University of Oxford
Department of Experimental Psychology



From cat to dog:
Investigating organisation in the infant lexicon

Suzy J Styles

Thesis submitted for the Degree of Doctor of Philosophy
Trinity Term 2008

Acknowledgements

This thesis would not have been possible without the help and support of a great many people. I can only thank a few here – the rest of you will find out later… Firstly, I would like to thank my supervisor, Kim Plunkett, for his great science, and the rigour he has helped me achieve in my work. My lab mates have been excellent, particularly Vanja Kovic and Loukia Taxitari, who started out their doctorates with me, and have been invaluable travelling companions along the way. The Teddy Hall MCR has been a great outlet for intellectual banter, where Cara Krmpotich has been a star of stars – it was always much easier with two of me around. And what would I have done without my rowing girls? I also would like to thank my local friends, whose fine British accents provided enthusiastic baby-talk to hundreds of delighted participants: Keri Smith, Sarah Fine, Anna Myers, Eugenie King, and Helen Bedford. Just where *is* that monkey?

Finally, there's Cynan. I can't thank him enough. And he knows it.

Short Abstract

From 'Cat' to 'Dog': Investigating Organisation in the Infant Lexicon

Suzy J Styles

St Edmund Hall, University of Oxford

Thesis submitted for the degree of Doctor of Philosophy, Trinity Term 2008

Do infants learn their early words in isolation? Or do they integrate new words into an inter-connected semantic system? Between the ages of one and two, infants develop a prodigious word-learning ability (P. Bloom, 2002; Fenson *et al.*, 1994; Hamilton *et al.*, 2000). Little is known, however, about how and when infants integrate their accumulating word-knowledge into an adult-like lexical network – one sensitive to relationships such as association and semantic category organisation.

For adults, context has a strong effect on the ease and speed of linguistic processing, and on behavioural responses to language. In lexical priming studies, both visual and auditory context are known to influence the speed of lexical processing (Antos, 1979; Meyer & Schvaneveldt, 1971; Radeau, 1983) and ambiguity resolution (Swinney, 1979), as well as guide predictive looks to contextually relevant areas of a complex scene (Huettig & McQueen, 2007; Kamide *et al.*, 2001). And while priming studies have been conducted with children (Betjemann & Keenan, 2008; Duncan & Kellas, 1983; Hartfield & Conture, 2006; Hashimoto *et al.*, 2007; McCauley *et al.*, 1976; Nation & Snowling, 1999; Radeau, 1983; Schvaneveldt *et al.*, 1977) and toddlers (Krackow & Gordon, 1998), the currently available methods are not applicable below the age of three.

This thesis presents a body of research which combines two experimental paradigms into a single new method which allows investigation of relationships within the developing infant lexicon. Over a series of experiments, this thesis develops a version of the infant-friendly Inter-modal Preferential Looking (IPL) task within the lexical priming paradigm. The method combines fine-grained temporal sensitivity with an intuitive, infant-friendly task. Priming effects were observed during lexical access for children two years-of-age and older, providing a promising new direction for investigations into the development of semantic organisation during development.

Long Abstract

From ‘Cat’ to ‘Dog’: Investigating Organisation in the Infant Lexicon

Suzy J Styles

St Edmund Hall, University of Oxford

Thesis submitted for the degree of Doctor of Philosophy, Trinity Term 2008

Do infants learn their early words in isolation? Or do they integrate new words into an inter-connected semantic system? Between the ages of one and two, infants develop a prodigious word-learning ability (P. Bloom, 2002; Fenson et al., 1994; Hamilton et al., 2000). Little is known, however, about how and when infants integrate their accumulating word-knowledge into an adult-like lexical network – one sensitive to relationships such as association and semantic category organisation.

This thesis presents a body of research which aims to combine two experimental paradigms into a single new method which will allow investigation of relationships within the developing infant lexicon. The aim is to develop a version of the infant-friendly Inter-modal Preferential Looking (IPL) task within the lexical priming paradigm. At the time this research was initiated, it was the only reported attempt to use behavioural methods to investigate lexical relationships between known words for children under the age of three.

For adults, context has a strong effect on the ease and speed of linguistic processing, and on behavioural responses to language. In experimental priming tasks, both visual and auditory context are known to influence the speed of lexical processing (Antos, 1979; Meyer & Schvaneveldt, 1971; Radeau, 1983) and ambiguity resolution (Swinney, 1979), as well as guide predictive looks to contextually relevant

areas of a complex scene (Huettig & McQueen, 2007; Kamide et al., 2001). Adult-like semantic organisation in the early stages of lexical development has been predicted by statistical models of lexicon growth (Hills *et al.*, 2008b; Steyvers & Tenenbaum, 2005). And while behavioural research has been conducted into the semantic systems of children (Betjemann & Keenan, 2008; Duncan & Kellas, 1983; Hartfield & Conture, 2006; Hashimoto et al., 2007; McCauley et al., 1976; Nation & Snowling, 1999; Radeau, 1983; Schvaneveldt et al., 1977) and toddlers (Krackow & Gordon, 1998), the currently available methods are not applicable below the age of three due to their high task demands. Experimental data were therefore required to establish *when relationships develop* in the infant lexicon, and *which kinds of relationship* provide organisation in the developing semantic system.

In order to develop a new experimental method, detailed examination was conducted into the composition of the lexicon in the second year of life. This was followed by a series of experiments designed to establish stimulus sequences and experimental parameters for a novel, behavioural task.

The first chapter of this thesis introduces the theoretical background to why lexical connectivity might be predicted in infancy, reviewing statistical models of network growth and semantic organisation in development. Behavioural priming studies conducted with children and toddlers are also discussed, along with electrophysiological evidence.

Chapter 2 is a norming study, detailing the size and composition of the British infant lexicon at 18 months-of-age (the youngest age of interest, and therefore the baseline in all stimulus selection). The chapter reports data from 179 infants, and is the first large-scale study to provide a detailed profile of lexicon composition within a narrow age range (17.5 months to 18.5 months). Norms for the number of words

comprehended and produced are reported, as well as comprehension rates for all of the words in the British CDI (Hamilton et al., 2000). The typical composition of the lexicon is described, and comprehension rates are compared to other lexical variables known to influence lexical access in online language production.

Chapter 3 describes the first attempt to integrate lexical priming into the inter-modal preferential looking (IPL) method for seventy infants at 18 and 24 months-of-age. The minimally-primed adaptation of IPL is conservative, with an auditory prime preceding a standard, full length (5,000ms) IPL trial. The prime-to-target inter-stimulus interval (ISI) was long (2,500ms), and visual stimuli appeared at the offset of the prime. The experiment primarily assessed whether infants would show typical lexical comprehension in a task which begins with a potentially confusing priming phrase. Three priming conditions are compared: unrelated, taxonomic, and compound (taxonomic + associative). Infants in both age groups showed robust effects of target recognition, which were not influenced systematically by the prime.

The experiment also assessed whether parental CDI reports of comprehension accurately predicted the items which would attract successful performance in the referent identification task. This sub-study identified an item-level match between parental report and target identification, demonstrating that parents do indeed know which words their children understand – contrary to a long-standing literature which has supposed otherwise (Stiles, 1994; Tomasello & Mervis, 1994).

Chapter 4 describes a pilot study for a truncated IPL task which would allow auditory primes to be presented directly before auditory targets. The words-before-pictures experiment presents an un-primed target word shortly before the onset of a pair of pictures, for eighty infants at 18 and 24 months-of-age. The experiment establishes experimental parameters for an appropriate target-to-test stimulus onset

asynchrony (SOA) and optimal picture presentation time. This chapter also establishes new measures for analysis of looking behaviour in the shortened task.

Chapter 5 describes a primed IPL experiment which combines an auditory prime with the truncated trial piloted in Chapter 4. In the quick-primed experiment, for one hundred and fifty eight infants at 18 and 24 months-of-age, two priming conditions are compared: related (taxonomic + associative) and unrelated. Two target-to-test stimulus onset asynchronies are also employed (400ms, 200ms). Robust priming effects are demonstrated for the older age-group, and mechanisms possibly underlying the primed processing effect are introduced – one mechanism relies on spreading activation in the lexicon, while the other relies on visual similarity.

Chapter 6 presents an experiment which attempts to tease apart the two possible processing accounts introduced in Chapter 5, by varying whether or not the target word is presented following the prime. One hundred and twenty four infants at 18, 24 and 30 months-of-age participated in this experiment. The results were difficult to interpret, but suggested that a lexical account of processing was appropriate for older infants.

The quick-primed IPL method established over the course of this thesis successfuly combines the adult lexical priming paradigm, with the infant inter-modal preferential looking (IPL) task. The method combines fine-grained temporal sensitivity with an intuitive, infant-friendly task. It is a flexible method, able to assess the speed and accuracy of lexical access during spoken language comprehension. Priming effects were observed during lexical access for children two years-of-age and older (although not for the younger age group), providing a promising new direction for investigations into the development of semantic organisation during development.

Table of Contents

Acknowledgements iii
Short Abstract v
Long Abstract vii
List of Figures xiii
List of Tables xvii
List of Abbreviations xix

Chapter 1 Introduction: How is the infant lexicon organised? 1

Lexical priming and lexicon organisation 2
Models of the early lexicon under construction 4
Priming before adulthood 8
Developing a lexical priming task for infants 15
A note on terminology 18
Statistical approach 20
Publications arising from this thesis 21

Chapter 2 Norming the 18-month-old lexicon 23

Introduction 24
Data sample 30
Norms for infants 31
Norms for words 42
Comprehension, age-of-acquisition, frequency and association 55
Word associations between O-CDI words 64
Establishing an experimental stimulus set 70
General discussion 72
Interim conclusions 76

Chapter 3 Minimally Primed IPL 79

Introduction 80
Method 86
3.1. Lexical comprehension and parental report 95
3.2 Priming conditions 119
General discussion 151
Interim conclusions 157

Chapter 4 Words-before-pictures: Picture duration 159

Introduction 160
Method 164
Results 171
Discussion 202
Interim conclusions 211

Chapter 5 Quick-Primed IPL 217

Introduction 218
Method 220
Results 227
General discussion 282
Interim conclusions 288

Chapter 6 Half Target – Quick Prime 289
Introduction 290
Method 295
Results 304
General discussion 334
Interim conclusions 341
Chapter 7 Summary & conclusions 343
Principle experimental findings 347
Future directions 355
References 361

Appendix A: Oxford CDI 1
Appendix B: CDI comprehension rate (Alphabetic) 11
Appendix C: CDI comprehension rate (by Category) 17
Appendix D: Word Associations 23
Appendix E: Laboratory Testing Booths 29
Appendix F: Screen Distortion 33
Appendix G: Stimulus Lists 35
Minimally Primed IPL (Chapter 3) 36
Words-before-pictures (Chapter 4) 37
Quick Primed IPL (Chapter 5) 38
Half Target Quick Primed IPL (Chapter 6) 39
Appendix H: Visual Stimuli 43
Minimally-Primed IPL (Chapter 3) 44
Words-before-pictures (Chapter 4) 45
Quick-Primed IPL (Chapter 5) 46
Half Target Quick Primed IPL (Chapter 6) 47

List of Figures

Figure 2.01. O-CDI individual comprehension plotted against production 35
Figure 2.02. M-CDI percentile comprehension plotted against production 36
Figure 2.03. Comprehension & production plotted against age-in-months 37
Figure 2.04. O-CDI composition of the 18m lexicon: CDI word groups 40
Figure 2.05. O-CDI composition of the 18m lexicon: Grammatical word class 42
Figure 2.06. O-CDI median comprehension rate: Grammatical word classes 45
Figure 2.07. O-CDI comprehension advantage: Grammatical word classes 46
Figure 2.08. O-CDI comprehension rate & production rate (Onomatopoeia highlighted) 47
Figure 2.09. O-CDI comprehension rate & production rate (other special cases) 48
Figure 2.10. Comprehension rates: M-CDI plotted against O-CDI 51
Figure 2.11. Comprehension advantage for word groups: Two populations 54
Figure 2.12. O-CDI 18m comprehension rate & Morrison age-of-acquisition 57
Figure 2.13. O-CDI comprehension rate & adult British noun frequency in 61
Figure 2.14. O-CDI comprehension rate & log-transformed adult frequency 62

Figure 3.01. Schematic of trial timing 91
Figure 3.02. A yoked picture-pair in three priming conditions 92
Figure 3.03. PTL and LLK according to lexical status (Broad) 101
Figure 3.04. PTL and LLK according to target status (Narrow) 108
Figure 3.05. CDI score and change-in-PTL according to target status (Narrow) 113
Figure 3.06. CDI score and change-in-LLK according to target status (Narrow) 114
Figure 3.07. Mean PTL in three priming conditions (18m) 123
Figure 3.08. Change in target preference according to sex and priming condition (18m) 125
Figure 3.09. Pre- and post-naming LLK (18m) 126
Figure 3.10. Histogram of latency over the course of the trial (18m) 130
Figure 3.11. Latency to switch in three priming conditions (18m) 131
Figure 3.12. Detailed trial time-course (18m) 135
Figure 3.13. Pre- and post-naming PTL (24m) 138
Figure 3.14. Change-in-PTL and change-in-LLK, according to sex (24m) 141
Figure 3.15. Pre-naming and post-naming LLK (24m) 142
Figure 3.16. Histogram of latency over the course of the trial (24m) 146
Figure 3.17. Latency to switch in three priming conditions (24m) 146
Figure 3.18. Detailed trial time-course (24m) 149
Figure 3.19. IPL screen configuration in two laboratories 156

Figure 4.01. Schematic of trial timing 167
Figure 4.02. PTL and LLK, according to age and timing condition 173

Figure 4.03. Macro- measures plotted against comprehension vocabulary 177
Figure 4.04. PTL plotted against LLK, according to age and timing condition 178
Figure 4.05. Percent of first fixations correct, according to age and timing condition 180
Figure 4.06. Histogram of time-to-fixate over the course of a trial 183
Figure 4.07. Time-to-fixate, according to age and timing condition 184
Figure 4.08. Histogram of time-to-disengage over the course of the trial 185
Figure 4.09. Time-to-disengage, according to age and direction of first fixation 186
Figure 4.10. Time-to-fixate plotted against time-to-disengage for each trial 189
Figure 4.11. Two RT measures plotted against CDI comprehension score 190
Figure 4.12. Relative within-pair picture interest in two macro-level measures. 191
Figure 4.13. Relative within-pair picture interest in two micro-level measures. 193
Figure 4.14. Means for each picture in two macro-measures, pooled across subject conditions 196
Figure 4.16. Best fitting models of variance for macro-level measures 201
Figure 4.17. Best fitting models of variance for micro-level measures 201
Figure 4.18. Two screen configurations and their impact on peripheral pictures 209

Figure 5.01. A yoked picture-pair in two priming conditions 222
Figure 5.02. Schematic of trial timing 226
Figure 5.03. PTL according to SOA group and prime condition (18m) 231
Figure 5.04. PTL according to sub-list and prime condition (18m) 233
Figure 5.05. LLK according to SOA group and prime condition (18m) 234
Figure 5.06. LLK according to sub-list and prime condition (18m) 236
Figure 5.07. Percent correct first fixations (18m) 239
Figure 5.08. Histogram of time-to-fixate over the course of the trial (18m) 240
Figure 5.09. Time-to-fixate according to look direction and SOA group (18m) 241
Figure 5.10. Time-to-fixate according to sub-list (18m) 243
Figure 5.11. Histogram of time-to-disengage over the course of a trial (18m) 245
Figure 5.12. Time-to-disengage according to SOA and prime condition (18m) 246
Figure 5.13. Time-to-disengage according to stimulus sub-list and priming condition (18m) 248
Figure 5.14. PTL according to SOA group and prime condition (24m) 251
Figure 5.15. PTL according to stimulus sub-list and prime condition (24m) 253
Figure 5.16. LLK according to SOA group and prime condition (24m) 254
Figure 5.17. LLK according to sub-list and prime condition (24m) 256
Figure 5.18. Percent correct first fixations according to prime condition (24m). 258
Figure 5.19. Histogram of time-to-fixate over the course of the trial (24m) 259
Figure 5.20. Time-to-fixate according to SOA group and prime condition (24m) 259
Figure 5.21. Time-to-fixate according to stimulus sub-list and prime condition (18m) 261
Figure 5.22. Histogram of time-to-disengage over the course of the trial (24m) 262
Figure 5.23 Time-to-disengage according to SOA group and prime condition (24m) 263
Figure 5.24. Time-to-disengage according to sub-list and prime condition (24m) 266
Figure 5.25. PTL and Δ-Prime for PTL according to age and lexicon size 269

Figure 5.26. Δ-Prime for time-to-disengage, split by fixation direction 270
Figure 5.27. PTL and Δ-Prime for PTL in relation to lexicon size and sex 272
Figure 5.28. Δ-Prime for time-to-disengage, compared tolexicon size and sex 272
Figure 5.29. Detailed trial time course (18m) 275
Figure 5.30. Detailed trial time course (24m) 276
Figure 5.31. PTL for target pictures in three interest levels 281

Figure 6.01. Two possible priming mechanisms 292
Figure 6.02. A yoked picture pair in four stimulus conditions 298
Figure 6.03. Schematic of trial timing 302
Figure 6.04. Onscreen attention device 303
Figure 6.05. PTL for younger and older infants according to condition 306
Figure 6.06. Percentage of first fixations to target 310
Figure 6.07. Histogram of time-to-fixate over the course of the trial 314
Figure 6.08. Time-to-fixate according to age group (pooled across look direction) 315
Figure 6.09 Time-to-fixate according to stimulus sub-list (18m) 317
Figure 6.10 Time-to-fixate according to trial condition (30m) 319
Figure 6.11. Histogram of time-to-disengage over the course of the trial 321
Figure 6.12. Time-to-disengage in three age groups 321
Figure 6.13. Detailed trial time course in four conditions (18m) 329
Figure 6.14. Detailed trial time course in four conditions (24m) 330
Figure 6.15. Detailed trial time course in four conditions (30m) 331
Figure 6.16. Target-First trials: Proportion of target fixations 332
Figure 6.17. Distracter-First Trials: proportion of target fixations 333

List of Tables

Table 2.1. Data collection methods in selected key studies 26
Table 2.2. Percentiles of CDI comprehension and production from three norming studies 33
Table 2.3. Number of words comprehended in each CDI word group 39
Table 2.4. Median comprehension rate for words in each CDI word group 44
Table 2.5. Comprehension of words in different sections of the CDI, according to two studies 52
Table 2.7. Correlations of comprehension rate and log frequency, with different dummy variables 63
Table 2.8. Number of word associations to and from O-CDI words in different word classes 69
Table 2.9. Stimulus set of concrete, generic nouns with lexical comprehension rates of above 50% 71

Table 3.1. Duration of prime phrases 89
Table 3.2. Duration of target phrases 90
Table 3.3. Number of trials according to lexical status (Broad) 99
Table 3.4. Four-way repeated measures ANOVA of LLK (Broad) 103
Table 3.5. Number of trials according to lexical status (Narrow) 105
Table 3.6. Four-way repeated measures ANOVA of PTL (Narrow) 107
Table 3.7. Four-way repeated measures ANOVA of LLK (Narrow) 110
Table 3.8. Three-way repeated measures ANOVA for PTL (18m) 124
Table 3.9. Three-way ANOVA for LLK (18m) 127
Table 3.10. Number of shifts following target onset (18m) 131
Table 3.11. Three-way ANOVA of latency to switch fixation (18m) 132
Table 3.12. Three-way ANOVA for PTL (24m) 139
Table 3.13. Three-way ANOVA for LLK (24m) 143
Table 3.14. Number of shifts following target onset (24m) 145
Table 3.15. Three-way ANOVA for Latency (24m) 147

Table 4.1. Number of participants in each subject group 165
Table 4.2. Duration of auditory stimuli 167
Table 4.3. Summary of target discrimination in four measures 203
Table 4.4. Summary of best-fitting categorical regression models of PTL 213
Table 4.5. Summary of best fitting categorical regressions models of LLK 214
Table 4.6. Summary of best fitting categorical regression models of time-to-fixate 215
Table 4.7. Summary of best fitting categorical regression models of time-to-disengage 216

Table 5.1. Duration of prime phrases 224
Table 5.2. Duration of target words 225
Table 5.3. Number of trials available in each subject group 229
Table 5.4. Three-way ANOVA of PTL (18m) 232
Table 5.5. Three-way ANOVA of LLK (18m) 235

Table 5.6. Number of first fixations to target and distracter according to prime type 237
Table 5.7. Four-way ANOVA of time-to-fixate (18m) 242
Table 5.8. Mean time-to-fixate according to SOA group and sub-list (18m) 244
Table 5.9. Four-way ANOVA of time-to-disengage (18m) 247
Table 5.10. Mean time-to-disengage according to SOA group and sub-list (18m) 249
Table 5.11. Three-way ANOVA of PTL (24m) 252
Table 5.12. Three-way ANOVA of LLK (24m) 255
Table 5.13. Number of first fixations to target and distracter according to prime type (24m) 257
Table 5.14. Four-way ANOVA of time-to-fixate distracters (24m) 260
Table 5.15. Three-way ANOVA of time-to-disengage from targets (24m) 264
Table 5.16. Three-way ANOVA of time-to-disengage from distracters (24m) 265
Table 5.17. Mean time-to-disengage according to SOA group and sub-list (24m) 267

Table 6.1. Duration of prime phrases 299
Table 6.2. Duration of target words 301
Table 6.3. Duration of dummy targets 301
Table 6.4. Three-way ANOVA of PTL (18m) 307
Table 6.5. Three-way ANOVA of PTL (24-30m) 308
Table 6.6. Three-way ANOVA of percentage of first fixations to target (18m) 311
Table 6.7. Four-way ANOVA of percentage of first fixations to target (24-30m) 312
Table 6.8. Three-way ANOVA of time-to-fixate (18m) 316
Table 6.9. Three-way ANOVA of time-to-fixate (30m) 319
Table 6.10. Three-way ANOVA of time-to-fixate (18m) 322
Table 6.11. Three-way ANOVA of time-to-fixate (24-30m) 323

List of Abbreviations

Methodological

IPL	Inter-modal preferential looking
VEx	Visual expectation
ISI	Inter-stimulus interval
SOA	Stimulus onset asynchrony
RT	Reaction time
PTL	Proportion of target looking: T/(T+D)
LLK	Difference in length of longest look: t-d

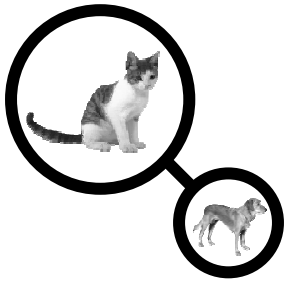
Statistical

N	sample population size
M	mean
SD	standard deviation
p	probability value
α	alpha level
n.s.	not significant
t	Student's t
F	Fisher's F
U	Mann-Whitney's U
z	Z-score
χ^2	*Chi*-squared
r	Pearson's r
ρ	Spearman's *rho*
d	Cohen's d
η^2	*Eta*-squared

For Jane and Howard,
who gave me the ball of string…

Chapter 1

Introduction: How is the infant lexicon organised?



Do infants learn their early words in isolation? Or do they integrate new words into an inter-connected semantic system? Between the ages of one and two, infants develop a prodigious word-learning ability (P. Bloom, 2002; Fenson et al., 1994; Hamilton et al., 2000). Little is known, however, about how and when infants integrate their accumulating word-knowledge into an adult-like lexical network – one sensitive to relationships such as association and semantic category organisation.

This thesis presents a body of research which aims to combine two experimental paradigms into a single new method which will allow investigation of relationships within the developing infant lexicon. The aim is to develop a version of the infant-friendly Inter-modal Preferential Looking (IPL) task within the lexical priming paradigm. At the time this research was initiated, it was the only reported attempt to use behavioural methods to investigate lexical relationships between known words for children under the age of three.

Lexical priming and lexicon organisation

For adults, context has a strong effect on the ease and speed of linguistic processing, and on behavioural responses to language. Both visual and auditory context are known to influence the speed of lexical processing (Antos, 1979; Meyer & Schvaneveldt, 1971; Radeau, 1983) and ambiguity resolution (Swinney, 1979), as well as guide predictive looks to contextually relevant areas of a complex scene (Huettig & McQueen, 2007; Kamide et al., 2001). The effect of prior context on performance in behavioural tasks is termed priming.

The lexical priming task developed by Meyer & Schvaneveldt (1976) presented participants with strings of letters on a screen. The strings were either real

words or pronounceable non-words. For each string, participants performed a lexical decision task, by pushing a button to indicate whether the string was a real word or not. Reaction times were measured, and the order of the stimuli was tightly controlled. When two words appeared on subsequent trials, the relationship between the words was manipulated so that half of the words were semantically related (e.g., *doctor – nurse*), and half were semantically unrelated (e.g., *library – nurse*). Reaction times to the second word (the target) were reliably influenced by whether the word pair was related; Responses were faster when a word was preceded by a related prime than they were when the same word was preceded by an unrelated prime.

The influence of semantic relationships in the decidedly non-semantic task suggested that the semantic memory system was directly connected to feature-level representations of words. In their account of primed processing mechanisms, the authors proposed that facilitation was produced by 'excitation' spreading between closely related items in the semantic system. When the prime was related to the target, the extra excitation lowered the amount of time needed to access the target word.

Spreading activation accounts of language processing (Anderson, 1983; Collins & Loftus, 1975) present the most widely recognised interpretation of lexical priming effects. These models propose that semantic knowledge can be represented as a network of nodes (words) linked by connections (relationships) through which activation flows during linguistic processing. These models suggest that bottom-up activation from visual or auditory input leads to lateral activation between word-level nodes in the network. Facilitation in priming tasks is thus produced by gradual bottom-up activation combining with lateral flow between nodes. In this model of processing, word-word relationships are structurally encoded in the network architecture, that is, the lexicon structure.

Priming experiments have demonstrated the cognitive reality of various types of relationship in the adult lexicon, including word association (Meyer & Schvaneveldt, 1971; Moss *et al.*, 1995; Shelton & Martin, 1992), taxonomy (meta-analysis by Lucas, 2000; McRae & Boisvert, 1998), shared semantic features (McRae *et al.*, 1997; Moss *et al.*, 1997), thematic relationships (Moss et al., 1995), phonological similarity (Radeau *et al.*, 1989) and even shared affective valence (Rossell & Nobre, 2004; Spruyt *et al.*, 2004). This leads to a complex model of the adult lexicon, in which each word exhibits a variety of connections to other words, some of which may exist at different levels of processing (e.g., phonological, semantic, affective). Despite the complex organisation of the adult lexicon, little is known about its development. Does network organisation begin in the early stages of lexical acquisition? And if so, is the structure of the early lexicon *adult-like* in its nature?

Models of the early lexicon under construction

One model of lexical network development comes from the field of computational linguistics. Steyvers and Tenenbaum (2005) tested the mathematical fit of three possible network structures to four natural language data sets. The three architectures were: 1) tree-structured hierarchical organisation; 2) unstructured (random) organisation; and 3) a 'small-world' structure of tightly clustered node neighbourhoods with a 'scale-free' pattern of attachment. While the first two architectures make no explicit claims about the development of the network, the third entails a pattern of preferential attachment for new nodes which results in a power law distribution of connectivity. This model is characterised by the majority of words

sharing few connections with other words, resulting in small-scale local clusters of nodes connected by a few highly connected 'hubs'.

The three possible architectures were tested against the statistical properties of networks created from four natural-language data sets: Two networks were created from the norms of the University of South Florida Free Association Norms (D. L. Nelson *et al.*, 1998); one network was created from the synonymy listings of Roget's Thesaurus (Roget, 1991); and the final network was WordNet (Miller, 1995). They found that all four data sets shared the statistical properties of the small-world, scale-free model, rather than a random or hierarchical organisation.

In addition, Steyvers and Tenenbaum were able to model node-by-node network growth according to the power-law attachment pattern, as each new word formed connections to others in the network. Having demonstrated the preferential attachment mechanism, they tested one of its main predictions: that words acquired earlier in development would have a 'head start' in the number of incoming and outgoing connections, compared to later acquired words. Words in the three adult networks did show a strong relationship between connectivity and age-of-acquisition (AoA). That is, words learned earliest were those which were deepest integrated into the semantic structure.

Given the deep integration of early acquired words, Steyvers & Tenenbaum's model offers good reason to predict that structured lexical organisation would be present in the early stages of lexicon development (i.e., infancy). The scale-free properties of their model also suggest that there is no discontinuity in the *nature* of network organisation over the course of development. This implies that where infant lexical organisation is observable, it would be predicted to be adult-like in its nature.

However, the authors acknowledge that their simplified statistical account of network development did not employ real words or ages of acquisition.

A recent investigation of infant lexical network growth tested the preferential attachment hypothesis against longitudinal reports of the composition of 20 infant lexicons, over a 15-month period (16-months to 30-months). The study by Hills, Maouene, Maouene, Sheya and Smith (2008b) tracked network growth across basic-level nouns from the MacArthur-Bates Communicative Development Inventory for Toddlers (Fenson et al., 1994), filled in at monthly intervals. The 130 nouns selected for comparison were those which appeared in both the McRae Semantic Feature Norms (McRae *et al.*, 2005), and in the University of South Florida Free Association Norms (D. L. Nelson et al., 1998). Two networks were generated according to the month-by-month CDI acquisition data. In the first network, newly acquired words were connected to others in the network if they shared functional features (e.g., used to eat food) or perceptual features (e.g., is red). In the second network, new words were connected if they had an attested word association. The shared feature network did not show the preferential attachment pattern predicted by Steyvers and Tenenbaum (2005), however the free association network did.

There are two ways of interpreting this finding. The possibility highlighted by the authors is that the adult feature norms were a poor match for toddler word-word relationships. This may be because adults' perspectives and typical experiences of objects differ from those of toddlers (e.g., for an adult, a teddy bear might share the perceptual feature 'is fluffy' with a French poodle, or a cloud; while for a toddler, teddy bears and biscuits may share the perceptual feature 'goes soggy in my mouth'). It is also possible that the feature norms are more variable than the word association norms due to the method of collection – each of the participants was instructed to

provide fourteen features for each word. By comparison, participants in the University of South Florida Free Association Norms were instructed to give *only the first* response which came to mind. Thus, differences in the method of norming may be responsible for the poor fit of the semantic feature norms. The authors imply that the association norms provide a better description of the infant data, as the association-based model matched Styvers and Tenenbaum's developmental prediction.

A second possibility, not explored in detail by the authors, is that changes in the organisation of the lexicon may occur during early development. The architecture of the infant lexicon may be organised radically differently during infancy, and yet shift gradually to an association-based organisation (including preferential attachment) with age. If this were the case, the adult network could still exhibit a relationship between age-of-acquisition and network connectivity, albeit on a larger time-scale.

A third possibility is that both network architectures (semantic feature based, and associative relationship based) exist in parallel. As suggested in a second model by Hills, Maouene, Maouene, Sheya & Smith (2008a), an architecture based on shared semantic features might help language learners develop an understanding of the way concepts relate in the world, providing evidence for categorical inferences. Connections based on association may emerge from the same system, as a product of experience, producing an associative super-structure in the network. The two parallel architectures may have qualitatively different properties, affecting the nature of spreading activation during language processing. Parallel architectures may account for why different kinds of semantic relationship have been demonstrated to influence primed language processing at different time scales (Moss et al., 1997; Shelton & Martin, 1992), and across different modalities (Moss et al., 1995).

These models of lexical network development predict that early lexicons will be semantically structured, and that the organisation of the network may well be adult-like. However, given the possibility of structural re-organisation during development, it is impossible to know whether a) scale-free, small-world architecture begins in infancy and continues, unchanged, into adulthood, b) large-scale re-organisation takes place in the early stages of development, and the scale-free, small-world characteristics of the adult lexicon develop gradually, or c) whether the scale-free, small-world properties of adult lexical networks develop in parallel with some other organisational structure. What is needed is evidence of which kinds of relationship have cognitive reality in the early developing lexicon. Lexical priming offers a possible platform for establishing behavioural evidence about the types of relationships in the infant lexicon, along with when they develop.

