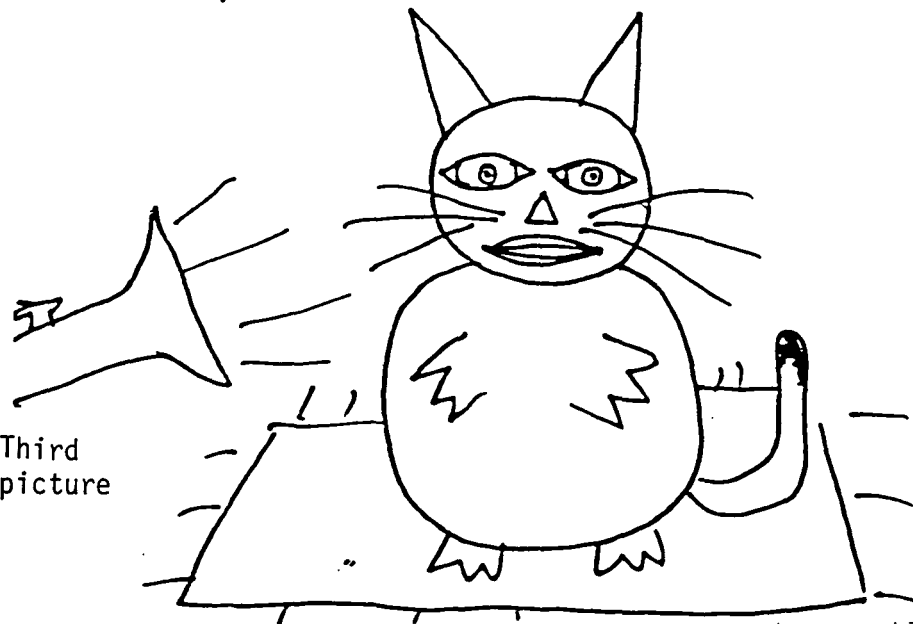
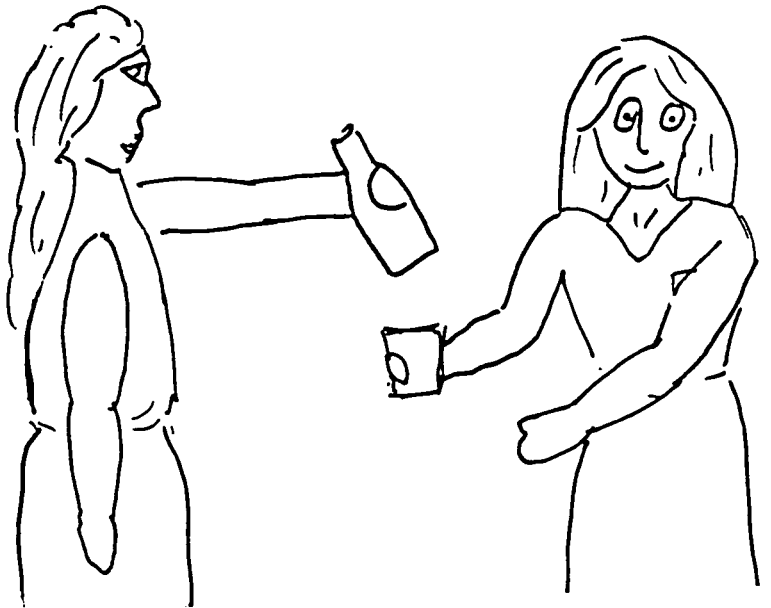


FIGURE 1.7 PICTURE STORY 7

First
picture



Second
picture



Third
picture

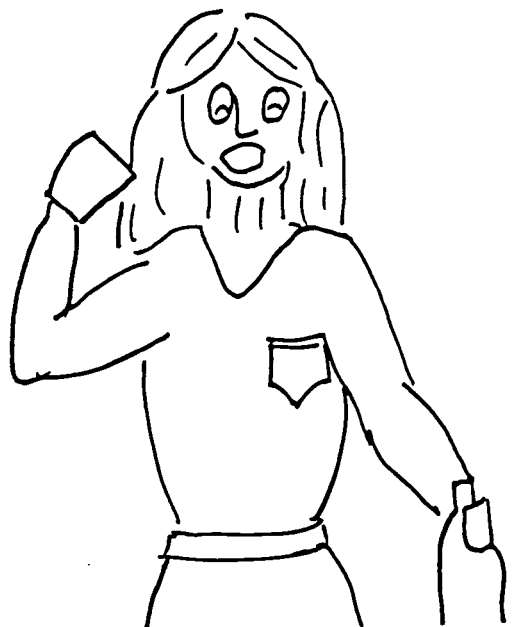
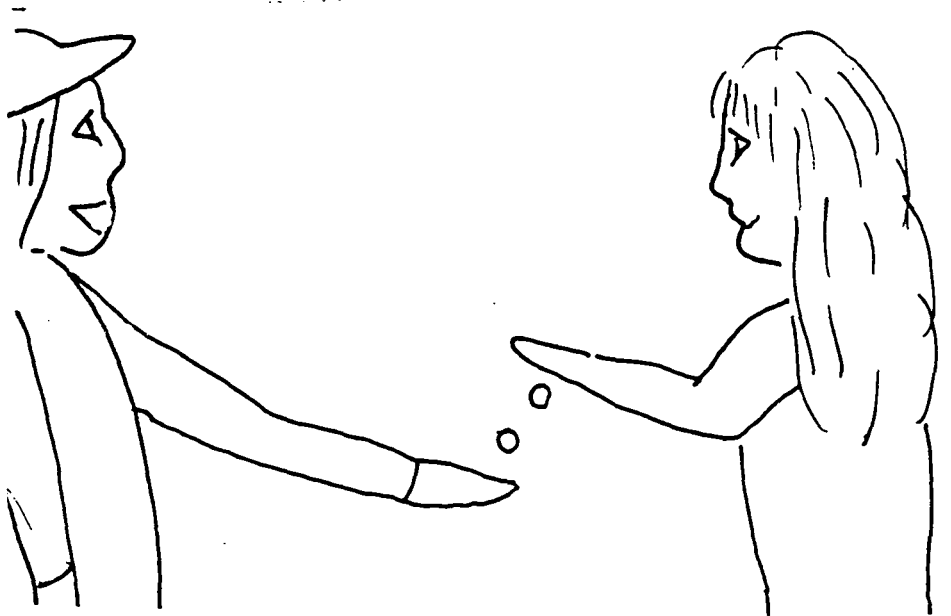


FIGURE 1.8 PICTURE STORY 8

First picture



Second picture



Third picture



teaching staff), determined the order in which they were presented for each pair of subjects.

A room which was familiar to the children was used for carrying out the experiment. An adjoining playroom to the nursery was used for the pre-schoolers and a room which was normally used for indoor games and watching videos was used for all other subjects.

Ignorant listeners were masked during the initial presentation of the slides and remained masked whilst the speaker was telling the story. Knowledgeable listeners were only masked whilst being told the story. Therefore, listeners in both conditions were masked whilst the speaker narrated the story. This was so that the speaker did not rely on deictic (pointing) gestures and to prevent listeners from providing non-verbal feedback to the speaker.

Phase 1: Instructions.

The following standardised instructions were given to each pair of subjects:

'We are going to play a little game. First, I will show you some pictures on the screen. These pictures make up a story. The first picture is the first part of the story, the next picture is the second part of the story and the third picture is the end of the story'.

At this point the experimenter presented the 3 slides of one of the practice stories.

'Now you are going to take it in turns to tell each other about the story in the pictures. I want to see how

good you are at telling each other stories. The story you tell must be the same as the story in the pictures'.

To subjects in the 'Listener knowledgeable' condition (LK), the following instructions were then given:

'When (speaker's name) tells the story, (listener's name) has to wear this listening mask. This is so that I can see who is doing the listening'.

To subjects in the 'Listener ignorant' condition (LI), the instructions were:

'(Listener's name) has to wear this mask. This is so that I can see who is doing the listening.' The listener was instructed to remove the mask only after the speaker had finished telling a story.

To subjects in both conditions, the experimenter then gave the final instructions:

'You must not begin to tell the story until you have seen the three pictures once. When I show you the pictures again, you can tell the story to (Listener's name).'

Phase 2: Running the experiment.

Practice stories were used until the experimenter was satisfied that subjects were performing at an optimum level. Each picture was shown on the projector screen for approximately 20 seconds on first presentation, but subjects were allowed as much time as they needed for each picture when they were actually telling the story.

The speakers' narratives were recorded on a tape recorder that was behind the projector. Thus, recording equipment

was hidden from the subjects' field of vision, behind the projector. The recordings were later transcribed for the analyses.

Scoring

Expressions were recorded separately for reference to MAIN and SUBSIDIARY animate referents. The expressions used to refer on first and second mention were scored into the following categories:

INDEFINITE:

Indefinite article - 'A man was fishing'

Other indefinite - 'There was THIS man fishing' (Any nominative indefinite 'There's a man there' was not counted as an indefinite score).

DEFINITE ARTICLE: - 'THE man was fishing'

PRONOUN: - 'HE was fishing' or demonstrative - 'THAT man was fishing'

No scores were given for references which involved only single words (eg. 'him' 'there' 'that one').

RESULTS

(Raw data and statistical analyses tables for each experiment are presented in the appendix)

A) MAIN RESULTS

First mention

Firstly, no subject persisted with single word answers or nominative expressions. Thus, each subject received a total score of 4 on first mention. Table 1.A shows the mean scores for indefinite and definite expressions in each age group, when introducing the main referent for the first time, in both LI and LK conditions.

(total out of 4) LI			LK		AGE MEAN	
AGE	INDEF.	DEF.	INDEF.	DEF.	INDEF.	DEF.
3-5	2.4	1.6	0.7	3.3	1.6	2.5
6-7	3.3	0.7	4.0	0	3.7	0.4
8-9	4.0	0	1.6	2.4	2.8	1.2
10-11	1.6	2.4	2.7	1.3	2.2	1.9
TOTAL MEAN	2.8	1.2	2.3	1.3	2.6	1.5

TABLE 1-A: MEAN INDEFINITE AND DEFINITE EXPRESSIONS ON FIRST MENTION.

Overall, there was a clear preference of indefinite expressions irrespective of the listener group and this was confirmed by analysis of variance, where a significant difference between indefinite and definite expressions was found ($F=32.12$, $df=1,72$, $p<0.001$). However, there was a significant interaction between indefinite/definite and listener group ($F=8.59$, $df=1,72$, $p<0.005$). The table also shows that most indefinite scores were from the 6 to 7 year olds and the most definite expressions were from the 3 to 5

year olds. Analysis of variance showed an interaction between age and indefinite/definite expressions ($F=21.57$, $df=3,72$, $p<0.001$). However, both the 3 to 5 year olds and the 8 to 9 year olds used more definite expressions in the LK condition than in the LI condition, the 6 to 7 year olds preferred indefinite reference, irrespective of listener condition and 10 to 11 year olds produced more indefinite expressions to a knowledgeable listener than to an ignorant listener. Analysis of variance showed a significant 3-way interaction between age, indefinite/definite reference and masked/unmasked conditions ($F=20.18$, $df=3,72$, $P,0.001$).

On first mention, only the 3 to 5 year old group used pronouns. In this age group, over 60 percent of definite descriptions involved pronouns in the LK condition and 50 percent of definite references involved pronominalisation in the LI condition. Analysis of variance yielded no significant difference between definite articles and pronouns from the 3 to 5 year olds on first mention ($F<1$), nor was there an interaction between definite article/pronoun and masked/unmasked conditions ($F<1$) in this age group.

B) SUBSIDIARY RESULTS

Second mention

All subjects used definite descriptions on second mention of the main referent. The only age differences were due to the type of definite expressions used: definite

articles or pronouns.

Table 1.B shows the mean scores out of 4 for using pronouns and definite articles on second mention, in each age group and for both LI and LK conditions.

AGE (out of 4)	LI		LK		AGE MEAN	
	DA	PRONOUN	DA	PRONOUN	DA	PRONOUN
3-5	2.2	1.8	1.6	2.4	1.9	2.1
6-7	3.2	0.8	3.2	0.8	3.2	0.8
8-9	4.0	0	3.8	0.2	3.9	0.1
10-11	4.0	0	4.0	0	4.0	0.0
MEAN	3.4	0.6	3.2	0.8	3.2	0.8

TABLE 1-B: MEAN SCORES FOR DEFINITE ARTICLES AND PRONOUNS ON SECOND MENTION.

Clearly, there was a low proportion of pronominal expressions compared with definite articles overall. A significant pronoun verses definite noun phrase difference was confirmed by analysis of variance ($F=258.62$, $df=1,72$, $p<0.001$). There also is little difference between LK and LI conditions and there was no significant interaction between listener condition and definite article/pronoun reference ($F=1.66$, $df=1,72$, NS). Moreover, pronominal reference was confined almost exclusively to the youngest groups, with there being a clear trend of decreasing pronominalisation with age. A significant interaction

between definite article/pronoun reference and age group was found ($F=38.76$, $df=3,72$, $p<0.001$) and Newman-Keuls revealed that this effect was, in fact, due to the difference between the 3 to 5 year olds against all other groups, as well as due to a difference between the 6 to 7 year olds against older groups. However, there was no significant 3-way interaction between age group, listener condition and definite NP/pronoun reference ($F=0.83$, $df=3,72$, NS).

Comparing reference used on first and second mention, a significant difference between first and second mention in the use of either indefinite or definite reference was significant using analysis of variance ($F=706.69$, $df=1,72$, $p<0.001$). Moreover, there was a significant interaction between age group and first/second mention choice of reference ($F=21.15$, $df=3,72$, $p<0.001$), due to the greater number of indefinite references overall for 6 to 7 year olds on first mention. There was also an interaction between listener condition and first/second choice of reference ($F=8.22$, $df=1,72$, $p<0.005$), due to the fact that there were less differences between first and second mention for LK conditions where there were more definites on first mention. Finally, there was a 3-way interaction between age, listener condition and first/second mention choice of reference ($F=19.95$, $df=3,72$, $p<0.001$). These effects are due to the fact that overall, the 6 to 7 year olds used predominantly indefinite expressions on first mention and thus, their first/second mention differences

are much greater for both LI and LK conditions than for other age groups. However, for the 10 to 11 year olds, there is only a strong difference between first and second mention for the LK condition (as indefinite references were preferred on first mention in the LK condition for this age group), whereas 3 to 5 year olds and 8 to 9 year olds only showed a first/second mention difference in the LI condition (as indefinite references were preferred on first mention in the LI condition for these age groups).

Comparing reference to main and subsidiary characters

Table 1.C presents mean indefinite expressions used on first mention out of 4, for references to both main and subsidiary characters, in each listener condition and for each age group.

(total out of 4) AGE	MAIN CHARACTER		SUBSIDIARY CHARACTER	
	LI	LK	LI	LK
3-5	2.4	0.7	2.3	1.0
6-7	3.3	4.0	3.4	3.4
8-9	4.0	1.6	3.5	1.8
10-11	1.6	2.7	2.8	2.5
MEAN	2.8	2.3	3.0	1.7

TABLE 1-C: MEAN INDEFINITE EXPRESSIONS FOR MAIN AND SUBSIDIARY REFERENTS ON FIRST MENTION.

The table shows little difference between indefinite scores on first mention for main characters and first mention indefinites for subsidiary characters, except for the 10 to 11 year old group in the LI condition, where there is a higher proportion of indefinites for subsidiary characters compared to main characters. Indeed analysis of variance showed no overall significant difference between main and subsidiary referents in the use of indefinite expressions ($F=0.43$, $df=1,72$, NS). However, there was an interaction between age group and main/subsidiary referents in indefinite usage ($F=4.79$, $df=3,72$, $p<0.005$) and a 3-way interaction between age, listener condition (LI/LK) and main/subsidiary referents in the use of indefinites ($F=10.18$, $df=3,72$, $p<0.001$).

Pronominalisation on first mention is confined all except for two cases in the 6 to 7 year old group, to the youngest group. In total, 3 to 5 year olds used 28 pronouns for first mention of main characters and 20 pronouns for first mention of subsidiary characters. Thus, there was little difference between main and subsidiary characters in the use of pronouns on first mention.

Comparison of results with previous experiments

Emslie's experiments 2 and 3 also manipulated whether the listener was ignorant or knowledgeable. She used children between 4 and 7 years and the results of her experiments are presented in table 1.D, together with the results for reference to the main character in this

experiment. Only first mention scores are included:

TABLE KEY:

- E1: Emslie's Experiment 1
 E2: Emslie's Experiment 2
 S: Results of current experiment.

AGE (percentages)	EXPERIMENT	KNOWLEDGEABLE		IGNORANT	
		INDEF.	DEF.	INDEF.	DEF.
3-5	S	17.5	72.5	60.0	40.0
4/5	E1	26.8	59.2	67.1	19.7
4/5	E2	39.4	50.5	58.8	30.7
6-7	S	100.0	0.0	82.5	17.5
6	E1	30.4	54.4	82.5	0.0
6/7	E2	30.1	50.7	54.2	28.5

TABLE 1-D: PERCENTAGE INDEFINITE AND DEFINITE SCORES IN EMSLIE'S EXPERIMENTS AND IN THIS EXPERIMENT (3 TO 7 YEARS).

The only clear difference between the scores in this experiment and the scores in Emslie's experiments are for the 6-7 year olds in the 'knowledgeable' listener condition: Whereas, Emslie found that her subjects used significantly more definites compared with indefinites in this condition, there is 100 per cent choice of indefinite

expressions to a 'knowledgeable' listener from the 6 to 7 year olds of this experiment.

DISCUSSION

It is clear that, except for 6 to 7 year olds, from 3½ years of age, children were influenced by whether the listener had seen the pictures or not when choosing between indefinite and definite expressions. The youngest children used mainly indefinite articles when the listener was 'ignorant' and used mainly definite articles when the listener was 'knowledgeable'.

However, the 6 to 7 year olds preferred the indefinite article irrespective of the knowledge of the listener: Listeners were treated as ignorant even when they had seen the pictures. It seems possible that 6 to 7 year olds were more influenced by their experience of listeners in a conventional story situation and consequently always treated them as if they were ignorant.

The finding that younger subjects used as many pronouns to first mention subsidiary referents as they did for the main referent suggests that young children will produce ambiguous definite reference because, when there is more than one potential referent, pronouns do not serve to refer to a specific referent unless that particular referent is the main character (thematic structure). Therefore, whilst definite expressions were preferred for second mentions of referents at all ages, one must be cautious about making any claims regarding the child's knowledge of anaphoric reference. This is because such

knowledge requires an awareness of the significance of 'specificity' in addition to an appreciation of listener 'familiarity'.

3.4 EXPERIMENT 2: The effects of using the experimenter as the listener and comparing story-telling with discussions about real-life experiences.

INTRODUCTION

An investigation of whether young children use their experience of everyday situations when considering the knowledge of the listener for indefinite or definite reference was carried out in the following experiment, where the experimenter was the listener under three experimental conditions. Using the story-telling task of experiment 1, young children were required to narrate a story to an experimenter who pretended not to have seen the pictures. For the other two conditions, an elicited-discussion procedure was used in order to get subjects to talk about real events that they had experienced outside the experimental situation. Thus, the actual knowledge of the experimenter was manipulated to see whether young children could treat an experimenter as ignorant when she genuinely was ignorant of the events being discussed.

It was hypothesised that, if young children do use pragmatic procedures, then they will be reluctant to treat the experimenter as ignorant for the story-telling task, but they will use indefinite expressions to first mention referents to an experimenter listener when she is genuinely ignorant of the referents being discussed: when those referents are part of the child's real experiences.

METHOD

Subjects

12 subjects (6 males and 6 females) took part in the experiment. All the children were from a playcentre of a primary school in Durham. The mean age of the children was 3.6 years (range: 3 - 4.4 years). A second experimenter (an undergraduate from Sheffield, who had never met the children before the experiment), volunteered to help out with the story-telling task.

Materials

Two of the sets of picture-stories (3 pictures per set) were selected from the stories used in experiment 1. A wide range of toys in the playroom were also required for the experiment. Finally, a tape recorder was used for recording the subjects' verbal responses.

Design and Procedure

Each subject was seen in a small room adjoining the playroom. This room was familiar to subjects, as it was where they hung up their coats and also, where they would often go to hear stories. The experiment manipulated two variables. These were:

- A. LISTENER FAMILIARITY (knowledgeable or ignorant)
- B. TYPE OF TASK (story-telling or discussing real experiences)

Although there are two conditions for each variable there was no story-telling/listener knowledgeable manipulation. Thus, there were three conditions altogether, with the

experimenter serving as the listener in each condition.

These three conditions were as follows:

1. STORY-TELLING (Ignorant Listener)
2. PLAYROOM: discussion of games/toys in playroom (Knowledgeable Listener)
3. BIRTHDAY: discussion of subject's last birthday (Ignorant Listener)

Each subject received all three experimental conditions. All subjects began with a pre-experimental phase of the 'playroom' condition, where the experimenter spent 40 minutes with the subject in the playroom (see below), but the order of presentation of the three conditions (the story-telling task, the 'birthday' condition and the experimental phase of the 'playroom' condition) was completely counterbalanced. The design of the experiment is shown in table 2.

ORDER OF PRESENTATION

SUBJECT	PRE-EXPTL SESSION		EXPERIMENTAL SESSION		
	1st	2nd	3rd	4th	
1 , 7	P1	P2	B	S	
2 , 8	P1	B	S	P2	
3 , 9	P1	S	P2	B	
4 , 10	P1	P2	B	S	
5 , 11	P1	B	S	P2	
6 , 12	P1	S	P2	B	

TABLE 2: COUNTERBALANCED ORDER OF PRESENTATION OF CONDITIONS AND PRE-EXPERIMENTAL SESSION.

TABLE KEY:

P1= PLAYROOM (phase 1)

B= BIRTHDAY

P2= PLAYROOM (phase 2)

S= STORY-TELLING

The three experimental conditions were as follows:

A. STORY-TELLING TASK (LISTENER IGNORANT)

The procedure was similar to that used in experiment one, except for the following differences:

1. The experimenter was the listener and she pretended to be ignorant.
2. Pictures were shown one at a time on a table which was in front of the subject, rather than being presented on a projector screen. This method was chosen because, owing to limited space, it was impractical to set up projector equipment in the experimental room.

The experimenter gave the following instructions: "You are going to look at three pictures. Picture one is the first part of a story, picture two is the middle of the story and picture three is the end of the story. I want you to look at the pictures and afterwards, I want you to look at the pictures again and tell me the story. I am going to cover over my eyes so that I won't know what is in the pictures, because I won't be able to see them."

The experimenter presented the pictures to the subjects face down. Each picture had a number on the back of it,

written in large letters. The pictures were numbered 1, 2 and 3 depending on whether they were the first, second or third picture of the story. The pictures were placed in front of the subject in consecutive order, so that s/he could only see the numbers of the pictures. The initial instruction given to subjects was "Do not say anything about the pictures yet. Just look at them carefully."

A second experimenter sat a few feet from the table to keep check on the actions of the child whilst the first experimenter was giving the instructions. If the child was seen by the second experimenter to be responding incorrectly, she waited until the child had finished telling the story and then reported back to the first experimenter, who explained the instructions again and re-tested the child using a different picture story.

After the subject had looked at the pictures individually for the first time and had turned each picture over again, so that only the numbers could be seen by the subject, the listening experimenter said, "Now, I want you to tell me the story that is in the pictures. Start with the first picture. Turn it over. Now tell me what happens in picture one." After the child had made reference to the first picture, the experimenter continued: "Now turn over picture two. What happens next in picture two." Finally, the first experimenter finished with, "Turn over picture three and tell me what happens at the end of the story."

Each subject was scored for only one set of

pictures. The experimenter's initial intention had been to counterbalance the two story sets across subjects, but some subjects were given the other story when they failed to respond appropriately to the initial instructions.

B. 'PLAYROOM' ELICITED-DISCUSSION Topic: Toys in the playroom (LISTENER KNOWLEDGEABLE)

Phase 1: Pre-experimental play session

The experimenter spent 40 minutes with each subject playing with any of the toys/games that were available in the playroom. The experimenter simply participated in any game or event that the child wished and played rather a passive role in the sense that she did not ask questions or encourage the child to do anything. Rather she let the child act spontaneously on the objects around her/him and joined in when the child began to include her in a game.

Phase 2: Experimental discussion session

In the experimental room, the experimenter asked questions to get subjects to refer to the games/toys that they had previously played with in the playroom. Care was taken to prevent subjects from using definite expressions merely as a result of leading questions: The experimenter did not mention the referents in the questions, otherwise the subject could have assumed that the experimenter was knowledgeable because she had just mentioned the referents, rather than because she had previously participated in the playroom activities with the child.

Two main questions were put to each subject, with the addition of some priming questions that were asked only if required. 'Question one' was asked to get subjects to mention at least one of the toys/games for the first time. 'Question two' was asked to encourage subjects to mention for a second time, one of the same toys/games that had just been mentioned. The questions were also asked in a way that would elicit complete sentences from subjects, rather than simply one word or single phrase answers. The questions were as follows:

QUESTION ONE: What did we do in the playroom today? Can you remember?

Primers: What else did we do? What were we playing with earlier on. Tell me about it?

QUESTION TWO: What else did we do today?

Primers: What did we do next? What was the funniest thing we did today? Can you tell me more about what we were playing with?

If a child failed to mention a referent after 10 minutes (for either question one or question two), a 'non-reference' response was recorded on the score sheet.

C. 'BIRTHDAY' ELICITED-DISCUSSION Topic: What subject did on his/her last birthday (LISTENER IGNORANT)

The subject was encouraged to talk freely about events/objects that were associated with their previous birthday.

Again, 2 main questions were given to each subject with the intention of getting subjects to refer to a referent twice. Additional primers were asked if necessary. In this condition, there is clearly more flexibility in the child's response and so, for 'question two', the type of question asked was adapted to suit the response of the subject to 'question one.' The questions were as follows:

QUESTION ONE: What happens on birthdays? What did you do on your last birthday?

Often subjects used possessive pronouns or proper nouns in response to this question, in which case the experimenter persisted with questioning until an article was used. However, if an article was not used within 10 minutes of questioning, the most recent expression was recorded.

Primers: What else did you do? What happened on your birthday. Can you remember?

QUESTION TWO: What else did you do on your birthday?

Primers: What did you do with the presents? Is there any present which you like best?

If within 10 minutes of questioning for each question (one and two), reference to an item had not been made by the subject, the child was given a 'non-reference' score.

Scoring

Owing to the more open-ended nature of the expressions elicited from subjects in the discussion conditions, expressions used on both first mention and second mention were classified into the following categories:-

Indefinite article

Definite article

Pronoun/demonstrative

Null article

Possessive pronoun

Proper noun (name)

Non-reference




RESULTS

MAIN RESULTS: FIRST MENTION

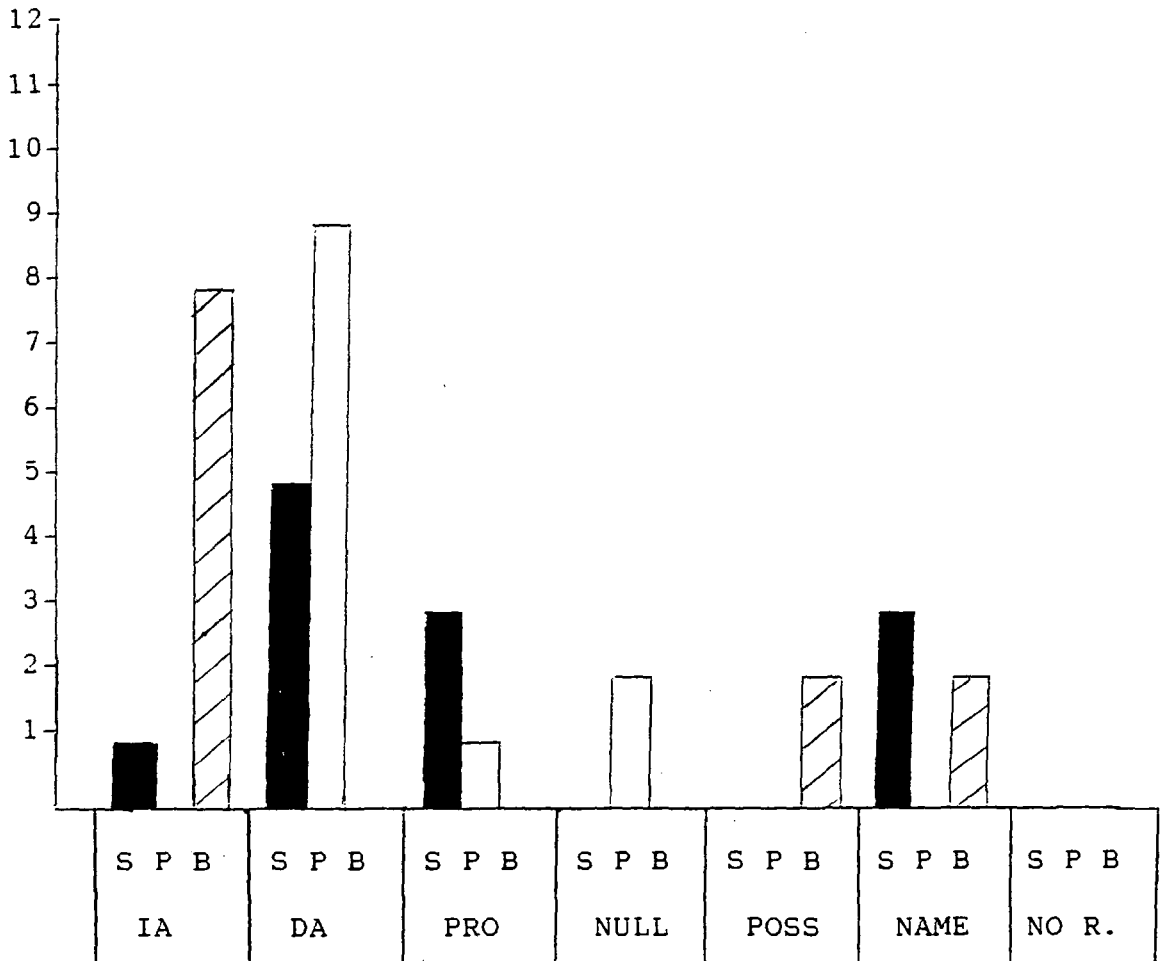
The total number of subjects using each category of expression on first mention of a referent in each of the three conditions is shown in figure 2-1.

It is clear from this figure that the pattern of results is different across the three conditions for first mentions, with both the story-telling and playroom conditions producing predominantly definite expressions and with the birthday condition producing predominantly indefinite articles. Excluding cases where possessive

FIGURE 2.1 TOTAL REFERENCE SCORES ON FIRST MENTION IN EACH CONDITION.

Key: S = Story-telling 
 P = Playroom 
 B = Birthday 

IA = Indefinite
 DA = Definite article
 PRO = Pronoun
 NULL = Null article
 POSS = Possessive pronoun
 NAME = Proper noun
 NO R. = No reference



pronouns were used, Cochran tests showed that, for first mentions, there was a significant difference between the 3 conditions in the use of indefinite articles ($Q=14.25$, $df=2$, $p<0.001$) and in the use of definite descriptions: definite articles, pronouns and proper nouns ($Q=9.56$ $df=2$, $p<0.01$).

a. Story-telling

The scores in the story-telling condition of this experiment were predominantly definite on first mention. Table 2-A shows percentage indefinite and definite scores for the story-telling condition as well as both the listener conditions of experiment one. (It is important to observe that the children in this experiment were on average a year younger than those in the youngest group for experiment 1).

(% scores)	Experiment 2	Experiment 1	
	Story-telling	'Ignorant'	'Knowledgeable'
Indefinites	8	60	17.5
Definites	92	40	82.5

TABLE 2.A PERCENTAGE INDEFINITE AND DEFINITE SCORES FOR STORY-TELLING BOTH IN THIS EXPERIMENT AND IN EXPERIMENT ONE.

Comparing the scores between this condition and those produced by infants in the ignorant listener condition of

experiment one, where young children preferred indefinite expressions, there is a clear difference in the scores between the two experiments. The scores in this condition are more like the listener knowledgeable condition of experiment one, where definites were the most popular form of reference.

b. Discussing Real experiences




As the results are clear from figure 2-1, no statistical analysis was required to show that there was a significant difference between the playroom and birthday conditions on first mention in the choice of indefinite/definite expressions: Clearly, definite expressions were preferred in the playroom condition (83%) and indefinite expressions were preferred in the birthday condition (67%).

SUBSIDIARY RESULTS: SECOND MENTION

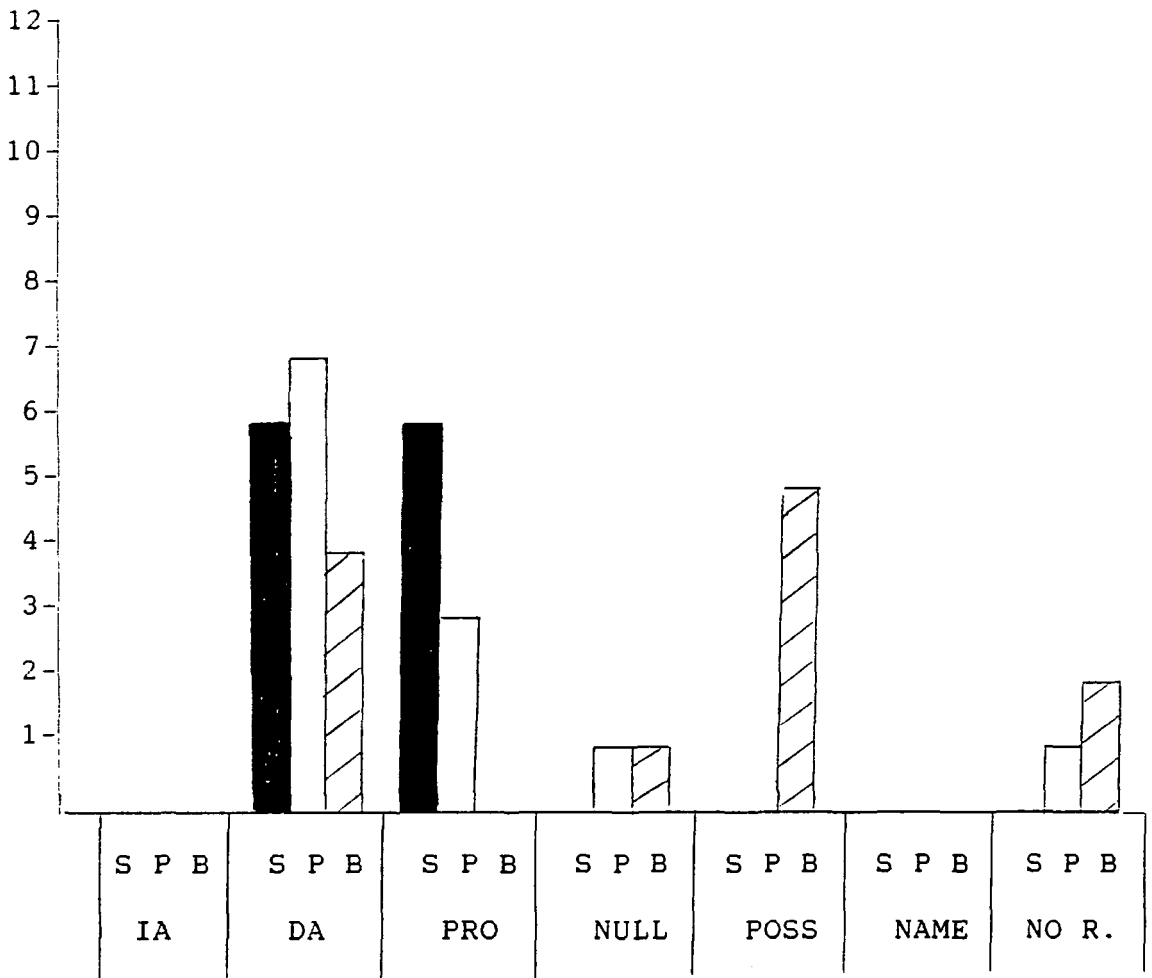
The total number of subjects using each category of expression on second mention is shown in figure 2-2. It is clear that expressions used did vary across conditions: mainly definite articles and pronouns for story-telling and playroom conditions, but possessive pronouns and definite articles for the birthday condition. However, a Cochran test did not reveal a significant difference between the conditions in the use of definite articles verses other definite descriptions, (including possessive pronouns) ($Q=2.0$, $df=2$, NS).

Finally, no subject used an indefinite expression in either of the three conditions on second mention.

FIGURE 2.2 TOTAL REFERENCE SCORES ON SECOND MENTION IN EACH CONDITION.

Key: S = Story-telling 
 P = Playroom 
 B = Birthday 

IA = Indefinite POSS = Possessive pronoun
 DA = Definite article NAME = Proper noun
 PRO = Pronoun NO R. = No reference
 NULL = Null article



DISCUSSION

The results from the story-telling task support the prediction that young children would be reluctant to treat an experimenter as an ignorant listener. Definite expressions were preferred on first mention in this condition, despite the fact that the experimenter covered her eyes.

However, the results also show that, when young children were encouraged to talk about real experiences for which they knew that the listening experimenter was genuinely ignorant, they did treat the experimenter as ignorant: They preferred to use indefinite expressions to introduce referents for the first time when talking to an experimenter who had not shared the experience with them, whereas definite descriptions were the predominant response when the experimenter had shared the experience/s.

A few issues need to be discussed further: In the birthday condition, the events occurred further back in time to the events in the playroom and picture-stories conditions. Therefore, it is possible that the child's sensitivity to the listener for real-events may be confounded by temporal factors: It may be that the child is more likely to use indefinite expressions for events they have more trouble remembering. Moreover, in the birthday condition, the experimenter listener and the events being discussed would never have been associated together before by the child. By contrast, in the picture-story situation and for the playroom condition, the listener and the referents were experienced together

in the same context (both in the playroom). Therefore, there is a chance that the listener's proximity to the referents is a factor which influences young children's ability to treat an experimenter as an ignorant listener.

Moreover, the change in procedure for the story-telling task in this experiment compared with the procedure used in experiment one, may have confounded results slightly: The fact that the pictures were in front of subjects in this experiment may have increased the use of deictic gestures/expressions and this may have exaggerated the strong preference for definite expressions in the story-telling condition of this experiment.

Bearing these issues in mind, the results here seem to suggest that young children do in fact use 'pragmatic procedures' in experiments looking at children's understanding of indefinite and definite reference (Karmiloff-Smith 1979). Therefore, it appears that it is the child's most expected view of the listeners' knowledge of the actual objects and events, which determines their choice of expression.

3.5 EXPERIMENT 3: The effects of story context on children's sensitivity to the listener's perception of referents.

INTRODUCTION

The previous experiments have shown firstly, that young children appreciate the social determinants of indefinite and definite reference and secondly, that young children's perception of the social requirements of the task is strongly influenced by their pragmatic knowledge.

The picture story-telling task is used again in the following experiment. However, in the second picture of each story, an object is depicted which is predictable given the context of picture one. I hypothesised that, as the young child appears to be strongly influenced by general knowledge in tasks looking at in/definite reference, their general knowledge of contexts is likely to have the greatest effect on their choice of expression, irrespective of the listener's direct perception of referents. Moreover, sensitivity to the listener manipulation will only be revealed when the referents cannot be found by pragmatic inferences. The same subjects were used as for experiment one, to see if there are any developmental effects of the influence of context on in/definiteness.

METHOD

Subjects

The subjects were the same 80 children who had taken

part in experiment one. Therefore, there were 20 subjects in each of the following age groups: 3 to 5 years (mean age 4.7), 6 to 7 years (6.6), 8 to 9 years (9.9) and 10 to 11 years (10.8) and each subject was already paired with a child of approximately the same age.

Materials

The materials that were used (projector apparatus, screen, tape recorder, mask) and the experimental layout was the same as for experiment one. The only difference was that 8 new story sets with 2 pictures making up each set, were constructed to replace the stories used in the first experiment. These picture stories were designed so that one referent introduced in picture two would be predictable given the story context from picture one. There were two stories for each of four different types of constructions. These different types have been labelled as LOCATIVE, ENTAILED, UNIVERSAL and SETTING. A description of each type, together with the two examples used in the experiment is presented below:

1. LOCATIVE:

E. Clark (1978, p88) argued that locative phrases are often definite. Emslie (1986, experiment 5) found evidence of children preferring definites for objects which served the function of a location. The locative referents in this experiment were SHELF and TABLE.

eg.

1a. A boy is carrying a book under his arm

He put the book on the shelf.

- 1b. A man is carrying a heavy box
He dropped the box onto the table.

2. ENTAILED:

Brown (1973) used the term, 'specified by entailment' for examples such as 'car - 'the' engine', where engine is a unique associate of a car and therefore, the introduction of 'car' presupposes the existence of 'engine'. Zehler and Brewer (1982) found that for objects which were context-unique (car - 'the' steering wheel) or context-intermediates, where a few items are available in a particular context (car - 'the' door), 2 and 3 year olds and adults used the definite article on first mention of the referent that was made available from context. The entailed referents in this experiment were TIN/LID (context unique) and CAR/DOOR (context intermediates).

eg.

- 2a. A lady is taking a tin out of the cupboard
She took off the lid and looked inside.

- 2b. A woman is leaning against a car
She opened the door with her key.

3. UNIVERSAL:

Two of Brown's (1973) definite uses were 'unique of all' (eg. the moon) and 'uniquely salient for a given social group' (eg. the constitution). Hawkins (1978) used the

term larger situation use of definite descriptions, where prior introduction of a referent can be assumed because of "general knowledge of the existence of certain types of objects in certain types of situations...." Thus, a definite noun phrase is possible where, within a certain society, culture or sub-culture, all members would expect only one of its kind to exist. SUN and QUEEN were the examples used in this experiment.

eg.

3a. A man is walking along with an umbrella
He became very hot when the sun came out.

3b. A boy is standing in a crowd.
He waved his arms as the queen went by.

4. SETTING:

The following corresponds to Brown's (1973) 'unique in a given setting' example (eg. the desk, the ceiling). Cruse (1980) argued that the 'associative anaphoric use' (Hawkins 1978) was the most common use of the definite article. When a noun phrase or verb phrase triggers the representation of a prototypical concept, for example 'shop', the scene is also depicting other objects which would be expected from the individual's knowledge of shops, such as a door, a counter, a shop assistant etc. The presupposed objects do not need to be introduced explicitly because the listener can infer that they exist from the setting. In this experiment PLAYGROUND/SANDPIT and CAFE/WAITER were the two examples used.

eg.

4a. A girl is standing in a playground/park

She started to play in the sandpit.

4b. A woman walked into a cafe

The waiter brought her a cup of tea.

(Picture stories are shown in figures 3.1A- 3.4B)

Design and Procedure

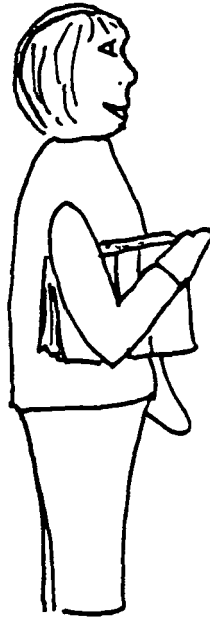
All 8 stories were presented for each pair of subjects, with each member of a pair acting as speaker for all four story types mentioned above. For each pair, one partner was the speaker for the SHELF, LID, SUN, SANDPIT examples and the other partner was the speaker for the TABLE, DOOR, QUEEN, WAITER examples. Cards numbered 1 to 4 were put into a hat and the order in which they were selected (by a member of teaching staff), determined the order in which the 4 story types were presented for each pair of subjects.

As for experiment one, the variable manipulated in this experiment was whether listeners had seen the pictures or not (Knowledgeable/ unmasked (LK) or Ignorant/masked (LI)). The subject pairs had already been assigned to these conditions in experiment one.

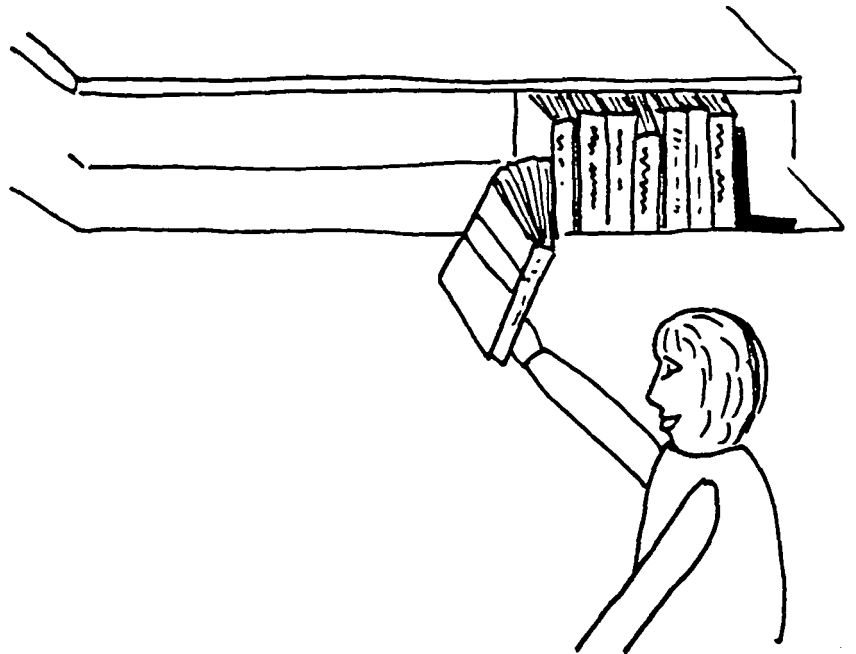
The experimental room was the same as for experiment one. The procedure was also the same as for experiment one, with each member of a pair alternating between the roles of speaker and listener. (See experiment one for full procedure and instructions).

FIGURE 3.1A PICTURE STORY 1A

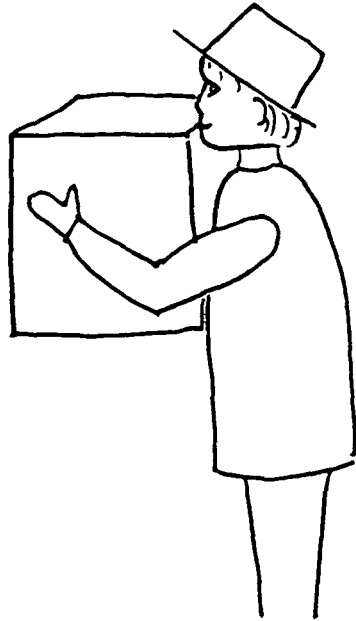
First
picture



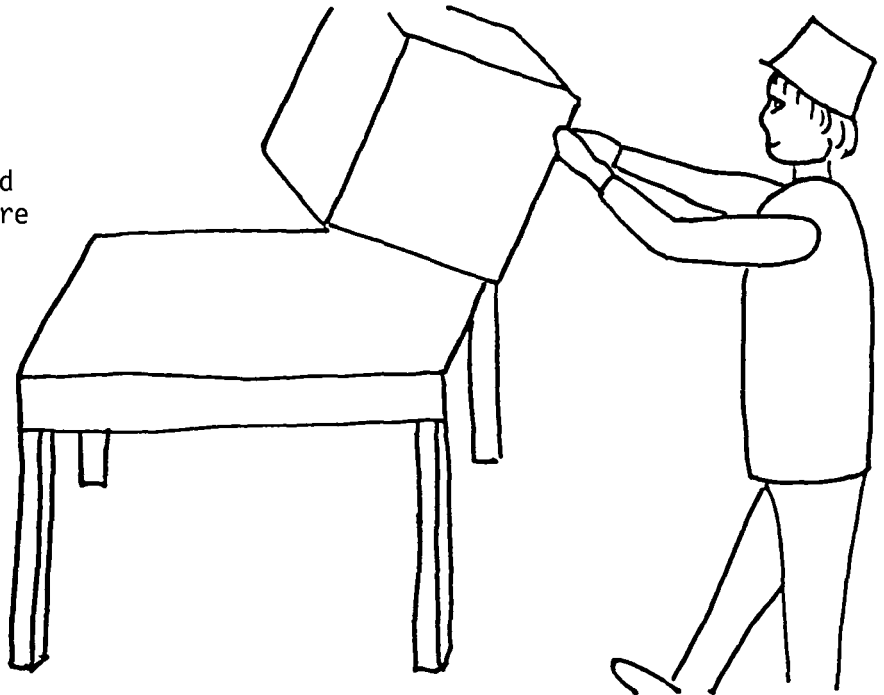
Second
picture



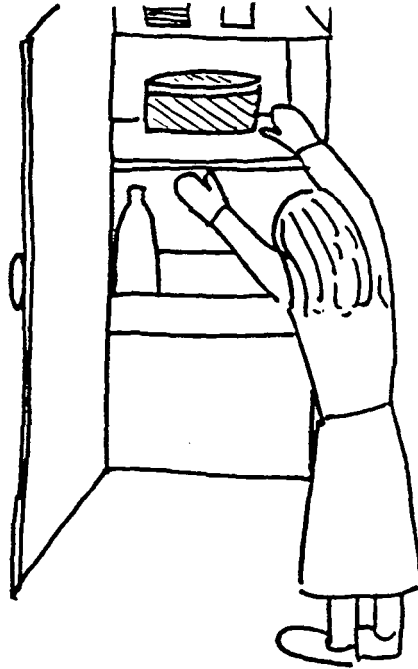
First picture



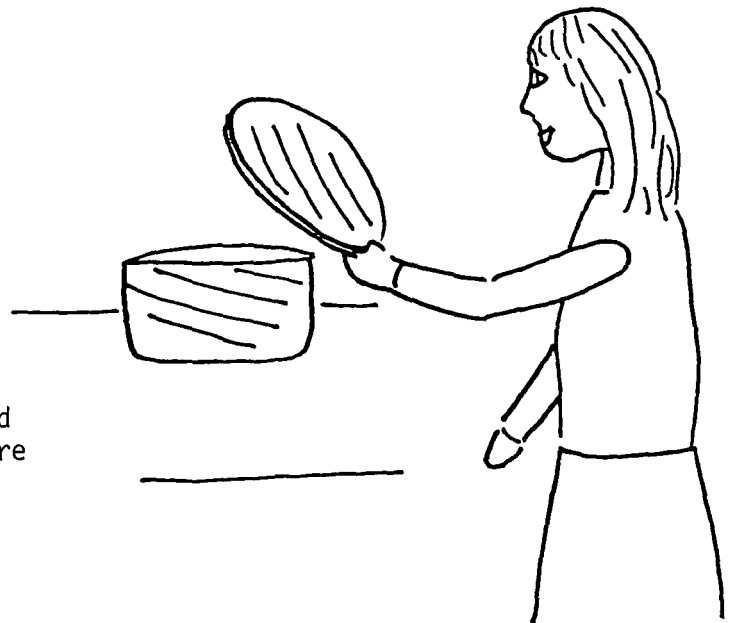
Second picture



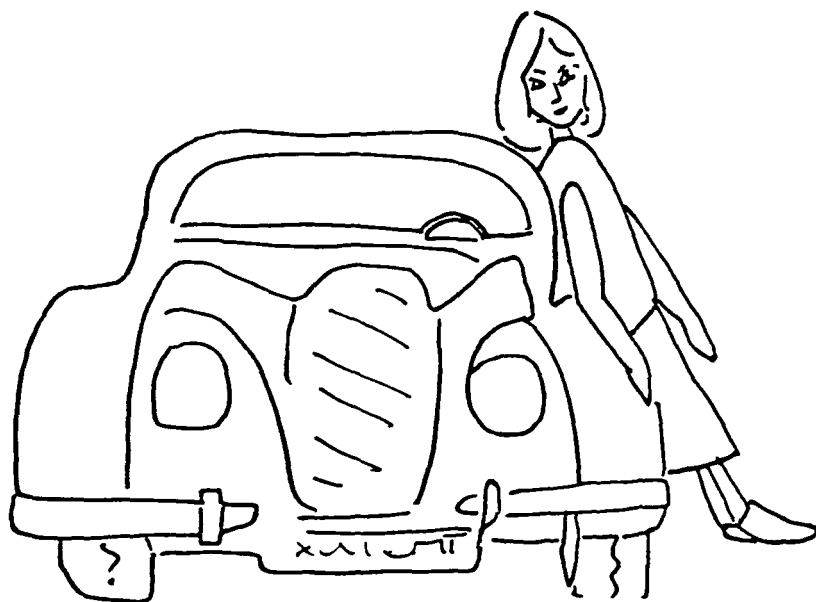
First
picture



Second
picture



First picture



Second picture

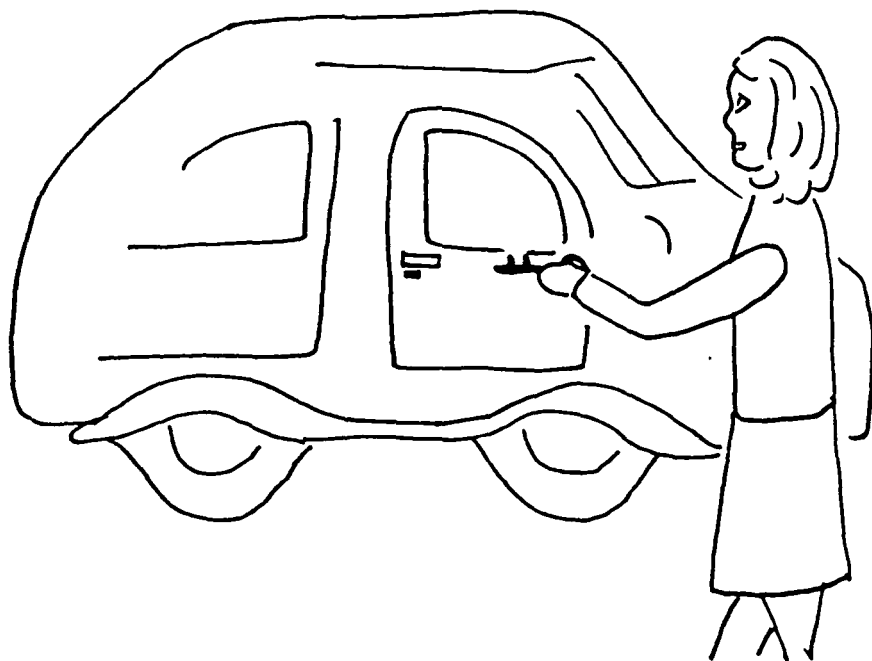
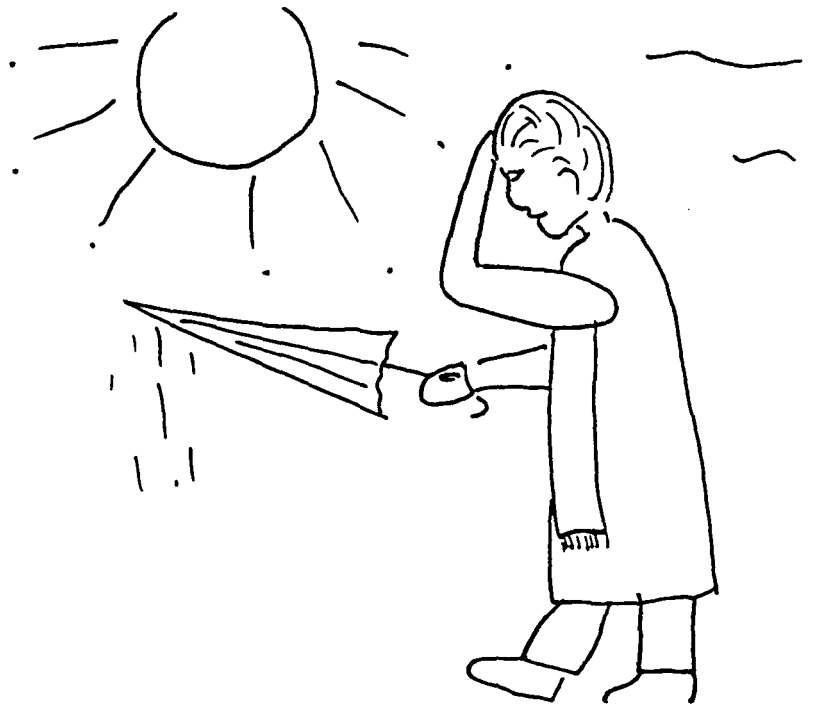


FIGURE 3.3A PICTURE STORY 3A

First
picture



Second
picture



First picture

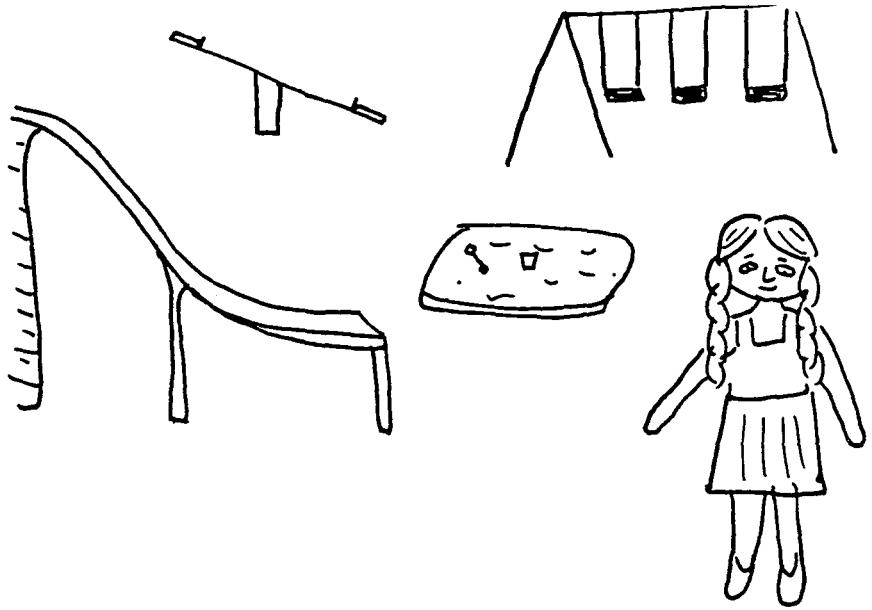


Second picture



FIGURE 3.4A PICTURE STORY 4A

First picture

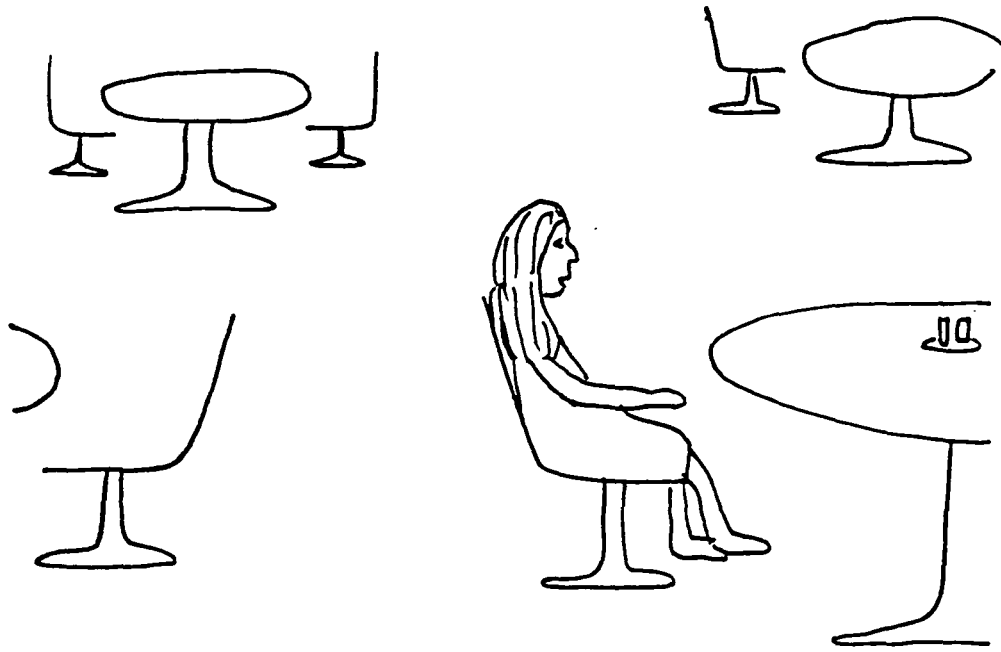


Second picture

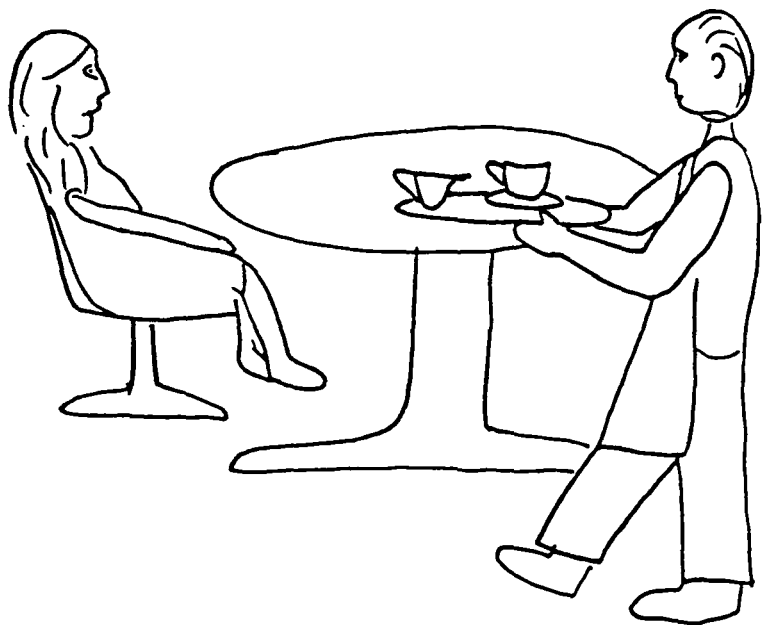


FIGURE 3.4B PICTURE STORY 4B

First
picture



Second
picture



Scoring

Interest was only in the expressions used for first mention of referents. The scores given for subject's expressions were out of the following:

1. Referents in picture 1 - INDEFINITE: a boy...a book
DEFINITE: eg. the boy/book,
he/it, his/its, Peter/Jane

2. Contextually predicted referent in picture 2 -

Scores were classified into the following categories:

CONTEXT DEFINITE: Definite noun phrase (the waiter, the sun)

pronouns, demonstratives (it, that)

INDEFINITE: A waiter, a sun etc.

RESULTS

A) MAIN RESULTS

Context referent in picture 2

A. Between Listener conditions

Although each subject received four picture-stories, some subjects received a score out of 3 rather than 4, for expressions used for picture 2. The reason for this was because, sometimes a subject failed to mention the waiter at all in the CAFE/WAITER story. For example, the subject would simply say, 'She had ordered a cup of tea' or 'She was brought a cup of tea' or even 'She drank some tea'. However, if a subject failed to mention 'waiter', but did

say something like 'the man brought her....', then a score was given, because this subject is clearly referring to 'the man who serves people in the cafe.' The fact that some subjects received a score out of 3 means that percentage scores were used for analysis.

The total (percentage) context definite scores for each age group in both the listener conditions is shown in table 3-A.

(Percentage scores)	LISTENER CONDITION.		
	IGNORANT	KNOWLEDGEABLE	MEANS
AGE			
3-5	89.5	97.4	93.5
6-7	83.8	85.0	84.4
8-9	70.3	84.2	77.3
10-11	83.8	63.9	73.9
MEANS	81.9	82.6	

TABLE 3-A: PERCENTAGE DEFINITES BY LISTENER CONDITION FOR EACH AGE GROUP TO REFERENT IN PICTURE 2.

It is clear that, in both LI and LK conditions, for all age groups, the highest percentage of scores were CONTEXT DEFINITES for first mention of the context referent. Moreover, the highest percentage of definite scores were from the youngest group and there appears to be a developmental trend towards fewer context definites with age, with this trend showing only in the LK condition

where scores decreased as a function of increasing age. However, in the LI condition the lowest scores were from the 8 to 9 year olds. Thus, the fewest 'context definite' scores for the 8 to 9 year olds were in the LI condition, whereas they were in the LK condition for 10 to 11 year olds. Analysis of variance showed a significant overall difference between context definites and other types of expressions ($F=188.31$, $df=1,72$, $p<0.001$). There was also a significant effect of age on context definite scores ($F=4.02$, $df=3,72$, $p<0.05$) and the Newman-Keuls analysis revealed that this effect was due to the difference between the scores for the 3 to 5 year olds against the scores in the two oldest groups. However, there was no significant effect of listener condition on context definite scores ($F=0.03$, $df=1,72$, NS). Nor was there an interaction between age group and listener condition ($F=2.64$, $df=3,72$, NS).

For mentioning the referent in picture 2 for the first time, those remaining scores that were not context definites were nearly all indefinite, apart from 4 scores in the youngest group which were due to reference using pronouns.

B. Between Stories

Combining both LI and LK conditions, table 3-B presents the percentage of context definite scores for each story type (locative, entailed, universal, setting) in each age group.

AGE (& scores)	LOCATIVE	ENTAILED	UNIVERSAL	SETTING
3-5	95.0	90.0	100.0	82.4
6-7	80.0	85.0	100.0	70.6
8-9	70.0	80.0	100.0	53.3
10-11	60.0	80.0	100.0	46.2
MEANS	66.3	83.8	100.0	63.1

TABLE 3-B: PERCENTAGE DEFINITES TO REFERENT IN PICTURE 2 ACROSS THE 4 STORY TYPES.

From this table, it can be seen that context definite scores were maximum for the universal story types. Scores were again very high for the entailed stories. However, for locative and setting story types, scores were lower as a function of a decrease in context definites with age.

Excluding those subjects who did not refer to the waiter for the setting story, Cochran tests showed that, whilst there was no difference in scores across those 4 types for the youngest group ($Q=6.23$, $df=3$, NS), there was a difference in the proportion of scores which were context definite across the 4 story types for all older age groups (6-7: $Q=8.05$, $df=3$, $p<0.05$, 8-9: $Q=13.96$, $df=3$, $p<0.005$, 10-11: $Q=11.14$, $df=3$, $p<0.05$).

Table 3-C presents the percentage scores which were

context definite for each individual story in each age group, keeping examples a. and b. separate for each story type.

(proportional scores)

AGE	LOCATIVE		ENTAILED		UNIVERSAL		SETTING	
	A)shelf	B)table	A)lid	B)door	A)sun	B)queen	A)sand	B)waiter
3-5	100	90	80	100	100	100	70	100
6-7	90	70	80	90	100	100	60	86
8-9	80	60	80	80	100	100	50	60
10-11	80	40	80	80	100	100	50	33
MEANS	88	65	80	88	100	100	58	59

TABLE 3-C: PERCENTAGE DEFINITES FOR REFERENT IN PICTURE 2 ACROSS ALL STORIES.

Table 3-C shows that the age decrease in context definite scores for the locative story type is clearest in the 'table' example. For the setting stories, although scores are altogether lower for the 'sandpit' example, the developmental drop in context definites is clearest for the 'waiter' example.

B) SUBSIDIARY RESULTS

Referents in picture 1

As each subject narrated a total of 4 stories and as there were 2 referents in picture 1 for each story, each subject

received scores out of 8 for the type of expression (indefinite or definite) used to refer to referents in picture 1. Table 3-D presents total indefinite and definite scores in each listener condition for all age groups.

(out of 80) AGE	<u>LISTENER IGNORANT</u>		<u>LISTENER KNOWLEDGEABLE</u>	
	INDEFINITE	DEFINITE	INDEFINITE	DEFINITE
3-5	52	28	10	70
6-7	61	19	54	26
8-9	52	28	30	50
10-11	40	40	54	26
MEANS:	44	29	37	43

TABLE 3-D: INDEFINITE AND DEFINITE SCORES IN EACH CONDITION FOR REFERENCE TO PICTURE ONE.

It can be seen from this table that indefinites were preferred for an ignorant listener and definites were preferred for a knowledgeable listener. However, this pattern only applies for the 3 to 5 year olds and the 8 to 9 year olds. 6 to 7 year olds preferred indefinites regardless of listener condition and 10 to 11 year olds preferred indefinites for knowledgeable listeners and showed no preference for ignorant listeners. These findings were confirmed by analysis of variance which showed an interaction between listener condition and

indefinite/definite expressions ($F=11.29$, $df=1,72$, $p<0.001$), an interaction between age and indefinite/definite expressions ($F=6.85$, $df=3,72$, $p<0.001$) and finally, a 3-way interaction between age, listener condition and indefinite/definite expressions ($F=7.79$, $df=3,72$, $p<0.001$).

DISCUSSION

The findings of this experiment show that, those same subjects who were sensitive to the social situation in picture one of this experiment and for experiment one were predominantly influenced by the context of the story for choosing an expression for the referent in picture two. Thus, the results do support the initial hypothesis that young children will let general knowledge define 'familiarity' regardless of the listeners' immediate perception, because from 3 years of age, children used mainly definite noun phrases on first mention of a referent that was predictable from context, even when listeners had not seen the referent before.

This overriding effect of context persisted in all age groups. However, with increasing age, the strength of context effects were reduced, with this age difference only appearing for some of the stories. The age effect seems to be due to older children being more reluctant to infer the existence of a referent when the particular context is flexible or when the context contains idiosyncratic features. For example, 'the sun' and 'the queen', for which context had the overriding effect at all

ages, are not predictable from the context of the story itself, but from knowledge of the 'world (society)' in general: People know only of one sun and one queen. Therefore, subjects are unlikely to be affected by the particular context of the current discourse, (except maybe that the boy in 'the queen' example is accepted as being English, so it is Queen Elisabeth that this boy is likely to be cheering). The fact that definite scores were also high at all ages for the 'entailed' stories may be because it is only the existence of the objects in picture one rather than the context in which these objects appear, that predict the referent in picture two.

However, the clear developmental decrease in context definites for the 'table' example can be accounted for by the fact that the context gave no clue as to whether a table should be present or not: Picture one gave no indication as to the possible whereabouts of the boy carrying the box and therefore, the appearance of the table in picture two would probably have come as a surprise to subjects. By contrast, for the 'shelf' example, books are normally expected to be located on shelves. It seems that, for older subjects, locatives are more likely to be definite if the location can be inferred from the particular context.

Similarly, with regards to the 'waiter' example, the developmental decrease may be due to the fact that, the context of picture one in this example does not automatically lead one to infer the existence of a waiter, because there are no clues as to the type of cafe being

referred to: whether it is self-service or waiter-service.

Therefore, the results suggest that context effects are so strong in very young children that they may fail to accomodate their general knowledge with novel or atypical features of the particular discourse context.

3.6 GENERAL DISCUSSION

The results showed that even the 3½ to 5 year olds adapted their choice of in/definite expression as a function of their perceived knowledge of the listener.

Experiments one and three showed that, when listener knowledge depended on the immediate situation (masking conditions), young children used indefinites preferentially for an 'ignorant' listener and definites for a 'knowledgeable' listener. However, older children did not follow this pattern and it seems that other factors were able to override the masking conditions for older subjects. One of these extraneous factors appears to have been the effects of story conventions. One convention for story-telling is to narrate certain events to an audience who have no knowledge of those events. This may explain why 6 to 7 year olds tended to prefer indefinite articles on first mention even when the listener was unmasked. In fact, most subjects in this age group chose to use expressions such as 'Once upon a time, there was a....' or 'One day, a....' (known as 'formal openings', Applebee, 1978) and often finished stories with the phrase 'lived happily ever after'. Thus, it seems that 6 to 7 year olds were more advanced in their understanding of conventions for story-telling and therefore, less concerned with the listener's actual knowledge of the referents in the experimental situation. This is also consistent with the findings that knowledge of story content precedes knowledge of story structure (Kontos, Mackley and Baltas 1986). By contrast, Emslie used film videos which

were much longer than the picture stories of this experiment. For one of her films which was particularly long, all subjects chose to use the definite article on first mention of one of the referents ('the tree'). Maybe, having to keep track of the events in Emslie's films prevented subjects from placing a strong emphasis on their role as story-tellers, in which case they may have been more influenced by the status of the listener in the immediate situation. Moreover, the instructions to subjects in Emslie's experiments were to 'tell the listener what is happening', with no mention of the word 'story' at all, suggesting that the results here were possibly due to the emphasis placed on story-telling. Nevertheless, even Emslie's 6 and 7 year olds used more indefinite expressions than did younger subjects, but this effect was probably obscured because Emslie did not look at the performance of children over 7 years of age. The current experiment showed that 8 and 9 year olds responded more like the youngest age group: more indefinites to a masked listener and more definites to an unmasked listener, suggesting that the 6 to 7 year olds are more impressed by story conventions than other age groups.

However, in experiment one when referring to the main character, 10 to 11 year olds used more definites to a masked listener, but when referring to subsidiary characters (experiment 1) or characters of equal significance (experiment 3 and Emslie's experiments 1986), the oldest group did prefer indefinites. Maybe, story

conventions can also account for this result: When there was one main referent, the oldest subjects appear to have been using the definite article to restrict the range of possible discourse entities to the only one relevant to the current context (cf. Johnson-Laird and Garnham, 1980, p.377).

Another extraneous factor may have been that a wide age range of subjects were used who were expected to respond in a similar way to the same sets of pictures. In fact, often the stories of the oldest group were quite elaborate in content, going beyond the information explicitly conveyed in the pictures. Looking at the examples below:

Example 1: (7 year old - LK)

- P1. One day a man went fishing.
- P2. A boy pointed to the fishing rod.
- P3. The man pulled the fishing rod out of the water and there was a fish on the end of the fishing rod.

Example 2: (11 year old - LK)

- P1. One fine day, a man decided to go to the park to do a spot of fishing.
- P2. Later on a boy who happened to be walking past, noticed that something was pulling on the man's rod and so he pointed to the rod so that the man would be able to catch whatever was on the end of it.
- P3. Luckily, he pulled the rod up in time. So he

thanked the boy and invited him home for some fish and chips.

Example 1 shows a 7 year old telling the story exactly as it was depicted on the slides. Example 2 shows that, for the same set of pictures, the 11 year old had constructed a personal interpretation of the events. Although the listener was unmasked, it is unlikely that the listener would have produced an identical story to that of the speaker when looking at the pictures, explaining why the oldest children often continued to treat the seeing partner as ignorant: using indefinites for an unmasked listener.

Therefore, whilst the youngest group and the 8 to 9 year olds were predominantly influenced by masking conditions in their choice of expression, factors such as story conventions and creativity may have affected the results of other age groups.

Nevertheless, there is evidence to suggest that the young child's choice of expression seems to be predominantly influenced by their general knowledge: Firstly, it appears that when a listener is one that young children regard as knowledgeable (for example, an experimenter), regardless of the listener's role as 'ignorant' in the experimental situation, young children will treat her as knowledgeable, unless there is evidence to make them believe that the experimenter is genuinely ignorant. Therefore, the results of experiment 2 do seem to support Emslie's (1986) criticism of Karmiloff-Smith's

findings as being a reflection of a failure to manipulate the social situation.

However, referents in the picture-stories were visible to the speakers, whereas in the real experiences conditions, the referents were absent and invisible to the listener. In previous experiments (Warden 1981, Emslie 1986), no difference was found between the child's sensitivity to the listener when referents were either visible or invisible. Moreover, an experiment I carried out in my undergraduate project revealed that whether the referents were visible or invisible had no single effect on the expressions children chose, but whether referents were typical of events in their everyday lives did influence their choice of expression, such that young children were unable to refer anaphorically when actions performed on objects were atypical of real-life events (for example, 'A cow pushes a car'). It seems that the child's inability to dissociate linguistic factors from their everyday experiences may account for many of the controversial findings between different experiments in the young child's understanding of indefinite and definite expressions. Secondly, even when a peer listener was unable to see the referent, the young child seemed to assume that a referent was familiar if it could be inferred to exist from the context, even when the particular discourse context contained idiosyncratic features that did not fit the prototype. This result supports McWhinney and Bates' (1978) claim that pre-schoolers "evidenced a fairly high level of sophistication in using world knowledge to make

judgements about newness". These results are also in agreement with the findings of Karmiloff-Smith (1979) and Zehler and Brewer (1982). Maybe Emslie's (1986) conclusion that "the extent to which subjects took their listener's general knowledge of contexts into account seemed limited" was a reflection of the fact that she only looked at the effect of general knowledge of contexts in one experiment and then she only used four referents and the contexts in which they were illustrated were not always clear.

There seem to be two possible explanations for these age differences in the use of context for determining reference. Firstly, the age decrease in context definites when contexts are more flexible may be due to age differences in the knowledge of the world in general.

A second possibility for the age differences may be that, what changes with development is an increasing concern to relate the particular context of discourse to knowledge of the world in general. For example, Barrett and Light (1976) found that in an experiment on children's drawings, children developed from being 'symbolists' (drawing the prototype) to 'intellectual realists' (drawing objects as they perceived them to be true in the 'world', without accomodating to differences in the picture) and finally, they became 'visual realists' (incorporating individual features of objects into their own perspective). When giving a story about the strange and magic inhabitant of a picture of an unusual house, only the older children (approximately 7 years) adapted to the atypical details and

drew the idiosyncratic features (intellectual realists).

It appears that young children are highly sensitive to social and pragmatic factors when making indefinite and definite reference and thus they appear to be aware of the 'familiarity' factor of in/definiteness. However, the finding in experiment one that younger subjects used pronouns ambiguously, suggests that young children are concerned with deictic reference: making objects in context salient (Karmiloff-Smith, 1979) rather than with anaphoric conservation of unique entities in a mental model (Emslie, 1986). Thus, it seems that young children can only consider the perspective of others when this depends on global knowledge factors (knowledgeable or ignorant listener), but not when they need to consider the particular type of knowledge that the listener possesses (specific/unique or non-specific representation of referents). The next chapter focuses predominantly on the 'specificity' factor.

CHAPTER 4

THE EFFECTS OF REFERENT 'SPECIFICITY'

4.1 ABSTRACT

Two experiments were carried out on children between 3 to 8 years of age. Experiment 4 manipulated whether a listener was knowledgeable or ignorant and whether referents were specific (singleton) or non-specific (identical). For experiment 5 listeners were always knowledgeable and a different task was used which aimed to make the referential array salient by increasing the size of the non-specific sets and by focusing on unmoved items rather than on the item to be referenced.

The children were sensitive to 'familiarity' but not 'specificity' for the task in experiment 4: They continued to use definite expressions when the listener was knowledgeable even when items were non-specific. However, the children did show a sensitivity to 'specificity' for the new task in experiment 5 and this sensitivity was greater in the older children.

It seems that young children are able to distinguish between specific and non-specific items in their choice of determiner. However, unlike their knowledge of 'familiarity', they do not seem to appreciate that 'specificity' is necessary for in/definite reference regardless of the salience of the referential context.

4.2 GENERAL INTRODUCTION

Two experiments are reported in this chapter. Whereas chapter 3 was concerned with listener 'familiarity', this chapter is primarily concerned with the 'specificity' of a referent for choosing between indefinite and definite reference. Interest here is whether children will choose their expression as a function of their knowledge of the 'specificity/non-specificity' of referents in the exophoric context.

In deciding between indefinite and definite expressions, a speaker needs to consider not only social-pragmatic factors (listener's explicit or implicit 'familiarity' with referents), but also the relationship between the intended referent and other referents in the referential context, or what was referred to in chapter one as 'specificity'. The diagram below shows the expressions required (indefinite/definite) across all 4 conditions of 'familiarity' and 'specificity'.

		<u>LISTENER</u>	
		IGNORANT	KNOWLEDGEABLE
<u>REFERENT</u>	NON-SPECIFIC	Indefinite	Indefinite
	SPECIFIC	Indefinite	<u>Definite</u>

As the diagram shows, definite expressions are only used when the listener is both familiar with and has specific

knowledge of the referent. Otherwise, indefinite expressions are necessary.

As mentioned before, Karmiloff-Smith was criticised by Emslie for failing to consider the child's interpretation of the listener's 'familiarity' with referents. However, Karmiloff-Smith was primarily concerned with relative-comparative factors rather than with social-pragmatic ones: In all her experiments she stipulated whether referents were specific ("different" from other referents) or non-specific ("identical" - more than one of its type in the referential array). Moreover, she often manipulated 'specificity' in her experiments. Thus, when Karmiloff-Smith consistently found that young children overused definite expressions, she was referring to the way that young children used definite expressions for non-specific items, rather than to the way that they used definites for ignorant listeners.

Therefore, Karmiloff-Smith concluded that the younger children were unable to use the articles with respect to the exophoric context: the relationship of intended referents to other referents in the array. Thus, she claimed that their overuse of definite articles indicated that they were using the definite article deictically: to point to a salient object.

However, Emslie (1986) partially replicated Karmiloff-Smith's 'hide and seek' experiment, with some changes to the experimental procedure. These changes were:

1. The experimenter had to close her eyes when hiding an object.

2. Only the subject/speaker could see the remaining items in the bag after an object had been hidden. These procedural changes were to ensure that the subject was treating the experimenter as ignorant of the object that had been hidden. Emslie claimed that it was possible that the children in Karmiloff-Smith's experiment could not accept that the experimenter had really forgotten what they had hidden and/or they believed that, like themselves, the experimenter could decipher the missing object from remaining items in the bag.

Emslie found that her change in procedure did serve to eliminate the overuse of definite expressions from the youngest group. Moreover, even the youngest subjects were able to discriminate between indefinite and definite reference on the basis of 'specificity' (singleton or identical): Although the listener was familiar with all the objects in the bag, young children preferred to use indefinite expressions when the hidden object was non-specific (identical item).

The previous experiments of this thesis have already manipulated listener 'familiarity' for specific items. Therefore, what still needs to be manipulated is the listener's 'familiarity' of referents for non-specific items.

Neither Karmiloff-Smith nor Emslie manipulated both listener 'familiarity' and referent 'specificity' in one experiment. Thus, for experiment 4, a peer listener is either familiar or unfamiliar with the referential context. Referents are either non-specific (more than one other

item) or specific (only one item). The task requires children to tell a peer which object is being moved by the experimenter. As for the experiments in chapter 3, the peer either can see or cannot see the referential context beforehand. If children are sensitive to the 'specificity' of referents when choosing between indefinite and definite expressions, then they are expected to continue to use indefinite expressions to refer to objects which are non-specific even when the listener has seen the referential context. However, 'familiarity' only overrides 'specificity' when the listener has not seen the referential context, in which case, all first references should be indefinite, irrespective of 'specificity'.

Owing to the contradictory findings in the literature concerning young children's sensitivity to the 'specificity' factor for determining reference, as well as the multiplicity of tasks that have been used by investigators making different predictions about children's abilities in this area, it seems that the problems for children are generally task-related. Thus, it appears that young children can use 'specificity' to determine reference, but they have trouble using 'specificity' in tasks which do not make the differences between referents salient for the child.

Those tasks which do seem to produce positive results are indirect problem solving tasks that require consideration of the entire referential context when deciding upon which object to refer to. For example, Emslie's version (1986) of Karmiloff-Smith's (1979) 'hide

and seek' task ensured that the child was in a position where it was necessary to work out which referent had been hidden, by looking at the objects remaining in the bag. Comprehension tasks, which focus less on the social aspects of the situation, also seem to produce more positive results than do production tasks (Karmiloff-Smith 1979). Finally, tasks which have generally increased the salience of an object in its referential context, by such procedures as having larger sets of identical items (see Karmiloff-Smith's experiment 12), have also produced better results.

Experiment 5 uses the procedure of experiment 4, except that peer listeners always see the referential context beforehand. The task increases the number of items which are identical and focuses on items that are not moved, as opposed to moved items and instructions emphasise the speaker's role in informing the listener of the correct object, rather than their role in telling the listener the object being moved. The task used in the previous experiment is an additional condition. If children are better at choosing between indefinite and definite reference on the new 'Context-salient' task than on the original 'Object-salient' task, then it would seem that children can use 'specificity' in determining reference, but fail spontaneously to see the need to consider the referential context in situations/tasks which do not make the entire array of objects salient. Alternatively, if children are just as poor at using indefinites for non-specific items in both tasks, the

results are more likely to be a reflection of the child's lack of appreciation of the influence of 'specificity' on reference form per se.

The main questions asked in this chapter are:

1. At what age do children begin to choose between indefinite and definite expressions as a function of referent 'specificity'?

2. At what age do children spontaneously consider the 'specificity' of referents for in/definite reference, irrespective of the salient features of that particular situation?

4.3 EXPERIMENT 4: The effects of the status of referents in the immediate context.

INTRODUCTION

For the following experiment, a peer listener was either familiar or unfamiliar with the referential context. Referents were either non-specific (more than one other item) or specific (only one item). Interest was in whether children would use indefinite expressions appropriately: using indefinites for all conditions except for the case where the listener was familiar with referents and the intended referent was specific.

METHOD

Subjects

48 children selected from a primary and junior school in Durham took part in the experiment. They were divided equally into 2 age groups (3 to 5 years [mean age: 4.7] and 6 to 8 years [mean age: 7.3]) and thus, there were 24 children in each group. Each subject was matched with another subject of approximately the same age, so that there were 12 dyads in each age group. 10 undergraduate students of Durham university volunteered for the adult procedure of the experiment.

Materials

2 sets of 2 identical objects (2 candles, 2 bricks) and 4 singleton objects (tree, ball, bus-stop, key) served as the

entire referential context for this experiment. A board with a coloured strip at the front was used for placing and moving objects. Table 4 below illustrates the experimental layout.

Objects -	X	X	a	b	c	d	Y	Y
-----------	---	---	---	---	---	---	---	---

Coloured strip.

LISTENER

SPEAKER

KEY:

X Y - Identical items

a b c d - Singleton items

Figure 4: EXPERIMENTAL LAYOUT

Recording equipment and a listener mask made up the remaining apparatus.

Design and Procedure

Two variables were manipulated. These were:

1. 'SPECIFICITY' OF REFERENT: Singleton (specific) or Identical (non-specific) referent.

2. LISTENER 'FAMILIARITY' WITH ARRAY: Listener knowledgeable

(familiar) or ignorant (unfamiliar) of array.

The 'familiarity' variable was a between-subjects factor and therefore, in each age group, dyads were equally and randomly divided into one of the two conditions (LK, LI). The 'specificity' variable was a within-subjects factor and therefore, all dyads received both conditions: Each pair referred to all the objects in the array.

The following procedure was carried out for each pair of children acting as subjects:

Phase 1: Familiarisation Period

For LI dyads, only the subject that was going to act as speaker was allowed to enter into the experimental room (playroom) at this point. For LK dyads, both the subjects of a pair were taken into the experimental room from the start of the experiment.

Subjects permitted to enter the room were shown the array of items positioned on the board, as presented in the diagram of the experimental layout. The subjects were asked to name each item and were then asked 'How many X's are there? These questions were to rule out any difficulties with the experiment which were due to a misunderstanding of the initial array of items.

At this point, for LI dyads, the experimenter went outside the room to blindfold the second subject and to escort the listener to her/his position in the room. At first, some subjects were reluctant to be a listener in the

LI condition. However, after being reassured that they would be able to see the objects afterwards, all subjects eventually felt comfortable as an ignorant listener.

Phase 2: Instructions

Each dyad was given the following instructions: "You (to L) will always wear the mask and I will move one of the items on the board onto a/this green line. Then, I want you (to S) to tell L which item I have moved, so that L will know exactly which object I have moved.

A few practice trials were run to ensure that subjects understood the experimental requirements.

Phase 3: Running the experiment

For LK dyads, the listener put on the mask at this point. The experimenter then moved items one by one onto the green strip and the speaker had to tell the listener which items were being moved. The experiment continued until all items had been moved onto the green strip. Therefore, there were 8 trials in total (8 individual items). The order in which the experimenter moved the objects was completely random.

Procedure for adult subjects

The 10 adult subjects were given the same task, but instead of being sorted into pairs, 5 subjects were told, 'pretend that you are talking to somebody who has seen the objects on the board, but who is now blindfolded' (LK condition) and 5 subjects were told, 'pretend you are talking to somebody who has never seen the items on the board and who

still cannot see the items' (LI condition).

The tape recorder was switched on just before the experimenter moved the first item. Speaker's expressions were subsequently transcribed for analysis

Scoring

Firstly, for all types of items, if a subject used a PRONOUN for reference, this was scored in a separate category.

For first identical items to be moved (non-specific items), subjects received either an INDEFINITE (a, another), a DEFINITE ARTICLE PLUS NOUN (DA + NOUN) or a DEFINITE ARTICLE PLUS ..NTH (DA + ..NTH) score. The latter category included definite expressions with a 'First' modifier.

For second identical items to be moved (the second of a pair of identical items), subjects received either an INDEFINITE, a DEFINITE ARTICLE PLUS NOUN (DA + NOUN), or a DEFINITE ARTICLE PLUS ...NTH (DA + ..NTH) score. The latter included 'SECOND', 'OTHER' or 'LAST' modifiers.

For singletons, subjects received either an INDEFINITE or a DEFINITE score.

RESULTS

Adult scores

Table 4-A presents adult indefinite and definite scores both for the first of two identical items to be moved and for the second of two identical items to be moved in each listener condition. As there were 2 sets of 2 identical items and 5 subjects in each listener condition, scores were out of 10 for both first and second identical items.

Out of 10	FIRST IDENTICAL		SECOND IDENTICAL	
	INDEF.	DEF.	INDEF.	DEF.
<hr/>				
LISTENER CONDITION				
LI	10.0	0.0	10.0	0.0
LK	9.0	1.0	0.0	10.0

TABLE 4-A: ADULT INDEFINITE AND DEFINITE SCORES FOR FIRST IDENTICAL AND SECOND IDENTICAL ITEMS IN EACH LISTENER CONDITION.

The table shows that adults preferred indefinite expressions for first identicals irrespective of the listener condition, whereas for second identicals, all scores were indefinite in the LI condition and all scores were definite in the LK condition.

Mann-Whitney tests showed no significant effect of listener condition on indefinite or definite scores for

first identical referents (U=10.0, NS, corrected for ties), but there was a significant effect of listener condition on indefinite and definite scores for second identical referents (U=0, <0.005, corrected for ties).

Finally, 8 of the definite scores for second identical items (LK condition) involved a modifier: 'The last X' or 'the second X'.

Table 4-B presents adult indefinite and definite scores for singleton items in each listener condition. As there were 4 singleton items and 5 subjects in each listener condition, scores are out of 20.

Out of 20	INDEFINITE	DEFINITE	TOTALS
LISTENER CONDITION			
LI	20.0	0.0	20.0
LK	2.0	18.0	20.0

TABLE 4-B: ADULT INDEFINITE AND DEFINITE SCORES FOR SINGLETON ITEMS IN EACH LISTENER CONDITION.

The table shows that for the LI condition, all scores were indefinite, whereas for the LK condition, most scores were definite. Analysis of variance showed no significant overall effect of indefinite verses definite expressions for singleton referents (F=1.0, df=1,8, NS), but there was a significant interaction between listener

condition and indefinite/definite expressions (F=81.0, df=1,8, p<0.001).

Children's scores

Firstly, no subject used a pronoun to refer to any of the items.

1. FIRST IDENTICALS

Each subject received a total of 2 scores for first identical items. Table 4-C presents overall total indefinite and definite scores, (keeping definite plus 'first' modifier expressions separate from definite article plus noun only expressions), in each listener condition and for both age groups. As there are 6 subjects per condition, scores are out of 12.

AGE	(Out of 12)	DEF.ART	DEF.ART +	TOTALS
LISTENER CONDITION	INDEF.	+ NOUN	MOD.+NOUN	
3-5				
LI	11.0	1.0	0.0	12.0
LK	1.0	11.0	0.0	12.0
6-8				
LI	10.0	0.0	2.0	12.0
LK	2.0	8.0	2.0	12.0

TABLE 4-C: INDEFINITE AND DEFINITE SCORES FOR FIRST IDENTICALS IN EACH LISTENER CONDITION AND IN EACH AGE GROUP.

The table shows that, despite the few cases of definite + 'first' modifier expressions from 6 to 8 year olds, for both age groups, indefinite expressions were preferred for the LI condition and definite expressions were preferred for the LK condition. (Note that definite + 'first' expressions are appropriate definites when the listener is either knowledgeable or ignorant. This is because, the adjectives 'first' and 'last' must always be preceded by definite articles and because, if a non-specific item is referred to as 'the first x', then this item becomes specific, due to the fact that it is the only object of that class which is moved first).

Wilcoxon tests revealed a significant difference between indefinite and definite (without modifier) expressions for 3 to 5 year olds, both in the LI condition ($z=-2.02$, $p<0.05$, 2-tailed) and in the LK condition ($z=-2.02$, $p<0.05$). However, for the 6 to 8 year olds, a significant indefinite verses definite (without modifier) difference was observed for the LI condition ($z=-2.20$, $p<0.05$), but not for the LK condition ($z=-1.26$, NS). However, this latter non-significant finding was completely due to the performance of one 6 to 8 year old subject in the LK condition, who used indefinite expressions for first identical items.

Nevertheless, Mann-Whitney tests showed significant effects of listener condition on indefinite scores for first identical items, in both the 3 to 5 year old group ($U=0.5$, $p<0.005$, corrected for ties) and the 6 to 8 year old group ($U=4.0$, $p<0.05$, corrected for ties).

However, Mann-Whitney tests revealed no significant effect of age group on indefinite expressions for first identical items, in either the LI condition (U=15.0, NS, corrected for ties), or in the LK condition (U=17.5, NS, corrected for ties).

2. SECOND IDENTICALS

Again each subject received a score out of 2 for second identical items. Table 4-D presents total indefinite and definite scores (again keeping 'the last/second' expressions separate) in both listener conditions and for both age groups.

AGE	(Out of 12)	DEF.ART	DEF. ART +	TOTALS
LISTENER CONDITION	INDEF.	+ NOUN	MOD. +NOUN	
3-5				
LI	9.0	1.0	2.0	12.0
LK	0.0	12.0	0.0	12.0
6-8				
LI	9.0	2.0	1.0	12.0
LK	5.0	6.0	1.0	12.0

TABLE 4-D: INDEFINITE AND DEFINITE SCORES FOR SECOND IDENTICALS IN EACH LISTENER CONDITION AND IN EACH AGE GROUP

Despite the 2 'the + last X' scores from 3 to 5 year olds

in the LI condition, 3 to 5 year olds preferred indefinite expressions in the LI condition and used definite expressions in the LK condition.

Whilst one subject in each condition used 'the + last X' expressions for the 6 to 8 year olds, indefinite expressions were again preferred in the LI condition. However, for the LK condition, 6 to 8 year olds do not show any clear preference when referring to second identicals: 40 percent of expressions were indefinite in this condition.

In fact, for the LK condition, the difference between indefinites and definites was significant, using Wilcoxon tests, for the 3 to 5 year olds ($z=-2.20$, $p<0.05$, 2-tailed), but not for the 6 to 8 year olds ($z=-0.27$, NS).

However, there was no significant difference between indefinite and definite (without modifier) expressions for either age group in the LI condition (3-5: $z=-1.83$, NS. 6-8: $z=-1.36$, NS).

Nevertheless, there was a significant effect of the listener condition for 3 to 5 year olds ($U=0.0$, $p<0.005$, corrected for ties), but not for the 6 to 8 year olds ($U=8.0$, NS).

3. SINGLETONS

Each speaker received a total of 4 scores for singleton items. Table 4-E presented total indefinite and definite scores in each listener condition and in each age group for singleton items. As 6 subjects were in each condition, totals are out of 24.

AGE (Out of 24)	INDEFINITE	DEFINITE	TOTALS
LISTENER CONDITION			
3-5			
LI	22.2	2.0	24.0
LK	0.0	24.0	24.0
6-8			
LI	24.0	0.0	24.0
LK	4.0	20.0	24.0

TABLE 4-E: INDEFINITE AND DEFINITE SCORES FOR SINGLETONS IN EACH LISTENER CONDITION AND IN EACH AGE GROUP.

The table shows that, for both age groups, there is a strong preference for indefinite expressions in the LI condition and for definite expressions in the LK condition. Analysis of variance showed that there was no overall difference between indefinite and definite expressions in singleton trials ($F=0.23$, $df=1,20$, NS). However, there was a significant interaction between listener condition and indefinite/definite scores ($F=100.23$, $df=1,20$, $p<0.001$), but there was no 3-way interaction between age, listener condition and indefinite/definite scores ($F=0.23$, $df=1,20$, NS).

DISCUSSION

This study has shown that, for non-specific items, 3½ to 8 year olds preferred definite expressions when the listener was knowledgeable. Therefore, the results suggest that, for children under 8 years of age, the most significant factor in determining their choice of in/definite expression is the listener's 'familiarity' with referents and they seem uninfluenced by relative-comparative factors.

The results also confirm those found from the experiments of the previous chapter: When introducing specific/singleton items to a peer listener, from 3½ years of age, indefinite expressions are preferred when the listener is unfamiliar with the referents and definite expressions are preferred when the listener is familiar with the referents.

6 to 8 year olds showed no particular preference for indefinite or definite expressions to refer to second identical items when the listener was knowledgeable, even though the listener could have inferred that there was only one of those items left to be moved. One possibility is that children of this age have problems distinguishing between 'specificity' and quantity of referents.

4.4 EXPERIMENT 5: The effects of increasing the salience of referents in the immediate context.

INTRODUCTION

For this experiment, listeners were always knowledgeable of referents and a task was used which differed slightly from the one used in experiment 4. For this new task, instructions emphasised that the speaker had to inform the listener of the correct object rather than telling the listener the object being moved. Also, it was remaining items, as opposed to the item currently being referred to, that was left visible to the speaking subject and finally, each set of 'identical' items was increased to 4. Therefore, this second type of task was intended to focus the child's attention on the referential context.

It was predicted that children would be better on the new task at using indefinite and definite reference to discriminate between specific and non-specific items, compared with their performance on the old task (the task used in experiment 4). This is what would be expected if it is the case that children can use 'specificity' in determining reference, but fail to spontaneously see the need to consider the referential context in situations/tasks which do not make the entire array of objects salient.

METHOD

Subjects

48 children who had not participated in the previous experiment acted as subjects in this experiment. All subjects were selected from schools in Durham. They were divided equally into 2 age groups (3 to 5 years [mean age: 4.4] and 6 to 8 years [mean age: 7.5]) and each subject was matched with another subject of approximately the same age. Thus, there were 12 dyads in each age group.

Materials

The 'Object-salient' task used the same items as for the previous experiment, with the addition of another set of 2 identical objects, so as to match up the number of items to be referenced between the two tasks. These items were: 3 sets of 2 identical items: 2 candles, 2 bricks, 2 stars. 4 singleton items: tree, ball, bus-stop, key. The board with the coloured strip that was used in experiment 4 was used again for this task.

The 'Context-salient' task used the following items:

2 sets of 4 identical items: 4 candles, 4 bricks.

2 singleton items: tree, ball.

A board without a coloured strip was used for this task. In addition, a plastic cup and a deep cardboard box were used for the 'context-salient' task.

Recording equipment and a listener mask made up the remaining apparatus. The experimental layout was the same

as for experiment 4.

Design and procedure

The variable manipulated in this experiment was 'TYPE OF TASK': 'Object-salient (OS)' or 'Context-salient (CS)'. This was a between-subjects factor and therefore, half the dyads in each age group received the OS task and the other half received the CS task. However, listener familiarity was not manipulated in this experiment. Instead, all listeners were knowledgeable.

The procedure was the same as for experiment 4, except that the 'Ignorant Listener' condition was excluded and, for the CS task, the instructions and procedure were altered in a way which served to increase the salience of the referential context for the speaker.

For the CS task, the instructions were, "You (to L) will always wear the mask and I want you (to S) to shut your eyes while I cover over one of the items on the board with this plastic cup. Then I want you (to S) to open your eyes and work out which object it is that I have covered over. Then I want you to say to L which object you think it is. You (to S) will get a point for saying the correct object to L. Then I will show you the covered object and put it in this box."

In order to try to make the length of the instructions as similar as possible between the two tasks, subjects given the OS task were first given the instructions used for experiment 4. Then they were given the final instruction, "You (to S) will get a point for

telling L which object I have just moved."

Thus, the differences in procedure for the CS task were; Firstly, the speaker had to close her eyes whilst the experimenter made the necessary changes to the array; secondly, the experimenter covered an item with the cup and finally, before the speaker had to close her eyes again for the next item, the experimenter uncovered the item and took it off the board (into the box, where it could not be seen by the speaker). The CS task continued until all referents had been covered and taken off the board. (This latter procedure was necessary to match the two tasks because, for the OS task, the speaker was referring to an array which decreased in size throughout the experiment and the second of two identical referents to be moved may have been referred to with a definite expression, because the listener could infer that this item was the only one of its type left still to be moved. Thus, moving each item off the board in the CS task, ensured that the tasks were matched in these respects).

The experiment continued until either all items had been moved onto the green strip (OS task) or until all the items had been moved off the board and into the box (CS task). Thus, there were 10 trials in total: 3 by 2 identicals plus 4 singletons for the OS task and 2 by 4 identicals plus 2 singletons for the CS task. The order in which the experimenter moved the objects was completely random. The tape recorder was switched on just before the experimenter moved the first item and speakers' expressions

were subsequently transcribed for analysis.

Scoring

Firstly, for all types of items, if a subject used a PRONOUN, this was scored in a separate category.

For non-specific items (identical objects that were not the last of their type to be moved/covered: 1st to 3rd items on CS task and 1st item on OS task), subjects received either an INDEFINITE (a, another), a DEFINITE ARTICLE PLUS NOUN (DA + NOUN) or a DEFINITE ARTICLE PLUS ..NTH (DA + ..NTH) score. The latter category included definite expressions with a 'First' modifier and, for the CS task, this also included 'second' and 'third' modifiers. (NOTE: if on the CS task, a subject used DA + OTHER modifiers for second and third identicals, these scores were included in the DA + NOUN category, because using the word 'other' does not make a physically non-specific referent become specific. Alternatively, 'the second X' expressions are appropriate because non-specific referents are made specific through the mention of the the order in which they are moved/covered).

For last identicals (the last identical item of its type left to be moved/ covered: 4th item on CS task and 2nd item on OS task), subjects received either an INDEFINITE, a DEFINITE ARTICLE PLUS NOUN (DA + NOUN), or a DEFINITE ARTICLE PLUS 'LAST' (DA + LAST) score. For the OS task the DA + LAST category also included 'OTHER' or 'SECOND'

modifiers.

For singletons, subjects received either an INDEFINITE or a DEFINITE score.

RESULTS

Firstly, no subject used a pronoun to refer to an item.

As there were different quantities of each type of item between the two tasks, all scores in the tables have been converted to percentages, so that comparisons between the tasks can be made.

1. NON-SPECIFIC ITEMS

For referring to non-specific items, subjects in the OS group received 3 scores (1st moved identical item for 3 sets), whereas subjects in the CS group received 6 scores (1st, 2nd and 3rd covered identical items for 2 sets). Table 5-A presents percentage scores for indefinites, DA + Noun and DA + ...nth expressions for non-specific items in each age group and for each type of task.

AGE (percentage scores)		DEF.ART	DEF. ART +
TYPE OF TASK	INDEF.	+ NOUN	...NTH
<u>3-5</u>			
Object Salient	5.6	94.4	0.0
Context Salient	50.0	50.0	0.0
<u>6-8</u>			
Object Salient	11.1	77.8	11.1
Context Salient	83.3	11.1	5.6

TABLE 5-A: PERCENTAGE INDEFINITE AND DEFINITE SCORES FOR NON-SPECIFIC ITEMS BY EACH TYPE OF TASK AND IN EACH AGE GROUP.

For reference to non-specific items, the table shows that, whilst, for the OS task, both age groups preferred definite expressions, for the CS task, the 3 to 5 year olds showed no difference between indefinite and DA + NOUN expressions and the 6 to 8 year olds showed a clear preference for indefinite expressions. Only the 6 to 8 year olds used DA +...NTH expressions, but this was a minority response for both types of task. Analysis of variance showed no significant difference between indefinites and DA + NOUN scores overall ($F=0.06$, $df=1,20$, NS). However, there was a significant interaction between type of task and indefinite/DA + noun expressions ($F=42.88$, $df=1,20$, $p<0.001$) as well as a significant interaction between age group and indefinite/DA

+ noun expressions ($F=13.24$, $df=1,20$, $p<0.005$). Finally, there was a significant 3-way interaction between age, task and indefinite/DA + NOUN scores ($F=7.12$, $df=1,20$, $p<0.05$) and this effect is due to the fact that, whilst both age groups used more indefinites for the CS task, the 6 to 8 year olds used a far greater number of indefinites in this task compared with the youngest group.

2. LAST IDENTICALS

For referring to last identicals, subjects in the OS group received 3 scores (2nd moved identical item for 3 sets), whereas subjects in the CS group received 2 scores (4th covered identical item for 2 sets). Table 5-B presents percentage scores for indefinites, DA + Noun and DA + last expressions for last identical items in each age group and for each type of task.

AGE (percentage scores)	DEF.ART	DEF. ART +	
TYPE OF TASK	INDEF.	+ NOUN	LAST
<u>3-5</u>			
Object Salient	0.0	100.0	0.0
Context Salient	25.0	75.0	0.0
<u>6-8</u>			
Object Salient	33.3	61.1	5.6
Context Salient	25.0	50.0	25.0

TABLE 5-B: PERCENTAGE INDEFINITE AND DEFINITE SCORES FOR LAST IDENTICALS BY TYPE OF TASK AND IN EACH AGE GROUP.

The table shows that, for last identical items, whilst a low proportion of DA + LAST scores were produced from the 6 to 8 year olds, most expressions involved a DA + NOUN in all conditions, irrespective of age and type of task. Nevertheless, the 6 to 8 year olds produced more indefinites than did the youngest group and 3 to 5 year olds produced 25 percent of indefinite expressions for the CS task compared with zero indefinite scores for the OS task. In fact, Wilcoxon tests only showed a significant difference between indefinite and DA + NOUN expressions for the 3 to 5 year olds on the OS task ($Z = -2.20$, $p < 0.05$, 2-tailed), but not for any other condition (OS task: 6-8, $Z = -1.05$, NS, 2-tailed, CS task: 3-5, $Z = -1.60$, NS, 2-tailed, 6-8, $Z = -0.94$, NS, 2-tailed).

For the OS task, Mann-Whitney tests revealed a significant effect of age on indefinite expressions ($U = 6.0$, $p < 0.05$, corrected for ties) as well as a significant effect of age on DA + NOUN expressions ($U = 39.0$, $p < 0.05$, corrected for ties). However, for the CS task, there was no significant effect of age on indefinite expressions ($U = 18.0$, NS, corrected for ties), but there was a significant effect of age on definite expressions ($U = 39.0$, $p < 0.05$, corrected for ties). However, this latter significant effect is due to the small proportion of DA + LAST expressions produced by the 6 to 8 year olds, which lowered the overall DA + NOUN expressions in this age group.

interaction between age and indefinite/definite expressions (F=1.0, df=1,20, NS).

Converting individual scores to percentages for each subject and using Wilcoxon tests to compare indefinite expressions used between singletons and non-specific items, no significant differences were found in either age group for the OS task (3-5: Z=-1.0, NS, 2-tailed, 6-8: Z=-1.0, NS, 2-tailed). However, differences were significant in both age groups for the CS task (3-5: Z=-2.20, p<0.05, 2-tailed, 6-8: Z=-2.20, p<0.05, 2-tailed).

DISCUSSION

The hypothesis that young children would show a sensitivity to 'specificity' when the referential context was made salient has been supported. When a task was used which required indirect problem solving to choose the correct, hidden referent before being able to actually refer to that object; when the quantity of objects in an identicals set was increased to 4 and when the instructions emphasised the importance of guessing the correct object, as opposed to telling the listener the object being moved, even subjects in the youngest age group (3½ to 5 years) were discriminative in their choice of expression as a function of referent 'specificity': Whilst definite expressions were preferred for singletons and for last-moved identical items, there was no such definite preference for non-specific items, despite the fact that listeners were always familiar with the referential array.

However, there was a clear age difference between the 3½ to 5 year olds and the 6 to 8 year olds in their sensitivity to 'specificity' on the 'Context Salient' task. Whereas, the youngest group failed to show a preference for either indefinite or definite expressions for non-specific items, the 6 to 8 year olds showed a clear preference of indefinites for non-specific items in the Context Salient task. These results suggest that, whilst the youngest group are aware of the effects of 'specificity' on form of reference, even when a task makes the referential context salient, it seems that the influence of listener 'familiarity' in determining reference still has a strong effect compared with the 'specificity' of referents. However, from 6 years of age, children do appear to let 'specificity' override 'familiarity' if the task favours it.

Also, as for experiment 4, there was a greater tendency for indefinite reference for last moved identical items compared with singletons. Again, this may be due to the child's initial confusion between quantity and 'specificity'.

These results suggest that young children can use 'specificity' to determine reference. The problem for children at this age appears to be due to an inability spontaneously to see the need to consider the referential context in situations/tasks which do not make the entire array of objects salient, as shown by the fact that in the original 'Object Salient' task, subjects continued to prefer definite expressions when the listener

had already seen the array, regardless of referential
'specificity'.

4.5 GENERAL DISCUSSION

The results of experiment four suggest that children under 8 years of age are poor at adapting their choice of indefinite or definite expression as a function of the 'specificity' of referents in the perceptual array and their choices of expressions are predominantly influenced by their experience of the listener's direct perception of the referents.

The question remains as to why young children used indefinite expressions appropriately for non-specific items in Emslie's 'hide and seek' task, but were insensitive to non-specificity in this experiment? There are 2 main differences between the 'hide and seek' procedure and the task used here which may account for these contradictory findings. Firstly, the speaking child was required to solve the problem as to which referent had been hidden by the experimenter in the 'hide and seek' task, whereas the task used in this experiment ensured that the speaker always knew which referent was being moved by the experimenter. Therefore, for this experiment, the child was probably providing expressions almost automatically, whereas, as the 'hide and seek' task required conscious decision making processes, the expressions chosen by the child were probably a result of being in a position where they were encouraged to focus their attention on the status of the object being referred to. Secondly, for the 'hide and seek' task, the speaker was made to focus on the remaining objects that were left in the bag, rather than on the object being mentioned, which was hidden from the

subject. By contrast, in this experiment, the speaker was encouraged to concentrate on the actual object that was moved onto the green strip. Therefore, it is possible that subjects were better at considering 'specificity' when the entire set of objects, as opposed to the individual object being referenced, was made a salient feature for the child.

In this experiment, the fact that the movement of one object set that object apart from the rest of the array, meant that this object was probably more salient to the child than the status of that object in relation to the remaining array of unmoved objects.

The results of experiment 5 showed that, using a task which increased the salience of the referential array did reveal a knowledge of 'specificity' in young children, although this was greater in children over 5 years of age. The fact that young children will show a sensitivity for 'specificity' when the referential array is made salient to them shows that they are aware of the differences between indefinite and definite expressions in labelling specific and non-specific reference. What they appear to lack is the knowledge that 'specificity' is significant for in/definite reference regardless of context.

Therefore, it is probably not accidental that even Karmiloff-Smith found more indefinite expressions from young children in the 'hide and seek' task compared with many of her other tasks and that her subjects were better at comprehension tasks and when four items made up the identicals set.

What seems to draw the line between the

performance of young children compared with that of adults is that young children do not appreciate the objective status of 'specificity' for indefinite verses definite reference: that it must necessarily hold across all situations and environments.

Finally, it is possible that development of an appreciation of the need to consider 'specificity' does proceed at the same time that the child comes to appreciate that 'specificity' is dissociated from quantity, because whilst indefinite reference for last moved identicals is unnecessary (the listener can infer that there is only one of its type left to be moved), children were more likely to use indefinites for last placed identical items than for singleton items.

One prediction which can be made from these findings is that young children will have trouble with anaphoric reference. This is because, anaphoric reference requires definite expressions to refer to an entity which is in the situation of discourse. Thus, the referential context for anaphoric reference is a mental set of entities (known as a 'mental model' by some investigators: Johnson-Laird, 1980, Emslie, 1986), as opposed to the physical array of objects that have been looked at in this chapter. In light of the current results, if the referential context is mental rather than physical, one might expect that young children will fail to show a sensitivity for 'specificity', because the referent set is unlikely to be salient. Children's production of anaphoric reference is the main issue of the next chapter.

CHAPTER 5

REFERENCE IN DISCOURSE PRODUCTION

5.1 ABSTRACT

Experiment 6 combined the variables of listener 'familiarity' and referent 'specificity' in a story-telling task and it was found that the listeners' 'familiarity' with referents had an overriding effect on young children's choice of indefinite and definite expressions, irrespective of the 'specificity' of the referents.

Experiment 7 involved getting subjects to tell stories in a referential communication task. When referents were physically different, young children were better at providing the relevant information to discriminate between referents than when referents were physically identical. Young children were particularly poor at using 'episodic' or 'character' information (for example, the clown who was skipping' or 'the one who likes ice-cream'), to discriminate between referents. Experiment 8 showed that young children would produce relative clauses when processing demands were reduced, but redundant relatives were as frequent as discriminative relatives. Experiment 9 involved subjects talking about actions performed on dolls. The results of experiment 7 were replicated with the additional finding that young subjects continued to give 'colour' information even when it was irrelevant. Experiment 10 showed the same developmental effects of discriminative reference in a more 'game-like' model village task, which compared 'elicited'

with 'spontaneous' comparisons between referents. Finally, experiment 11 showed that young children would produce pronouns when they were insufficient for unambiguous anaphoric reference.

It appears that young children only use information which is salient to them, both for reference using definite articles and for pronominal reference, whereas older children use information necessary for unambiguous anaphoric reference, regardless of salience.

5.2 GENERAL INTRODUCTION

The experiments in this chapter are intended to look at indefinite and definite reference used by different age groups in discourse. Whilst the experiments in chapter three used story-telling tasks, interest was in the speaker's sensitivity to the social-pragmatic conditions of discourse. Chapter 4 was concerned with relative-comparative factors ('specificity' of referents), but failed to look at children's references in a discourse situation. Nevertheless, both 'familiarity' and 'specificity' are significant for indefinite and definite reference in discourse.

Moreover, definite reference in discourse may be anaphoric: referring to an entity introduced either explicitly or implicitly in previous discourse. The literature which looks at the age when children can understand anaphoric reference covers a very broad field. Some investigators have looked at children's ability to make the distinction between indefinite and definite

articles (Karmiloff-Smith 1979, Warden 1976, Emslie and Stevenson 1981). These studies have either been concerned with the 'specific/non-specific' distinction or with the social features of communication (listener's 'familiarity/ non-familiarity'). Other studies have concentrated on the nature of the definite expressions that children use (Emslie 1986, Hamburger and Crain 1982): For anaphoric reference, it is necessary that the speaker incorporates presuppositional information in the definite predicate so that the intended referent is specific for the listener. The short text below presents an example of anaphoric reference.

"A girl walked over to the counter. Another girl came into the shop. The girl who was at the counter was doing her weekly shopping."

The entire definite description necessary to identify the specific referent in this example was a noun phrase plus a relative clause (The girl who was at the counter). In the first sentence, two referents of the same gender were introduced and therefore, the presuppositional information concerning each of the girl's respective actions was necessary for the listener to unambiguously identify a specific referent.

Studies concerning the form of definite expression used by children for anaphoric reference have been controversial with regards to whether results are due to information processing factors, not related to knowledge of anaphoric reference (eg. memory limitations) and/or whether results are confounded by methodological flaws, such as

using contexts where anaphoric reference would be redundant.

The six experiments reported in this chapter are concerned with the way that 'familiarity' and 'specificity' are used for producing reference in discourse.

Experiment 6 manipulates the 'familiarity' of the listener when telling picture-stories and the need to refer to either specific or non-specific referents. It was expected that young children would overuse definite expressions inappropriately for non-specific items when listeners are familiar. Moreover, any lack of appreciation of 'specificity' should result in ambiguous anaphoric reference when the definite description requires further presuppositional information for making the referent specific. These results were predicted because, as the referential context is less likely to be a salient feature in a discourse model compared with a concrete array of objects, young children are unlikely to focus on the relationship between the intended referent against the entire referential set.

Experiment 7 looks at anaphoric reference more closely, comparing different types of 'specificity' information (physical verses episodic) necessary for unambiguous reference. Investigators have found evidence to suggest that understanding ambiguity of reference is a developmental phenomenon (Robinson and Robinson 1978, Flavell 1985). Others have found evidence to show that when tasks are made less complex and instructions are clearer, young children's performance is improved considerably. The

question remains as to whether the conflicting findings shown between different tasks/procedures are a reflection of the fact that children can only produce unambiguous reference when the necessary information required for reference is salient to the speaker, just as the procedural differences in tasks looking at children's production of in/definite reference is a function of the salience of the exophoric context. Thus, one might expect young children to produce less ambiguous anaphoric reference when 'specificity' information is salient, such as 'Physical Characteristics' (colour, size), but will show greater ambiguity when 'specificity' information is not salient 'Episodic' information. Moreover, as this task meant that it was necessary to provide unambiguous anaphoric reference for the correct performance of the listener, other effects such as the listeners' judgements regarding attributions of reference failure are also looked at.

In one of her experiments, Emslie (1986) used a story-telling task with three conditions: Referents were either singletons (two referents of different genders), similars (two referents of the same gender but of different sizes) or identicals (two referents of the same gender and same sizes). She found that young children's performance for the similars and identicals versions was considerably poorer than for the singletons. Emslie argued that this difference was due to a particular linguistic difficulty in using relative clauses and partitives, rather than due to a general problem with anaphoric reference.

The literature concerning children's understanding

and use of relative clauses is undoubtedly controversial. Slobin (1966) argued that self-embedded relatives ('the man who came last Saturday was very friendly') are universally more difficult for children to process than right-branching relatives ('the box fell onto the lady who was wearing the red hat') and therefore, he proposed the 'Interruption Hypothesis': that self-embeddedness is difficult, due to the fact that the main clause has to be held in memory whilst the relative clause is being processed. This claim has been empirically supported by some investigators (Menyuk 1969, Slobin and Welsh 1973), whilst Sheldon (1974) found evidence that conflicted with Slobin's hypothesis. However, Sheldon and others have independently found evidence to suggest that 4 to 5 year olds are using particular strategies to interpret relatives in comprehension acting-out tasks and that this is what accounts for their errors in relative interpretation. For example, Sheldon's 'Parallel Function Hypothesis' stated that, children will choose the noun phrase that is in the same grammatical position (subject, object) as the relative noun phrase to be the referent of the relative. Others have found evidence to suggest that children use other strategies in interpreting relatives: extraposition hypothesis, conjoined clause strategy (Tavakolian 1978) and the bird-in-hand strategy (Legum 1975).

However, in a longitudinal study, Hamburger (1980) found evidence of 'protorelatives' in 2 year olds, such as "This my did it" to refer to a particular picture, which he interpreted as meaning something like 'This is my thing

such that I did it'.