Priming before adulthood

To date, only a small number of experimental studies have investigated whether relationships between words are observable in the language processing of young children or toddlers. Barriers to the use of standard adult methodologies in early development have included participants' variable reading ability, limited attention spans, and relatively small vocabularies. Further complications for toddlers include the lack of meta-linguistic knowledge required for a lexical decision task (e.g., whether a string of sounds is a 'word'), and for infants, the inability to understand complex instructions. Priming studies including children, and child-friendly adaptations of the standard adult priming task are summarised below.

Priming studies incorporating reading

One of the earliest studies with school-aged children (Schvaneveldt et al., 1977) replicated the basic adult priming effect, in a visual lexical decision task. Seven-year-olds and nine-year-olds were shown slides containing words and non-words. Children in both age groups were instructed to provide lexical decisions about all words presented (both primes and targets), and showed faster reaction times when the target word was preceded by an associatively related prime. In addition, weaker readers and smaller vocabulary children showed larger effects of priming condition (although cf Chapman, Chapman, Curran and Miller's (1994) argument that larger priming effects can be an artefact of slower RTs).

Priming adaptations without reading

Radeau (1983) developed an auditory version of the lexical decision task, in which primes and target were presented acoustically to 6- to 7-year-olds and adults. They made lexical decisions about the pairs of words. Both children and adults responded faster to associated word pairs than to unrelated word pairs.

Nation and Snowling (1999) used the auditory lexical decision task with school-aged children classed as 'poor comprehenders', who were asked to make a lexical decision about each word they heard. Word pairs were taxonomically or thematically related, and half of the word pairs shared a strong word association. Participants from ranged from 6 to 11 years-of-age, and a control group was matched for reading comprehension ability. Both taxonomic and thematic relationships generated faster reaction times, with enhanced priming when the words were also associated. For poor comprehenders, however, taxonomic relationships only generated faster reaction times in the presence of an associative relationship.

Betjemann and Keenan (2008) used a visual lexical decision task with school-aged children with reading disabilities, along with age matched and reading-age matched controls. They presented five kinds of prime-target relationship: unrelated, associated, taxonomic (synonym), phonological/graphemic, combined phonology-and-semantic (e.g., *float – boat*), and identity (e.g., *boat – boat*). Control children demonstrated faster reaction times for all related conditions, and were fastest in the identity condition. Children with reading disabilities only showed faster reaction times in the phonologically related conditions. They repeated the experiment using an auditory lexical decision task, and found the same pattern of results – indicating that the children with reading disabilities showed the deficit in semantic facilitation even outside the reading modality.

Adaptations without lexical decision: Picture naming

McCauley, Weil and Sperber (1976) introduced a primed picture naming task, in which school-aged children were asked to name pictures, presented sequentially. Pairs of pictures were controlled for associative and for taxonomic relationships. Six-year-olds demonstrated faster naming latencies for the target picture when it was preceded by an associated prime. 8 year olds demonstrated faster naming latencies for both associative and taxonomic relationships.

Hartfield and Conture (2006) used auditory primes in a picture naming task for three- to five-year-olds who stuttered. Compared to a control condition where a tone was presented prior to picture naming, picture naming latencies were slower when primed by a physically related word (e.g., *ball – sun*), than when primed by a functionally related word (e.g., *shine – sun*). Latencies for semantically related primes (e.g., *star – sun*) fell between these values. Age matched children with fluent speech

showed no difference between conditions, a finding which may have been due to the relatively long inter-stimulus interval (700ms) between the offset of the auditory prime and the onset of the target picture.

Adaptations without lexical decision: Other tasks

Duncan and Kellas (1983) used a primed *category membership task* developed by Rosch (1975), with school aged children and adults. In this task, participants made a decision about whether a pair of simultaneously presented pictures came from the same category. An auditory prime (the category name) preceded the picture pair in half of the trials. Picture typicality and stimulus degradation were also manipulated across conditions. All participants showed faster responses to intact stimuli than to degraded stimuli. When the picture pairs were different (two different pictures from the same category, or two different pictures from different categories), 8-year-olds, 10-year-olds, 12-year-olds and adults showed faster latencies in primed trials, faster latencies for typical stimuli and faster latencies for intact stimuli. For physically identical picture pairs, all groups except the youngest showed faster response latencies to typical stimuli in primed trials when compared to un-primed trials, and slower response latencies to atypical stimuli.

Hashimoto, McGregor and Graham (2007) used a primed *object decision task* in which school-aged children and adults made judgments about whether pictures depicted 'real' objects, or 'made up' objects. Pairs of pictures were presented in four conditions, a taxonomic condition, a thematic condition (where pictures shared temporal, spatial or causal relations), a perceptual similarity condition, and an unrelated control. 6-year-olds, 8-year-olds and adults showed faster object decision latencies in the taxonomic and thematic conditions, compared to the perceptual

condition and the control. The only difference observed between taxonomic relationships and thematic relationships was that children were better able to describe the thematic relationships in a follow-up task.

Krackow and Gordon (1998) introduced a primed *paired associate memory task*, in which 3- to 4-year-olds listened to a list of eight word pairs. At the end of the presentation, children were given each cue, and asked to recall the paired response they had heard previously. In this task, associative relationships primed accurate recall, while taxonomic relationships did not.

Together, these studies demonstrate the existence of adult-like lexical relationships in the of lexicons children of school age, and below. The majority of studies assessed associative relationships, evidence of which was observed in children as young as 3 years-of-age (Krackow & Gordon, 1998). In normally developing children, associative priming effects appear to be more robust than taxonomic relationships, as they produced larger priming effects (Nation & Snowling, 1999), and in some tasks, were observed at earlier ages (McCauley et al., 1976). Where different kinds of non-associative relationships were employed, little difference in the degree of priming was generally observed (Hashimoto et al., 2007), except in the performance of children with poor reading comprehension, for whom thematic relationships were more robust than taxonomic, in the absence of association (Nation & Snowling, 1999).

While providing encouraging evidence that lexical relationships may be observable at even younger ages, the applicability of these methods to younger participants remains limited by task demands: All of the above methods required a meta-linguistic decision or a verbal response of some kind (e.g., picture naming). To extend the lexical priming task to even younger participants would require an intuitive

experimental task without complex instructions, which could also accommodate the high temporal accuracy typically required to observe priming effects.

Passive lexical processing in infancy

One approach which satisfies the intuitive task condition, is observing electrophysiological brain activity during passive language listening. A small number of researchers have sought to compare infant patterns of brain activity with established adult activation patterns which index semantic processing. In adult event-related potentials (ERPs), the amplitude of the 'N400 component' is understood to index semantic congruency (Kutas & Hillyard, 1980), semantic relatedness (Federmeier & Kutas, 1999) and category organisation (Heinze *et al.*, 1998). Recent developmental studies have sought an infant analogue of this semantic component.

Friedrich & Friederici (2004, 2005a) presented infants with nouns while they looked at a single picture on screen. Infants' event-related potentials following the onset of the spoken word included a component similar to the adult N400, which was large in the presence of a mismatch between the picture and the label, small in the presence of a match, and midway between these values for a mismatch drawn from the same semantic category. Thus infants looking at a picture of a cat were less sensitive to a mismatch when they heard a word like *dog* than they were when they heard a word like *chair*. Friedrich and Friederici interpret their findings as a demonstration of adult-like graded semantic sensitivity. One limitation of this word-picture mismatch task is that it investigates electrophysiological activity during static picture processing, and therefore does not inform our understanding of sequential language processing, i.e., whether the ease or speed of lexical processing would be affected (primed) by spreading activation from a preceding word.

An additional complication of the ERP method, is that it is difficult to take into account whether individual infants understood, or were familiar with, words used in the test. In these studies, there was no control for comprehension of individual items. Consequently, the appearance of graded sensitivity to a 'within category mismatch' (e.g., hearing *dog* while looking at a picture of a cat) may actually be the outcome of pooling data of different types. That is to say, infants who understood both *dog* and *cat* may have registered the word as a 'mismatch' to the picture. Yet, infants who only understood *cat* but did not know *dog*, may have accepted the pairing as a 'match' through overextension. A mean drawn from these two populations would produce a value midway between 'match' and 'mismatch', thus giving the appearance of graded sensitivity.

Friedrich and Friederici (2005b) also extended the infant ERP paradigm to an auditory-only sentence-processing task in which 19- and 24-month-olds heard simple subject-verb-object sentences. Semantically restrictive verbs were used to generate sentences in which the final noun was either semantically congruent, or incongruent with the preceding verb. The larger N400 component for incongruent sentences was consistent with similar findings in the adult literature. This auditory-only paradigm provided evidence of adult-like responses to semantic violation during *sequential* language processing.

Following the auditory-only procedure, Torkildsen et al. (2007) conducted a sequential priming ERP study with 24-month-old infants. The authors presented auditory primes and targets with a gap of 1200ms. The authors found the ERP signal following the onset of the target word differed according to whether the prime and target were members of the same category. The finding suggested that infants' on-line lexical processing is indeed influenced by preceding linguistic context. Torkildsen et

al. acknowledge, however, that it is difficult to tell whether the variations in ERP signature were due to spreading activation between connections at a lexical or conceptual level of representation. They also acknowledge that it is unclear what functional effect these ERP signatures might have on the ease or speed of linguistic processing. Thus although children two years and under show patterns of electrical activity which are similar to adult language processing effects, without corroborating behavioural evidence it is impossible to know whether the same brain signature has the same cognitive effect.

Developing a lexical priming task for infants

It thus remains to be seen whether adult-like primed processing effects can be replicated in an on-line behavioural task for infants, in which item-level sensitivity is achievable. If infants' online language processing is affected by verbal context in sequential word presentation (as indexed, for example, by lexical comprehension), it would provide strong support for a model of lexicon development which features adult-like inter-connectivity from a very early age. The goal, therefore, is to develop an infant-friendly behavioural task in which stimulus presentation can be systematically varied to produce a priming task, and which uses ease and speed of lexical comprehension as an index of online language processing.

The intermodal preferential looking (IPL) task established by Golinkoff et al. (1987) is a free-looking task for infants, in which looking behaviour is monitored while a pair of images is presented, and auditory stimuli are introduced. When one of the images is labelled, infants' looking behaviour shows an increase in preference for the named image (J Stephen Reznick, 1990). The IPL task is a flexible framework

which is sufficiently sensitive to demonstrate infants' sensitivity to auditory manipulations (Fernald *et al.*, 2001; Mani & Plunkett, 2007; Swingley & Aslin, 2007; White & Morgan, 2008), as well as to item typicality (Meints *et al.*, 1999). Contemporary implementations of IPL also employ offline frame-by-frame analysis of video recordings, and thus have high temporal accuracy. By bringing together the adult lexical priming task and the infant-friendly IPL task, the proposed primed IPL task is expected to provide a platform for detailed investigations into the kinds of relationships in the infant lexicon, when they develop, and how they relate to word learning.

For the current body of research, the age range of interest is the latter half of the second year. This age is around the time of the so-called 'word spurt' or 'naming explosion' (L. Bloom, 1973; Ganger & Brent, 2004; Goldfield & Reznick, 1990; K. Nelson, 1973) . Infants at this age are experienced word learners, comprehending something in the order of 100 words (Fenson et al., 1994). From a theoretical perspective, this means their lexicons are sufficiently large that different kinds of semantic relationship are possible between words. From an experimental perspective, it should be possible to construct a sizeable stimulus pool for infants in this age range, as their vocabularies are sufficiently large to provide substantial overlap between the particular words individuals understand.

In order to develop appropriate materials for use in a priming study for 18- to 24-month-olds, Chapter 2 is a norming study, which details the size and composition of the British lexicon at 18 months-of-age (the youngest age of interest, and therefore the baseline for all stimulus selection). The chapter reports data from 179 infants, and is the first study of British infants' lexical comprehension to provide a detailed profile of lexicon composition within a narrow age range (17.5 months to 18.5 months).

Norms for the number of words comprehended and produced are reported, as well as comprehension rates for all of the words in the British CDI (Hamilton et al., 2000). Typical composition of the lexicon is described, and comprehension rates are compared to other lexical variables known to influence lexical access in online language production.

Chapter 3 describes the first attempt to integrate lexical priming into the inter-modal preferential looking (IPL) method for seventy infants at 18 and 24 months-of-age. The minimally-primed adaptation of IPL is conservative, with an auditory prime preceding a standard, full length IPL trial. The prime-to-target inter-stimulus interval (ISI) is long, and visual stimuli appear during this ISI. The experiment primarily assesses whether infants will show typical lexical comprehension in a task which begins with a potentially confusing priming phrase. Three priming conditions are compared: unrelated, taxonomic, and compound (taxonomic + associative). The experiment also assesses whether parental CDI reports of comprehension accurately predict the items which attract successful performance in the referent identification task.

Chapter 4 describes a pilot study for a truncated IPL task which would allow auditory primes to be presented directly before auditory targets. The words-before-pictures experiment presents an un-primed target word shortly before the onset of a pair of pictures, for eighty infants at 18 and 24 months-of-age. The experiment establishes experimental parameters for an appropriate target-to-test stimulus onset asynchrony (SOA) and optimal picture presentation time. This chapter also establishes new measures for analysis of looking behaviour in the shortened task.

Chapter 5 describes a primed IPL experiment which combines an auditory prime with the truncated trial piloted in Chapter 4. In the quick-primed experiment,

for one hundred and fifty eight infants at 18 and 24 months-of-age, two priming conditions are compared: related (taxonomic + associative), and unrelated. Two target-to-test stimulus onset asynchronies were also employed. Robust priming effects are demonstrated, and mechanisms possibly underlying the primed processing effect are introduced – one mechanism relies on spreading activation in the lexicon, while the other relies on visual similarity.

Chapter 6 presents an experiment which attempts to tease apart the two possible processing accounts introduced in Chapter 5, by varying whether or not the target word is presented following the prime. One hundred and twenty four infants at 18, 24 and 30 months-of-age participated in the half-target experiment.

A note on terminology

There is a great deal of ambiguity in the terminology used to talk about meaning in psycholinguistic research. By way of clarification, the term *semantic* is taken to mean 'meaning based,' in a broad sense, with *semantic relationships* referring to any kind of meaningful relationship between words or concepts in language.

By contrast, *taxonomic relationships* refer to those relationships in a classical hierarchy of language based categories (e.g., *dog* is a daughter of *animal*, *dog* and *cat* are sisters). *Thematic* relationships are non-taxonomic relationships based on an event schema, and can include, for example spatial relationships (e.g., *hat – head*) and functional relationships (e.g., *broom – floor*). *Association* refers only to lexical relationships which have been experimentally elicited in word association tasks (e.g., Moss & Older, 1996; Postman & Keppel, 1970). Thus association, theme and

taxonomy are kinds of semantic relationship, and association can co-occur with either taxonomic or thematic relationships.

The term *lexical* is used in this thesis to refer to items and processes which occur in the lexicon, and therefore rely on activation of a word-level node for their effects. Thus *lexical priming* refers to the process which influences the ease and speed of activation of a word-form during comprehension (hearing or reading), or production (naming or reading aloud). This is to be distinguished from other kinds priming which may generate behavioural effects at different levels of processing. For example, negative spatial priming (e.g., Amso & Johnson, 2005) functions extra-linguistically, while perceptual priming using pictures (e.g., Matsukawa *et al.*, 2005) may operate at a conceptual level, outside the lexicon *per se*. This contrasts with the terminology of many prominent researchers (e.g., McNamara, 2005), who use the term 'semantic priming' to refer to all word-word priming techniques where the variable under manipulation is meaning-based. However, given the historic ambiguity in the term 'semantic' (which can be used to mean 'generally meaning based', 'taxonomic and/or associative,' 'taxonomic but not associative,' 'any meaning-based non-associative relationship,' 'non-associated taxonomic sisterhood only,' etc.), the less ambiguous term 'lexical priming' is preferred here. While some researchers reserve this term for effects which occur at the phonological level of processing, in the current document, the term *phonological priming* will provide disambiguation.

Consistent with this use of the term *lexical*, the *lexicon* is understood to be comprised of items (words) which have both form and meaning. In the infant lexicon, a newly acquired word may consist of an imprecise phonological form paired with a weak representation of the word's meaning. Any concept for which no word has been acquired is not considered to be part of the infant's system of representing meaning in

language. Similarly, a familiar word for which the infant has no conceptual representation is not considered to have entered the lexicon.

Lexical organisation, or *lexicon structure* thus refers to the organisation of the elements (words) which constitute the lexicon. Relationships between words are understood to exist in the lexicon at more than one level of processing. Given the semantic focus of this thesis, *lexical organisation* will typically refer to the semantic level.

Statistical approach

A number of different analysis techniques were employed over the course of this thesis. In the experimental chapters, where many dependent measures are employed throughout the thesis, and a systematic approach to analysis was used. For multi-factorial analyses, a repeated-measures Analysis of Variance (ANOVA) was conducted when the experimental design employed more than one trial phase (Chapter 3), and where the data satisfied the conditions for normality and sphericity. If the data were not appropriately distributed, a difference measure was calculated, and subjected to a univariate ANOVA, or in one case (Chapter 3) linear regression. Univariate ANOVAs were also employed in experimental designs with only a single trial phase. If variance was not sufficiently homogeneous for analysis by univariate ANOVA, the data were subjected to simple comparisons (Chapters 4, 5, 6). Simple comparisons were conducted by independent samples t-test where the data were normally distributed, and otherwise, according to non-parametric tests (Mann-Whitney's U test for small samples, and its equivalent, Wilcoxon's rank sums test for

large samples). Where comparison was made to 'chance,' a one-sample t-test was conducted.

Publications arising from this thesis

Some of the material in this thesis has been published elsewhere. Material from the first half of Chapter 2 (concerning the comparison of parental report with IPL measures of lexical comprehension) is currently in press with *Journal of Child Language*, where it appears with the title: 'What is 'word understanding' for the parent of a one-year-old? Matching the difficulty of a lexical comprehension task to parental CDI report', co-authored by Kim Plunkett. Selected material from Chapter 5 has been accepted for publication in the inaugural edition of the journal of the UK Cognitive Linguistics Association, *Language and Cognition*, where it appears with the title: 'How do infants build a semantic system?', co-authored by Kim Plunkett, which is due to appear in May 2009. Selected material from Chapters 5 and 6 appears as part of a paper in the *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, with the title 'Priming and lexical interference in infancy', co-authored by Natalia Arias-Trejo and Kim Plunkett. This paper is discussed in Chapter 7, and the citation is included in the list of References.

Chapter 2

Norming the 18-month-old lexicon



Before investigating the detailed *organisation* of the infant lexicon, it is necessary to establish norms for the size and item-level detail of 18-month-olds' lexicons. The process of norming is both theoretically and methodologically important: From a theoretical perspective, it is difficult to develop appropriate models of lexicon organisation, without knowing the size or scope of the information contained; From a methodological perspective, it is important to establish which words are typically understood by infants within the age range of interest, so as to select the most appropriate experimental stimuli for test. This chapter describes a norming study carried out on a large sample of vocabulary surveys, drawn from the same geographic region as the planned experimental population.

Introduction

Documenting the infant vocabulary has a long history. The earliest scientific descriptions of infant development included details about the first words comprehended and produced, along with idiosyncratic word meanings and overextensions in production (Chamberlain & Chamberlain, 1904; Darwin, 1877; Taine, 1877). Later linguistic researchers documented individual infants' linguistic development in detailed language diaries, which described the words used by infants, along with detailed information about the social and linguistic contexts in which they occurred (Leopold, 1948; e.g., Moore, 1896; K. Nelson, 1973). Over the past three decades, cross sectional and longitudinal information about the size and word types of the infant lexicon has also been collected in a variety methods, including structured parental interviews, vocabulary checklists, videotaped observation sessions, and comprehension tests.

Data collection methods used in a number of key studies are summarised in Table 2.1. Note that earlier works, with their labour-intensive longitudinal data collection methods, documented vocabulary development in relatively few infants. Given the growing recognition of the value of parental report, many studies now employ a parental checklist based on the cued recognition of large numbers of items. The cued recognition approach has been shown to be more comprehensive than free recall of items (Snyder *et al.*, 1981). Checklist approaches also demonstrate a high test-retest reliability (Fenson et al., 1994; J Steven Reznick & Goldfield, 1994), and correlate well with other developmental measures including observed language production (Bretherton *et al.*, 1983; J Steven Reznick & Goldfield, 1994), experimental comprehension tasks (Bretherton et al., 1983; Ring & Fenson, 2000), and established indexes of early developmental such as the Bayley infant development scales (Benedict, 1979; Phillip S Dale *et al.*, 1989) and the Peabody Picture Verification task (Bretherton et al., 1983).

The MacArthur-Bates Communicative Development Inventories (Fenson et al., 1994), are now the most widely used parental checklists. The two communicative development inventories (CDIs) are designed for children at different ages, with the *Words and Gestures* checklist documenting the comprehension and production of infants aged 8- to 16-months, and the *Words and Sentences* checklist documenting production for children aged 16- to 30-months. The MacArthur-Bates CDIs have been normed over a number of populations in the USA (Philip S Dale & Fenson, 1996; Feldman *et al.*, 2000). These checklists have also been used to produce the LEX Database (Philip S Dale & Fenson, 1996), which lists the rate at which each word in the MacArthur-Bates checklist is comprehended or produced at different ages.

Table 2.1. Data collection methods in selected key studies

			Design		Diary		Interview		Checklist		Expt. Test		Observ.		
Study	N	Age (months)	Longitudinal	Cross-Sectional	COMP	PROD	COMP	PROD	COMP	PROD	COMP	PROD	COMP	PROD	other
Moore (1896)	1	0-24	Y	-	<15m	Y	-	-	-	-	-	Y	-	-	
Chamberlain et al., (1904)	1	0-33	Y	-	-	Y	-	-	-	-	-	-	-	-	
Nelson (1973) (first 50 words)	18	12-24	Y	Y	-	Y	-	Y	-	-	Y	Y	-	Y	Bayley
Goldin-Meadow (1976)	12	14-26	-	Y	-	-	-	-	-	-	Y	Y	-	-	
”	3	14, 19, 22	Y	-	-	-	-	-	-	-	Y	Y	-	-	
Benedict (1979) (first 50 words)	8	9-15, 17-21, 29	Y	-	<10m	<10m	Y	Y	>10m	-	Y	-	Y	Y	Bayley
Snyder, Bates & Bretherton (1981)	32	14	-	Y	-	-	Y	Y	Y	Y	-	-	-	-	
Bretherton, McNew et al. (1983)	30	20, 28	Y	-	-	-	Y	Y	Y	Y		Y	Y	Y	PPVT
Dale, Bates, Reznick, Morriset (89)	225	20	-	Y	-	-	-	-	-	Y					Bayley
Goldfield & Reznick (1990)	18	14-22	Y	-	-	Y	-	-	-	Y	-	-	-	-	
Bates et al. (1994)	673	8-16	-	Y	-	-	-	-	Y	Y	-	-	-	-	
”	1,130	16-30	-	Y	-	-	-	-	-	Y	-	-	-	-	
”	217	8-30	Y	-	-	-	-	-	Y	Y	-	-	-	-	
Fenson et al. (1994)	659	8-16	-	Y	-	-	-	-	Y	Y	-	-	-	-	
”	1,130	16-30	-	Y	-	-	-	-	-	Y	-	EOWPVT	-	-	Bayley
Reznick & Goldfield (1994)	24	14-22	Y	-	-	Y	-	-	-	Y					
Harris et al. (1995)	6	6-24	Y	-	Y	>10m	Y	>10m	>15m	-	Y	-	Y	>10m	
Ring & Fenson (2000)	43	20-30	-	Y	-	-	-	-	-	Y	Y	Y	-	-	
Feldman et al., (2000)	2,156	10-13	Y	Y	-	-	-	-	Y	Y	-	-	-	-	-
”	2,156	22-25	Y	Y	-	-	-	-	-	Y	-	-	-	-	-
Hamilton et al. (2000)	669	12-25	-	Y	-	-	-	-	Y	Y	-	-	-	-	

Notes. Names of standardised developmental tests are given where appropriate ‘Bayley’ refers to the Bayley Infant Development Scales. PPVT refers to the Peabody Picture Vocabulary Test. EOWPVT refers to the Expressive One Word Picture Vocabulary Test.

The current study employs a British language checklist, the Oxford CDI, which was adapted from the vocabulary sections of the *MacArthur-Bates CDI: Words and Gestures*. The Oxford CDI has been normed over a large British population (Hamilton et al., 2000).

Since the earliest reports of linguistic development in infancy, there has been a general consensus that lexical comprehension precedes production. That is to say, infants show signs of comprehending particular words and phrases prior to producing their first words (Benedict, 1979; Harris *et al.*, 1995; Taine, 1877), comprehension of individual words typically precedes production of the same word (Benedict, 1979; Bretherton et al., 1983; Philip S Dale & Fenson, 1996; Moore, 1896; Snyder et al., 1981), and infants and toddlers typically comprehend more words, in total, than they produce (Benedict, 1979; Fenson et al., 1994; Goldin-Meadow *et al.*, 1976; Hamilton et al., 2000; Moore, 1896; Snyder et al., 1981). Furthermore, researchers interested in the referential nature of infant word-use have noted that infants' comprehension typically contains less over-extensions of word meaning (P. Bloom, 2002; Goldin-Meadow et al., 1976; Huttenlocher, 1974; Moore, 1896), and can be less context-bound than the production vocabulary (Harris et al., 1995; Huttenlocher, 1974), especially once the lexicon expands beyond the first twenty-or-so words (Snyder et al., 1981). This means that the developing lexicon contains more words, and more accurate representations, than the infant can reliably demonstrate in production.

Lexical production is generally assumed to catch up with comprehension shortly after the onset of the so-called 'word spurt' (e.g., Goldin-Meadow et al., 1976), but this may not actually be the case. To take an example from the adult lexicon, it is possible for an adult to have no difficulty comprehending the term *deoxyrhibonucleic acid*, and to have a clear representation of its meaning, and yet

stumble, or stall in their production of the term. In this case, comprehension accuracy is ahead of production accuracy (cf. the common stumble, *sta-stis-tics*, and the well-known tip-of-the-tongue phenomenon (Brown & McNeill, 1966)). Although productive vocabulary clearly approaches comprehension in the lexicon, the two need not perfectly align.

Given this life-long mismatch between comprehension and production, it is interesting to note that most cross-sectional studies of the developing infant lexicon dispense with quantifying comprehension after about 16 months-of-age. In Table 2.1, for example, only a small number of studies assessed comprehension beyond 16 months-of-age, and most of those employed observational language samples or small sample behavioural measures, rather than large-scale checklists. This follows the tradition of the early diarisers, who tended to document emergent behaviours, sometimes at the expense of continuing behaviours. For example, following the onset of word recognition, less attention is typically paid to articulatory development in non-speech utterances (Moore, 1896; Taine, 1877), and following the emergence of multi-word utterances, less attention is typically paid to the number or type of words in the lexicon (L. Bloom, 1973; Clark, 1973; Moore, 1896). Some parents have also reported it to be 'too difficult' to keep detailed records of comprehension once the lexicon is sizeable (Benedict, 1979; Harris et al., 1995; J Steven Reznick & Goldfield, 1994). It is in this tradition that comprehension has been omitted from the now-standard *MacArthur-Bates CDI: Words and Sentences* (Fenson et al., 1994), aimed for use with children aged 16-30 months. It is worth bearing in mind, however, that the tools deemed 'too difficult' generally required extremely detailed information about the context in which each word was encountered, *every time* the child exhibited evidence of comprehension. By contrast, the checklist format is succinct and time-

saving, and it should therefore be possible to continue assessment of comprehension beyond the first 16 months. The Oxford CDI (Hamilton et al., 2000) is one example of a vocabulary checklist which assesses both comprehension and production, and is used with children up to 36 months-of-age.

This chapter therefore attempts to address the following omissions in the literature concerning comprehension vocabulary above 16 months-of-age: a) only a single study reports the size of the *comprehension* vocabulary in a large sample of infants above 16 months-of-age (Hamilton et al., 2000), and this is also the only lexicon study investigating British infants; b) no studies focus on the detailed composition of the comprehension lexicon at 18 months-of-age over a large sample of infants; c) no norms exist for the likelihood of comprehension of individual words at 18 months-of-age.

The primary aims of the current chapter are therefore to provide descriptive information about the *number* and *type* of words in the British infant lexicon at 18-months-of-age, and also to provide norms for the likelihood of particular words being comprehended by British infants at this age. The norming of infant lexicon size will be valuable for identifying infants in future studies whose performance may lie outside the normal range of lexical comprehension. The norming of individual words will allow a pool of age-appropriate test items to be established for use in future studies. In addition, an index of possible word associations between items in the infant lexicon will be established, for manipulation in future priming studies, and comprehension rates for individual words will be compared to previously established norms for age-of-acquisition and surface frequency. The comparison between infant norms and norms for older language users allows a valuable insight into the developmental trajectory of the lexicon over the life course.

It is expected that, as in previous vocabulary research, comprehension vocabulary will exceed production vocabulary, and the two measures will correlate for individuals. In the comparison of norming studies, it is expected that the current study will produce values which are consistent with previously reported norms for British 18-month-olds assessed with the same tool. These norms may well differ from those collected in North American norming studies.

Data sample

The Oxford CDI (Hamilton Plunkett & Schaffer: 2000) is a British adaptation of the MacArthur-Bates Communicative Development Inventory (Fenson et.al: 1994). The standard list contains 416 words likely to be encountered by infants under three, and is typically used to assess lexical comprehension and production between 11 and 26 months-of-age. The list of words was originally selected and adapted from the vocabulary sections of the *MacArthur-Bates CDI: Words and Gestures* (Fenson et al., 1994). Following the format of the MacArthur-Bates CDI (M-CDI), next to each word are two circles in which care-givers are asked to mark whether their child "understands" or "understands and also says" each word. The two options allow simultaneous assessment of lexical comprehension, and production, while also allowing the caregiver to leave blank any words which are unknown at the time of assessment. At the end of the form is a space to mark 'extra words' which may be particular to the individual. Some recent versions of the Oxford CDI (O-CDI) include a regular set of 'extra words'. In the current study, the standard 416 words are assessed, with no weight given to supplementary 'extra words'. The paper document is eight pages long, including a single page of instructions. The full CDI is given in

Appendix A. Parents also note their infant's date of birth, and the date on which they filled out the CDI. Age in months is calculated from the dates provided[1].

Over five hundred parents were requested to fill in CDIs shortly before an experimental visit, where they participated in a play session involving a variety of short IPL tasks. Some of these CDIs were also included in a study by Hamilton, Plunkett and Schafer (2000), which reported cross-sectional norms for comprehension and production over the age range 12-25 months. The fully anonymised data were collated in digital form, and are publicly available on the website of the Oxford BabyLab:

http://psyweb.psy.ox.ac.uk/babylab/cdi.html

(last access for current study: 17 Feb 2006)

The full sample contained 548 CDIs, returned by caregivers of infants aged 9.5 to 27.2 months-of-age. All caregivers filled out individual CDIs between July 1990 and November 2001. Across the data set, 179 CDIs were in the range 17.5 to 18.5 months-of-age. The mean age of this sample was 17.99 months (SD = 0.25), with a range from 17.52 months to 18.48 months.

Norms for infants

This section investigates the typical size and organisation of the British lexicon at 18 months-of-age. These norms describe percentiles for comprehension and

[1] Age in months is calculated according to the number of days since the infant's date of birth, expressed as follows: age-in-months = 12*days / 365.25. The measure is calculated according to the *average* length of a month, removing the complication of calendar months varying in duration (in a non leap-year, for example, if ages are calculated in calendar months, July infants will be 3 days older than February infants for 7 months of the year, and younger, for 5).

production of the 416 words in the Oxford CDI. Norms are compared to other studies, and the detailed composition of the typical lexicon is described.

Lexicon size: Comprehension & production

Eighteen-month-olds understood a median of 174 words out of a possible 416 in the Oxford CDI (O-CDI). Reported comprehension ranged from 24 words (at 18.01 months), to 400 words (at 17.98). For the sake of comparison with other studies, production norms were also calculated: 18-month-olds produced a median of 33 words, ranging from none (four infants, mean age of 17.83 months) to 307 words (one infant at 17.85 months). Percentile norms for 18-month-olds are given in Table 2.2, along with norms reproduced from two large studies: the Hamilton et al. (2000) Oxford CDI (O-CDI) study of British infants, and the Fenson et al. (1994) MacArthur-Bates CDI (M-CDI) study of North American Infants. As previously noted, the 18-month age group falls shortly after the transition from the 'infant' checklist to the 'toddler' checklist (M-CDI: *Words & Gestures*, and M-CDI: *Words and Sentences*, respectively) in the North American norming studies. No measures of comprehension are therefore available in the North American norms. Production norms for 18-month-olds are reproduced for comparison, although note that these were collected in the M-CDI: *Words and Sentences*, which contains a much longer checklist than the O-CDI. For comparison, raw CDI scores were converted to a percentage of the total number of words in the checklist used.