Moreover, Hamburger and Crain (1982), presented young children with mini-stories involving identical referents (two referents of the same gender) and they used an elicited production procedure which involved asking subjects questions, to which they had to answer by giving a noun phrase rather than a complete sentence. They found that 73 percent of 4 year olds used relative clauses to discriminate between identical referents in this experiment. Interestingly, in a similar task by Potts, Carlson, Cocking and Copple (1979), 69 percent of infants produced relatives when they did not have to complete a whole sentence. Hamburger and Crain argued that the reason why children failed to use relative clauses in many previous experiments was because, only one referent had been in the array and thus, relative clause information would have been redundant. This evidence makes two claims regarding children's use of relatives: Firstly, young children have the linguistic ability to produce relative clauses when information processing demands do not restrict performance and secondly, young children are sensitive to the supporting context of relatives because they use relative clauses only when there is more than one potential referent. However, from an example of Hamburger and Crain's story-question pairs: "There were two brothers. One boy ate dinner and the other boy went to bed without dinner. Which boy was sad?" (Typical response: the one that didn't eat dinner), it seems possible that relatives were used in their

experiment merely because they were elicited from subjects in a choice decision task which drew the child's attention to the differences between potential referents. In addition, Hamburger and Crain did not manipulate felicity conditions within the same experiment. Thus, in experiment 8 half of the story-question pairs abide by felicity conditions (more than one referent available) and the remaining half are unfelicitous (only one available referent). It was hypothesised that in an elicited production task, young children will produce relative clauses when information processing demands are reduced. It was also hypothesised that, if children do not appreciate 'specificity' for definite reference, they should fail to obey felicity conditions: They will use relative clauses redundantly.

As for experiment 7, experiment 9 compares reference for specificity attributes that are salient (colour) with those that are not salient (character), but an elicited production task is used. This task is intended to reduce any problems of always having to use relative clauses or having to refer to two referents. Moreover, real objects are used so as to eliminate any potential clues to reference which were present in the pictures of experiment 7. Interest was in whether young children would only show a sensitivity for 'specificity' for definite references when the differences between potential referents are a salient feature, as reflected by their improved performance only in the 'colour' difference condition and their overuse of colour information even when

it is unnecessary for specific reference.

Experiment 10 is an attempt to tie together all the issues so far: 'non-specific/specific' reference and first/second mentions. Some investigators (Asher and Oden 1976) cite evidence to support that children's poor performance on referential communication tasks is due to difficulties the child has in making the actual comparisons between the referents. However, other investigators have produced the opposite findings (Robinson and Robinson 1978, Flavell 1968, 1970, Deutch and Pechman 1982, Roberts and Patterson 1983) that, even when a task is simplified so that young children can make the necessary comparisons, they do not spontaneously discriminate between referents.

Moreover, referential communication tasks have often been criticised as being unnatural. Investigators using observational methods have sometimes argued that the artificiality of the referential communication task may have led to spurious results which reflect children's reactions to experiment-generated procedures. Yet, it is difficult to test young children's understanding of ambiguity in everyday settings, because in natural situations, the context in which references are made often serve to disambiguate reference and it is difficult to dissociate deictic reference from other forms of definite reference.

Therefore, experiment 10 uses a 'model village' which aims to bring about active participation between speaker and listener and compares elicited discrimination with spontaneous discrimination. Pronoun

production is also looked at.

Whilst Karmiloff-Smith looked at children's production of pronouns, she used a quasi-experimental method: Children told narratives from picture-stories. Nevertheless, Karmiloff-Smith was basically interested in the way that children acquired the 'thematic subject strategy' rather than in their production of pronouns in ambiguous contexts. One naturalistic study (Van Hekken, Vergeer and Harris 1980) has concentrated on children's spontaneous production of ambiguous pronouns to invisible objects, (when reference cannot be disambiguated by deictic gestures), but there appears to be no experimental study to date which concentrates on children's production of referentially ambiguous pronouns.

However, due to the probability that children will only use pronouns when referents are in focus plus the fact that, if there is more than one referent in focus, pronominal reference can often be disambiguated by lexical and/or pragmatic factors, it is probably rare that one would find ambiguous pronominal reference in young children, irrespective of whether they appreciate the importance of 'specificity' for definite reference or not. Therefore, to show that this latter knowledge does apply to all forms of definite reference, including pronominal reference, experiment 11 explicitly manipulates lexical and pragmatic pronominal ambiguity in a context where more than one referent is in focus.

It was predicted that, as shown from their performance with definite articles, young children's use of

pronouns will fail to vary across those situations where a referent is disambiguated either pragmatically or lexically and where reference is ambiguous, because potential individuals are of the same-gender and neither are more plausible given the pragmatic context.

Whilst a number of issues are being looked at in this chapter, the basic questions being asked are:

1. Can young children's lack of sensitivity to 'specificity' for in/definite reference be applied to reference in discourse?
2. Is the development of 'specificity' appreciation reflected in definite anaphoric reference, where the definite predicate must include sufficient information to make the referent specific for the listener?
3. Are young children able to choose the appropriate definite expression when the necessary features of the referential context are made salient, in the same way as for immediate reference to concrete objects: Is the development of 'specificity' for in/definite reference one of spontaneously appreciating the necessary significance of specificity for appropriate and unambiguous reference?
4. Are any developmental differences a reflection of general in/definite expressions, or do they depend on the type of linguistic determiner used: Do the same patterns occur for pronominal reference as they do for reference using definite articles?

5.3 EXPERIMENT 6: Expressions used in a story-telling task as a function of Social and Comparative factors.

INTRODUCTION

As for experiment 4, the main two variables manipulated in this experiment were the 'familiarity' of the listener and the 'specificity' of referents. However, unlike experiment 4, this experiment involves a story-telling task and therefore, the speaker is required both to introduce and to refer anaphorically to referents in discourse.

In accordance with the experimental findings so far, it was hypothesised that young children will show a sensitivity to the listener's general 'familiarity' of referents, but they will fail to use indefinite expressions to introduce a non-specific item to a listener who is familiar with the referents. This result was predicted because story-telling is a discourse phenomenon, as opposed to one which requires reference to the immediate visible situation and as the referential context is less likely to be a salient feature in a discourse situation, young children are unlikely to focus on the relationship between the intended referent against the entire referential set.

Moreover, it was also predicted that, as young children do not seem to appreciate the significance of 'specificity' for definite reference, they will fail to mention a referent for the second time using exactly that information necessary for anaphoric reference.

METHOD

Subjects

32 children between 3½ and 5 years of age were used as subjects. There were 17 males and 15 females. None of the subjects had participated in any of the previous experiments. They were selected from the 'infants' of a primary school in Durham. The children were sorted into pairs of approximately equal ages.

Materials

4 new stories were constructed. Unlike the stories for experiments one and two, there was no main referent in each story. In the stories here, there were 2 animate characters per picture, and therefore, 2 pictures rather than 3 made up each story, as no subsidiary referent needed to be introduced. For each story, an additional picture was constructed which depicted the main animate and inanimate referents that were to appear in the story (referent picture).

A verbal translation of the 4 picture-stories is presented below:

- A. 1. A man is in a car waving. Another man is standing on the pavement/road waving back.
2. The man who was driving opens the door of the car and the other man walks over the road and gets into the car.

Referents: Inanimate/specific - road, car/door

Animate/non-specific - 2 men

B. 1. There are two ladies in a kitchen. One of them is sitting at a table eating something. The other lady is bringing a pot of tea over to the table.

2. The lady with the teapot then pours out some tea and the other lady picks up a cup from the table.

Referents: Inanimate/specific - pot of tea, table

Animate/non-specific - 2 ladies

C. 1. A girl is sitting on a chair with a teddy bear. Another girl is kneeling on the floor with a ball.

2. The girl sitting on the chair then throws the teddy bear up into the air and the girl with the ball drops it on the floor.

Referents: Inanimate/specific - teddy, ball

Animate/non-specific - 2 girls

D. 1. A girl runs towards a jump and another girl holds a flag.

2. The girl who is running jumps the pole and the other girl waves the flag.

Referents: Inanimate/specific - jump, flag

Animate/non-specific - 2 girls

As the 4 examples show, for each story, there were 2 animate/identical referents and 2 inanimate/specific referents that needed to be mentioned twice: They were presented in both pictures of the story. (Picture stories are shown in figures 6.1 - 6.4).

Tape recording equipment, projector, screen and a listener mask made up the remaining apparatus.

Design and Procedure

The experiment was conducted in the corner of the school's main hall. This was a place that was very familiar to all the children because it is usually used for assemblies, music lessons, P.E classes and dinner. The hall was not used for any of the usual purposes whilst the experiment was taking place.

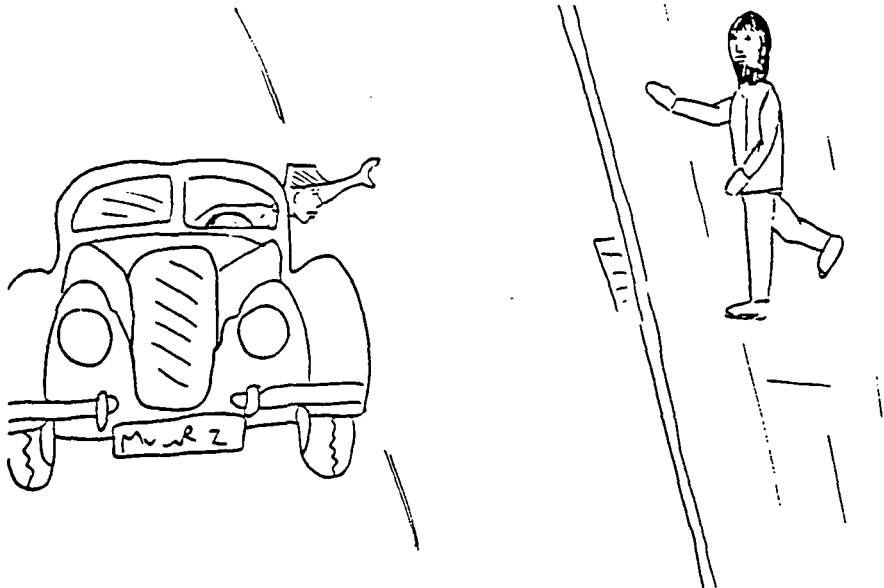
Altogether, 2 variables were manipulated in this experiment. These were:

1. FAMILIARITY: Listener ignorant (masked) or knowledgeable (unmasked).
2. SPECIFICITY: Referents specific (inanimate/single) or non-specific (Animate/double).

Each dyad was assigned to either the LI (listener ignorant) or the LK (listener knowledgeable) condition, but as each story contained inanimate (single) and animate (double) referents, each subject received both conditions of the 'specificity' factor. Each subject in a dyad received 2 of the 4 stories. The 2 stories presented to

FIGURE 6.1 PICTURE STORY A

First picture

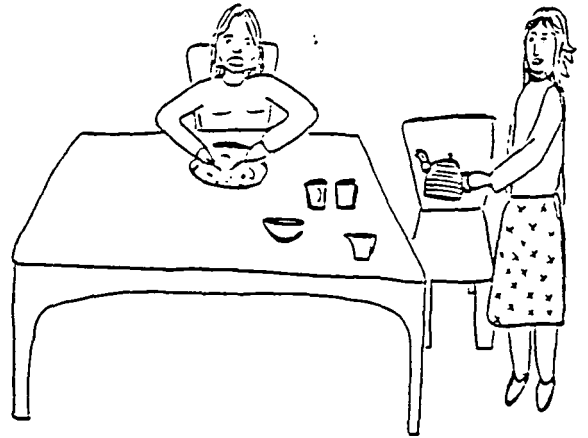


Second picture



FIGURE 6.2 PICTURE STORY B

First picture



Second picture

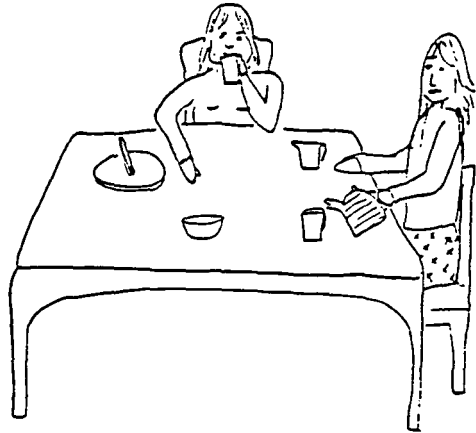
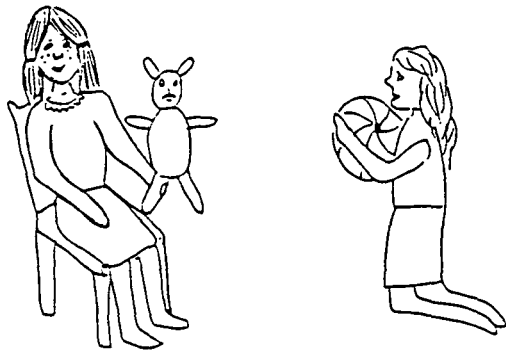


FIGURE 6.3 PICTURE STORY C

First
picture



Second
picture

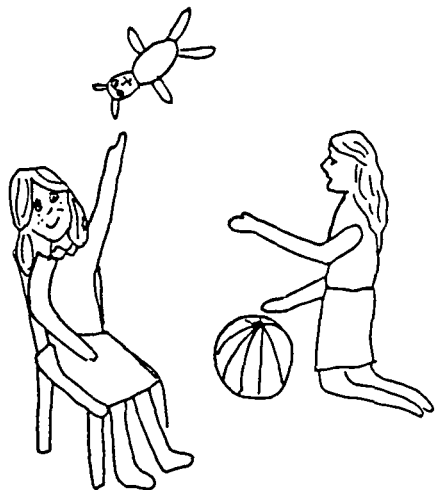
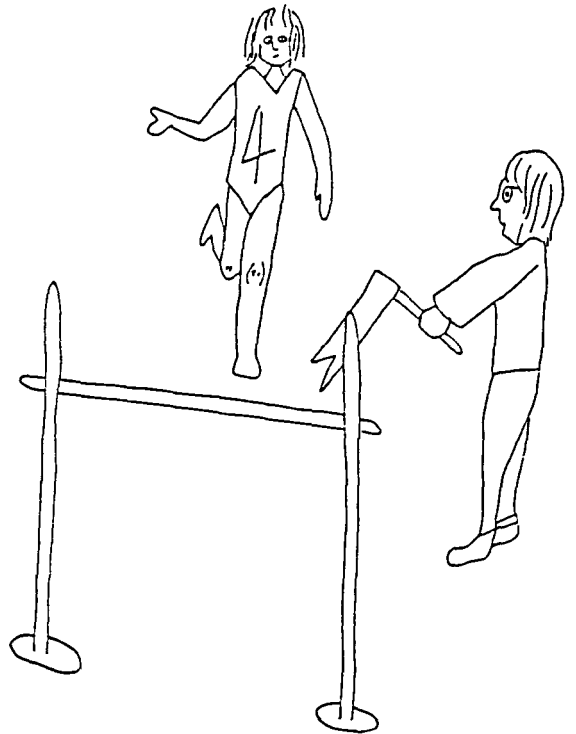
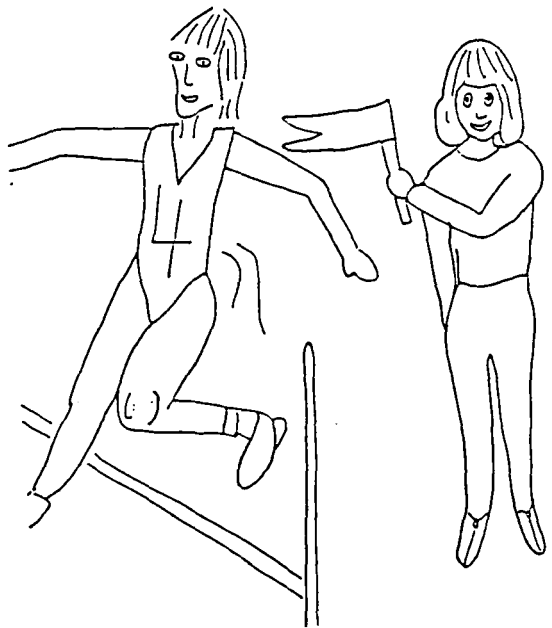


FIGURE 6.4 PICTURE STORY D

First
picture



Second
picture



each subject in a pair were randomly selected, as the order in which slides were put in the projector were selected from cards numbered 1 to 4 that were picked out of a hat by a member of the teaching staff.

Each dyad received the following instructions:

'One of you is going to be the listener and the other one is going to be the speaker. You will both take it in turns. I am going to show you some pictures on the screen. The first picture will show some people and objects and these are going to be in the story that I will show you next. The next picture will show the same people and objects but they will be doing something. This will be the first part of a story. The last picture will show the same people and objects again for the end of the story. I want you to look at the pictures carefully to see what is happening in them. Afterwards, I will show you the pictures again one at a time and you will then tell the story to (listener) as clearly as you can'.

For the LI condition, the listener had to wear the mask throughout the entire trial, that is, whilst the speaker looked at the pictures both initially and for the production of the story. For the LK condition, the listener put on the mask after the 'referent picture' had been shown, and therefore, even the knowledgeable listener never, at any point in the experiment, saw either the first or second story pictures. The reason why the LK listener only saw the 'referent picture' was so that the listener would be familiar with the critical referents and

therefore, animate/double referents would remain non-specific for them. Thus, specificity and familiarity were clearly differentiated in this experiment. Moreover, as the LK only saw referents without witnessing their role in the story, the type of definite expression used for anaphoric reference was just as important for knowledgeable listeners as for ignorant listeners. This is because, for double referents, it was necessary to use definite expressions with further presupposing information in order to reduce the subject of the immediate discourse to only one possible referent.

Each subject in a pair received a practice story. These were selected from experiment 1 and the only difference was that in this experiment a 'referent picture' was included. Speaker and listener swapped roles after each trial and the experiment was completed when each subject had narrated 2 of the 4 stories.

Speaker's utterances were recorded and subsequently transcribed for analysis.

Scoring

The expressions used by subjects were put into the following categories:

INANIMATES/SPECIFIC

First mention:

Indefinite: (excluding naming ie. 'It's a X')

Context Definite: When the context implied the existence
of the referent eg. 'A car - the road'

Possessive Pronoun: Her cup

Definite article, pronoun, demonstratives: It, this, the
Null / indeterminate

Second mention: Indefinite, definite, null/ indeterminate

ANIMATES/NON-SPECIFIC

First mention:

Indefinite Noun-phrases: 'A x....another x' OR '2 x's.
One (of the) x, the other x'.

Definite Noun-phrases: 'the x...the other x'

Definite-Single: failure to inform listener of quantity of
referents: The x...the x OR he...he..

Null / indeterminate

Second mention:

Indefinite: 'A x...another x'

DISCRIMINATIVE: Definite article with partitive, modifier
or relative clause eg 'The first x..the
second x' 'The x who has/was' etc..

NON-DISCRIMINATIVE; A. Definite article plus 'other':
'The x...the other x'

B. Definite article without modifier
or pronouns: x.....the x' or
'He.....he etc.

Null / indeterminate

RESULTS

A) MAIN RESULTS: FIRST MENTION

Animate / non-specific references

A total score of 2 (1 for each story) was given for each subject and there were no indeterminate scores. Thus, each group received a total of 32. Table 6.A presents percentages on first mention for non-specific references in each group.

GROUP	INDEFINITE	DEFINITE	DEFINITE
	NP	NP	SINGLE
LI	81.25	9.38	9.38
LK	21.88	75.0	3.13
MEANS:	51.57	42.19	6.26

TABLE 6.A: PERCENTAGE OF EXPRESSIONS ON FIRST MENTION FOR ANIMATES/NON-SPECIFIC ITEMS

As there were few definite-single scores in either group, definites are taken together for analysis. Clearly, there is a preference for indefinite as opposed to definite expressions in the LI group, whereas definites clearly exceed indefinites in the LK group. A Wilcoxon test showed a significant difference between indefinite and definite scores in both the LI group ($z=-2.55$, $p<0.05$, 2-tailed) and the LK group ($z=-2.20$, $p<0.05$, 2-tailed). Mann-Whitney

tests showed a significant effect of listener condition on indefinite scores ($U = 35.5$, $p < 0.0005$, 2-tailed, corrected for ties).

Inanimate / specific references

Each child received a score out of 4 (2 in each story) for inanimate referents. One child in the LK group gave an indeterminate response, so the total score was 64 for the LI group and 63 for the LK group. Percentages for each possible category of expression for LI and LK groups are shown in table 6.B.

GROUP	INDEFINITE	DEFINITE	CONTEXT DEFINITE	POSSESSIVE PRONOUN
LI	50.0	6.25	23.44	20.31
LK	6.35	60.32	22.22	11.11
MEANS	28.18	33.29	22.83	15.71

TABLE 6-B: PERCENTAGES OF EXPRESSIONS ON FIRST MENTION FOR INANIMATES/SPECIFIC ITEMS

Table 6-B shows that, for the LI condition, the most popular response was indefinite, 23% of expressions were context definites and 20% were possessive pronouns. For the LK condition, most expressions were definite, with also a fair number of context definites and some

possessive pronouns. Comparing indefinite expressions with definites (excluding context definites), analysis of variance revealed no significant difference between indefinite and definite scores ($F < 1$). However, there was an interaction between indefinite/definite scores and listener group ($F = 65.82$, $df = 1, 30$, $p < 0.001$).

Considering those scores which were appropriate given the familiarity of the listener: indefinites, possessive pronouns and context definite for the LI group and definites, context definites and possessive pronouns for the LK group, analysis of variance failed to find a significant difference between listener groups in appropriateness of reference ($F < 1$).

B) SUBSIDIARY RESULTS: SECOND MENTION

Animate / non-specific references

As there were only 3 indeterminate scores on second mention and these all involved null article responses, these scores have been included in the table below. Thus, there is a total of 32 scores in each group. Table 6.C presents the percentages for expressions on second mention for animate referents in each group.

GROUP	INDEF.	DISCRIMINATIVE	NON-DISCRIMINATIVE	DA+ 'OTHER' DA/PRO	NULL
LI	28.13	3.13	31.25	28.13	9.38
LK	0.0	0.0	40.62	59.38	0.0
MEANS:	14.07	1.57	35.94	43.76	4.69

TABLE 6.C: PERCENTAGES OF EXPRESSIONS ON SECOND MENTION FOR ANIMATES/NON-SPECIFIC ITEMS

The table shows that only 1.6% of scores were discriminative and this was due to one case in the LI group. Most scores involved single pronouns or definite articles without modifiers and almost as many scores involved definite articles plus 'other' modifiers.

Inanimate / specific references

Again, each child received a total of 4 scores on second mention for inanimate referents. The indeterminate responses have been included as these were all null article responses. Thus, again the total score for each group was 64. Table 6.D presents the percentages for each group on second mention.

GROUP	INDEFINITE	DEFINITE	NULL	POSS. PRONOUN
LI	6.25	81.25	7.81	4.69
LK	9.38	90.62	0.0	0.0
MEANS:	7.82	85.94	3.92	2.35

TABLE 6.D: PERCENTAGE OF EXPRESSIONS ON SECOND MENTION FOR INANIMATES/SPECIFIC ITEMS

Clearly, definite descriptions were preferred in both groups. These results were confirmed by analysis of variance, which showed a significant difference between indefinites and definites overall ($F=208.33$, $df=1,30$, $p<0.001$), but no interaction between indefinite/definite and listener group ($F=0.33$, $df=1,30$, NS).

Taking the scores for animate and inanimate references together, analysis of variance revealed an interaction between indefinite/definite and first/second mention ($F=97.68$, $df=1,30$, $p<0.001$) and a 3-way interaction between group, indefinite/definite and first/second mention ($F=28.44$, $df=1,30$, $p<0.001$).

DISCUSSION

The results supported those found in the previous experiments, that the young speaker's choice of indefinite or definite expression on first mention of a referent depended on the knowledge of the listener: 'Ignorant' listeners received predominantly indefinite expressions and 'knowledgeable' listener's received predominantly definite expressions on first mention. Moreover, virtually all definites for the ignorant listener were context definites, such as 'the jump' after referring to 'a high-jumper', 'the road' in the context of a car and 'the table' in the context of a kitchen, supporting the results of experiment 3 that young children use their pragmatic knowledge to define the knowledge of the listener. Possessive pronouns (for example, 'her teddy' 'her ball' 'his car') were also appropriate regardless of listener condition.

However, the 'familiarity' variable had such an overriding effect that subjects were insensitive to the 'specificity' variable: When referents were non-specific (animate referents), although subjects usually made it clear to the listener that there was more than one person of the same-gender and age in the story, the young child persisted with definite expressions for a knowledgeable listener, despite the fact that non-specific referents should be referred to with indefinite expressions.

Although virtually all second mentions were definite, many of these expressions for animate referents were ambiguous. In fact, only one subject referred to an

animate referent using a relative clause: 'The one with the teddy threw it up and the other one threw the ball on the floor'. Many second references did not even serve to inform the listener that they were referring to one of two possible individuals, as the 'other' modifier was not always included after the article.

Clearly then, young children were not using 'specificity' in deciding upon the appropriate expression necessary for unambiguous anaphoric reference.

5.4 EXPERIMENT 7: Using information for unambiguous reference I.

INTRODUCTION

In the following experiment, the salience of the information needed to make a referent specific for anaphoric reference was manipulated. The Referential Communication procedure was applied to a picture story-telling task in order to investigate this issue. Available referents were both animate and of the same gender, to see whether subjects would use complex noun phrases (for example, Definite article plus modifiers/partitives or a relative clause) where more presuppositional information is required to make a referent specific because referents cannot be differentiated on the basis of gender.

Secondly, referents could either be differentiated on the basis of salient physical attributes ('the big boy') OR on the basis of non-salient episodic events ('the boy who was playing with the train'). It was predicted that young children would produce many ambiguous anaphoric utterances as a result of their failure to use further presuppositional information. It was expected that performance would be improved in young children when salient physical attributes provided the necessary 'specificity' information, compared with when non-salient episodic information was necessary to make the referent specific for the listener. Thirdly, in accordance with Robinson and Robinson (1983), one would expect young

listeners to blame themselves rather than the speakers when incorrect pictures are selected as a result of ambiguous messages.

METHOD

Subjects

120 children and 16 parents selected from an 'infants' and a 'junior' school in Durham were used as subjects and none of them had been used in any of the previous experiments. The children were divided into 3 age groups (3 to 5 years [mean age: 4.5], 6 to 8 years [7.6] and 9 to 11 years [10.3]) and there were approximately an equal proportion of females to males in each of these groups. The parents were a random selection of the children's mothers. For all age groups (including parents), each subject was paired with another subject of approximately the same age.

Materials

A series of picture stories was constructed, (two pictures per story). The Experimental stories involved two animate referents of the same-gender and the Distractor stories involved only one animate referent. The purpose of the distractors was to see if children could refer anaphorically when only simple definite noun phrases were required (definite article plus noun, for example 'the boy') as well as to provide a control that any preference the listener had in choosing a picture was a function of the message, rather than due to learning the pattern that

they should select a picture depicting two persons, (which would enable them to pick the correct picture 50% of the time by chance alone, as opposed to the 33% probability of choosing correctly by chance when listeners can choose the picture of a single referent). Below is an example of a translation of an experimental story and a distractor story:

EXPERIMENTAL:

Picture 1 There are two clowns. The one in the green spotty suit is skipping and the one in the red stripy suit is playing with a ball.

Picture 2 The clown in red has turned round and the clown in green is scratching his head.

DISTRACTORS:

Picture 1 There is a boy holding a trumpet and a book.

Picture 2 Now he is holding a cricket bat and a ball.

6 experimental stories were constructed (three for each of the two experimental conditions. See design below) and 2 distractor stories were constructed. For each story, in addition to the picture depicting the second part of the story, two other selection pictures were constructed, each of which differed from picture 2 in certain aspects. For the experimental stories, the two additional selection pictures were SINGLE (one referent is performing both actions) and OPPOSITE (two referents with each one doing that action that the other referent is doing

in the original picture 2). For the distractor stories, the two additional selection pictures were DOUBLE/SAME OBJECTS (two animate referents, each playing with one of the objects/toys which are being played with by only one referent in the original picture 2) and DOUBLE/DIFFERENT OBJECTS (two animate referents, each playing with objects that are not depicted in the original picture 2). Providing two choice pictures with two animate referents as alternatives to the correct single picture in the distractor stories, provides a matched control against experimental stories, where there is always two pictures of two referents and one picture of one referent. Thus, referring again to the example stories shown above, the selection pictures would be as follows:

EXPERIMENTAL:

1. Picture 2 (A red stripy clown with his back facing us and a green spotty clown scratching his head).
2. SINGLE: One green and red (spotty and stripy) clown with his back turned and scratching his head.
3. OPPOSITE: A GREEN SPOTTY clown with his back facing us and a RED STRIPY clown scratching his head.

DISTRACTOR:

1. Picture 2 (A boy holding a cricket bat and a ball)
2. DOUBLE-SAME OBJECTS: two identical looking boys, one holding a cricket bat and the other one holding a ball.
3. DOUBLE-DIFFERENT OBJECTS: two identical looking boys, one banging a drum and the other one holding a flag.

Thus, for each story, there was a picture 1 and two versions of all three possibilities of picture 2 (one for the speaker and one for the listener).

(Picture stories for Physical and Episodic conditions, Distractor stories and selection pictures are shown in figures 7.1A - 7.8B).

In addition to the picture stories and selection pictures, a dividing screen with an attached microphone was used to separate the speaker and listener who were seated on opposite sides of a table. The screen was intended so that the listener would not be able to see the pictures in front of the speaker, but the height of the screen was low enough so that each subject could see the top of their partner's head and thus would remain constantly aware of the person on the opposite side of the screen. The experimenter sat behind a second table (used for when both partners have to get together for the post-selection questioning. See procedure) and this was placed at the side of the communication table. A tape recorder was also placed on this second table. The layout of the experiment is illustrated in figure 7.

FIGURE 7.1A PHYSICAL STORY A

Red and stripy clown + green and spotty clown.

①

First picture



②

Second picture

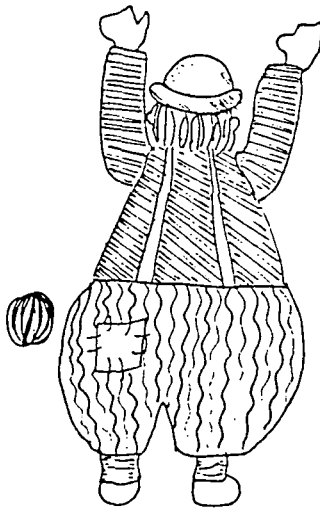


FIGURE 7.1B SELECTION PICTURES FOR PHYSICAL STORY A

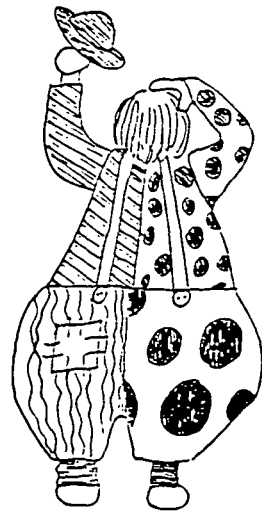
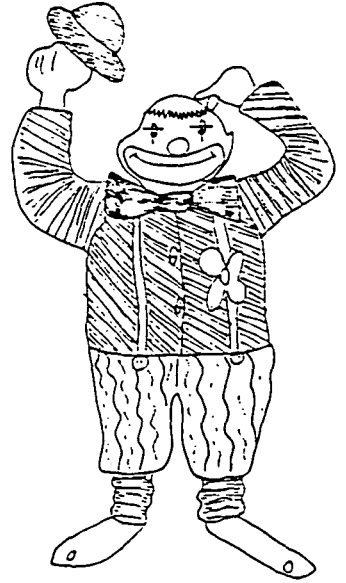


FIGURE 7.2A PHYSICAL STORY B

Tall boy with green trousers + small boy with red jumper.

①

First picture



②

Second picture



FIGURE 7.2B SELECTION PICTURES FOR PHYSICAL STORY B

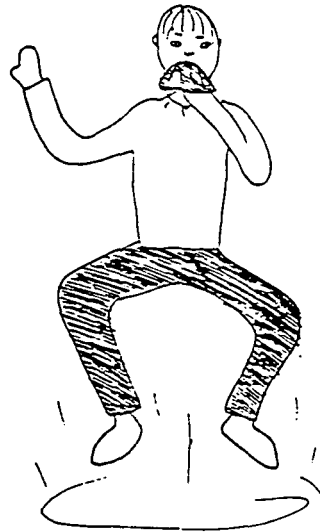
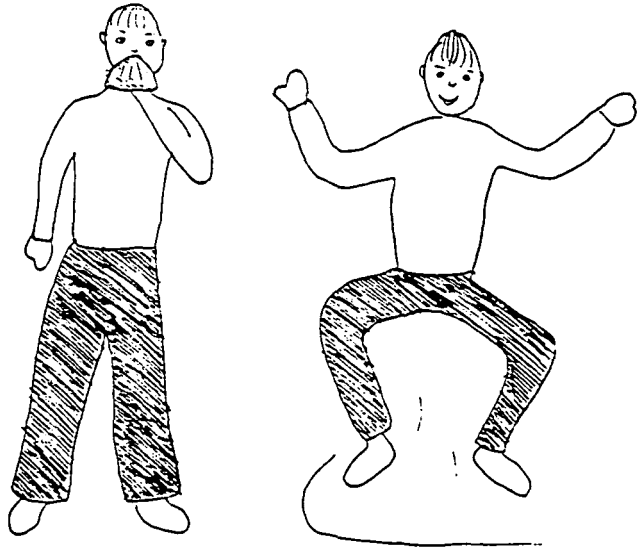


FIGURE 7.3A PHYSICAL STORY C

Fat man with green shirt + thin man with red shirt.

First picture



2

Second picture



FIGURE 7.3B SELECTION PICTURES FOR PHYSICAL STORY C

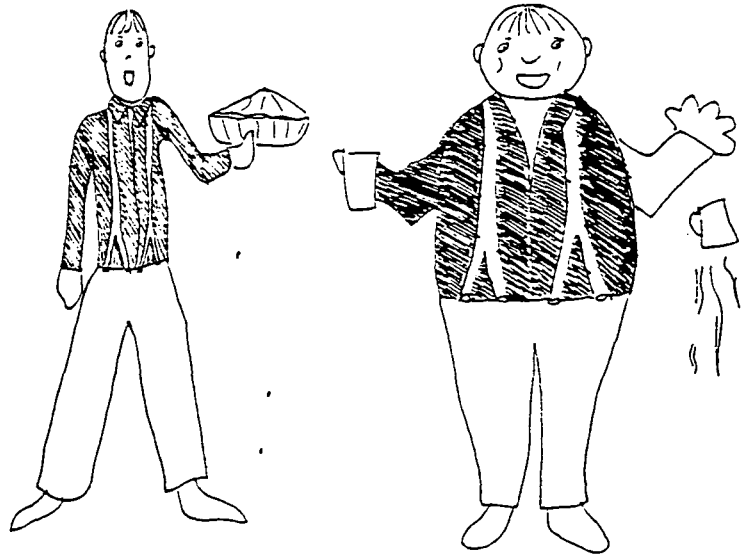


FIGURE 7.4A EPISODIC STORY A

Two green and stripy clowns.

First
picture



Second
picture

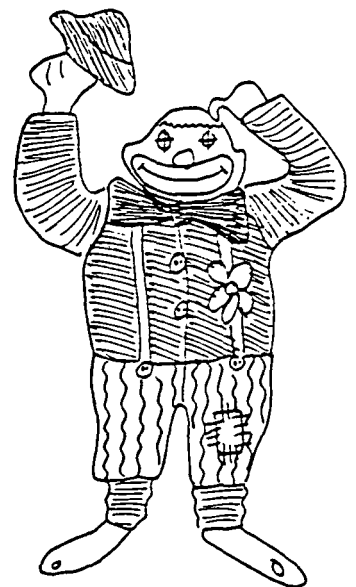


FIGURE 7.4B SELECTION PICTURES FOR EPISODIC STORY A

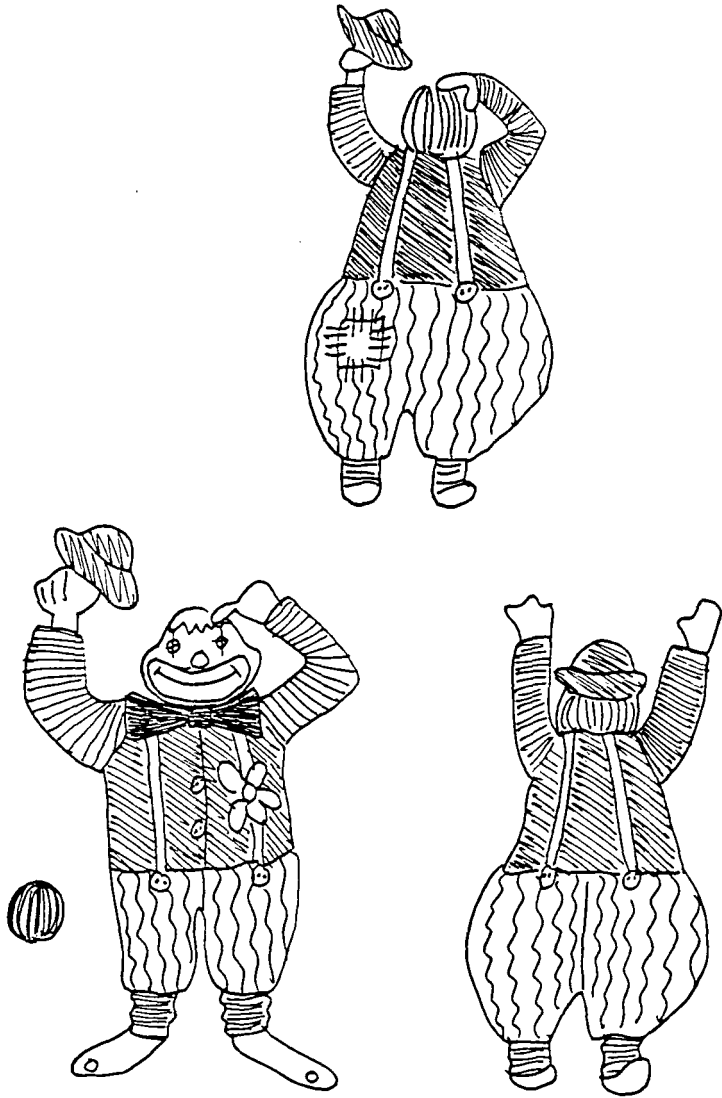


FIGURE 7.5A EPISODIC STORY B

Two same size boys with green jumpers.

First
picture



Second
picture

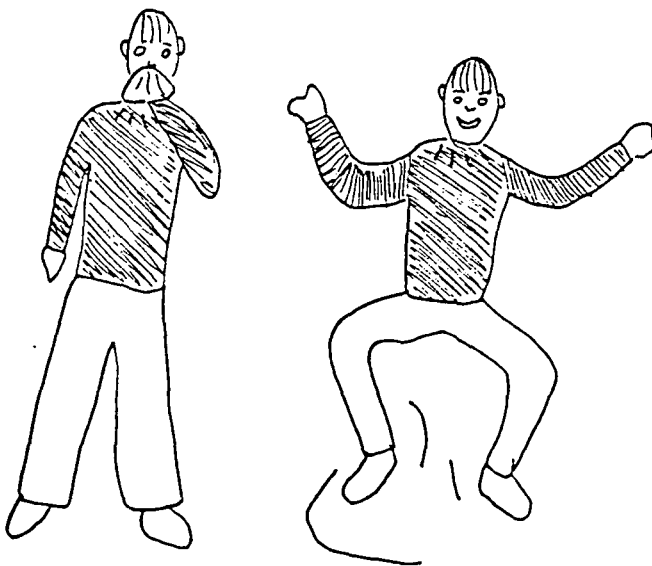


FIGURE 7.5B SELECTION PICTURES FOR EPISODIC STORY B

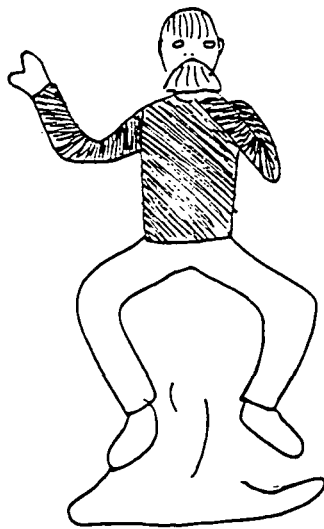
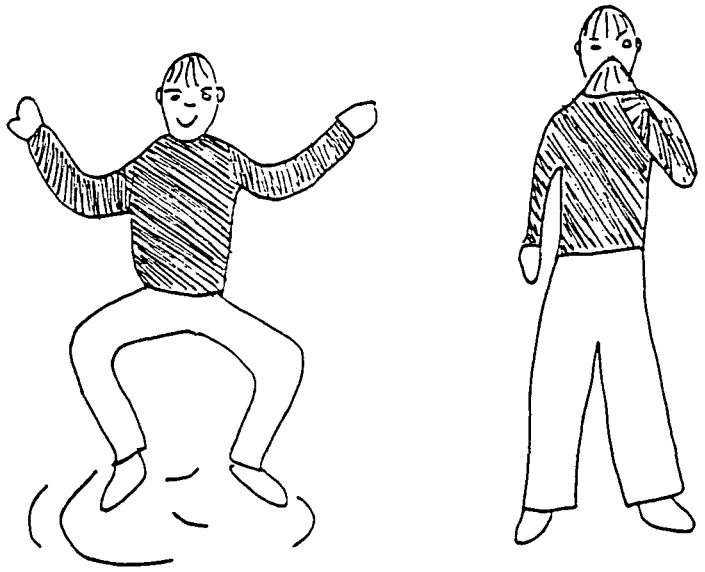


FIGURE 7.6A EPISODIC STORY C

Two same size men with red trousers.

First
picture



Second
picture

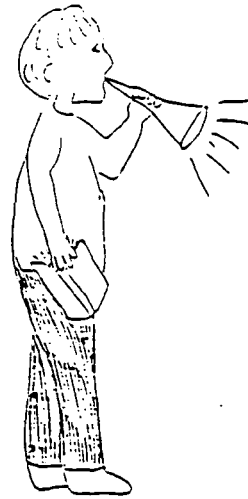


FIGURE 7.6B SELECTION PICTURES FOR EPISODIC STORY C



FIGURE 7.7A DISTRACTOR STORY A

First
picture



Second
picture

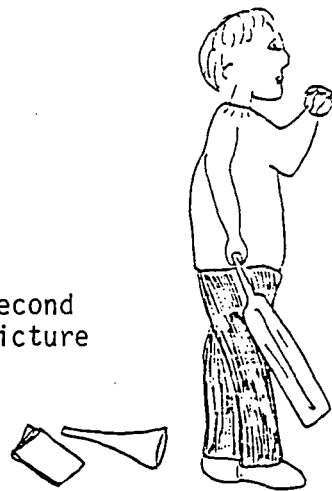
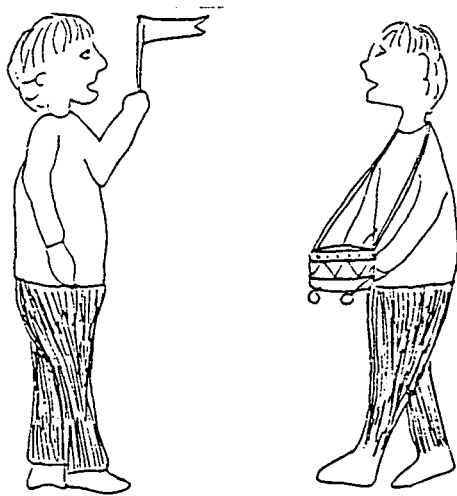


FIGURE 7.7B SELECTION PICTURES FOR DISTRACTOR STORY A



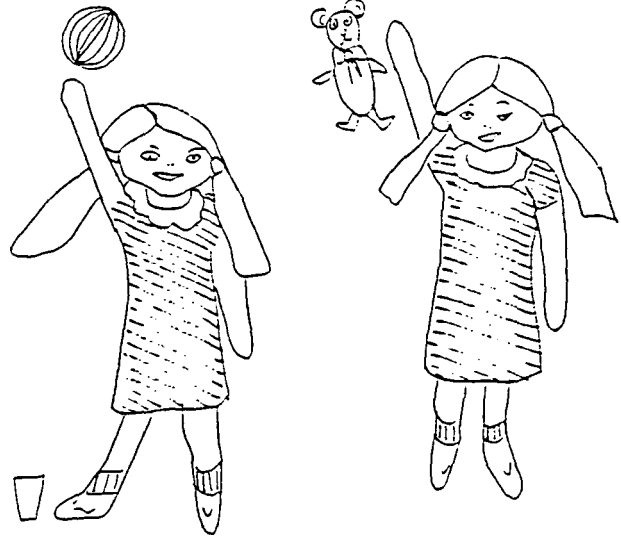
First
picture



Second
picture



FIGURE 7.8B SELECTION PICTURES FOR DISTRACTOR STORY B



EXPERIMENTER

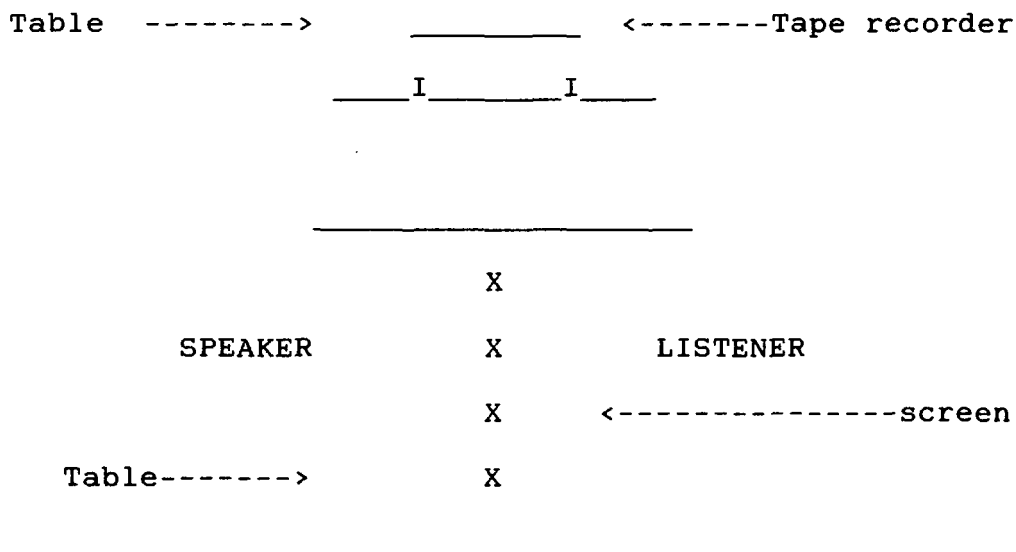


FIGURE 7: LAYOUT OF EXPERIMENT.

Design and Procedure

The experiment was carried out in familiar rooms for all subjects: the school library for the children over 5 years and the painting room for the youngest group. Each pair of subjects were assigned either to the 'Physical' (P) or the 'Episodic' (E) condition. The only difference between these groups was that, for the P group, the experimental stories involved referents which differed in at least two physical attributes (eg. a clown in a red stripy suit and a clown in a green spotty suit) and for the E group, the same stories involved referents which were physically identical in all aspects (2 clowns both in green stripy suits).

One member of a pair acted as speaker and the other member acted as listener throughout the entire experiment.

The experimental layout was as shown in figure 7 (see materials).

Phase 1: Instructions

The instructions given to subjects were as follows: "You are going to play a little game together. I want you to help each other. Will you help (listener) and will you help (speaker)? You are both going to have pictures to look at, but you must not show your pictures to each other.

Firstly, don't say anything yet. Just look at these pictures. First look at this one and then look at this one. Can you see that this picture is the first part of a story and this picture is the second part of the story? The people in the first picture are the same as the people in the second picture, except that they are doing different things."

Phase 2: Reference to picture one

The experimenter then gave a pair of pictures to the speaker and asked her to look at them carefully. The experimenter then turned the second picture over, so that the first picture was the only one in view to the speaker. The experimenter then said, "I want you (talking to speaker) to tell X (listener) what is happening in this picture. You must speak very clearly because X (listener) will not know what is in the picture will she?"

For this part of the procedure, interest was in the way that the speaker introduced referents to a listener who was unfamiliar with them. The speaker's utterances were

recorded for later transcription.

Phase 3: Reference to picture two

The experimenter took away picture one from the speaker, turned picture 2 back over and gave the speaker 2 additional pictures (the two alternative choice pictures). These were placed above the original picture, so that the speaker did not get confused as to which was the correct picture 2. The listener was also given the same three pictures as the speaker (3 selection pictures), but was not informed which of the pictures was the correct one. It was made clear to speakers that listeners had the same three pictures in front of them as they did and that listeners did not know which one was really the second part of the story. The experimenter said to the speaker, "I want you to tell (listener) what is happening in the correct picture 2, so that she knows about the second part of the story. I want you to speak very clearly because (listener) has got to choose the picture that you are talking about." The experimenter introduced the referents in picture one again (in case the speaker had not already done so appropriately and to ensure that any problems with picture 2 were neither due to memory failure nor due to the tendency of subjects to treat picture 2 as a new story as a result of the disturbance between reference to pictures 1 and 2). The experimenter then finished with "and then... Now tell (listener) what is happening next in picture 2". (Note: For the episodic condition, the two referents in the OPPOSITE selection picture could be distinguished from

the correct picture 2 either because the object that an individual was playing with in picture one was illustrated beside the appropriate referent on each of the selection pictures or because the left-right spatial positioning of the referents was the same in picture 2 as it was for picture 1).

The Experimenter told listeners that if they were not sure which of the 3 pictures the speaker was talking about, they could ask questions until they were sure which one they were going to choose. Throughout the experiment, the listeners were constantly reminded of their opportunity to intervene and thus were discouraged from taking a passive role in the experiment.

Phase 4: Post-selection questions

After the listener had selected a picture, the experimenter took the correct picture 2 from in front of the speaker, as well as the picture selected by the listener and placed both pictures in the middle of the second table. Both subjects were then asked to leave their seats and to walk around to the front of this second table, so that they were both facing the pictures. The experimenter then asked both subjects, "Is this picture just like this picture?" After a response had been made, the experimenter then asked the question "Is this picture NOT like this picture?" The latter question was to control for any response bias. (Notice the words same or different were not used, because it has been previously found (Karmiloff-Smith 1979, Emslie 1986) that young subjects interpret 'same' to mean 'same

kind' (similar) rather than 'same one' (identical).

When alike pictures were judged the same, the experimenter said, "Are you sure? Look carefully. Is this picture just like (not like) this picture? If, after further questioning, it was clear that neither subject had noticed the differences between the two pictures, the experimenter provided the appropriate feedback, "No they are not the same because..."

Finally, if the listener had not chosen the picture that was the correct picture 2, the experimenter then asked a question to see which subject the listener would attribute the mismatch of the pictures to (speaker or listener/herself): The experimenter asked the listener, "Why did you choose this picture and not this picture?"

The participants were first given practice stories so that they were familiar with the requirements of the task. For the experiment proper, each pair of subjects received 3 experimental stories (either P or E versions) and the 2 distractor stories. The order of presenting all 5 of these stories was randomly selected for each pair of subjects.

Speaker's utterances were recorded as well as any questions from the listener and the answers to post-selection questions.

Scoring

The scoring was as follows:

Speaker scores

EXPERIMENTALS

A) DISCRIMINATIVE REFERENCE (second mention only)

Indefinite: "A girl...a(nother) girl"

Definite/single: failure to mention that more than one referent "She..She" "The girl..the girl"

Definite/double: "The girl.....the other girl"

Definite/discriminative: Use the appropriate information to discriminate between referents, for example, "The spotty clown...the stripy clown" (P group) OR "The clown who was playing with the rope.....the other clown/the one who was playing with the ball....." (E group).

Indeterminate

B) INDEFINITE/DEFINITE REFERENCE (first and second mention):

Indefinite: "2 girls, one...one" "a girl....another girl"

Definite: All definite noun phrases or pronouns for reference to each animate referent

Mixed: A mixture of indefinite and definite expressions
"The girl....a(nother) girl"

Indeterminate

DISTRACTORS (first and second mention)

Indefinite: "A boy..."

Definite "He" "The boy"

Indeterminate

Listener scores

EXPERIMENTAL STORIES:

Correct (DC)

Opposite Double (DO)

Single (S)

DISTRACTORS:

Correct (single) (SC)

Double, correct objects (D)

Double, wrong objects (DX)

RESULTS

Unfortunately, due to a fault in some of the recording, some of the data were lost from the 9 to 11 year old group. Therefore, for this group, there are scores for 7 subjects in the P group and for 6 subjects in the E group. Thus, as the numbers of subjects per group is unequal, percentage scores are shown in the tables.

A) MAIN RESULTS

Speaker scores

DISCRIMINATIVE REFERENCE (experimental stories only).

For each definite reference used on second mention of the experimental stories, a score for discriminative reference in either of the 3 possible categories was given: Single (she...she, the girl...the girl), Double (the girl...the other girl) and Discriminative (PHYSICAL GROUP: the spotty clown. EPISODIC GROUP: the clown who was playing with the rope)

Physical group

Table 7-A shows the percentage scores in each discriminative category, for each age group and for the PHYSICAL group only.

AGE(Totals)	DISCRIMINATIVE	DOUBLE	SINGLE
3-5 (24)	37.9	25.0	37.5
6-8 (25)	100.0	0.0	0.0
9-11(19)	89.5	10.5	0.0
PAR.(11)	100.0	0.0	0.0
MEANS	81.8	8.9	9.4

TABLE 7-A: DISCRIMINATIVE ABILITY ON SECOND MENTION IN PHYSICAL GROUP.

It appears from the table that most expressions were discriminative, apart from in the youngest age group, where there were many expressions in both 'double' and 'single' categories and particularly in the 'single' category.

Episodic group

Table 7-B shows the percentage scores in each discriminative category, for each age group and for the EPISODIC group only.

AGE(Totals)	DISCRIMINATIVE	DOUBLE	SINGLE
3-5 (26)	7.7	65.4	26.9
6-8 (27)	51.9	44.4	3.7
9-11(18)	77.8	16.7	5.5
PAR. (6)	100.0	0.0	0.0
MEANS	59.4	31.6	9.0

TABLE 7-B: DISCRIMINATIVE ABILITY ON SECOND MENTION IN EPISODIC GROUP.

The oldest children's group and the parent group gave mainly discriminative expressions with very low percentages of non-discriminative scores. The youngest group gave mainly non-discriminative 'double' expressions, with a very low percentage of discriminative scores. 6 to 8 year olds gave more discriminative expressions than other categories, but almost 50% of their expressions were

non-discriminative, and these were virtually all from the 'double' category.

Therefore, the 'Episodic' group's discriminative scores were much poorer than for the 'Physical' group. However, the youngest group were particularly poor in the 'Episodic' group and the 6 to 8 year olds were poorer than older age groups for the 'Episodic' condition, whereas they were no poorer than older age groups for the 'Physical' condition. Analysis of variance did find an interaction between Physical/Episodic groups and discriminative/non-discriminative scores ($F=10.17$, $df=1,53$, $p<0.005$). However, there was no 3-way interaction between age, Physical/Episodic group and discriminative/non-discriminative scores ($F<1$) and this latter non-significant finding is due to the fact that all ages except the youngest group showed ceiling effects in the 'Physical' group and thus all groups were slightly poorer in the 'Episodic' group, despite the fact that the greatest difference between 'Physical' and 'Episodic' scores were for the youngest two groups.

Taking 'Physical' and 'Episodic' groups as a whole and combining 'double' and 'single' non-discriminative scores, analysis of variance found a significant difference between discriminative and non-discriminative scores ($F=19.15$, $df=1,53$, $p<0.001$) as well as an interaction between age group and discriminative/non-discriminative scores ($F=13.12$, $df=3,53$, $p<0.001$).

Improvements over trials

Cochran tests were carried out to see if discriminative scores improved between first to third trials. However, no significant differences across trials were found in either 'Physical' or 'Episodic' groups: (PHYSICAL: $Q = 1.17$ $df=2$, NS. EPISODIC: $Q = 0.50$, $df 2$, NS).

B) SUBSIDIARY RESULTS

Speaker scores

INDEFINITE/DEFINITE REFERENCE

With the exception of a few 'mixed' scores from the younger subjects on first mention, expressions were either 'indefinite' or 'definite'. Each subject received 5 scores (one for each experimental and distractor story). The percentages of scores for each age group, on first and second mention are presented in table 7-C.

	<u>FIRST MENTION</u>			<u>SECOND MENTION</u>	
<u>AGE(totals)</u>	<u>INDEF</u>	<u>DEF</u>	<u>MIXED</u>	<u>INDEF</u>	<u>DEF</u>
3-5 (100)	62.0	34.0	4.0	10.0	90.0
6-8 (100)	85.0	13.0	2.0	8.0	92.0
9-11 (65)	81.5	18.5	0.0	4.6	95.4
PAR. (40)	100.0	0.0	0.0	17.5	82.5
<u>MEANS</u>	<u>82.1</u>	<u>16.4</u>	<u>1.5</u>	<u>10.0</u>	<u>90.0</u>

TABLE 7-C: PERCENTAGE INDEFINITE/DEFINITE SCORES ON FIRST AND SECOND MENTION FOR EACH AGE GROUP.

The table shows that, on first mention, scores were predominantly indefinite, whereas they were predominantly definite on second mention. The only age differences were due to the youngest group producing more definites on first mention compared with the other groups. Analysis of variance showed that there was an overall significant difference between indefinite and definite scores, ($F=5.97$, $df=1,57$, $p<0.05$) due to the greater percentage of definite scores overall. There was also an interaction between first/second mention and indefinite/definite scores ($F=372.32$, $df=1,57$, $p<0.001$) and an interaction between age group and indefinite/definite scores ($F=9.2$, $df=3,57$, $p<0.001$). Finally, there was a significant 3-way interaction between age, first/second mention and indefinite/definite scores ($F=2.93$, $df=3,57$, $p<0.05$).

Listener scores

A. Choices of pictures

EXPERIMENTAL STORIES

Physical group

Table 7-D shows percentages for each type of picture choice in each age group, for the 'Physical' group.

AGE(totals)	CORRECT	INCORRECT	
	(DOUBLE)	(OPPOSITE)	(SINGLE)
3-5 (30)	66.7	23.3	10.0
6-8 (30)	90.0	10.0	0.0
9-11(21)	100.0	0.0	0.0
PAR.(12)	100.0	0.0	0.0
MEANS	89.2	8.3	2.5

TABLE 7-D: PERCENTAGE SELECTION SCORES IN 'PHYSICAL' GROUP.

Most choices were correct in all age groups. However, the youngest group still selected an incorrect picture for just over 30 percent of selection choices. These findings are in accordance with the speaker's discriminative scores, except that in the youngest group, children chose more correct than would be expected from speaker's discriminative scores and most of the youngest groups' choice errors were due to the referents being 'opposite' to the original, whereas most errors in speakers utterances were due to the failure to identify 2 individuals (single). Combining the scores for both types of incorrect choices, analysis of variance showed a significant difference between correct and incorrect choices overall ($F=124.2$, $df=1,27$, $p<0.001$) and there was also an interaction between age group and correct/incorrect choices for the 'physical' group ($F=5.0$, $df=3,27$, $p<0.01$).

Spearman correlations between the clarity of the speaker's message and the selection of the correct picture by the listener were impossible for the 2 oldest groups where the scores were nearly perfect for both speaker and listener. However, for the younger 2 groups, there was a significant correlation between speaker clarity and listener correctness for the 3 to 5 year olds ($r=-0.64$, $p<0.05$), but not for the 6 to 8 year olds ($r=0.42$, NS).

Episodic group

Table 7-E shows percentages for each type of picture choice in each age group, for the 'Episodic' group.

AGE(totals)	CORRECT	INCORRECT	
	(DOUBLE)	(OPPOSITE)	(SINGLE)
3-5 (30)	33.3	56.7	10.0
6-8 (30)	63.4	33.3	3.3
9-11(18)	88.8	11.2	0.0
PAR.(12)	100.0	0.0	0.0
MEANS	71.4	25.3	3.3

TABLE 7-E: PERCENTAGE SELECTION SCORES IN 'EPIDODIC' GROUP.

The table shows that most picture choices were correct in the episodic group. However, the number of correct choices was not as large as for the physical group as there is a larger developmental increase in correct choices in this group compared with the physical group. Most incorrect

choices were due to choosing two referents, but with opposite roles to the original. Combining the scores for both types of incorrect choices, analysis of variance showed a significant difference between correct and incorrect choices overall ($F=17.11$, $df=1,26$, $p<0.001$) and there was also an interaction between age group and correct/incorrect choices for the 'episodic' group ($F=8.22$, $df=3,26$, $p<0.001$).

Spearman correlations between the clarity of the speaker's message and the correctness of the listener were not possible for the oldest group, where both sets of scores were near perfect. However, for all other age groups, no significant correlation was found between speaker clarity and listener correctness (3-5: $r = 0.0$, 6-8: $r = 0.08$, 9-11: $r = 0.25$).

Furthermore, comparing 'physical' and 'episodic' groups, analysis of variance tests revealed a significant interaction between physical/episodic group and correct/incorrect choices ($F=8.2$, $df=1,53$, $p<0.01$), due to the greater proportion of correct choices overall in the 'physical' group. However, despite the greater developmental effect in the 'episodic' group, there was no 3-way interaction between age group, physical/episodic group and correct/incorrect choices ($F=1.47$, $df=3,53$, NS).

DISTRACTORS

In the distractor task all expressions (except possibly one ambiguous expression in the 9 to 11 year old group), referred to a single referent. The percentages of scores

in each age group for each type of picture selected by the listener are presented in table 7-F.

AGE(totals)	SINGLE		DOUBLE	
	(CORRECT)	(CORRECT OBJECTS)	(WRONG OBJECTS)	
3-5 (40)	65.0	2.5	32.5	
6-8 (40)	90.0	10.0	0.0	
9-11(26)	100.0	0.0	0.0	
PAR.(16)	100.0	0.0	0.0	
MEANS	88.8	8.1	3.1	

TABLE 7-F: PERCENTAGE DISTRACTOR SELECTION SCORES.

For all age groups, the correct 'single' picture was the most popular selection from the listener. Only the youngest two groups selected any other picture and this was extremely rare for 6 to 8 year olds. Just under a third of choices in the youngest group also involved choosing the wrong objects. Wilcoxon tests comparing correct verses incorrect listener choices showed no significant difference for the youngest group ($Z=-1.89$, NS), but there was a significant difference in correct verses incorrect choice scores from all older age groups for distractor stories (6-8: $Z=-3.52$, $p<0.0005$, 9-11: $Z=-3.18$, $p<0.005$, parents: $Z=-2.52$, $p<0.05$, 2-tailed).

B. Relationship between picture choice and speaker's discriminative ability.

Distractors are not dealt with in this section because the question of discriminative ability did not arise for the distractor stories.

Physical group

Table 7-G presents the percentage of each type of picture chosen by the listener as a function of DISCRIMINATIVE SCORES from the speaker in the 'physical' group. Each age group is shown separately.

	<u>DISCRIMINATIVE SCORES</u>		
	CORRECT	INCORRECT	
AGE(totals)	(DOUBLE)	(OPPOSITE)	(SINGLE)
3-5 (9)	66.7	33.3	0.0
6-8 (25)	92.0	8.0	0.0
9-11(17)	100.0	0.0	0.0
PAR.(11)	100.0	0.0	0.0

TABLE 7-G: PERCENTAGE LISTENER CHOICES AS A FUNCTION OF DISCRIMINATIVE SCORES FROM THE SPEAKER IN THE PHYSICAL GROUP.

When the speaker referred discriminatively, the majority of pictures chosen were correct, although a few choices from the younger subjects were of the 'opposite' type.

Although nearly all non-discriminative and single reference scores in this condition were from the youngest

group, 73% of these expressions still resulted in the correct choice of picture from the listener.

Episodic group

Table 7-H presents the percentage of each type of picture chosen by the listener as a function of DISCRIMINATIVE SCORES from the speaker in the 'episodic' group. Each age group is shown separately.

AGE(totals)	<u>DISCRIMINATIVE SCORES</u>		
	CORRECT	INCORRECT	
	(DOUBLE)	(OPPOSITE)	(SINGLE)
3-5 (2)	(1 CASE)	(1 CASE)	0.0
6-8 (14)	71.4	21.4	7.2
9-11(14)	92.9	7.1	0.0
PAR.(6)	100.0	0.0	0.0

TABLE 7-H: PERCENTAGE LISTENER CHOICES AS A FUNCTION OF DISCRIMINATIVE SCORES FROM THE SPEAKER IN THE EPISODIC GROUP.

In the 'episodic' group, discriminative expressions, which were virtually all confined to the older groups of subjects, usually lead to the correct picture being chosen by the listener. The 6 to 8 year olds made a few wrong choices, mainly of the 'opposite' type.

C. Locus of blame

Picture mismatch was scored as being either due to the listener's fault or due to the speaker's fault. A mismatch was due to the SPEAKER if the listener had chosen the 'opposite' picture when the speaker had made non-discriminative 'double' reference or when the listener chose any incorrect picture when the speaker produced 'single' reference. A mismatch was due to the LISTENER if the speaker had given discriminative reference and the listener chose incorrectly or when the speaker produced non-discriminative 'double' reference and the listener had chosen the 'single' referent picture. Table 7-I shows the 'locus of blame' scores from the listener, after an incorrect picture had been selected by the listener. The table shows the relationship between whose fault it actually was that the wrong picture was selected (speaker or listener) and the locus of blame given by the listener. Scores are expressed as percentages of all ascriptions of selection mismatch in each age group.

Obviously with such perfect discriminative scores from the older groups, we can only look at the performance of the 3 to 5 and 6 to 8 year old groups. For both the 3 to 5 year olds and 6 to 8 year olds, it seems that most pairs blamed the listener, irrespective of whether it was the speakers'

fault.

AGE	REAL FAULT	BLAME SPEAKER	BLAME LISTENER	NEITHER
3-5	SPEAKER (23)	43.5	52.2	4.3
	LISTENER(7)	26.6	71.4	0.0
6-8	SPEAKER (7)	14.3	71.4	14.3
	LISTENER(7)	42.9	57.1	0.0
9-10	SPEAKER (1)	-	-	-
	LISTENER(1)	-	-	-

TABLE 7-I: ASCRIPTION OF BLAME FROM THE LISTENER FOR PICTURE MISMATCHING.

DISCUSSION

Looking at the physical group first, the largest differences were between the 3-5 year olds and the 6-8 year olds. Only about a third of expressions were discriminative in the youngest group, owing to a failure to refer to two individuals, for example, "The clown is scratching his head..and he (referring to a different clown) is turned round," as well as due to a failure to inform the listener which of the referents performed each particular action, for example "The clown scratched his head and the other one turned round." This result supports the first hypothesis that younger subjects would not be able to use explanatory modifiers and relevant

discriminative information for making an intended referent specific for the listener.

Also as predicted, performance was poorer for the episodic group than for the physical group. Whilst all ages were poorer in the episodic group, the differences in scores between the youngest and older groups was far greater than in the physical group, owing to the almost complete failure of any subject in the youngest group to refer unambiguously in the episodic condition. Furthermore, the 6 to 8 year olds, who had referred unambiguously at ceiling level in the physical group, only produced just over 50 percent of unambiguous discriminative references in the episodic group. Thus, the developmental effects in the episodic condition were stronger than those in the physical condition. As there was no evidence of any improvement in discriminative performance over trials, it seems that the results do reflect the young child's failure to appreciate the need for using 'specificity' information for anaphoric reference, when differences between referents are not salient physical characteristics.

The listeners' choice of pictures for 'distractor' stories showed that subjects were not choosing a picture of two referents merely because they had learned an expected pattern of always choosing a picture with 2 referents: Almost all choices in distractor trials were for a single referent.

There does seem to be some relationship between the speakers' messages and the listeners' choice of pictures, suggesting that ambiguity did relate to outcome. However,

the listeners' comprehension needs to be more closely manipulated before any conclusions can be made.

As to the locus of blame, although I was unable to measure this for the older subjects, who scored practically at ceiling level on second mention, the youngest two groups (3 to 5 years and 6 to 8 years) showed a strong tendency to blame the listener, irrespective of whether it was the speaker's or the listener's fault. Therefore, even for 6-8 year olds who did give more unambiguous messages than the youngest group, there was still a strong tendency to blame listeners when the message was ambiguous, even though this meant blaming themselves for communication failure.

Finally, as further support for the previous findings, despite a slight improvement with age, even the youngest group were sensitive to the perceptual knowledge of the listener in their choice of indefinite or definite expression, when introducing a referent for the first time.

5.5 EXPERIMENT 8: Using Relative Clauses for Anaphoric reference.

INTRODUCTION

The following experiment is an attempt to replicate the findings of Hamburger and Crain (see general introduction) using a similar elicited-production task. It was hypothesised that the same children who failed to produce relative clauses in the previous experiment, would be more likely to produce relatives in this experiment.

It was also predicted that as young children are insensitive to the significance of 'specificity' for definite reference, the increase in relative production in the following experiment will be equally as effective in both necessary (more than one potential referent) and redundant (only one potential referent) versions, in which case it would seem that young children are not showing a sensitivity to felicity conditions in their use of relative clauses.

METHOD

Subjects

40 children who had acted as subjects in experiment 7 served again as subjects for this experiment. These were the 20 children in the 3 to 5 year old group and the 20 children in the 6 to 8 year old group. Thus, there were approximately an equal proportion of males to females in both age groups.

Materials

12 mini-stories with 12 matching questions were constructed similar to those used by Hamburger and Crain. Each story could be presented in either the 'same' referents or 'different' referents version. An example of one of the story-question pairs, shown in both its 'same' and 'different' versions, is presented below:

Same

There were two boys at a party. One of the boys sat on his own and the other boy started to dance. Who went home first do you think? (Expected response: The one who was sitting on his own).

Different

There was a girl and a boy at a party. The boy sat on his own and the girl started to dance. Who went home first do you think? (Expected response: The boy).

6 of the stories required answers referring to the most recently mentioned individual and the other 6 stories required answers referring to the first mentioned individual. This was to control for any effects of memory or subject/object biases in subjects.

(See appendix H for mini-stories)

Recording equipment was also required for the experiment.

Design and procedure

The variable manipulated in this experiment was whether stories were felicitous (same referents) or unfelicitous (different referents) with regards to relative production. The former were felicitous because there was more than one potential referent that a simple definite noun phrase (the boy) could refer to, whereas the latter were unfelicitous because a simple definite noun phrase would be sufficient to refer unambiguously to one referent and therefore, relative information would be redundant.

The experiment was presented as a repeated measures task, with subjects receiving all 12 stories, 6 in the 'same' version and 6 in the 'different' version. As for which stories were presented in each condition, for each subject, the first six numbers between 1 to 12 that were picked from a hat by a member of the teaching staff were presented in the 'same' version and remaining stories were in the 'different' version.

Irrespective of same/different referent conditions, story-question pairs were read to each subject in consecutive order from numbers 1 to 12. The general procedure was as follows:

Each subject was seen individually in the school painting room (3 to 5 years) or the main hall (6 to 8 years). Both these locations were familiar to subjects within each respective age range. The subjects were told that they must listen carefully while the experimenter read them some short stories. They were told that, after each story, they would be given a question to answer. They were

first given a few practices with some non-experimental stories to ensure that they understood the experimental requirements.

If a subject unambiguously referred to one referent, but their choice of referent was the reverse of the expected response or, if it was unclear whether subjects were referring to information conveyed in the story or not, (such as, 'the good girl', when the context made no such explicit reference to good or bad behaviour'), the subject was asked to justify her/his response: The experimenter would say for example, "Why do you choose the good girl". These requests for justification were a check that subjects were really listening to the context sentences and not merely repeating information verbatim, in which case one would not be able to conclude that they were referring anaphorically.

The tape recorder, which was positioned out of the subjects' field of vision, behind the experimenter, was switched on just before the experimenter read the first story and was switched off after the subject had answered the question to story 12. Subjects' answers were later transcribed for analysis.

Scoring

Answers to questions were scored into the following categories:

- A. Indefinite or definite

- B. Form of reference:

RELATIVE WHO/THAT: "The boy who was sitting on his own"
RELATIVE WHAT: "The boy what was sitting on his own"
DESCRIPTIVE MODIFIER: "The weak boy"
'OTHER' MODIFIER: "The other one"
SIMPLE DEFINITE: "The boy" "He did"
DEMONSTRATIVE: "That one"
NO REFERENCE TO CONTEXT: "The one who was best" (plus inability to justify with reference to context).

RESULTS

Firstly, although it is not presented in the table, all subjects used definite noun phrases when referring back to one of the individuals, except in one instance where a 4 year old subject used an indefinite article in the different-gender condition. As this response also included a relative, 'A girl that was working hard', it has been included as a relative that/who score amongst the definite relative scores.

'SAME' REFERENCES

The scores for each age group have been classed as either unambiguous (relative (that/who), relative (what), modifier), ambiguous (other modifier, demonstrative, simple noun phrase) or no reference to context.

Table 8-A presents totals of each category of expression in the 'same' condition, for both age groups. As there were 20 subjects per group and each subject received a score of 6 in this condition, scores are out of 120 for each age group.

AGE	UNAMBIGUOUS		AMBIGUOUS		NO CONTEXT		
	RELATIVE 'Who/That'	MODIF. 'What'	'OTHER'	SIMPLE NP	DEMONS		
3-5:	62	8	34	2	8	0	6
6-8:	89	2	29	0	0	0	0
TOT:	151	10	63	2	8	0	6

TABLE 8-A: EXPRESSIONS USED TO ANSWER QUESTIONS IN THE 'SAME' CONDITION

The table shows that, in both age groups, most expressions were unambiguous and of the relative (who, that) type, with a fair proportion of unambiguous reference using modifiers "The weak one". It can also be seen from the table that, for a few answers, subjects used 'what' instead of 'who/that' when using relative clauses. Very few expressions were ambiguous or failed to refer to context and these were all from the youngest group. In fact, only two subjects in the youngest group were unable to make any form of unambiguous discrimination. One of these subjects always gave answers which bore no relationship to the preceding information given by the experimenter (no context) and gave some form of irrelevant personal modifier, for example, 'the nice one' and 'the best one',

without justification. The other either said 'the other one' or used a simple noun phrase, 'the boy', that is, they ambiguously referred to the referent they had in mind without verbally discriminating between the referents. A third subject gave ambiguous answers two thirds of the time, but for the rest of his answers he made some form of unambiguous discrimination.

Some of the unambiguous scores above include cases where subjects chose a referent which was unexpected given the context, or used information that had not been explicitly conveyed in the passage. These were accepted as unambiguous because the experimenter was satisfied that subjects were able to justify their choices with respect to the context.

Analysis of variance showed a significant difference between unambiguous and ambiguous scores ($F=206.98$, $df=1,38$, $p<0.001$), but there was no interaction between age group and unambiguous/ambiguous reference ($F=3.06$, $df=1,38$, NS).

'DIFFERENT' REFERENCES

The scores for each age group have been classed into each category as either unambiguous (relative (that/who), relative (what), modifier, other modifier, simple noun-phrase), ambiguous (demonstrative) or no reference to context. Note that for the 'different' condition, the only form of reference which was ambiguous was a demonstrative, which failed to indicate the gender of the referent.

Table 8-B presents totals of each category of expression in the 'different' condition, for both age groups. Again scores are out of 120 for each age group.

	UNAMBIGUOUS				AMBIGUOUS NO CONTEXT		
AGE	RELATIVE	MODIF	'OTHER'	SIMPLE	DEMONS		
	'Who/That'	'What'		NP			
3-5:	59	10	32	1	6	7	5
6-8:	84	0	34	0	2	0	0
TOT:	143	10	66	1	8	7	5

TABLE 8-B: EXPRESSIONS USED TO ANSWER QUESTIONS IN THE 'DIFFERENT' CONDITION

The table shows that again, for both age groups, most expressions involved relatives of the that/who type and both groups produced a fair number of modifiers. There were only 7 ambiguous expressions and 5 unrelated to context answers, both from the youngest group. Interestingly, those subjects in the youngest group who used simple noun phrases and 'other' modifiers were the same subjects who used these expressions for the 'same' condition.

Analysis of variance showed a significant difference

between unambiguous and ambiguous scores ($F=433.84$, $df=1,38$, $p<0.001$), but there was no interaction between age group and unambiguous/ambiguous reference ($F=3.21$, $df=1,38$, NS).

Table 8-C presents each category of expressions for both 'same' and 'different' referent conditions taking the scores of both age groups together.

total out of 240

	WHO	WHAT	MODIF	'OTHER'	SIMPLE NP	DEMON.	NO CONTEXT	SUM
SG:	151	10	63	2	8	0	6	= 240
DG:	143	10	66	1	8	7	5	= 240

TABLE 8-C: TOTAL 'SAME' AND 'DIFFERENT' SCORES IN EACH CATEGORY.

The table shows that, despite the small proportion of demonstratives in the different condition compared with the complete lack of demonstratives in the same condition, the proportions of scores in each category are approximately the same for both same and different conditions: Most scores are for Who/that relatives, about a quarter include modifiers and there is a low proportion of scores in other categories in both same and different conditions.

Using analysis of variance to compare each type of expression between conditions, no significant differences between 'same' and 'different' conditions were found for either form of expression (F= who: 1.76, what: 0, modifier: 0.34, 'other': 0.2, simple NP: 0, demonstrative: 2.01, no context: 1.0, df=1,39, all non-significant).

DISCUSSION

Firstly, the prediction that the same young children who failed to use relatives in experiment 7, would produce a fair number of relatives in this experiment was supported: Despite the slight improvement in relative production between the infants group and the 6 to 8 year olds, even the youngest group produced over 86 percent of unambiguous references when there were two potential referents (both the same/same gender), and excluding those appropriate expressions which involved the use of modifiers instead of relatives: 'The weak one', the majority of expressions still involved the use of relatives (58.3 %) in the youngest group.

Moreover, the prediction that the higher proportion of relatives from younger subjects in this experiment would be no different between felicitous (same) and unfelicitous (different) conditions was also supported. Thus, whilst the task used in this experiment served to elicit relatives from young children, the fact that they were used just as often when they were redundant provides no evidence regarding the young child's knowledge that relatives must be used when they supply the only available

information for making referents specific.

Finally, the results support the previous findings that, when referents are familiar to the listener (experimenter), definite reference is clearly preferred to indefinite reference.

5.6 EXPERIMENT 9: Using information for unambiguous reference II

INTRODUCTION.

The following experiment uses an elicited-production procedure to control for processing difficulties. There are four conditions each of which vary according to the type of information required to make reference unambiguous.

The main hypotheses were that, even when subjects are encouraged to minimise redundancy and when reference cannot be disambiguated on the basis of pragmatic context, young children were expected to show poorer discriminative ability than older children, because previous evidence suggests that they are less able to appreciate the necessity of using the appropriate linguistic information (for example 'colour', 'character' or 'gender' information), to make a referent specific for a listener. Also, young children were expected to perform better when referential differences were based on colour (similar condition), than when differences were based on character information (identicals condition): Young children were expected to use colour information because it was salient to them. By contrast, older children and adults were expected to use both colour and character information, but only when each piece of information was necessary: not when information was irrelevant or redundant.

A third prediction was that younger children will perform very well in different-gender conditions (two referents of different gender), where 'gender' information

alone is sufficient to discriminate between the referents. However, young children were expected to show no differences in their choice of reference between same-gender conditions (two referents of the same gender) and different-gender conditions, as shown by their failure to make the distinction between discriminative and redundant information. However, with development, it was hypothesised that subjects would be more likely to use information discriminatively across all 4 conditions (same/different gender, similars/identicals), as this is the performance one would expect if subjects can appreciate the necessity of using definite descriptions unambiguously.

METHOD

Subjects

96 children and 32 adults were subjects. The children were selected from a primary and secondary school in Durham and were divided into 3 age groups (3 to 5 years [mean age: 4.4], 6 to 8 years [6.9] and 9 to 11 years [10.3]), with approximately half male and half female subjects in each group. The adults were a random selection of mothers of the children. Each subject was randomly paired with another subject of the same age, so that there were 16 pairs in each subject group.

Materials

A selection of dolls, varying across the two dimensions of GENDER (male or female) and COLOUR (red or green), served as the referents in this experiment.

A mask for the listener, a stop-watch for timing the subjects' responses, as well as tape recording equipment and a microphone for recording the expressions used by the speaker, made up the remaining experimental apparatus.

Design and procedure

A. DESIGN

Two main variables were manipulated in this experiment.

These were:

a) GENDER: SAME OR DIFFERENT.

For this variable, either 2 same-gender or 2 different-gender dolls were selected.

b) DISTINGUISHING FEATURES: SIMILARS OR IDENTICALS.

For this variable, dolls were either physically SIMILAR (different colours - same character) or they were physically IDENTICAL (same colour - different characters).

Half the subject pairs in each age group were assigned to the similars group and half were assigned to the identicals group. Each pair received 12 experimental trials (6 trials in each of the same-gender and different-gender conditions). All trials were presented in a random order.

For each trial, the type of character information used was selected randomly from the 8 character types that were used in the pretest (The pretest is described below).

Also, with regards to which of the two dolls was moved by the experimenter (see phase 3), this was chosen randomly

on each trial.

B. PRETEST

When introducing the character information, the experimenter used one of 8 possible character references. The action sequence that E subsequently carried out on one of the dolls depended on the character information that was given. This was to avoid any possibility that one doll may be more plausible than another in the identicals condition (for example, if the character information given to one of the dolls was that, s/he was 'happy', then the child may have been biased to assume that this doll would be the one who 'put his/her hands in the air'). Thus, the character references were matched with an action sequence and a pretest preceded the experiment in order to ensure that pragmatic bias had been eliminated. The pretest involved six 4 year old subjects and six 8 year old subjects. For example, I asked 'Here are two girls. This one likes ice-cream. Which girl do you think would go for a walk? This one or this one?' The 8 character types, the 4 action sequences and the number of subjects referring to the most popular doll are presented in table 9 for each paired association:-

CHARACTER TYPE	ACTION SEQUENCE	MOST POP. DOLL CHOSEN		
		4 YRS	8 YRS	MEAN(max.6)
A1 feel happy	goes for a walk	6	6	6.0
A2 feel sad	,, ,, ,, ,,	5	6	5.5
B1 is an artist	puts hands in air	6	5	5.5
B2 is a scientist	,, ,, ,, ,,	6	5	5.5
C1 likes icecream	sits on the floor	5	7	6.0
C2 likes chocolate	,, ,, ,, ,,	6	5	5.5
D1 dislikes horses	jumps up and down	6	7	6.5
D2 dislikes cats	,, ,, ,, ,,	5	6	5.5

TABLE 9: PRETEST SCORES

Neither age group in the pretest showed any strong preference for one doll over the other. Overall the range of the mean of the highest scores was from 5.5 to 6.5.

Notice that the character types were made up of both positive and negative attributes. This was to control for any bias towards choosing the referent with a positive character. Also, for the identicals group, in order to refer to the doll who was not given the character information, a subject would have had to use negatives (for example, 'the one who doesn't like ice-cream' or would have had to assume that the doll possessed the opposite attribute (for example, 'the one who is sad', after being informed that the other doll is happy).

C. PROCEDURE

Phase 1: Familiarisation Period

Each pair of subjects was taken into the experimental room, (school reading room) and they were seated side by side facing the experimenter, who was seated on the other side of the table. The experimenter delivered the following instructions to both subjects:- 'We are going to play a game. I want you to pretend that these dolls are real people with thoughts and feelings, just like you and me. Can you see that these are boys and these are girls? Sometimes they are green and sometimes they are red.'

'We will also pretend that they can move just like us too.' (E shows subjects that the dolls can move their arms, legs and can walk etc.) 'Do you want to have a go?'

The children were given a chance to practise moving the dolls before the experimenter continued with the instructions.

One subject was then assigned the role of speaker and the other subject was assigned the role of listener.

Phase 2: Introduction of two referents and character information

The experimenter picked out two of the dolls and covered all those which were not selected. Then the experimenter introduced character information. This involved either attributing the same character to both dolls (similar), for example, E may have said 'Pretend that THEY BOTH like ice-cream,' or attributing a character to only one of the dolls (identicals), for example,

'Pretend that THIS ONE likes ice-cream.'

Phase 3: Referring to one of the two dolls

The listener was blindfolded and the speaker was told to watch the experimenter perform some action with one of the two dolls that had been selected. For example, E made one of the dolls go for a walk. The speaker was then given a question (in this case 'Who went for a walk?') and had to tell the listener what happened as quickly as possible after the 'ready' signal. The important point to emphasise to the child was that they must tell the listener as quickly as possible. The stop-watch was held up for the speaking subject to see. The emphasis on a time constraint was to prevent subjects from producing more information than was necessary.

Phase 4: Listener performance

After the speaker had answered the question, the scarf was again removed from the listener. (The speaker wore the mask at this point, so as to prevent any non-verbal cues being given to the listener). The listener was told to perform the action on that toy which the speaker had just told them about. The two dolls were then returned to the initial array and the procedure was run again for a total of 12 trials.

Before the experiment proper, 2 practice trials were given, to ensure that performance was not an artifact of subjects misunderstanding the experimental requirements.

D. SCORING

The tape recordings were subsequently transcribed and the speaker was scored for the information supplied when referring to one referent (gender, colour, character, spatial).

For each trial, a score of '0' was given if the speaker made reference within 10 seconds, a score of '1' was given for a duration of between 10 to 20 seconds and a score of '2' represented a duration of over 20 seconds for making reference.

Finally, the listener received a score of 1 for selecting the referent that was moved by the experimenter.

RESULTS

1. OVERALL DISCRIMINATIVE ABILITY

Table 9-A presents mean scores in each subject group, for discriminative reference. A discriminative score was given for the use of 'colour' in the 'similar' group and for the use of 'character' information in the 'identicals' group. In addition, for the different-gender condition (DG), a discriminative score was ALSO given when a subject used 'gender' information.

In the ID group, there were some cases of subjects using 'spatial' information ('Spatial' discrimination was when a subject said something like 'the one on the left' or 'the one nearest to me/you'). These references have also been included as discriminative scores for the ID

group. Nevertheless, whilst about 25 percent of ID discriminative scores were due to spatial discriminations, for each age group, most discriminative scores in the ID group were due to 'character' references.

AGE GROUP	SIMILARS	IDENTICALS	OVERALL MEAN (out of 6)
<u>SAME-GENDER</u>			
3-5:	5.0	0.9	3.0
6-8:	5.8	3.3	4.6
9-11:	6.0	5.3	5.7
Parents:	6.0	5.3	5.7
<u>MEANS:</u>	5.7	3.9	4.8
<u>DIFFERENT-GENDER</u>			
3-5:	6.0	5.1	5.6
6-8:	6.0	4.9	5.5
9-11:	6.0	6.0	6.0
Parents:	6.0	6.0	6.0
<u>MEANS:</u>	6.0	5.5	5.8

TABLE 9-A: DISCRIMINATIVE SCORES IN EACH CONDITION.

Analysis of variance showed that there was a significant effect of age on discriminative ability ($F=13.29$, $df=3,56$, $p<0.001$). There was also an interaction between age and SIM / ID groups ($F=6.64$, $df=3,56$, $p<0.001$).

There was also a significant difference between SG

and DG conditions ($F=26.06$, $df=1,56$, $p<0.001$), an interaction between age and SG / DG conditions ($F=6.35$, $df=3,56$, $p<0.001$) and an interaction between SG / DG and SIM / ID groups ($F=13.14$, $df= 1,56$, $p<0.001$).

Inspection of table 9-A suggests that these results are due to the following effects: Performance improved with age in the ID group but not in the SIM group (possibly due to a ceiling effect). Similarly, performance improved with age in the SG but not DG condition (again possibly due to a ceiling effect). The only condition with any clear developmental effects was the Identicals/same-gender condition.

2. 'SIMILAR' AND 'IDENTICAL' GROUPS.

a) Positive and negative relatives.

Table 9-B presents the total discriminative 'character' references (Identicals group), comparing scores for positive relatives ('the one who likes ice-cream') with negative relative scores ('the one who doesn't like ice-cream'), in both SG and DG conditions, in each subject group. (The scores for SG and DG conditions have been pooled because, there was no significant interaction between SG/DG and positive/negative relative scores ($F=1.67$, $df=1,28$, NS)).

(Total out of 48) DISCRIMINATIVE CHARACTER SCORES

<u>AGE GROUP</u>	<u>POSITIVE</u>	<u>NEGATIVE</u>	<u>DIFFERENCE</u>
3-5:	6	5	1
6-8:	18	17	1
9-11:	29	26	3
Parents:	30	29	1

TABLE 9-B: DISCRIMINATIVE CHARACTER SCORES: POSITIVE AND NEGATIVE RELATIVES.

It is clear from the table that there was very little difference in discriminative character references between positive and negative scores for all age groups. Analysis of variance found no significant difference between positive and negative scores ($f=0.54$, $df=1,28$, NS) and no interaction between positive/negative scores and age group ($F=0.06$, $df=3,28$, NS).

b) Time to respond

Firstly, table 9-C shows the mean duration of responding for each age group. An overall mean across all trials was recorded for each subject, so that each subject had a score out of two (0=0 to 10 seconds, 1=10 to 20 seconds and 2=over 20 seconds). The actual times were not recorded because interest was in the way that timing reduced redundancy: Time scores were subsidiary results, taken as a consequence of the redundancy control.

AGE	(Mean scores out of 2)
3-5:	0.8
6-8:	0.2
9-11:	0
Parent:	0

TABLE 9-C. TIME TO MAKE A RESPONSE.

It is clear that there was a decrease in time with age. However, the main differences were between the youngest group against all other groups. Mean scores show that 3 to 5 year olds were closer on average to 1 (10 to 20 seconds), whereas other groups were all closer on average to a 0 score (under 10 seconds).

c) Irrelevant Information

IRRELEVANT information includes any information that did not serve to make reference discriminative. This information may have been used either as insufficient information, when the speaker failed to provide discriminative reference, or as redundant information, when the speaker was at the same time providing discriminative information.

Table 9-D presents the mean scores for discriminative and irrelevant uses of 'colour' and 'character' information in both SIM and ID groups. Spatial discriminations have been excluded. (Again, the scores for

SG and DG conditions have been pooled because the only differences between these two conditions were very slight and confined to the parent group. SG and DG comparisons will be looked at later in section 3).

<hr/>			
AGE GROUP		(Mean scores-total out of 6)	
<u>SIMILARS</u>	COLOUR	CHARACTER	OVERALL
	(Discriminative)	(Irrelevant)	MEANS
3-5:	4.8	0.3	2.6
6-8:	5.9	0.4	3.2
9-11:	6.0	0.2	3.1
Parents:	4.9	0	2.5
<hr/>			
MEANS:	5.4	0.2	
<hr/>			
<u>IDENTICALS</u>	CHARACTER	COLOUR	OVERALL
	(Discriminative)	(Irrelevant)	MEANS
3-5:	0.7	3.3	2.0
6-8:	2.2	4.0	3.1
9-11:	3.5	1.9	2.7
Parents	3.7	1.3	2.5
<hr/>			
MEANS:	2.5	2.7	
<hr/>			

TABLE 9-D: DISCRIMINATIVE AND IRRELEVANT COLOUR AND CHARACTER SCORES (OUT OF 6).

For the SIM group, table 9-D shows that all subject groups used a high proportion of discriminative 'colour', whereas all ages used a very low proportion of irrelevant 'character' information. The scores for the ID group

showed an increase in use of discriminative 'character' information with age, whereas there was a decrease in use of irrelevant 'colour' information with age. Overall, analysis of variance revealed a significant difference between discriminative and irrelevant scores ($F=99.23$, $df=1,56$, $p<0.001$). Moreover, there was a significant interaction between SIM/ID groups and discriminative/irrelevant scores ($F=107.97$, $df=1,56$, $p<0.001$): Discriminative scores exceeded irrelevant scores in the SIM group only. Also, there was an interaction between age group and discriminative/irrelevant scores ($F=7.08$, $df=3,56$, $p<0.001$): Discriminative scores increased with age, while irrelevant scores decreased.

Finally, there was a significant 3-way interaction between age group, SIM/ID groups and discriminative/irrelevant scores ($F=5.42$, $df=3,56$, $p<0.005$). Inspection of table 9-D suggests that this effect is due to the fact that, the 2-way interaction between SIM/ID and discriminative/irrelevant scores is confined to the youngest age groups: For the two oldest age groups, discriminative scores exceeded irrelevant scores in the ID condition, as well as in the SIM condition.

3) 'SAME-GENDER' AND 'DIFFERENT-GENDER' COMPARISONS.

Comparing 'same-gender' and 'different-gender' conditions, the only effects were for the parent group, who used more colour information in the SIM group and more character information in the ID group when referents were of the

same-gender compared with when they were of different gender. Analysis of variance showed that whilst there was no significant difference between SG and DG overall ($F=3.18$, $df=1,56$, NS) and there was no significant interaction between SG /DG and SIM/ID conditions in the use of discriminative 'colour' and 'character' ($F=0.1$, $df=1,56$, NS), there was a significant interaction between age and SG/DG effects ($F=3.25$, $df=3,56$, $p<0.05$).

Contrastive expressions in different-gender condition

Table 9-E presents percentage CONTRASTIVE scores in Different-Gender conditions. A CONTRASTIVE score is a discriminative score that involved either 'gender' without other discriminative information ('colour' or 'character/spatial'), or 'colour', 'character/spatial' discriminations that did not involve mentioning 'gender'.

* SCORES	SIMILARS		IDENTICALS	
	Gender only	Colour only	Gender only	Char/spatial only
3-5:	22.9	50.0	68.3	12.2
6-8:	2.1	22.9	22.5	37.5
9-11:	0.0	31.3	33.3	45.8
Parents	37.5	0.0	52.1	12.5
MEANS	15.6	26.1 (20.9)	44.1	27.0 (35.6)

TABLE 9-E: CONTRASTIVE EXPRESSIONS (DIFFERENT GENDER).

Firstly, the table shows that there were more CONTRASTIVE expressions in the ID group than in the SIM group. Moreover, for the SIM group, all age groups except parents preferred 'colour only' to 'gender only'. In the ID group, for the youngest and oldest subjects, 'gender only' was preferred to 'character only', whereas the reverse is the case for the intermediate groups.

Thus, for both conditions, only the parent group chose more 'gender only' expressions than both 'colour only' and 'character only' expressions.

4). LISTENER SCORES

Table 9-F shows overall means of correct listener scores (choosing the doll that E had moved). The table also shows the relationship between the listeners' correctness of choice and speakers' discriminative scores.

(Mean scores out of 12)

AGE	SPEAKER DISCRIMINATIVE	LISTENER CORRECT	DIFFERENCE SCORES
3-5:	8.5	9.8	1.3
6-8:	9.9	10.7	0.8
9-11:	11.6	11.1	-0.6
Parent:	11.6	12.0	0.4

TABLE 9-F: LISTENER SELECTION OF REFERENTS

There is a developmental increase in choosing the

correct dolls. Analysis of variance showed a significant age effect on listener correctness ($F=6.24$, $df=3,60$, $p<0.001$) and the NEWMAN-KEULS analysis revealed that the differences were between the youngest group against the two oldest groups, as well as due to the difference between 6 to 8 year olds and parents. Moreover, Pearson correlations between discriminativeness and listener scores were significant ($r=0.61$, $p<0.000$). Taking each group separately, whilst no analysis is needed to show that there was a correlation for the parents, there was a significant speaker-listener correlation for the youngest two groups (3 to 5: $r=0.63$, $p<0.005$, 6 to 8: $r=0.74$, $p<0.001$), but there was no significant correlation between speaker and listener performance for the 9 to 11 year olds (9 to 11: $r=-0.18$, NS). This is due to the fact that 9 to 11 year olds showed a slight negative trend: Discriminative scores slightly exceeded listener correctness.

DISCUSSION

As predicted, there was an effect of age group on discriminative scores. 6 to 8 year olds produced better scores, but performance was not at a very high level until subjects were 9 years of age or more. However, supporting the second hypothesis, the youngest subjects were very good in the 'similar' group, but they were very poor at using character information in the 'identicals' group, where there was a highly significant age effect of discriminative ability. There was no evidence to suggest that poverty of character

references was a reflection of difficulties in using negative relatives ('the girl who doesn't like ice-cream'): There was no difference in performance between constructions which required positive relatives and those which required negative relatives.

Therefore, it is clear that parents were discriminative. 9 to 11 year olds were also discriminative, but they did sometimes rely on the salience of the colour characteristic. 6 to 8 year olds showed some discriminative use of 'character', but they still relied heavily on the salience of colour and used many irrelevant 'colour' modifiers. However, the youngest group were clearly non-discriminative, using 'colour' when it was both relevant and irrelevant and failing to use 'character' information in any condition. The performance of the younger subjects in the 'identicals' group was greatly facilitated in the 'different-gender' condition, supporting the prediction that, when young children could rely on gender information to discriminate between referents, their performance would be improved.

It was also predicted that younger subjects would be using gender information non-reflectively, whereas older subjects would be using gender discriminatively: using 'gender' only in the 'different-gender' conditions. There seems to be little evidence for the discriminative use of 'gender' at any age. However, only the parents' scores in the different-gender condition were more likely to involve 'gender' only than either 'colour' or 'character' only and the parents were the only group to use less 'colour' and

'character' information to discriminate between referents in the 'different-gender' condition, as compared with the 'same-gender' condition. These findings suggest that parents were better at appreciating the discriminative use of 'gender' in 'different-gender' conditions than were any of the children.

Finally, there was a strong relationship between the speakers' discriminative scores and the listeners' correctness scores, suggesting that discriminative information did actually serve to affect the performance of the listener.

5.7 EXPERIMENT 10: Model Village - Identification and anaphoric reference.

INTRODUCTION

The following experiment used a task which did not require the use of character information and which aimed to minimise production restrictions (for example, subject-embedded relatives). A pretest was given to see whether any problems in using relevant discriminating information is a reflection of problems with comparison activity per se or is due to problems in making spontaneous comparisons for disambiguating reference.

Moreover, a more age-appropriate task was used compared with those used in previous experiments and referential communication tasks. A goal-directed game was used to encourage motivation from both the speaker and the listener. This game involved the experimenter moving items onto a model village and the speaker was required to inform the listener about the items which were moved, the object of the game being to match the boards of both speaker and listener.

The aim of the experiment was to look at the way that children would discriminate between a wide variety of expressions: indefinite expressions, definite noun phrases, definite noun phrases requiring extra presuppositional information (modifiers, relatives) and pronouns, in a variety of contexts. In fact, the experiment looks at expressions used for a range of referents (animate/inanimate, male/female), which differ

in their degree of specificity in the 'world' (singletons, similars and identicals) and which are referred to over a range of discourse contexts (referents in focus, plausibility and presuppositional context (number of referents)).

The hypotheses were, firstly, that even when subjects can make the comparisons between referents when it is specifically requested of them (pretest), there will be a developmental increase in the ability to spontaneously use extra presuppositional information when this information is needed to make reference unambiguous (for example, physical discriminations 'the small cow', quantity discriminations 'the second car' and spatial or temporal discriminations, 'the lady who is by the tree in the field' or 'the lady that you picked up last'). Moreover, it is predicted that the latter form of discrimination (spatio-temporal) will be most difficult for younger subjects, because this information is not salient to them. However, the size and quantity of referents is predicted to be used more because it is salient information to subjects anyway, regardless of the use of this information for unambiguous reference.

Secondly, contexts in which pronouns are preferred will be compared with those for which definite articles are preferred, to see if there are any developmental differences in the type of definite expressions used in different contexts. One speculative prediction was that young children will be more likely to use pronouns ambiguously (when the referent is not in focus or when

there is no pragmatic information available) than will older subjects.

Finally, younger subjects are expected to prefer inappropriate definite expressions for the first placed non-specific items when the listener is familiar with the referents, especially as there are only 2 objects in a set of non-specific items and as it is the experimenter who is moving the items, which therefore, fails to make 'non-specificity' salient for the subjects.

METHOD

Subjects

128 subjects carried out the experiment. The children were from Durham schools and the adults were a mixed selection of the children's parents and undergraduate students from Durham University. Subjects were divided into 4 age groups, with 32 subjects per group: 3 to 5, [mean age: 4.6] 6 to 8 [7.4], 9 to 11 [10.5] and adults (parents and undergraduates). Subjects were matched into pairs of approximately equal ages, so there were 16 pairs of subjects in each age group. Finally, there were approximately an equal proportion of males to females per group.

Materials

2 boards of a model village were constructed, each of which depicted a FIELD, ROAD, POND, BRIDGE and BUS STOP. Two sets of items were provided. Each set contained 2 identical trees, 2 identical women, 2 identical cars, 2

identical sheep, 1 big cow, 1 small cow and a single man. A screen to separate the listener and speaker was also required and the screen had a microphone attached to it. (See figure 10.1 for materials and layout). Recording equipment made up the remaining experimental apparatus.

Design and procedure

One subject in each pair acted as speaker and the partner acted as the listener. All pairs received exactly the same pattern for reference. There were approximately 45 references to be made altogether, with one or two objects referred to on each move made by the experimenter. An interpretation of each reference for moved items is shown below. The numbers indicate the objects referred to.

A.A /one of the tree(s) in the middle of the field.

1 2

B.A woman / one of the women next to the tree.

3 4

C.The second / other tree in the park near the pond.

5 6

D.The woman walking up to the other tree. She jumps up and

7 8 9

[down

E.She walks over the bridge.

10 11

F.A / one of the sheep stands on the bridge.

12 13

G.The small cow stands by the tree in the park.

14

15

H.The second/other sheep stands at the bus-stop.

16

17

I.The big cow stands at the tree in the field.

18

19

J.The small cow walks over the bridge.

20

21

K.The second/other woman stands by the first tree/the tree

22

23

that is in the middle of the field. She steps backwards.

24

L.The man stands next to this woman/her/the woman in the

25

26

[field.

M.A / one of the cars is going along the road.

27

28

N.The car / it stops by the bridge.

29

30

O.The lady on the bridge walks up to the car.

31

32

P.The man walks up to the tree in the park. He steps

33

34

35

[backwards.

Q.He/ the man walks up to the lady by the tree

36

37

R.The big cow walks over the bridge.

38

39

S.The man walks over it/the bridge.

40

41

T. The second car drives up to the car by the bridge.

42

43

U. The big cow stands behind the second car.

44

45

For subjects who mentioned the referents which were predicted (as shown above), each reference belonged to one of the categories of expressions presented below, where each category and sub-category is shown with its associated referents.

The categories of expressions are presented below, together with the references required for each category:

NON-SPECIFIC (ONE OF TWO): 1,3,12,27. eg. 'A woman/one of the women.'

PRONOUN: 'He' 'It'

a. only referent in focus: 7,10,25,36,41

b. 2 in focus but pragmatic differences: 13,29,41

DEFINITE ARTICLE + NOUN: 'The man'

a. not in focus: 2,6,11,17,21,28,30,33,40.

b. not in focus but unique in world: 25,33 (eg. 'the man')

c. 2 in focus and no pragmatic differences: 4,32

DEFINITE ARTICLE + SIZE + NOUN: 'The big cow'

not in focus: 14,18,20,38,44.

DEFINITE ARTICLE + QUANTITY + NOUN: 'The second lady'

a. not in focus: 5,16,22,42.

b. 2 in focus and no pragmatic differences: 8,45.

DEFINITE ARTICLE + SPATIAL RELATIVE: 'The lady by the tree
in the field'

not in focus: 15,19,23,31,34,37,43

Also, 9, 10, 26, 36, 40 - referents in high focus (brought
into focus by being the main referent in two previous
actions).

The procedure was as follows:

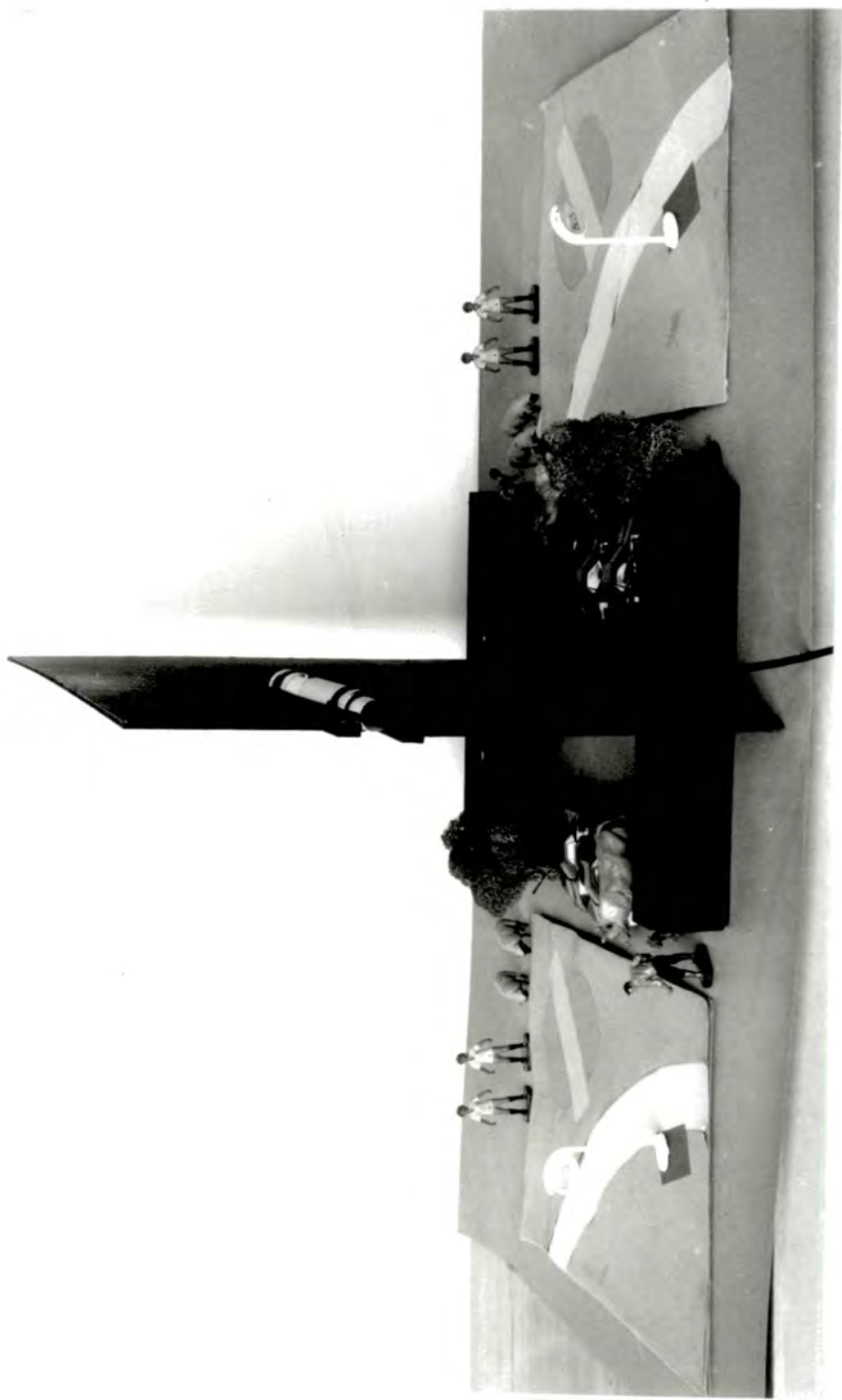
Phase 1: Introduction to materials

Both subjects of a pair were given identical 'model
villages' and objects. At this stage, there was no screen
between the children, so that both subjects could see that
the materials in front of their partner were the same as
their own.

Phase 2: Pretest

Both subjects, in turn, were then asked about the
illustrations on the board and about the items. For
example, E asked 'Can you tell me what this is?'. Are
there any more like this? Are there any differences
between them. These questions were asked several times for
all objects, to see whether subjects were able to compare
objects and mention differences and matches between items.

FIGURE 10.1 EXPERIMENTAL MATERIALS AND LAYOUT



Then, looking at each partner in turn, the experimenter placed two of the items in different positions on the board and asked the subjects about the differences between the objects. This was to see if they could compare the objects on the board and appreciate differences in spatial positioning.

Phase 3: Instructions

If the subjects were able to make these comparisons, one subject was then assigned the speaker (S) role and the other subject was assigned the role of the listener (L). However, subjects were not told of their role. Instead, the following instructions were given:

'I am going to put this screen in the middle of you both, so that you cannot see what (S's name) is doing and you cannot see what (L's name) is doing. I am going to move objects onto your (to S) board, so L will not be able to see what I am doing. Will she? Why? So I want you to help L by telling her what I am doing with the objects. Then I want L to move the objects, so that L moves the same things that I do. You (to S) must tell L to move exactly the same objects as I do, because at the end, you will both get a sweet if this board looks the same as this board. So you must help each other. S, you must tell L clearly the objects that I am moving and L, you must tell S if you do not understand what S is saying: You must ask S to tell you more if you think that she hasn't told you enough'.

A sweet was given to both partners at the end of the

experiment irrespective of the actual outcome, but the use of a reward was intended to encourage both partners to produce identical models and to eliminate any competitive element in the game, such as the speaker trying to cheat and fool the listener.

Phase 4: Moving objects and reference

Several practice trials were carried out, (moving single objects onto the board and then checking the behaviour of both partners). This was to ensure that the subjects clearly understood what the task required of them. All objects were then removed from both the speaker's and the listener's boards. The experimenter proceeded to move individual objects as indicated by the standard pattern, in order to observe the speakers' choices of expressions for the 42 critical references. After each movement of an item, the speaker was expected to refer to either one or two referents (The object moved and usually the location of the moved object). The experimenter regularly reminded listeners to pay attention to speakers and to ask questions if they were unsure or if they thought that speakers were unclear.

Scoring

The type of information given by the speaker for each referent was recorded and subsequently transcribed. Also, any questions and verbal intervention from the listener was recorded.

Finally, the position of each object in the

listener's final display was recorded two dimensionally on a sheet representing the model village.

1. SPEAKERS EXPRESSIONS.

For each of the critical references, the speakers' initial expressions (before any listener intervention) were put into the following categories:

Indefinite

Definite article

Pronoun

Definite + size modifier.

Definite + 'other'/'second' modifier.

Definite + spatial information.

Null or non-referring expressions.

2. LISTENERS COMMENTS.

These were assigned to the following and a note was made of the reference number:

General comments: 'I don't know' 'I can't do that'

General questions: 'Tell me more?'

Process questions: 'Which one?' 'Do you mean X or Y?'

Perceptual comments: 'Speak up' 'I can't hear you'.

RESULTS

A) MAIN RESULTS

1. Pretest

No subject had to be excluded from the experiment as a result of performance on the pretest: All subjects, in all

age groups, were able to make the physical, quantity and spatial comparisons between items, when it was explicitly requested of them.

2. Discriminative reference

A discriminative score was given when a subject appropriately used definite expressions plus extra information in order to differentiate between referents: 'the small cow', 'the first car', the girl in the field'. As each subject's narrative varied slightly, discriminative scores were each converted to a percentage out of those contexts where discrimination was necessary to prevent ambiguity. The mean percentage discriminative scores for each age group are shown in table 10-A.

<u>AGE GROUP</u>	<u>MEAN PERCENTAGE DISCRIMINATIVE SCORES</u>
3-5	39.3
6-8	68.9
9-11	79.9
ADULT	93.4

TABLE 10-A: MEAN PERCENTAGE DISCRIMINATIVE SCORES.

It is clear from the table that discriminative ability increased with age, with the adults achieving almost maximum performance and with the 3 to 5 year olds showing less than 50 per cent appropriate discriminations. Analysis of variance confirmed that there was a significant

effect of age group on percentage discriminative scores ($F=50.91$, $df=3,60$, $p<0.0000$) and the Newman-Keuls analysis revealed that the differences lay between all age groups.

Table 10-B presents the mean percentage scores out of those contexts where physical (the small cow), quantity (the first car) and spatio-temporal (the girl in the field OR the girl you just moved) discrimination was necessary.

AGE	PHYSICAL	QUANTITY	SPATIO-TEMPORAL	MEANS
3-5	58.0	30.7	33.9	40.9
6-8	79.0	64.0	63.8	68.9
9-11	82.3	64.3	89.4	78.7
ADULT	93.9	89.4	97.4	93.6
MEANS	78.3	62.1	71.1	70.5

TABLE 10-B: MEAN PERCENTAGE DISCRIMINATIVE SCORES FOR PHYSICAL, QUANTITY AND SPATIO-TEMPORAL DISCRIMINATIONS.

The developmental increase in discriminative ability occurred for all three types of discrimination. However, physical discrimination constituted the least developmental effect, with quantity and spatio-temporal discriminations producing stronger developmental effects. Analysis of variance showed a significant difference overall between physical, quantity and spatio-temporal percentage discriminative scores ($F=12.88$, $df=2,120$, $p<0.001$). There was also a significant interaction between age group and

discrimination type on percentage discriminative scores (F=3.54, df=6,120, p<0.005).

3. Redundant reference

Redundant expressions were where a subject used extra information in contexts which did not require such information: 'the red car' (when both cars were red), 'the first man' (when only one man) and 'the small cow who is in the park' (when there was only one small cow). Redundant scores for each subject were converted to a percentage of the total number of discriminative scores produced and the mean percentage of redundancy to discriminative ability is presented for each age group in table 10-C.

AGE	MEAN PERCENTAGE REDUNDANCY SCORES
3-5	5.3
6-8	19.0
9-11	26.7
ADULT	28.7
MEAN	19.9

TABLE 10-C: MEAN REDUNDANCY SCORES AS A PERCENTAGE OF DISCRIMINATIVE SCORES

The table shows that, with increasing age, there was an increase in referential redundancy, but with 9 to 11 year olds and adults producing approximately the same percentage of redundant expressions. This effect was confirmed by

analysis of variance ($F=8.0$, $df=3,60$, $p<0.0001$) and Newman-Keuls analysis revealed that the effect was due to the difference between 3 to 5 year olds against all other groups.

Table 10-D shows the total number of different types of redundancy scores for each age group (Physical, Quantity or Spatio-temporal).

AGE	PHYSICAL	QUANTITY	SPATIO-TEMPORAL
3-5	3	0	2
6-8	4	1	37
9-11	8	0	63
ADULT	3	0	88
MEANS	4.5	0.25	50.0

TABLE 10-D: TOTAL REDUNDANCY SCORES ACROSS REDUNDANCY TYPE (PHYSICAL, QUANTITY AND SPATIO-TEMPORAL).

The table shows that from 6 years of age, the majority of redundancy was spatio-temporal. Analysis of variance found a significant difference between types of redundancy ($F=90.93$, $df=2,120$, $p<0.001$) and an interaction between age group and redundancy type ($F=14.87$, $df=6,120$, $p<0.001$).

Although there was no fixed number of appropriate

spatio-temporal scores (appropriateness varied as a function of the expressions chosen by each speaker throughout the game), in order to compare discriminative and redundant scores, table 10-E shows the total and mean number of spatio-temporal discriminative scores (before converted to percentages) and total and mean number of spatio-temporal redundant scores per age group.

Totals (mean) SPATIO-TEMPORAL REFERENCES.

<u>AGE</u>	<u>DISCRIMINATIVE</u>		<u>REDUNDANT</u>	
3-5	24	(1.5)	2	(0.1)
6-8	80	(5.0)	37	(2.3)
9-11	119	(7.4)	63	(4.0)
ADULT	91	(5.7)	88	(5.5)
<u>MEANS</u>	<u>78.5 (4.9)</u>		<u>47.5 (3.0)</u>	

TABLE 10-E: SPATIO-TEMPORAL REFERENCES: DISCRIMINATIVE AND REDUNDANT.

The table shows that discriminative spatio-temporal scores exceeded redundant spatio-temporal scores in all age groups. Using Wilcoxon tests, the differences between these scores were significant for all age groups, except for adults (3-5: $Z=-3.06$, $p,0.005$, 6-8: $Z=-2.61$, $p<0.01$, 9-11: $Z=-2.41$, $p<0.05$, Adults: $Z=-0.50$, NS, 2-tailed tests).

4. Pronominal reference

Table 10-F shows the total number of appropriate and inappropriate usage of pronouns. Again the total number of possible uses of an appropriate pronoun varied as a function of each individual narrative. An appropriate usage of a pronoun was to refer to an item which was in the current focus of the discourse representation or, when other referents were also in focus, the intended referent needed to be disambiguated on the basis of plausibility or lexical marking. Any other use of a pronoun was scored as inappropriate.

<u>AGE GROUP</u>	<u>PRONOMINAL REFERENCE</u>	
	<u>APPROPRIATE</u>	<u>INAPPROPRIATE</u>
3-5	28	1
6-8	40	0
9-11	46	0
ADULT	113	0

TABLE 10-F: TOTAL APPROPRIATE AND INAPPROPRIATE PRONOMINAL REFERENCE

The table shows that, for all age groups, when pronominal reference was made, it was used appropriately. There was a strong age increase in the use of pronouns, with adults using pronouns approximately four times as often as 3 to 5 year olds. This age increase in appropriate pronominal reference was confirmed by analysis of variance ($F=9.8$, $df=3,60$, $p<0.0000$) and Newman-Keuls analysis showed that this effect was due to the difference

between adults scores compared with the other three groups.

Five referents had been brought into high focus (by being mentioned twice previously). Although high focus pronominal scores are included in table 10-F scores, table 10-G presents only the mean scores for using either a pronoun or a definite article when a referent was in HIGH FOCUS.

(means out of 5)	<u>REFERENTS IN HIGH FOCUS</u>	
<u>AGE</u>	<u>PRONOUN</u>	<u>DEFINITE ARTICLE</u>
3-5	0.6	4.1
6-8	0.4	4.4
9-11	1.1	3.9
ADULT	2.0	2.9
<u>MEANS</u>	1.3	3.8

TABLE 10-G: PRONOUNS AND DEFINITE ARTICLES FOR REFERENTS IN HIGH FOCUS.

The table shows that, for referents in high focus, the majority of references involved using definite articles and there were few pronominal expressions. This was confirmed by analysis of variance ($F=83.92$, $df=1,60$, $p<0.001$). However, the adults produced more pronouns and less articles compared to other groups and there was a significant interaction between age and type of reference (pronoun, definite article) for referents in high focus

(F=4.83, df=3,60, p<0.005).

B) SUBSIDIARY RESULTS

5. Indefinite expressions

Table 10-H presents the means and mean percentages of predicted indefinites (total out of 4) and unpredicted indefinites (total out of 42). A predicted indefinite score was for first mention of the first of two non-specific/identical referents to be placed on the board. An indefinite expression involved either using an indefinite article plus a noun (eg. A tree) or involved using partitives (eg. One of the trees).

<u>AGE GROUP</u>	<u>PREDICTED</u>		<u>UNPREDICTED</u>		MEAN TOTAL
(Mean)Mean%	(out of 4)		(out of 42)		
3-5	(0.6)	15.6	(1.4)	3.3	2.0
6-8	(1.1)	28.1	(1.2)	2.8	2.3
9-11	(2.6)	65.6	(1.4)	3.3	4.0
ADULT	(3.4)	82.8	(0.6)	1.3	3.6
MEAN TOTALS	(1.9)	48.0	(1.2)	2.7	

TABLE 10-H: PREDICTED AND UNPREDICTED INDEFINITE EXPRESSIONS.