Table 2.2. Percentiles of CDI comprehension and production from three norming studies

British 18-month-olds									
O-CDI Comprehension	5th	10th	25th	Median	75th	90th	95th	Total	*N*
Raw score	71	92	135	174	240	301	332	416	179
% CDI	17.1%	22.1%	32.5%	41.9%	57.7%	72.4%	79.8%	100%	179
% CDI (Hamilton et al., 2000)	-	21.7%	32.4%	44.5%	56.5%	70.9%	-	100%	90
O-CDI Production	5th	10th	25th	Median	75th	90th	95th	Total	N
Raw score	2	7	14	33	74	135	169	416	179
% CDI	0.5%	1.7%	3.4%	7.9%	17.8%	32.5%	40.6%	100%	179
% CDI (Hamilton et al., 2000)	-	1.1%	2.8%	6.1%	14.8%	29.9%	-	100%	90
North American 18-month-olds									
M-CDI Production	5th	10th	25th	Median	75th	90th	95th	Total	N
Raw score (Fenson et al., 1994)	-	15	32	80	157	222	-	680	80
% CDI (Fenson et al., 1994)		2.2%	4.7%	11.8%	23.1%	32.6%		100	80

Notes. In both comparison studies, figures were originally reported as growth curves over a number of ages. Figures from Fenson et al. (1994) were interpolated from the plotted values for 18-month-olds. Figures from Hamilton et al. (2000) were established via personal communication with the second author. *N* refers to the number of CDIs collected in each study.

Comprehension norms show a very close correspondence in the two British studies. Percentiles differed by less than three percentage points – equivalent to eleven words. Production means also exhibited a close correspondence, differing by three percentage points (thirteen words). It is worth bearing in mind that approximately half of the CDIs in the current sample were also part of the database used in the Hamilton et al. (2000) norming study, making a close correspondence between studies highly likely.

Comparing the current data to the production results of the M-CDI norming study, there is an apparently large difference between the raw scores for production. The North American sample lead production at every level, with an increasing lead for larger vocabulary infants – at the lower 10th percentile North American 18-month-

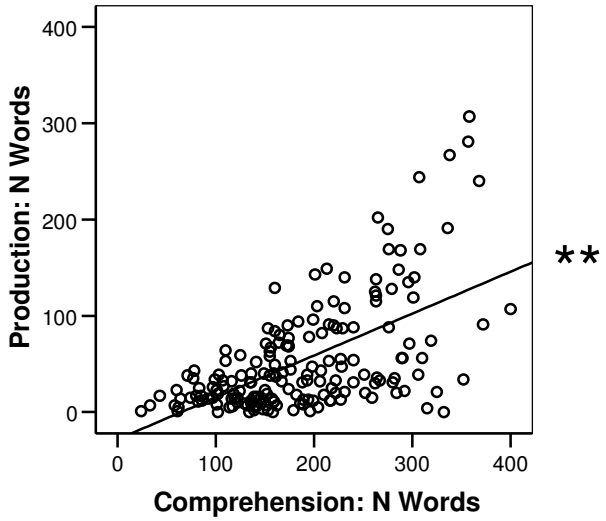
olds had a head-start of eight words in production, and at the upper 90th percentile, 87 words. Hamilton et al. (2000) also noted the tendency for North American norms to exceed British norms across development. However, recall that the M-CDI used for this age-group (18- to 30-months) is systematically different to the O-CDI (9- to 36-months). In fact, the vocabulary sections of the toddler M-CDI: *Words and Sentences* contains almost twice as many items as the M-CDI: *Words and Gestures*, upon which the O-CDI was based. Comparison of percentage scores reveals that North American infants lead production by less than six percentage points at each level (equivalent to 22 words in the O-CDI, or 36 words in the M-CDI). This difference is larger than the difference between the two British norming studies. Given that the two British studies exhibit a three percent difference (despite sharing a large number of participants), it is possible that the observed difference between British and American norms is an artefact of naturally high variability of the data. Note also that the sample sizes are quite different, with the current study sampling more than twice as many infants as the 18-month-olds sampled in the North American norms. However, it is worth bearing in mind that the cross-cultural difference was consistent and unidirectional, unlike the British studies, which varied differentially across the vocabulary scale.

Individual comprehension & production

The CDI comprehension score for each individual is plotted against their CDI production score in Figure 2.01, where the upward trend in the distribution indicates that production and comprehension increase together. The correlation between comprehension and production is highly reliable ($\rho(179) = 0.54$, $p < 0.001$), demonstrating that infants who say more words are typically those who understand more words. This finding fits well with previous large-scale norming studies, in which

similar correlations have been reported: Fenson et al. (1994) reported a correlation of .53 for infants 8 to 16 months-of-age (age partialled out); and Ring and Fenson (2000), .53 for infants between 20 and 30 months-of-age.

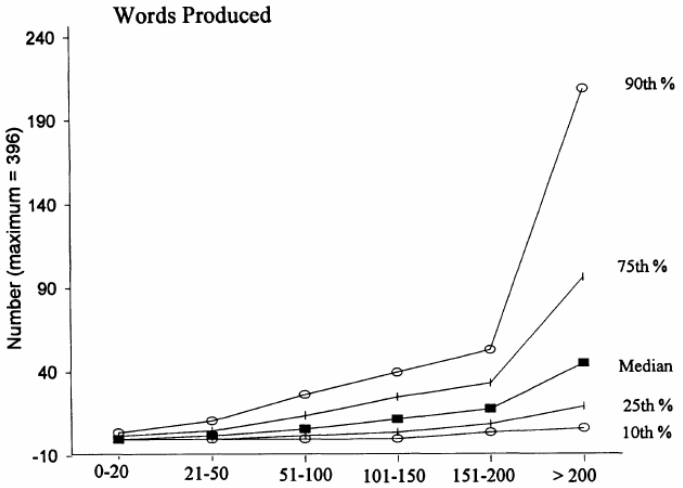
Figure 2.01. O-CDI individual comprehension plotted against production



Notes. CDI scores are measured out of 416 possible words in the Oxford CDI. Solid line indicates linear fit. **p<.01

In Figure 2.01, the data do not distribute evenly about the line describing the linear fit. In fact, no values occur above the diagonal intercept. This distribution is striking, as it suggests that no infants were reported to produce more words than they comprehended. Note also the similarity of this distribution to values presented in the MacArthur-Bates CDI norming study (Fenson et al., 1994), reproduced as Figure 2.02. Although this plot summarises data from infants across a broad age range, the pattern is quite similar, with the very highest performing infants almost equal on both scales, and no infants demonstrating a performance vocabulary larger than their comprehension lexicon.

Figure 2.02. M-CDI percentile comprehension plotted against production



Notes. Percentiles CDI Comprehension scores from infants aged 8-16 months. Reproduction of Figure 14 from Fenson et al. (1994)

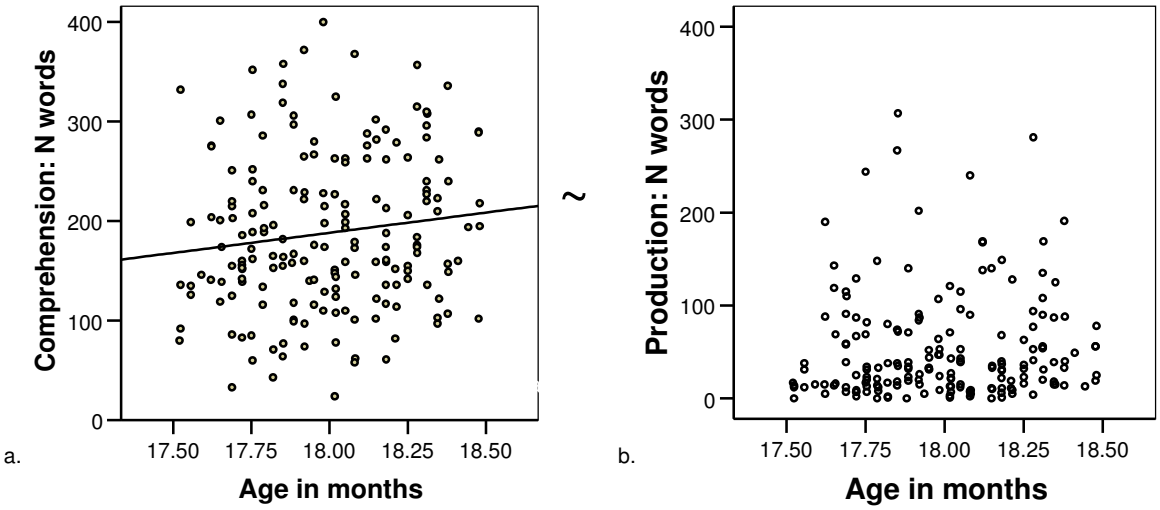
The distributional pattern found here (no children producing more words than they understand) is a necessary artefact of the CDI design. Parents are instructed to use the first circle to mark "words that your infant **understands but does not yet say**", and the second circle to mark "words that your child **understands and also says**" (original bolding & underline). The instruction explicitly precludes parents from indicating a word which is said before it is understood. While it seems intuitively obvious that individual words would be comprehended before they are produced, this is not necessarily the case – a point which will be returned to in the discussion.

Age & lexicon size

Age in months is plotted against vocabulary size in Figure 2.03. In Figure 2.03a, large variability in comprehension is evident, with broadly distributed scores throughout the age-range 17.5-18.5. Age-in-months did not correlate with infants' ranked comprehension score ($\rho(179) = .13$, n.s.). There was evidence of a weak trend

in the expected direction, which barely approached the alpha level of .05, with a *p*-value of .079. In Figure 2.03b, the skew towards lower production scores is apparent, along with a broad distribution across the age range. Ranked production scores did not correlate with ranked age in months ($\rho(179) = .11$, n.s.), and no trend was evident.

Figure 2.03. Comprehension & production plotted against age-in-months



Notes. CDI scores are measured out of 416 possible words in the Oxford CDI. Solid line indicates linear trend. ~p<.10

The weak relationship between comprehension and age-in-months suggests that within such a narrow age bracket, within-group analysis of age is unlikely to be fruitful in future studies.

Composition of the lexicon

Different researchers have used a variety of coding schemes to classify words in the infant lexicon. The current study will primarily focus on the nineteen word groupings of the CDI, as a way of interrogating the relationship between the tool and the infant lexicon. Some of the word groupings are super-ordinate semantic categories (e.g., *animals*), some correspond to grammatical word classes (e.g., *action words*:

verbs, *descriptive words*: adjectives), and some are rough groupings which help to disambiguate between different word meanings (e.g., *fish (animal)* or *fish (food)*), and to cue parents' memory for words which might be experienced in a given domain (e.g., *furniture and rooms, outside*). The word groups of the Oxford CDI are (in order of appearance): *animal sounds*, *animals*, *vehicles*, *toys*, *food and drink*, *body parts*, *clothes*, *furniture and rooms*, *outside*, *household items*, *people*, *games and routines*, *action words*, *descriptive words*, *time*, *pronouns*, *prepositions* and *quantifiers*. The word groups of the CDI contain different numbers of items, ranging from six words (*question words*) to 68 words (*action words*).

For each word group, the median number of words comprehended is given in Table 2.3. More than half of the words were reported as understood by 18-month-olds in the word groups *animal sounds, animals, vehicles, toys, food and drink, body parts and games and routines*. The percentage of words comprehended generally decreases in the word groups towards the end of the checklist. These word groups contain a smaller number of items, and contain more abstract, grammatical parts of speech.

Another way of considering these data is to look at the word groups as a proportion of the total vocabulary. Figure 2.04 plots the proportion of the median vocabulary which is made up of the different word groups, next to the proportion of the 416-word CDI devoted to each word group. In this plot the word groups have been ordered according to the major grammatical word classes represented in the CDI: nouns (*animals, vehicles, toys, food and drink, body parts, clothes, furniture and rooms, outside, household items, people*) represented by the red, yellow, and pink groups at the base of the chart; verbs (*action words*) represented by white; adjectives (*descriptive words*) represented by pale blue; the grammatical word groups (*question words, time words, pronouns, prepositions* and *quantifiers*) represented by greys;

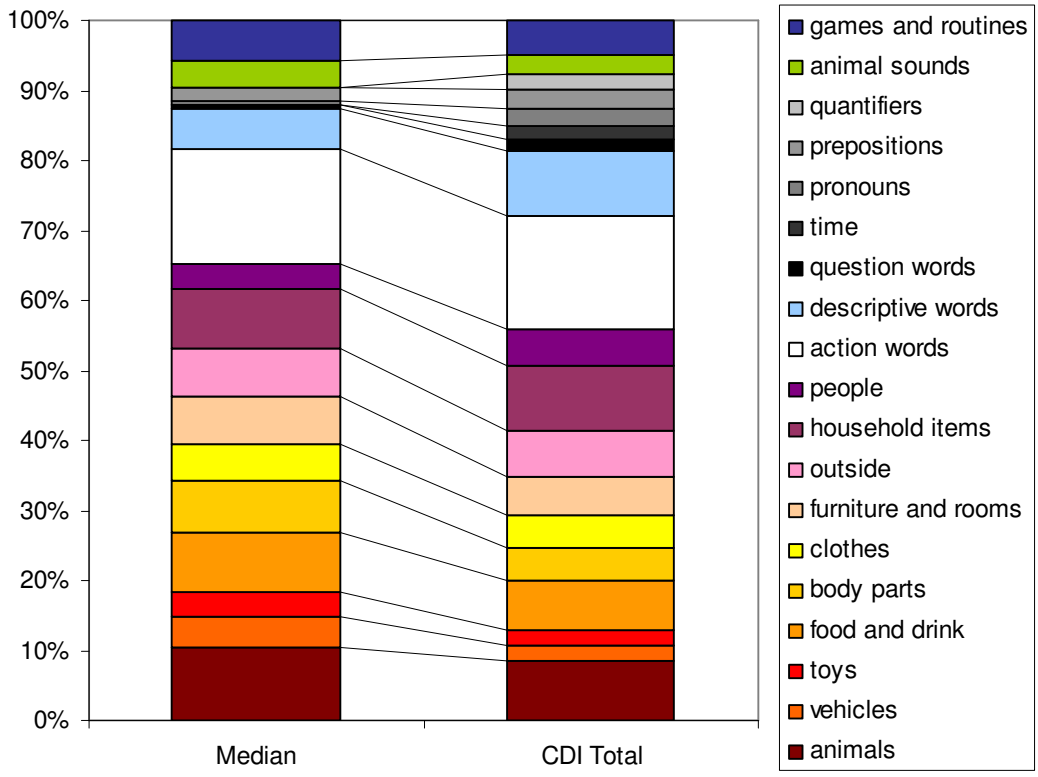
onomatopoeia (*animal sounds*) represented by pale green, and a mixed group (*games and routines*), represented by dark blue. The 'mixed group' contains words classed as interjections in adult language (e.g., *hello, night-night*), as well as nouns describing daily routines (e.g., *bath, tea*) and common games (e.g., *pat-a-cake, peekaboo*).

The words in the mixed group are likely to hold a variety of functions in the infant lexicon (some are likely to hold the status of ritualised utterances, interjections or performatives, while others may function as nouns). It was decided not to separate this group of words for the current analysis, given the difficulty of deciding what status each item was likely to hold in the 18-month-old lexicon.

Table 2.3. Number of words comprehended in each CDI word group

CDI Word Group	Median	%	Range	Total
animal sounds	7	58.3	0-12	12
animals	18	51.4	0-35	35
vehicles	8	80.0	2-10	10
toys	6	66.7	0 - 9	9
food and drink	15	51.7	1-29	29
body parts	13	65.0	0-20	20
clothes	9	47.4	0-19	19
furniture and rooms	12	52.2	0-23	23
outside	12	44.4	0-27	27
household	15	38.5	0-39	39
people	6	28.6	1-21	21
games and routines	10	50.0	2-20	20
action words	29	42.6	1-68	68
descriptive words	10	25.6	0-29	39
question words	1	16.7	0 - 6	6
time	0	0.0	0 - 8	8
pronouns	1	9.1	0-11	11
prepositions	3	27.3	0-11	11
quantifiers	0	0.0	0 - 9	9

Figure 2.04. O-CDI composition of the 18m lexicon: CDI word groups



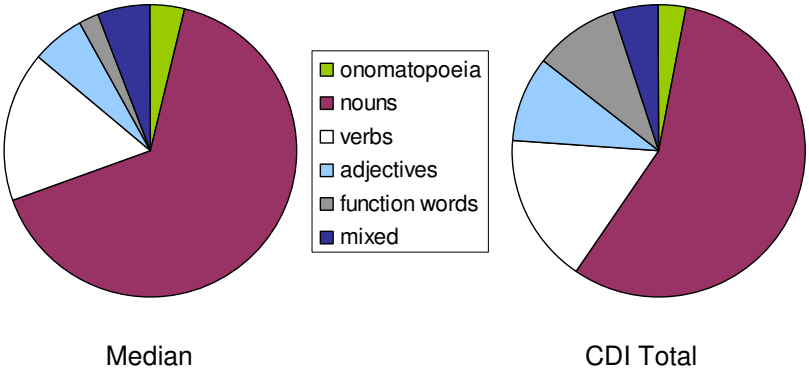
Notes. The 'median vocabulary' is created as the total of the median values for each word group, and therefore represents a hypothetical 'median infant'.

In Figure 2.04, the generally low slope of the lines joining the two plots shows a general correspondence between the composition of the O-CDI and the reported median comprehension vocabulary. This pattern was confirmed by a correlation between the proportion of the O-CDI devoted to each word type and the proportion of words of each type reported in the median comprehension vocabulary ($\rho(19) = .903$, $p < .001$). The plot also graphically demonstrates that a larger proportion of words are reported as comprehended in the noun classes vehicles and furniture, than the proportion included in the CDI form, while proportionally less words were reported as comprehended in the grammatical word groups question words, time, pronouns, and quantifiers.

To make the relationship between the grammatical word classes more explicit, Figure 2.05 plots the composition of the 18-month vocabulary according to grammatical word classes. Sixty five percent of the words in the median vocabulary for 18-month-olds were nouns, seventeen percent verbs, while each of the other word classes accounted for less than ten percent. The composition of the O-CDI form is provided on the right of Figure 2.05 for comparison, where the proportion of the form dedicated to nouns appears to be smaller (56%), and the proportion allocated to adjectives and the grammatical classes appear to be larger (10% each). In a one-tailed test of correlation (expecting longer word lists to pattern with higher reports in comprehension), the number of words in each CDI word group did not significantly correlate with the median number of words understood in each word group ($\rho(6) = .72$, n.s.). A trend was evident, which approached the alpha level of .05, with a *p*-value of .054.

These comparisons demonstrate that the median lexicon at 18 months-of-age is similar to the organisation of the Oxford CDI. A tighter correlation might suggest that the composition of the Oxford CDI was 'leading' or 'restricting' the responses collected (although this is clearly true for large vocabulary children, who are reported to know the majority of the listed words). Within this age range, the composition of the lexicon suggests that the different grammatical classes might be valuable for identifying individual differences between infants in the tails of performance. In particular, nouns might be more diagnostic for finding differences between infants with low comprehension scores, while adjectives and function words might be more diagnostic for identifying differences between high scorers.

Figure 2.05. O-CDI composition of the 18m lexicon: Grammatical word class



Notes. The 'median vocabulary' is created as the total of the median values for each word group, and therefore represents a hypothetical 'median infant'.

The finding that nouns make up the largest word class reported in the lexicons of 18-month-olds is uncontroversial (Benedict, 1979; Goldin-Meadow et al., 1976; Huttenlocher, 1974; Leopold, 1948). Having validated this finding in the sample population of interest (British 18-month-olds), the studies included in this thesis will focus on noun comprehension, as this word class provides the largest number of words for selection of test stimuli.

Norms for words

The rate of comprehension was calculated for each word in the O-CDI, as the percentage of the 179 infants in the 18 month age group (17.5 to 18.5) reported to comprehend the word. These norms follow the format of the LEX database of norms (Philip S Dale & Fenson, 1996), created from the data collected in the M-CDI norming study of North American infants and toddlers (Fenson et al., 1994). The norms in the current study here are compared with the LEX 2005 comprehension norms (Philip S Dale & Fenson, 1996), and the comprehension rate of different word types is described.

Comprehension norms

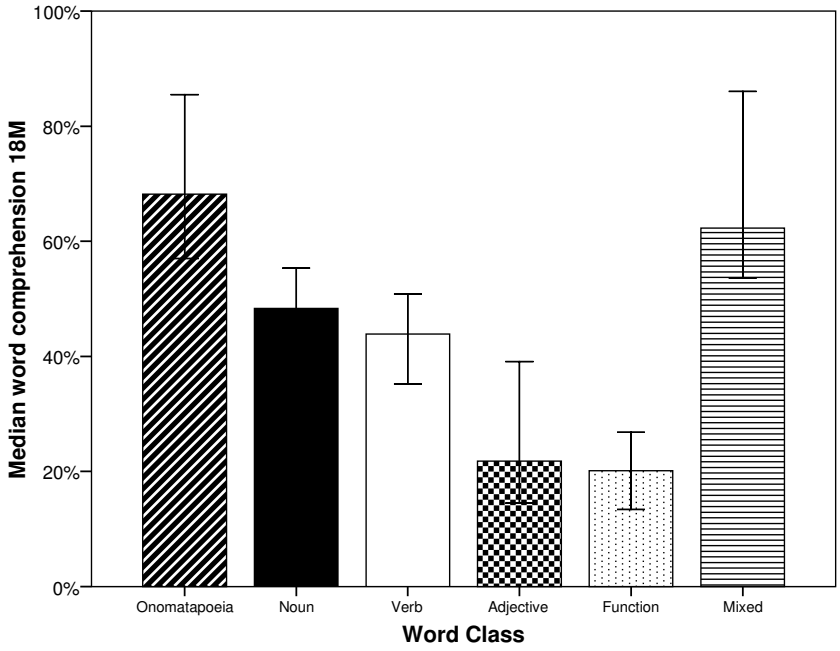
The full list of comprehension rates for O-CDI words is given in the appendices where words are presented in two orders: alphabetical order (Appendix B), and in descending order of comprehension within each word group (Appendix C). Comprehension rate was not normally distributed across the 416 standard words in the O-CDI, with a skew towards lower values. The median comprehension rate was 43.6% (for the word *picture*). The interquartile range (25^{th} to 75^{th} percentile) was from 23.0% to 64.7%, and extremes (5^{th} and 95^{th} percentile) fell at 7.7% and 88.9%. The best known word for 18-month-olds was *no*, known by 99.4% of 18-month-olds. The least known CDI words at 18-months were *old* and *fine*, each known by 1.7% of infants.

Table 2.4 summarises the median comprehension rate of items, according to the word groups of the CDI. In this table it is clear that words in different groups differed in their typical comprehension rate. Figure 2.06 displays comprehension rate of words according to grammatical word class. Comprehension rate differed systematically across the word classes ($\chi^2(5) = 68.9$, $p < .001$). Note that the word group for onomatopoeia (*animal sounds*), and the mixed word group (*games and routines*) make up very small proportions of the reported lexicon (6% and 4%, respectively), yet the words in these groups have the highest rates of comprehension.

Table 2.4. Median comprehension rate for words in each CDI word group

CDI Word Group	Median %	Min%	Max%	Total words
animal sounds	68.2	50.3	88.3	12
animals	54.2	3.9	98.3	35
vehicles	67.9	26.8	96.6	10
toys	60.3	52.0	97.8	9
food and drink	44.1	13.4	90.5	29
body parts	72.6	16.8	92.7	20
clothes	48.0	10.6	98.3	19
furniture and rooms	61.5	5.0	93.3	23
outside	39.1	4.5	77.7	27
household	39.7	8.9	87.2	39
people	16.2	5.6	98.9	21
games and routines	62.3	23.5	99.4	20
action words	43.9	7.3	92.7	68
descriptive words	21.8	1.1	84.9	39
question words	18.7	2.8	58.7	6
time	7.8	2.8	20.7	8
pronouns	16.2	5.6	33.0	11
prepositions	40.2	22.9	67.0	11
quantifiers	18.7	6.1	82.7	9

Figure 2.06. O-CDI median comprehension rate: Grammatical word classes



Notes. Whiskers show +/- 95% confidence interval.

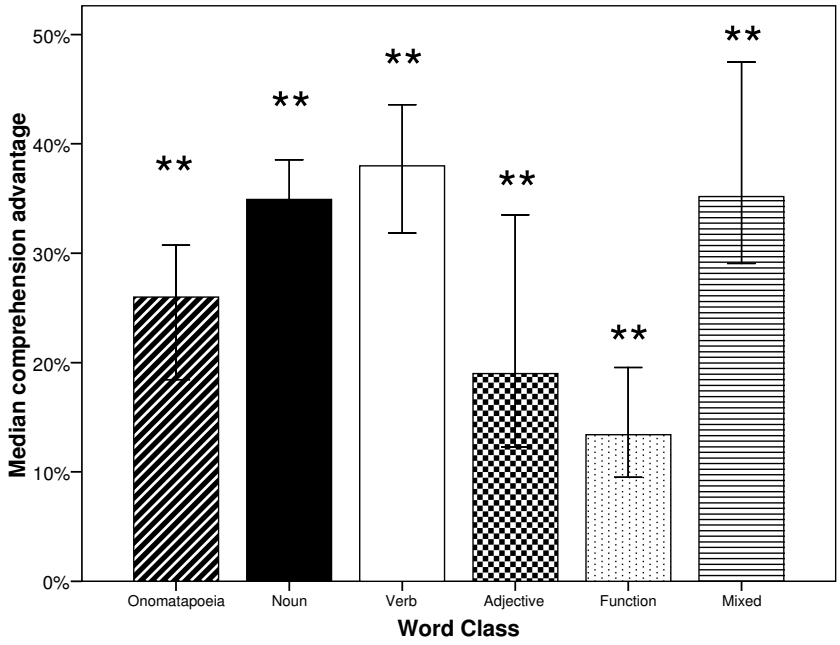
Production norms

For the purposes of comparison, production norms for words in the CDI were also calculated, as the percentage of infants reported to comprehend and also produce each item. As the focus of the current research involves the organisation of the comprehension lexicon, production norms are summarised. The median ranked words (*swim*, *blue*, *pretty*, *yellow*, *donkey*, *kitchen*, *don't*, *picture*, *plate*, *work*, *away*, *off*, *on*, *my*) were produced by 7.3% of the 179 infants in the sample, falling below the mean rate of production (12.9%), demonstrating a large skew towards low rates of production. Words ranged in production from 0% (8 words) to 80.5% (*daddy* and *mummy*). No words were universally produced.

Comprehension rate & production rate

The median comprehension rate (43.6%) was higher than the production rate (7.3%), a pattern confirmed by Wilcoxon's signed rank test ($z(416) = 17.67$, $p < .001$). Figure 2.07 illustrates the comprehension advantage (comprehension rate minus production rate) in the different word classes. The advantage for comprehension over production was present in all word groups of the CDI, and comprehension advantage differed significantly between the word classes ($\chi^2(5) = 54.8$, $p < .001$). As can be observed in Figure 2.07, comprehension advantage for the function words is quite small, while comprehension advantage for nouns, verbs and the mixed group is much higher.

Figure 2.07. O-CDI comprehension advantage: Grammatical word classes

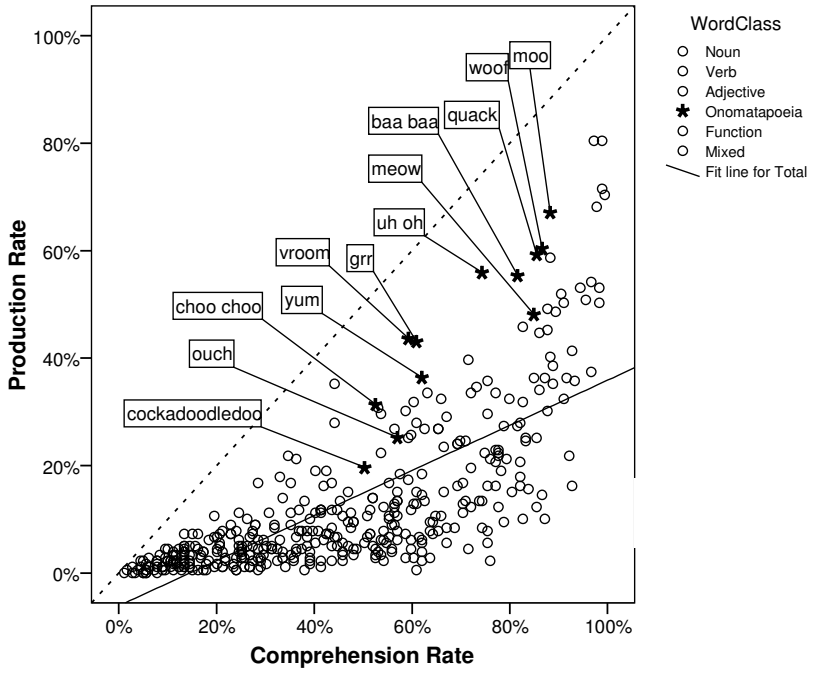


Notes. Advantage calculated as comprehension rate minus production rate for each word. One sample t-tests compare comprehension advantage to chance value of 0%. **p<.01. Whiskers show +/- 95% confidence interval.

Figure 2.08 plots the comprehension rate against the production rate for each word in the O-CDI. The distribution of this plot mirrors the distribution of Figure

2.03, with no values falling above the diagonal intercept. The rate of comprehension for each word correlated with rate of production ($\rho(416) = 8.03$, $p < .001$). This plot highlights words in the onomatopoeia class (*animal sounds*), as all of the words in this class lie very close to the diagonal intercept, demonstrating almost equal rates of production and comprehension. No other class of words shows this distributional pattern.

Figure 2.08. O-CDI comprehension rate & production rate (Onomatopoeia highlighted)

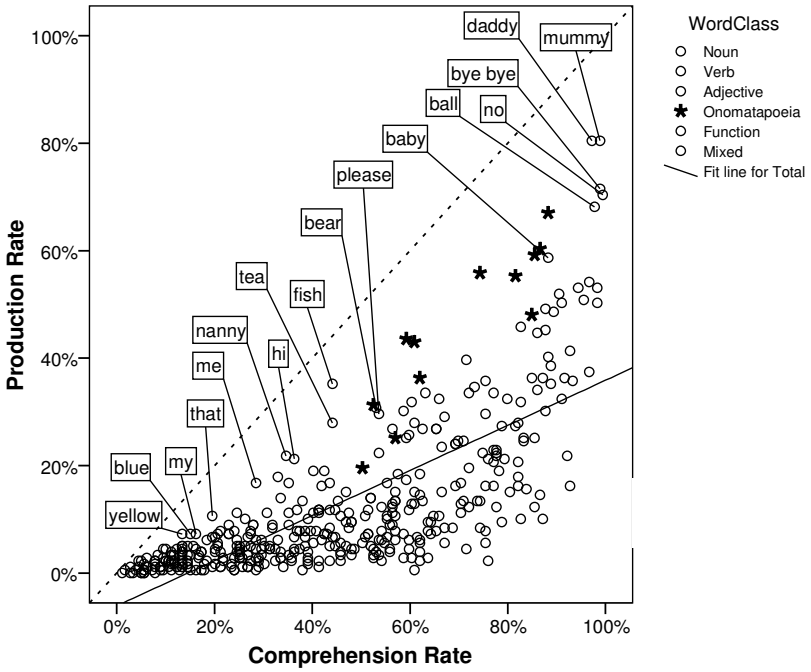


Notes. Dashed line indicates maximum possible correspondence between comprehension and production. Solid line indicates linear fit for total sample. Onomatopoeia words highlighted.