The table shows that overall, scores for predicted indefinites were fairly low, but there was a developmental improvement in these scores, with adults performing as predicted 82.8% of the time and with 3 to 5 year olds

showing only 15.6% of predicted indefinites. There were a few scores for unpredicted indefinites in all subject groups, with the least amount appearing in the adults group. However, the overall percentage of unpredicted indefinites was very low in all age groups. Analysis of variance showed a significant difference between percentage predicted and unpredicted indefinite scores (scores expressed as a percentage of the total possible predicted and total possible unpredicted indefinites. $F=171.71$, $df=1,60$, $p<0.001$) and there was an overall effect of age group on indefinite expressions ($F=14.99$, $df=3,60$, $p<0.001$). Finally, there was an interaction between age group and percentage scores for predicted/unpredicted indefinite expressions ($F=21.45$, $df=3,60$, $p<0.001$). The interaction is due to the fact that, with age, there was an increase in the proportion of predicted indefinite scores and a decrease in the proportion of unpredicted indefinites.

6. Display outcome

A display error was scored whenever an item on the listener's final display was not in the same position as the matching item on the speaker's board. The errors were also compared with ambiguous expressions from the speaker. Table 10-I presents the mean number of display errors per age group, together with the mean proportion of errors which were due to speaker ambiguity.

AGE	MEAN DISPLAY ERRORS	MEAN PROPORTION DUE TO AMBIGUITY
3-5	3.8	2.7
6-8	1.6	0.6
9-11	1.2	0.5
ADULT	1.0	0.5
MEANS	1.9	1.1

TABLE 10-I: MEAN DISPLAY ERRORS AND MEAN PROPORTION OF ERRORS DUE TO AMBIGUITY.

The table shows a decrease in display errors with age, with there being on average, approximately 4 errors from the youngest subjects and on average, only 1 error from adults. A Kruskal-Wallis one-way analysis of variance confirmed that there was a significant effect of age on display errors (Chi-square=26.27, $p < 0.0000$, corrected for ties). Looking at the proportion of errors due to ambiguity, the errors were mainly confined to the 3 to 5 year olds and there is only a low proportion of ambiguity related error in either of the other 3 groups. A Kruskal-Wallis one way analysis of variance test confirmed that the effect of age on display errors due to ambiguity, was significant (Chi-square=26.04, $p < 0.0000$, corrected for ties). Finally, a negative correlation between discriminative scores and display errors due to ambiguity was significant, using a Spearman rank correlation test ($r = -0.72$, $p < 0.000$).

7. Listener intervention

Table 10-J presents the mean number of comments, responses and/or questions from the listener, in each age group.

<u>AGE</u>	<u>LISTENER INTERVENTION</u>
3-5	2.4
6-8	2.0
9-11	1.8
ADULT	1.1
MEANS	1.8

TABLE 10-J: MEAN LISTENER INTERVENTION SCORES

The table shows a decrease in intervention with age, but this effect was not significant, using analysis of variance ($F=1.91$, $df=3,60$, NS).

Table 10-K shows the mean listener interventions when speakers were ambiguous, as a percentage of the speaker's total ambiguity score, for each age group. The table also shows mean PROCESS questions asked, when speakers were ambiguous, as a percentage of the speakers' ambiguity scores.

INVERVENTION AS A FUNCTION OF AMBIGUITY		
AGE	OVERALL INTERVENTION	PROCESS QUESTIONS
3-5	10.8	6.5
6-8	20.7	10.5
9-11	18.6	14.0
ADULT	35.7	35.7
MEANS	21.5	17.1

TABLE 10-K: LISTENER INTERVENTION AND PROCESS QUESTIONS AS A FUNCTION OF SPEAKER AMBIGUITY.

At all ages, there was a low percentage of ambiguity related intervention. However, the table shows a developmental increase in intervention generally and an even larger developmental increase in the asking of process questions, when speakers were ambiguous. The age effects were confirmed by analysis of variance (General intervention: $F=3.2$, $df=3,55$, $p<0.05$. Process questions: $F=4.87$, $df=3,55$, $p<0.005$) and the Newman-Keuls test showed that these age effects were both due to the oldest group.

DISCUSSION

The findings of this experiment supported the hypothesis concerning discriminative ability: There was a clear increase in discriminative ability with age and these effects were due to a developmental increase in the ability

to appreciate the need to discriminate between referents to prevent ambiguity, rather than due to the young child's problem in carrying out the comparisons per se. The evidence to support this conclusion is that all subjects, including those who showed poor discriminative ability, performed maximally on the pretest when physical, quantity and spatial comparisons between referents were explicitly requested of them.

The developmental increase in spontaneous discriminations was due mainly to the increase in quantity and spatio-temporal discriminations with age. It was suggested that physical differences between the sizes of the two 'cows' was a salient feature for younger children. However, the fact that quantity discriminations required representing both the array ON the board and the array OFF the board, may account for why, contrary to expectations, quantity did not turn out to be a salient feature in this experiment.

It was also found that redundancy increased with age and this was confined almost exclusively to spatial redundancy, but this effect appears to be due to the fact that older subjects provided more information generally than did younger subjects. Moreover, the increase in pronominal reference with age seems to reflect the fact that older subjects were more concerned with exactness and perfection with regards to the positioning of the listener's referents.

It was predicted that, as in the previous experiments, young subjects would have problems using

indefinite references to first mention the first of two non-specific (identical) referents to be placed on the board. In fact, this is what was found. The youngest two groups were very poor at using indefinites where they were expected. However, at all ages, predicted indefinites were always higher than unpredicted and inappropriate indefinites.

Speaker ambiguity was clearly related to listener performance: The listeners final display was less likely to match the speakers with decreasing age. Furthermore, 3 to 5 year olds were less likely to question the speaker despite the fact that younger listeners received more ambiguous messages. Moreover, older children and adults used a fair proportion of process questions (Which woman are you talking about?), when references were ambiguous, suggesting that knowledge of referential ambiguity is acquired at approximately the same time for both comprehension and production.

5.8 EXPERIMENT 11: Pronominal reference as a function of lexical and pragmatic factors.

INTRODUCTION

A major difference between full definite noun phrases and pronouns is the degree of presuppositional information each provides. If there is more than one referent available, a full noun phrase is often necessary to discriminate between referents. However, if reference can be disambiguated by thematic roles (the main referent is often referred to with a pronoun) and possibly through pragmatic information (plausibility), pronominal reference is acceptable.

Therefore, the following experiment looked at children's production of pronouns. The factors manipulated were lexical and pragmatic ambiguity in a context where more than one referent was in focus. It was predicted that, as with definite articles, young children will fail to vary their use of pronouns as a function of ambiguity: Children were not expected to use pronouns discriminatively across 'lexical' and 'pragmatic' conditions.

METHOD

Subjects

48 children from a Durham school were divided equally into 2 age groups to act as subjects (3 to 5 years [mean age: 4.11] and 6 to 8 years [mean age: 7.4]). Also, 10 adult undergraduates from Durham University were used for the pilot test. In each group (including the adults),

there was an equal number of males and females. Each of the children were matched with a subject of approximately the same age. Therefore, for the child groups, there were 12 subject dyads per group.

Materials

A picture depicting 2 male and 2 female individuals with their names printed underneath (John, Peter, Jane, Susan) was required for introducing the referents in the experiment.

A set of cards were also designed. Each illustrated a house with 2 of the above referents standing in front of it. Each picture also illustrated a ball. The pictures were so constructed that the experimenter could manipulate the movement of the ball and the individuals from the back of the card: The experimenter was able to slide the ball from behind, so that it appeared to the child that the ball was leaving the hands of one of the individuals and was being caught by the other person. Also the experimenter was able to move either of the individuals onto their side, so that it appeared that they had fallen over.

(See figures 11.1 - 11.3 for examples of pictures).

A listener mask and recording equipment made up the remaining apparatus.

Design and Procedure

For each trial, 2 referents were brought into focus, but depending on the type and action of these individuals, the

subject's use of a pronoun would either be appropriate or inappropriate. Two variables were manipulated. These were:

1. LEXICAL CONSTRAINT: same or different gender.
2. PRAGMATIC ASSOCIATIONS: plausible or neutral.

The first factor, 'lexical constraint', meant that reference was ambiguous lexically (both of same gender) or unambiguous lexically (each of different gender). This factor was manipulated as the experimenter chose the picture for that particular trial. If the picture was of John and Peter OR of Jane and Susan referents were of same-gender. However, pictures of Jane and Peter OR John and Susan were used on different-gender trials. For example:

1ST REFERENCE: Peter and Jane/John are in front of the house'.....

The second factor, 'pragmatic associations', meant that reference was ambiguous pragmatically (no plausible referent) or unambiguous pragmatically (one plausible referent). This factor was manipulated as the experimenter moved the card for the speaker to refer to the individuals for the third time. If the movement was of a person falling over, neither referent was more plausible (neutral pragmatics), but if the movement was of a person throwing the ball back again, one referent was more plausible: that referent who currently had the ball in

their possession. For example:

2ND REFERENCE: Peter throws the ball to John

3RD REFERENCE: He (John) throws it back again / falls over.

Second reference to individuals always involved the same action: X throwing the ball to Y. The purpose of second reference was to provide the pragmatic context for the third reference (ie. He throws it back again) as well as to bring the referents into focus, because pronouns are normally used to refer to referents which are in the focus of discourse.

As the experiment used a repeated measures design, each subject received each condition ($2 \times 2 = 4$) of the experiment and received 3 trials for each condition: $3 \times 4 = 12$ experimental trials per subject.

The procedure was as follows:

Phase 1: Introduction to pictures and referents

Each pair of subjects was taken into the experimental room. This was the school video room and was thus, very familiar to all the children. Both subjects were shown the picture of John, Peter, Jane and Susan. Then, they were shown the series of cards depicting 2 individuals in front of a house and they were, in turn, asked to name the referents on each card, until the experimenter was satisfied that they were

able to recognise and remember the names of the four individuals.

The subjects were then shown that the experimenter could move the back of a card to make the individuals throw the ball to each other and/or to make one of them fall over. The subjects were given a chance to play with the moving cards if they requested it.

Phase 2: Instructions

One of the subjects acted as listener and was blindfolded throughout the experiment. However, after 6 trials, the listener was allowed to remove the mask and play with the cards for awhile before continuing with the remaining trials. This rest period was to prevent restlessness or boredom from the subjects, especially the listener.

The experimenter then delivered the following instructions: "I am going to show you these pictures and make certain things happen (eg. let one person throw the ball to the other person or make somebody fall over). Now, as (L) cannot see what is going on, I want you (S) to tell her/him which picture I have picked up and what is happening. Therefore, (L) will know what is going on even though s/he cannot see".

Phase 3: Reference to individuals and actions

Each trial invited the speaker to use 3 sentences.

FIGURE 11.1 THE FOUR INDIVIDUALS SERVING AS THE REFERENTS

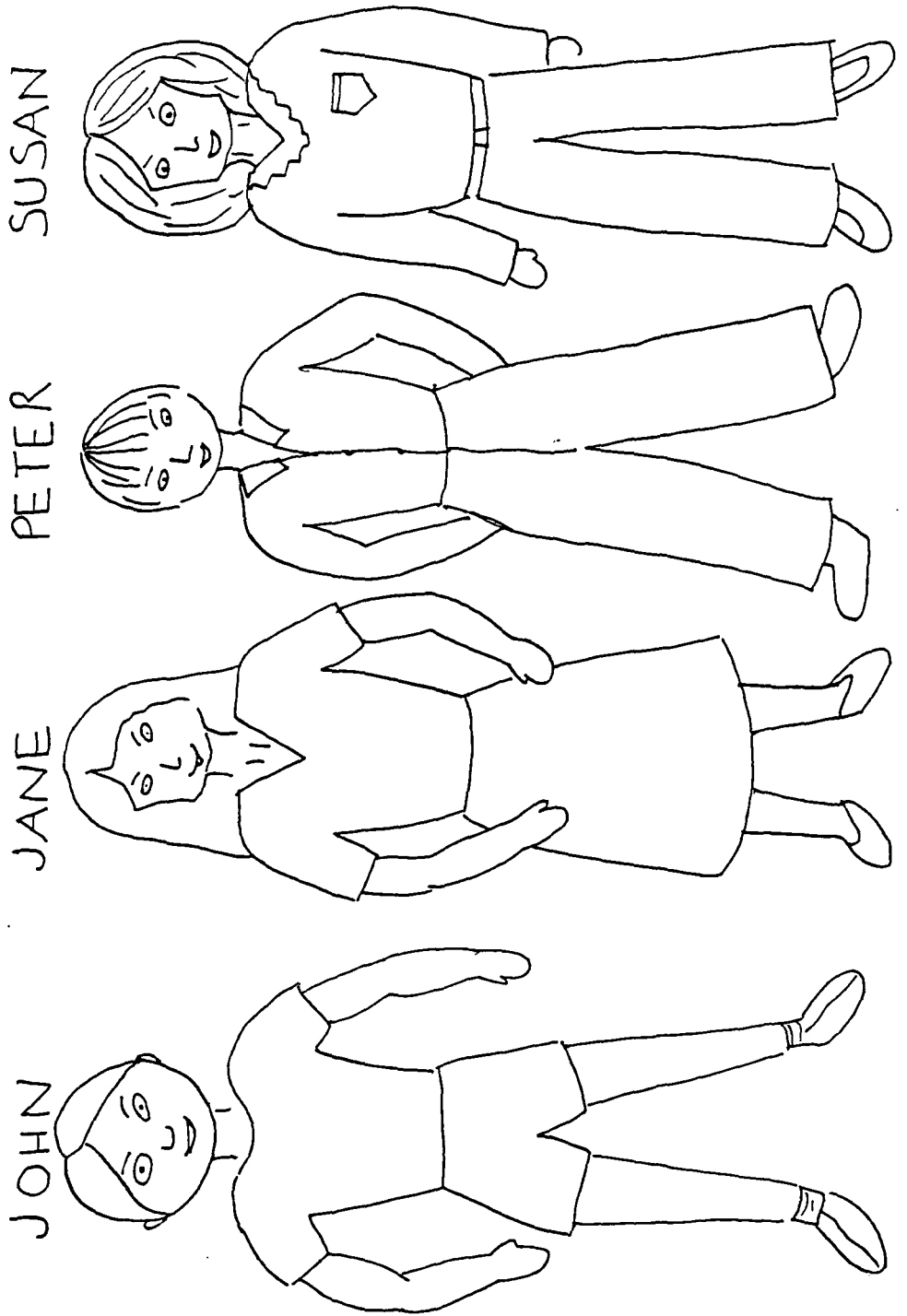


FIGURE 11.2 PICTURE SHOWING 'SECOND REFERENCE'

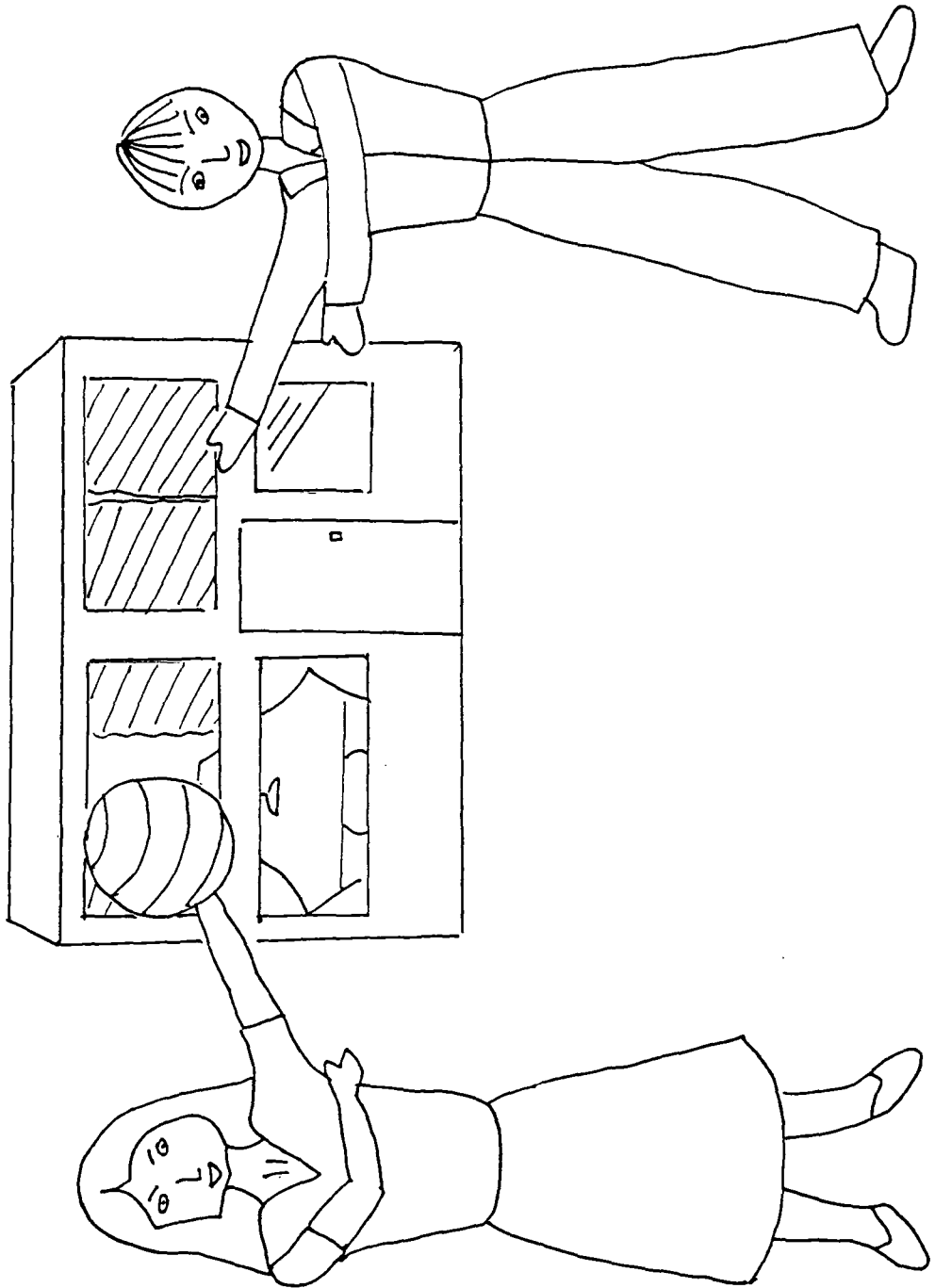
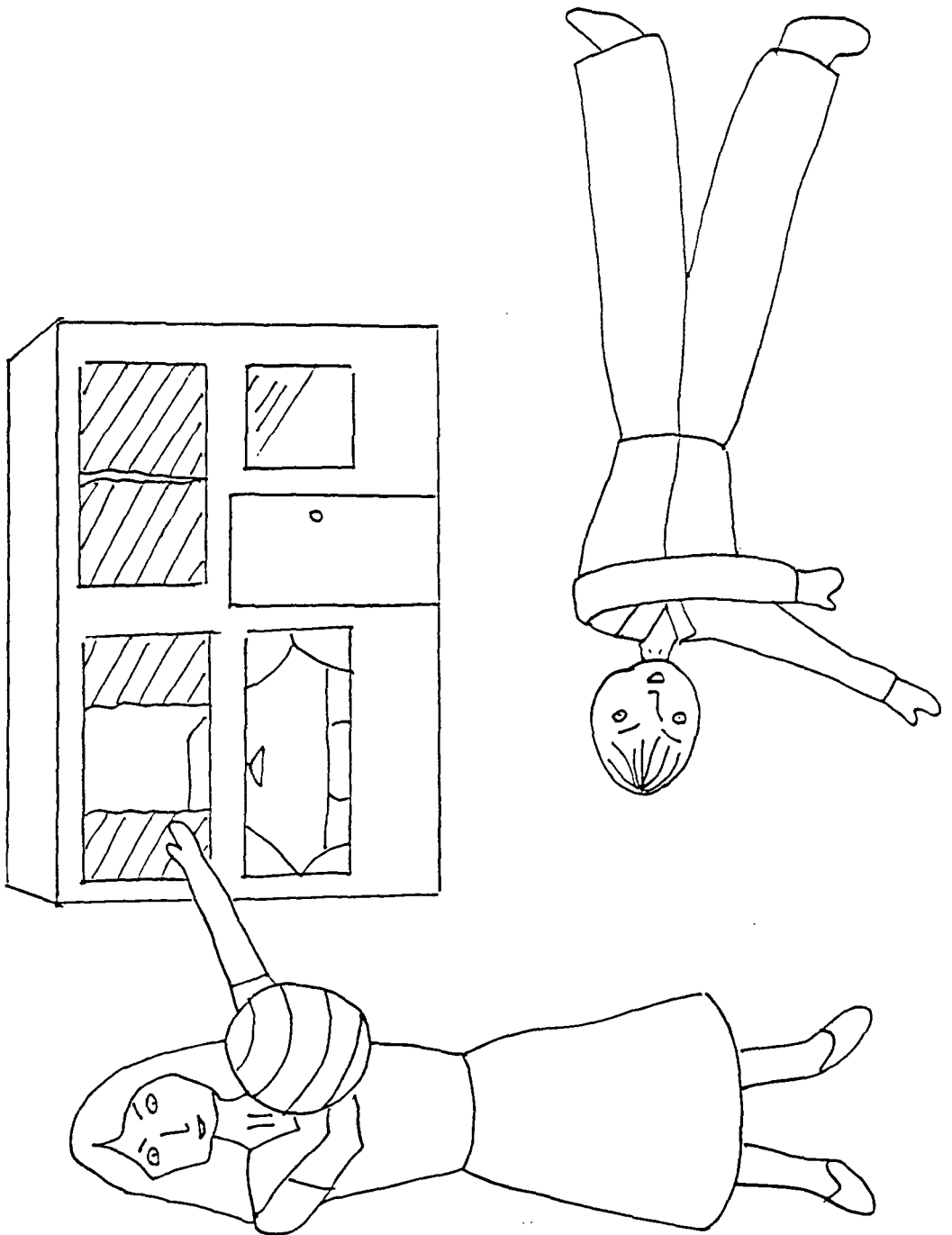


FIGURE 11.3 PICTURE SHOWING 'THIRD REFERENCE' - NEUTRAL
PRAGMATICS



REFERENCE 1

Firstly, the experimenter picked up one of the pictures. At this point the speaker was required to introduce this picture to the listener. For example, "This is the house with Jane and Peter in front of it". The picture chosen determined the lexical condition of the trial: whether referents were of the same or different gender.

REFERENCE 2

Next, the experimenter moved the ball so that it looked as if one of the individuals was throwing it to the other. At this point, the speaker had to describe the action to the listener. For example, "Peter throws the ball to Jane". This sentence provided the pragmatic context for the next reference and brought the individuals into focus.

REFERENCE 3

Finally, the experimenter either moved the ball so that it looked as if the person who currently had the ball was throwing it back to the initial possessor OR she moved one of the individuals so that they were falling over. For example, either "She throws the ball back again" (plausible condition) or "He falls over" (neutral pragmatics condition).

Thus, it is only first and third reference that varied according to the conditions of the experiment: Whether referents were of the same or different gender and whether one was more plausible than the other.

A number of practice trials were given to subjects to ensure that they understood what the experiment required of them. All 12 trials were presented randomly for each subject. Interest was in the type of reference subjects chose for reference 3: pronouns, proper names, definite articles, other expressions.

It should be mentioned at this point that, initially, 2 of the 4 year old subjects used a pronoun for second reference when referents were of the same gender. As pronominal reference was only expected for reference 3, after the referents had been brought into focus, these subjects were excluded and 2 further pairs of 3 to 5 year olds were selected to make up an equal number for both age groups: 12 pairs per group.

PILOT STUDY

Prior to the experiment, 10 adult undergraduates were asked to carry out the experiment, with each subject pretending that s/he was speaking to a blindfolded listener. 3 trials for each of the 4 conditions ($3 \times 4 = 12$) were also given to each adult subject, although a rest period was not needed.

Adults preferred to use proper names rather than pronouns. However, despite the fact that pronouns were the least form of definite expression in all conditions, no adult subject used a pronoun when a referent was both pragmatically and lexically ambiguous: Same gender and

neutral plausibility. Most pronouns were used when one referent was plausible, (irrespective of lexical ambiguity). The mean scores out of 3 for adults in the pilot test are shown in the table 11 below:

(means out of 3)	SAME-GENDER	DIFFERENT-GENDER	MEAN TOTALS
PLAUSIBLE	1.4	1.3	<u>1.4</u>
NEUTRAL PRAGS.	0.0	0.9	<u>0.5</u>
MEAN TOTALS	<u>0.7</u>	<u>1.1</u>	

TABLE 11: RESULTS OF PILOT STUDY ON ADULT UNDERGRADUATES.

Analysis of variance showed an effect of pronominal use as a function of 'pragmatic associations' ($F=30.38$, $df=1,9$, $p<0.001$), but no effect of pronominal use as a function of 'lexical constraints' ($F=3.69$, $df=1,9$, NS). However, there was a significant interaction between 'pragmatic associations' and 'lexical constraints' in the use of pronominal reference ($F=5.63$, $df=1,9$, $p<0.05$) and this effect was due to the fact that a small number of pronouns were used in the neutral condition, but only for the different-gender condition. These results show that adults did not use pronouns ambiguously: when a referent could not be identified lexically or pragmatically. Moreover, they were reluctant to rely on lexical information alone to disambiguate reference.

RESULTS

All subjects correctly named the referent on first and second reference, including those who were selected after it was decided to exclude the results of two of the initial subjects.

All subjects used definite reference for third reference and few subjects (3 cases) used a simple definite noun phrase such as 'the boy'. As this latter form of reference was only ever used when referents were of different gender and therefore was never used when reference was ambiguous lexically, scores were treated as binary: pronouns or other expressions. As nearly all other expressions involved mentioning proper names, (as for first and second reference), the other category is simply referred to as 'names'.

When pronouns were used on third reference they were never incorrect lexically: When the intended referent was male, the subject always used 'he' and when the intended referent was female, the female gender pronoun, 'she' was always used.

As each subject received 12 trials, table 11-A presents mean scores for pronouns and names used in each age group out of 12.

AGE	PRONOUNS	NAMES	TOTAL
3-5:	5.3	6.7	=12
6-8:	2.9	9.1	=12
TOTALS	8.2	15.8	

TABLE 11-A: MEAN PRONOUN AND NAME SCORES IN EACH AGE GROUP.

It is clear from this table that, as for the adult sample in the pilot study, names were preferred to pronouns. However, the difference was smaller for 3 to 5 year olds who used more pronouns overall than 6 to 8 year olds.

Analysis of variance showed that there was a significant difference between the use of pronouns and names overall ($F=9.84$, $df=1,22$, $p<0.005$) but an interaction between age group and pronoun/name scores was not quite significant ($F=3.65$, $df=1,22$, NS).

Table 11-B presents the mean production of pronouns for each age group in each of the 4 experimental conditions.

AGE	SAME-GENDER		DIFFERENT-GENDER	
	PLAUSIBLE	NEUTRAL	PLAUSIBLE	NEUTRAL
3-5:	0.9	1.9	1.3	1.2
6-8:	0.8	0.9	0.8	0.4
MEANS	0.9	1.4	1.1	0.8

TABLE 11-B: MEAN PRONOUN SCORES IN EACH CONDITION: LEXICAL CONSTRAINT AND PRAGMATIC ASSOCIATIONS.

The table shows that there was very little difference across all 4 conditions in the choice of a pronoun. The only differences were that the 3 to 5 year olds used most of their pronouns in the same-gender/neutral (ambiguity) condition and the 6 to 8 year olds used pronouns least in the different gender/neutral condition. However, analysis of variance failed to show a significant difference in the use of pronouns across the 4 conditions ($F=2.17$, $df=3,66$, NS), nor was there a significant interaction between age group and the use of the pronoun across the 4 conditions ($F=1.08$, $df=3,66$, NS).

Analysis of variance also failed to yield a significant effect of lexical constraint ($F=1.74$, $df=1,22$, NS). Nor was there an interaction between age group and lexical constraint ($F=0.05$, $df=1,22$, NS). Moreover, there was no significant effect of pragmatics ($F=0.67$, $df=1,22$, NS).

Nor was there an interaction between age and pragmatics ($F=2.06$, $df=1,22$, NS).

However, there was an interaction between lexical and pragmatic effects ($F=4.80$, $df=1,22$, $p<0.05$). Yet, there was no 3-way interaction between age, lexical and pragmatic conditions ($F=0.95$, $df=1,22$, NS). Looking at table 11-B, it is clear that the interaction between lexical and pragmatic effects is due to the fact that, in both age groups, for the neutral pragmatics condition, more pronouns were used in the same-gender condition compared with the different-gender condition.

In fact, comparing same-gender and different-gender pronouns in each pragmatic condition revealed a significant effect of lexical condition in the neutral condition, ($F=5.26$, $df=1,22$, $p<0.05$), but not in the plausible condition ($F=0.28$, $df=1,22$, NS).

However, lexical condition by age group interactions were not found to be significant, in either the neutral condition ($F=0.21$, $df=1,22$, NS), or the plausible condition ($F=0.78$, $df=1,22$, NS).

DISCUSSION

Firstly, the results show that, even from three years of age, young children were able to discriminate between male and female gender markers when using pronouns: All pronouns were 'he' when referring to a male referent and 'she' when referring to a female referent. Thus, any ambiguity as a result of lexical marking does not appear to be a reflection of the young child's inability to

appreciate lexical markers for gender per se.

The hypothesis that young children would use pronouns ambiguously was supported: Whilst no adult used a pronoun when reference could not be disambiguated either lexically or pragmatically, the majority of the pronouns used by the 3 to 5 year olds were in the same-gender/neutral condition and for both age groups (3 to 8 years), more pronouns occurred for same-gender referents compared with different-gender referents in the neutral pragmatics condition. Therefore, the results show no evidence of young children appreciating the need to consider the 'specificity' of a referent when using pronouns in discourse, because they produced as many (and even more) pronouns when referents were not made specific through lexical markers or pragmatic inferences and thus, many pronouns were ambiguous.

The adults showed a stronger sensitivity to pragmatics than to lexical constraints: They were more reluctant to rely solely on lexical marking for disambiguating reference and the main factor determining their choice of pronoun was whether the referent was plausible or not. Again, neither of the child groups showed this preference. In fact, the 3 to 5 year olds were more likely to use pronouns in the neutral pragmatics condition.

In any case, these findings do show how difficult it is to manipulate those situations where pronouns will be preferred to other forms of definite expressions in discourse or narrative-tasks, because even adults preferred

full noun phrases when pronouns would have been sufficient.

5.9 GENERAL DISCUSSION

The findings of this chapter have shown that the young child fails to appreciate 'specificity' in discourse and will only produce unambiguous anaphoric reference when the information required to make a referent specific is salient to them, just as their performance was shown to be dependent on salience for exophoric reference.

Moreover, despite the lexical and structural constraints of pronominalisation, it appears that the understanding and production of unambiguous definite reference in discourse follows the same developmental trend for pronouns as it does for reference using definite articles.

Firstly, experiment 6 showed that young children's choice of indefinite or definite expression in discourse was determined solely by the listener's 'familiarity' of the referents. This supports the prediction that young children are unlikely to appreciate 'specificity' in discourse, because the referential context is not salient.

Moreover, on second mention, the children sometimes used pronouns ambiguously. Also, children failed to use further presuppositional information, such as 'the man who is driving opens the door and the other man crosses the road', to discriminate between animate referents. However, this may be because the listener was able to infer who was the most plausible referent from the context. For example, 'She poured it out', where the verb 'poured' implies that (it) refers to a liquid of some sort (tea) and the person holding the tea-pot is likely to be the person

doing the pouring (she). Moreover, this experiment did not emphasise the importance of successful communication.

However, experiment 7, which did use a task that required unambiguous reference for a successful outcome, showed that young children failed to provide the necessary disambiguating information and performance was worse when this information was episodic (the one who was skipping) than when required information was physical (the red one).

The fact that even older subjects and adults preferred indefinites on first mention and definites on second mention is probably due to the fact that, unlike experiment 1, experiment 7 did not mention the word 'story' and subjects were more likely to introduce referents in a non-narrative form, such as, 'In the first picture, there is a (one) boy and.....'.

Nevertheless, there is evidence to suggest that the youngest group were having problems with the particular task used in experiment 7 and were less motivated to participate. For example some of the youngest subjects judged that the pictures were alike when they were different from each other, suggesting that the listener sometimes failed to compare referents correctly (whether they understood the need for comparison or not). Moreover, for the 'distractors', some of the young listeners selected a picture with incorrect objects/toys, suggesting that these subjects were having problems due to lack of attention, memory difficulties or misunderstanding of instructions. In fact, even if the

message was good, younger subjects were less likely to choose the correct picture than older subjects. Also, the possibility remains that young children did not provide presuppositional information in the episodic group merely because the production of two subject-embedded clauses placed too great an information processing burden on them (Slobin 1966).

Moreover, experiment 7 may have produced artificial findings as a result of the complexity of the task for the younger subjects, such as having two referential arrays (one for speaker and one for listener) and having to make comparison and judgements after each story. Nevertheless, although the 6 to 8 year olds did not appear to suffer from task complexity, they were still poor at producing discriminative references when this required using non-salient episodic information. Furthermore, listeners of this age still tended to blame themselves (as listener) when outcome failure was a result of an ambiguous message. This result suggests that children under 9 years of age are still relying on salience rather than 'specificity'.

Other factors which may have either confounded or obscured any genuine referential abilities in experiment 7 were that the second picture of the stories often provided clues as to what the intended referent should be. For example, the toys used in the first picture were often presented as being on the floor or in the background of the second picture in order to make the second picture look as if it followed on from the first. Also, referents were

arranged spatially in the same way in the second picture as they were in the first picture and so maybe subjects were matching order of mention with left to right positioning. Moreover, sometimes a listener could have easily used pragmatic inferences to decide which referent performed a particular action in the second picture. For example:- 'The thin man (who was previously holding a jug) spilt some water on the floor, and the fat man (who had eaten a big cake) ate ANOTHER even bigger cake'.

It is also possible that discriminative scores in the episodic group were overestimated for the older subjects, because they preferred to use spatio-temporal features to discriminate between referents and it seems more likely that these characteristics were salient than were previous actions/episodes. In any case, it cannot be concluded affirmatively that older subjects were being discriminative unless redundancy is looked at: If subjects use colour and episodic information when it is redundant, then it does not necessarily mean that the subject understands the need to use this information to prevent ambiguity. In fact there is experimental evidence to show that adults are often redundant in discourse (Whitehurst 1976, Deutch and Peckmann 1982).

The finding that listener choices were often better than would be expected given the discriminative scores of the speaker (particularly in the episodic group), is probably because the type of expression speakers had used to refer to the first picture affected the listener's expectations regarding picture 2. For example, for the

youngest group, listeners never chose a single picture for the experimental stories presumably because speakers often referred to more than one referent for picture one. Furthermore, as many of the older subjects used spatio-temporal expressions to discriminate between referents in the episodic condition, their higher than expected listener scores may be due to subjects inferring that the first mentioned referent was the one on the left and vice-versa.

Experiment 8 showed that, when performance limitations were eliminated, even infants would produce relative clauses. However, in contrast to Hamburger and Crain's claims, young children were just as likely to use relatives when they were redundant (already one referent available) as when they were necessary (2 potential referents). Thus, the findings of this experiment do not indicate that the problems in experiment 7 were due solely to processing difficulties and the increase in relatives in experiment 8 is probably due to the fact that subjects had to focus on the unique characteristics of each individual in order to answer the questions.

However, experiment 9 showed that, even when pragmatic, memory and salience factors were controlled, up to the age of about 9 years, children did not use relatives or other information to make reference unambiguous, when the information required was not salient to them (character information). In fact, even the youngest group were influenced by salience in this experiment, whereas in experiment 7, the youngest group

were poor at mentioning even physical differences between referents. Perhaps this result reflects a more accurate picture of the performance of young children when memory difficulties, problems with relative clauses and/or difficulties due to the unfamiliarity of the story-telling task have been eliminated.

'Gender' was used as minimal referential information for identifying a referent. Those cases which did not involve using gender used one anaphora, for example, 'The red one' or 'the one who likes ice-cream', as opposed to 'The girl' and therefore, it is not surprising that there was little evidence of any discriminative use of 'gender'. However, 'gender' was used slightly more in 'different-gender' conditions than in 'same-gender' conditions, at all ages, which does suggest that subjects were showing some discriminative use of 'gender', (or it could be that differences in the gender of the dolls were more salient than dolls being of the same gender).

Nevertheless, the parents' discriminative scores in the 'different-gender' condition, were more likely to involve 'gender' only than either 'colour' or 'character' only, which suggests that parents were better at appreciating the discriminative use of 'gender' in 'different-gender' conditions than were any of the children. In fact, only the parents used less 'colour' and 'character' information to discriminate between referents in the 'different-gender' condition, as compared with the 'same-gender' condition.

There were occasions, especially from the 3 to 5

year olds and from the parent group, where the timing element served either to bring about a rushed, babbled reponse, or served to inhibit responding completely. For the parents, this was probably due to their fears of being tested and of embarrassing themselves. Most subjects over 3 to 5 years responded on average in under 10 seconds. The youngest group responded in an average time of 10 to 20 seconds. One cannot say whether this time would have been any different if there had not been a time constraint. At least subjects did not give long endless stories, describing every character and event in intimate detail, as has sometimes been found. Thus, redundancy scores were from categories which were the most salient and which were the most difficult for subjects to omit completely. Hence, the majority of redundancy scores were due to an over-usage of 'colour' information.

Although ambiguity was related to outcome, the younger listeners performed slightly better than would be predicted from discriminative scores, particularly in the 'identicals / same-gender' condition. This is consistent with the fact that, for expressions that were not discriminative, there was a 50 per cent chance that the listener would pick the correct one just by guessing. This explains why there was a stronger correlation between discriminative scores and listener scores for older groups.

Finally, it is possible that the performance of the younger subjects reflected difficulties in using 'character' information. However, experiment 10 showed that although young subjects could make necessary

comparisons which were explicitly requested, they made few spontaneous comparisons when differences between referents were not salient (spatio-temporal differences).

In experiment 10, the younger group seldom mentioned physical characteristics redundantly, such as 'the (red) car', probably because it was only the DIFFERENCES between similar referents (big cow / calf) which were salient features for young children.

Although it was quantity and spatio-temporal discrimination which constituted the largest developmental improvements, contrary to the hypothesis, that quantity would be salient for the children, 9 to 11 year olds were almost as good as adults on spatial discriminations, but were still considerably poorer on quantity discriminations. A possible explanation for this difference is that quantity discrimination requires the representation of two arrays (the one on the board and the one off the board), for example 'the other car', meaning the car that is still remaining off the board. By contrast, spatial discrimination only required representing one array: the one on the board.

In previous experiments, higher redundancy has been equated with increased effects of salience. However, in experiment 10, there was no time constraint and the increase in redundancy with age may be due to the fact that children use the principle of least effort (Whitehurst 1976) and/or due to the fact that increasing redundancy reflects an increased sensitivity to the task requirements: As the goal of the task was to match the

boards of speaker and listener, redundancy could be to 'double check' that listener's were moving the same objects into the same positions. In fact, older subjects tended to emphasise redundant expressions and used stronger intonation. For example, one speaking adult delivered the following redundant information:

'Now the man who should currently be standing next to the tree in the field, moves across to the park, next to the other tree'.

It is clear from the phrase 'should currently be', that the adult speaker was checking that the listener was sure that she was talking about the 'man'. As Shantz (1981) argued, 'Redundant messages are often classed as poor messages. Yet a good speaker may be redundant precisely because the speaker knows that it is difficult for a listener to get the criterial information with only one oral presentation (p. 100)'.

The fact that each referent set was small (2 items), that there were two arrays (array off board and array on board), plus the passive role played by children with regards to choice of items, served to lower the salience of the referential array for the child and therefore, in accordance with the previous findings, seems to explain why young subjects failed to show an appreciation of referential 'specificity'. However, using definites for non-specific items was rarely confusing for the listener, except for 'the sheep', where reference was ambiguous with

regards to whether the speaker intended to refer in the singular or the plural.

Although there were few unexpected indefinites, when they were used by subjects over 5 years, they were mostly for the second identical referent to be placed on the board, but this did not seem to cause problems as the listener could infer that the intended referent was a non-specific item that is left unplaced on the board. Anyway, these indefinites were mostly for references to 'a tree' and it makes sense pragmatically that, if there is a tree in the field and a tree in the park, they must both be different trees because trees do not normally move.

The increase in pronominal usage with age seems to be a reflection of the general age increase in utterance length, rather than due to an increase in the ability to use pronouns appropriately, because older subjects and adults often used a pronoun to provide further information about the exact location of the referent just moved. In this experiment (and probably in many everyday uses), pronominal usage is often unambiguous just by chance alone. In those few cases where there was more than one referent in focus and when they shared the same lexical properties, the pragmatic information already served to disambiguate reference. For example, 'One of the cars drives on the road and it moves along to the bridge', where 'it' can only refer to the car, as it is implausible for a road to move.

However, there was some evidence to suggest that young subjects are reluctant to use pronouns at the

beginning of a sentence: Although pronominal reference was low in all age groups, even when referents were in high focus, the two youngest groups showed a non-significant preference to use a pronoun in object position. Older subjects were more likely to use pronouns towards the end of the text, presumably because increasing pronominal usage relates to increasing presuppositional information. The fact that the older subjects were more likely to use a pronoun to refer to something in high focus compared with the younger subjects, seems to contradict Clibbens' claim that young children are more likely to use a pronoun for a main character than are older children. Nevertheless, Clibbens did make the distinction between a main referent and the only referent available and found that, at all ages, subjects would prefer a pronoun to refer to the only referent. In this experiment the 'high focus' referent could often be disambiguated lexically and pragmatically and therefore, often it was probably the case that this was the only referent in focus.

Therefore, experiment 10 again showed an increase in discriminative ability with age. However, having to represent two arrays may have complicated the task and the fact that unambiguous reference could have been produced by chance alone, obscures any effects of pronominal reference.

Although experiment 11 was intended to manipulate contexts where pronouns would either be ambiguous or unambiguous through pragmatic or lexical information, pronouns were the minimum form of expression at all

ages: Proper names were usually preferred despite the fact that the referents had been brought into focus and complete ambiguity (lexical and pragmatic) only occurred for 25 percent of the trials. Moreover, pronominal reference decreased with age. The lowest proportion of pronouns was from the adults and the 3 to 5 year olds produced more pronouns compared with the 6 to 8 year olds. Thus, it appears that pronominal reference is the exception rather than the norm for these types of stories. Maybe, these mini-stories were insufficient to bring referents into high focus. In fact, Clibbens' results, that adults did not discriminate between main and subsidiary characters when choosing pronouns, may also be interpreted in this way: A main character in Clibbens' experiment probably did not meet the criteria for a high focused referent for older subjects and therefore, they used a definite noun phrase to refer to the main character just as they did for the subsidiary character.

In any case, the results of experiment 11 clearly showed that, in contrast to adults, young children did not use pronouns discriminatively as a function of ambiguity, suggesting that the development of a knowledge of 'specificity' for anaphoric reference applies to pronominal reference in the same way as for reference using definite articles.

The next chapter uses comprehension tasks. Experiment 12 looked at the comprehension of definite noun phrases and experiment 13 investigated the comprehension of pronouns. Ambiguity was manipulated by

looking at children's manual and/or verbal responses to definite references from the experimenter.

CHAPTER 6

INTERPRETING ANAPHORIC REFERENCE

6.1 ABSTRACT

Experiment 12 looked at whether children would choose a referent as a function of the ambiguity of the definite expression used by the experimenter. Experiment 13 looked at whether children would show a sensitivity to pronominal ambiguity in their choice of response.

Both experiments showed that young children would make pragmatic inferences in assigning reference. Moreover, even young children were good at using the lexical 'gender' markers of pronouns (experiment 13). However, in both experiments, young children were predominantly influenced by pragmatic factors at the expense of noticing a conflict between either explicit and implicit information (experiment 12) or lexical and pragmatic information (experiment 13).

Moreover, in both experiments, when reference was ambiguous, there was a developmental increase in the detection of ambiguity and as experiment 13 used a task which always required a manual response, this age effect cannot be merely due to the young children's 'performative bias' per se.

These results suggest that the comprehension of ambiguity follows the same developmental trend as for the production of definite reference, both for pronominal reference as well as for reference using definite articles. It seems that young children are overimpressed by

non-linguistic factors at the expense of concentrating on whether the linguistic information refers unambiguously and thus, they do not appear to be appreciating the importance of 'specificity'.

6.2 GENERAL INTRODUCTION

The experiments in chapter 5 have looked at children's production of messages in tasks which require discriminative reference for a listener. In experiment 7, it was found that this ability was age related to the listener's attributions of reference failure and in experiment 10, listener intervention was poorer in younger subjects. Thus, it appears that both the comprehension and production of message adequacy appear at roughly the same time. In fact, research has shown that the age at which a child produces unambiguous messages in a referential communication task is correlated with the age at which a child asks for specific requests from a speaker when a message is inadequate (Robinson and Robinson 1978).

Unfortunately, in experiments 7 and 10, the listeners' responsiveness to ambiguity could not be looked at in detail, because the ambiguity of a message was not specifically manipulated, but depended entirely on the ability of the speaker. Therefore, as speaker-listener dyads were age matched and as speaker performance increased with age, it was rarely the case that older listeners had to respond to an ambiguous message and so, it was impossible to look for developmental changes. Thus, it could not be affirmatively concluded that the performance

of younger listeners was due to the fact that their knowledge of message ambiguity was less developed than older subjects. Therefore, this chapter aims to look at listener performance more closely.

The two experiments reported in this chapter are comprehension tasks intended to see whether understanding definite anaphoric reference is developed at the same rate and in the same way as is production. Experiment 12 looks at the comprehension of ambiguous definite noun phrases and experiment 13 looks at ambiguous pronoun comprehension.

So far it has been argued that the problem for the young child is in understanding definite reference, that is, they are unable to use the appropriate information to make a referent specific for the listener. If production and comprehension ability co-develop, then one would expect that the young listener will be unable to judge any unsuccessful reference as being due to the inadequacy of the message itself.

Moreover, Ackerman (1981,1986), has shown that younger children display a 'performative bias', which implies that young children always prefer to choose a referent, rather than to ask the speaker specific questions or to blame a speaker when a message is inadequate. Thus, experiment 12 looks at children's comprehension of definite reference where a referent can either be chosen, due to explicit mention or through pragmatic inference, or cannot be chosen owing to message ambiguity. It is predicted that, whilst young children will be able to make pragmatic inferences, they will fail to detect ambiguity

and show a 'performative bias' (choose a referent rather than question the experimenter). Moreover, it is expected that young children will justify their choices by non-linguistic means whenever possible, owing to their inability to understand that choice of a referent depends on whether the linguistic information is ambiguous or not.

Wykes argued that the problem for children was in making the relevant inferences with pronouns, and not due to their lack of ability to make inferences per se. However, she never explicitly tested the precise inference needed to interpret the pronoun. For example to test the inference for the following sentence pair, 'Jane found Susan's ball. She gave it to her', she asked questions such as, 'If Jane needed Susan's ball, what would Jane do? Secondly, she did not explicitly test premise memory: The effects of making a more complicated inference in another of her experiments may have been due to the added complexity of the task, rather than due to adding an extra step to the inference process. Moreover, it seems that the examples she used may have biased subjects to the 'subject' antecedent, because the object noun phrase was embedded in the first sentence and different verbs were used when the 'subject' was the antecedent compared with when the 'object' was the antecedent (Subject: 'found', 'had', 'bought' and 'mended'. Object: 'need', 'want', 'ask for' and 'like'): It may be that for 'object' reference, subjects tended to focus on the association between the object of need and the 'subject' who required it, and were therefore, biased to the 'subject' referent. Moreover, a

bias to the subject may have also been caused by misinterpreting some of the sentences as if they were in the past tense (For example, 'Jane wanted Susan's pen because, she had (in the past), given it to her (Susan)'). In fact, subjects were not asked to act out the first sentence in this experiment and therefore, it is impossible to know how the sentences had been interpreted by the children.

Furthermore, Tyler and Marslen-Wilson's results (1982, 1983), that children cannot use the lexical information carried by the pronoun until 7 years of age, may have been a consequence of the type of tasks that they used (mispronunciation detection tasks): It is possible that children had difficulties because of their slower speed of assigning pronouns, rather than due to their inability to assign the pronoun per se. In fact, their better performance when a full noun phrase was used may have been because subjects were merely matching the noun phrase in sentence two with a noun phrase in sentence one. In any case, it seems that a more complicated inference was required for interpreting the pronoun, than for interpreting full noun phrases.

Experiment 13 uses a similar task to Wykes' acting-out experiment (1981, 1983), whilst avoiding the potential extraneous factors described above, to look at children's comprehension of pronominal reference.

Pronouns are generally used in different contexts to definite articles: Pronouns are normally preferred to full definite noun phrases when the referent is in focus and/or

when one referent is more plausible than another. Nevertheless, experiment 11 showed that young children would produce ambiguous pronouns just as they have been shown to produce ambiguous reference using definite articles. Thus, it is predicted that, as both pronouns and definite noun phrases must refer to a specific referent, the comprehension of referential ambiguity will follow the same pattern with pronouns as has been found with definite articles. Moreover, experiment 13 provides a means of responding to ambiguity whilst, at the same time, making a choice response. It is hypothesised that young children will be able to choose a referent through pragmatic inferences and even when they can use lexical information, they will fail to detect pronominal ambiguity, despite the fact that the task controls for any possible 'performative bias' in young children.

The main issues raised in this chapter are:

1. Does comprehension follow the same pattern as for production? Do young children fail to detect referential ambiguity in the same way as they produce ambiguous anaphoric reference.

2. Are any developmental differences a reflection of the children's knowledge of definite reference in general or does comprehension depend on the type of linguistic determiner used: Do the same patterns occur for the comprehension of pronominal reference as they do for understanding reference using definite articles?

6.3 EXPERIMENT 12: Comprehension of definite
references as a function of 'specificity' and pragmatic
context.

INTRODUCTION

This experiment required children to choose objects in response to definite references made by the experimenter.

As it has already been shown (experiment 3) that even 4 year olds are able to make plausible inferences in defining 'familiarity' for definite reference, one variable manipulated in this experiment provided a test of the listeners' use of plausible inferences when choosing a referent. This variable ('PRAGMATIC ASSOCIATION'), manipulated whether a message was plausible in relation to the objects and events being referred to.

Another variable in this experiment was the 'specificity' of reference produced for the listener. This variable ('REFERENTIAL SPECIFICITY'), manipulated whether the message made specific reference or whether it failed to discriminate between the two individuals which were familiar to the listener.

If listener performance does parallel the speaker's level in children of approximately the same age, one would expect young listeners to be able to choose the correct referent by making plausible inferences, but they should have difficulty when definite references are ambiguous and when pragmatics is insufficient to disambiguate reference. In fact, for the latter form of reference, it was predicted that, young listeners would show a performative bias and

only by 9 years of age would listeners question speakers when messages were ambiguous.

METHOD

Subjects

64 subjects selected from Durham schools were used for the experiment. 16 subjects were in each age group, with there being 4 age groups in total: 3 to 5 years [mean age: 4.6], 6 to 8 years [mean age: 7.2], 9 to 11 years [10.1] and parents. As in the previous experiment, the parent group consisted of a random selection of the mothers of the children. In the other age groups, there was approximately an equal proportion of males to females. Some subjects had participated in a previous experiment. However, none of the groups contained exactly the same sample of subjects as in any previous experiment.

Materials

2 male dolls and 2 female dolls were required for the experiment and there was one red doll and one green doll of each gender.

16 cards illustrating objects/toys were constructed. The cards were pictures of 8 larger objects and 8 smaller objects, with each picture being associated with a toy in another picture, so that each larger object was matched with a smaller one. Matched pairs were related functionally in the sense that, if the toys represented by the picture cards were to be used for their 'normal' purpose, one larger toy together with its matched smaller

toy would be required for this function/game to be carried out properly.

The actual matched cards were as follows:

1a	BUCKET	-	1b	SPADE
2a	BAT	-	2b	BALL
3a	PAINTS	-	3b	PAINTBRUSH
4a	DRUM	-	4b	STICKS
5a	GARAGE	-	5b	CAR
6a	BLACKBOARD	-	6b	CHALK
7a	NEST	-	7b	BIRD
8a	HOSPITAL	-	8b	NURSE

Before the experiment proper, twelve adults, six 7 year olds and six 5 year olds (subjects not used for the actual experiment) were presented with all the cards displayed randomly on the tray. They were then simply asked to arrange all the objects into pairs, matching them in the way that they thought would be most likely. All adults and children matched the objects into the pairs predicted by the experimenter (as shown above) and therefore, the pragmatic association between object pairs was accepted as being a consistent one for subjects of all ages. (See figure 12.1 for pictures).

Design and Procedure

For the experiment, a reading room in the school building was used. Therefore, each subject was familiar with the surroundings whilst acting as a subject.

Each subject was shown either the 2 male dolls or 2 female dolls and the gender of dolls shown was

FIGURE 12.1 PICTURES ON CARDS AS EXPERIMENTAL MATERIALS

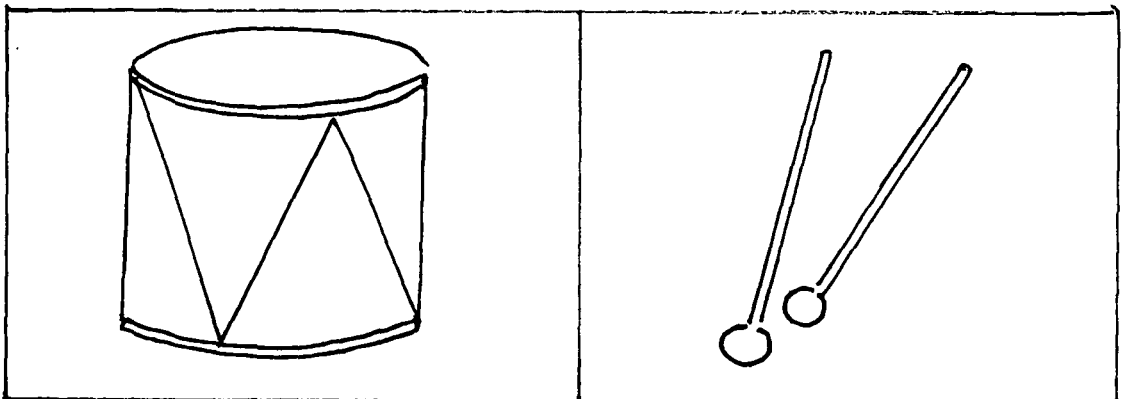
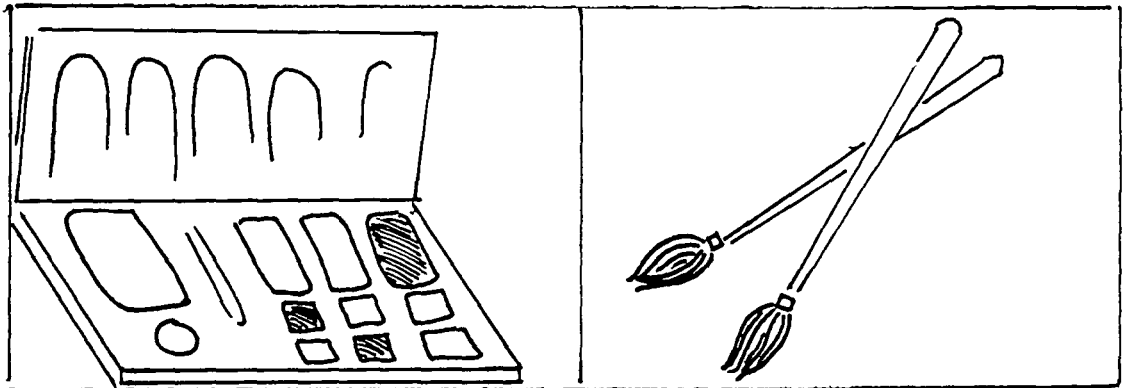
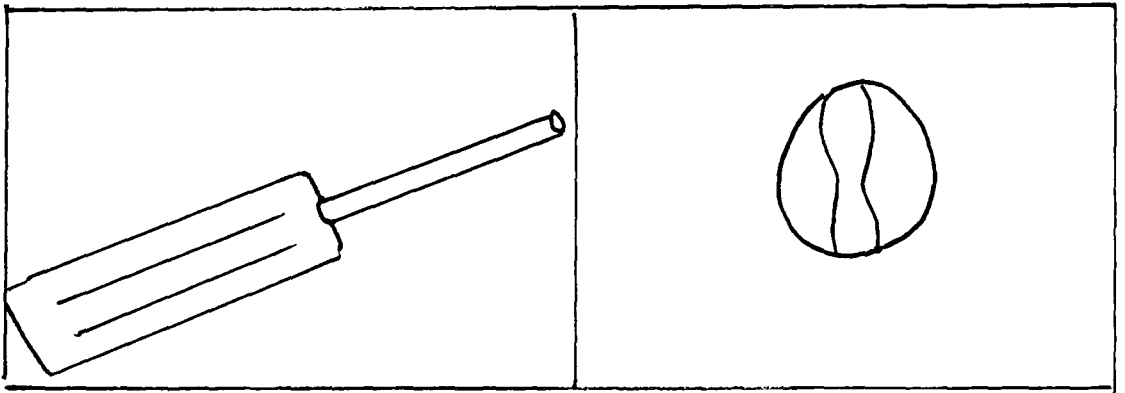
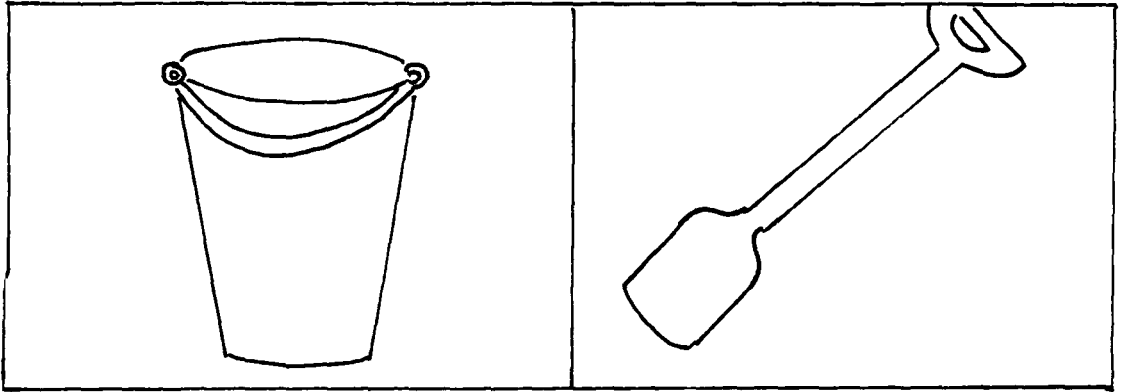
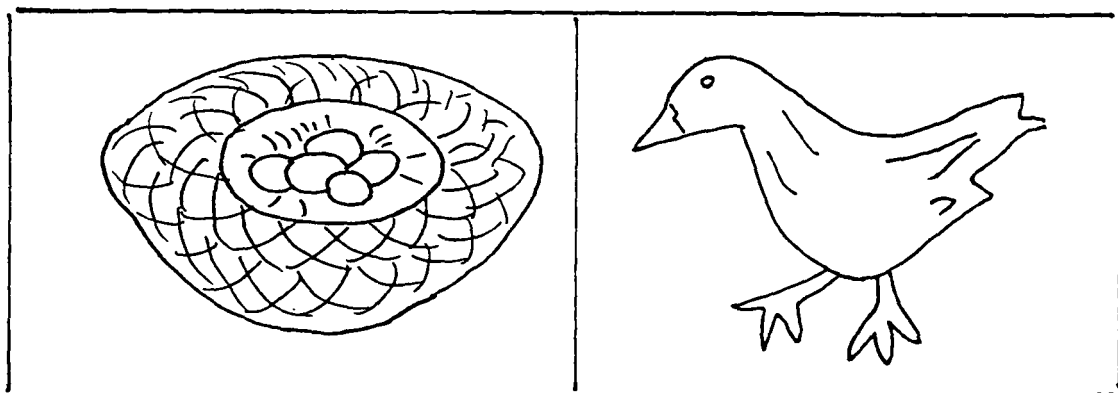
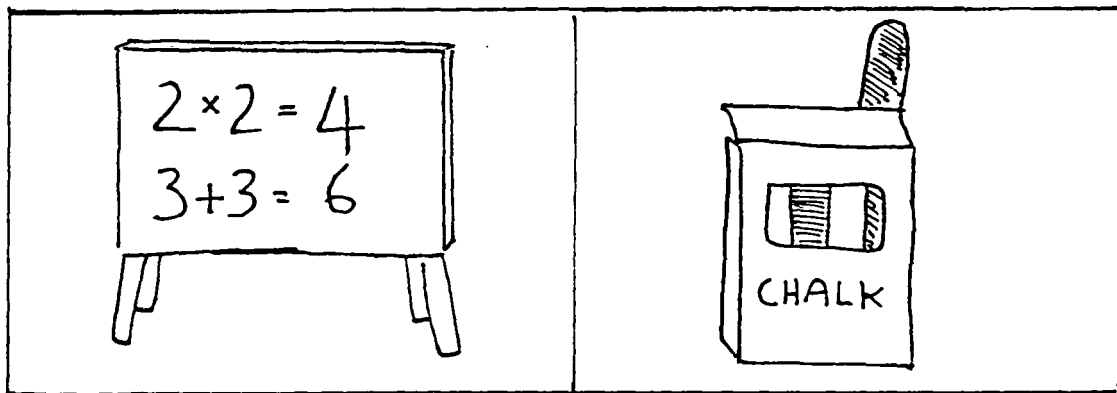
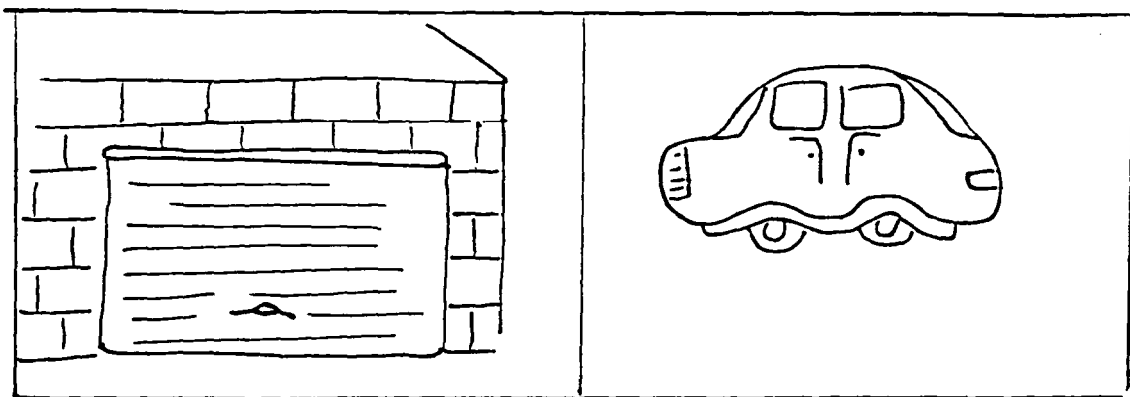


FIGURE 12.1 PICTURES ON CARDS AS EXPERIMENTAL MATERIALS

continued....



counterbalanced across both age and gender of subjects. Thus, each subject saw 2 'same-gender' dolls of different colours (one red and one green doll). As experiment 9 showed that gender provided the minimum referential information, it was decided that dolls should be of the same gender so that reference which only included gender information would always be linguistically ambiguous.