For comparison, Figure 2.09 highlights non-onomatopoeia words falling close to the maximum possible comprehension/production correspondence. These words come from a variety of word classes, and include words for people (*mummy, daddy, baby*, *nanny*), the grammatical word classes (*me, my*, *that*), a small number of common nouns (*ball, fish, tea*), adult interjections (*bye bye, no, hi, please*), as well as

the adjectives *blue* and *yellow*. These last two words are of some interest, as they represent half of the colour words included in the standard O-CDI.

Figure 2.09. O-CDI comprehension rate & production rate (other special cases)



Notes. Dashed line indicates maximum possible correspondence between comprehension and production. Solid line indicates linear fit for total sample. Non-onomatopoeia words near the maximum correspondence are highlighted.

Comparison of word norms from two studies

As the Oxford CDI (O-CDI) was adapted from the vocabulary sections of the *Words and Gestures* version of the MacArthur-Bates CDI (M-CDI), the majority of words are common to the two tools. This allows an almost direct comparison of the normed values obtained in the current study, with the LEX 2005 Database (Philip S Dale & Fenson, 1996), which gives values for the mean comprehension of the 396 words in the M-CDI. The M-CDI sample contains item-level data from 63 16-month-olds (30 male, 33 female) included in the Fenson et al. (1994) developmental norms for North American English (recall that the M-CDI *Words and Gestures*, which

records both comprehension and production, is employed for typically developing infants up to 16 months-of-age). The current O-CDI sample contains data from 159 18-month-olds, some of whom were included in the Hamilton et al. British English (2000) developmental norms for British English. As the age of interest in the current study is 18 months-of-age, an age-advantage is expected for the British infants. However, given that Hamilton, Plunkett and Schafer (2000) reported North American CDIs to be systematically higher than British CDIs, it is possible that the advantage will be moderate.

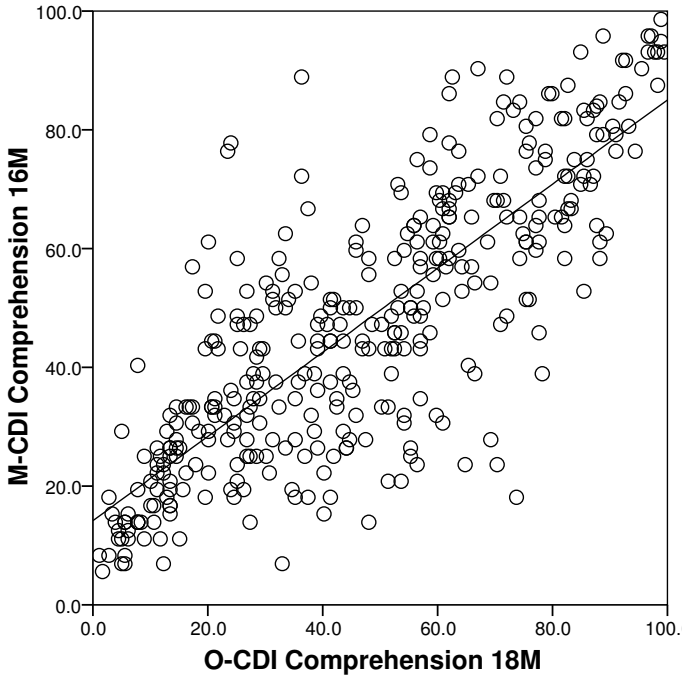
The O-CDI and the M-CDI contain 340 identical words, and 9 which differ in a superficial way (such as *yum* for *yum yum*, *kitten* for *kitty*, or *bubble* for *bubbles*). There are also 22 words for which a British dialect equivalent directly replaces a North American term (e.g., *cot* for *crib*, *nappy* for *diaper*, *trousers* for *pants*), including 8 where a second British alternative appears alongside a word common to both dialects (e.g., *truck / lorry* for *truck*, *bunny / rabbit* for *bunny*). In two cases, a single British term was equivalent to two North American words, and in these cases comprehension of North American words *cracker* and *cookie* was averaged for comparison to the British word *biscuit*, and comprehension of *stove* and *oven* was averaged for comparison to the (freely varying) British term *cooker / stove / oven*. The M-CDI contained 11 words which had no equivalent in the O-CDI (*bug*, *noodles, cheerios*, *raisin*, *beads*, *owie (body-part)*, *home*, *backyard*, *sitter's-name*, *cute*, *yucky*). The O-CDI contained 46 words which had no equivalent in the M-CDI including 35 words from the common word categories (*spider*, *boat*, *chips*, *jam*, *tea (food)*, *nail (body-part)*, *bin*, *dummy*, *jug*, *mug*, *bucket*, *wall*, *doctor*, *policeman*, *friend*, *tea (games and routines)*, *catch*, *call*, *carry*, *cuddle*, *cut*, *drop*, *find*, *have*, *hear*, *know*, *like*, *make*, *scratch*, *tell*, *green*, *nasty*, *yellow*, *sad* and *again*), and the category of

prepositions (*away*, *back*, *down*, *in*, *inside*, *off*, *on*, *out*, *there*, *under* and *up*), which was only included in the O-CDI. Altogether, there were 370 words available for comparison.

Between CDIs, two words were categorised into different sections of the form (*glasses*, a 'small household item' in the M-CDI and 'clothing' in the O-CDI, and *teddy bear*, an 'animal' in the M-CDI and a 'toy' in the O-CDI). The O-CDI classification was treated as categorical.

Comprehension rates of individual words in the two norming studies is plotted in Figure 2.10, where it is clear that words which attracted high comprehension scores in the M-CDI also attracted high comprehension scores in the O-CDI study. Comprehension was correlated between the two studies ($\rho(370) = .804$, $p < .001$). The median O-CDI word was comprehended by 46% of British 18-month-olds (range: 1% to 99%). The median M-CDI word was comprehended by 47% of North American 16-month-olds (range: 6% to 99%). Despite the British two-month age advantage over the M-CDI infants, no difference in comprehension scores was evident ($z(370) = 1.01$, n.s.). This finding is consistent with the general advantage of North American CDIs over British CDIs reported in Hamilton et al. (2000). However, it is possible that while the overall size of the lexicon is the same between the two populations, the internal structure of the lexicon may be quite different.

Figure 2.10. Comprehension rates: M-CDI plotted against O-CDI



Notes. Circles indicate comprehension rates for individual words common to both the MacArthur Bates CDI: Words and Gestures (M-CDI) and the Oxford CDI. Solid line indicates linear fit.

For this analysis, the CDI words were split into the word groups of the O-CDI. Comprehension rates in each group are given in Table 2.5, where there are clear differences between the two populations in some of the word groups: Notably, a*nimal sounds, animal words, body parts* and *vehicles* were understood by a greater proportion of British 18-month-olds than North American 16-month-olds, while *action words, adjectives, household objects, pronouns* and *time words* were understood by a greater proportion of North American 16-month-olds than British 18-month-olds.

Table 2.5. Comprehension of words in different sections of the CDI, according to two studies

	Communicative Development Inventory				
	O-CDI at 18M	M-CDI at 16M			
CDI Section	*M (SD)*	*M (SD)*	*Z*	*N*	*p*
Action Words	49.1 (18.6)	54.7 (17.4)	-3.19**	54	.001
Adjectives	31.9 (21.7)	36.9 (21.6)	-2.98**	35	.003
Animal Sounds	70.3 (14.6)	61.4 (13.5)	-2.12*	12	.034
Animals	50.1 (27.4)	37.3 (23.1)	-3.50**	34	<.001
Body Parts	68.1 (18.1)	61.5 (18.2)	-2.42**	19	.016
Clothes	50.3 (26.7)	48.4 (23.0)	-0.12	19	.904
Food & Drink	49.9 (24.5)	48.9 (22.1)	-0.37	26	.713
Furniture & Rooms	53.7 (26.6)	53.0 (17.8)	-0.76	23	.447
Games & Routines	65.4 (25.0)	72.5 (17.3)	-1.21	19	.227
Household Objects	45.3 (20.7)	50.6 (19.5)	-2.16*	35	.031
Outsides	38.3 (22.7)	36.3 (20.5)	-0.82	25	.412
People	35.8 (31.7)	38.8 (30.1)	-1.20	18	.231
Prepositions	40.6 (15.2)	N/A	-	11	-
Pronouns	17.1 (8.6)	33.4 (13.4)	-2.94**	11	.003
Quantifiers	22.9 (24.7)	25.2 (17.4)	-1.40	8	.161
Question Words	22.5 (21.5)	29.2 (17.2)	-1.36	6	.173
Time	10.1 (6.4)	17.8 (10.1)	-2.52**	8	.012
Toys	72.7 (17.8)	69.6 (16.3)	-0.65	9	.515
Vehicles	65.2 (22.4)	50.8 (23.6)	-2.31**	9	.021

Notes. Where differences in word section occurred between forms, the Oxford CDI was used as the classifier. Comparisons made where word (or dialectial equivalent) is found in both CDI forms. Wilcoxon's Signed Ranks test compares comprehension for words according to the two norms. **p<.01, *p<.05. No prepositions were included in the M-CDI.

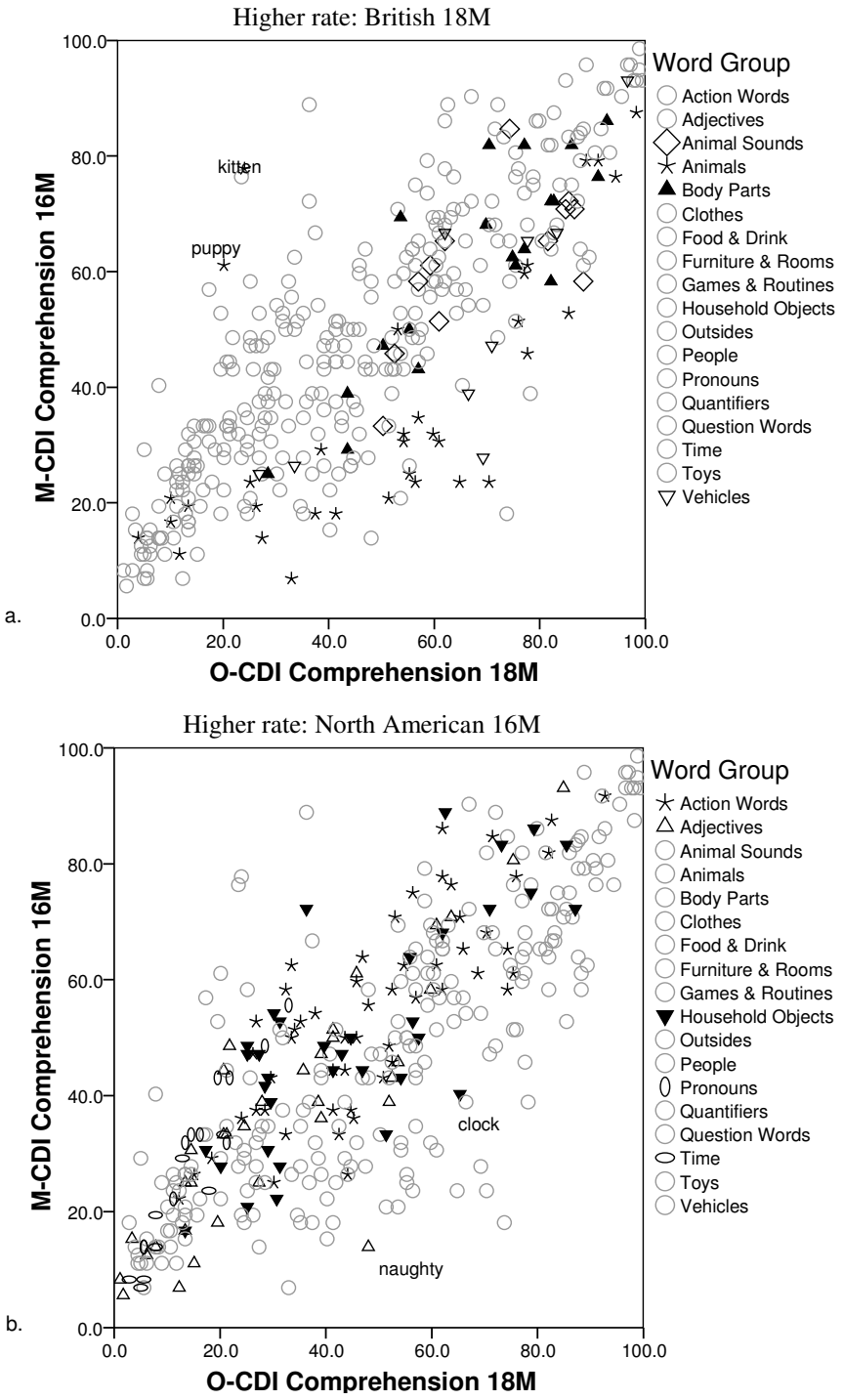
These relationships are graphically depicted in Figure 2.11, which reproduces the data from Figure 2.10, with those word groups highlighted which demonstrated significant advantage for one population or the other. In Figure 2.11a, the distribution of word groups showing an advantage for the British 18-month-olds can be seen separately from the other word groups. Note that two items in the *animal* group show an uncharacteristic advantage for the North American 16-month-olds. These two items were *kitten* (*kitty* in the M-CDI) and *puppy*. In Figure 2.11b, the word groups showing an advantage for the North American 16-month-olds are depicted. In this

case, two words showing an uncharacteristic advantage for British 18-month-olds are *clock*, and *naughty*.

The particular groups of words showing advantage are of some interest. It is interesting to note that three of the five word groups showing a North American advantage are from more abstract and relational word classes than the names of animals and vehicles. Action words and adjectives are also the two largest word groups in both CDIs, as can be seen quite clearly in from the larger number of highlighted words in Figure 2.11b. The different distributions of comprehension across the CDIs may reflect legitimate differences in the linguistic environments of the children in two different populations – it may be the case that North American infants hear and respond to more action words and more adjectives than British infants (who in turn are exposed to, and respond to more vehicles and animal names). Alternatively, it may be a difference in the nominal/referential 'style' of individuals in the two sample populations (e.g., Snyder et al., 1981).

A third possibility is a difference in reporting biases. Much has been made of the likelihood of parental overestimation of comprehension in the North American literature (Feldman et al., 2000; Tomasello & Mervis, 1994), and it may be the case the overestimation is more likely for North American parents when the word is relatively abstract, while they are more conservative in their report of concrete referential words which they may be able to 'test' for themselves. On the flip-side, this suggests that British parents are more conservative in their estimation of non-concrete referential words which they are unable to test directly.

Figure 2.11. Comprehension advantage for word groups: Two populations



Notes. 'Higher rate' determined according to results of Wilcoxon signed rank test (Table 2.04). Comprehension is calculated as the proportion of infants in the age group who are reported to 'understand' or 'understand and also say' each word.

One final possibility which may tie the lack of age-based enhancement for 18-month-olds to the internal differences in the word types showing advantage, concerns

the development of the MacArthur-Bates CDI itself. As a research tool developed over several years, it may be uniquely attuned to specific features of the infant population its former versions were piloted on. The tool may have been specifically developed to mirror North American developmental trends. The authors of the MacArthur-Bates CDIs (Fenson et al., 1994) describe gradual adjustments of the tool over a number of years. The authors describe curtailing some sections of the CDI and expanding others on the basis of information from a large pilot study. These 'refinements' may have given statistical weight to parts of the lexicon in which North American infants (the sample population) outperform British infants – take, for example, the inclusion of *time words* (in which the North American sample outperform the British sample), but the exclusion of prepositions (which had a median comprehension rate of 40.2% for the British sample, but were untested in the North American sample). It is therefore possible that the North American numerical advantage reflects a tighter correspondence between the assessment tool and the population norm, for those items on which both groups were assessed.

Comprehension, age-of-acquisition, frequency and association

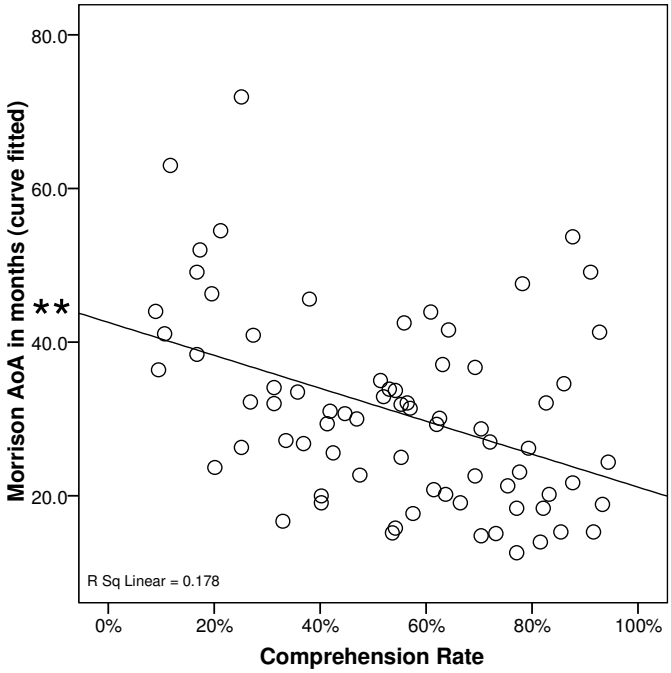
In this section, the norms for comprehension rate are compared to previously established lexical norms for later childhood and adult language. This provides valuable information about factors which may need to be controlled for in the selection of future test stimuli.

Comprehension rate & Morrison age-of-acquisition

Recent studies have proposed that some frequency effects in lexical access may be due to underlying differences in the age at which words are first learned, and thus, their integration into the lexicon. For example, Morrison and colleagues (Morrison & Ellis, 1995; Morrison *et al.*, 1992) have demonstrated that when age-of-acquisition (AoA) is included as an independent variable in a picture naming task, frequency effects disappear, although this effect may not hold true for all types of word (Boulenger *et al.*, 2007). In addition, age-of-acquisition has been shown to interact with priming effects in adults (Catling & Johnston, 2006), and to affect lexical comprehension in school age children (Funnell *et al.*, 2006), making it an important variable for consideration in the planned experimental studies.

In a large norming study of British children's AoA (Morrison *et al.*, 1997), 280 children between the ages of two-and-a-half, and eleven years took part in a picture naming study. They were asked to name a sample of 297 pictures representing concrete nouns, many of which were also included in the O-CDI. A standardised age-of-acquisition was then extrapolated from the observed data according to two methods: a curve fitting procedure and a production threshold of 75% competence. A small number of adult participants also provided subjective AoA Ranks along a 7-point scale, according to when they thought they had probably acquired each word. Note that the curve fitting procedure is not able to provide values for every word tested.

Figure 2.12. O-CDI 18m comprehension rate & Morrison age-of-acquisition



Notes. Age-of-acquisition extrapolated from experimental data according to a curve fitting procedure. Solid line indicates linear fit. ***p*<.01

108 of the words in the Morrison study overlapped with the 232 nouns in the O-CDI. Figure 2.12 plots the comprehension rate at 18-months against the objectively established age-of-acquisition norms from the Morrison study. This plot shows a downward trend, with early age-of-acquisition patterning with high rate of comprehension. Comprehension rate was negatively correlated with observed AoA values computed according to both methods (curve fitted: $\rho(74) = -.40$, $p < .001$; 75% threshold: $\rho(108) = -.49$, $p < .001$). A correlation was also evident between each word's comprehension rate and the 'rated' age-of-acquisition established in the same study ($\rho(108) = -.64$, $p < .001$).

This correlation demonstrates that the lower the comprehension rate at 18-months, the later a word is ordinarily 'acquired'. As the Morrison AoA norms are based on experimentally derived data, this finding provides external validity to the comprehension rates established in this study: The words which *parents report* to be

understood by less 18-month-olds are *demonstrated* to be acquired later by the majority of children. In addition, the correspondence between comprehension rate and AoA assessed in early childhood suggests that ranking of comprehension rate changes little over a period of some years. This finding suggests that comprehensive vocabulary norming will not require replication for other age groups in this series of experiments.

Comprehension rate & Leech word frequency

Word frequency is also a strong predictor of reaction time in adult lexical access (e.g., Savin, 1963), and is thus considered important in the design of adult priming studies (Meyer & Schvaneveldt, 1976). During lexical development, however, it is unclear whether word frequency in caregiver language plays a role in the acquisition of individual words, and the subsequent ease of lexical access during infancy. Morrison et al. (1997) demonstrate a relationship between adult word frequency and experimentally derived estimates of age-of-acquisition, which suggests that frequency in caregiver speech might be involved in the age at which a word is acquired. To give a simple example, the word *leg*, has a surface frequency of 5089 tokens in the 10-million-word British National Corpus[2] ("The british national corpus, version 3", 2007), and is acquired quite early (understood by more than half of the 18-month-olds in the current study). By contrast, *lemon*, has a frequency of 1209 tokens, and is reported in the Morrison norms to be acquired at around 3-4 years, while *leopard* (225 tokens) is a word for 5-6 year olds, and *lepidoptera* (36 tokens), was successfully acquired by the adult author about a year ago. Across the four words *leg,*

[2] The surface frequencies were generated in a simple SARA search, via the online interface of the British National Corpus. The frequencies available in this simple search are non-lemmatised, and are generated from both spoken and written materials.

lemon, leopard and *lepidoptera* decreasing frequency is accompanied by increasing age of acquisition.

However, the correlation between frequency and age-of-acquisition does not indicate direction of causation. It is impossible to tell whether a) word frequency causes some items to be acquired in a particular order; b) age-of-acquisition causes some items to be used with higher frequency in adult language; or c) whether some other mediating factor (cognitive, social or linguistic) is responsible for both the age of acquisition and the frequency of a word's use in adult language. An investigation of the relationship between adult word frequency and infant comprehension rate could help to address these questions by providing information about whether a relationship exists between frequency and age-of-acquisition in the early stages of lexicon development (*leg, lemon*), or whether the relationship is driven by the large number of low frequency words acquired much later in life (*leopard*, *lepidoptera*). If no linear relationship exists between frequency and comprehension rate in the early lexicon, then it suggests that the frequency is not causally involved in the order of acquisition.

Two complications make direct investigation of frequency difficult in the current context. Firstly, word frequencies for infant directed speech have yet to be extensively documented. Where child-directed word frequencies do exist, they tend to be relevant to children of early school age, and include common primary school reading materials (e.g., Zeno *et al.*, 1995). Secondly, although corpora of infant- and toddler-directed speech do exist (e.g., MacWhinney, 2000), they are based on relatively small numbers of parent-child pairs. This limits the applicability of infant frequency data to other households. Given the limitations in the published literature, adult norms are considered here. It is worth noting, however, that pre-verbal infants overhear a large amount of language which is not directly addressed to them. Parents

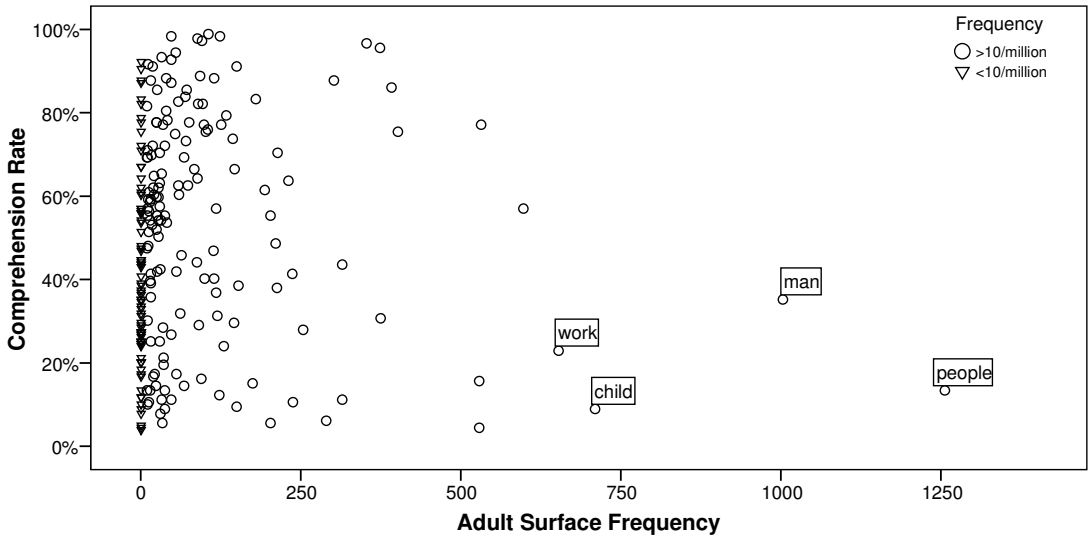
will interact with other adults in their presence, and infants will overhear a variety of speech registers through media sources such as radio and television. Thus, if only for exploratory investigation, comparison of infant lexical comprehension with a standard adult word frequency was conducted. For this comparison, only nouns, the largest grammatical class in the 18-month-old lexicon, were investigated.

Frequencies were taken from a list of common nouns with frequencies above 10 per million words in *Word Frequencies in Written and Spoken English, based on the British National Corpus* (Leech *et al.*, 2001). This particular list of frequencies was selected over better known word frequencies (Kucera & Francis, 1967; Thorndike & Lorge, 1944) for two reasons: Firstly, because the Leech norms relate to the variety of English spoken by parents of the target population; and secondly, because the frequencies represent *contemporary* word usage. 153 of the 232 nouns in the CDI were listed in the Leech frequencies, as words with a total frequency of over 10 tokens per million.

Figure 2.13 plots the comprehension rate of nouns in the O-CDI against their frequency of occurrence in the Leech norms from the British National Corpus (Leech et al., 2001). For illustrative purposes, words which fell below the reporting level of the Leech frequency norms are replaced with a dummy variable of 1 token per million. Along the X-axis, the frequency of nouns in adult British English shows the power law distribution described by Zipf's law (Zipf, 1935). The comprehension rate for nouns in the CDI is spread fairly evenly across the scale. There is also a moderate downward trend for words with frequencies of over five hundred tokens per ten million words: The four highest frequency nouns in adult English fell below 50% comprehension in the infant lexicon, and are labelled in Figure 2.13. However, note that where power-law distributions exist, an apparent downward trend in an un-

adjusted figure may be an artefact of the greater distance between items in the upper extreme of the scale.

Figure 2.13. O-CDI comprehension rate & adult British noun frequency in



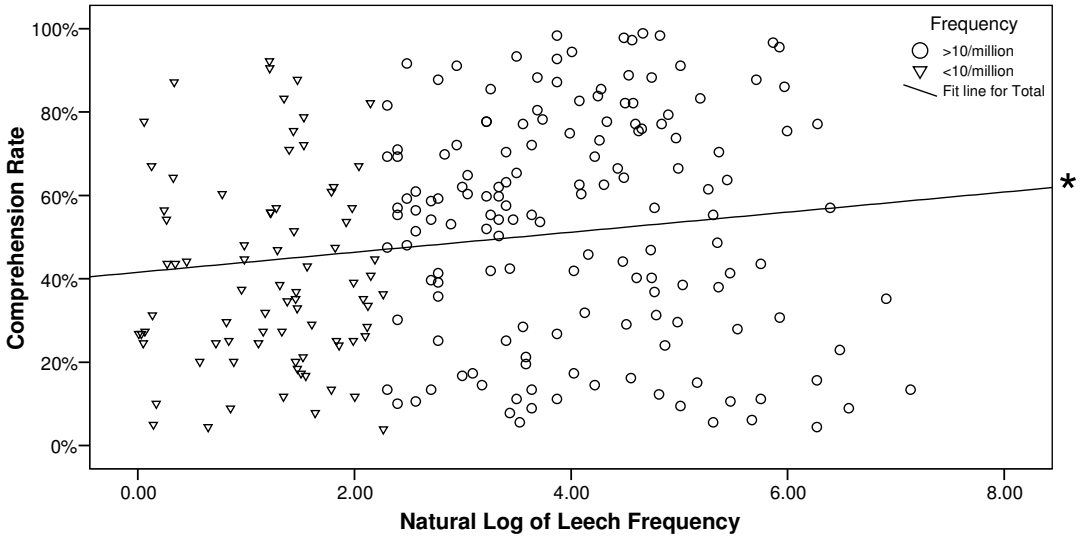
Notes. Nouns with frequencies below 10 per million words replaced with the dummy frequency of 1 tokens per million for illustrative purposes. Five highest frequency words are labelled.

To accommodate the non-Gaussian distribution, the natural log of the Leech frequency was compared to the O-CDI comprehension rate. When only those nouns reported with a frequency of over ten tokens per million were included in the analysis, no correlation was evident between comprehension rate and frequency in adult language ($\rho(153) = -.006$, n.s.). However, if the lower frequency words differ systematically in their comprehension rate from higher frequency words, the missing values could mask a legitimate linear relationship between variables. In fact, the median comprehension rate for words with a frequency above 10 per million was 57.0%, while the median rate for words with a frequency below 10 per million was 36.3%. This difference was significant ($U(232) = 4404$, $p = .001$), indicating that the dropout was systematic and needs to be accounted for.

To address the discrepancy, nouns with frequencies lower than 10 per million were allocated dummy values. Two dummy values were used, the value of 1 word per million (0 in the log transformed frequency variable), and a set of random numbers uniformly distributed from zero to the lowest log-transformed value. The random number procedure was conducted ten times. In Figure 2.14, comprehension rate is plotted against the natural log of Leech frequencies for all words, including one series of randomly generated dummy values. In this plot, an upward trend in the distribution is evident when both high and low frequency nouns are considered, showing that higher frequency nouns in adult language are those which have a higher comprehension rate in infant language.

The results of comparisons between the comprehension rate and the log transformed frequency values (with different dummy values) are given in Table 2.7. There was a significant correlation between adult noun frequency and comprehension rate in all comparisons.

Figure 2.14. O-CDI comprehension rate & log-transformed adult frequency



Notes. Dummy variables for nouns with frequencies below ten tokens per million were randomly generated values, uniformly spread across the range from zero to the lowest log transformed value. *p<.05

Table 2.7. Correlations of comprehension rate and log frequency, with different dummy variables

Dummy Variable	ρ	*N*	*p*
one	.173**	232	.008
random $_a$	.153**	232	.020
random $_b$	.168**	232	.010
random $_c$	.164**	232	.012
random $_d$	.161**	232	.014
random $_e$	.164**	232	.013
random $_f$	.165**	232	.012
random $_g$	.180**	232	.006
random $_h$	.156**	232	.017
random $_i$	.162**	232	.014
random $_j$	.154**	232	.019

Notes. Random dummy variables uniformly distributed between 0 and 2.29. **p<.01

This finding suggests that the relationship between frequency and acquisition (as indexed either by AoA or comprehension rate) exists even in early stages of lexicon development. Thus frequency in adult language could be causally involved in the acquisition of particular words. However, note that the current data do not rule out the other possible causal directions (AoA causes high frequency in later language production; or a third factor mediates both).

Clearly, this finding should be treated with caution, as the dummy values may not be representative of the actual frequency distribution for items with frequencies below the Leech cut-off. However, it is interesting to note that a relationship was evident even though the sample of adult language (the British National Corpus) was not infant directed, nor was it exclusively spoken language. Further research would be needed to clarify whether this is also true for other parts of speech (including the more

abstract, grammatical parts of speech), and whether an even tighter relationship would be evident if age-appropriate frequency counts were available.

Word associations between O-CDI words

As noted in the Introduction, one of the lexical variables known to produce priming effects in the adult lexicon is word association (WA). Once again, an adult WA norm has been selected for comparison in this study. The selection was made on the basis of three reasons: lack of appropriate infant / toddler association norms; debate about qualitative reorganisation of the childhood associative lexicon; and the relevance of adult norms in the infant linguistic environment, including caregiver speech. Each of these arguments is outlined in detail below.

Word association in early development

In a typical word association task, participants are asked to report the first word which comes to mind when presented with a given stimulus word. For example, given the stimulus *cat*, the majority of English participants will spontaneously come up with the word *dog* as a response. The task can be administered in written or oral form, which makes it attractive to researchers working with pre-literate populations. However, the task is not as simple as it appears – the participant performs a meta-linguistic task, by generating a word in a context devoid of meaningful linguistic interaction. It is questionable whether very young children are able to interpret fully, or to follow an instruction of this nature. Furthermore, when young children respond, it is difficult to interpret how much the demands of production (coupled with the artificial nature of the task) tax infants' cognitive resources. Thus any apparent

differences between the WAs of children and adults may be due to a) different task strategies, or b) different production demands. For example, a child who produces a 'follow on' association such as *kick – ball*, may be turning the experimenter's speech into an utterance which makes more socio-pragmatic sense to the child than the original stimulus did, and a child who produces a 'clang' association such as *blue – blow* may be relying on the phonological level of representation as a mean of lowering their processing demands. Both of these response types have traditionally been taken as hallmarks of the child association pattern. It is thus important to acknowledge that the demands of the task are different for children – the associations they produce *may not* reflect underlying cognitive relationships in the same manner as they do for adults.