Subjects were required to pick up toys as instructed by the experimenter and to give one to each of the two dolls. Then the subject was told to give an additional toy to only one of the dolls. The type of information given at this point was manipulated across the two main variables of the experiment.

These were as follows:

1. Linguistic: REFERENTIAL SPECIFICITY
2. Non-linguistic: PRAGMATIC ASSOCIATIONS

The variable 'Referential Specificity' manipulated whether the experimenter provided the listener with specific reference to one of the dolls, (definite descriptions which included a colour modifier, 'the red boy....'), or whether the experimenter failed to give definite descriptions that specifically referred to a referent (providing a definite noun phrase without a colour modifier, 'the boy...').

The variable 'Pragmatic Associations' manipulated whether the pragmatic context allowed the listener to make a plausible inference to decide which doll was being referred to (for example, if the green girl had a garage and the red

girl had a bucket, then it must be the RED girl who also played with the spade. This choice was made by inferring that the bucket is associated with the spade). In other cases, the non-linguistic context failed to allow the listener to make a plausible inference (for example, neither of the girls above were more likely to have also played with the chalk).

There were 5 conditions, each defined according to the variables of 'Referential Specificity' and 'Pragmatic Associations'. Each condition was varied only according to the type of sentence used by the experimenter when referring to just one of the dolls picking up a second toy. Given the context of a RED girl with a BLACKBOARD and a GREEN girl with a BIRD, the 5 conditions were varied in the second sentence as follows:

(+R+P) REFERENTIAL SPECIFICITY AND PRAGMATIC ASSOCIATIONS: in this condition, there was both specific reference to one doll and pragmatic cues provided by the context biasing a choice to the same doll. For example, 'The RED girl also plays with the CHALK'.

(+R-P) REFERENTIAL SPECIFICITY ONLY: For this condition, the experimenter made specific reference to one of the dolls, but the context failed to provide any pragmatic cues (toys could not be related). For example 'The RED girl also plays with the spade'.

(-R+P) PRAGMATIC ASSOCIATIONS ONLY: For this condition, the experimenter failed to make specific reference to one of the dolls but the context provided pragmatic cues (2 toys could be related). For example, 'The girl also plays with the CHALK'.

(-R-P) AMBIGUITY CONDITION: The experimenter failed to make specific reference and the non-linguistic context failed to provide any pragmatic cues. Hence, in this condition, the subject was unable to use either linguistic or non-linguistic information to choose one of the dolls and thus, reference was completely ambiguous. For example 'The girl also plays with the spade'.

(R OR P) CONFLICT CONDITION: The experimenter made specific reference to one of the dolls, but the context provided pragmatic cues, biasing a listener's choice to the other doll. For example 'The RED girl also plays with the BIRD', (with 'bird' matching the nest, which is currently in the possession of the green girl').

The experiment was a repeated measures design and therefore, each subject received all of the above 5 conditions.

The procedure was as follows:

Phase 1: Familiarisation Period

Each subject was seen individually. The experimenter and subject were seated at the opposite sides of a table.

The experimenter presented the tray of randomly assorted picture cards on the table, with all the objects facing the subject. The subject first had to name all the objects and to hand them to the experimenter one by one. Then the experimenter presented either 2 male or 2 female dolls to the subject. The subject was told to point to each doll as it was named by the experimenter. The experimenter named each doll by mentioning both colour and gender information (for example, 'This is a red girl and this is a green girl'). This was to ensure that subjects were able to identify the dolls on the basis of these characteristics.

Phase 2: Instructions

The instructions were then given to the subject, who was told to pretend that the two dolls were real children and that the card pictures were real toys. They also had to pretend that the children were playing with some toys and that the experimenter would tell them which toys each child was playing with.

Finally, the experimenter said, 'sometimes I will say something which is silly, because you won't know which boy/girl that I am talking about. If I say a silly sentence, I want you to tell me that it is silly and then I will give you another one'. This point was made very clear to the subject and the experimenter checked to make sure that the subject had fully understood the instructions by subsequently asking questions such as, 'So what are you going to do if you cannot choose one of the dolls'?

Phase 3: Reference to both dolls

For each trial, the experimenter first delivered a pair of introduction sentences and the toys mentioned were varied across each trial. For example:

'One of the girls (boys) picks up a (nest).'

'The other girl (boy) picks up a (blackboard).'

The subject was required to pick up each toy mentioned and to place each one in front of a separate doll.

Phase 4: Reference to one doll only

For the second part of each trial, the experimenter delivered a further sentence which required the subject to pick only one of the dolls to play with a second toy. After the sentence, the subject had to say whether it was possible to choose the doll the experimenter was talking about or whether the sentence was 'silly'. If the subject said that it was possible to pick out a doll, s/he had to place the mentioned toy beside the doll that s/he thought was being referred to. The second part varied according to the 5 experimental conditions.

When specific reference was made, the experimenter used a colour modifier and when pragmatic associations were possible, the toy mentioned related to a toy that one of the dolls had already picked up.

Phase 5: Justifications

For the final part of each trial, the experimenter said 'Good! Can you tell me why you gave the (toy) to this

boy/girl rather than to this one'? This question was intended to look at whether subjects could justify their choice of a particular referent. The question was always preceded by the exclamation 'Good', for the reason that subjects may have thought that they had selected the wrong doll and that the justification question was really a second chance at choosing correctly. If a subject had said that the sentence was 'silly' and did not choose one of the dolls, the experimenter instead asked, 'Will you tell me why you think the sentence was silly'?

At the end of a trial, the toys were placed back on the tray and the experimenter said 'Now they are going to play a different game'.

At first, a subject was given 2 practice trials corresponding to the +R+P condition. This was to check that they understood what they were required to do. If and when this performance was satisfactory, the subject was presented with 4 trials for each of the 5 conditions, thus making a total of 20 trials per subject.

To prevent carry over effects or confounding, all conditions were presented randomly, with the 4 trials for each condition being separated by at least 3 intervening trials. For the second part of each trial, for +R-P, +R+P and R OR P conditions (those involving specific reference), the colour of the doll referred to was counterbalanced across the 4 trials for each condition. In the -R+P condition, the doll which was most plausible was also counterbalanced across trials.

(This was unnecessary again in the +R+P condition, because the plausible referent had to be the same as the one specifically referred to and it was unnecessary in the R OR P condition because the plausible referent had to be the one that was not specifically referred to).

Scoring

For each subject, the experimenter recorded whether they were initially able to name the dolls and toys correctly and whether they always selected the correct objects.

For the second part of each trial (reference to only one of the dolls), the response of a subject was recorded as one of 4 possibilities. These were:

1. saying the sentence was 'silly'.
2. choosing the doll specifically referred to.
3. choosing the most plausible doll.
4. choosing a doll neither specifically referred to nor plausible.

For justification responses, each justification was scored into one of the following categories:

REFERENTIAL: eg. 'because this one was actually said by the experimenter'.

PRAGMATIC: eg. 'because this toy matches that one'

OTHER PRAGMATIC: any other pragmatic inference not manipulated in the experiment.

PHYSICAL: colour, pattern or size matching (for example, 'the doll is the same colour as the toy', 'the doll has two toys of the same colour, (shape or size)', or 'the doll only had a smaller toy and therefore, I gave her a bigger toy'.

MESSAGE: (only if the subject said the reference was silly). Saying that the experimenter had not given enough information or any justification which actually referred to the message or the ambiguity of reference.

NO JUSTIFICATION: Unable to explain a response.

OTHER: Any other justification that cannot be put into a category.

Returning to the general hypotheses of the experiment, it was predicted that all children should be able to use pragmatic associations and justify their answers with pragmatic inferences. Moreover, they should have no problems when the message is linguistically unambiguous: when specific reference is made. However, it was predicted that there would be a developmental improvement in the ability to say that a message was 'silly' when reference was both pragmatically and linguistically ambiguous, as well as a developmental increase in the ability to justify this kind of non-response as due to the message failing to provide sufficient information. Finally, it was expected that younger subjects would show a tendency to justify by

non-linguistic factors whenever it was possible to do so, even when reference was specific.

RESULTS

Firstly, all subjects were able to name the dolls correctly: No subject had trouble reporting the gender and colour of the dolls. Additionally, all subjects were able to recognise all the toy objects illustrated on the cards and no subject selected a wrong card, without making a hasty correction, at any time throughout the experiment.

1. DISCRIMINATIVE CHOICE SCORES.

Table 12-A shows the mean choice scores for all conditions of Referential Specificity and Pragmatic Associations, except for the R OR P (conflict) condition: Thus, the table shows age means for the following conditions, +R+P, +R-P, -R+P, -R-P.

<u>CORRECT CHOICES</u>					
MEAN SCORES	(max. score = 4)				AGE
<u>AGE GROUP</u>	<u>+R+P</u>	<u>+R-P</u>	<u>-R+P</u>	<u>-R-P</u>	<u>MEANS</u>
3-5	4.0	2.94	2.75	1.25 (2.75)	2.74
6-8	4.0	3.81	3.19	1.63 (2.38)	3.16
9-11	4.0	4.0	3.38	2.44 (1.56)	3.46
PARENT	4.0	3.94	3.19	3.38 (0.63)	3.63
<u>MEANS</u>	<u>4.0</u>	<u>3.67</u>	<u>3.13</u>	<u>2.18</u>	

TABLE 12-A: MEAN CHOICE SCORES FOR +/- PRAGMATIC ASSOCIATIONS AND +/- REFERENTIAL SPECIFICITY CONDITIONS.

(For the -R-P condition, the scores in brackets are for choosing one of the referents, despite the fact that reference was ambiguous: 'PERFORMATIVE BIAS').

Clearly, all subjects chose the correct doll (the one both specifically referred to and the one most plausible) in the +R+P condition. Scores were not far behind in the +R-P condition, but scores were slightly less overall in the -R+P condition. The scores for the -R-P condition were clearly the poorest overall. A difference between the scores in +R and -R conditions was found to be significant ($F=111.01$, $df=1,60$, $p<0.001$) and a significant difference was also found between +P and -P conditions ($F=22.87$, $df=1,60$, $p<0.001$). There was also a significant interaction between +R/-R and +P/-P conditions ($F=5.14$, $df=1,60$, $p<0.05$), due to the fact that the differences between +R and -R were slightly larger than the +P/-P differences.

Overall, there was an improvement in scores with age and this was confirmed by analysis of variance ($F=11.28$, $df=3,60$, $p<0.001$).

There are clearly no age differences in the +R+P condition. In the +R-P condition the 3 to 5 year olds' scores were slightly poorer than scores for the older groups and analysis of variance did find a significant effect of age for +R-P scores ($F=6.16$, $df=3,60$, $p<0.001$), which Newman-Keuls showed was due to the differences between all older groups against the 3 to 5 year old group.

The scores in the -R+P condition were similar for all

ages and analysis of variance showed no significant age differences in the -R+P condition ($F=0.75$, $df3,60$, NS).

For the -R-P condition, there was clearly a developmental increase in correct responses (SILLY SENTENCE) and thus a developmental decrease in PERFORMATIVE BIAS. Analysis of variance showed a significant age by response ('silly sentence'/performative bias') interaction in the -R-P condition ($F=7.46$, $df3,60$, $p<0.001$).

Finally, an age by +R/-R interaction was found ($F=2.86$, $df=3,60$, $p<0.05$) and an age by +P/-P interaction was also found to be significant ($F=4.59$, $df=3,60$, $p<0.01$). This result is due to the fact that younger groups produced larger effects owing to their poorer correct scores in the -R-P (ambiguity) condition. However, a 3-way interaction between age, +R/-R and +P/-P was not found to be significant ($F=1.60$, $df=3,60$, NS).

Table 12-B presents mean choice scores for the R OR P (conflict) condition, where a score was given either for choosing the doll explicitly referred to (REFERENTIAL SPECIFICITY), or for choosing the doll which was most plausible due to the association of the toys (PRAGMATIC ASSOCIATIONS).

MEAN SCORES (max. score = 4)			
AGE GROUP	REFERENTIAL SCORES	PRAGMATIC SCORES	AGE MEANS
3-5	3.06	0.94	2.0
6-8	3.81	0.19	2.0
9-11	3.94	0.06	2.0
PARENT	3.75	0.25	2.0
OVERALL MEANS	3.64	0.36	

TABLE 12-B: MEAN CHOICE SCORES FOR R OR P (CONFLICT CONDITION). SCORES FOR DOLL EITHER SPECIFICALLY REFERRED TO OR MOST PLAUSIBLE.

It is clear from the table that most subjects frequently chose the dolls which were specifically referred to, in all age groups.

Analysis of variance found a significant difference between REFERENTIAL and PRAGMATIC scores in the R OR P condition ($f=201.43$, $df1,60$, $p<0.001$).

REFERENTIAL SPECIFICITY choices were very similar to those in the +R-P condition (see table 12-A). Indeed, comparing REFERENTIAL choice scores between these two conditions failed to yield a significant result ($F=0.12$, $df1,60$, NS). Nor was there found to be an age by '+R-P' verses 'R OR P' condition interaction for REFERENTIAL choice scores ($F=0.52$, $df3,60$, NS).

2. JUSTIFICATION SCORES

Tables in this section show proportional scores because any differences in justification scores for a certain choice response is a function of the total number of subjects who made that response.

A. Referential Specificity and Pragmatic Associations (+R+P).

Justification Scores for the +R+P condition are shown in table 12-C. Most scores involved making a PLAUSIBLE inference, but parents used twice as many REFERENTIAL justifications as PLAUSIBLE inferences.

MEAN SCORES	(overall age total = 1.0)		OTHER JUSTIF. OR
AGE GROUP	REFERENTIAL	PRAGMATIC	NO JUSTIF.
3-5	0.25	0.70	0.05
6-8	0.13	0.86	0.02
9-11	0.33	0.67	0
PARENT	0.64	0.36	0

TABLE 12-C: JUSTIFICATION SCORES FOR +R+P, WHEN CHOICES WERE TO DOLL SPECIFICALLY REFERRED TO AND MOST PLAUSIBLE.

Analysis of variance showed that there was a significant difference between these two justifications ($F=12.85$, $df1,60$, $p<0.001$), as well as an age by justification interaction ($F=6.05$, $df3,60$, $p<0.001$).

B. Referential Specificity only (+R-P).

Justification Scores for the +R-P condition are presented in table 12-D. It is clear that the most popular justification for the correct choice was to mention the fact that the doll chosen was clearly referred to by the experimenter (REFERENTIAL). However, the 3 to 5 year olds produced less of these justifications than other groups and over a third of their justifications were PHYSICAL (for example, 'because this one is green and the car is green', 'because this one only has a small toy, so I gave him a bigger one.' etc.). At all ages OTHER PRAGMATIC justifications were rare.

MEAN SCORES (age tot.=1.0)		OTHER		
AGE GROUP	REFERENTIAL	PRAGMATIC	PHYSICAL	NO JUSTIF.
3-5	0.49	0.09	0.36	0.06
6-8	0.83	0.08	0.09	0
9-11	0.98	0.02	0	0
PARENT	0.95	0.05	0	0

TABLE 12-D: JUSTIFICATION SCORES FOR +R-P, WHEN CHOICES WERE TO DOLL SPECIFICALLY REFERRED TO.

Analysis of variance comparing only REFERENTIAL and OTHER justifications showed a significant difference between the use of these justifications when all age groups were taken together ($F=93.69$, $df1,58$, $p<0.001$). Also, an interaction between age group and type of justification

(referential or other) was found to be significant (F=12.29, df3,58, p<0.001).

C. Pragmatic Associations only (-R+P).

Justification Scores for the -R+P condition are presented in table 12-E. The table shows that all age groups preferred PRAGMATIC inferences to explain the reasons for choosing the most plausible doll.

MEAN SCORES	(overall age total = 1.0)		NEITHER OR
AGE GROUP	PRAGMATIC	PHYSICAL	NO JUSTIF.
3-5	0.83	0.04	0.13
6-8	0.87	0.11	0.02
9-11	0.98	0.02	0
PARENT	1.0	0	0

TABLE 12-E: JUSTIFICATION SCORES FOR -R+P, WHEN CHOICES WERE TO MOST PLAUSIBLE DOLL.

Again, analysis of variance failed to find a significant effect of age in the use of pragmatic inferences (F=2.14, df3,55, NS).

D. Ambiguity condition (-R-P).

Justification scores for correct responses (SILLY MESSAGE) in the -R-P condition are shown in table 12-F. It can be seen that most scores were OTHER PRAGMATIC (clearly inferences of their own as they were not manipulated in the

experiment). There was also a developmental decrease in NO JUSTIFICATION scores, and finally, older groups, as well as using OTHER PRAGMATIC justifications, gave a fair proportion of MESSAGE justifications (mentioning that the experimenter had given them insufficient information), especially in the oldest two groups.

MEAN SCORES (overall age total = 1.0)

AGE GROUP	MESSAGE	OTHER PRAGMATIC	NO JUSTIFICATION.
3-5	0.0	0.58	0.42
6-8	0.29	0.50	0.21
9-11	0.43	0.47	0.10
PARENT	0.36	0.57	0.07

TABLE 12-F: JUSTIFICATION SCORES FOR AMBIGUITY CONDITION (-R-P) WHEN CORRECT RESPONSE (SILLY SENTENCE).

Comparing all justification categories, a significant difference in the use of these justifications was found ($F=5.47$, $df2,90$, $p<0.01$). However, no age by justification interaction was found to be significant ($F=1.45$, $df6,90$, NS).

E. Conflict condition (R OR P).

Looking at scores for the justification of choices in the R OR P condition, table 12-G shows that, for the majority choices (those scores which were REFERENTIAL), justifications were mainly REFERENTIAL,

except for the youngest group, where a fair proportion of justifications were PRAGMATIC and PHYSICAL.

MEAN SCORES AGE GROUP	(overall age total = 1.0)			NEITHER OR
	REFERENTIAL	PRAGMATIC	PHYSICAL	NO JUSTIF.
3-5	0.45	0.26	0.27	0.02
6-8	0.76	0.10	0.11	0.02
9-11	0.85	0.05	0.06	0.03
PARENT	0.95	0.05	0	0

TABLE 12-G: JUSTIFICATION SCORES FOR R OR P WHEN CHOICES WERE TO DOLL SPECIFICALLY REFERRED TO.

Analysis of variance found a significant difference between REFERENTIAL and PRAGMATIC justifications ($F=125.3$, $df1,57$, $p<0.001$) and an age by justification interaction was also found to be significant ($F=7.614$, $df3,57$, $p<0.001$).

DISCUSSION

Although their performance was slightly poorer than for older subjects, 3 to 5 year olds were reasonably good at choosing a referent that was referred to 'specifically'. However, a fair proportion of the youngest group justified a correct choice by resorting to physical justifications, as if they were unaware that explicit reference made by the experimenter had led them to choose the correct referent and thus, they looked for some non-linguistic explanation for their response. All older groups however, clearly

preferred to mention that the doll was chosen because the experimenter had just told them to choose that one and thus, they were aware that the information provided by the speaker had led them to choose a specific referent.

Subjects were very good at making pragmatic inferences and were able to justify their choice of referent accordingly. Therefore, the hypothesis that even from 4 years of age, children are very capable of making plausible inferences, is supported.

When reference was both lexically and pragmatically unambiguous (+R+P), all age groups, except parents, justified their choice by making a pragmatic inference, suggesting that plausibility of reference is more significant to children than is the structure of the sentence used to make reference. On the other hand, parents showed a greater preference for referential justifications, which suggests that only for adults did the message itself play a more significant role than non-linguistic contextual factors.

However, when specific reference conflicted with plausibility, at all ages, subjects preferred to choose the referent which was 'specifically' referred to. Nevertheless, whilst all older groups justified this response by saying that it was the doll whom the experimenter had told them to select, the 3 to 5 year olds tended either to revert to making the plausible inference about the doll whom they had not selected, or gave other non-linguistic, physical justifications.

Thus, whenever a referent was 'specifically' referred

to, the youngest group had no trouble in choosing the correct doll. However, pragmatic factors were more important for them, and as it seems that they were not aware that their choice was influenced by the verbal information given to them by the experimenter, they gave non-linguistic contextual justifications.

Complete ambiguity of both the linguistic and the pragmatic information resulted in a developmental increase in saying that the message was 'silly', suggesting that younger subjects did in fact, show a performative bias: They preferred to guess rather than to not respond at all. Whenever younger subjects did complain about the message, they had problems trying to justify their response: They either failed to give any justification at all or they used a plausible inference of their own. There was a developmental increase in the ability to justify the failure to choose a referent by mentioning problems with the actual message. There was also a developmental decrease in not being able to give any justification at all. Thus, it seems that, when reference cannot be understood by using the pragmatic context, young children are unable to recognise ambiguous definite reference.

6.4 EXPERIMENT 13: Comprehension of pronominal reference as a function of lexical and pragmatic ambiguity.

INTRODUCTION

The experiment to follow was an attempt to overcome the potential methodological flaws of previous studies on children's comprehension of pronouns. A comprehension task was used similar to the one used by Wykes. Both lexical and pragmatic ambiguity were manipulated in order to test children's use of pragmatic inferences and lexical constraints. Also, the experiment aimed to find out the age that children could detect ambiguity. Pairs of sentences were presented, with the pronoun mentioned in the second sentence. The order in which referents were mentioned in the first sentence was also manipulated, to see if there were any developmental differences in using particular strategies and to see if they were used only when no other information (pragmatic or lexical) was available.

It was predicted that infants would be able to make pragmatic inferences and would also be sensitive to lexical constraints. Moreover, it was predicted that pragmatic inferences would serve as the main factor in young children's choices of antecedent, whereas older subjects would be more concerned with linguistic factors, as reflected by an ability to detect referential ambiguity.

METHOD

Subjects

48 children, 16 in each age group, were selected from a Durham primary school, to act as subjects. None of these children had participated in the previous comprehension experiment. There were three age groups: 3 to 5 years [mean age: 4.4], 6 to 8 years [7.4] and 9 to 11 years [10.3] and there were approximately an equal number of males and females in each group. All children were judged to be of average ability. 16 parents also acted as subjects (12 mothers and 4 fathers) and they had children in the school whose ages varied between 3 and 11 years.

Materials

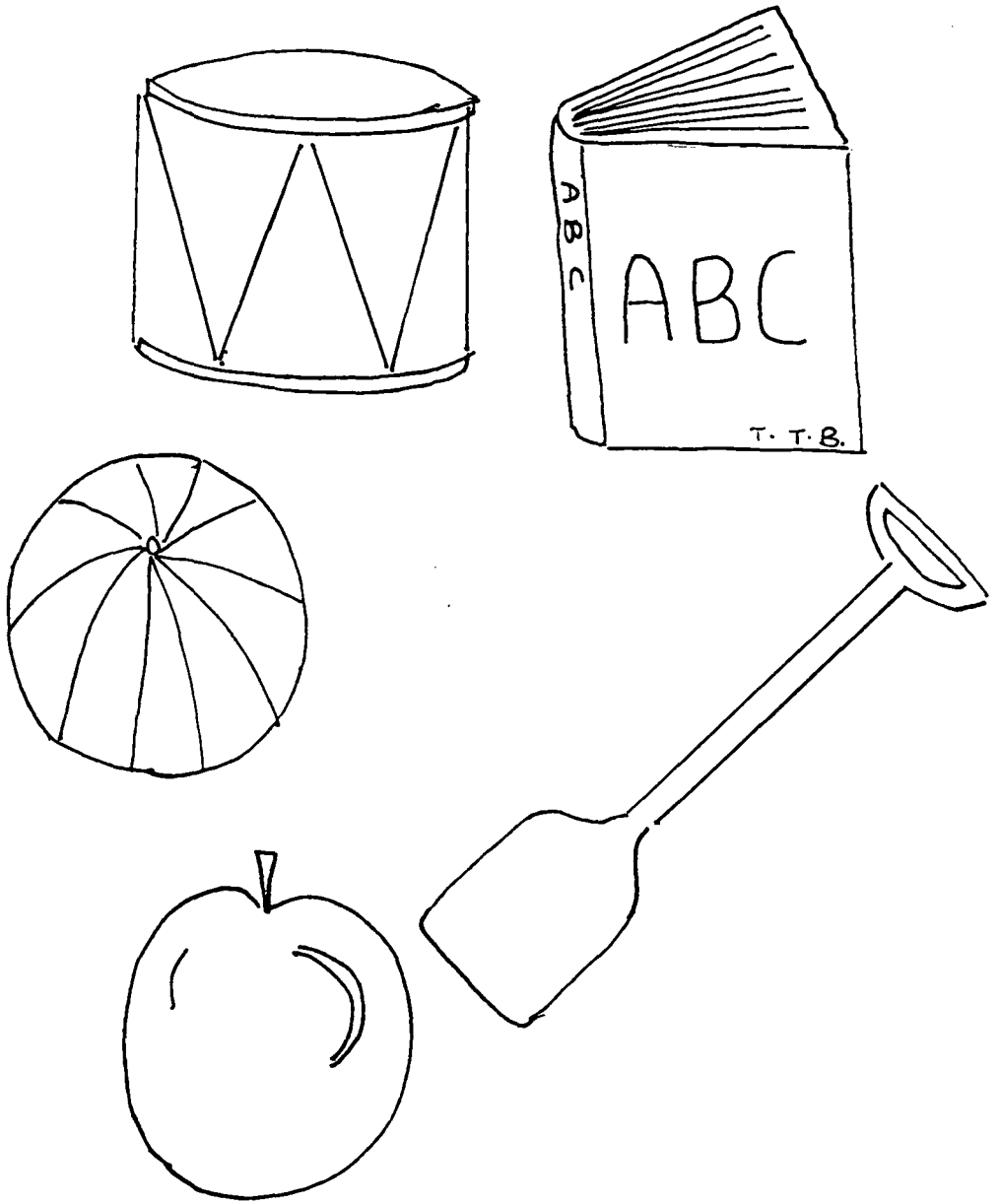
4 dolls (2 male and 2 female) and 5 cardboard shapes depicting individual objects: BOOK, SPADE, BALL, APPLE and DRUM served as the experimental materials (see figure 13.1). There were also 5 shallow boxes, 4 of which had names printed on in large letters: These represented the names of each of the dolls (Susan, Jane, Peter and John). The fifth box had the word 'somebody' printed on it.

Sentence-pairs were written on individual cards to be used as the experimental stimuli. Examples are presented in the Design and Procedure section (see below).

Design and Procedure

The overall task involved presenting the child

FIGURE 13.1 PICTURES ON CARDS AS EXPERIMENTAL MATERIALS



with 3 dolls: two dolls of one gender and one doll of the other gender. For each trial the experimenter read an introductory sentence (S1), which required the subject to pick up two objects and to put each one into one of the boxes. A second sentence containing the pronoun was then read (S2) and this was followed by a question about who the pronoun referred to, to which the child was required to respond by choosing one of the boxes.

The sentence-pairs varied across the two main variables of the experiment. These variables were:

1. Lexical marking for gender: LEXICAL CONSTRAINT.
2. Plausibility of the referent: PRAGMATIC ASSOCIATIONS.

The variable 'lexical constraint', depended on whether the first sentence was about same-gender or different-gender referents. For the former, the pronoun in the second sentence was lexically ambiguous, but for the latter, the pronoun lexically disambiguated reference through gender information.

The variable 'pragmatic associations', depended on whether the action described of the referent in the second sentence was pragmatically related to the object given to one of the dolls from the first sentence (plausible) or whether the described action was unrelated to either of the objects given to each of the dolls (neutral).

Altogether there were 5 conditions and a description of

each is presented below:

(+L+P): LEXICAL AND PRAGMATIC - S1 referred to different-gender dolls, so that lexical gender marking in S2 constrained choice of referent AND this referent was the most plausible, due to the association between the action mentioned in S2 and the object given to this doll after S1. For example:

S1. SUSAN picks up a BALL and JOHN picks up a spade.

S2. Then, SHE GOES OUT TO PLAY IN THE PARK.

(+L-P): LEXICAL CONSTRAINT - S1 referred to different-gender dolls, so that lexical gender marking in S2 constrained choice of referent BUT S2 was pragmatically neutral. For example:

S1. SUSAN picks up a ball and JOHN picks up a spade.

S2. Then, SHE puts it back again.

(-L+P): PRAGMATIC ASSOCIATIONS - S1 referred to same-gender dolls, so that lexical gender marking in S2 did NOT constrain choice of referent BUT S2 described an action which made one referent more plausible, given the object s/he already had in her/his box. For example:

S1. Susan picks up a BALL and Jane picks up a spade.

S2. Then, she GOES OUT TO PLAY IN THE PARK.

(-L-P): AMBIGUITY CONDITION - S1 referred to same-gender dolls, so that lexical gender marking in S2 did NOT constrain choice of referent AND S2 was pragmatically neutral. For example:

S1. Susan picks up a ball and Jane picks up a spade.

S2. Then, she puts it back again.

(L OR P): CONFLICT CONDITION - S1 referred to DIFFERENT-gender dolls, so that lexical gender marking in S2 constrained choice of referent BUT S2 described an action which made the OTHER referent more plausible, given the object s/he already had in her/his box. For example:

S1. SUSAN picks up a ball and JOHN picks up a spade.

S2. Then, SHE STARTS TO DO SOME GARDENING.

For S1, the verb 'picks up' was always used. In S2, for +P conditions (+L+P, -L+P and L OR P), the sentences described actions which were related pragmatically with one of the objects that had been picked up. The sentences used to associate an action with each object were as follows:

1. BALL - S/he goes out to play in the park.
2. SPADE - S/he starts to do some gardening.
3. BOOK - S/he begins to tell a story.
4. APPLE - S/he feels hungry.
5. DRUM - S/he wakes up the neighbours.

For all conditions that involved -P (+L-P and -L-P), S2

was randomly selected from the following three 'neutral sentences'.

1. S/he is going to see a film tomorrow.
2. S/he puts it back again.
3. S/he opens the door of a cupboard.

Before the experiment was run, a pretest was given to 6 infants (mean age = 4.6), 6 juniors (mean age = 7.3) and 6 parents, neither of who took part in the actual experiment. The purpose of the pretest was to ensure that pragmatic associations were genuinely judged to be more plausible by the subjects. They were each presented with single sentences such as "The girl with the drum, started to tell a story", and the subjects had to say whether the sentence 'could happen', was 'unlikely to happen' or 'could never happen'. All objects were presented across the 5 possible pragmatic sentences and for each neutral sentence. A score of 2 was recorded for 'could happen', 1 for 'unlikely to happen' and 0 for 'could never happen'.

It was clear that, for each age group, the action predicted to be associated with a particular object achieved higher scores than sentences that were not predicted to be associated with that object, but for the 'neutral sentences', it was clear that pragmatic associations between all sentences and all objects were judged as unlikely or impossible, as scores for these sentences were very low. The scores for each age group are pooled in table 13 below:

(SCORES OUT OF 36)		SENTENCES					:	NEUTRAL		
OBJECTS	1	2	3	4	5	:	N1	N2	N3	
1	*34*	6	16	2	17	:	6	7	11	
2	22	*32*	19	6	16	:	12	16	11	
3	18	8	*36*	9	8	:	8	12	5	
4	6	8	12	*34*	4	:	2	9	13	
5	6	7	4	13	*36*	:	10	8	10	

TABLE 13: PRETEST RATINGS ACROSS ALL SENTENCES WITH EACH OBJECT.

For the actual experiment, the objects were all selected randomly, and the order in which dolls were mentioned in S1 was also selected randomly for each trial.

The child was required to perform S1 as well as S2 to ensure that any problems in answering the question were not due to memory failure. A repeated-measures design was used and thus, each subject was presented with 20 experimental trials (4 trials in each of the five conditions) and all trials were presented in a random order.

In each condition, except for -L-P (as there was no correct antecedent in this condition), the 'correct' pronominalised referent was either the first or the second referent to be named (two trials as first mentioned and two trials as second mentioned referent). This was to control for the possibility that choice of a referent may be a reflection of a particular strategy preference, such

as choosing the most recently mentioned referent (minimal distance strategy).

The experiment was run in the school reading room which was familiar to all the children. The procedure was as follows:

Phase 1: Introduction to materials

Each subject was seen individually. The child was presented with 3 dolls, who were each positioned below the box on which their name was printed. The 'somebody' box was also presented. Whether subjects received the 3 dolls, Susan, John, Jane OR Peter, John, Jane was counterbalanced across the age and gender of the subject. They were asked to name each doll and were told that the 'somebody' box did not belong to anybody in particular and could belong to any of the dolls. They then had to name the 5 objects on the tray. The experimental layout was as shown in figure 13.

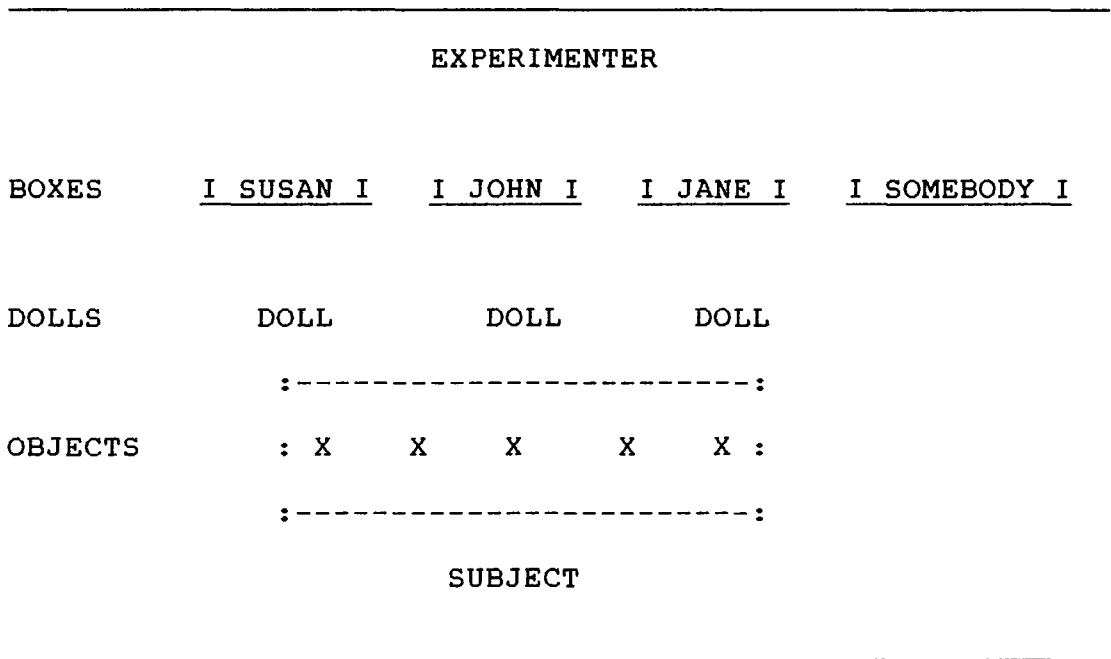


FIGURE 13: EXPERIMENTAL LAYOUT

Phase 2: Instructions

Only when the child could name the dolls and objects correctly were they given the instructions. Each subject was told to pretend that the dolls were real people and that the objects were real objects. They were then told that the experimenter would ask them to pick up an object and to put it into one of the boxes. Subjects were then given some practice with S1 sentences. The experimenter then said that she would go on to talk about only one person and that the subject would be asked to choose the box of the person that they thought was being referred to. 2 sentence-pairs (+L+P type) were first given as practice trials. For example:

S1. Susan picks up a book and John picks up a ball.

S2. She starts to tell a story.

Question. Who starts to tell a story?

The subjects were also told that if they thought that it was impossible to choose one particular doll, they were to point to the 'somebody' box. To check that subjects understood this instruction, the experimenter finally asked, "So, what will you do if you cannot choose one of the dolls"?

Phase 3: Running the experiment

When subjects had shown that they understood the experimental requirements, they were presented with the 20

experimental trials. These were read with neutral intonation and without any emphatic stress on the pronoun.

Scoring

The experimenter recorded the responses given to each question. Scores for each condition were as follows:

+L+P: Correct score was doll whose gender matched pronoun and who was most plausible. Any other response was incorrect. If a child pointed to the 'somebody' box, this was recorded separately.

+L-P: Correct score was doll whose gender matched pronoun. Any other response was incorrect. If a child pointed to the 'somebody' box, this was recorded separately. (In this condition, if a child chose a doll not mentioned in S1, but who matched the pronoun for gender, then a correct score was given).

-L+P: Correct score was doll who was most plausible. Any other response was incorrect. If a child pointed to the 'somebody' box, this was recorded separately.

-L-P: Correct score was to choose the 'somebody' box, as reference was both pragmatically and lexically ambiguous.

L OR P: An L score was given if a doll was chosen whose gender matched the pronoun. (If chose referent not mentioned in S1, but which matched the gender of the

pronoun, an L score was also given). A P score was given if the child chose the most plausible doll. Any choice of a referent not matching the pronoun gender and not mentioned in S1 was an incorrect response. If a child pointed to the 'somebody' box, this was recorded separately.

When a particular doll was selected which had been mentioned in S1, the sentence position of the selected doll (first or second mentioned) was also recorded by the experimenter.

Finally, any incorrect choices of a named referent or object in S1 were also noted.

RESULTS

After the familiarisation procedure, all subjects were able to choose the correct objects and the correctly named dolls.

When making a choice after the pronominalised sentence, if a subject chose the doll that had not been mentioned in the first sentence, they received a score in the +L-P condition and they received a 'lexical' score in the 'L OR P' condition when the gender marking of the pronoun matched the gender of the chosen referent. However, choosing the unmentioned doll in any other condition or when gender did not match lexical marking, an incorrect response was recorded. However, only one subject (a 4 year old), tended to choose the referent that was not named in the first sentence, and did so irrespective of condition and gender marking.

1. Choice of first or second named referent.

To see whether subjects used a particular strategy when choosing a referent after the pronominalised sentence (choice depends on the order in which dolls were named in the first sentence), table 13-A presents mean choice scores for the first and second named dolls in each condition, for each age group.

Mean score out of 4.

<u>AGE</u>	<u>+L+P</u>	<u>+L-P</u>	<u>-L+P</u>	<u>-L-P</u>	<u>L OR P</u>	<u>MEAN</u>
<u>CHOOSING FIRST MENTIONED REFERENT</u>						
3 - 5	2.0	1.88	1.94	0.56	2.0	1.7
6 - 8	2.0	2.13	1.94	1.19	2.0	1.9
9 - 11	2.0	1.94	1.88	0.56	1.33	1.5
Parent	2.0	1.88	2.06	0.14	1.89	1.6
<u>MEAN</u>	<u>2.0</u>	<u>2.0</u>	<u>2.0</u>	<u>0.6</u>	<u>1.8</u>	<u>1.7</u>
<u>CHOOSING SECOND MENTIONED REFERENT</u>						
3 - 5	1.88	1.88	1.94	2.69	1.88	2.1
6 - 8	2.0	1.88	2.06	2.25	1.73	2.0
9 - 11	2.0	2.06	2.13	2.56	1.11	2.1
Parent	2.0	2.13	1.94	1.86	2.0	2.1
<u>MEAN</u>	<u>2.0</u>	<u>2.0</u>	<u>2.0</u>	<u>2.3</u>	<u>1.7</u>	<u>2.1</u>

TABLE 13-A: CHOOSING FIRST OR SECOND MENTIONED REFERENT.

At all ages, there were more choices for the last mentioned referent than for the first mentioned referent and analysis of variance confirmed that this effect was significant ($F=16.65$, $df=1,60$, $p<0.001$). There was no significant interaction between age and choice of first/last referent ($F=0.91$, $df=3,60$, NS).

It appears that the last referent preference is due to the -L-P condition, at all ages. In fact, a significant interaction between first/last choice and condition was found ($F=23.89$, $df_{4,240}$, $p<0.001$), but a 3-way interaction between age, first/last and condition was not found to be significant ($F=1.61$, $df=12,240$, NS).

2. Discriminative choice scores

Table 13-B presents the mean number of correct choices in each of the first 4 conditions: +L+P, +L-P, -L+P, -L-P. A correct score for each condition was as follows:

+L+P : choose referent as marked lexically by pronoun (male or female gender) AND who is most plausible given the object that has been picked up.

+L-P : choose referent as marked lexically by pronoun (male or female gender).

-L+P : choose referent who is most plausible given the object that has been picked up.

-L-P : choose 'somebody' box.

mean score out of 4.

<u>AGE GROUP</u>	<u>+L+P</u>	<u>+L-P</u>	<u>-L+P</u>	<u>-L-P</u>	<u>MEAN</u>
3 - 5	3.81	3.0	3.81	0.56	2.8
6 - 8	4.0	3.63	3.94	0.56	3.0
9 - 11	4.0	3.94	3.88	2.25	3.5
Parent	4.0	3.88	3.94	3.13	3.7
<u>MEAN</u>	<u>4.0</u>	<u>3.6</u>	<u>3.9</u>	<u>1.6</u>	<u>3.3</u>

TABLE 13-B: DISCRIMINATIVE CHOICE SCORES FOR +-LEXICAL AND +-PRAGMATIC CONDITIONS.

There was clearly a difference in the scores between conditions. The highest overall scores were for the +L+P condition and the scores for the -L+P condition were not far behind. Interestingly, the 3 to 5 year olds' mean score was identical for +L+P and -L+P conditions. At all ages there were slightly less scores in the +L-P condition, particularly for the youngest group and the poorest scores were in the -L-P condition for all ages. A significant difference between +P and -P was found ($F=194.67$, $df=1,60$, $p<0.001$), as well as a significant difference between +L and -L scores ($F=108.29$, $df=1,60$, $p<0.001$). Also, an interaction between +P/-P and +L/-L was significant ($F=113.19$, $df=1,60$, $p<0.001$). This effect was due to the fact that, the +P/-P differences were only significant in the -L conditions and +L/-L differences were only significant in -P conditions.

Overall, there was an increase in scores with age and this was mainly due to the -L-P condition. Also, there was a developmental improvement in the +L-P condition from the youngest group up to the 9 to 11 year old group. Analysis of variance showed a significant effect of age ($F=14.39$, $df=3,60$, $p<0.001$). Moreover, there was a significant interaction between age and +P/-P ($F=16.82$, $df=3,60$, $p<0.001$) and a significant interaction between age and +L/-L ($F=6.23$, $df=3,60$, $p<0.001$) due to the decrease in +P/-P and +L-L differences with increasing age. Finally, there was a significant 3-way interaction between age, +P/-P and +L/-L ($F=7.67$, $df=3,60$, $p<0.001$) and this effect

was due to the fact that the age by +P/-P interaction only applied to -L conditions and the age by +L/-L interaction only applied in the -P conditions.

Table 13-C presents the mean number of choices which were either 'pragmatic', 'lexical' or 'somebody' in the 'L OR P' condition.

mean score out of 4.

AGE GROUP	PRAGMATIC	LEXICAL	SOMEBODY
3 - 5	3.69	0.19	0.0
6 - 8	2.69	0.81	0.50
9 - 11	0.44	0.94	2.63
Parent	0.38	1.81	1.81
<u>MEAN</u>	1.8	0.9	1.2

TABLE 13-C: DISCRIMINATIVE CHOICE SCORES FOR CONFLICT CONDITION (L OR P).

Overall, there was a preference for a 'pragmatic' choice and the least popular choice was a 'lexical' one. Analysis of variance confirmed that there was a significant difference between types of choices ($F=4.66$, $df=2,120$, $p<0.05$).

The strong preference for a 'pragmatic' choice decreased with age. By contrast 'lexical' and 'somebody' choices increased with age. However, 9 to 11 year olds clearly preferred a 'somebody' response, whereas parents

produced the same proportion of 'lexical' choices as 'somebody' choices. No overall age effect was significant ($F=1.0$, $df=3,60$, NS), but a significant interaction was found between age and response ($F=14.19$, $df=6,120$, $p<0.001$).

DISCUSSION

The results showed that all subjects were very good at using the pragmatic relationship between the action of a doll and the toy in the doll's box and were able to make the appropriate pragmatic inferences in assigning the pronoun to a referent.

Although there was some age increase in the use of lexical information for choosing a referent, the majority of responses in the +L-P condition, at all ages, were to the referent which matched the lexical marking of the pronoun for gender.

The findings of the conflict (L OR P) condition showed that, whilst young children can use lexical information, they are more influenced by their knowledge of events in the world, to the extent that they may overuse pragmatic information even when linguistic information points to the contrary.

As for the -L-P condition, it is interesting that this produced the poorest scores at all ages. However, there was a clear developmental increase in the ability to appreciate the ambiguity of reference and to choose the 'somebody' box in this condition. Between 3 to 8 years, there were very few responses to the 'somebody'

box and even from 9 to 11 years, only about 50 percent of responses were of this kind. In fact, only the parent group gave the correct response significantly more than any other response in this condition. Therefore, as for ambiguous definite reference with the definite article, the failure to appreciate ambiguous pronominal reference is developed much later than is the use of other information to decide upon a referent.

Finally, at all ages, those subjects who did choose a specific referent in the -L-P condition, appear to have been using a simple heuristic device: choosing the most recently named referent (cf. Chomsky (1969) 'minimal distance strategy'). This suggests that both children and adults use similar strategies in assigning pronouns and these strategies are only used when there is no other information (lexical and pragmatic) available.

6.5 GENERAL DISCUSSION

It appears that when ambiguous references do not complicate the task for the listener, the comprehension of referential ambiguity is developed along the same lines as is the production of unambiguous definite reference: Whilst young listeners used pragmatic inferences to assign reference, they failed to recognise referential ambiguity.

Moreover, even when young children could use the lexical information carried by pronouns, the development of an understanding of unambiguous definite reference in discourse applied to pronouns as well as definite articles.

Looking first at the results of experiment 12, at all ages, subjects were more likely to base their choice of a referent on the explicit linguistic information provided by the speaker, even when pragmatic information pointed to the contrary. However, when justifying their responses, all child groups preferred to refer to the pragmatic context, especially children under 9 years of age, whereas parents preferred to refer to the linguistic information.

Moreover, when reference was neutral with respect to plausibility, the youngest subjects, even if the correct doll had been selected, looked puzzled when asked to justify their response, suggesting that they could not reflect on the reasons for their choice of doll. Therefore, it seems that specific reference brought about an almost automatic choice from the young listener, but they had difficulty explaining the reasons for their response.

3 to 5 year olds even preferred to invent some

plausible explanation, rather than to justify their response on the basis of the linguistic information. Therefore, young children did not appear to be consciously aware that they were using linguistic information and tried to give plausible explanations even when the pragmatic context was unlikely to be the cause of their response.

When the 6 to 8 year olds chose incorrectly, they were less likely to look for some plausible inference of their own, perhaps because they knew that these inferences would be less plausible than those manipulated in the experiment (associations between toys). This could explain why they looked for other non-linguistic factors, such as physical matching to justify their choice of a referent.

Unlike the children, for those few cases where parents made unexpected pragmatic inferences, it appears from their justifications that these inferences were just as plausible as those manipulated in the experiment. The examples below demonstrate the difference between 'other pragmatic' justifications by 3 to 5 year olds compared with those used by parents:-

3 to 5 years: after choosing the doll with the garage and giving her the spade - (after pausing) 'because Daddy keeps the spade next to the freezer in the garage'

Parent: after choosing the doll with the car and giving him the bucket - (confidently) 'I chose that one because he could be playing at washing the car.'

From these justifications, it seems that the young child used a rather subjective explanation for her choice of referent and pondered over the request for a justification as if trying to think of a connection between the objects, no matter how plausible it was. Alternatively, the parent did not hesitate and the justification given seems to be just as plausible as the association between the bucket and spade (especially for an adult whose experience with buckets is more likely to be related to cleaning than playing in the sand).

A possible reason for why other pragmatic associations were not revealed in the pretest is because the pretest required that each object was paired with one of the other objects. Therefore, even if one object had two plausibly associated objects, the decision as to which should be chosen as a pair depended on how it related to all other objects on the tray. For example, if a subject had paired the bucket with the car, they would have been left at the end with a spade and a garage: two objects which are far less plausibly related.

Young children did not respond appropriately when reference was ambiguous in experiment 12. This finding supports the claim that the development of understanding ambiguity is a gradual process, because this requires the ability to reflect on linguistic information apart from the pragmatic context to which it normally refers. However, the fact that younger subjects tended to show a 'performative bias' (Ackerman 1981) may imply that their failure to detect ambiguity was due to performance factors

rather than due to their problems in appreciating 'specificity'. Moreover, if the predominant response of young children was to choose any doll, then it is possible that they chose a doll when reference was ambiguous not because they were unaware of the ambiguity, but merely because they thought they should make a guess, even if the message was ambiguous.

One potential criticism of this experiment is that the pragmatics only and the conflict conditions were unnatural, (as well as the intended ambiguity condition). In fact, a handful of subjects at all ages said that sentences were 'silly' in the pragmatics only condition. In normal conversation, reference is either in accordance with pragmatic information or it relies solely on the explicit linguistic information for unambiguous reference. However, references with definite articles rarely rely on pragmatic information alone as is sometimes the case for pronominal reference. Therefore, it is important to bear in mind that performance may have been affected by the artificiality of some of the sentences that were used.

Nevertheless, the youngest group often justified their response that pragmatics only references were silly by making the pragmatic inference that they had just failed to show, suggesting that they detected an incompleteness in the message, but they were not sure of the reasons for their failure to respond and therefore, they justified by making the inference as if they HAD chosen the correct doll.

Finally, any conclusions derived from justification

scores must be treated with caution because, it is possible that young children and adults interpreted these requests very differently: Children may have felt as if further questioning was due to their performance being insufficient, whereas adults were probably more likely to treat the questions as explicitly requesting some kind of explanation for their choice of doll.

Turning now to experiment 13, the results failed to support Tyler's conclusion that the problem for young children was in using the relevant lexical information, because young children always correctly interpreted pronoun gender marking. Thus, it seems that the type of tasks used by Tyler were probably a better measure of the young child's speed of performance rather than being a reflection of their understanding of pronouns per se. However, in the conflict condition, 3 to 8 year olds were more likely to use the pragmatic information as opposed to choosing the lexically marked referent. Thus, these findings fail to support Wykes' conclusion that the problem in interpreting pronominal reference for young children is an inferential one.

At all ages, subjects were influenced more by pragmatic factors in experiment 13 compared with experiment 12. However, the type of inferences that were required were different for each of these experiments: In experiment 12, a pragmatic inference meant forming an association between two toys (for example, GARAGE-CAR), whereas, in this experiment, an inference required forming a relationship between an object and an action (for

example, SPADE-GARDENING). It seems that the latter type of inference was more natural than the former, due to the fact that the functional relationship between pairs of toys is more flexible than the function of a particular object per se: It is more likely that a child may play with two different types of toys, but it is absurd to imagine, for example, somebody eating an inedible object.

Thus, in experiment 13, the pragmatic information can probably be accepted as universal, whereas in experiment 12, there was room for personal interpretation and inferences were more likely to depend upon the particular individuals, as well as the age of the individuals making them.

Unlike the younger children, 9 to 11 year olds clearly noticed when there was a conflict because they pointed to the box not related to one particular individual (the 'somebody' box). Parents would also point to the 'somebody' box, but they were equally as likely to choose the referent which was lexically marked. It is possible that, whilst adults were aware of the conflict, many still used the lexical information because they were aware that this information always constrains the assignment of a pronoun, whereas plausibility is more flexible, ambiguous and relative to the particular situation and participants involved.

The results of experiment 13 showed that the young child's problem in detecting referential ambiguity is not due to a 'performative bias' per se, because even when the correct response in the ambiguity condition was to

point to the 'somebody' box, young children preferred to choose a box belonging to a particular individual. Whilst there was a developmental increase in choosing the 'somebody' box in the ambiguity condition, those subjects at all ages who did choose a specific referent in this condition appear to have been using a 'minimum distance strategy': choosing the most recently named referent. This suggests that both children and adults use similar strategies in assigning pronouns and these strategies are only used when there is no other information (lexical and pragmatic) available.

One potential flaw of this experiment is again, the atypicality of the sentence-pairs that were used. There has been experimental work on adults and work from Artificial Intelligence researchers to suggest that the critical factor in determining whether a pronoun is used depends on whether the referent is in the focus of attention. It was intended that the examples used in this experiment would be neutral with respect to discourse topic and theme. However, as individuals were mentioned in separate clauses, it is possible that the first mentioned referent was not in focus and this may account for the second mentioned referent bias in the ambiguity condition.

A possible means of bringing both referents into focus in future experiments would be to mention them both in the same clause, such as "Jane gave Susan the ball".

Altogether these results suggest that, for definite reference using either pronouns or full noun phrases, young children have problems detecting

ambiguity, just as they have difficulties producing unambiguous definite expressions themselves. Thus, knowing that a referent must be 'specific' for definite reference seems to be a general developmental ability that is not yet acquired in young children.

Nevertheless, the findings showed that young children were very good at using lexical constraints and this provides an example of how linguistic knowledge can be acquired independently of knowledge of referential ambiguity. For the first experiment of the next chapter, interest is in the age when children will appreciate when a noun phrase cannot refer to a c-commanding pronoun in a sentence (binding principle C), as well as whether the understanding of syntactic constraints are acquired before the child can actually appreciate that both syntactic and non-linguistic factors need to be considered when interpreting the reference of pronouns. The final 2 experiments in the next chapter were an attempt to look at the way that older children and adults use linguistic and non-linguistic factors for anaphoric reference.

INTERACTIONS BETWEEN LINGUISTIC AND NON-LINGUISTIC FACTORS.

7.1 ABSTRACT

Experiment 14 was concerned with referent choice as a function of pragmatic context and syntax (principle C). Experiments 15 and 16, using only older children (8 to 11 years) and adults, looked at story completions as a function of semantic (number of available referents) and pragmatic (plausibility of referents) contexts. Experiment 16 also aimed to eliminate possible completion biases that were due to the way that the stories were constructed.

Experiment 14 showed that, whilst even young children were influenced by plausibility and were sensitive to principle C of binding theory, only from about 9 years of age would subjects demonstrate an awareness of both a syntactic and a non-linguistic factor influencing reference, as shown by their failure to discriminate between answers to 'linguistic' questions and answers to 'pragmatic' questions when syntax and pragmatics were contradictory. By contrast, young children were biased to only one form of interpretation (either syntax or pragmatics) in the 'conflict' condition.

The results of experiment 15 supported Altmann and Steedman's theory, that anaphoric reference requires an interaction between semantic context and syntactic possibilities, because choice of completion (relative or non-relative) depended on the 'referential support' of the number of available referents (semantics), rather than on

the plausibility of a referent (pragmatics). The findings were replicated in experiment 16, suggesting that results are not an artefact of the performance biases induced by the materials.

Altogether, the results of this chapter suggest that the problem for young children is due to their inability to consider more than one possibility for anaphoric reference (linguistic and non-linguistic). Older children do appear to use both linguistic and non-linguistic factors when processing anaphoric expressions. Moreover, the most significant contextual factor influencing their preference of linguistic construction appears to be the semantic context (number of available referents) as opposed to the pragmatic context (plausibility).

7.2 GENERAL INTRODUCTION

Three experiments are reported in this chapter which are concerned with the way that both linguistic and non-linguistic factors are involved for definite reference. Chapters 4, 5 and 6 showed that young children have problems appreciating 'specificity'. A possible explanation for this difficulty may be that young children are less able to integrate the context of a discourse with the particular linguistic construction that is required and knowledge of 'specificity' for definite reference requires knowing that the type of linguistic construction used depends on the particular contextual information which makes a referent specific. This knowledge is not the same as knowing the linguistic constructions for definiteness

per se. For example, whilst young children can produce relative clauses when memory limitations are controlled, they may not appreciate that this form of definite construction is related to the referential context (whether a referent is specific or not).