In addition to this general concern about what child word associations index, a further complication exists for toddlers and infants. As word association is a production task, it is not possible to administer with pre-linguistic infants, or young toddlers with limited production skills. Thus no association norms exist for the ages of interest in this series of experiments (18-months to 30-months).

Childhood re-organisation: Qualitative of quantitative?

There has been debate in the word association literature about whether children's word associations are qualitatively different to adults' word associations, and, if so, what does this mean for the organisation of the developing lexicon. From the very first reports of children's word associations, differences have been reported between the WAs produced by children and adults. The differences have been characterised by Woodrow and Lowell (1916) as contiguity for children (such as the 'follow on' example, above), and semanticity (such as similarity or contrast) for

adults, who produced associations primarily from the same word class. Brown & Berko (1960) described a gradual change from heterogeneous to homogeneous (in terms of word class alone), while Ervin (1961) described the shift as syntagmatic to paradigmatic (with reference to how the different relationships were learned). Other researchers have attempted to identify semantic components involved in the shift (D. Entwisle, 1966; MacNeil, 1966), leading Francis (1972) to characterize the shift as wholesale semantic *reorganisation*. Other developmental researchers have tied word association norms to other aspects of semantic development, and characterise the change as thematic to taxonomic.

However, most studies describing changes between adult and child word associations focus on salient differences, without mentioning the overwhelming similarities. A cursory glance over the Woodward & Lowell norms (1916) reveals that many of the high-strength associations are constant throughout childhood and into adulthood, and a number of associations generated by each word are common throughout development. Even 'clang' associations are reported in adult norms (e.g., *dog - fog* from the Birkbeck WA norms (Moss & Older, 1996), albeit at a lower frequency. This suggests that differences in association norms over time are more quantitative than qualitative.

Furthermore, the tendency to provide 'child-like' or 'adult-like' word associations varies according to a variety of non-age-related factors. Some examples are listed below:

- Children generate more 'adult-like' associations if there is a high proportion of nouns in the test set, due to a general bias for giving noun responses (D. R. Entwisle *et al.*, 1964).

- Adults generate more 'childlike' associations in response to low frequency words (Deese, 1962), and also in oral administration (Palermo, 1971).
- High socio-economic status children generate more 'adult-like' associations in oral administration (D. Entwisle, 1966).
- French 'workmen' generate more 'childlike' associations than French male University students (Rosenzweig, 1970).
- Children and middle-aged adults generate 'childlike' responses, while only university undergraduates generate 'adult-like' responses (Smiley & Brown, 1979).

Furthermore, when models of semantic memory are considered (for review, see Nelson (1977)), both adults and children can demonstrate variable preferences for taxonomic or associations thematic, depending on the task (Bremner *et al.*, 2007; Lin & Murphy, 2001; Murphy, 2001; Waxman & Braun, 2005).

Word associations of the caregiver

Finally, it should be noted that caregiver word associations are likely to be highlighted in the presence of the language-learning infant. Whether an association is taxonomic (*cat - dog*), thematic (*brush – teeth*) or collocation-based (*all – day*), if it is salient for the caregiver, it may attract special attention in the caregiver's speech or behaviour. This point, taken together with the others, suggests that adult word associations may be equally appropriate as child word association norms for identifying potential relationships in the infant lexicon.

LEXICON NORMS (18M)

A collection of studies published together as the *Birkbeck Word Association Norms* (Moss & Older, 1996) was selected for comparison in this section. Altogether, 1,243 young adults participated in the studies, generating norms for 2,462 words. This collection of association norms was selected for two reasons. Firstly, it contains normed word associations for British English, the variety of English most relevant to the proposed test population. Secondly, as word associations norms from both adults and children have been demonstrated to vary from generation-to-generation (Jenkins & Russell, 1960; Koff, 1963), it was considered important to select contemporary norms. Note, for example, that the Postman and Keppel (1970) norms include associations generated by undergraduates 60 years ago (while these norms may perfectly align with the associations of contemporary grandparents, they no longer coincide with the associations of parents of infants).

Of the 416 words in the O-CDI, 247 words were used as probe words in the Birkbeck word association task (Moss & Older, 1996). Every probe returned at least one association *to another word listed in the O-CDI*. Table 2.8 summarises the number of words tested for association in the Birkbeck studies, according to word class, and the number of different associations directed to another O-CDI word. Nouns were the largest class of CDI words used as probes, followed by verbs. Given the apparently special status of the onomatopoeia words discussed in this chapter, it is unfortunate (if unsurprising) that no onomatopoeia words were used as probes in the Birkbeck studies. In Table 2.8, the word class which generated the highest number of forward associations to other words in the CDI was verbs, followed by nouns, adjectives and the mixed-class words, and finally then the group of function words.

Table 2.8. Number of word associations to and from O-CDI words in different word classes

	Noun	Verb	Adjective	Function	Mixed	Onom.	Total
words tested	133	56	25	28	5	0	247
forward inter-associations	953	466	180	160	36	-	1796
mean rate per word	7.2	8.3	7.2	5.7	7.2	-	7.2
response words	192	62	31	35	11	4	335
reverse inter-associations	1020	320	186	296	55	5	1822
mean rate per word	5.3	5.1	6.0	8.5	5.0	1	5.4

Notes. Discrepancy between total forward and reverse associations due to duplication where homonyms exist in the CDI (e.g., *glass+es (household object)*, *glasses (clothing)*), and where alternate words senses are listed separately (e.g., *fish (food), fish (animal)*).

Table 2.8 also lists the number of CDI words which were generated as responses to O-CDI probes, along with how many times reverse associations were generated, according to word class. The words attracting the highest rate of associations per response word were in the closed classes of function words (*out* generated seven forward O-CDI associations, but received 36 reverse associations from other O-CDI words; and *all* generated three forward associations, but received 33). Despite the fact that no onomatopoeia words were used are probes, some were generated in response to words from other word classes. This table illustrates that when responses are limited only to those words which are common in infant language, the remaining adult associations begin to demonstrate some of the features of word-associations generated by children. For example, many of these restricted adult associations cross grammatical word classes, many include 'follow-on' associations, or 'clang' associations, including onomatopoeia responses. Appendix D gives a list of inter-associations between nouns in the O-CDI.

Establishing an experimental stimulus set

One of the goals of this current chapter was to establish a set of words for use in future behavioural tasks with children as young as 18 months-of-age. Having established that the noun class is the largest in the reported lexicon at this age, and that a relatively large number of noun-noun inter-associations are available for experimental manipulation, all that remains it to clarify which nouns will be suitable for use in future studies. The proposed experimental method for the remainder of this thesis is the IPL method, which relies on the visual presentation of pictures. All nouns in the CDI were thus coded for whether their referents are concrete, generic, objects which were easily imagable and visually isolatable from their environments.

A second criterion was that items selected for test should be comprehended by the majority of infants invited to participate in a testing session. A cut-off comprehension rate of 50% was imposed to satisfy this condition. Table 2.9 gives the list of stimulus words which satisfy both of these conditions. Future stimulus items will be drawn from this pool.

Table 2.9. Stimulus set of concrete, generic nouns with lexical comprehension rates of above 50%

Animals		Food & Drink		Household Objects	
dog	98	banana	90	toothbrush	87
cat	94	biscuit	88	telephone	85
duck	91	milk	87	cup	79
bird	88	apple	77	spoon	79
cow	86	bread	72	key	73
bunny/rabbit	78	juice	71	brush	71
pig	78	toast	69	bin	69
horse	77	cheese	63	clock	65
fish	76	water	60	light	62
sheep	70	cake	55	bottle	62
lion	65	Outside		bowl	58
bee	61	flower	78	towel	56
chicken	60	tree	67	hoover/vacuum	55
monkey	57	swing	64	watch	54
frog	56	house	57	plate	51
butterfly	55	slide	57		
mouse	54			Vehicles	
elephant	54	Furniture & Rooms		car	96
bear	54	bath/bathtub	93	aeroplane/airplane	83
tiger	51	door	87	pushchair	77
		bed	83	bicycle/bike	70
Clothes		chair	81	bus	69
shoe	98	stairs	80	train	67
sock	92	TV	74	boat	63
nappy	92	cot	72	lorry/truck	62
hat	88	high chair	67		
coat	78	table	64		
bib	75	window	61		
trousers	56	refrigerator/fridge	56		
boots	54				
pyjamas	54				

General discussion

The aims of this chapter were to establish a set of norms relevant to the proposed test population, as a means of understanding the scope of information contained in the infant lexicon, and to provide standardised norms (lexical comprehension, word association) to aid in future experimental design. The method chosen was a detailed investigation of a large number of CDIs, within a relatively restricted age-band. Previously established norms concerning the size of the British lexicon at 18-months were replicated, with only a small degree of variation. These norms also follow patterns previously reported in the literature, including more words in comprehension than in production (Benedict, 1979; Fenson et al., 1994; Goldin-Meadow et al., 1976; Snyder et al., 1981), and lower reports of comprehension for British infants than for North American infants (Hamilton et al., 2000). The composition of the lexicon was also investigated in detail, according to the word groupings of the O-CDI, and to grammatical word classes. The comprehension rates generated in this chapter are, to the best of my knowledge, the only comprehension norms for children above 16-months-of-age, as well as the only comprehension rates for early words in British English.

Comprehension rates for individual nouns correlated with both age-of-acquisition and adult frequency in British English. These findings suggest that comprehension rate may be an appropriate experimental control in future studies, where age-of-acquisition data are not available, and representative infant-directed frequencies have yet to be established.

As this chapter is based on analysis of comprehension and production data collected by CDI, it is worth addressing the advantages and limitations of data

collected in this manner. On one hand, checklists allow detailed information to be collected from a much larger population that would be possible with experimental or observational techniques (Fenson et al., 1994). However, questions have been raised about the veracity of the checklist method, and in particular, about asking parents to report on lexical comprehension. Stiles (1994) points out that while lexical production is an observable behaviour, comprehension must be *inferred*, while Tomasello and Mervis (1994), suggest that comprehension is likely to be overestimated. From a different angle, Feldman et al. (2000) are concerned by the wide variability evident in large populations of CDIs, and by a tendency for people from lower socio-economic backgrounds to generate higher CDI reports (Feldman et al., 2000; Fenson et al., 1994). In addition, Robinson and Mervis (1999) have demonstrated that the standardised checklist only records *a sample* of the possible words that infants and toddlers are learning, and thus systematically under-reports total vocabulary size.

With regard to variability in CDI reports, Fenson, Bates, Dale, Goodman, Reznick and Thal (2000) convincingly argue that infants, not CDIs, are responsible for the observed variability: There is large variability between the linguistic development of different children, and CDIs reflect this. One advantage of large-scale norming studies is that they provide a good indication of *how much* variation is typical in a given sample population. Furthermore, the reliable correspondences between multiple administrations of the CDI (J Steven Reznick & Goldfield, 1994) and comparing CDI scores to the results of other tools (e.g., Bretherton et al., 1983) demonstrate that CDIs are a valuable source of information about the linguistic development.

As far as sampling-based underestimation is concerned, Robinson and Mervis (1999) conclude that vocabulary scores should be treated as representative, not

comprehensive. In this respect, the advantage of using a standardised tool, is that all participants will be subject to the same limitations (i.e., maximum vocabulary of 416 words, maximum noun-score of 232 words). CDIs must be expected to produce somewhat asymmetrical sampling between individuals. However the norming study carried out here suggests that the word classes of the CDI could be used to provide greater sensitivity in the tails of the distribution (e.g., assessing nouns and verbs for low vocabulary infants, and assessing function words for high vocabulary infants). The final, and perhaps most theoretically contentious issue – whether parents can accurately infer comprehension from their infants' behaviour – will be addressed in the following chapter.

One final limitation in the current O-CDI tool remains to be discussed. Where comprehension was plotted against production (Figure 2.03, Figure 2.08, Figure 2.09), it became clear that no words are reported as produced before they are comprehended. While this observation seems intuitively logical, it mirrors a constraint imposed by the design of the CDI. With only two response categories (comprehend, or comprehend-and-produce), this design could inhibit valid parental responses. For example, at the end of an experimental session with the author, a two-year-old was choosing from a range of coloured t-shirts to take home as a gift. She pointed to a red shirt and said "Green!" Her mother commented that she frequently uses colour words (and appeared know understand the property of visual contrastive to which they applied), but that she did not known *which* word referred to *which* colour – a clear case of accuracy in production preceding accuracy in comprehension. According to the response categories of the CDI forms (comprehend, or comprehend-and-produce) the colour words were omitted from her lexicon, as she was not judged by her mother to 'understand' properly.

This is likely to be a common confound in the domain of colour, but perhaps also in other less-concrete domains such as adjectives (e.g., *hot*), pronouns (e.g., *I, he, her*) and prepositions (e.g., *on*, *in*). Furthermore, it is interesting to note that zealous overextension of a more concrete word such as *lorry*, is often marked as 'understanding'. As one parent mentioned to the author, "Everything's a 'lorry' this week!" And yet, for more abstract words, inaccurate mapping can be judged as failure to comprehend.

Bloom (2002) reviews a similar abstract learning phenomenon concerning the acquisition of early number words: Low non-one numbers may be produced and partially understood for some time, before they can be mapped accurately to the correct number of objects. Their nature as descriptive words is understood, as well as their use for contrasting different *quantities* of collected objects, yet the precise numeracy of each word takes longer to establish. Bloom's account highlights the gradual nature of word learning. It raises the related question – when is a word understood *well enough* to be considered as "understood"? – another issue which will be addressed in the following chapter.

One final point of interest arising form the current norming study was the special status of onomatopoeia in the 18-month-old lexicon. This group of 'animal sounds' (which included 'ouch' 'uh oh' and 'yum'), was the nearest to the linear maximum for comprehension-production correspondence, which indicated that they were reported to be produced by almost the same number of infants as they were comprehended by. Their rate of comprehension was also high (with every word known by 50% to 88% of infants), and despite not being used as adult word association probes, they received backward association *even in the adult WA norms*. Although space and time do not permit an exploration of onomatopoeia in the current

series of studies, it is worth highlighting their potential value in future experimental work with infants in this age range.

Interim conclusions

Lexical comprehension norms for a specific population were established from a large sample of parental checklists. The 179 CDIs comprise the largest reported sample of 18-month-olds. The norms for infants and for words were given external validation by their close correspondence to previously reported norms, which included comprehension rates in North American norms, and age-of-acquisition norms for a large number of nouns. This finding adds further evidence that parental report is a valuable indicator of language performance. Furthermore, given the strength of the relationship between comprehension rate and other normed variables known to affect lexical access in adults, comprehension rate is expected to be a useful surrogate variable in the absence of comprehensive age-of-acquisition norms or infant-directed frequency counts.

The predominance of nouns in the 18-month-old lexicon makes this word class the most fertile for systematic investigation of lexical comprehension in this age group. Nouns were also the word class for which the largest number of inter-associations were attested. The experiments in the remainder of this thesis will focus exclusively on nouns. Finally, the largest number of attested inter-associations occurred between taxonomic category co-ordinates. Although this may well be due to a typical taxonomic bias in the undergraduate population who provided the adult norms, it is reasonable to expect caregivers of future participants to highlight similar

associations in the presence of their children. In the priming manipulations which follow, all associated word pairs will also be taxonomic co-ordinates.

Chapter 3

Minimally Primed IPL



This chapter describes the first in a series of experiments designed to incorporate a priming phase into an intermodal preferential looking (IPL) design, as a means of investigating organisation in the infant lexicon – that is to say, when and whether relationships are evident between words in the infant lexicon. The experiment is a conservative adaptation of a standard IPL task, designed firstly to assess whether infants can show lexical comprehension when a (potentially confusing) 'priming phase' precedes a standard lexical comprehension task, and secondly, to assess whether priming effects are evident in this weakly primed scenario. An additional feature of this experiment is that it allows investigation of the item level accuracy of parental report of comprehension, by assessing whether parental CDI reports of 'understanding' accurately predict the words for which infants demonstrate comprehension.

Introduction

The intermodal preferential looking (IPL) task established by Golinkoff et al. (1987) is a free-looking task for infants. IPL is a valuable experimental method which taps into a motor system which infants are able control very early in development. Fantz (1963) demonstrated that when visual stimuli were presented one-at-a-time, new-born infants selectively fixated some stimuli longer than others, even within the first two days following birth. His experiment not only showed that very young infants have control of gaze duration, but that differences in the timing of fixations could be used to infer *discrimination* between different types of visual stimuli. Spelke (1976) brought the interpretation of gaze information to cross-modal events, demonstrating that when two videos were presented side-by-side (a woman playing

peekaboo or a woman playing a percussive instrument), infants looked significantly longer at the video which matched the accompanying soundtrack. Thomas, Campos, Shucard, Ramsay & Shucard (1981) adapted the cross modal approach to the comprehension of familiar words, assessing eye-movements towards an array of four objects, one of which was named. Following these developments, the inter-modal preferential looking (IPL) procedure was formalised by Golinkoff et al. (1987). Pairs of pictures are presented on a screen, and an auditory instruction directs infants to look at one of the pictures, by giving its name. Infants' eye movements are monitored, and the duration of gaze is used to assess infants' ability to pick out a named 'target' from an unnamed 'distracter'. In this procedure, infants can sit comfortably with a familiar caregiver, while the experimenter and any other confederates are concealed from view – thus eliminating any possible 'clever Hans' effects which might be caused by the experimenter inadvertently cuing the infant. Reznick (1990) incorporated a 'baseline' free-looking period for each infant, so that a-priori picture preferences could be quantified. He demonstrated that infants systematically increase their looking to named pictures following the onset of a familiar label.

Subsequent studies (e.g., Fernald *et al.*, 1998; Schafer, 1998; Schafer & Plunkett, 1998; Schafer *et al.*, 1999; Swingley *et al.*, 1999) have incorporated the visual baseline into the body of the IPL trial, by presenting the picture-pair before the onset of the target word. In these contexts, it is also referred to as a pre-naming phase (e.g., Houston-Price *et al.*, 2007; Mani & Plunkett, 2007; Meints et al., 1999), as relative preference before and after naming can be compared. Within-trial 'baselining' offers two statistical advantages over the original trial design. Firstly, if target and distracter looking are similar prior to the onset of naming, above chance preference for the target following the onset of naming must be due to the effect of the name.

Alternatively, if pre-naming values deviate from chance, but a *change* in behaviour is observed following naming, the change provides evidence of the effect of the name.

Since the development of the IPL method, significant advances have been made to the accuracy of coding methods, with offline frame-by-frame coding of video data allowing greater temporal accuracy in fixation data. An additional advantage of offline coding schemes is that they allow genuinely 'blind' coding of behavioural states, by removing coding-biases which could be generated when auditory stimuli are overheard. A number of studies (Bailey & Plunkett, 2002; Houston-Price *et al.*, 2005; Mani *et al.*, 2008; Meints et al., 1999; Meints *et al.*, 2002) employ a 'standard' timeline of approximately five seconds of paired picture presentation, with the target word beginning half way through picture presentation. The standard timeline provides equal pre- and post-naming periods for comparison.

This chapter reports a conservative adaptation the IPL task – a five-second timeline preceded by a short auditory priming phrase. Three types of prime were included in the current study: a neutral prime, which was unrelated to the target; a taxonomic prime, which was a basic-level sister of the target, and a compound prime which was related to the target by association as well as taxonomy. The first half of the chapter demonstrates that parental report of comprehension is sufficiently accurate to pick out those words for which infants will be able to demonstrate lexical comprehension in this adaptation of the a IPL task. The second half of the chapter builds on this finding, by assessing the three priming conditions in detail.

In McNamara's (2005) comprehensive review of the behavioural priming literature, he summarises fundamental principles of stimulus arrangement in lexical priming designs: a) to isolate the effect of the prime independent of the target, each target should appear in each priming condition; b) to avoid strategic effects, memory

effects or short-term dependencies, repetition of items should be avoided unless separated by long periods of time; c) counterbalancing of target conditions can be achieved between stimulus lists for different participants, providing that list is included as a factor in analysis; and d) primes should also be counterbalanced across lists where possible. As infants in the age range of interest (18 to 24 months-of-age) are known to have relatively short attention spans, a long experimental study was considered inappropriate. In order to follow the principle of non-repetition, *primed* IPL tasks will therefore require a 'one-shot' design, in which no stimulus is encountered on more than one occasion. This differs from standard IPL practice, where repetition of items is often considered necessary in order to counterbalance labelling across both items within a picture pair (e.g., Houston-Price et al., 2007), or to 'normalise' the data from an individual infant across multiple presentations of the same stimulus (e.g., Fernald et al., 2001). In a short experimental session, repetition would confound *existing* relationships in the lexicon, with possible *associative learning* generated during the test session. In a 'one-shot' IPL design, there is no confound with learning, as counterbalancing can be achieved across stimulus lists.

A 'one-shot' design necessarily includes a large number of different stimuli – a feature which has advantages and disadvantages. Compare, for example, the four targets (*doggie, doll, tree,* and *truck*) of Swingley and Fernald's (1999) IPL study, in which each item was presented four times, to the 36 discrete priming relationships included in the current experiment. The former study provides evidence about differences in the speed of processing particular words such as *tree*. In the current study, by contrast, a pattern of responses generated by a diverse set of primes is statistically *more likely* to be applicable to novel tokens of the same type; the findings are more generalisable. On the other hand, a disadvantage of using a large number of

stimuli is that words with lower comprehension rates will need to be included to complete the stimulus-set, and individual differences between which items are understood can be predicted to generate experimental noise.

A final advantage of the 'one-shot' design, is that it allows visual attention to the target and distracter to be assessed separately. This assessment is important for investigations of which items individual infants comprehend. If an infant does not understand the target word, but knows the name of the distracter, she might be able to infer the correct referent by rejecting the name-known object – a process known as mutual exclusivity (Markman, 1989; Merriman & Bowman, 1989) or N3C (novel-name nameless-category: Golinkoff *et al.*, 1994). Single presentations of picture pairs allows investigation of such processes, as the lexical status (known, unknown) and the naming condition (named, unnamed) can be assessed independently.

Before detailed investigation of the priming conditions, the 'one-shot' design will be used to address two important methodological questions:

1. Can infants demonstrate typical lexical comprehension performance (looking more and longer at labelled pictures following naming), when the IPL task follows a potentially confusing priming phase?

2. Are parental reports of comprehension informative about the words infants know (and can respond to) in this potentially confusing task?

Concerning the issue of lexical comprehension, given that the priming scenario requires the infant to sit in semi-darkness while the priming phrase is presented, it is important to ascertain whether infants will be able to maintain

attention during picture presentation, following a period with no visual stimulation. The priming phrase also mentions an absent object, which is not depicted throughout the experiment. It is thus unclear whether infants will be able to show their normal ability to discriminate the named target from the unnamed distracter following the potentially distracting auditory non-sequitur.

Concerning the issue of parental report, it is important to establish whether individual prime and target words are known to the infant at the time of test, so as to establish whether or not they exhibit evidence of a relationship *between* pairs of words. From a theoretical perspective, investigating a possible relationship between two items is problematic if one of the items is not yet represented in the lexicon. From an experimental perspective, the inclusion of unknown words would create a bias toward unrelated prime types, and would be likely to mask genuine priming effects. Eliminating this confound is particularly relevant for infants with small vocabularies, for whom the bias would be greatest. If parental report can be treated as a reliable indication of which words are known to the infant at the time of test, it can be used as an experimental control in future priming studies.

The first part of the chapter will address these issues by assessing whether parental report is sufficiently accurate to identify those trials in which infants are able to demonstrate comprehension in a (potentially confusing) primed IPL task. This section involves only 18-month-old infants, as infants in this age group had a mix of test items reported as known and unknown, thus allowing a relatively balanced sample for analysis of parental report. The second part of the chapter will address the behaviour of both 18-month-olds and 24-month-olds according to the priming conditions of the minimally-primed IPL task.

Method

Participants were recruited from a database of parents who had previously expressed an interest in participating in developmental studies. All families who visited the laboratory were given a small gift, or a contribution towards their travel expenses, regardless of whether their infant participated in, or completed the task. Infants were allocated to a stimulus list on arrival. Forty 18-month-olds were tested (24 males, 17 females; mean age: 18.0 months, range: 17.6 to 18.7). Five 18-month-olds were removed without replacement (three for parental failure to return the completed CDI form, one for parental interference during test, and one for equipment malfunction). One infant was removed and replaced at the time of test, for failure to complete. Thirty 24-month-olds were tested (14 males, 16 females; mean age: 24.1 months, range: 23.3 to 24.6). Five infants were removed without replacement following test (two for interference from a family member during test, one for data corruption, one for experimenter error, and one for equipment failure). Following exclusions, there were 35 18-month-olds (20 male, 15 female) and 25 24-month-olds (10 male, 15 female) for whom both behavioural data and vocabulary information were available.

Lexical pre-test

In the week before visiting the laboratory, primary caregivers of all subjects filled out the 416-word Oxford CDI (Hamilton et al., 2000), introduced in the previous chapter. Parents brought their completed CDIs with them to their testing session. Words marked as either ‘understood’ or ‘understood and also said’ were coded as ‘known’. The CDI pre-test was used to establish which stimulus items were

reported as known to each individual. For 18-month-olds, a total comprehension score was also calculated out of the 416 words in the checklist.

Materials

42 words were selected for use in the current study. 12 'target' words labelled a picture while it appeared on screen, 12 'distracter' words had referents which remained unnamed while they appeared alongside a target, and 18 words were used as auditory 'primes', which were presented prior to the appearance of the pictures. All words used in the current experiment were selected from the stimulus pool established in the previous chapter (Table 1.7), in which all words had an 18-month comprehension rate of above 50% and were basic-level object names, with referents which were easily imageable.

For visual stimuli, colour photographs were chosen as the most appropriate images for the following reasons: a) Fantz (1963) demonstrated that newborn infants look longer at visually complex stimuli than at visually simple stimuli, a finding which suggests that simplified line drawings (such as clip art) would be less perceptually salient for infants than richly featured images; b) Callaghan (2000) showed that infants are better at matching real objects to realistic illustrations of objects (both colour paintings, and detailed, shaded, black-and-white drawings), than to schematic line drawings, suggesting that realistic pictures act as better representations of stimulus items than line drawings; and c) Simcock and Deloache (2006) demonstrated that toddlers are better at re-enacting actions from photographic story books than from a re-drawn pencil version of the same book, suggesting that photographic materials offer the best representations of real objects.

Visual stimuli were thus created from high quality digital photographs of animals and objects. Photographs were photo-touched and edited by an experienced graphic designer to correct colour, contrast, and lighting of images; to mask and repair irrelevant details (e.g., a horse's bridle, long grass covering a bear's feet); and to remove images from their original backgrounds. All pictures were presented on a plain 5% grey background. Photographs depicted animals and vehicles in canonical profile view and other objects from the perspective of greatest clarity. Photographs of some artefacts were colour-rotated and enhanced to generate bright contrasting colours within a picture pair. Where it was possible to achieve a natural effect, target colours for artefacts were bright red, blue, yellow and green.

Yoked picture-pairs were made up of two images from the same semantic category (animals, vehicles, food, clothing, tableware, furniture), which shared no attested word association in adult language (Moss & Older, 1996), and which did not share a phonological onset or rhyme. The 12 target-distracter pairs were: *horse-lion*, *cow-chicken*, *dog-monkey*, *mouse-bear*, *banana-toast*, *cake-bread*, *bus-lorry*, *bicycle-pushchair*, *coat-sock*, *shoe-pyjamas*, *cup-bowl*, and *bed-highchair*. Pictures are given in Appendix H.

For audio stimuli, a single digital recording session was conducted in a sound-attenuating booth on DAT tape sampling at 44.1 kHz. A minimum of three tokens of each auditory stimulus were produced by a female native speaker of British English, using high-affect child directed speech. The single best token of each stimulus was manually selected for clarity, typicality and affect, and edited to remove head and tail clicks, and to align the onset of the target word to 2500ms after the start of the trial.

As infants at 18 months-of-age are faster at understanding words in sentences than in isolation (Fernald & Hurtado, 2006), recordings were made with primes and

targets occurring at the end of a short sentence. The prime was the final word in a simple sentence (e.g., *Yesterday, I saw a cat!*), in which it was preceded by a semantically appropriate verb, common in overheard speech (lemmatised adult frequencies of above 144 tokens per million words according to Leech et al., 2001). The prime phrases and their durations are given in Table 3.1. Prime phrases were right-aligned to the end of a 3000ms priming phase at the beginning of the trial.

Table 3.1. Duration of prime phrases

Prime phrase	Duration (ms)
Yesterday, I saw an aeroplane	2313
Yesterday, I ate an apple	1868
Yesterday, I ate a biscuit	2183
Yesterday, I found a boot	2284
Yesterday, I saw a car	2107
Yesterday, I saw a cat	2262
Yesterday, I ate cheese	2279
Yesterday, I bought a cot	2235
Yesterday, I saw an elephant	2301
Yesterday, I found a hat	1894
Yesterday, I saw a pig	1846
Yesterday, I bought a plate	2205
Yesterday, I saw a rabbit	2126
Yesterday, I saw a sheep	2009
Yesterday, I bought a spoon	2397
Yesterday, I saw a table	1804
Yesterday, I saw a train	2275
Yesterday, I found some trousers	2432
M (*SD*)	2157 (194)

Notes. Onsets and offsets of tokens were established using visual waveform editing software.

To maintain infants' attention during the experiment, four different 'carrier' phrases were used to introduce the target words: 'Look at the…!', 'Which one's the…?', 'Can you see the…?' and 'Where's the…?'. The phrases were recorded in

their entirety, to provide natural prosodic contour in each target phrase. The durations of the target phrases are given in Table 3.2.

Table 3.2. Duration of target phrases

	Carrier				
Target	Look …!	Which …?	Can you…?	Where's …?	*M (SD)*
banana	1355	1701	1214	1415	1421 (205)
bed	1314	1213	1312	1314	1288 (50)
bicycle	1535	1439	1676	1375	1506 (130)
bus	1262	1437	1414	1506	1405 (103)
cake	1385	1501	1291	1246	1356 (113)
coat	1276	1563	1494	1329	1416 (135)
cow	1493	1312	1259	1454	1380 (112)
cup	1151	1527	1339	1022	1260 (221)
dog	1283	1406	1187	1051	1232 (150)
horse	1417	1460	1464	1404	1436 (30)
mouse	1392	1577	1680	1324	1493 (164)
shoe	1283	1389	1453	1308	1358 (78)
M (SD)	1346 (107)	1460 (128)	1399 (163)	1312 (147)	1379 (65)

Notes. Onsets and offsets of tokens were established using visual waveform editing software.

Procedure

After a few minutes of 'settling in', infants sat on a caregivers' lap facing a rear-projection screen in an IPL booth. Parents were asked to wear headphones and to close their eyes during the procedure, which lasted approximately one and a half minutes. For 3000ms, the screen remained dark while the priming phrase began. At the offset of the prime, a picture-pair appeared on-screen for 5000ms, during which the target phrase began (e.g., *Look at the dog!*). The target word was timed to begin 2500ms into the picture presentation period of the trial, creating a pre- and a post-naming phase of equal length. A schematic of the trial timing is given in Figure 3.01.

The experiment was 'infant timed,' in that each trial was manually initiated when the infant was attending to the screen. All infants saw the same 12 picture pairs, in which no item was repeated. Trial order was randomised on presentation, and target side was counterbalanced across stimulus lists. Infants sat approximately 90cm from a screen with a display area 790mm wide. Images were 320mm wide, together occupying a visual angle of 48º and separated by a gap of 155mm (10º). A photograph of the booth and a diagram of the screen configuration are given in Appendix E (Booth 1).

Figure 3.01. Schematic of trial timing



Origin (0ms) indicates the onset of the test-phase of the trial, which is equally divided into a pre-naming and a post-naming phase. Wave-forms from actual stimulus tokens, as produced by GoldWave visual waveform editing software (1995, v5.10).