Therefore, many investigators have recently concentrated on the interactions which take place between linguistic and non-linguistic factors when processing discourse and it is now widely recognised by various researchers (including some linguists), that it is virtually impossible to look at linguistic parsing out of the non-linguistic and pragmatic context in which it is processed.

Experiment 14 looked at children between 3 and 11 years of age and adults (parents and students), testing their knowledge of principle C of binding theory. Basically, this rule states that a noun phrase is not co-referential with a pronoun/noun phrase which c-commands it.

From the spontaneous remarks made by Balfour's subjects (1983) it appears that only adult subjects were aware of a contradiction when pragmatics was the opposite of syntactically constrained 'non-identity'. Thus, it seems that young children were not aware that both linguistic and non-linguistic factors were involved in assigning pronominal reference. Therefore, for experiment 14, it is predicted that, even when children have no trouble obeying principle C and when they can make pragmatic inferences to assign reference, they will fail to

show an awareness of the presence of both factors when the pragmatic context conflicts with syntax.

However, the type of non-linguistic information which is thought to be relevant for anaphoric processing differs between investigators: Whilst Marslen-Wilson and Tyler (1982) were primarily interested in the way that pragmatic context influenced reference, Crain and Steedman (1985) / Altmann and Steedman (1988) argued that the actual number of available referents (semantic context) was significant in influencing choice of definite construction.

Nevertheless, if linguistic and non-linguistic interactions are a reflection of 'specificity', then it is possible that both these types of non-linguistic context are significant for choosing the appropriate definite construction: The number of referents available is important for determining whether further presuppositional information is required to make a referent specific. However, inferences can also narrow reference down to one or more potential antecedents. For example, given the sentences:

Bill was watching Harry on the roundabout.

Then he decided to have a turn.

Using the pragmatic inference that, 'a turn' means going on the roundabout, plus the fact that Harry had already had a turn, one could infer that Bill is the antecedent of the pronoun because, although there are 2 available referents, one is clearly more plausible than the other.

Thus, it would appear that producing definite reference requires the following type of processing and not necessarily in this particular order: A. Ruling out those constructions which are unacceptable syntactically (eg 'The man' cannot refer to an antecedent which c-commands it) and lexically (eg 'She' cannot refer to a male). Thus, one is left with a number of definite constructions which are acceptable linguistically. B. Rating the 'specificity' of referents: whether other potential referents are available and C. Rating the plausibility of referents: whether there are other plausible referents. In other words, syntactic and lexical factors, semantic factors (discourse model) and pragmatic factors must all interact with each other to ensure successful definite reference. The interaction may be a weak one as suggested by Crain et al, where context determines which of the alternative linguistic constructions to rule out or it may be a strong one where context determines the choice of construction.

Experiment 15 looks at the way that older children and adults choose to complete mini-stories after presented with a word which is ambiguous with regards to the type of linguistic construction which could follow it. For example, 'that' may be followed by a relative clause (that was poor) or a complement construction (that it was a shame). The reason why this task is chosen in preference to the Crain et al's grammaticality judgement task is because, as I have previously criticised with the Marslen-Wilson et al experiments (see experiment 13), reaction time may be an

indication of factors extraneous to the experimental manipulations and therefore, getting subjects to choose a construction themselves may be a more accurate reflection of the normal procedures that people use in deciding upon form of definite reference. For example, the fact that subjects showed effects of context on judgement times in the Crain et al task may be because some constructions were more familiar, owing to their greater frequency of use in everyday contexts, rather than because subjects were aware that construction type was actually a reflection of the 'specificity' of referents in context.

It is predicted that adults choices of completions will be a reflection of the context: either the semantic context (the particular number of available referents) and/or the pragmatic context (whether one referent is plausible or not) and that they will prefer relatives when there is more than one available (and/or when no referent is plausible) and produce more non-relatives when there is already only one referent available (and/or when there is a plausible referent). Interest is also in whether both semantic and pragmatic context will affect completions, (in which case, one would expect relatives to appear only when there are two potential referents and where neither one is plausible) or whether one context (semantic or pragmatic) is more influential than another.

Experiment 16 is a partial replication of the previous experiment but it controls for any leading biases in the stories, which may have obscured pragmatic effects in the previous experiment. Thus, experiment 16

avoids using context relatives and any contextual information which may bias subjects away from a particular response in their efforts to prevent repetition and over-simplicity. Hypotheses are the same as for experiment 15: It is predicted that type of completion (relative or non-relative) will be a reflection of whether reference is made specific or not through contextual (semantic and/or pragmatic) information.

The main questions raised in this chapter are:

1. Is there evidence of a developmental increase in the ability to use both linguistic and non-linguistic factors for definite reference?
2. Does an appreciation of 'specificity' for definite reference require a particular type of interaction between pragmatic, semantic and syntactic knowledge?

7.3 EXPERIMENT 14: Understanding definite reference as a function of Principle C of binding theory and pragmatic context

INTRODUCTION

The experiment to follow was carried out in an attempt to manipulate linguistic (principle C) and non-linguistic (plausibility) factors in a comprehension task using a similar procedure to the Balfour experiment. However, instead of single sentences, sentence pairs were used, so that the correct referent was mentioned in an additional non-pronominalised sentence. Pragmatic context was manipulated using the same three conditions as Balfour, with the non-pronominalised sentence also varying in terms of pragmatics and in accordance with its pronominalised sentence. The questions only required Yes or No answers in the hope that this might reveal any differences between young children and adults in their choice of referent when pragmatics conflicted with principle C.

As a test of Chomsky's claim that young children prefer to choose a referent not mentioned in the pronominalised sentence for all cases of backward anaphora regardless of syntactic constraints, two control conditions were given. For these conditions, sentence pairs were used where reference was not constrained syntactically. Yet half of these control examples involved backward anaphora and half involved forward anaphora.

Also, as well as having examples where the non-pronominalised sentence preceded the pronominalised one

(Anaphoric context), there were also examples where the non-pronominalised sentence succeeded the pronominalised one (Cataphoric context). Umstead and Leonard (1983) have found that cataphoric references were the hardest sentences for children between 3 to 5 years of age to interpret. Therefore, this variable provided a test of whether subjects could choose an antecedent not mentioned in the pronominalised sentence, even when the antecedent appeared later, in a subsequent sentence.

In order to test claims regarding tangibility of referents, another variable was manipulated to see whether the performance of younger subjects was better when concrete objects were used than when the individuals referred to were hypothetical.

The main predictions were that Balfour's findings would be replicated: All children (even 4 year olds) would be able to make pragmatic inferences. However, if young children still need to learn exceptions to non-identity for backward anaphora (presumably due to 'parameter setting'), only at about 6 years of age would children be able to appreciate non-identity as a function of principle C when there was no pragmatic support. However, the youngest children were expected to show an over-application of non-identity to constructions where backward anaphora was not constrained by non-identity (preceeding pronoun was in the subordinate clause).

Also, for those children who do appreciate principle C, younger children were predicted to show no evidence of an explicit awareness of a conflict

between principle C and pragmatic context, as reflected by their tendency to choose the same response either for all those constructions which only supported pragmatics or for all those constructions which only supported non-identity. However, adults and older subjects were expected to show an awareness of the possibility of both interpretations as revealed by their strong either 'yes' or 'no' biases in the condition where principle C and pragmatics were in conflict. This is the result expected if young children cannot appreciate both contextual factors and linguistic factors simultaneously.

It was suggested that the problems in the Widowsky and Berkowitz and Widowsky experiments were due to the fact that their sentences were pragmatically biased towards intra-sentence co-reference and so if younger subjects are more influenced by pragmatics when it conflicts with syntax, then it is not surprising that the latter experiments failed to show the young child demonstrating a knowledge of syntax. It was expected that, rather than having difficulties with referents that were not tangible, it would be the conflict between syntax and context that would show the strongest developmental effects and therefore, it was predicted that there would be no differences between tangible and intangible conditions.

METHOD

Subjects

48 children were selected from a primary school in Durham. There were 16 children in each age group (3 to 5 [mean

age: 4.0], 6 to 8 [7.2], 9 to 11 [9.11]), with approximately equal numbers of males and females per group. 16 mothers were selected, who had children in the school ranging from 4 to 11 years of age and 16 undergraduate students (9 females and 7 males) from Durham University were selected from the psychology and engineering departments. For each of the 3 child groups, half the subjects were allocated to the 'tangible' condition and half were in the 'intangible' condition.

Materials

Model dolls of Jane and Susan were required for the tangible group. BBC Computers were set up for running the experiment on parents and students. 48 sentence-pairs were constructed for the older groups (9 to 11 years and adults) and 48 sentence-pairs were constructed for the two youngest groups (see appendix N for sentence-pairs).

Each sentence-pair contained one pronominalised sentence (containing both a pronoun and a proper noun) and one non-pronominalised sentence (containing a proper noun for a referent not 'named' in the pronominalised sentence).

Design and Procedure

There were 3 types of SENTENCE CONSTRUCTIONS used in the pronominalised sentence and these were as follows:

1. NON-IDENTITY - Pronoun precedes proper noun (pronoun in main clause) Example: She was happy that Susan had passed the exam.

2. UNRESTRICTED BACKWARD ANAPHORA - Pronoun precedes proper noun (pronoun in subordinate clause) Example: As she was happy, Susan had passed the exam.

3. UNRESTRICTED FORWARD ANAPHORA - Proper noun precedes pronoun (pronoun in subordinate clause) Example: Susan was happy that she had passed the exam.

The POSITION OF CONTEXT SENTENCE was varied across two conditions: The non-pronominalised sentence either preceded (Anaphoric context) or followed (Cataphoric context) the pronominalised one. The reasons for two types of context sentences was to ensure that choosing the referent outside the pronominalised sentence was not due to a general bias of referring to the individual in the preceding sentence (anaphoric context) because, an appreciation of structural non-identity implies an awareness that the referent can be anywhere (anaphoric or cataphoric) except in the pronominalised sentence. An example of each context position of the non-identity construction type are as follows:

ANAPHORIC - 'Jane was reading a book. She was happy that Susan had passed the exam'.

CATAPHORIC - 'She was happy that Susan had passed the exam. Jane was reading a book'.

Each sentence-pair varied over three conditions of PRAGMATIC CONTEXT. These were as follows, using anaphoric, non-identity types to present an example of each:

NEUTRAL - no pragmatic bias to one referent over another. 'Jane was reading a book. She was happy that Susan had passed the exam'.

PLAUSIBLE OUTWARD - pragmatic bias to referent outside pronominalised sentence and incompatible with referent inside pronominalised sentence. 'Peter was on the beach. He built a sandcastle while Paul was paddling'.

IMPLAUSIBLE OUTWARD - pragmatic bias to referent inside pronominalised sentence and incompatible with referent outside pronominalised sentence. 'Teresa had no money. She bought a ticket before Wendy got in the train'.

There were 3 conditions of SENTENCE CONSTRUCTION, 2 conditions of CONTEXT POSITION and 3 conditions of PRAGMATIC CONTEXT ($3 \times 2 \times 3 = 18$ conditions). However, there were 4 examples of each condition for non-identity constructions (3 pragmatic context by 2 context position: $4 \times (3 \times 2) = 24$), and there were 2 examples of each condition for each unrestrictive construction ($2 (2(2 \times 3) = 12 + 12)$), so altogether there were 48 sentence-pairs ($24 + 12 + 12 = 48$).

However, the 48 sentence-pairs presented to the youngest

two age groups were different to those presented to the older groups. The reason for this was because the ability to use contextual information depends on whether the speaker has the necessary general knowledge for making an inference in the first place. In previous experiments, the fact that adults' general knowledge is different to that of younger children was reflected by the pragmatic inferences they used in the experiment (see experiment 12). Therefore, for the 3 to 5 year olds and 6 to 8 year olds, sentence-pairs were constructed with the aim of using information which was more typical of the young child's general experiences. Moreover, so that memory factors did not limit the amount of data that could be obtained, the sentence-pairs given to young children were shorter and less complex than those given to older children and adults.

Subjects given the tangible condition received all sentence-pairs using the same two individuals (Susan and Jane), whereas subjects in the intangible condition received a different pair of same-gender individuals on each trial.

Altogether then, apart from age, there were 4 variables manipulated in the experiment: sentence construction, context position, pragmatic context and tangibility. Tangibility was a between-subjects variable, whereas the others were repeated-measures factors.

To prevent any experimenter or situational confounding, the order that subjects were tested was counterbalanced across tangible and intangible groups.

As subjects received quite a large number of trials, each child was seen on two separate occasions, in order to prevent any confounding of results due to fatigue or boredom.

The procedure was as follows:

Phase 1: Familiarisation Period

For subjects in the tangible group, Jane and Susan dolls were first introduced, whereas no dolls were introduced to the child in the intangible condition. As in the Balfour experiment, children were initially asked questions concerning single sentences with the mention of only one name. For example, 'Jane watches T.V. Who watches T.V?' and became more complex: 'You saw Susan hiding. Who was hiding?' When the child became familiar with this task, the 48 sentence-pairs were presented in a random order.

Phase 2: Running the experiment (6 to 11 year olds)

For all children, except for the youngest group, (reasons for the exception of this group are explained later), the experimenter read each sentence-pair followed by 3 questions: Referential, Memory and Control questions, which were presented in a random order. For example, given the sentences:

'Jane was reading a book. She was happy that Susan had passed the exam'.

- questions were asked about one of the two referents. The question sometimes required a NO response and sometimes required a YES response. This was to control for any 'acquiescence bias' in answering questions. (Whether a yes or no response was required was counterbalanced for each question type over trials in each condition).

The REFERENTIAL question asked about the referent of the pronoun, (either 'Was Susan happy?' or 'Was Jane happy?').

The MEMORY question asked about the noun phrase in the non-pronominalised sentence when there was an anaphoric context and asked about the noun phrase in the pronominalised sentence when there was a cataphoric context. This question served as a check that the children could remember the referents mentioned earlier in the sentence-pair, (either 'Was Jane reading a book?' or 'Was Susan reading a book?').

The CONTROL question asked about the noun phrase in the pronominalised sentence when the context was anaphoric, and asked about the noun phrase in the non-pronominalised sentence when the context was cataphoric. This question was a further check that children were not responding randomly to the referential question, as well as serving to prevent any bias caused by mentioning some of the events and not others, (either 'Did Susan pass the exam?' or 'Did Jane pass the exam?').

Whenever, a subject gave an incorrect response to a memory or control question, the sentence-pair was re-read and all

3 questions were repeated. The experimenter recorded answers to referential questions on the recording sheet. For all sentence-pairs a score was given for choosing the referent outside the pronominal sentence and for not choosing the referent inside the pronominal sentence. No and Yes scores were recorded separately.

For the tangible condition, a response could have been either verbal ('Jane' or 'Susan') or non-verbal (pointing to one of the two dolls). Thus, whether responses were verbal or manual, was also recorded by the experimenter, for this group.

Procedure for the 3 to 5 year olds

For 3 to 5 year olds, after initially testing 7 subjects from this group using the procedure above, it was observed that all subjects showed a particularly high acquiescence bias (nearly all responses were 'yes'). Therefore, 7 new subjects were selected to make up the numbers for this group and the procedure was changed in a way which appeared to reflect more accurately the young child's referential knowledge. Thus, the procedure for the 3 to 5 year olds involved firstly, asking the referential question before presenting each pair of sentences and asking the question again after the sentences were presented. Then either the memory or control question was given to check that the subject could remember the explicit information from one of the sentences. If the child failed on the memory/control question, the sentence-pair was presented again at the end of the experiment. This change

in procedure appeared to be more appropriate for younger subjects because, a) asking the referential question before reading the sentences primed the child to pay attention to the sentences, while searching for a referent and b) asking one memory/control question after the referential question meant that young children's ability could be tested whilst they could still remember the information from the sentence-pairs.

Adults procedure

Students and parents were only presented with the intangible condition. They read the sentence-pairs and questions on a computer screen. The reason for this procedural difference was because there was evidence of problems with experimenter-adult interaction in previous experiments. For example, adults often appeared to perceive the experiment as being threatening and were concerned that their performance might have exposed them unfavourably (see chapter 5 discussion). The subjects remained anonymous in the present experiment, as subjects were able to select their own subject number, to cross the number off the list and to enter it into the computer without the experimenter being present. The task was self-paced: After a subject had indicated that they had read a sentence-pair, the screen cleared and the 3 questions were presented individually on the screen in a random order. The subject had to make a response before the next question or sentence-pair was presented and they were allowed as much time as they needed to answer either

'yes' 'no' or 'don't know'. All three question types were asked to ensure that adults were reading the sentence-pairs properly and were not just responding randomly. Answers to questions were later transcribed for analysis and it was decided that, if any subject failed to answer the memory and control questions correctly for more than 8 trials, they would not be included in the analysis.

Finally, a post-experimental, multiple-choice questionnaire was given to all adult subjects (see appendix N for questions). This was to see whether subjects perceived any of the sentence-pairs as being ambiguous or complicated and to see if they regarded any of the questions as being difficult. Also, the questionnaire checked that the results were not confounded by fatigue or boredom.

RESULTS

Firstly, no subject had to be excluded from the analysis, because all the children eventually answered control and memory questions correctly and the highest incidence of memory errors for any adult subject was 4.

The results were divided into the following 5 sections:

1. NON-IDENTITY RESTRICTION [Pro..NP. Pronoun in main clause].
2. UNRESTRICTED BACKWARD ANAPHORA [Pro....NP. Pronoun in subordinate clause].
3. UNRESTRICTED FORWARD ANAPHORA [NP....Pro. Pronoun in subordinate clause].

4. YES/NO RESPONSES

5. QUESTIONNAIRE RESPONSES OF ADULTS.

1. NON-IDENTITY RESTRICTION

A. Pragmatic Context and Position of Context.

Table 14-A presents the mean number of outward scores (choosing the referent in the non-pronominalised sentence and failing to choose the referent in the pronominalised sentence) for non-identity trials in each subject group and with scores in each pragmatic context (plausible, neutral and implausible).

Scores out of 8. Random scores = 4. Inward bias <4.

SUBJECT GROUP	PLAUSIBLE	NEUTRAL	IMPLAUSIBLE	MEANS
3-5	7.9	6.3	2.6	5.6
6-8	7.4	6.2	1.7	5.1
9-11	6.4	6.1	2.8	5.1
PARENT	5.6	5.0	3.4	4.7
STUDENT	6.4	6.4	4.7	5.8
<u>MEANS</u>	<u>6.7</u>	<u>6.0</u>	<u>3.0</u>	<u>5.3</u>

TABLE 14-A: NON-IDENTITY SCORES AS A FUNCTION OF PRAGMATIC CONTEXT.

The table shows that the overall trend was for outside responses to depend on plausibility of context: The highest outside scores were in the plausible condition and the neutral scores were higher than the scores in the implausible condition. However, the largest change in

scores was in the implausible condition and there was an overall outside preference even in the neutral condition. In fact, there was no difference between plausible and neutral scores for the student group. However, analysis of variance showed a significant effect of pragmatic context overall ($F=134.50$, $df=2,158$, $p<0.001$) as well as a significant difference between neutral and plausible conditions ($F=24.57$, $df=1,75$, $p<0.001$) and a significant difference between neutral and implausible conditions ($F=166.02$, $df=1,75$, $p<0.001$).

Whilst the highest scores were from the student group, the lowest scores were from the parent group in the plausible and neutral conditions.

Within each pragmatic condition, the youngest groups produced higher scores in the plausible condition than older subjects. Comparing plausible and neutral conditions, subject group differences were found to be significant overall ($F=6.55$, $df=4,75$, $p<0.001$) and there was also an interaction between plausible/neutral differences and subject group ($F=3.77$, $df=4,75$, $p<0.01$). Within subject group comparisons showed that this effect was due to plausible and neutral differences in the two youngest age groups only (3-5: $F=19.57$, $df=1,15$, $p<0.001$. 6-8: $F=20.59$, $df=1,15$, $p<0.001$), but not for the older groups (9-11: $F=1.0$ Parent: $F=2.8$ Student: $F<1$).

There were clear age differences in performance for the implausible condition. However, whilst outward scores showed a developmental increase in the implausible condition, the poorest scores for this condition were from

the 6 to 8 year olds. Comparing implausible and neutral conditions, subject group differences were found to be significant overall ($F=5.76$, $df=4,75$, $p<0.001$) and there was also an interaction between implausible/neutral differences and subject group ($F=5.89$, $df=4,75$, $p<0.001$). Within subject group comparisons showed that, although the largest differences were for the child age groups, even adult groups revealed a significant difference between implausible and neutral conditions (3-5: $F=53.12$, 6-8: $F=303.75$, 9-11: $F=26.68$. $df=1,15$, $p<0.001$. Parents: $F=9.64$, Students: $F= 9.30$. $df=1,15$, $p<0.01$).

Table 14-B shows mean outward scores for each subject group in each of the 2 positions of context: anaphoric/cataphoric).

Scores out of 12. Random scores = 6. Inward bias <6.

SUBJECT GROUP	ANAPHORIC	CATAPHORIC	MEANS
3-5	9.0	7.8	8.4
6-8	9.1	6.2	7.7
9-11	8.7	6.5	7.6
PARENT	8.6	5.4	7.0
STUDENT	9.7	7.9	8.8
MEANS	9.0	6.8	7.9

TABLE 14-B: NON-IDENTITY SCORES IN EACH CONTEXT POSITION.

Firstly, it is clear that anaphoric outward scores exceeded cataphoric outward scores in all conditions and analysis of variance found a main effect of context position on outward non-identity scores ($F=62.62$, $df=1,75$, $p<0.001$).

The position of context was found to be significant for all pragmatic contexts (plausible: $F=12.01$, $df=1,79$, $p<0.001$. neutral: $F=30.83$, $df=1,79$, $p<0.001$. implausible: $F=8.72$, $df=1,79$, $p<0.001$).

The position effect occurred in all subject groups and there was no interaction between context position and subject group ($F=1.41$, $df=4,75$, NS). Nor was there a 3-way interaction between context position, pragmatic context and subject group ($F=1.74$, $df=8,150$, NS).

B. Tangibility of referents

Analysis of variance showed no overall effect of tangibility ($F=1.28$, $df=1,42$, NS), nor an interaction between tangibility and position of context ($F=0.57$, $df=1,42$, NS). However, an interaction between tangibility and pragmatic context was significant ($F=3.57$, $df=2,84$, $p<0.05$), presumably because there were more outward scores in the implausible condition for the tangible group (67) compared with the intangible group (47).

The only differences between age groups was that 3 to 5 year olds produced higher scores in the tangible group for the neutral cataphoric condition. (Tangible - 30 Intangible - 20. In fact, these scores were higher than in the neutral anaphoric condition where there were 27

outward scores in the tangible condition). Also, 9 to 11 year olds produced less scores in the tangible group for the plausible anaphoric condition (Tangible - 23 Intangible - 30). These 'tangible' scores were slightly lower than for the plausible cataphoric condition (25) for this age group.

Overall, no age by tangibility interaction was found ($f=0.97$, $df=2,42$, NS), but there was a 3-way interaction between age, tangibility and pragmatic context ($f=4.16$, $df=4,84$, $p<0.005$).

Manual responses were only present in the tangible condition for 3 to 5 year olds. However, there was only a very small proportion of manual responses even in this age group.

2. UNRESTRICTED BACKWARD ANAPHORA

A. Pragmatic Context and Position of Context.

Table 14-C shows the outward scores for unrestricted backward anaphora in each pragmatic context, for each subject group.

Scores out of 4. Random scores = 2. Inward bias <2.

SUBJECT GROUP	PLAUSIBLE	NEUTRAL	IMPLAUSIBLE	MEANS
3-5	3.2	2.2	0.7	2.0
6-8	3.2	2.5	0.3	2.0
9-11	3.4	2.4	1.1	2.3
PARENT	2.9	1.6	1.1	1.9
STUDENT	2.5	0.9	0.9	1.4
MEANS	3.0	1.9	0.8	1.9

TABLE 14-C: UNRESTRICTED BACKWARD ANAPHORA AS A FUNCTION OF PRAGMATIC CONTEXT: OUTWARD SCORES.

There were clear differences between pragmatic conditions, with outward scores being the highest in the plausible condition, approximately unbiased in the neutral condition and showing a stronger inward bias in the implausible condition.

However, using analysis of variance, there was a significant subject group effect on scores in the neutral condition ($F=8.45$, $df=4,75$, $p<0.0001$) and the Newman-Keuls revealed that this was due to the parent and student groups, who produced less outward scores in the neutral condition than other age groups. However, there were no subject group effects in the plausible condition ($F=2.44$, $df=4,75$, NS) and whilst there were significant subject group effects in the implausible condition ($F=2.72$, $df=4,75$, $p<0.05$), there were no two groups that differed significantly on the Newman-keuls test.

Table 14-D presents the outward scores for unrestricted backward anaphora for both context positions in each subject group.

Scores out of 6. Random scores = 3. Inward bias <3.

SUBJECT GROUP	ANAPHORIC	CATAPHORIC	MEANS
3-5	3.3	2.8	3.1
6-8	3.5	2.5	3.0
9-11	4.1	2.8	3.5
PARENT	3.4	2.2	2.8
STUDENT	3.1	1.2	2.2
<u>MEANS</u>	3.5	2.3	2.9

TABLE 14-D: UNRESTRICTED BACKWARD ANAPHORA SCORES IN EACH CONTEXT POSITION: OUTWARD SCORES.

At all ages, there were more outward scores in the anaphoric position than in the cataphoric position, particularly the students, who produced a very low number of outward scores mainly in the cataphoric condition. Analysis of variance showed a significant position of context effect ($F=46.85$, $df=1,75$, $p<0.001$) and a significant subject group effect ($F=4.50$, $df=4,75$, $p<0.005$), but no significant interaction was found between subject group and context position on outward scores

($F=1.93$, $df=4,75$, NS).

Overall there were clearly more outward scores for non-identity than for unrestricted backward anaphora (UBA) in all conditions and in all subject groups. However, there seemed to be less differences between 'non-identity' and 'UBA' scores in the plausible condition.

Analysis of variance confirmed that there was a significant interaction between Non-identity/UBA scores and pragmatic context ($F=10.65$, $df=2,150$, $p<0.001$). Also there was an interaction between Non-identity/UBA and subject group ($F=6.79$, $df=4.75$, $p<0.001$), due to 9 to 11 year olds and parents showing more outward scores for UBA in the plausible condition.

B. Tangibility of referents

The only differences were for the 9 to 11 year olds in the cataphoric plausible condition, where more outward scores were produced in the intangible group (15) compared with the tangible group (10). Only this effect was significant on the Mann-Whitney tests ($U=15.5$. Corrected for ties $z=-2.03$, $p<0.05$).

3. UNRESTRICTED FORWARD ANAPHORA.

A. Pragmatic Context and Position of Context.

Table 14-E shows the outward scores for each pragmatic context in each subject group, for unrestricted forward anaphora.

Mean scores out of 4. Random scores=2		Inward bias<2		
SUBJECT GROUP	PLAUSIBLE	NEUTRAL	IMPLAUSIBLE	MEANS
3-5	2.8	0.5	0.2	1.2
6-8	1.8	1.4	0.4	1.2
9-11	1.8	1.1	0.5	1.1
PARENT	2.3	1.5	0.8	1.5
STUDENT	2.4	0.9	0.7	1.3
<u>MEANS:</u>	<u>2.2</u>	<u>1.1</u>	<u>0.5</u>	<u>1.3</u>

TABLE 14-E: UNRESTRICTED FORWARD ANAPHORA SCORES AS A FUNCTION OF PRAGMATIC CONTEXT: OUTWARD SCORES

The table shows that the overall preference was for an inward response. There were slight differences between the conditions: Again, the outward scores in the plausible condition exceeded the neutral condition, which itself exceeded the scores in the implausible condition. However, the inward preference did override plausibility to some extent, as random totals appeared in the plausible condition and a clear inward bias was shown in the neutral condition. Analysis of variance confirmed that the differences between pragmatic contexts were significant ($F=115.55$, $df=2,150$, $p<0.001$) and whilst there was no significant subject group effect overall ($F=1.42$, $df=4,75$, NS), there was a significant interaction between pragmatic context and subject group ($F=4.57$, $df=8,150$, $p<0.001$).

This interaction is due to the small differences between neutral and implausible conditions for 3 to 5 year olds and students.

Table 14-F presents the outward scores across the anaphoric and cataphoric context positions, in each subject group.

Mean scores out of 6. Random scores=3 Inward bias<3

SUBJECT GROUP	ANAPHORIC	CATAPHORIC	MEANS
3-5	1.5	2.0	1.8
6-8	1.9	1.7	1.5
9-11	1.9	1.5	1.7
PARENT	2.9	1.7	2.3
STUDENT	2.6	1.4	2.0
<u>MEANS:</u>	<u>2.2</u>	<u>1.7</u>	<u>1.9</u>

TABLE 14-F: UNRESTRICTED FORWARD ANAPHORA SCORES AS A FUNCTION OF POSITION OF CONTEXT: OUTWARD SCORES.

The scores showed little anaphoric/cataphoric differences and the largest differences were for the two adult groups, who showed a stronger outward preference in the anaphoric condition. Again, analysis of variance revealed a significant effect of context position ($F=10.38$, $df=1,75$, $p<0.005$). There was no significant overall subject group effect ($F=1.42$, $df=4,75$, NS), but a significant context position by subject group interaction

was found ($F=3.83$, $df=4,75$, $p<0.01$).

Overall, there was clearly far less outward scores for unrestricted forward anaphora (UFA) than for non-identity in all conditions and in all subject groups. Again, there seemed to be less differences between these two scores in the plausible condition than for other pragmatic contexts and this was due particularly to the adult group.

Analysis of variance revealed an interaction between Non-identity/UFA scores and pragmatic context ($F=22.0$, $df=2,150$, $p<0.001$) and a 3-way interaction between subject group, pragmatic context and non-identity/UFA scores ($F=5.40$, $df=8,150$, $p<0.001$).

Outward scores were clearly higher for UBA than for UFA. The only exceptions appear to be for implausible conditions in all subject groups and for all conditions in both the adult groups, where UBA/UFA differences were small. Analysis of variance showed a main effect of UFA verses UBA ($F=62.91$, $df=1,75$, $p<0.001$), an interaction between UFA/UBA and pragmatic context ($F=5.36$, $df=2,150$, $p<0.01$), an interaction between UFA/UBA and subject group ($F=5.08$, $df=4,75$, $p<0.001$) and a 3-way interaction between UFA/UBA, pragmatic context and subject group ($F=4.28$, $df=8,150$, $p<0.001$).

B. Tangibility of referents

Mann-Whitney tests confirmed that there were no significant differences between tangible and non-tangible groups in

either of the above 6 conditions (pragmatic context and position of context), in any subject group.

4. YES/NO SCORES IN NON-IDENTITY TRIALS

This excludes 3 to 5 year olds, who were not required to give yes and no answers.

Table 14-G presents the mean number of YES scores in each pragmatic context and position of context and for each subject group for 'non-identity' trials only.

Scores out of 4. Random scores = 2. NO bias <2.

T = Total (out of 8. Random = 4. NO bias <4)

anap = anaphoric cat = cataphoric

SUBJECT GROUP	PLAUSIBLE			NEUTRAL			IMPLAUSIBLE			MEANS	
	anap	cat	T	anap	cat	T	anap	cat	T	anap	cat
6-8	2.0	2.1	(4.1)	2.2	2.1	(4.3)	2.4	1.9	(4.3)	2.2	2.0
9-11	2.2	1.9	(4.1)	2.3	1.8	(4.1)	3.1	2.4	(5.5)	2.5	2.0
PAR.	2.2	1.8	(4.0)	1.9	2.1	(4.0)	1.0	1.3	(2.3)	1.7	1.7
STUD.	2.3	1.9	(4.2)	2.1	2.1	(4.2)	1.3	0.7	(2.0)	1.9	1.6
MEANS:	2.2	1.9		2.1	2.0		2.0	1.6		(2.1)	(1.8)
T:	(4.1)			(4.1)			(3.6)			(3.9)	

TABLE 14-G: YES SCORES AS A FUNCTION OF PRAGMATICS AND CONTEXT POSITION FOR NON-IDENTITY EXAMPLES.

Whilst there was not a yes/no preference in the neutral and pragmatic conditions, there was a slight NO preference in

the implausible condition, as shown by a significant effect of pragmatic context on yes responses using analysis of variance ($F=7.52$, $df=2,120$, $p<0.001$).

However, there was a significant effect of subject group on yes scores ($F=14.63$, $df=3,60$, $p<0.001$) and an interaction between subject group and pragmatic context ($F=15.59$, $df=6,120$, $p<0.001$). These effects are due to the greater number of YES responses in the 9 to 11 year old group and the greater number of NO responses from the parents and students in the implausible condition. This was confirmed when comparing pragmatic context in each subject group. There was a significant effect of pragmatic context in all groups except for the 6 to 8 year olds (6-8: $F=0.27$, $df=2,30$, NS. 9-11: $F=13.5$, $df=2,30$, $p<0.001$. Parents: $F=11.31$, $df=2,30$, $p<0.001$. Students: $F=26.16$, $df=2,30$, $p<0.001$).

Analysis of variance also showed a significant effect of context position ($F=13.42$, $df=1,60$, $p<0.001$), and whilst there was no significant interaction between pragmatic context and context position ($F=1.06$, $df=2,120$, NS), there was a significant interaction between subject group and position of context in yes/no responses ($F=3.10$, $df=3,60$, $p<0.05$). This effect is due to the fact that 9 to 11 year olds showed a stronger YES bias in anaphoric examples than for cataphoric examples.

Finally, no effects of tangibility on yes/no scores were significant (Main effect: $F=0.32$, $df=1,28$, NS. Age by

tangibility: $F=0.01$, $df=1,28$, NS. Pragmatics by tangibility: $F=0.24$, $df=2,56$, NS. Context position by tangibility: $F= 0.9$, $df=1,28$, NS).

5. QUESTIONNAIRE RESPONSES - ADULTS ONLY

Responses to the questionnaire from the adult groups were as follows (Tables 14-H - 14-K)

QUESTION 1. DID YOU FIND THAT SOME OF THE SENTENCES WERE:

A. COMPLICATED	65.6 (21)
B. OK	34.4 (11)
C. EASY	0 (0)

TOTAL: PERCENT (SUM) -	100 (32)
------------------------	----------

TABLE 14.H QUESTION ONE

Chi-square tests showed a significant difference between scores for A,B and C ($\chi^2=20.69$, $df=2$, $p<0.0000$), but no significant difference between A and B ($\chi^2=3.13$, $df=1$, NS).

QUESTION 2. DID YOU FIND THAT SOME OF THE SENTENCES WERE:

C. AMBIGUOUS	93.8	(30)
B. NOT VERY TYPICAL	6.3	(2)
A. ALWAYS VERY COMMON	0	(0)

TOTAL: PERCENT (SUM) -	100	(32)
------------------------	-----	------

TBALE 14.I QUESTION 2

Chi-square tests showed a significant difference between scores for A, B and C ($=52.75$, $df=2$, $p<0.0000$) and a significant difference between B and C ($=24.5$, $df=1$, $p<0.0000$).

QUESTION 3. DID YOU FIND THAT SOME OF THE QUESTIONS WERE:

C. DIFFICULT	65.6	(21)
A. OK	28.1	(9)
B. EASY	6.3	(2)

TOTAL: PERCENT (SUM) -	100	(32)
------------------------	-----	------

TABLE 14.J QUESTION 3

Chi-square tests showed a significant difference between scores for A, B and C ($=17.31$, $df=2$, $p<0.0000$) and a significant difference between A and C ($=4.8$, $df=1$, $p<0.05$).

QUESTION 4. DID YOU FIND THAT DURING THE EXPERIMENT YOU WERE:

A. TIRED AND BORED	46.9 (15)
B. CONCENTRATE ALL TIME	37.5 (12)
C. BECAME TIRED, BORED	15.6 (5)

TOTAL: PERCENT (SUM) -	100 (32)
------------------------	----------

TABLE 14.K QUESTION 4

Chi-square tests showed neither a significant difference between scores for A, B and C ($=4.94$, $df=2$, NS), nor a significant difference between A and B ($=0.33$, $df=1$, NS).

Overall, these findings show that adult subjects mostly found some sentence-pairs ambiguous and quite a few found some of them complicated. Most adults also found some questions difficult. However, there is no evidence that boredom and fatigue were the main problems encountered by adults.

DISCUSSION

As predicted, subjects of all ages made use of pragmatic context when making decisions about referents in both restrictive and unrestrictive trials. Thus, the evidence suggests that even 3 year olds are capable of making pragmatic inferences in choosing a referent and the finding that pragmatic support actually improved

non-identity constrained by principle C, shows that context could be used to support syntactic principles at all ages.

Contrary to Balfour's finding that children under 6 years of age tended to overapply non-identity to unconstrained backward anaphora examples, the results of the above experiment showed that even the youngest children (3 years) could appreciate principle C. The two pieces of evidence to support this conclusion were that, firstly, young children preferred to choose the referent in the non-pronominalised sentence when the pronoun preceded the referent, even when there was no pragmatic support in favour of one referent over another, showing that this reflected linguistic rather than non-linguistic knowledge. Secondly, whilst all children preferred an outward response when the pronoun preceded the referent and an inward response when the referent preceded the pronoun, even the youngest children preferred an outward response more when the pronoun preceding the referent was in the main clause rather than in the subordinate clause. This finding shows that the children were using structural notions, such as c-command, in making decisions about co-reference, because they were more likely to show non-identity when the pronoun was higher in the phrase structure tree than the referent and this effect did override the directional bias. This evidence suggests that 'parameter setting' is not as significant as is the young child's high level knowledge of syntactic restrictions.

However, the prediction that only from 9 years of age would children be able to integrate syntax and context when

both were in conflict was supported. This was shown by the fact that children in the youngest two groups failed to show an awareness of the presence of two competing interpretations and instead, they relied solely on either a non-linguistic interpretation or used purely structural criteria. By contrast, the oldest children showed some evidence of being able to accept that there was both a linguistic and a contextual interpretation because they produced a 'yes' bias in the implausible condition on non-identity trials. Also, adults showed a strong 'no' bias in this condition, which suggests that they were able to appreciate the conflict between syntax and pragmatics. It seems that the young child does not appreciate that both linguistic and non-linguistic factors must be compatible for successful and unambiguous definite reference.

The poorer non-identity scores in cataphoric conditions at all ages is probably a reflection of the atypicality of cataphoric constructions, which may have led subjects to infer that the pronoun referred to some individual not mentioned in the story.

Also, as predicted, there was no evidence of tangibility of referents facilitating non-identity performance of the youngest children. Thus, it is concluded that the comprehension of principle C is an aspect of knowledge that is inherent in young children and which is used regardless of the objects of referents themselves.

Finally, there were strategy preferences in all subject groups. There were also some age differences, but

it is difficult to make any conclusions about whether these are developmental or are merely a reflection of the tasks used in this experiment and differences in the way that subjects' perceived the experimental requirements.

7.4 EXPERIMENT 15: Story completions I: as a function of contextual information about available and/or plausible referents.

INTRODUCTION

For the experiment to follow, older children (8 to 11 years) and adults were asked to complete mini-stories, where the last word given was syntactically ambiguous with regards to whether a relative or a non-relative construction should follow.

If the Crain and Steedman/ Altmann and Steedman model (A&S) is correct, that semantic context is the main factor in solving syntactic ambiguities for anaphoric reference, one would expect completions to depend on specific discourse context: When there is more than one potential referent a RELATIVE CONSTRUCTION will be used, whereas, when there is only one referent available a NON-RELATIVE CONSTRUCTION will be used. However, if Marslen-Wilson and Tyler's theory is correct, that pragmatic context is the major factor determining anaphoric reference, then one would expect completions to depend on pragmatics: When the referent was more plausible than the other a NON-RELATIVE would be used, but when the referents were of equal plausibility, a RELATIVE would be used.

METHOD

Subjects

32 subjects were used for the experiment. 16 of these were children aged between 8 and 11 years (mean age: 9.0) who were selected from a primary school in Durham and the remaining 16 subjects formed the adult group. The latter were mainly undergraduate students from Durham University, but 6 adults were non-student volunteers from a 'parents' group in Durham. There was approximately an equal number of males and females in each of these 2 groups.

Materials (See appendix O for incomplete stories).

96 incomplete stories (plus two practice stories) were constructed as the experimental stimuli. Each incomplete story contained 2 complete sentences and one incomplete sentence. One of the complete sentences introduced one item/individual and the other complete sentence mentioned two other items/individuals. The referent mentioned before the ambiguous word in the third incomplete sentence referred to one of the 2 referents in the sentence which introduced 2 items/individuals.

A tape-recorder and score sheets were also required for the experiment.

Design and Procedure

A. DESIGN

The 2 main variables manipulated in this experiment were:

1. SEMANTIC CONTEXT: number of potential referents (1 OR 2)

The two item/individual sentence contained referents that were either the same (same-gender or same item/individual) or different (different-gender or different item/individual). Therefore, there was either one potential referent or two potential referents for the noun phrase preceding the ambiguous word in the incomplete sentence.

2. PRAGMATIC CONTEXT: whether one referent plausible or not

Either the pragmatic information about one of the two items/individuals was coherent with the information in the incomplete sentence or the information given about neither item/individual was related pragmatically to the information in the incomplete sentence. Therefore, there was either one plausible referent or no one plausible referent for the item/individual referred to before the ambiguous word in the incomplete sentence.

Therefore, there were 4 conditions varying as a function of semantic context (1 or 2 potential referents) and pragmatic context (plausible or neutral). An example of stories used in each condition are presented below: Pragmatic associations are shown in brackets. Potential referents as well as incomplete reference in the final sentence are underlined.

PLAUSIBLE

ONE REFERENT

A farmer had (just planted some vegetables).

There was a pig (that trod over his carrots) and a horse that was asleep.

The farmer (shouted) to the pig that.....

TWO REFERENTS

Betty (found an empty tennis court)

She saw a boy (from the tennis club) and another boy from the swimming club.

Betty told the boy that.....

NEUTRAL

ONE REFERENT

A headmaster was praising somebody for getting ten out of ten.

A teacher was sitting by the window and a pupil was sitting by the door.

The headmaster told the pupil that.....

TWO REFERENTS

Jack was holding a cloth

There was a table with crumbs on it and another table with chocolate on it.

Jack cleaned the table with.....

Each subject received both conditions of SEMANTIC CONTEXT. However, the PRAGMATIC CONTEXT was a between-subjects factor and so, subjects were randomly assigned to either 'plausible' or 'neutral' conditions within each subject group.

Another factor which was counterbalanced within all conditions was whether the ambiguous word in the incomplete sentence was a repetition of the word used to discriminate between individuals/items in the 2 referent sentence. The purpose of this was to ensure that subjects were using the context constructively, rather than simply repeating the phrase used in a previous sentence, as was possible for the 'repeated' examples. A 'repeated' and a 'non-repeated' example is shown below:

Repeated

Jane had lost her mummy.

She saw a man that was selling sandwiches and another man that was selling balloons.

Jane told the man that.....

Non-Repeated

David kept the football in his room.

A friend was looking for it in the garage and another friend was looking for it in the shed.

David showed his friend that.....

When one referent was plausible or when there was only one

potential referent, the position (order of mention) of intended and unintended referents in the 2 referent sentence was counterbalanced across all conditions.

To see whether any effects were particular to the type of ambiguous word used, there were 2 types of constructions ('That' and 'With') and there was an equal number of each of these constructions across all conditions. An example of each of these two constructions is shown below:

THAT Construction

Peter won a prize.

He saw a man that had an angry face and another man that had a smile on his face.

Peter told the man that.....

WITH Construction

A man was wearing heavy boots.

He saw a bird with a long tail and another bird with a short tail.

The man trod on the bird with.....

Finally, to control for possible confounding of results due to memory limitations or strategy preferences, such as referring to the first mentioned or most recently mentioned referent, the order in which the 2 complete sentences were presented was reversed for half the subjects, so that in each of the above conditions, half the subjects received the 1 item/individual sentence first and the other half

received the 2 item/individual sentence first. Thus, for the 'With' construction example above, sentence one and two would be reversed for half the subjects, so they would receive the following mini-story instead:

A man saw a bird with a long tail and another bird with a short tail.

He was wearing heavy boots.

The man trod on the bird with.....

Controlling for all the factors above, the number of within subject conditions was 16 (4 factors by 2 conditions) and therefore, each subject received 16 experimental stories to complete.

Each subject also received 16 filler stories. The rationale for using fillers was to prevent subjects from discovering the variables being manipulated in the experiment. The fillers were also a check to see whether subjects did refer back to the contextual information or not. For the fillers, subjects were required to complete each story after an unambiguous word (either 'because' or 'in'), which was intended to get subjects to refer to the item/individual mentioned in the one referent sentence. In order to match these filler stories as best as possible with the experimental stories, the referents in the 2 item/individual sentence were either the same or were different. Also, the order of the two complete sentences matched the order that these sentences were presented for

the experimental stories.

Therefore, each subject received 32 incomplete stories (16 experimental and 16 filler stories) and these were presented in a random order.

B. PROCEDURE

Each subject was seen individually by the experimenter. Subjects were instructed to listen carefully to each story and were told to complete the last sentence of the story when the experimenter raised her index finger. Two practice stories were first presented. This was to ensure that subjects understood what was required of them. As for the experimental and filler stories, they were each made up of 2 sentences plus one incomplete sentence. The two practice stories are presented below:

P1. Jane went to a party.

There were games and fizzy drinks for everyone.

Jane was very happy and she.....

P2. Peter picked up a pen.

He wrote a letter to his friend.

Peter told his friend that he.....

All experimental stories and fillers (16 + 16 = 32) were presented in a random order. The experimenter read each incomplete story, without emphatic stress (particularly for the final incomplete sentence) and allowed the subject as

much time as they required to complete each story. Completions were recorded on the tape recorder to be subsequently transcribed for analysis.

Each subject was seen on two occasions, because, as the procedure was fairly long, each subject was given a rest period after receiving the first half of the stories. This was to eliminate any possible changes in performance that could be due to the effects of fatigue or boredom.

SCORING

Completions were scored into one of the following categories:

A. Relative completion- using the information from the sentence containing two individuals/items (eg. the man trod on the bird with the long tail).

B. Non-relative context completion - using the information from the sentence containing one individual/item : either a prepositional phrase for a 'with' construction (eg. the man trod on the bird with his heavy boots) or a complement clause for a 'that' construction (eg. the girl told the man that she was lost).

C. No reference to passage - eg 'the man trod on the bird with all his might'.

D. Indeterminate - eg 'the man trod on the bird with.....because he's mean'.

RESULTS

Virtually all completions for the fillers were coherent with context: made reference either explicitly or implicitly to information presented in the previous sentences. The mean scores out of 8 for reference to context were as follows: BECAUSE: 8-11= 7.9. Adult= 7.9. IN: 8-11= 7.9. Adult= 7.6.

Table 15-A presents total scores for relative and non-relative completions, as a function of both semantic and pragmatic context for children and adult groups.

	<u>SEMANTIC CONTEXT</u>			
	RELATIVE		NON-RELATIVE	
AGE GROUP	1 REF.	2 REF.	1 REF.	2 REF.
<hr/>				
8 TO 11 YEARS:				
<u>PRAGMATIC CONTEXT</u>				
PLAUSIBLE	19	36	45	28
NEUTRAL	19	24	45	40
<hr/>				
ADULTS:				
PLAUSIBLE	3	12	61	52
NEUTRAL	17	33	47	31

TABLE 15-A: RELATIVE AND NON-RELATIVE COMPLETIONS AS A FUNCTION OF BOTH THE SEMANTIC AND PRAGMATIC CONTEXT.

It appears that, regardless of pragmatic context, both

groups used more relatives for more than one referent and more non-relatives for one referent. This was confirmed by analysis of variance which showed a significant effect of semantic context (1 or 2 referents) on completion type (relative or non-relative) ($F=33.54$, $df=1,28$, $p<0.001$), but there was no 3-way interaction between age, semantic context and type of completion ($F=0.14$, $df=1,28$, NS).

Moreover, for neither age group was there a 3-way interaction between semantic context, WITH/THAT construction and completion type (8-11: $F=0.33$, $df=1,14$, NS. Adult: $F=0.43$, $df=1,14$, NS). Adults used more relatives for the neutral condition and more non-relatives for the plausible condition, regardless of semantic context, but these pragmatic effects were small compared with the strong semantic effects.

The only pragmatic effects for the children were slight and the reverse of the adults: more relatives when plausible and more non-relatives when neutral. In fact, analysis of variance showed no interaction between pragmatic context and type of completion overall ($F=2.06$, $df=1,28$, NS), but there was a significant 3-way interaction between age, pragmatic context and type of completion ($F=8.59$, $df=1,28$, $p<0.01$).

Analysis of variance failed to find a 3-way interaction between semantic context, pragmatic context and type of completion ($F=0.38$, $df=1,28$, NS). However, there was a 4-way interaction between age, semantic context, pragmatic context and type of completion ($F=5.48$, $df=1,28$, $p<0.05$), owing to the fact that the slight 'reverse' effect

of pragmatics for the children was confined to when the semantic context contained two potential referents. Also, for both age groups, there was a 3-way interaction between pragmatic context, WITH/THAT construction and type of completion (8-11: $F=7.71$, $df=1,14$, $p<0.05$. Adult: $F=23.2$, $df=1,14$, $p<0.001$).

In order to analyse the source of the 3-way interaction above, table 15-B presents total scores for relative and non-relative completions as a function of pragmatic context, looking at 'with' and 'that' stories separately, in both children and adult groups.

AGE GROUP	RELATIVE			NON-RELATIVE		
	PLAUS.	NEUT.	total	PLAUS.	NEUT.	total
<u>WITH</u>						
8-11	38	41	(79)	26	23	(49)
ADULTS	5	42	(47)	59	22	(81)
<u>THAT</u>						
8-11	17	2	(19)	47	62	(109)
ADULTS	10	8	(18)	54	56	(110)

TABLE 15-B: RELATIVE AND NON-RELATIVE COMPLETIONS AS A FUNCTION OF PRAGMATIC CONTEXT, FOR 'WITH' AND 'THAT' STORIES.

For 'with' stories, there seems to be a pragmatic

effect only for adults: more relatives for neutral contexts and more non-relatives for plausible referents. Analysis of variance revealed a significant interaction between pragmatic group and type of completion for 'with' stories ($F=16.72$, $df=1,28$, $p<0.001$). Also, there was a 3-way interaction between age, pragmatic context and completion type for 'with' stories ($F=12.08$, $df=1,28$, $p<0.005$). For 'that' stories, there seems to be no pragmatic effect for the adult group. However, for the children, there is an effect opposite to that predicted: more relatives for plausible referents and more complements for neutral contexts. However, analysis of variance revealed no significant interaction between pragmatic group and type of completion for 'that' stories ($F=3.33$, $df=1,28$, NS), nor a 3-way interaction between age, pragmatic context and completion type for 'that' stories ($F=1.95$, $df=1,28$, NS).

The table also shows that for 'with' stories, there was a preference for prepositional phrases overall for the adults. However, for the children, there was an overall relative completion preference. Analysis of variance revealed no significant difference between relative and preposition completions for 'with' constructions ($F=0.04$, $df=1,28$, NS), but there was an interaction between completion type and age ($F=10.7$, $df=1,28$, $p<0.005$). For 'that' stories, there was a strong preference for complement phrases overall, for both subject groups. Analysis of variance revealed a significant difference between relative and complement completions for 'that'

constructions ($F=95.5$, $df=1,28$, $p<0.001$), but there was no interaction between age and completion type ($F=0.01$, $df=1,28$, NS).

Furthermore, analysis of variance showed an interaction between THAT/WITH constructions and completion type for both age groups (8-11: $F=85.71$, $df=1,14$, $p<0.001$. Adult: $F=12.83$, $df=1,14$, $p<0.005$).

Analysis of variance showed that there were no effects of sentence order (order of presentation of 2 completed sentences) on the results. For both age groups, there was no interaction between sentence order and type of completions (8-11: $F=2.84$, $df=1,12$, NS. Adult: $F=2.65$, $df=1,12$, NS), nor was there a 3-way interaction between semantic context, sentence order and completion type (8-11: $F=0$, $df=1,12$, NS. Adult: $F=0.02$, $df=1,12$, NS). Finally, there was no 3-way interaction between between pragmatic context, sentence order and type of completion (8-11: $F=0.71$, $df=1,12$, NS. Adult: $F=1.99$, $df=1,12$, NS).

DISCUSSION

The results showed a clear effect of semantic context on type of completion for both children (8-11 years) and adults: More relatives were used when there was more than one potential referent and more non-relatives were used when there was only one potential referent. This finding supports Altmann and Steedman's claims that the number of referents in the discourse model has an effect on construction type (relative/complement) in that the former must referentially support the latter.

Pragmatic effects were never as strong as semantic effects and they depended on both type of ambiguous word (with/that) and subject group (children/adults). Adult subjects showed an effect of pragmatic context on type of completion only for 'with' stories, using more relatives for neutral context and more prepositional phrases when there was a plausible referent. No such effects were observed for 'that' constructions. The children (8 to 11 years) also showed a slight effect of pragmatic context on type of completion, but this effect was the reverse of that predicted: more relatives when there was a plausible referent and more complement phrases when the pragmatic context was neutral. Moreover, this effect was confined only to 'that' constructions and only when the semantic context portrayed two potential referents. Therefore, only the results from the 'with' stories with adults show any support for Marslen-Wilson and Tyler's claims regarding the significance of pragmatic context.

Overall, preferred completion type was a function of the ambiguous word (with/that) and subject group (children/adults): 'With' stories brought about more relatives from children and more prepositions from adults, whereas 'that' stories rarely produced any relatives from either subject group.

However, the possibility remains that results were at least partly a reflection of the way the different constructions were presented, together with differences in the way that the two subject groups perceived the task and this may have obscured any genuine pragmatic effects. For

example, many of the 'that' stories used relative clauses in the context sentences. Thus, it is possible that if subjects were reluctant to use repeated-relatives, it would account for the very strong complement bias for 'that' stories, regardless of pragmatic context.

For 'with' constructions, a prepositional phrase using the information provided in the story usually meant repeating a word used previously, which often made the completion very obvious and probably too obvious for subjects to feel that this was the required response. In fact, spontaneous comments from the children revealed that they were often reluctant to use a prepositional phrase because they regarded the completion as being too simple. Such comments were, for example, "You've already said that she dried him with a towel" and "What? Shall I just say it again"? The following is an example of one of the few prepositional phrases that were used by the children:

'A man was wearing heavy boots.....The man trod on the bird with his HEAVY BOOTS'

However, adults obviously felt less confined to the explicit information conveyed in the text compared with the children, because many of their prepositional phrases were more indirectly related to the information presented in the story. An example of a typical adult prepositional phrase for 'with' constructions was as follows:

'A man was wearing heavy boots.....The man trod on the

bird with his HIS CLUMSY BIG FEET'

These age differences possibly account for why, in contrast to the adult subjects, the children showed a slight relative preference for 'with' stories and also may explain why any pragmatic context effects were obscured for the children.

7.5 EXPERIMENT 16: Story completions task 2.

INTRODUCTION

The aim of the following experiment was to partially eliminate any potential confounding induced by the materials used in experiment 15, as the way that the stories were constructed in the previous experiment may have resulted in artificial completion biases and hence may have also obscured any genuine effects of pragmatic context.

Also, in order to match the contextual variables (semantic and pragmatic), a repeated measures design was used: Each subject received each condition for both semantic context and pragmatic context.

Another change in the procedure of this experiment was that half the subjects received the stories visually (questionnaire) and the remaining subjects received the stories verbally (as for the previous experiment). Thus, it was possible to look at whether results were affected by the way that the task was presented.

METHOD

Subjects

36 subjects were used for the experiment, none of whom had participated in the previous experiment. 18 of these were children aged between 8 and 11 years (mean age: 9.9) and were selected from a Durham school. The remaining 18 subjects, the adult group, were an assortment of 9 undergraduate students from Durham

University and 7 parent volunteers. There was approximately an equal number of males and females in each of these 2 groups.

Materials (See appendix P for incomplete stories)

32 incomplete stories were constructed as the experimental stimuli. As before, each incomplete story contained 2 complete sentences and one incomplete sentence. 18 questionnaire sheets were also constructed. A tape-recorder and score sheets were again required for the verbally presented procedure.

Design and Procedure

A. DESIGN

The main difference between this experiment and experiment 15 was that the stories had been changed slightly for the current experiment in an attempt to prevent completion biases.

WITH CONSTRUCTIONS

In the previous experiment it was found that there was a relative bias from the children for 'with' constructions. It was suggested that this was due to the fact that repeating the contextual information would have made a prepositional phrase completion too simple and obvious. Therefore for the 'with' stories of this experiment, a prepositional phrase would involve making a choice between referents presented in the context. For example:

A girl carried some scissors and some glue.

She had a box that had a broken lid and a box that had a good lid. She mended the box with

(Prepositional phrase would be either 'the scissors' or 'the glue')

Paradoxically, despite the child's relative bias in the previous experiment, it was suggested that using relatives in the context sentences may have discouraged subjects from using relative completions. It is very difficult to avoid using relatives for 'with' constructions. However, it was decided to use a variety of different relatives in context sentences. Thus, four types of constructions were compared: no relatives, "That" relatives, "Who" relatives and "Which" relatives.

THAT CONSTRUCTIONS

Unlike the 'with' stories, using the contextual information for a non-relative (complement) completion for 'that' stories did not involve merely a simple repetition, because unlike prepositional phrases, complements are complete clauses rather than simple phrases. For example: 'Jane had lost her mummy....Jane said to the man that SHE had lost her mummy'.

However, there was a very strong complement bias from both children and adults for 'that' stories in the previous experiment. It was suggested that this was possibly due to the fact that subjects were reluctant to repeat relatives that had already been used in the context sentences.

Therefore the 'that' stories of this experiment did not use any context relatives. For example:

Jane had lost her mummy at the fair.

By the gate, a man was selling sandwiches and another man was selling balloons. Jane said to the man that.....

As for the previous experiment, the 2 main variables manipulated were:

1. SEMANTIC CONTEXT: number of potential referents (1 OR 2)
2. PRAGMATIC CONTEXT: whether one referent was plausible or not.

An example of stories used in each condition are presented below: Pragmatic associations are shown in brackets. Potential referents as well as incomplete reference in the final sentence are underlined.

PLAUSIBLE

ONE REFERENT

A farmer had (just planted some vegetables).

His pig was (treading on his carrots) and his horse was asleep.

The farmer (shouted) to the pig that.....

TWO REFERENTS

Betty (wanted somebody to partner her in a game of tennis).
She saw a boy (at the tennis club) and another boy at
the swimming club.

Betty told the boy that.....

NEUTRAL

ONE REFERENT

A headmaster was praising somebody for getting ten out of
ten.

A teacher was sitting by the window and a pupil was
sitting by the door.

The headmaster told the pupil that.....

TWO REFERENTS

Jane had lost her mummy at the fair.

By the gate, a man was selling sandwiches and another
man was selling balloons.

Jane said to the man that.....

Each subject received each condition for both SEMANTIC
CONTEXT and PRAGMATIC CONTEXT.

When one referent was plausible or when there was only one
potential referent, the position (order of mention) of
intended and unintended referents in the 'relative
information sentence' was counterbalanced across all
conditions.

An equal number of 'That' and 'With' constructions were presented across all conditions.