Prime conditions

Three stimulus lists were created in which 12 yoked picture-pairs occurred, with the same picture named as the target in each list. Each target occurred with three different primes: a neutral, or unrelated prime (e.g., *cot – dog*); a taxonomic prime which was related by taxonomy alone (e.g., *elephant – dog*); and a compound prime, related by taxonomy and by association (e.g., *cat – dog*). The prime conditions are illustrated in Figure 3.02. Distracters shared no phonological or associative relationship with the prime nor with the target in their yoked pair. All picture pairs were taxonomic 'sisters' from the same super-ordinate semantic category.

Figure 3.02. A yoked picture-pair in three priming conditions

Unrelated Prime	Taxonomic Prime	Compound Prime
Yesterday I bought a cot!	Yesterday I saw an elephant!	Yesterday I saw a cat!
		
Look at the dog!	Look at the dog!	Look at the dog!
(Assoc Strength = 0)	(Assoc Strength = 0)	(Assoc Strength = 67%)

Notes. Association strength from Moss and Older (2000), c.f., Appendix D.

In this priming scenario, it is possible that the prime will affect picture looking before the target word is named. If this is the case, the prime would be expected to generate similar pre-naming interest in the target and the distracter in two of the three priming conditions: the unrelated condition, where neither picture is related to the prime; and the taxonomic condition, where both pictures are taxonomically related to the prime. In the compound prime condition on the other hand, both pictures are taxonomically related to the prime, but the prime also shares the additional relationship of association with the target. In this condition, the prime might be expected to generate greater pre-naming interest in the target. If, on the other hand, the prime influences picture looking *after* the onset of the target word, then the two related conditions would be expected to differ from the unrelated condition. Furthermore, if infants show graded sensitivity to the two related prime types, with a larger priming effect in the compound prime condition, this would mirror the adult 'associative boost' demonstrated by Moss, Ostrin, Tyler & Marslen-Wilson (1995).

All infants saw an equal number of trials in each of the three priming conditions, and across infants, all targets occurred in all the three priming conditions. Each stimulus list was duplicated to present the same stimulus combinations on the

opposite side of the screen. Target carriers were counterbalanced across the six stimulus sets according to a Latin Square order. Trial order was randomised on presentation. The stimulus lists are given in Appendix G.

Scoring & Measures

Infants' eye movements were monitored by small cameras located above the two image presentation areas. Recordings were digitally captured as a split-screen image during test. Blind manual coding was conducted offline frame-by-frame at a temporal accuracy of 40ms using Noldus Observer software. All coding was conducted by an experienced coder (previously assessed inter-coder reliability of above 95% correspondence: $r(48) = .97$, $p < .0005$). Looks to left and right were automatically re-combined with trial information to generate fixation data in terms of target and distracter. Looks to the target and distracter were analysed separately for the pre- and post-naming phases of the trial.

Both large- and small-scale timing measures were calculated: two macro-level measures assessed the relative preference for the named target picture during the course of the whole trial, and a micro-level measurement of reaction time (RT) assessed the speed of infant eye-movements following the onset of the target word.

Macro-measures: The proportion of target looking (PTL) is the total amount of time spent looking at the target (T) as a proportion of the total amount of time spent looking at both pictures (T+D). It can be represented as T/(T+D). PTL represents relative picture interest over the whole picture presentation period, excluding time

spent switching between pictures, blinking, or looking away from the screen area.[3] The difference in longest look (LLK) is the difference, in milliseconds, between the single longest look to the target (t) and the single longest look to the distracter (d). It can be represented as t-d. As highlighted by Schafer (1998), both measures are relative assessments of target looking, which take into account the *dependent* nature of looking behaviour in a dual display (i.e., any time spent looking at one picture cannot be spent looking at the other). However, the dependency between target and distracter looking is not absolute, as both measures discount time spent blinking, shifting or looking away from the screen. In addition the difference in longest look compares only two discrete looks.

The 'linking hypothesis' for the macro-level measures is this: If infants distribute their fixations randomly between pictures for the duration of the trial, then relative measures assessing looking over the test phase would 'even out' across trials, creating means of similar value for targets and distracters. However, if the name of the picture induces a systematic visual preference for named targets over unnamed distracters, then the pattern of behaviour would be consistent with infants mapping spoken words to target pictures. As argued by Aslin (2007), however, the duration of accumulated fixations is difficult to interpret (does more looking imply continuous interest, effortful processing, or blank staring?). For this reason, the micro-structure of the trial is also valuable, as it provides information about *rapid responses* to suddenly appearing stimuli and about the processing load which particular stimulus combinations may create.

[3] A 'fixation' is defined as 2-or-more consecutive frames ($\geq$ 80ms) in which the eye is still. A 'look' is defined as the total duration of contiguous fixations to one picture area. Note that onset of a look is taken from the time of fixation, not the onset of the saccade preceding fixation.

Micro-measures: Latency is a reaction time measure which assesses the time taken to shift from the currently fixated picture to the alternative picture. In the current experiment, latency is measured from the onset of the target word to the subsequent fixation. The linking hypothesis for a latency measure is that linguistic processing leads to a 'match' or a 'mismatch' between the spoken word and the currently fixated object. Infants typically fixate an alternative picture faster when there is a 'mismatch' between the label and the object than when there is a 'match' (Swingley & Fernald, 2002), and differences in the speed of distracter rejections are understood to index the speed of lexical processing as the spoken word unfolds (Fernald et al., 1998; Swingley et al., 1999).

3.1. Lexical comprehension and parental report

This section of analysis focuses on whether infants are able to demonstrate lexical comprehension in the modified (primed) version of a standard IPL task, and whether their ability to demonstrate comprehension matches parental report of which words are 'understood' by each infant. While addressing a methodological question expected to be valuable for future priming research, this analysis also addresses a long-standing issue in the literature concerning parental reports of language comprehension. As noted in the previous chapter, some researchers have questioned the validity of parental reports of comprehension, given that comprehension can only be inferred from behaviour (Stiles, 1994), a problem that is compounded by the lack of clear definitions of what *level of understanding* should be deemed 'comprehension' (Tomasello & Mervis, 1994). There is also little evidence currently available about the level at which parents interpret 'understanding' when filling out a CDI form. To

address the questions of reporting validity and reporting level, external evidence is needed of whether words reported as 'understood' elicit behaviour consistent with lexical comprehension, while words reported as 'not understood' fail to produce this behaviour. Inter-modal preferential looking (IPL) provides a platform for testing these questions in a controlled environment where contextual cues (such as daily routine), and social cues (including joint attention and body language) cannot be used to disambiguate the meaning of a word. To date, however, there has been little systematic investigation of IPL performance at the item level, or whether parental report and IPL agree.

Houston-Price, Mather & Sakkalou (2007) report two IPL experiments in which parental CDI report *did not* predict infants' performance in identifying a named 'target' image. In Houston-Price et al.'s experiments, parents of infants at 1;3, 1;6, and 1;9 filled out two questionnaires (a vocabulary inventory and an object familiarity survey) in the fortnight preceding an experimental session. Experimental stimuli were then tailored to individual infants' lexicons. Items were selected for test if they had been judged as 'familiar' to the infants (encountered more than once a week), and from these items, equal numbers of name-known image pairs and name-unknown image pairs where created. Pairs were presented two or four times, with each image acting equally often as a named target, and as an unnamed distracter. Their results showed an increase in target looking following the onset of naming, a finding consistent with lexical comprehension. However, they found *no difference* between responses to words that were pre-judged to be name-known or name-unknown.

The authors gave three possible reasons for their finding: a) parental failure to observe overt instances of comprehension, such as pointing, b) parental inability to monitor more subtle signs of comprehension, such as eye-gaze, and c) a possible

mismatch in the threshold of 'comprehension' assessed in the laboratory task and in parental judgement. As previous research has shown that the process of refining word meanings continues well after a word is reported as understood (Meints et al., 1999; Meints et al., 2002), the third possibility (mismatch between task and report) is considered in more detail. To give an example of mismatch between CDI threshold and IPL task, an infant who is familiar with socks might have a weak representation of the word, which only includes contextual information; associating the word with an indoor scenario, with a particular room in the house, with getting dressed, or perhaps even with feet. Prior to test, her parents may have judged 'sock' to be 'unknown', as she cannot ordinarily pick out a sock from her trousers. And yet, given the choice between a sock and a lorry, this infant may be able to use her contextual knowledge to infer the correct referent. That is to say, knowing that 'sock' is a household word may be sufficient to correctly infer that the sock is a more likely referent than the lorry. Such a case would represent a mismatch between the level of word knowledge assessed by the parent (sock-trousers), and the difficulty of the particular task (sock-lorry). It may be possible to tease apart the question of inaccuracy or comprehension 'level', in the previously reported literature by introducing a more difficult lexical comprehension task.

Compared to the Houston-Price et al. experiment, the current experiment assesses infants of the same age, but involves a more difficult task. The 'one shot' design means that that every picture is unfamiliar on every trial, and infants have only one chance to demonstrate comprehension. In addition, all picture pairs are from the same semantic category, making it more difficult for an infant to pick out the referent of a weakly-known word (such as 'sock' in the example above). While both the 'one-shot' design and the taxonomic pairs were originally adopted for manipulation of

priming conditions, these features also provide a valuable comparison in terms of task difficulty. In addition, the priming phase might also make the task more difficult for infants in their second year, as hearing the name of an absent object might be predicted to confuse infants if a picture of the named object is not subsequently displayed. Data were pooled across priming conditions for the analysis of lexical status (known, unknown) in the current section. The analysis of parental report was only conducted for infants in the 18-month age group, as there was a mix of stimulus words reported as known and unknown for this age group. 24-month-olds were reported to know a much larger proportion of the stimulus items, and are not included.

Results

Following test, trials were categorized according to parental report of target status (known, unknown) and distracter status (known, unknown), resulting in four trial types: target known, distracter known; target known, distracter unknown; target unknown, distracter known; target unknown, distracter unknown. As the identification of trial types was a post hoc assessment, there was a possibility of asymmetrical distribution between the trial types. The data were therefore analysed twice: In the 'broad' analysis, data from all 35 participants were included. Pre-tests confirmed a distributional bias in the broad analysis towards knowing the target (fifteen infants, for example, knew more than 10 of the 12 target words, and of them, six infants knew all 12 words). To account for the bias, in a more stringent, 'narrow analysis', the sample was restricted to those infants whose trials were spread across all four trial types. 16 participants fulfilled this criterion. The same pattern of results was obtained in both the broad and the narrow analyses, both of which are given in full.

Broad Analysis (N = 35)

Of the twelve target words, infants were reported to understand a mean of nine words, ranging from three words for a single infant, to all twelve words, known by six infants. Thus most infants (19) knew more than half of the words used to label target pictures. The target word known by the largest number of subjects was *dog* (known by 34 of 35 subjects), and the relative order was as follows: *dog*, *shoe*, *cow*, *banana*, *bed*, *coat*, *bicycle*, *cup*, *bus*, *horse*, *mouse*, *cake*. Target words were known by a mean of 26 infants, ranging from 13 (*cake*), to 34 (*dog*).

Trials were removed from analysis when infants did not fixate both images at least once during the trial. According to this criterion, 15 trials were removed from the original 420. The number of valid trials available in each of the different lexical status conditions is given in Table 3.3.

Table 3.3. Number of trials according to lexical status (Broad)

		Target Status	
		Known	Unknown
Distracter Status	Known	199 (49%)	40 (10%)
	Unknown	100 (25%)	66 (16%)

The lexical status of target and distracter was matched in 67 percent of target-known trials, and in 63 percent of target unknown trials. Thus the majority of trials contained no competition between the lexical status of target and distracter.

PTL (Broad Analysis)

Individual means for the proportion of target looking (PTL) were calculated for trials in each of the four lexical status conditions. Five outlying means were removed to achieve normal (Gaussian) distribution in all subsets of the data. Figure

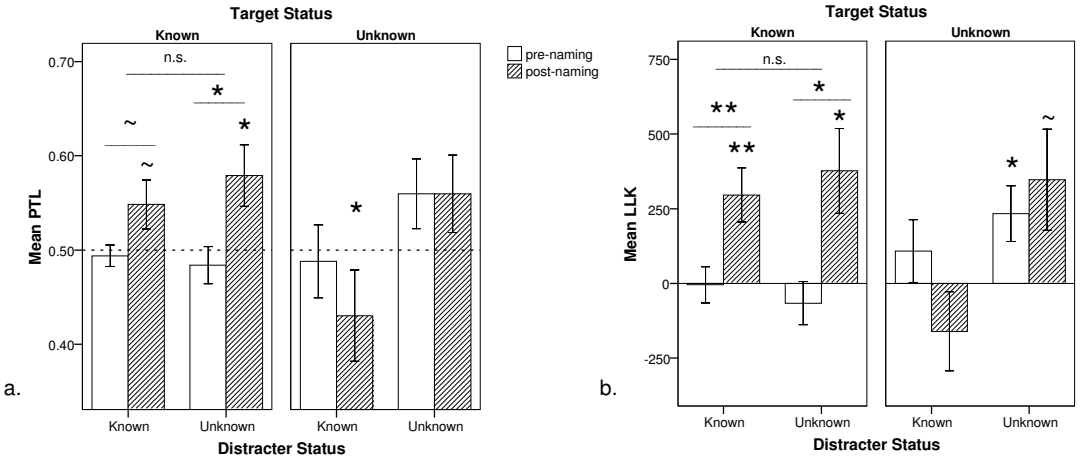
3.03 plots the pre- and post-naming values for PTL according to the status of the target and distracter. In this plot, differences between pairs of bars indicate a change in the picture preference following the onset of naming. In all conditions, the pale bars (pre-naming) are close to the chance value of .50. In the target-known trials, PTL increases following naming, while in the target unknown trials this is not the case. When both words are unknown, no change is evident, and when only the distracter is known, there is a tendency towards distracter looking following the onset of the target word.

To quantify the observed effects, parametric analyses were conducted. The asymmetrical distribution of trials across the four lexical status conditions created unequal variance in the subsets of the data, making them unsuitable for analysis in a repeated measures ANOVA. However, as the data were normally distributed at the overall level, and within each subset of the data, a stepwise linear regression of change in PTL (post-naming minus pre-naming) was instead conducted. The regression of PTL change included target status (known, unknown), distracter status (known, unknown) and sex (male, female) as potential regressors. A model in which target status was the sole regressor was the only model which produced a significant fit to the data ($F(1,102) = 6.88$, $p = .01$, adjusted $r^2 = .05$), with target status significantly predicting the change in PTL ($t = 2.62$, $p = .01$).

Simple effects were used to clarify the effect of target status. In target known trials, the mean change in PTL was .08 ($SD = .20$), and in target unknown trials the mean change was -.05 ($SD = .29$), demonstrating that change in PTL was significantly larger for target-known trials than for target unknown trials ($t(66.1) = 2.62$, $p = .017$, $d = 0.60$)[4].

[4] Adjusted degrees of freedom and *t*-value, as variances were not equal between groups.

Figure 3.03. PTL and LLK according to lexical status (Broad)



Notes. Proportion of target looking (PTL) and difference in longest look (LLK) according to target status. Horizontal lines indicate 'chance' value. Single sample t-tests comparing PTL to chance are marked above each bar. Paired sample t-tests comparing pre-naming and post-naming PTL are marked above pair of bars. Independent samples t-tests (marked above comparison) compare change in PTL according to distracter status. ** $p<.01$, *$p<.05$, ~$p<.10$. Whiskers show +/- one standard error.

Furthermore, in target-known trials, there was a systematic increase target looking following naming (pre-naming: $M = .49$, $SD = .09$; post-naming: $M = .56$, $SD = .17$; $t(61) = -2.93$, $p = .005$, $d = -.57$). When the target was unknown no change was evident (pre-naming: $M = .52$, $SD = .17$; post-naming: $M = .47$, $SD = .23$; $t(40) = 1.14$, n.s., $d = .26$). This pattern of results is consistent with infants comprehending only those words their parents reported as known.

LLK (Broad Analysis)

Individual means for the difference in length of longest look (LLK) were calculated for each of the four lexical status conditions. Ten means were removed to achieve normal distribution at the overall level, and within each subset of the data. The difference in longest look (LLK) is depicted in each of the four lexical status conditions in Figure 3.03b, where LLK increases following naming in the target known conditions, but does not increase in the target unknown conditions. In this

measure, as in PTL, a moderate decrease was evident in target looking following naming (a tendency to prefer the distracter) where the target was unknown and the distracter, known.

A four-way repeated measures ANOVA was conducted to investigate the effect of target-status (known, unknown), distracter status (known, unknown), sex (male, female) and naming (pre-naming, post-naming), on PTL. The results of the ANOVA are reported in Table 3.4, where there was a main effect of distracter status, qualified by an interaction between target status and distracter status. The expected interaction between target status and naming was also found.

Simple effects were conducted to clarify the pattern of effects and interactions. Concerning the overall effect of distracter status, mean LLK (pooled across trial phase) was 88ms when the distracter was known (*SD* = 230ms), and 204ms when the distracter was unknown (*SD* = 396ms), indicating that target looking was *lower* when infants knew the name of the distracter. When assessed independently, the comparison was not significant ($t(96) = 1.62$, n.s., $d = 0.33$), and barely approached the alpha level of .05 with a *p*-value of .11.

This finding indicates that infants tended to look less at a picture when its competitor was name known. However, the trend was an overall effect and did not interact with naming. Concerning the interaction between target status and distracter status, when the target was known, knowing the name of the distracter had no effect on LLK (distracter known: *M* = 146ms, *SD* = 277ms; distracter unknown: *M* = 155ms, *SD* = 418ms; $t(62) = 0.11$, n.s., $d = 0.03$). However, when the target was unknown, knowing the name of the distracter significantly reduced LLK (distracter known: *M* = -26ms, *SD* = 319ms; distracter unknown: *M* = 291ms, *SD* = 349ms; $t(32) = 2.75$,

p = .01, d = 0.97). Both distracter effects were at the level of overall picture interest, and did not interact with naming.

Table 3.4. Four-way repeated measures ANOVA of LLK (Broad)

Source (between S)	*df*	*F*	*p*	partial η^2
target status (T)	1	0.01	.94	<.001
distracter status (D)	1	4.48*	.04	.047
sex	1	<0.01	.98	<.001
T x D	1	4.73*	.03	.050
T x sex	1	0.57	.45	.006
D x sex	1	0.09	.77	.001
T x D x sex	1	0.61	.44	.007
S w/in group error	90	(248845)		
Source (within S)	***df***	***F***	***p***	**partial η^2**
naming	1	2.52	.12	.027
naming x T	1	6.63*	.01	.065
naming x D	1	2.05	.16	.022
naming x sex	1	0.01	.77	<.001
naming x T x D	1	0.31	.58	.003
naming x T x sex	1	0.01	.93	<.001
naming x D x sex	1	0.03	.87	<.001
naming x T x D x sex	1	0.59	.45	.006
error (naming)	90	(358690)		

Notes. Mean square error given in brackets. *p < .05.

As for the critical interaction between naming and target status, when the target was known, there was an increase in LLK following naming (pre-naming: M = -36ms, SD = 375ms; post-naming: M = 336ms, SD = 671ms; t(63) = 3.59, p = .001, d = 0.50). When the target was unknown, there was no change in LLK (pre-naming: M = 175ms, SD = 407ms; post-naming: M = 108ms, SD = 679ms; t(60) = 0.46, n.s., d = 0.11). This last finding indicates that while the lexical status of the distracter had some overall effect on the amount of target looking (particularly

when the target was reported as unknown), knowing the name of the target was the only factor responsible for the increase in target looking after the onset of the target word. This finding is consistent with the results of the linear regression model for PTL, as both analyses demonstrate that infants can discriminate the named target from the unnamed distracter when they are reported to know the word, but do not show evidence of lexical comprehension when the target is reported as unknown.

Taken together, the PTL and LLK findings from the broad analysis suggest that across thirty nine 18-month-olds, parental report predicts those trials in which infants will be able to demonstrate lexical comprehension (as shown by an increase in target looking following the onset of the target name), and discriminate them from trials in which they will not be able to demonstrate lexical comprehension.

However, recall that the full sample of 39 infants contained a heavy bias towards knowing the target (c.f., Table 2.1, with six infants reported to know all twelve target words). It is possible that the pattern of effects and interactions demonstrated here is driven by some factor other than the item-level sensitivity of parental report. As CDI report has been shown to correlate with other measurements of general intelligence and linguistic development (e.g., Benedict, 1979; Bretherton et al., 1983; Phillip S Dale et al., 1989) it is possible that high-vocabulary infants were simply better at picture-finding in the IPL task. As a greater proportion of high-scorers' trials would (by chance) end up in the target-known condition, it is possible that high-scorers lead the trend due to their general linguistic intelligence, not according to which words they were reported to know. In order to account for this possibility, a more stringent sub-sample of the original 39 infants was selected for the narrow analysis.

Narrow Analysis (*N*=16)

For the narrow analysis, infants were only included if they provided trials in each of the four lexical status conditions. 16 participants fulfilled this condition, and between them, the number of trials included in each condition was much more closely balanced. Table 3.5 gives the number of trials in each condition. There were nine males and seven females in the narrow analysis. The mean receptive CDI score for infants in the narrow analysis was 113 words of a possible 416 (*SD* = 53), ranging from 44 words to 187 words. Note that this range is relatively low compared to the percentile norms in the previous chapter. The over representation of low vocabulary infants is the result of balancing the sub-sample for number of words known and unknown.

Table 3.5. Number of trials according to lexical status (Narrow)

		Target Status	
		Known	Unknown
Distracter Status	Known	43 (22%)	35 (18%)
	Unknown	61 (32%)	53 (28%)

Between subjects in the narrow condition, the number and range of targets reported as known and unknown was closely balanced, as was the number and range of distracters reported as known and unknown: Infants included in the narrow analysis knew a mean of seven of the twelve target words, ranging from three words (known by one infant) to nine words (two infants), and a mean of five of the twelve distracter words, ranging from two words (one infant), to seven words (two infants). The best known target words were *dog* and *shoe*, known by 15 of 16 subjects. Target words were known by a mean of seven infants (range: 3 to 15). The best-known distracters were *sock* and *pushchair*, known by 11 of the 16 infants. Distracters were also known

by a mean of seven infants (range: 2 to 11). Of the sixteen infants' 192 trials, the target was reported as known in 54% of trials, while the distracter was known in 41% of trials.

PTL (Narrow analysis)

Participant means were calculated for PTL in each trial type. As data from the balanced sub-sample satisfied the required distributional symmetry, a four-way repeated-measures ANOVA was conducted comparing the effects of target status (known, unknown), distracter status (known, unknown) and sex (male, female) on naming (pre-naming, post-naming) in PTL. The results of the ANOVA are presented in Table 3.6, where main effect of distracter status is evident, along with interactions between distracter status and target status, and distracter status and sex. Of particular interest in the current analysis, there was also an interaction between target status and naming.

The pattern of effects and interactions was clarified by simple comparisons. As in the broad analysis, the main effect of distracter status generated less looking to targets when the distracter was known (distracter known: $M = .46$, $SD = .12$; distracter unknown: $M = .55$, $SD = .10$; $t(62) = 3.43$, $p = .001$, $d = 0.87$), but this effect was qualified by target status, such that PTL only differed between distracter conditions when the target was unknown (distracter known: $M = .41$, $SD = .13$; distracter unknown: $M = .57$, $SD = .10$; $t(30) = 3.81$, $p = .001$, $d = 1.39$).

Table 3.6. Four-way repeated measures ANOVA of PTL (Narrow)

Source (between S)	*df*	*F*	*p*	partial η^2
target status (T)	1	0.38	.54	.007
distracter status (D)	1	11.49	.001**	.170
sex	1	1.63	.21	.028
T x D	1	5.90	.02*	.095
T x sex	1	2.68	.11	.046
D x sex	1	5.49	.02*	.089
T x D x sex	1	0.12	.74	.002
S w/in group error	56	(0.02)		
Source (within S)	***df***	***F***	***p***	**partial η^2**
naming	1	0.27	.61	.005
naming x T	1	5.42	.02*	.088
naming x D	1	2.42	.13	.041
naming x sex	1	0.29	.87	.001
naming x T x D	1	0.04	.84	.001
naming x T x sex	1	0.13	.72	.002
naming x D x sex	1	0.83	.37	.015
naming x T x D x sex	1	2.04	.16	.035
error (naming)	56	(0.04)		

Notes. Mean square error given in brackets. **p<.01, *p<.05.

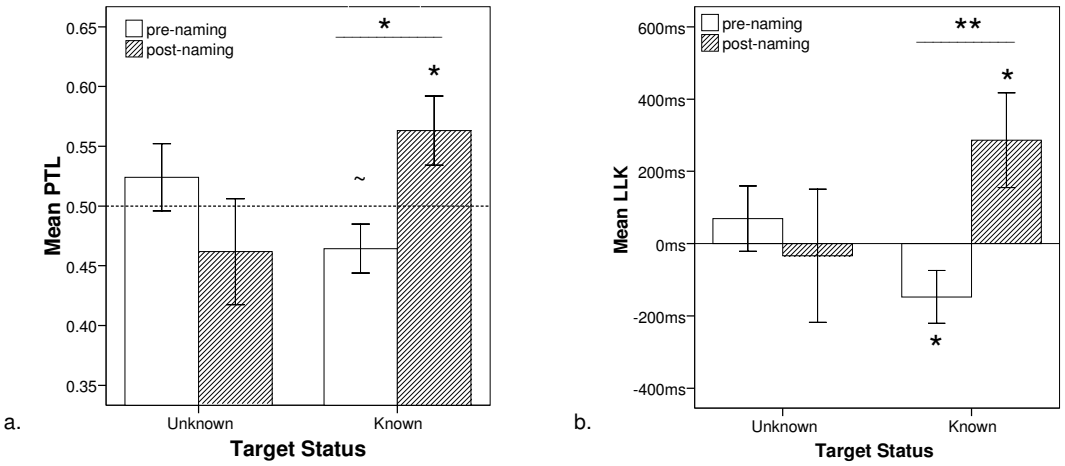
PTL remained the same when the target was known (distracter known: M = .50, SD = .10; distracter unknown: M = .53, SD = .09; $t(30)$ = 0.93, n.s., d = 0.34). The effect of distracter status was also qualified by sex, where only males showed the effect of distracter status on their PTL (distracter known: M = .42, SD = .13; distracter unknown: M = .56, SD = .11; $t(34)$ = 3.71, p = .001., d = 1.27), while females did not show a difference in PTL (distracter known: M = .51, SD = .95; distracter unknown: M = .53, SD = .08; $t(26)$ = .81, n.s., d = 0.32). These effects occurred at the overall level, and did not interact with naming.

Concerning the critical interaction between naming and target status, there was an increase in PTL in those trials where the target was reported as known

(pre-naming: M = .46, SD = .12; post-naming: M = .56, SD = .16; $t(31) = 2.63$, p = .01, d = 0.62), but no change was evident in those trials where the target was unknown (pre-naming: M = .52, SD = .16; post-naming: M = .46, SD = .25; $t(31) = 1.12$, n.s., d = 0.26). This response pattern is consistent with infants discriminating the target on only those trials where the target marked as 'understood'.

As was seen in the broad analysis, distracter status had an effect on target looking – it affected looking at the overall level of general picture interest, but did not have an effect on the change between pre- and post-naming looking. Only knowing the target influenced whether there was an increase in target looking following naming. As differences caused by distracter status and sex were not found to influence the relationship between target status and naming, the data were pooled for presentation in Figure 3.04.

Figure 3.04. PTL and LLK according to target status (Narrow)



Notes. Participant means for proportion of target looking (PTL) and difference in longest look (LLK). Horizontal lines indicate chance value for each measure. One sample t-tests comparing mean to chance are marked above each bar. Paired sample t-tests comparing pre-naming to post-naming are marked above each pair of bars. **p<.01,* p<.05, ~p<.10. Whiskers show +/- one standard deviation.

LLK (Narrow analysis)

Individual means for LLK in the narrow analysis (N = 16) are depicted in Figure 3.04b, where pre-naming values are close to the chance value of 0ms and post-naming values are higher (at about 300ms), but only when the target is reported as known. LLK was subjected to a four-way repeated-measures ANOVA comparing the effect of target status (known, unknown), distracter status (known, unknown) and sex (male, female) on LLK, according to trial phase (pre-naming, post-naming). The results of the ANOVA are given in Table 3.7, where a main effect of distracter status is evident, along with an interaction between distracter status and sex. Three further interactions did not achieve significance, but suggested trends in the data: target status with naming, target status with distracter status, and target status with sex.

Simple effects clarified the pattern of effects and interactions. As in previous analyses, knowing the name of the distracter significantly reduced target looking (distracter known: M = 554ms, SD = 96ms; distracter unknown: M = 215ms, SD = 67ms; $t(62) = 2.92$, $p = .005$, $d = 0.74$). This effect was qualified by the interaction between target status and distracter status, whereby distracter status did not create a difference when the target was known (distracter known: M = 2ms, SD = 457ms; distracter unknown: M = 137ms, SD = 380ms; $t(30) = 0.91$, n.s., $d = 0.33$), but reduced target looking when target was unknown (distracter known: M = -257ms, SD = 606ms; distracter unknown: M = 292ms, SD = 368ms; $t(30) = 3.10$, $p = .004$, $d = 1.13$).

Table 3.7. Four-way repeated measures ANOVA of LLK (Narrow)

Source (between S)	*df*	*F*	*p*	partial η^2
target status (T)	1	0.06	.82	.001
distracter status (D)	1	7.74**	<.01	.121
sex	1	0.92	.34	.016
T x D	1	3.44~	.07	.058
T x sex	1	3.15~	.08	.053
D x sex	1	4.02*	.05	.067
T x D x sex	1	0.03	.86	.001
S w/in group error	56	(400554)		
Source (within S)	***df***	***F***	***p***	**partial η^2**
naming	1	1.46	.23	.025
naming x T	1	3.67~	.06	.062
naming x D	1	2.20	.14	.038
naming x sex	1	0.50	.82	.001
naming x T x D	1	0.06	.81	.001
naming x T x sex	1	0.46	.50	.008
naming x D x sex	1	0.05	.82	.001
naming x T x D x sex	1	0.33	.57	.006
error (naming)	56	(565095)		

Notes. Mean square error given in brackets. **p<.01, *p<.05. ~p<.10

The effect of target status also tended to interact with sex, and further investigation revealed that only males showed a significant effect of distracter status (distracter known: M = -274ms, SD = 585ms; distracter unknown: M = 266ms, SD = 399ms; $t(34) = 3.23$, $p = .003$, $d = 1.11$), while females did not show a difference according to whether the distracter was known or unknown (distracter known: M = 61ms, SD = 436ms; distracter unknown: M = 148ms, SD = 347ms; $t(26) = .59$, n.s., $d = 0.23$). Sex also tended to interact with target status, although further investigation did not reveal any significant effect of target status for males (target known: M = 110ms, SD = 511ms; target unknown: M = -117ms, SD = 606ms;

$t(34) = -1.21$, n.s., $d = 0.42$) or for females (target known: $M = 18$ms, $SD = 268$ms; target unknown: $M = 191$ms, $SD = 476$ms; $t(26) = 1.19$, n.s., $d = 0.47$).

Investigating the critical trend between target status and naming revealed that when the target was reported as known, there was a significant increase in LLK (pre-naming: $M = -147$ms, $SD = 412$ms; post-naming: $M = 286$ms, $SD = 742$ms; $t(31) = 2.85$, $p = .008$, $d = 1.02$), and when the target was unknown, there was no change in LLK (pre-naming: $M = 69$ms, $SD = 512$ms; post-naming: $M = -34$ms, $SD = 1043$ms; $t(31) = 0.49$, n.s., $d = 0.17$). Although the target-by-naming interaction did not achieve statistical significance in the ANOVA, the high significance of the pre- to post-naming change in the target known condition demonstrates that the effect of target naming was reliable. As neither distracter status nor sex interacted with pre- to post-naming change in LLK, data were pooled across these conditions for presentation in Figure 3.04b.

These findings demonstrate that even when the sample size is reduced to a small sample of infants (whose trials were evenly spread across the four lexical status conditions), that parental report accurately predicted the group of trials in which infants were able to demonstrate lexical comprehension (as measured by two standard IPL macro-level measures), and discriminated them from those trials in which infants were not able to demonstrate lexical comprehension. While distracter status did affect target interest overall, the effects of distracter status were most obvious for boys, and when the target word was not known. Even though these effects were quite pronounced, they did not interact with the effect of target status on naming.

One additional possibility remains for interpretation of these results. If IPL indexes infants' general linguistic ability, regardless of item, then infants with larger vocabularies may still be leading the trend in the narrow analysis. To address this

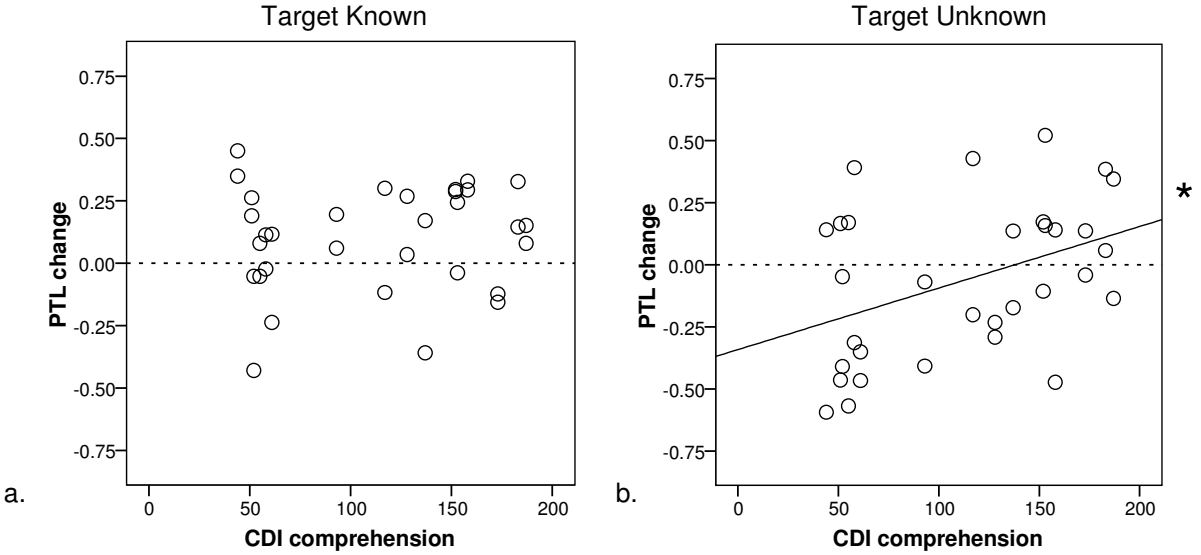
possibility, an analysis was conducted to evaluate the impact of vocabulary size on performance, using change-in-PTL following naming as the dependent measure. It was predicted that if change-in-PTL indexed general task performance, then infants at the high and the low end of the vocabulary scale would perform differently for words reported as 'known', with higher performance for large vocabulary infants. If this were the case, for 'known' words, change-in-PTL would be expected to co-vary with infants' CDI comprehension score.

Vocabulary size and change in target preference (Narrow analysis)

CDI comprehension score is plotted against change in PTL in Figure 3.05, for each infant included in the narrow analysis. Figure 3.05a depicts the spread of values when the target is reported as known. In this plot, the values tend to be above chance, and the distribution is constant throughout the CDI range. No correlation was evident ($\rho(32) = .09$, n.s.). Infants were also separated into two CDI groups according to a median split of their comprehension score. No difference in PTL change was evident between low- and high-vocabulary infants (low CDI: $M = .12$, $SD = .20$; high CDI: $M = .08$, $SD = .23$; $t(30) = 0.61$, n.s., $d = 0.22$). This finding indicates that for known words, infants across the vocabulary range presented the same pattern of results: an increase in target looking ($t(31) = 2.63$, $p = .01$, $d = 0.47$) which was above the level predicted by chance.

Figure 3.05b depicts comprehension score and PTL change when the target is reported as unknown. Here a positive relationship is evident, and infants with the smallest CDI scores appear to *decrease* their looking to the named target when the target is unknown. This pattern was confirmed by a significant negative correlation between CDI comprehension score and change-in-PTL ($\rho(32) = .37$, $p<.05$).

Figure 3.05. CDI score and change-in-PTL according to target status (Narrow)



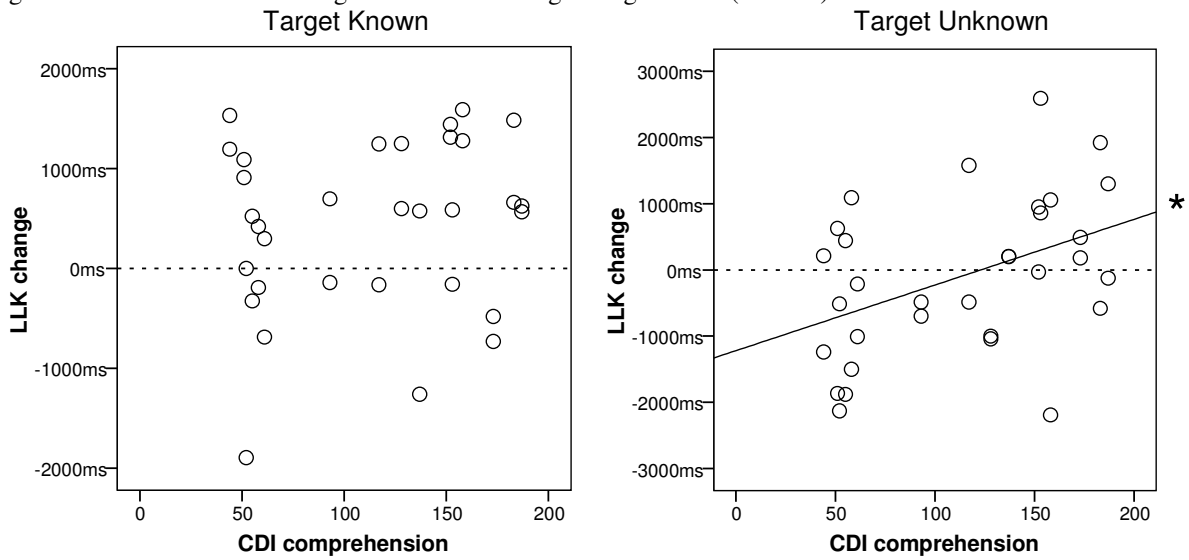
Notes. Change in PTL represents the proportion of target looking (PTL) in the post-naming period minus the pre-naming period. Dashed horizontal lines indicates 'chance' value of 0. Solid line indicates linear fit for total, where significant. *p<.05.

According to a median split of the CDI scores, the PTL change was larger (in the negative direction) for infants with lower CDI scores (low CDI: M = -.16, SD = .34; high CDI: M = .04, SD = .27) although the comparison did not achieve significance (t(30) = 1.86, n.s., d = 0.68). A trend was evident, however, approaching the alpha level of .05 with a p-value of .07. This moderate trend suggests that infants with smaller vocabularies tend to look to the *distracter* following naming when the target is unknown.

Figure 3.06 depicts the relationship between LLK change and CDI comprehension score, where the same pattern was found: For target known trials, no correlation (ρ(32) = .09, n.s.) and no difference between median split groups (low CDI: M = 282ms, SD = 870ms; high CDI: M = 584ms, SD = 850ms; t(30) = 1.00, n.s., d = 0.36); For target unknown trials, a positive correlation (ρ(32) = .37, p = .04) and a non-significant trend for low CDI infants to show a reduction in LLK (low CDI: M = -504ms, SD = 1091ms; high CDI: M = 299ms, SD = 1183ms; t(30) = 2.00, n.s.,

d = 0.73), where the p-value of .055 approached the alpha level of .05. That is to say, smaller vocabulary infants tended to look to the distracter following naming.

Figure 3.06. CDI score and change-in-LLK according to target status (Narrow)



Notes. Change in LLK represents the longest look difference (LLK) in the post-naming period minus the pre-naming period. Dashed horizontal line indicates 'chance' value of 0. Solid line indicates linear fit for total, where significant. *p<.05.

Taking these two analyses together, both PTL and LLK demonstrate that when the target is reported as known, lexical comprehension (as indexed by an increase in target looking following naming) does not pattern with CDI score. That is to say, small vocabulary infants perform just as well as large vocabulary infants in the referent identification task, but this is only the case when their parents report the items as known. These findings demonstrate that the general pattern of target discrimination for known words was not driven by general linguistic development, but by accurate parental identification of items attracting performance consistent with lexical comprehension.

When targets were reported as unknown, by comparison, there was a tendency for target looking to decrease following naming, but only for the smaller vocabulary infants. As the data in this comparison were pooled across distracter conditions, it is

likely this trend was driven by cases in which the target was unknown and the distracter, known – a trend which was evident in Figure 3.03. In other words, small vocabulary infants tended to look at a name-known picture when they heard an unfamiliar label, even though the picture did not match the label.

Recall that mutual exclusivity (ME) or N3C would predict that infants would reject a known distracter when they hear an unknown target. The trend observed here is thus in opposition to the standard description of mutual exclusivity. However, recent research has shown that ME-style responding requires a certain amount of processing time. Mather and Plunkett (In Press) for example, report an IPL experiment in which a novel object was presented alongside a familiar object, while a novel word was heard. Post-naming preference for the novel object did not appear on the first presentation of object pairs, but *was* observed on subsequent presentations. The current results are consistent with their findings, as infants did not show ME-style target preference following naming, when exposed to pictures for a single presentation. The trials reported here might illustrate an early stage of referent checking, prior to the emergence of a mutually exclusive response.

Summary: Does Mummy know which words are understood?

In summary, target-known trials showed a significant increase in target preference following the onset of naming. By contrast, the target-unknown trials showed no systematic change in preference, with preference for both pictures remaining similar throughout. Knowing the name of the distracter also affected the looking pattern at the overall level, but did not interfere with the general effect of naming on target recognition. This finding demonstrates a match between an IPL task and CDI report for a particular population of infants (British 18-month-olds), whereby

parental report of individual words predicted which words would attract looking behaviour consistent with lexical comprehension.

Importantly, these findings were reliable when all infants were included in the analysis, and when the subjects were reduced to a small, balanced sub-sample of original participants. It was also demonstrated that performance in target known trials was equally reliable for large- and for small- vocabulary infants, indicating that the observed effect was not simply led by general linguistic ability, but by the particular items listed as 'known'.

An interaction between sex and distracter status was also observed, in which boys were more likely to show reduced target looking when the distracter was name known: In LLK, the reduction in target looking was only evident for boys, and when the target was unknown. This effect was only apparent in the narrow analysis, where the discrepancy between number of boys and girls was more pronounced (ten males, six females) than in the broad analysis (twenty males, fifteen females), suggesting that it may have been a statistical artefact, produced by the discrepancy in sample size. In analysis of both PTL and LLK, sex-by-distracter interactions influenced target looking at the overall effect, and did not interact with infants' pattern of looking more and longer at known targets following the onset of naming – nor with parents' ability to pick out the items which attracted increases in target looking.

These findings differ from the results of Houston-Price et al. (2007), which showed performance consistent with lexical comprehension for all items, whether the target had been rated as known or unknown. The Houston-Price et al., experiment shares many similarities with the current experiment: both experiments assessed lexical comprehension at 18 months-of-age using the same CDI, with the same instruction sheet; the IPL experiments contained trials of the same duration, presented

on-screen in the same manner, with target words occurring in similar pre-recorded phrases; infants sat with their parents, who were given the same instructions, and the experimenter was not present during test. Thus the discrepancy in results must lie, not in general methodological differences, but in the particular stimulus combinations.

Two key differences in stimulus organisation can be highlighted, both of which affect the difficulty of the task: distracter selection, and stimulus repetition. In the current study, all picture-pairs contained a target and a distracter from the same semantic category. In the pair, 'sock-trousers', for example, both images were clothing, made of fabric, likely to be encountered in the routine 'getting dressed,' which typically occurs indoors. Infants therefore needed to know more about target words in order to distinguish them from this type of distracter: Knowing that 'sock' is related to getting dressed would be insufficient to pick out a sock from trousers. By comparison, if a picture of a sock were paired with a picture of a lorry, a lower level of knowledge about the referent of the word 'sock' would allow the lorry to be rejected. Thus the current task assesses lexical comprehension at the level of *sock-trousers*, while the Houston-Price et al. task (2007), assesses the less difficult level of referent identification, *sock-lorry*.

An additional level of difficulty in the current experiment was provided by the 'one-shot' design, in which infants encountered each item only once. There are a number of reasons why a single presentation of picture pairs should be predicted to make an IPL task difficult. To give just one example, considering that infants have only a brief exposure to the test exemplars prior to naming (2,500ms), and only a short amount of time (2,500ms) to demonstrate comprehension following naming, if an infant needed longer to decide whether a novel exemplar matched a familiar label, they would not be able to demonstrate comprehension in a 'one shot' experiment.

One final difference in stimulus organisation remains. In the Houston-Price et al. task, all stimuli were selected from a pool of items which had been reported to be encountered at least weekly. In the current experiment, stimuli were selected from a pool of words reported as 'understood' by a large cohort of eighteen-month-olds. The cohort comprehension data provided an approximation of item familiarity at this age (on the logic that items typically encountered frequently are more likely to be understood). However, no direct measure of object familiarity was included in the current design. This was, in part, to ensure that all infants were able to view the same set of stimuli. It is therefore possible that object familiarity played some part in infants' differential responding to words reported as known and unknown. As item familiarity, item repetition and distracter-type are confounded between the two studies, it is impossible to draw direct conclusions from their comparison. A more systematic investigation of these elements (i.e., a one-shot IPL design pitting object familiarity against distracter-type) would be needed to clarify the exact contribution of each. Despite this limitation, the results of the current experiment still contribute new insights to the kind of IPL task which can match parental CDI reports, and the accuracy with which both IPL and CDI approximate the nebulous concepts of 'comprehension,' and 'understanding.'

Interim conclusions

1. Infants can identify named targets in an unfamiliar environment, in a 'one shot' task when a potentially confusing priming phrase precedes the standard looking task – as long as the target word is reported as understood.

2. Parental report is sufficiently accurate to separate those items which infants will be able to identify (the words they know) from those items which they will not be able to identify (the words they don't know) in the current IPL environment.

3. CDI reports of comprehension are a valuable predictor of performance in an experimental test of comprehension at this age.

4. Both CDI and IPL demonstrate item-level accuracy, thus validating the use of item-by-item exclusions in future priming studies.

3.2 Priming conditions

Having established that parental report is a reliable indicator of which words infants know well enough to recognise in an unfamiliar environment, and that infants are able to demonstrate the normal pattern of target recognition for their age group in a primed IPL design, it remains to be seen whether differences between the three priming conditions are evident in infants' performance. The three priming conditions (unrelated, taxonomic and compound) were selected to assess two possible relationships which might exist in the early lexicon. The relationship of taxonomy (belonging to the same category) is one which infants may be able infer from their environment, through recognition of shared properties such as visual features (e.g., faces on animals, or textured surfaces of clothes), or shared semantic features (e.g., 'can eat it', for food, or 'can ride in it' for vehicles). The relationship of association

(as documented in formal word association tasks) might develop in the infant lexicon through basic associative learning, such as hearing two words in close proximity (e.g., 'raining cats and dogs') or encountering two objects at the same time (e.g., bowl and spoon). As noted in the previous chapter, the majority of adult associations between words common in the 18-month-old lexicon occur between taxonomic sisters. It was thus decided to combine taxonomy and association in the compound prime condition, as there were insufficient words available to independently manipulate association. Given that adults typically demonstrate an associative 'boost' in priming studies when association is combined with taxonomy (Moss et al., 1995), it was also predicted that this priming condition might generate a larger priming effect than the taxonomic prime condition.

In adapting the adult priming paradigm to the infant IPL methodology a number of complications were encountered. As the standard IPL design includes a pre- and a post-naming phase, infants are exposed to two images during the inter-stimulus interval (ISI) between the offset of the prime and the onset of the target word. It is possible that following a priming phrase containing 'cat', an infant might begin looking at a target picture of a 'dog' prior to the onset of the target word. This would confound lexical priming with visual similarity, as it would be unclear whether 'cat' had primed a lexical representation of a 'dog' via spreading activation in the lexicon, or whether 'cat' generated a semantic representation of a cat which shared resembled the dog picture.

In order to avoid such visual biases in the pre-naming target bias, taxonomic sisters were selected as picture-pairs, ensuring that the prime would be an equally good visual match to *either* picture in the pair. Following target naming however, if the auditory prime was still active via spreading activation in the lexicon, or via some

short-term strategic memory process, then its effect might become evident in the magnitude of target preference, or in the ease and speed of target word processing. As infants have already been shown to demonstrate lexical comprehension in this task, the question of whether their performance is influenced by the priming condition remains to be addressed.

The previous section established that parental report in the Oxford CDI was a valuable index of lexical comprehension behaviour in 18-month-olds. CDI report will be used as an item-by-item exclusion criterion for the remainder of the chapter, and later experiments reported in this thesis. Only trials in which both the prime and the target were reported as known will be included in analysis. Given the high 'level' of comprehension identified by parental report, this lexical exclusion criterion may be considered too strict, or overly conservative (as infants may have some level of word-object knowledge, without meeting their parents' criterion for 'understanding'). However, as the current series of experiments were designed to assess relationships *between* items in the lexicon, it was considered inappropriate to test for such relationships between items whose status in the lexicon was unclear. Thus only words known well enough to be identified by parents were included.

Data from both 18-month-old and 24-month-olds are assessed. There was a notable difference in the number of stimulus words reported as known for the two age groups. The resulting asymmetry in the number of exclusions generated differences in the variance between the two age groups. Therefore data from each age group are analysed separately.

Results: 18-month-olds

To ensure that differences between individuals did not outweigh possible priming effects, CDI information was collated to check that infants provided valid trials (prime known, target known) in at least two of the three priming conditions. Of the 35 18-month-olds for whom both CDI and testing data were available, five individuals were excluded for this reason, leaving 30 18-month-olds in the subject pool. 28 infants provided trials in the unrelated prime condition, 22 provided trials in the taxonomic prime condition, and 28 provided trials in the compound prime condition. Of the 360 valid trials available for analysis, the prime was known in 251 trials (70%), and the target, 281 trials (78%). Both the prime and the target were known in 220 trials (88%). The exclusion criteria are stringent, but allow assessment of only those trials in which both words are known, and therefore, between which a relationship might have developed in the lexicon.

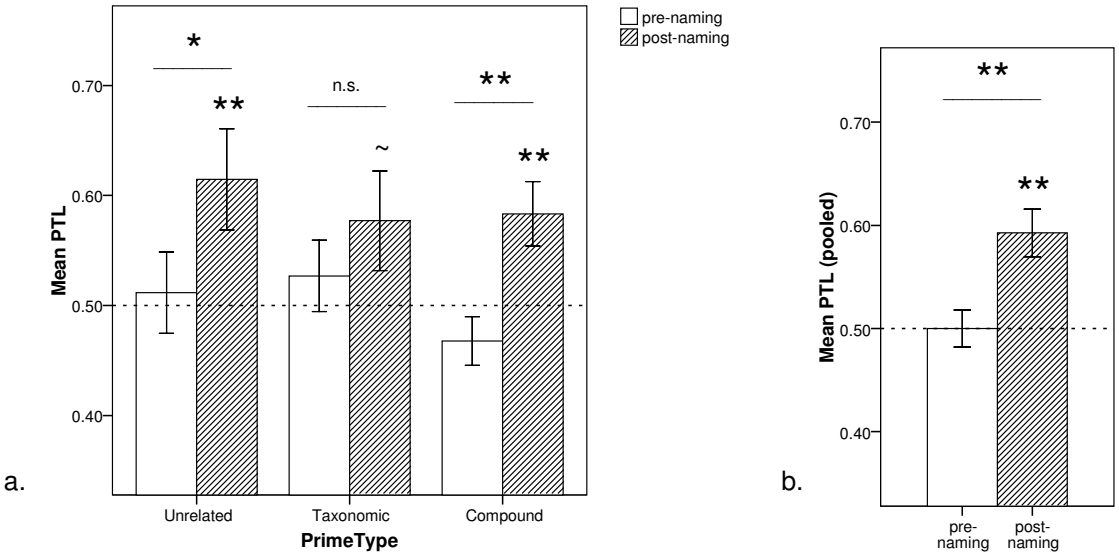
PTL (18-month-olds)

The mean proportion of target looking (PTL) in each priming condition was calculated separately for each infant. Mean PTL is plotted in Figure 3.07a, where PTL is approximately at chance prior to target naming, and increases following naming in all three priming conditions. There is little apparent difference between priming conditions.

To quantify the influence of priming condition and stimulus list, PTL data were subjected to a three-way repeated measures ANOVA comparing the effect of priming condition (unrelated, taxonomic, compound), and stimulus list (List A, List B, List C) with the effect of naming (pre-naming, post-naming). The results are given

in Table 3.8, where a main effect of naming is apparent, with no further effects or interactions.

Figure 3.07. Mean PTL in three priming conditions (18m)



Notes. Participant means for proportion of target looking (PTL). a. depicts the three priming conditions separately. b. depicts the data pooled across priming condition. One-sample *t*-tests comparing mean to chance value of .5 are marked above each bars. Paired-sample *t*-tests comparing pre-naming and post-naming PTL are marked above pairs of bars. ~p<.10, *p<.05, **p<.01. Whiskers show +/- one standard error.

Data were collapsed across list and priming condition for simple comparison of the pre- and post-naming PTL values. Pooled data are depicted in Figure 3.07b. Pre-naming PTL did not differ from chance (M =.50, SD =.16, $t(77) = 0.01$, n.s., $d < .001$), while post-naming PTL was significantly above chance (M = .60, SD = .20, $t(77) = 4.01$, $p < .001$, $d = 0.45$). The increase from pre- to post-naming was significant ($t(77) = 3.11$, $p = .003$, $d = 0.43$). Given that all trials included in priming analysis were target-known, this analysis re-states the results of the previous section (from means generated across prime conditions). The finding is consistent with lexical comprehension following naming, and together, the PTL findings suggest that the priming condition did not influence post-naming change in target preference.

Table 3.8. Three-way repeated measures ANOVA for PTL (18m)

Source (between S)	*df*	*F*	*p*	partial η^2
prime type	2	0.68	.51	.019
list	2	1.15	.32	.032
prime type * list	4	0.61	.66	.034
S w/in group error	69	(0.03)		
Source (within S)	***df***	***F***	***p***	**partial η^2**
naming	1	10.02**	.002	.127
naming * prime	2	0.53	.59	.015
naming * list	2	1.91	.16	.052
naming * prime * list	4	0.43	.79	.024
error (naming)	69	(0.04)		

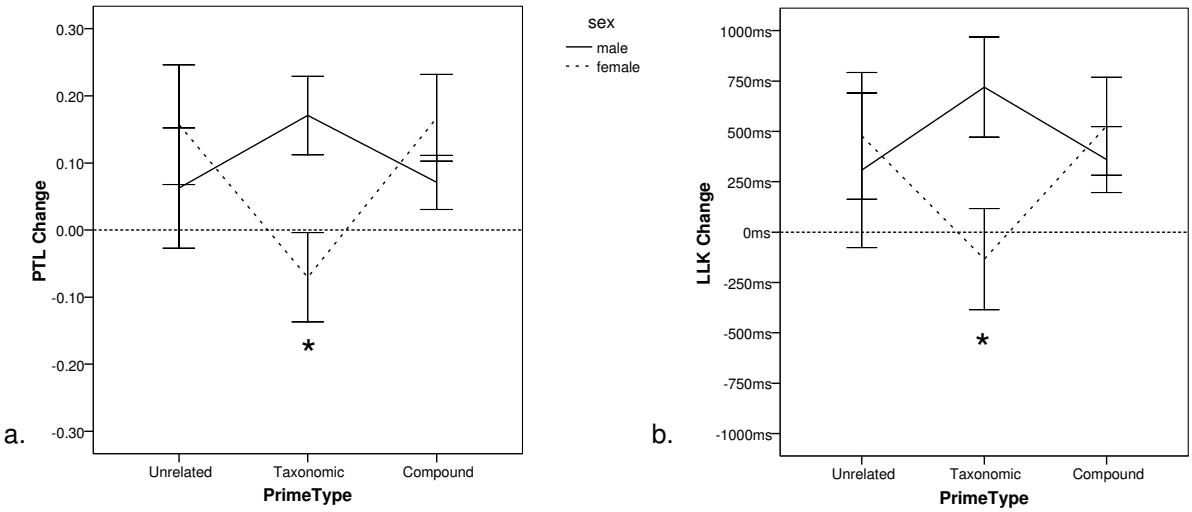
Notes. Mean square error given in brackets. **p<.01

An additional comparison was conducted to check for possible sex differences. As list was found not to interact with prime condition or naming, participants were pooled across list. Change-in PTL is plotted according to prime condition and sex in Figure 3.08a, where the change is *below* the chance value of 0ms for females in the taxonomic prime condition. As the number of male and female participants differed, PTL failed to achieve equality of variance required for analysis by repeated measures ANOVA. Instead, change-in-PTL (post-naming minus pre-naming) was assessed in a two-way ANOVA comparing priming condition (unrelated, taxonomic, compound) and sex (male, female). Main effects were not observed for priming ($F(2, 72) = 0.49$, n.s., partial $\eta^2 = .014$), nor for sex ($F(1, 72) = 0.49$, n.s., partial $\eta^2 = .001$). There was, however, a significant interaction between priming and sex ($F(2, 72) = 3.32$, $p = .04$, partial $\eta^2 = .084$). Simple

comparisons revealed that change-in-PTL differed between the sexes in the taxonomic prime condition (male: M = .17, SD = .19; female: M = -.07, SD = .22; $t(20)$ = -2.72, p = .01, d = 1.22), where females showed a smaller change in PTL than males. There was no difference between the sexes in the unrelated prime condition (male: M = .63, SD = .36; female: M = .16, SD = .31; $t(26)$ = 0.47, n.s., d = 0.29), nor in the compound prime condition (male: M = .07, SD = .16; female: M = .17, SD = .23; $t(26)$ = 1.30, n.s., d = 0.51).

The female data were investigated further with simple comparisons. Change in PTL was smaller in the taxonomic prime condition (M = -.07, SD = .22) than in the unrelated condition (M = .16, SD = .31). The comparison did not reach significance ($t(21)$ = 2.01, n.s., d = 0.88). However, a trend was evident, with a p-value of .057 approaching the alpha-level of .05. Change in PTL was also smaller in the taxonomic prime condition than in the compound prime condition (M = .17, SD = .23), and this comparison was significant ($t(22)$ = 2.55, p = .02, d = 1.09).

Figure 3.08. Change in target preference according to sex and priming condition (18m)

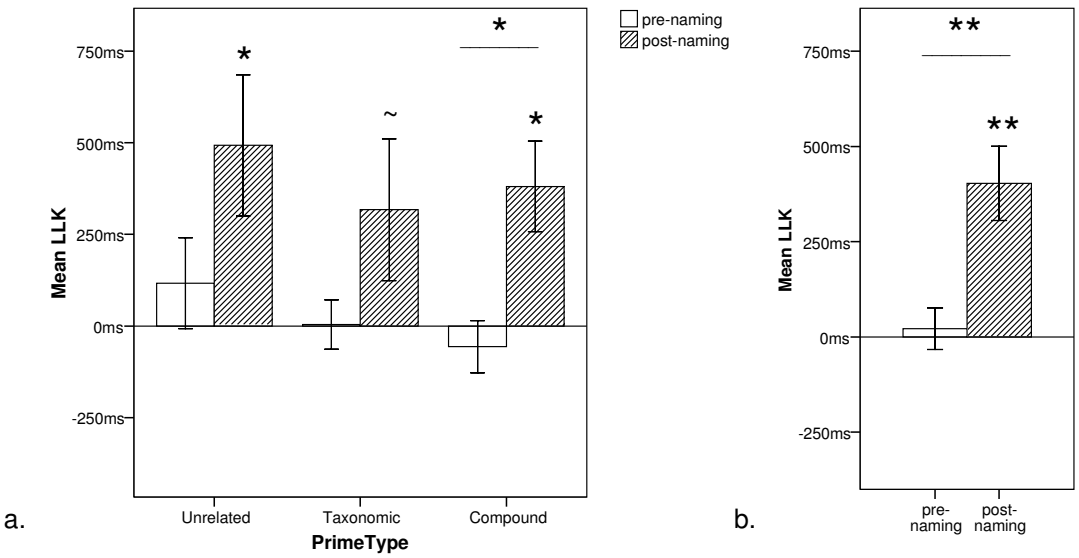


Notes. Results of independent samples t-tests are marked below sex comparison. *p<.05

LLK (18-month-olds)

Two outliers were removed to achieve normal distribution within all subsets of the LLK data. These means represented single trials from two participants, and both were more than two standard deviations from the mean within their priming condition (-2000ms, and 1680ms). Participant means are plotted in Figure 3.09a, where pre-naming LLK in all three conditions is close to the chance value of 0ms, and post-naming LLK is larger. The data did not satisfy the equality of variance-covariance matrices required for analysis by repeated measures ANOVA, however, change-in-LLK satisfied the homogeneity of variance required for univariate ANOVA analysis. A three-way ANOVA was conducted comparing the effect of prime condition (related, taxonomic, compound), list (List A, List B, List C) and sex (male, female) on LLK change (calculated as post-naming minus pre-naming).

Figure 3.09. Pre- and post-naming LLK (18m)



Notes. One sample t-tests comparing mean to the 'chance' value of 0ms are given above each bar. Paired samples t-tests comparing pre- and post-naming values are given above each pair of bars. $^{**}p<.01$, $^{*}p<.05$, $\sim p<.10$. Whiskers show +/- one standard error.

The results of the ANOVA are given in Table 3.9, where no main effect of prime or sex was evident. An interaction between prime type and sex was evident. A near-significant main effect of list was also apparent, with a *p*-value of .07 approaching the alpha-level of .05. These findings suggest that stimulus list may have played a moderate role in the magnitude of target preference shown following naming. Given that the target items did not differ across lists, it is tempting to interpret the effect of list as a systematic effect of the combination of auditory primes with targets. However, it is also possible that the rigorous exclusion criteria for trials (both prime and target understood) generated an unequal pattern of exclusions across stimulus lists.

Table 3.9. Three-way ANOVA for LLK (18m)

Source (between S)	*df*	*F*	*p*	partial η^2
prime type	2	0.54	.59	.018
list	2	2.67~	.08	.084
sex	1	0.29	.59	.005
prime type * list	4	0.55	.70	.036
prime type * sex	2	4.19*	.02	.126
list * sex	2	1.52	.23	.050
prime type * list * sex	4	0.20	.94	.013
S w/in group error	58	(.06)		

Notes. Mean square error given in brackets. *p<.05. ~p<.10

Change-in-LLK is plotted separately for the sexes in the three priming conditions in Figure 3.08b, where lower values for females in the taxonomic condition are again apparent. Concerning the significant interaction between prime type, list and sex, simple comparisons revealed that in the taxonomic prime condition, change-in-LLK was smaller for females (M = -134ms, SD = 795ms) than for males (M = 719ms, SD = 826ms), and this difference was significant ($t(19) = 2.41$, $p = .03$, $d = 1.10$). No

difference was observed between the sexes in the unrelated prime condition (males: M = 307ms, SD = 1533ms; females: M = 477ms, SD = 1044ms; $t(25) = 0.32$, n.s., $d = 0.13$), nor in the compound prime condition (males: M = 360ms, SD = 632ms; females: M = 526ms, SD = 876ms; $t(26) = 0.58$, n.s., $d = 0.23$).

For female infants, as in the analysis of PTL, change in LLK was smaller in the taxonomic condition (M = -133ms, SD = 795ms), than in the unrelated condition (M = 477ms, SD = 1043ms), although this comparison was not significant ($t(19) = 1.50$, n.s., $d = 0.69$). Change in LLK was also smaller in the taxonomic condition than in the compound prime condition (M = 526ms, SD = 876ms), a comparison showing a non-significant trend ($t(21) = 1.86$, n.s., $d = 0.81$), with a p-value of .08 approaching the alpha-level of .05. This pattern parallels the interaction between sex and prime type found in PTL. Note that there were less females than males in the current 18-month-old sample, and that following lexical exclusion, the taxonomic priming condition had the smallest number of trials. Thus the subset of the data with the smallest statistical weight is also the subset with the most divergent results.