As the previous experiment found no effects of the order in which the 2 complete sentences were presented, all subjects received the 'non-relative information sentence' first and the 'relative information sentence' second.

Each subject also received the 16 filler stories that had been used in the previous experiment. Therefore, each subject received 48 incomplete stories (32 experimental and 16 filler stories) and these were presented in a random order, both on the questionnaires and when they were presented verbally by the experimenter.

B. PROCEDURE

For verbal presentations, the procedure was the same as for the previous experiment and the same practice trials were used. For the questionnaire (see appendix), each subject remained anonymous and the instructions at the top of the paper were as follows:

PLEASE READ EACH OF THE FOLLOWING INCOMPLETE MINI-STORIES. FOR THE THIRD SENTENCE OF EACH STORY YOU WILL SEE THAT THE SENTENCE HAS NOT BEEN COMPLETED. YOUR TASK IS TO COMPLETE THE FINAL SENTENCE OF EACH STORY.

Scoring: As for the previous experiment.

RESULTS

Again, the 'fillers' scores showed that virtually all completions were coherent with context: made reference either explicitly or implicitly to information presented in the previous sentences. For reference to context, the mean scores out of 8, were as follows: BECAUSE: 7.7. IN: 7.7.

Experimental Stories

The completion scores for 'With' and 'That' constructions have been looked at separately.

WITH CONSTRUCTIONS

Table 16-A presents total scores for relative and preposition completions, as a function of both semantic and pragmatic context for 'with' constructions. Children (8 to 11 years) and adult scores are shown separately.

	RELATIVE			PREPOSITION		
	1 REF.	2 REF.	TOTAL	1 REF.	2 REF.	TOTAL
<u>PLAUSIBLE</u>						
8-11	35	58	(93)	37	14	(51)
ADULT	8	34	(42)	64	38	(102)
<u>NEUTRAL</u>						
8-11	50	66	(116)	22	6	(28)
ADULT	20	30	(50)	52	42	(94)
<u>TOTAL</u>						
8-11	(85)	(124)	(209)	(59)	(20)	(79)
ADULT	(28)	(64)	(92)	(116)	(80)	(196)

TABLE 16-A: RELATIVE AND NON-RELATIVE COMPLETIONS AS A FUNCTION OF BOTH THE SEMANTIC AND PRAGMATIC CONTEXT FOR 'WITH' CONSTRUCTIONS.

Firstly, there is a clear effect of semantic context in both age groups: more relatives when two potential referents and more prepositional phrases when one referent. Analysis of variance showed a significant effect of semantic condition on completion preference (8-11: $F=40.03$, $df=1,16$, $p<0.001$ Adult: $F=11.55$, $df=1,16$, $p<0.005$).

For the children, there also seems to be an interaction between pragmatic context and completion type: more relatives when neutral and more prepositional phrases when one plausible referent. This effect was confirmed by analysis of variance ($F=24.05$, $df=1,16$, $p<0.001$). However,

for the adults, pragmatic effects were only small for 'with' stories and no significant interaction between pragmatic context and completion type was found for the adult group ($F=1.21$, $df=1,16$, NS).

The table shows that the strong effect of semantic context on completion type for 'with' stories occurred for both plausible and neutral conditions. For the children, the pragmatic effects occurred when there was both one referent and when there were two referents available and no interaction was found between semantic and pragmatic context for the children ($F=1.32$, $df=1,16$, NS). However, the slight effect of pragmatic context on completion preference for adults was only confined to contexts where there was only one available referent and this was confirmed by analysis of variance which showed that there was a significant interaction between semantic and pragmatic context on type of completion for adults ($F=6.52$, $df=1,16$, $p<0.05$).

It is clear that, whilst the children preferred relative clauses, adults preferred prepositional phrases twice as often as relatives for 'with' constructions. The difference between relative and preposition completions was significant for both age groups using analysis of variance (8-11: $F=35.50$, $df=1,16$, $p<0.001$ Adult: $F=18.94$, $df=1,16$, $p<0.001$).

There were no effects of presentation form on completions (verbal versus questionnaire group): There was no

interaction between semantic context and form of presentation (8-11: $F=3.18$, $df=1,16$, NS Adult: $F=2.28$, $df=1,16$, NS). Nor was there an interaction between pragmatic context and form of presentation (8-11: $F=0.05$, $df=1,16$, NS Adult: $F=1.21$, $df=1,16$, NS). Moreover, there was no 4-way interaction between presentation type, semantic context and pragmatic context on completion scores (8-11: $F=0.68$, $df=1,16$, NS Adult: $F=0.41$, $df=1,16$, NS) and finally, there was no overall interaction between completion preference and form of presentation of sentences (questionnaire verses verbal) in either age group (8-11: $F=1.89$, $df=1,16$, NS Adult: $F=0.11$, $df=1,16$, NS).

Table 16-B presents total relative scores as a function of type of context: no relative, 'that', 'who' or 'which' relatives, for 'with' constructions, keeping children's and adult's scores separate.

RELATIVE SCORES				
	'THAT'	'WHO'	'WHICH'	NO RELATIVE
8-11	43	51	61	58
ADULT	18	20	22	32
TOTAL	61	71	83	90

TABLE 16-B: RELATIVE SCORES AS A FUNCTION OF CONTEXT (NO RELATIVE, 'THAT', 'WHO' OR 'WHICH' RELATIVES), FOR 'WITH' CONSTRUCTIONS.

There appears to be little difference in relative completion scores between 'that', 'who' and 'which' relative contexts for 'with' stories. Most relative completions occurred when there was no relative used in the context. However, this effect only applies to the adult group. In fact, differences in scores as a function of context form was only significant for the children using analysis of variance (8-11: $F=7.32$, $df=3,48$, $p<0.001$ Adult: $F=2.51$, $df=3,48$, NS). This effect for the children is due to the fact that relatives were used less when 'that' relatives appeared in context than for any other context form. Finally, there was no interaction between context form and presentation form (questionnaire, verbal) on relative completions (8-11: $F=1.37$, $df=3,48$, NS Adult: $F=1.12$, $df=3,48$, NS).

THAT CONSTRUCTIONS

Table 16-C presents total scores for relative and complement completions, as a function of both semantic and pragmatic context for 'that' constructions. Children and adult scores are shown separately.

	RELATIVE			COMPLEMENT		
	1 REF.	2 REF.	TOTAL	1 REF.	2 REF.	TOTAL
<u>PLAUSIBLE</u>						
8-11	18	51	(69)	54	21	(75)
ADULT	4	46	(50)	68	26	(94)
<u>NEUTRAL</u>						
8-11	32	42	(74)	40	30	(70)
ADULT	0	26	(26)	72	46	(118)
<u>TOTAL</u>						
8-11	(50)	(93)	(143)	(94)	(51)	(145)
ADULT	(4)	(72)	(76)	(140)	(72)	(212)

TABLE 16-C: RELATIVE AND NON-RELATIVE COMPLETIONS AS A FUNCTION OF BOTH THE SEMANTIC AND PRAGMATIC CONTEXT FOR 'THAT' CONSTRUCTIONS.

Again, for both children and adult groups, there is a clear effect of semantic context: more relatives for two potential referents and more complements for one referent. This effect was confirmed by analysis of variance (8-11: $F=45.94$, $df=1,16$, $p<0.001$ Adult: $F=45.11$, $df=1,16$, $p<0.001$).

For 'that' stories, whilst there is no clear effect of pragmatic context on completion type for the children, as confirmed by analysis of variance ($F=0.87$, $df=1,16$, NS), pragmatic effects for the adults were the reverse of those for 'with' stories: more relatives for the plausible

condition and more complements for the neutral condition. This effect of plausibility on completion type for the adults was found to be significant ($F=31.14$, $df=1,16$, $p<0.001$).

The table also shows that for 'that' stories, whilst there was no completion preference for the children, adults clearly preferred complement completions compared with relative completions, using 3 times as many complements as relatives. In fact, analysis of variance showed no significant difference between relative and complement scores for the 8-11 year olds ($F=0.01$, $df=1,16$, NS), but there was a significant difference for the adults ($F=38.37$, $df=1,16$, $p<0.001$).

Again, there was no interaction between semantic context and presentation form (8-11: $F=0.22$, $df=1,16$, NS Adult: $F=1.41$, $df=1,16$, NS). Also, there was no interaction between presentation form and pragmatic context (8-11: $F=0.04$, $df=1,16$, NS Adult: $F=3.46$, $df=1,16$, NS). Finally, there was no interaction between presentation form and type of completion overall for the children ($F=1.41$, $df=1,16$, NS), despite the fact that there were more relatives used by the 'verbal' presentation group (80) compared with the 'questionnaire' group (63) for the children. Similarly, there was no interaction between presentation form and type of completion overall for the adult group ($F=1.20$, $df=1,16$, NS).

Another effect shown in table C is that the strong influence of semantic context for 'that' story completions applied to both plausible and neutral conditions. However, the reverse effect of pragmatic context for adults mainly applied to contexts where there were 2 available referents. In fact, the children showed a 'reverse' pragmatic effect when there were two available referents, but tended to show the 'usual' pragmatic effect (more relatives when neutral and more complements when plausible) when there was only one available referent. There was an interaction between semantic and pragmatic context on type of completion for both age groups (8-11: $F=14.59$, $df=1,16$, $p<0.005$ Adult: $F=22.26$, $df=1,16$, $p<0.001$) as well as a 4-way interaction between presentation form, semantic context and pragmatic context on completion scores (8-11: $F=0.69$, $df=1,16$, NS Adult: $F=5.57$, $df=1,16$, $p<0.05$).

In order to account for the 4-way interaction above, table 16-D presents total relative and complement scores in each condition for 'that' stories, in the adult group only, showing the 'questionnaire' group and the 'verbal' group separately. It is clear from this table that the reverse pragmatic effects occurring when there were two potential referents were due to the 'verbal' presentation group and not the 'questionnaire' group.

ADULTS.	RELATIVE		COMPLEMENT	
	1 REF.	2 REF.	1 REF.	2 REF.
<u>PLAUSIBLE</u>				
QUESTIONNAIRE	2	24	34	12
VERBAL	2	22	34	14
<u>NEUTRAL</u>				
QUESTIONNAIRE	0	18	36	18
VERBAL	0	8	36	28

TABLE 16-D: RELATIVE AND NON-RELATIVE COMPLETIONS AS A FUNCTION OF BOTH THE SEMANTIC AND PRAGMATIC CONTEXT FOR 'THAT' CONSTRUCTIONS, LOOKING AT QUESTIONNAIRE AND VERBAL PRESENTATION SEPARATELY: ADULT GROUP.

DISCUSSION

Just as for the previous experiment, the clearest and most reliable effects appear to be due to the semantic context for both 'THAT' and 'WITH' types of constructions.

Although the 'with' stories had been reconstructed so that any relative bias due to the materials themselves would be eliminated, children still preferred relatives rather than prepositional phrases for 'with' constructions.

For 'that' stories, it seems that the new materials (which avoided using context relatives) did serve to increase relative completions, although the largest

increase in relatives was shown by the children, especially those receiving 'verbal' presentation. The average age of children in this experiment was slightly higher than for experiment 15 (9.9 compared with 9.0) and this may partly account for the overall increase in relatives in this experiment. Nevertheless, the reduction of complements still failed to reveal any clear effects of pragmatic context on completion preference for 'that' stories. In fact, for both children and adults, there was some evidence of pragmatic effects for 'that' stories which were the reverse of those predicted. In any case, for the adults this effect was only due to the 'verbal' presentation group and such effects were confined predominantly to contexts where there was more than one referent.

Whilst, pragmatic effects were as predicted for the children's 'with' stories (more relatives for neutral condition and more prepositional phrases for plausible condition), any pragmatic effects for adult's 'with' stories were very slight and confined to the semantic context of only one potential referent. Thus, the adult pragmatic effects cannot be due to their appreciation of the need to disambiguate reference when there was more than one potential referent.

Therefore, it appears that mature individuals are reluctant to rely on pragmatic information to disambiguate reference, presumably because plausibility is flexible and unreliable, whereas semantics is an absolute indication of ambiguity.

The fact that the results showed little evidence of subjects using less relatives when there was a plausible referent and more relatives when there was no plausible referent (except for the children on 'with' stories), means that there is insufficient support for Tyler and Marslen-Wilson 's (1977) claim that pragmatic context has the largest influence on anaphoric processing.

It seems that the strongest effect is 'semantic context'. Therefore, the results support Crain and Steedman (1985) and Altmann and Steedman's (1988) theory of anaphoric processing, which states that it is the semantic context (number of available referents) that interacts with syntactic possibilities. The evidence here that similar interactions between semantic context and completion possibilities can be shown using story completion tasks, weakens any criticisms of A&S's theory which argue that their findings were simply a reflection of using on-line reaction time tasks.

7.6 GENERAL DISCUSSION

The findings of this chapter have shown firstly, that young children are capable of the syntactic constraint of definite reference known as principle C of binding theory: Even 4 year olds were more likely to reject that a noun phrase was co-referential with a pronoun which preceded it in the main clause (higher in the phrase structure tree), as they preferred to choose a referent outside the sentence. Moreover, as it was the older children (9 to 11 year olds) who showed the greatest outward response on unrestrictive backward anaphora examples, it is possible that, rather than the overapplication of BAR being a parameter setting phenomenon, (at least for the older group of children), the outward bias for backward anaphora may be due to a preference of certain strategies, such as the minimum distance principle. This explanation would also account for the finding that outside preferences for backward anaphora were reduced in cataphoric examples, where the nearest referent would be the one mentioned in the pronominalised sentence.

Furthermore, young children could use pragmatic factors to determine reference even when reference was not constrained syntactically. The fact that younger subjects were more influenced by pragmatic context than were older subjects may be due to the use of different types of sentences for younger groups compared with older groups: It is possible that the more difficult constructions for older groups were also more ambiguous pragmatically.

However, young children failed to show any knowledge of the fact that both linguistic and non-linguistic factors are responsible for influencing reference because, when principle C conflicted with pragmatics, children under 9 years of age failed to show any evidence of an awareness of the conflict. Instead, they either accepted only the linguistic interpretation (3 to 5 year olds) or only the pragmatic interpretation (6 to 8 year olds). The fact that 3 to 5 year olds were less influenced by plausibility in the implausible condition than were the 6 to 8 year olds may be due to differences in procedure between these two groups: 3 to 5 year olds often chose a referent before the referential question had been asked for the second time by the experimenter and only when they were asked the memory or control questions, did they stop to think about the content of the story. Interestingly, there was no evidence of 3 to 5 year olds producing more memory errors in the implausible condition than in any other condition, which suggests that the youngest children made no effort to integrate their initial choice of a referent with the content of the story, when thinking about their answers to the memory and control questions.

Alternatively, the 6 to 8 year olds spent more time thinking about the first question that was asked (either referential, memory or control question) and spent less time thinking about the second and third questions regardless of the order in which referential, memory and control questions were presented. By the time the second and third questions were asked, they had presumably already

disambiguated reference after reflecting on the content of the story. Thus, in the implausible condition, their concern for plausibility tended to override principle C.

Older children and adults showed a preference for the same response (either 'yes' or 'no') for questions in the conflict condition, suggesting that they were aware of the presence of two possible interpretations: pragmatics and syntax. The 'yes' preference of 9 to 11 year olds, regardless of which referent had been chosen, shows that they were aware of the possibility of both a linguistic and a non-linguistic interpretation. The fact that adults preferred a 'no' response, regardless of chosen referent, suggests that they were aware that neither referent could be selected when there were both interpretations (linguistic or pragmatic).

Moreover, answers to the subsequent questionnaire, showed that adults most often claimed that some of the sentences were unnatural and ambiguous and that some of the questions were difficult to answer, which further suggests that adults were aware of the conflict in some of the stories.

These findings seem to be in accordance with the spontaneous comments made by subjects in Chomsky's and in Balfour's experiments. However, as adults are more likely than children to communicate with and criticise an adult experimenter, spontaneous comments alone are insufficient evidence for making any affirmative conclusions. The Yes/No biases found in this experiment provide stronger evidence for concluding that, with development, subjects become

increasingly able to consider both linguistic and pragmatic factors when interpreting discourse.

The students showed the largest difference between plausible and implausible conditions: Whilst showing a clear preference for the most plausible referent in the 'plausible' condition, they showed a slight bias against the most plausible referent in the 'implausible' condition. This finding, together with the fact that students failed to show any difference between neutral and plausible conditions, suggests that for students, the linguistic constraint was more important to them than the pragmatic context.

This result contrasts with the results of Balfour's adult subjects who were more influenced by pragmatics for the conflict sentences. However, in this experiment, adults interacted with a computer. This procedure may have encouraged students to see the task as one which required objectivity and therefore, one which required paying attention to the linguistic features more than the content of the stories. Presumably, parents are less likely to be faced with a situation which requires objectivity and this may account for the differences between parents and students in this experiment: When pragmatic and linguistic factors were in conflict, parents were more pragmatic and students were more linguistic.

Only the 3 to 5 year olds showed no differences between anaphoric and cataphoric conditions on neutral trials, a result which conflicts with Umstead and Leonard's finding (1983) that 3 to 5 year olds find

cataphoric pronouns harder than anaphoric ones. However, it seems possible that cataphoric examples were less natural than anaphoric examples and therefore, subjects inferred that the pronoun referred to some individual not mentioned in the story. The lack of any effects of context position in the youngest group may therefore be due to the fact that the type of task given to this group meant that they had to choose a particular referent rather than simply to either agree or disagree with the experimenter. This would also be consistent with the results showing that 3 to 5 year olds made more outward choices for cataphoric examples in the tangible group compared with the intangible group, because for the former, the subject's choice of a referent was restricted to either one of the two available referents.

Despite the presence of objects in the tangible condition, only two subjects (both in the 3 to 5 year old group) relied on manual responses when referents were tangible, but after three trials (one of these subjects) and 7 trials (the other of these subjects), manual responses were no longer solely used as a response without linguistic confirmation and after the eighth trial (the first subject) and the twelfth trial (the second subject), manual responses were no longer used at all: The children relied solely on linguistic responses. Thus, whilst tangibility served to improve scores slightly in the implausible condition, the objects themselves were of less importance to subjects than were the roles and position of individuals in the stories.

This experiment provides a clear example of the problems encountered when testing a wide range of subjects and it is necessary to consider these procedural variations as potential sources of any observed group differences in scores, before making any affirmative conclusions about underlying developmental effects.

Nevertheless, at first, the results of experiment 14 appear to be consistent with Goodluck's (1988) sequential model. She claimed that lexical and syntactic analyses occur initially and only later is this information integrated into a representation of the entire discourse. She supported this claim with experimental data which showed that, whilst young children could use contextual and linguistic information for pronominal reference, they failed to use pragmatic and discourse information to help them out with sentence-level syntactic operations. There is also experimental evidence to support this claim (Stevenson and Pickering, 1987).

Thus, Goodluck argued that the integration of pragmatic and discourse based information into the syntactic parser of a sentence develops only in the school years because the immature processor has less capacity than the adult processor and thus, as the use of discourse information is the last analysis to be done, it is the inefficient use of the latter, which tends to cause problems. Other investigators have also stressed the importance of considering processing limitations, when testing children's understanding of anaphoric devices (Stevenson 1988a).

However, the story-completions experiments showed that it was the particular informational context (number of available referents) as opposed to the general pragmatic context (plausibility of referents) which had the greatest influence on anaphoric reference for older children and adults, because relative completions were preferred for contexts where there was more than one available referent, whereas non-relative completions were used mainly when there was only one referent available. This supports the theory that, rather than 'independent' processing (Goodluck, 1988), anaphoric definite reference requires a 'weak interaction' between linguistic (structural) and non-linguistic (contents of discourse model) factors. This was the theory proposed by Crain and Steedman (1985) and Altmann and Steedman (1988).

By contrast, experiment 15 provided little support for Tyler and Marslen-Wilson's strong interactionist theory, because any effects of pragmatic context on choice of completion were small and inconsistent.

Although, for experiment 16, stories were reconstructed in an attempt to rule out potential experiment-generated biases, the only differences were that relatives were increased for 'that' stories and the children showed an effect of pragmatic context for the 'with' stories but not for the 'that' stories.

Moreover, in experiment 16, adults in the 'verbal' presentation condition showed the 'reverse' effect of pragmatics for 'that' stories, whereas this effect was only shown for the children in experiment 15. One

possibility for the 'reverse effect' is that subjects were unlikely to choose a referent if neither one was plausible and therefore, they were not likely to use a relative construction when they were unsure which individual should be referred to. The fact that these reverse pragmatic effects were not revealed for the 'with' constructions may have something to do with the possibility that 'with' is a more natural word for allowing a specific referent to be chosen even when the context does not determine which of the available individuals to refer to. Also, the fact that these effects only occurred for the 'verbal' presentation group for adults may be because questionnaires enable subjects to make their own choices more freely than in the interview situation.

Nevertheless, whilst the effects of semantic context were clear cut in both experiments, there seemed to be no consistent effects of pragmatic context on type of completion.

If logical modularity is compatible with parallel processing (Crain and Steedman/Altmann and Steedman) then it becomes difficult to determine whether any developmental differences in the ability to effectively integrate linguistic and contextual information would be due to processing factors per se or due to differences in an ability to appreciate the need to integrate all factors to prevent ambiguous definite reference. Thus, with regards to the Altmann and Steedman model and the results of the latter experiments, it seems that Goodluck's evidence for a processing explanation is insufficient.

CHAPTER 8
DISCUSSION

8.1 Summary of results

The experiments reported in the previous five chapters were investigating children's understanding and use of indefinite and definite referring expressions. The first three experiments were reported in chapter 3 which looked at 'the effects of listener familiarity'.

The results of experiment 1 showed that even children approaching 4 years of age would adapt their choice of expression as a function of the listener's perception: Ignorant listeners received predominantly indefinite expressions and knowledgeable listeners received mainly definite expressions. However, the results of experiments 2 and 3 showed that the young child's choice of expression was influenced most strongly by their general knowledge: Firstly, when the experimenter was a listener (experiment 2), regardless of the listener's role as ignorant in the picture-story task, young children treated her as knowledgeable, using definite articles to mention referents for the first time. However, they were able to treat an experimenter as ignorant when an elicited-discussion procedure was used, presumably because the child was able to believe that the experimenter was genuinely ignorant. Thus, it appears that when experimenters have acted as listeners in experiments looking at young children's references (Karmiloff-Smith 1979), the results may have been partly a reflection of young children's expectations

about the knowledge of the listener, rather than being an indication of their knowledge of the listener as it is defined in the experimental task (Emslie 1986).

As further evidence for the major influence of the child's general knowledge, experiment 3 showed that, even when a peer listener was unable to see referents, young children were likely to assume that they were knowledgeable if the referent was predictable from the story's context. Moreover, young children were so influenced by context in their choice of expression that, even when the context in the pictures was atypical of real life situations, they were still more strongly influenced by prototypicality, whereas older subjects were able to accept that some of the pictures portrayed idiosyncratic contexts.

These results suggest that young children are highly sensitive to social and pragmatic factors in their choice of expressions. Thus, they appear to be aware of the significance of 'familiarity' in determining in/definite reference.

Chapter 4 described two experiments that investigated 'the effects of referential 'specificity''. Both experiments involved reference to an exophoric context (a concrete array of objects), where some objects were specific (singleton) and some were non-specific (identicals). Experiment 4 manipulated both listener 'familiarity' and referential 'specificity', whereas experiment 5 manipulated the salience of the referential array, by increasing the size of non-specific sets and

by requiring indirect reference that involved looking at the remaining objects to discover what the hidden object was. The purpose of the latter experiment was to see whether sensitivity to 'specificity' in determining choice of expression was a function of whether the child's attention was focused onto the similarities and differences between the available referents.

The results of experiment 4 showed that children under 8 years of age were poor at adapting their choice of expression as a function of 'specificity': They failed to use more indefinites for non-specific items or use more definites for specific items. They were mainly influenced by the listener's ability to see referents in their choice of expression. However, when the referential array was made salient to them (experiment 5), the young child did show a sensitivity to 'specificity': When the listener was familiar with the array, speakers used more indefinite articles for non-specific items and more definite references for specific items.

These findings suggest that, whilst young children can discriminate between non-specific and specific items in their choice of indefinite or definite expression when the referential array is salient to them, they appear to lack the knowledge that 'specificity' is significant for in/definite reference in all situations. These results are also consistent with Karmiloff-Smith's claim (1979) that tasks looking at indirect reference lead to a greater production of indefinite expressions from young children and possibly, they also account for why Emslie

observed young children showing an ability to discriminate between non-specific and specific items in her version of Karmiloff-Smith's 'hide-and-seeK' task. Moreover, the improvement in performance for comprehension tasks and when there was a larger number of items in a non-specific set that was shown by Karmiloff-Smith's subjects may be accounted for by the fact that the referential array was more salient to the young children in these tasks. Furthermore, the evidence that non-specific/specific discrimination was better in Emslie's 'farmyard' experiment when subjects were able to choose an item for themselves, as opposed to when the experimenter chose the item, may be because playing an active role in choosing a referent helps the children to focus on the referential array and thus, may help to increase their sensitivity to 'specificity'. Finally, the results of Emslie's 'balancing' task may also be due to the fact that children were given the chance to focus on all the items on the tray and thus, the referential set was salient to them.

The six experiments in chapter 5 looked at 'reference in discourse production.' Experiment 6 manipulated both 'familiarity' and 'specificity' in a picture-story experiment. Experiments 7 to 9 were mainly concerned with anaphoric reference for referential communication. Experiment 10 used a more naturalistic 'model village' communication task and looked at a number of different expressions including pronominal reference. Experiment 11 looked at pronominal reference in more detail, manipulating lexical and pragmatic ambiguity.

The results confirmed the findings from the previous chapter that, just as for exophoric reference, young children were insensitive to 'specificity' in discourse. They failed to use indefinite expressions for non-specific items and they produced ambiguous anaphoric reference, failing to use relevant modifiers and/or relative clauses to make a referent specific for the listener (experiments 6, 7, 9 and 10). The results of experiment 7 were that young children failed to discriminate between a pair of individuals for anaphoric reference when relative clauses were required. Experiment 8 then showed that these results were probably confounded by processing difficulties because, using the same procedure as Hamburger and Crain, young children could use relative clauses when they only had to produce noun phrases as opposed to complete sentences. However, a similar proportion of children used relative clauses to discriminate between referents when they were of different-gender as when they were of the same-gender, despite the fact that, for the former condition, they were redundant (gender information alone would have sufficed). Thus, it appears that 4 year olds preferred relative clauses in experiment 8 because they needed to use their general knowledge about characteristics in order to answer the questions in the first place.

The results of experiment 9 showed that even when pragmatic, memory and salience factors were controlled, up to the age of about 9 years, children did not spontaneously use relative clauses to make reference unambiguous when

the information required was not salient to them (character information). These results suggest that the redundant relatives in experiment 8 were a reflection of the type of task used (choice-decision task), which served to make the information supplied by the relative clause salient for the children. Experiment 10 showed that even when young children could produce discriminations that were explicitly requested of them, they were still unable to make spontaneous discriminations for anaphoric reference. Emslie claimed that young children can use relevant discriminating 'colour' information when the task is one that clearly requires unambiguous reference and when misperception of the instructions, competitive and boredom factors are prevented. However, looking at the task where Emslie did find evidence of 'colour' modifiers (the 'balancing task'), it seems possible that subjects could merely have been using colour because they could see the mis-match between their own and their listener's array, rather than because they appreciated the need to use colour to prevent a mis-match between the arrays.

Moreover, the same findings were found for pronominal reference and despite the young child's sensitivity to lexical constraints and pragmatic factors (experiments 10 and 11), they provided ambiguous pronouns when references could not be disambiguated by lexical or pragmatic means (experiment 11). The lack of ambiguity for pronominal reference in experiment 10 was accounted for by the fact that, in the 'model village' task, the context usually allowed reference to be disambiguated through

thematic and/or pragmatic factors and thus, it was rarely the case that pronouns would have been ambiguous.

Experiment 7 also showed that young listeners tended to blame themselves for inadequate messages and experiment 10 showed that despite the greater number of ambiguous references from younger speakers, the young listeners seldom questioned their partners.

Chapter 6, 'Interpreting Anaphoric Reference' used comprehension tasks that explicitly manipulated ambiguity of reference. The tasks involved subjects moving items and giving them to dolls. Experiment 12 was a test of ambiguous reference using definite articles and experiment 13 looked at the comprehension of ambiguous pronominal reference.

These experiments showed that the comprehension of referential ambiguity followed the same developmental pattern as for the production tasks. For both pronominal reference and for reference using definite articles, the young listener failed to recognise referential ambiguity that was due either to lack of lexical constraints (experiment 13) or pragmatic information (experiments 12 and 13). This was revealed by their choice of one particular referent (experiments 12 and 13) and their non-linguistic justifications (experiment 12). Moreover, their failure to detect ambiguity was not merely a reflection of young children's 'performative biases' because, even when detection of ambiguity required a manual response (experiment 13), detection was still poor for young children.

The findings of chapters 5 and 6 suggest that children lack a knowledge of 'specificity' for in/definite reference in discourse just as they do for immediate reference. It appears that exophoric and anaphoric ambiguity in young children is related to salience factors, suggesting that the young child lacks an appreciation of the necessity of 'specificity' for determining in/definite reference.

Finally, chapter 7 involved three experiments which were attempts to look at differences between the child's understanding of linguistic constraints and non-linguistic factors in isolation compared with their understanding of the interaction between linguistic constraints and non-linguistic factors, as it seems that only the latter shows an understanding of the 'specificity' requirement of in/definite reference. Moreover, the way that context and linguistic restrictions interact in processing anaphoric reference was also looked at in experiments on older children and adults.

Experiment 14 looked at the comprehension of principle C of binding theory and manipulated whether or not pragmatic contexts were consistent with linguistic constraints. Experiments 15 and 16 were story completion tasks aiming to look at whether anaphoric processing in older children and adults was a function of both contextual and syntactic factors. The experiments were also interested in the type of context which had the strongest influence on choice of completion: either Crain and Steedman (1985) and Altmann and Steedman's (1988) theory

that adults use the particular contextual information about the number of available referents to decide upon choice of definite construction, or Marslen-Wilson and Tyler's theory (1982, 1983) that anaphoric mapping involves the same processes as discourse in general, where pragmatic context has the overriding effect.

The findings of this chapter suggest that young children show an awareness of certain linguistic constraints (principle C) and can use general knowledge to make pragmatic inferences well before they show an ability to consider linguistic factors and non-linguistic factors simultaneously. This is reflected by their lack of sensitivity to a conflict between linguistic non-identity and pragmatic context (experiment 14). However, from 9 years of age, it seems that children's performance is affected by two discrete components: linguistic and non-linguistic, as reflected by their failure to support only the pragmatic or only the linguistic interpretation when both possibilities were in conflict.

However, the results of experiments 15 and 16 suggest that, whilst pragmatic factors may have some effects in determining form of expressions, the major factor influencing the processing of anaphoric reference is the number of available referents introduced in the preceding context. When there was more than one available referent, 9 year olds and adults preferred relative completions to non-relative completions, but when there was only one referent available, non-relatives were preferred to relatives. Whilst there were some slight effects of



pragmatics, semantic effects tended to override any effects of plausibility.

From the results of experiment 14 it appears that the child's lack of appreciation of 'specificity' is reflected in the way that young children fail to appreciate both linguistic constraints and the non-linguistic context simultaneously.

The findings of experiments 15 and 16 suggest that knowledge of 'specificity' for definite reference does require an interaction between linguistic constraints and the particular context (number of available referents).

8.2 General Conclusions: Children's use and understanding of indefinite and definite reference.

The experiments of this thesis suggest that young children are aware that 'familiarity' is an essential determinant of in/definite reference, because when they defined the listener as knowledgeable they preferred definite expressions, whereas they preferred indefinite expressions when they defined the listener as ignorant. Moreover, they appear to be strongly influenced by their general knowledge in interpreting listener 'familiarity', because they preferred definite expressions when they were talking to an experimenter in a picture story task and when the context of a particular story implied that a referent already existed and was therefore, known to the listener. Thus, young children are highly sensitive to the social and pragmatic context when choosing between indefinite and definite reference, suggesting that the 'familiarity'

requirement of indefinite and definite distinctions is fully acquired at around . . . 4 years of age.

However, the young child appears to be unaware that 'specificity' is an essential determinant of in/definite reference. Whilst they already seem to be able to use indefinites preferably for a non-specific item and definites preferably to a specific item, they only use this information when the set of referents is salient to them.

This lack of appreciation of the requirement of 'specificity' is reflected even more strongly in discourse, where reference to a specific entity often requires a more complex definite noun phrase involving extra presuppositional information. Any unique features of referents in exophoric contexts are often salient because exophoric reference involves concrete objects and physical differences are readily perceivable. By contrast, for anaphoric contexts, differences between referents are not often the most salient features of the situation (ie. personality, character and episodic differences) and therefore, the fact that the young child fails to use spontaneously the relevant information for unambiguous anaphoric reference, even when the task is simplified to control for general processing difficulties, suggests that they are not aware of the need to use 'specificity' information for anaphoric reference, regardless of whether the information is salient or not.

Moreover, this developmental lag in the appreciation of 'specificity' seems to be a reflection of a general cognitive deficit, because the same effects were found

across different types of comprehension and production tasks and across different forms of definite reference (pronouns and definite articles).

One conclusion that can be made from these experiments is that certain aspects of pronominal reference develop at roughly the same time as they do for reference with definite articles. At least, using pragmatic inferences, appreciating the listener's knowledge and understanding ambiguity for anaphoric reference appear to be more general conceptual abilities, which apply to both forms of reference. By contrast, experiment 11 highlighted that there are particular aspects of pronominal reference which do not apply to the use of definite articles, such as the use of heuristic strategies (for example, the 'minimum distance principle') and lexical constraints.

Finally, it appears that anaphoric mapping involves a 'weak interaction' between the particular informational context (number of referents available) and possible syntactic constructions and thus, the evidence to suggest that a child has problems dissociating linguistic constraints from context, leads me to suggest that the problems for young children in failing to appreciate 'specificity' is that they are unable to allow their linguistic knowledge to interact with non-linguistic knowledge when using and interpreting indefinite and definite reference. If this is the case, it appears that any processing difficulties that young children have are a reflection of a general cognitive deficit: that of failing

to consider both linguistic and non-linguistic factors simultaneously.

8.3 Evaluation of Theoretical Perspectives regarding indefinite and definite reference.

A. Definition

In the introduction, it was concluded that the crucial factors for defining indefinite and definite reference were FAMILIARITY and SPECIFICITY. From the findings of the experiments in the current investigation, it appears that these factors are what children over 8 years of age and adults do actually use for indefinite and definite reference. However, the adult's interpretation of 'familiarity' often went beyond those features of the immediate situation. Older subjects were better at integrating immediate situational features with general knowledge. Moreover, adults often used the pictures as a general background only and thus, they probably treated a listener as ignorant when the stories they produced went beyond the information presented in the pictures. Also, their definitions of 'familiarity' seemed to depend on conventional story procedures, such as only treating a listener as ignorant if there was no main character or referent being talked about. Nevertheless, regardless of these differences, older subjects were discriminative in their choice of reference as a function of listener 'familiarity'.

With regards to 'specificity', it was clear that, when the listener was familiar with the referential array,

older children and adults still used indefinite expressions for non-specific items, even when the task did not serve to increase the salience of the referential array. Moreover, their knowledge of 'specificity' was clearly shown by their unambiguous anaphoric references, again regardless of the salience of information required to differentiate between potential referents. Finally, mature subjects did not choose one particular referent when a message was ambiguous.

Thus, the evidence from the current experiments suggests that 'familiarity' and 'specificity' are those factors that adults are sensitive to in their interpretation and use of indefinite and definite reference. Therefore, it appears that these factors are a valid criterion for measuring indefinite and definite knowledge.

B. Processing

It was argued in the introduction that Emslie's results (1986) failed to reflect knowledge of anaphoric reference, because she interpreted children's referring expressions from the way that referents were presented in the 'world', rather than from the way that referents are represented in the listener's 'model'. Yet, Emslie used the notion of a 'mental model' (Johnson-Laird, 1983, Stenning, 1978) to account for her findings. A 'mental model' stands for the way that speakers and listeners represent referents when social and pragmatic aspects of discourse are considered. It seems that in Emslie's effort to include potential

disparities between speaker and listener models, she failed to also interpret her results as showing whether her subjects were considering the 'specificity' of referents in a model.

Moreover, 'mental models' have no structure: They do not represent the structural constraints of language (c-command and local domain). Thus, in order to look at the type of information used by children in their use of definite descriptions, it is necessary to consider the way that adults process discourse for anaphoric reference.

The results of this investigation seem to support most strongly the processing theory of Crain and Steedman (1985) and Altmann and Steedman (1988). They put forward an on-line parallel 'weakly interactive' model, which suggests that the specific informational context (number of available referents) determine which of those linguistically constrained alternative definite constructions to rule out and that this weak interaction is incremental (as each word/phrase is encountered). The above investigators supported their theory from reaction time tasks on adults that showed 'garden path' effects only occurring when context did not 'referentially support' the final construction: when there were 2 potential referents and a complement construction was used (eg. Two ladies...the doctor told the lady that she was to return) and when there was 1 potential referent and a relative was used (eg. A lady...the doctor told the lady that was

ill to return). For the former, the context and construction are incompatible because reference is ambiguous, whereas for the latter, the context and relative are incompatible because reference is redundant.

Their experiments showed that the number of referents in discourse has an important effect on the way that language is processed. Sidner (1986) has also argued that it is necessary that all potential referents are considered when processing discourse. Thus, without an exhaustive check of all potential antecedents, reference may be ambiguous. I propose that, in order to process discourse in the way suggested by Crain et al, it is necessary to carry out the following procedures when using definite expressions for anaphoric reference:

1. To use the entirety of the listener's current discourse model, so that an exhaustive comparison of all referents is carried out. An exception to this may be when there are referents in focus.

2. To use the minimum available presupposed information for making one (and not more than one) referent in the listener's discourse model coreferential with the definite expression. If the intended referent is the only one in focus, a pronoun would be sufficient. However, if there are other referents also in focus or in the current discourse model, a full definite noun phrase is necessary to identify the intended referent.

From the final two experiments of this thesis, it appears that adults show a similar sensitivity to the effects of the informational context on form of completion after an ambiguous word. Relatives were preferred when there were 2 available referents and non-relative completions were preferred when there was only one referent available. Thus, it appears that a sensitivity to specificity for anaphoric reference requires this particular 'weak interaction' between semantic context and linguistic form.

However, if anaphoric processing follows the same principles for pronominal reference as it does for reference using full noun phrases, one might wonder how pronouns fit into the Crain and Steedman model. Stevenson and Pickering (in preparation) have suggested ways in which principles A and B of binding theory might interact with the contents of a discourse model, just as permissible structural analyses (relative/complement) can interact with semantic context. They also include lexical marking as another semantic component which can rule out potential syntactic readings for pronouns. Clearly, an experiment incorporating pronominal ambiguities as well as structural ambiguities is required to see if Altmann and Steedman's theory can be applied to anaphoric reference in general.

Moreover, as I will continue to discuss in the next section, I suggested that the same kind of interaction occurs for exophoric reference, the only difference being that the particular informational context is extra-linguistic as opposed to being in the context of

discourse.

Finally, there was little evidence for the primary role of pragmatic context in anaphoric mapping (Marlen-Wilson and Tyler 1982,1983) so presumably, whilst plausibility can reduce/eliminate semantic ambiguity, it is not a strong factor and adults seem to prefer to use extra presuppositional information when there are doubts about reference. If it is the case that pragmatics has no direct effect on syntactic ambiguity, then the findings of Ferriera and Clifton (1987) are probably due to the fact that they used a pragmatically biased context and hence found a non-interactive result: that context had no affect on syntactic 'garden pathing'. In fact, Stevenson and Pickering (in preparation) suggest that pragmatics can only indirectly affect syntactic choice by solving semantic ambiguities. Perhaps people are more sensitive to pragmatics when referents are highly plausible and when non-referents are highly implausible. Further investigation into this possibility should hopefully help to resolve this issue.

C. Development

Some investigations of children's knowledge of indefinite and definite reference were described in the introduction and it was concluded that there was not yet any affirmative evidence regarding the issue of whether young children use definite expressions 'egocentrically'. The evidence from the current investigation suggests that young children are not 'egocentric' and the results here support the findings

of Emslie and Stevenson (1981) and Emslie (1986) that even infants will use an indefinite article when introducing referents to an 'ignorant' listener. Therefore, it seems that the negative findings of previous studies are due to their failure to consider the young child's perception of the knowledge of an experimenter listener (that they are normally expected to be familiar with referents, regardless of their role in the experiment) and/or failed to consider the young child's use of pragmatic inferences in interpreting familiarity (that for contexts which predict referents, the listener can be treated as knowledgeable).

Moreover, the claim that understanding indefinite and definite reference requires a developmental progression from deictic to exophoric and finally to anaphoric reference (Lyons 1979, Karmiloff-Smith 1979) seems to be supported. At least the development from deictic to phoric reference does seem to reflect the results of this investigation. This evidence is shown by the fact that, for tasks which did not require comparing and differentiating between entities, the young child had no trouble referring to an entity in the immediate situation, their choice of reference form only being dependent on the social context: whether the listener was ignorant or knowledgeable. However, when reference required considering the referential array of objects or the number of referents in discourse, the young child failed to refer appropriately. Thus, the evidence that 'familiarity' is acquired before 'specificity' appears to be consistent with the claim that deixis is acquired before phoric reference.

However, as I argued earlier, the problems for children under 9 years seem to be due to 'specificity' generally, be it a function of reference to an exophoric array or anaphoric reference in discourse and thus, I suggest that any evidence of an intermediate phase of knowledge of exophora being acquired before anaphora is probably a fallacious conclusion, which is a consequence of the fact that exophoric contexts are normally more salient than anaphoric ones. Referring back to Hawkins' theory of definiteness, he argued that all definite descriptions must be inclusive within a 'shared set' regardless of whether this set consists of the contents of the discourse context (anaphoric) or of the immediate situation (exophoric). As the current investigation uses 'specificity' rather than inclusiveness to define definite reference (for reasons which are discussed in the chapter 1), I should like to make a similar argument to Hawkins by claiming that 'specificity' applies to contexts involving physical and concrete items as well as to contexts involving individuals and events which are less immediate or even hypothetical (current discourse models). Moreover, as physical differences between items are more likely to be salient features compared with episodic or psychological differences between individuals, the former information is more likely to be used by young children for this reason alone, regardless of the fact that this information makes the intended referent specific for the listener.

From the findings of the current experiments, I

suggest that the intermediate stage of development, rather than being that of acquiring exophoric reference, is one of implicitly knowing that non-specific reference involves indefinite expressions and that specific reference involves definite expressions. In fact, even the youngest children (3 to 5 year olds) did seem to be able to discriminate between specific and non-specific reference in this way. What they still seem to lack is an explicit awareness of the significance of 'specificity' for all forms of indefinite and definite reference: that definite expressions can only be used if they refer specifically.

The young child's reliance on salience factors and the evidence that young children fail to appreciate that linguistic factors and non-linguistic factors must both be considered for definite reference, seems to be in accordance with the claim that a general cognitive deficit of young children is their failure to dissociate 'appearance' from 'reality' and their failure to treat a situation meta-linguistically: to disembed language from context. Furthermore, it seems that the young child's failure to appreciate 'specificity' is a consequence of a subjective 'propositional attitude'. In brief, it seems that the child's problem stems from a failure to consider more than one piece of information at the same time (language and context). Instead young children are only concerned with that piece of information which is most salient to them in a given situation.

D. Summary of Conclusions

From the results of the current investigation it appears that the factors of 'familiarity' and 'specificity' are necessary for understanding indefinite and definite reference. Moreover, for processing anaphoric reference, where the information in the definite predicate is significant for making a referent specific, a 'weak interaction' between linguistic possibilities and semantic context appears to be necessary. The current findings also show that young children can use social and pragmatic factors in determining in/definite reference. Thus, it seems that even young children are aware of the significance of 'familiarity'. However, the evidence shows that young children have problems with anaphoric reference and fail to consider 'specificity' unless information is salient to them. These results, plus the findings which show young children failing to consider linguistic and non-linguistic factors simultaneously, suggest that the 'weak interaction' model of anaphoric processing is not being used by young children. An attempt to analyse the processes used by young children when interpreting definite expressions is discussed in the next section.

8.4 A proposed model of processing anaphoric definite reference in young children.

The Altmann and Steedman (A&S) model (1988) discussed above is a parallel, fine-grained, weakly interactive model of processing. Basically, this theory states that, whilst all potential constructions derived from syntax are parsed

simultaneously, the contents of the discourse model (semantic context) is significant for ruling out possibilities so that only one remains. Moreover, this 'weak interaction' occurs after each small unit of discourse is processed (word or phrase). For example, given the incomplete sentence, 'the doctor told the patient that..', at this point, there are two phrase structure possibilities that the listener can follow: a relative (... that was in bed) or a complement (...that she could get up). However, if the preceding sentences had been, 'The doctor was talking to two of his patients. One of them was standing by a sink and the other one had just woken up', in accordance with A&S's theory of 'referential support', the listener would be expected to rule out the complement construction and proceed with the relative construction, because A&S claim that in the context of 2 available referents a relative is chosen, whereas in the context of one available referent a complement is chosen.

However, the referent of a definite expression depends not only on syntactic and semantic knowledge, but also on general knowledge (pragmatics).

The results of experiment 15 and 16 suggest that pragmatic context does not apply to A&S's theory and the 'weak interaction' mentioned above applies predominantly to the effects of semantic context (number of available referents). Therefore, whilst pragmatic inferences are sometimes necessary to interpret discourse, pragmatics is not often relied on exclusively to determine form of definite anaphoric reference. This probably explains why,

in the preceding experiments, adults were reluctant to rely solely on pragmatic information for determining reference, as pragmatic information is not capable of ruling out possibilities.

Therefore, the 'weakly interactive' model proposed by A&S appears to be a more accurate account of anaphoric processing than the 'strongly interactive' model proposed by Tyler and Marslen-Wilson T&M (1977), who argued that pragmatic context itself can influence choice of reference. T&M probably failed to reveal the effects found by A&S because they manipulated pragmatic context as opposed to semantic context.

Whilst A&S failed to manipulate pragmatic context, the results of experiments 15 and 16 here show that semantic context rather than pragmatic context is significant for the 'weak interaction'.

The evidence showing that young children overuse definite expressions for non-specific referents when talking to a 'knowledgeable' listener, plus the results that show young children producing and failing to recognise ambiguous anaphoric reference as well as their lack of sensitivity to a conflict between linguistic and non-linguistic factors for definite reference, suggest that young children lack an understanding of the significance of 'specificity' for definite reference.

In fact, there is empirical evidence which shows young children failing to use the context for making decisions about which linguistic structure to use: Balfour (1983) found that young children failed to point out the

absurdity of sentences for which structural non-coreference restrictions conflicted with pragmatic context. Stevenson and Pickering (1987) found that children failed to use context to help them with decisions of pronominal reference constrained by syntax and gender.

As Goodluck's (1988) attempt to account for developmental effects in terms of a sequential model: the young child having insufficient processing capacity to integrate discourse based information into early syntactic and lexical analyses, appears to be insufficient, it seems that it is necessary to interpret the developmental effects as due to young children not using the A&S parallel weakly interactive model.

The evidence does suggest that young children are able to make pragmatic inferences in determining reference. However, it seems that, rather than using the pragmatic context to choose between possible forms of reference, young children are only using the context to decide upon a referent. Moreover, young children appear to be using the pragmatic context to help them choose a referent, but they will choose a referent anyway irrespective of whether it is more plausible than the others. Therefore, whilst it may appear that young children are using a 'strongly interactive' model, such as the one proposed by M&T, I argue that as young children seem to be using pragmatic information to help them choose a referent, rather than to influence their choice of a particular form of definite reference (ie. relative or non-relative), they cannot be using a 'strongly interactive' model either.

Given the evidence that young children's choice of construction seems to be predominantly influenced by the salient features of that particular situation, I suggest that young children follow the first analysis which comes to mind. Thus, they are using 'independent' processing rather than interactive and parallel processing and context is not being used to decide upon form of definite reference.

Furthermore, I speculate that, in addition to salience (the most readily available information as a function of contextual information), young children will be influenced by the most readily available information as a function of linguistic knowledge in their choice of definite construction. Thus, if experiments 15 and 16 were adapted to suit the needs of younger children, one might expect their completions to depend on 'minimal attachment' (Frazier 1987), because simple heuristics do not require an interaction of any kind between context and syntax. This strategy states that when parsing a sentence, the reader constructs a phrase structure that has the minimum number of possible nodes. Presumably, this structure would be the one most available to the young child. If young children are interpreting definite expressions in the way proposed by Frazier, one would also expect them to parse only one analysis and to regress backwards if they have chosen the wrong referent, in which case 'garden path' effects should occur whenever the most complicated analysis is the correct one. However, there is evidence to show that young children will prefer relatives in certain contexts (cf. experiment 8). Therefore, young

children may prefer a more complex phrase structure if the context makes that information more salient.

Finally, the question remains as to whether the problems for young children stem from processing limitations per se (adults have a more complex processor), or are a reflection of a more general cognitive inability: to appreciate the need to dissociate linguistic from non-linguistic aspects of discourse.

The PERFORMANCE explanation could account for age differences if it was found that children had difficulties holding both context and all syntactic possibilities in memory simultaneously. However, there is some evidence (experiments 8 to 10) to show that, even when information processing burdens are reduced, young children still fail to show a sensitivity to context in their choice of definite construction. The COMPETENCE explanation is consistent with evidence that young children hold 'subjective propositional attitudes' and fail to dissociate their own intentions from the particular words that they use.

8.5 Suggestions for further work

Given a task similar to those used for experiments 15 and 16 and given the evidence above, one would expect young children to be primarily influenced by linguistic structure and use linguistic strategies ('Minimal Attachment' Frazier, 1978), irrespective of context. Therefore, one would expect young children to prefer non-relative constructions regardless of whether the context provided referential support or not.

Obviously, a means of testing young subjects which controls for task complexity, memory limitations and which uses pragmatic knowledge more appropriate to the age of the subjects, is necessary.

The experimenter could read incomplete narratives to the child, such as, Two girls (A girl and a boy) were playing. One (the girl) had an ice-cream and the other one (the boy) had some chocolate. A lady came over and cleaned the face of the girl with a towel/the ice-cream.

These narratives could be followed by questions such as:

Was the lady calling the girl who had an ice-cream?

Was the lady calling the girl (boy) who had some chocolate?

Are you sure

Plausibility could also be manipulated. Given the evidence above, one would expect young children to always choose a referent and to say that they are sure, regardless of whether whereas the narrative was completed with a relative or not.

Another claim that has been made in the current investigation is that 'specificity' is the factor that children have problems with and this applies to both exophoric and anaphoric reference. One way of testing for this is to use an Object-location task (Wimmer and Perner type task) and to manipulate a) exophoric or anaphoric reference and b) salience of 'intension' or 'extension'. The prediction would be that children will show similar results for both exophoric and anaphoric reference and that the only differences would be

due to salience: One would expect better performance when the intension was salient (emphasising the knowledge of the seeker) than when the extension was salient (showing subjects the actual location).

Further work may therefore serve to provide stronger support for the claims made in the current investigation regarding children's knowledge of indefinite and definite reference.

ANNEXE:

COMMENTS ON EXPERIMENTAL RESULTS WHERE THE ROBUSTNESS OF ANALYSIS-OF-VARIANCE IS UNCERTAIN

Some of the ANOVA tests in this thesis were used on data in circumstances where it is unclear whether parametric statistics are robust, owing to:

- a) the nature of the measurement underlying the 'scores' (percentages or proportions)
- b) floor and/or ceiling effects, resulting in flat distributions in one or more of the experimental conditions

Presented below is a comment on each of these particular analyses, which should help to confirm the results already found and to resolve any likely uncertainties with regards to the power of the ANOVA statistics used in these situations.

A) The examples below are instances where the robustness of ANOVA statistics are questionable, but where the unanalysed scores alone show obvious and definite results (refer to relevant tables listed below). Therefore, even if one feels that the assumptions for the use of parametric statistics are not sufficiently met, one is confident that the results do actually represent the clear-cut effects shown in the data:

Experiment 1:

(table 1-B) - comparing definite articles and pronouns on second mention as a function of age group

(table 1-C) - comparing indefinite expressions for main and subsidiary referents on first mention as a function of age group

Experiment 4:

(table 4-B) - comparing indefinite and definite 'singleton' scores as a function of listener condition for the adult group

(table 4-E) - comparing indefinite and definite 'singleton' scores as a function of listener condition and age group

Experiment 5:

(table 5-A) - comparing indefinite and definite scores for 'non-specific' items by type of task and age group

(table 5-C) - comparing indefinite and definite scores for 'singleton' items by type of task and age group

Experiment 6:

(table 6-B) - comparing indefinite and definite scores as a function of listener condition for inanimate/specific items on first mention

(table 6-D) - comparing indefinite and definite scores as a function of listener condition for inanimate/specific items on second mention

Experiment 7:

(table 7-C) - comparing first verses second mention indefinite and definite scores as a function of age group

(table 7-D) - comparing correct and incorrect 'selection' scores in the 'Physical' group as a function of age group

(table 7-E) - comparing correct and incorrect 'selection' scores in the 'Episodic' group as a function of age group

Experiment 10:

(table 10-A) - comparing type of redundancy scores (physical, quantity and spatio-temporal) in each age group

Experiment 11:

(table 11) - comparing pronoun scores across all 4 conditions same/different gender and plausible/neutral for the pilot study

Experiment 12:

(table 12-F) - comparing type of justification for the ambiguity condition when the correct response was given (saying 'silly sentence') as a function of age group

B) There were some uncertain uses of ANOVA on scores which did not display such clear-cut results. For these cases, the addition of statistical confirmation would indeed be a valuable source of information. Therefore, appropriate non-parametric tests have been carried out on such data in order to confirm the present findings. This should help to eradicate any reservations held on the basis of having used dubious ANOVA tests. The results of these non-parametric statistics are presented below:

Experiment 1:

(table 1-A) - comparing indefinite scores between LK and LI conditions on first mention for each age group (Mann-Whitney U tests and Median tests)

3-5 years (Mann-Whitney U test)

mean rank

6.50 - LK(10)

14.50 - LI(10)

U=10 W=65 p<0.005 (2-tail)

Median tests were chosen for the 2 intermediate age groups, because all subjects had a ceiling score of 4 in one of the conditions for these age groups and therefore, scores cannot be ranked.

6-7 years (Median test)

	LK	LI
>median	0	0
<median	10	10
median=4	p=1 ns.	

8-9 years (Median test)

	LK	LI
>median	0	10
<median	10	0

median=3.5 p<0.001

10-11 years (Mann-Whitney U test)

mean rank
13.20 - LK(10)
7.80 - LI(10)
U=23 W=132 p<0.05 (2-tail)

Experiment 3:

(table 3-A) - comparing 'context definites' between all age groups (Mann-Whitney U tests)

3-5 by 6-7 years

mean rank
22.83 - 3-5(20)
18.17 - 6-7(20)
U=153.5 W=456.5 ns

3-5 by 8-9 years

mean rank
24.75 - 3-5(20)
16.25 - 8-9(20)
U=115 W=495 p<0.01 (2-tail)

3-5 by 10-11 years

mean rank
25.83 - 3-5(20)
15.18 - 10-11(20)
U=93.5 W=516.5 p<0.005 (2-tail)

6-7 by 8-9 years

mean rank
22.10 - 6-7(20)
18.90 - 8-9(20)
U=168 W=442 ns

6-7 by 10-11 years

mean rank
23.02 - 6-7(20)
17.98 - 10-11(20)
U=149.5 W=460.5 ns

8-9 by 10-11 years

mean rank
21.55 - 8-9(20)
19.45 - 10-11(20)
U=179 W=431 ns

(table 3-D) - comparing indefinite scores A) between LI and LK conditions in each age group (Mann-Whitney U tests) and B) between the 4 age groups (Kruskal-Wallis 1 way ANOVA test):

A) Mann-Whitney U tests

3-5 years

mean rank

6.15 - LK(10)

14.85 - LI(10)

U=6.5 W=61.5 p<0.001 (2-tail)

6-7 years

mean rank

8.90 - LK(10)

12.10 - LI(10)

U=34 W=89 ns

8-9 years

mean rank

7.20 - LK(10)

13.80 - LI(10)

u=17 w=72 P,0.05 (2-tail)

10-11 years

mean rank

12.65 - LK(10)

8.35 - LI(10)

U=28.5 W=126.5 ns

B) Kruskal-Wallis 1-way ANOVA

mean rank

29.60 - 3-5(20)

53.63 - 6-7(20)

36.72 - 8-9(20)

42.05 - 10-11(20)

chi-square=11.77 p<0.01

Experiment 7:

(tables 7-A and 7-B) - A) comparing discriminative scores in 'physical' and 'episodic' groups for each age group (Median tests) and B) comparing discriminative scores across the 4 age groups (Kruskal-Wallis 1-way ANOVA test)

A) Median tests

3-5 years

	P	E
> median	6	1
< median	4	9
median=0	p<0.05	

6-8 years

	P	E
> median	0	0
< median	9	10

9-11 years

	P	E
> median	0	0
< median	7	6

parent group

	P	E
> median	0	0
< median	4	4

median=1 (values all = or < median)

B) Kruskal-Wallis 1-way ANOVA test

mean rank
16.27 - 3-5(20)
34.82 - 6-8(19)
38.08 - 9-11(13)
43.50 - parents(8)
chi-square=26.66 p<0.000†

Experiment 9:

(table 9-A) - comparing 4 age groups' discriminative scores in the same gender/identicals condition (Kruskal-Wallis 1-way ANOVA).

same gender/identicals condition

mean rank
6.13 - 3-5(8)
14.81 - 6-8(8)
21.88 - 9-11(8)
23.19 - adult(8)
chi-square=17.84 p<0.0005

Experiment 12:

(table 12-A) - comparing appropriate responses in each condition +/- 'referential specificity' and +/- 'pragmatic associations' for each age group (Wilcoxon matched-pairs signed ranks tests)

3-5 years

+R v-R
mean rank
7.25 - 12(-R lt +R)
4.00 - 1(-R gt +R)
3(-R = +R)
z=-2.90 p<0.005 (2-tail)

+P v-P

mean rank
6.50 - 12(-P lt +P)
0.00 - 0(-P gt +P)
4(-P = +P)
z=-3.06 p<0.005 (2-tail)

6-8 years

+R v-R
mean rank
8.00 - 15(-R lt +R)
0.00 - 0(-R gt +R)
1(-R = +R)
z=-3.41 p<0.001 (2-tail)

+P v-P
mean rank
26.18 - 41(-P lt +P)
25.25 - 10(-P gt +P)
 13(-P = +P)
z=-3.85 p<0.0001 (2-tail)

9-11 years

+R v-R
mean rank
6.50 - 12(-R lt +R)
0.00 - 0(-R gt +R)
 4(-R = +R)
z=-3.06 p<0.005 (2-tail)

+P v-P
mean rank
6.40 - 10(-P lt +P)
7.00 - 2(-P gt +P)
 4(-P = +P)
z=-1.96 p<0.05 (2-tail)

parent group

+R v-R
mean rank
6.50 12(-R lt +R)
0.00 0(-R gt +R)
 4(-R = +R)
z=-3.06 p<0.005 (2-tail)

+R v-P
mean rank
5.07 - 7(-P lt +P)
8.50 - 5(-P gt +P)
 4(-P = +P)
z=-0.28 ns