Finally, for comparison of pre- and post-naming LLK, data were pooled across priming condition, sex and list. The pooled data are presented in Figure 3.08b. Pre-naming LLK did not differ from chance (M = 22ms, SD = 476ms, $t(75) = 0.40$, n.s., $d = 0.05$), indicating equal preference for the target and distracter prior to naming. Post naming LLK was significantly above chance (M = 403ms, SD = 846ms, $t(75) = 4.15$, $p<.001$, $d = 0.48$), indicating longer looks to the target. The increase in target looking from pre-naming to post-naming was significant ($t(75) = 3.26$, $p = .002$, $d = 0.50$).

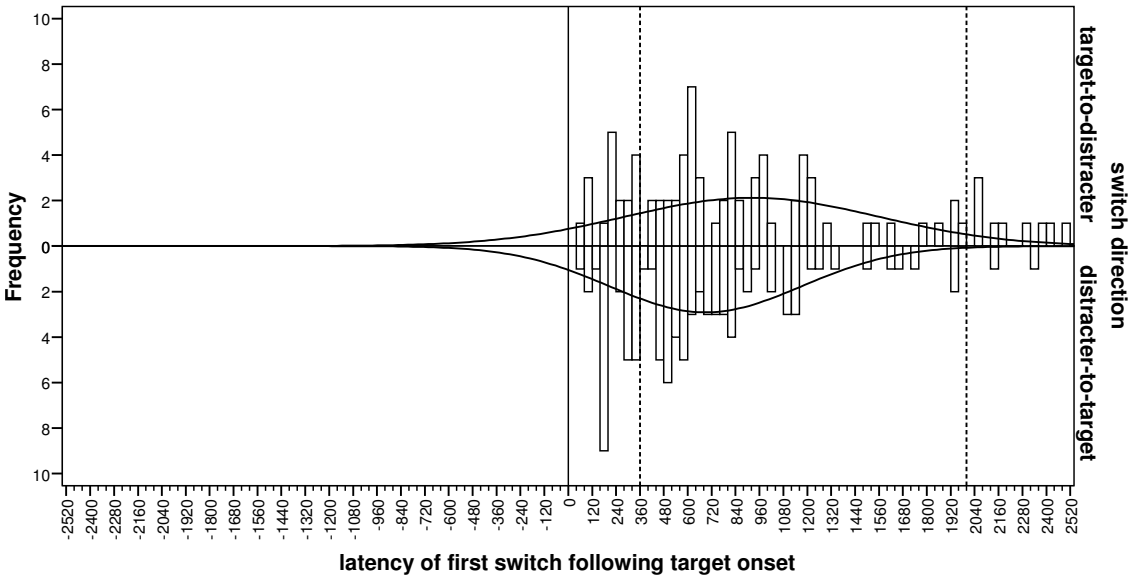
Latency & Time Course

Latency is a measure of reaction time, measured from the onset of the target word to the first frame in which the alternate picture is fixated (i.e., the saccade has been landed). Distribution of latency to switch is plotted in Figure 3.10 for valid latency trails (in which 18-month-olds knew both the prime and the target, fixated each picture in the display at least once, and were fixating one of the pictures at the onset of target presentation). This figure also illustrates an age-appropriate window for latency analysis established by Fernald Swingley & Pinto (2001)[5]. The early 'spike' in pre-window values is understood to represent trials in which an eye movement was planned or initiated prior to the onset of the target word. The upper bound of the latency window is understood to exclude late shifts which do not constitute 'responses' to the target word (for full discussion, see Fernald *et al.*, 2008).

An overall trend is evident towards earlier latencies when the infant is fixating the distracter at the onset of the target word. This is consistent with previous research demonstrating that infants in their second year are motivated to shift their gaze quickly when they hear a known-label which mismatches the currently fixated object (Swingley & Fernald, 2002). The difference between the latency of shifts away from target and distracter can be treated as a measure of target discrimination.

[5] The precise window in the Fernald et al. (2001) study began at 333ms, according to the 33ms frame-rate of the NTSC video capture system used in the Stanford Centre for Infant Studies. The window is adapted to 360ms, for the 40ms frame-rate of the PAL video capture system in the Oxford BabyLab.

Figure 3.10. Histogram of latency over the course of the trial (18m)



Notes. Time scale illustrates entire picture presentation time. Origin (0ms) indicates onset of target word. Dashed vertical bars at 360ms and 2000ms depict show latency 'window' of Fernald, Swingley & Pinto (2001). Normal curve plotted over all valid latency trials.

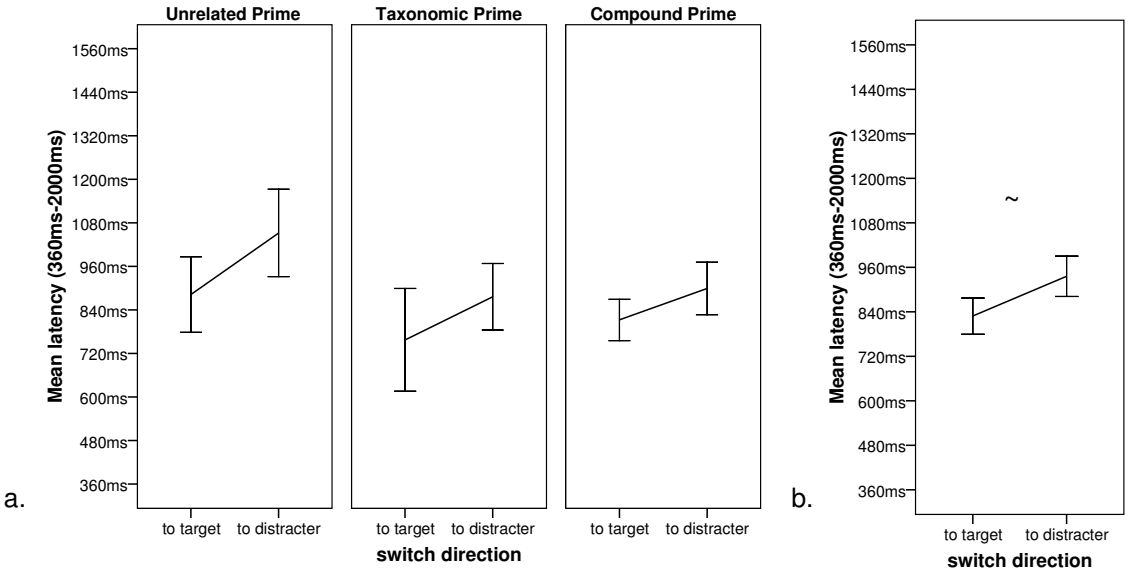
Table 3.10 summarises the distribution of latency values according to the latency window (360ms to 2000ms), and includes the number of trials in which no shift occurred before the end of the trial. In the majority of trials (87%), infants shifted their gaze before picture offset. A larger percentage of infants did not initiate a shift before the end of the trial if they were fixating the target at naming (19%), compared to if they were fixating the distracter (6%). Of those trials in which a switch was made, a larger proportion of infants landed a fixation by the end of the latency window when shifting from the distracter to the target (97%), than when shifting away from the target to the distracter (89%). The pattern of more and faster distracter rejections is consistent with infants quickly detecting the mismatch between distracter picture and the unfolding target word.

Table 3.10. Number of shifts following target onset (18m)

	Early shift	Latency Window	Late shift	No shift
	<360ms	360-2000ms	>2000ms	>2500ms
Target-to-distracter	18 (17%)	58 (55%)	9 (9%)	20 (19%)
Distracter-to-target	25 (27%)	61 (65%)	2 (2%)	6 (6%)

Figure 3.11a depicts latencies within the window of 360ms to 2000ms according to priming condition, where switches toward targets tend to be moderately faster than switches away from targets. Very little difference between priming conditions is apparent. To investigate possible priming effects, a four-way ANOVA was conducted comparing the effect of switch direction (to target, to distracter), prime condition (unrelated, taxonomic, compound) and list (List A, List B, List C) on latency.

Figure 3.11. Latency to switch in three priming conditions (18m)



Notes. Latency measured from target onset to the frame where first saccade away from fixated picture is landed. Latency window: 360ms-2000ms. Whiskers show +/- one standard error. Wilcoxon's tests compare latency to switch in target and distracter direction. ~p<.10. Whiskers show +/- one standard error.

Latency values were not normally distributed within the latency window, but satisfied the symmetry requirements for analysis by univariate ANOVA. The results

of the ANOVA are given in Table 3.11. There were no significant main effects or interactions. A moderate trend toward look direction was observed. The effect did not achieve significance, but the *p*-value of .07 approached the alpha level of .05. As illustrated in Figure 3.11b, when the data were pooled across priming conditions, switches to targets (*M* = 828ms, *SD* = 381ms) were moderately faster than looks to distracters (*M* = 936ms, *SD* = 415ms). This comparison did not achieve significance in a non-parametric Wilcoxon's rank sums test ($z(58, 61) = 1.77$, n.s.). A trend was evident, however, as the *p*-value of .08 approached the alpha-level of .05.

Table 3.11. Three-way ANOVA of latency to switch fixation (18m)

Source (between S)	*df*	*F*	*p*	partial η^2
switch direction (dir)	1	3.25~	.074	.031
prime type	2	1.74	.181	.033
list	2	0.73	.483	.014
dir * prime type	2	0.30	.742	.006
dir * list	2	1.44	.241	.028
prime type * list	4	0.86	.490	.033
dir * prime type * list	4	0.75	.558	.029
S w/in group error	101	(163273)		

Notes. Mean square error given in brackets. ~$p<.10$

This finding demonstrates that while 18-month-olds tended to discriminate named targets from unnamed distracters in the speed of their first switch, the tendency was not reliable when assessed using a previously established latency window. No priming effect was observed in this measure. Given that target preference was reliable for 18-month-olds in both PTL and LLK, it is surprising that target discrimination was unreliable in this measure. It is possible that infants were unable to demonstrate target discrimination due to features of the current experimental design (e.g., confusion caused by prime, non-repetition of items), or because of the number of trials lost to

the lexical exclusion criterion. It is also possible that the previously established latency window does not fit the distribution of latencies produced by infants in the current task. Recall that in Figure 3.10 and Table 3.10, a large proportion of the switches were initiated before the beginning of the latency window (22% of all trials). It is possible that features of the current experimental configuration might generate a faster switching pattern than the latency analysis window employed here.

To investigate the changes in the looking pattern over the course of the trial, the detailed time-course of trials were also investigated. In Figure 3.12a the number of trials in which infants were fixating the target is plotted against the number of trials in which infants were fixating the distracter. These two measures are not entirely dependent as infants can look away from the screen at any time, hence the horizontal asymmetry of the lines (as indicated by the very low values at the onset of picture presentation, and the slight decline in mean values over the second half of the trial).

At the beginning of the trial, no infants were fixating the screen. Review of videos indicated that during the priming phrase infants tended to look at the speaker, which was centrally mounted above the screen. From the onset of the pictures there was a rapid increase in screen looking over the first 500ms of the picture presentation period. Looking to targets and distracters was approximately equal until the onset of the target word at 2500ms. Over the 500ms following the onset of the target word there was a further increase in target looking. There were more infants fixating the target than the distracter for the remainder of the trial.

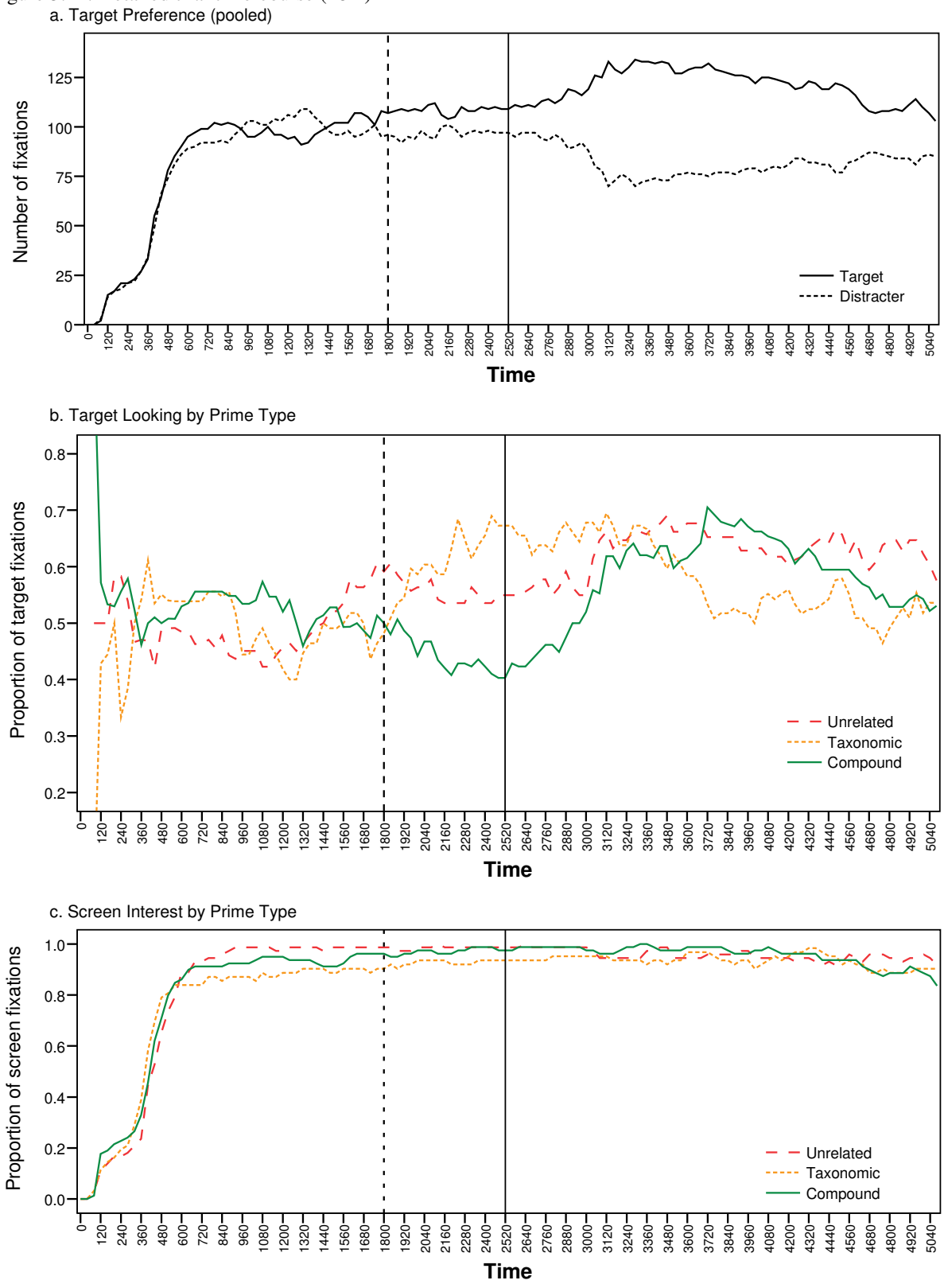
To establish whether the time courses differed in the three priming conditions, the number of target fixations in each frame was converted to a proportion of all fixations, for each priming condition. The proportion of target fixations over the course of the trial are plotted in Figure 3.12b. In this plot it is apparent that the

priming conditions begin to differ during the presentation of the target carrier – some 500ms before the onset of the target word itself. By target onset, the proportion of target fixations was very close to chance in the unrelated prime condition, but appears to be a larger in the taxonomic prime condition, and smaller in the compound prime condition.

This pattern could help to account for the interaction of sex and priming in the macro-level measures, where females showed no change in PTL or LLK following naming in the taxonomic condition. Figure 3.12b illustrates that the lack of change in the macro level measure could have been caused by inflated target looking immediately prior to naming in this condition. In addition, the behaviour of infants during the target carrier may also help to account for the weak target discrimination in the latency measure – if a large number of infants have recently initiated shifts from one picture to another, it is unclear whether they would be able to demonstrate a 'normal' latency response.

It is unclear what might have motivated differences in target looking prior to naming. It is possible that the apparent change is an artefact of the sample population and their pattern of lexical exclusions. It is also possible that differences in the visual interest or comprehension rate of particular target words may have interacted with the pattern of trial dropout to produce the observed pattern.

Figure 3.12. Detailed trial time-course (18m)



Notes. Origin (0ms) indicates the onset of the picture pair. Solid vertical line indicates onset of target word at 2500ms. Dashed vertical line indicates mean onset of target carrier phrase at 1808ms.

An alternative possibility is that co-articulation in the target carrier-phrases may have directed some infants to the target word prior to onset. That is to say, phonological cues to the identity of the target word may have been available prior to the onset of the first phoneme in the target word. Co-articulatory cues, such as partial vowel harmony or assimilation of the penultimate word towards the place of articulation of the first target phoneme, might have provided better clues to the identity of some targets than others. Investigation of the data from 24-month-olds is predicted to help to clarify the issue: If a similar pattern is observed in the performance of older infants (who are expected to provide a larger number of valid trials), it will provide validation of the 18-month-old pattern in a separate population.

A final way of considering the time course of the trial involves the total interest in the pictures areas of the screen. Figure 3.13 depicts the number of trials in which infants are fixating either picture, as a proportion of the number of trials in that condition. In this plot it is clear that screen interest is very similar in all three priming conditions.

Summary: 18-month-olds

For the current analyses, which included only name known test items, 18-month-olds showed reliable target discrimination in two standard macro-level measures of target looking, with increases in the proportion and duration of target looking following the onset of the label. The relationship of the auditory prime to the target did not interact with the effect of the label on the looking behaviour. Nor did it reliably influence looking behaviour in either trial phase. Target discrimination was unreliable in the latency measure, although a trend in the expected direction was evident, with moderately faster switches from distracter-to-target than from target-to-

distracter. No influence of the priming condition was evident in 18-month-olds' latency to switch.

When change in target looking was considered, a prime-by-sex interaction was observed, with females showing less increase in target preference following naming, only in the taxonomic condition. Detailed investigation of the time course of the trials suggested that the priming effects may have been due to a moderate fluctuation in the looking pattern towards the end of the target carrier phrase – prior to the onset of the target word. This pattern of behaviour was not predicted, and is possibly an artefact of the small sample size and data loss due to the lexical exclusion criterion. Investigation of a second infant population for whom the data loss is less extensive is predicted to aid interpretation of this finding.

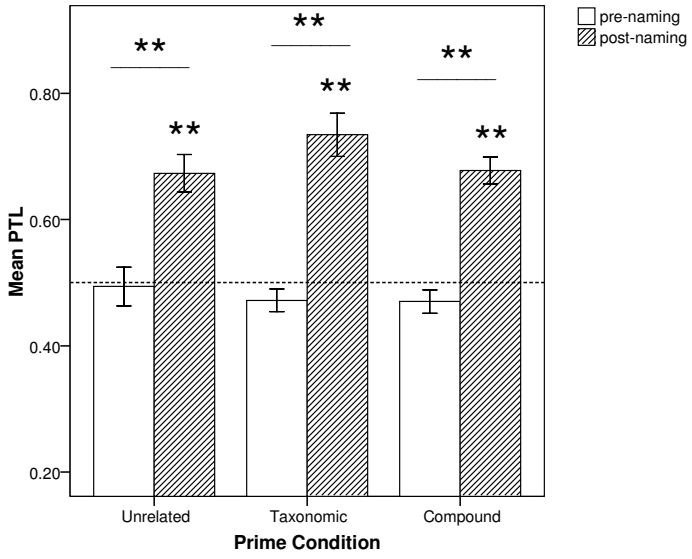
Results: 24-month-olds

Of the 300 trials from 25 participants, the prime was known in 289 trials (96%), and the target was known in 292 trials (97%). There were 283 trials in which both the target and the distracter word were known, meaning that 94% of the original trials were available for analysis. Following lexical exclusions, trials were evenly distributed across the three trial types, with 95 trials in the unrelated prime condition, 94 in the taxonomic prime condition, and 94 in the compound prime condition. All 24-month-olds contributed valid trials to all three priming conditions. Of the 25 participants available for analysis, 10 were male, and 15, female. This asymmetry was expected to complicate straightforward analysis of sex effects.

PTL (24-month-olds)

Participant means for the proportion of target looking (PTL) were calculated separately for the pre- and the post-naming period of trials in each of the three priming conditions. One mean was removed from analysis to achieve normal distribution across all subsets of the data. The mean represented four trials from a single infant in the unrelated priming condition. The value was one standard deviation below the mean, but fell well below the majority of values, generating a negative skew. Figure 3.13 depicts PTL in the pre- and the post-naming period of trials in the three priming conditions, where it is clear that pre-naming values in all three conditions are close to the chance value of .50, and post-naming values are higher. Very little difference between priming conditions is apparent.

Figure 3.13. Pre- and post-naming PTL (24m)



Notes. Proportion of target looking (PTL). Horizontal dashed line indicates 'chance' value of .50. Single sample t-tests comparing PTL to chance are marked above individual bars. Paired sample t-tests comparing pre- and post-naming PTL are marked above pairs of bars. **p<.01. Whiskers show +/- one standard deviation.

To investigate the increase in PTL following naming, a three-way ANOVA comparing the effect of naming (pre-naming, post-naming), prime condition

(unrelated, taxonomic, compound), and stimulus list (List A, List B, List C) was conducted on PTL. Table 3.12 gives the results of the ANOVA, where a main effect of naming was found, with no additional effects or interactions. Participant means were pooled for simple comparison. Mean PTL in the pre-naming phase was .48 (*SD* = .11), slightly below the chance value of .5. This value did not differ significantly from chance (*M* = .48, *SD* = .11, *t*(73) = -1.63, n.s., *d* = -0.19), although the comparison indicated a trend. Mean PTL in the post naming phase was significantly above chance (*M* = .69, *SD* = .14, *t*(73) = 11.67, *p* < .001, *d* = 1.36), and the increase from pre- to post-naming was significant (*t*(73) = 10.22, *p* < .001, *d* = 1.42). This indicates that 24-month-olds reliably increased their target preference following naming, and looked more at the named target following naming, but did not do so before. This behaviour is consistent with lexical comprehension of named items, and is consistent with parental report of test items as 'known'.

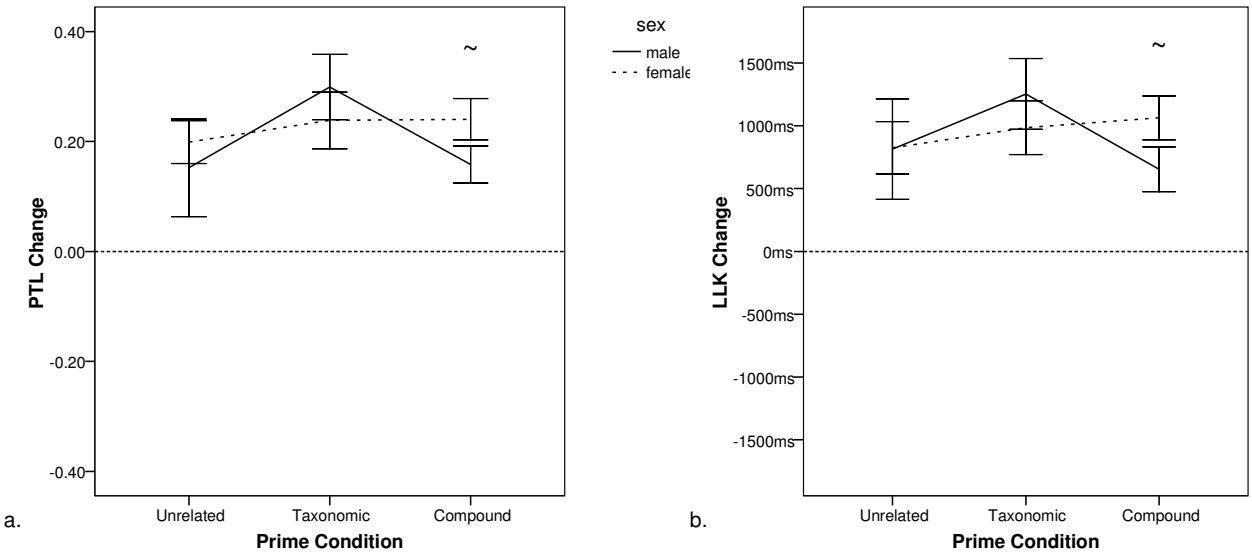
Table 3.12. Three-way ANOVA for PTL (24m)

Source (between S)	*df*	*F*	*p*	partial η^2
prime type	1	0.59	.56	.019
list	2	0.22	.80	.007
prime type * list	2	0.12	.32	.072
S w/in group error	62	(0.02)		
Source (within S)	***df***	***F***	***p***	**partial η^2**
naming	1	99.79**	<.001	0.62
naming * prime	2	0.90	.41	0.03
naming * list	2	0.26	.77	0.01
naming * prime * list	4	1.11	.36	0.07
error (naming)	62	(0.02)		

Notes. Mean square error given in brackets. **p<.01

Change-in-PTL according to sex and prime condition is depicted in Figure 3.14a, where very little difference between the prime conditions is apparent. Analysis of sex was conducted on change-in-PTL (post-naming minus pre-naming), pooled across list. Simple comparisons revealed that no differences were evident between the sexes at the overall level (male: M = .20, SD = .21; female: M = .23, SD = .16; t (72) = 0.54, n.s., d = 0.13), nor within two of the three priming conditions: the unrelated condition: (males: M = .15, SD = .28; females: M = .19, SD = .15; t (22) = 0.53, n.s., d = 0.23); and the taxonomic condition: (males: M = .30, SD = .19; females: M = .24, SD = .20; t (23) = 0.76, n.s., d = 0.28). In the compound priming condition, on the other hand, female 24-month-olds tended to show larger increases in PTL than male 24-month-olds (female: M = 0.24, SD = .14; male: M = .16, SD = .11). This comparison did not reach significance (t (23) = 1.52, n.s., d = 0.63), with a p-value of .10. One-way ANOVAs revealed that change-in-PTL did not differ significantly across priming conditions for females (F(2, 41) = .028, n.s., partial η^2 = .014), nor for males (F(2, 27) = 1.64, n.s., partial η^2 = .109), indicating that the interaction of sex with prime condition should be treated as a minor trend.

Figure 3.14. Change-in-PTL and change-in-LLK, according to sex (24m)



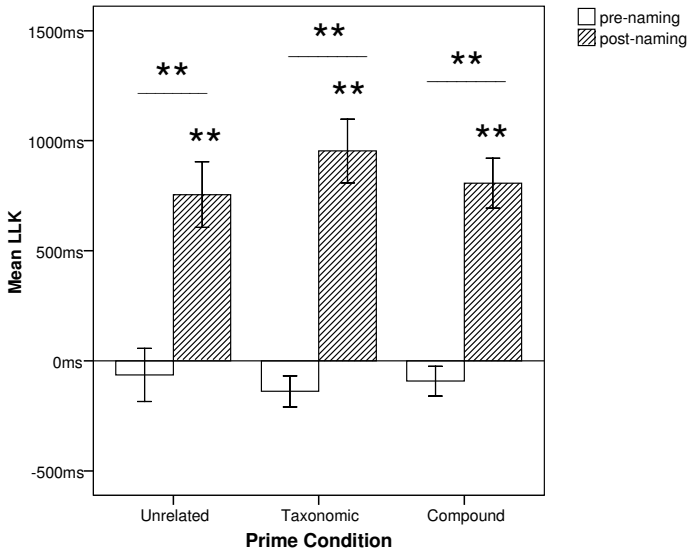
Notes. Change in proportion of target looking (PTL), and change in the difference in the length of longest look (LLK) calculated as post-naming minus pre-naming. Horizontal dashed line indicates chance value for each measure. Independent samples t-tests comparing the sexes are marked above the comparison. ~p = .10. Whiskers show +/- one standard error.

LLK (24-month-olds)

Participant means were created for the difference in single longest look (LLK) in each trial phase, and each priming condition. Figure 3.15 depicts pre- and post-naming LLK in the three priming conditions. As in PTL, pre-naming values in all three conditions are close to the chance value of 0ms, and post-naming values are higher, at approximately 800ms. Again, little difference is apparent between the three priming conditions.

A three-way repeated measures ANOVA was conducted, comparing the effect of naming (pre-naming, post-naming), prime condition (unrelated, taxonomic, compound), and stimulus list (List A, List B, List C) on LLK. Table 3.13 gives the results of the ANOVA, where a main effect of naming was found, with no additional effects or interactions.

Figure 3.15. Pre-naming and post-naming LLK (24m)



Notes. Difference in length of longest look (LLK). Origin indicates chance value of 0ms. Single sample t-tests comparing LLK to chance are marked above individual bars. Paired sample t-tests comparing pre- and post-naming LLK are marked above pairs of bars. **p<.01. Whiskers show +/- one standard deviation.

Simple effects demonstrated that LLK pooled across the priming conditions did not differ significantly from the chance value of 0ms (M = -98ms, SD = 444ms, $t(74)$ = -1.92, n.s., d = -0.22), although there was a trend towards distracter looking (LLK below 0ms) in the pre-naming phase, with a p-value of .059 approaching the alpha-level of .05. Post-naming LLK was significantly above chance (M = 839ms, SD = 681ms, $t(74)$ = 10.67, p < .001, d = 1.23). Post-naming LLK was also significantly higher than pre-naming LLK ($t(74)$ = 9.56, p < .001, d = 1.32). This pattern mirrors the findings in the other sections, and is consistent with reliable lexical comprehension.

Table 3.13. Three-way ANOVA for LLK (24m)

Source (between S)	*df*	*F*	*p*	partial η^2
prime type	1	.08	.92	.003
list	2	.02	.98	.001
prime type * list	2	.72	.59	.043
S w/in group error	63	(312549)		
Source (within S)	***df***	***F***	***p***	**partial η^2**
naming	1	85.29**	<.001	.575
naming * prime	2	0.37	.69	.012
naming * list	2	0.01	.99	<.001
naming * prime * list	4	1.01	.37	.065
error (naming)	63	(358251)		

Notes. Mean square error given in brackets. **p<.01

Post hoc analysis of sex revealed the same pattern of change-in-LLK as the analysis of change-in-PTL, and is depicted in Figure 3.14b. There was no significant difference between males and females when the data were pooled across priming condition (males: M = 907ms, SD = 952ms; females: M = 957ms, SD = 763ms; t (73) = 0.25, n.s., d = 0.06). When considered in the different priming conditions, no difference was found in either the unrelated prime condition (males: M = 814ms, SD = 1264ms; females: M = 824ms, SD = 807ms; t (23) = 0.02, n.s., d = 0.01) or the taxonomic prime condition (males: M = 1253ms, SD = 890ms; females: M = 985ms, SD = 832ms; t (23) = 0.77, n.s., d = 0.32). There was a weak tendency in the compound priming condition, for females to show a greater LLK increases than males (males: M = 655ms, SD = 564ms; females: M = 1062ms, SD = 675ms; t (23) = 1.58, n.s., d = 0.66), but the comparison did not reach significance (p-value of .13), and was not supported by ANOVA comparison of the priming conditions within each sex: No difference was found between the priming conditions for females (F (2, 42) = 0.82,

n.s., partial $\eta^2 = .04$), nor for males (F (2, 42) = 1.64, n.s., partial $\eta^2 = .11$). Note that the moderate sex trend for 24-month-olds is different to the sex differences observed for 18-month-olds: 18-month-old females produced significantly smaller changes in target preference in the taxonomic condition; 24-month-old males tended to show less change in target preference in the compound prime condition.

Latency & trial time course

Of the 283 valid trials (prime and target known), there were 17 trials in which infants were not fixating either picture at the onset of the target word. These trials were removed from analysis. At target onset, infants were fixating the target in 159 trials (60%) and the distracter, in 107 trials (40%). In total, there were 64 trials in which the infant did not shift to the other picture before the end of the trial. There were thus 266 trials in which infants made a shift from one picture to another before the end of the trial. Of these shifts, 102 were from distracter to target, and 100 were from target to distracter.

Table 3.14 gives the distribution of latency values according to an age-appropriate latency analysis window established by Fernald, Perfors and Marchman (2006). When 24-month-olds were fixating the distracter picture at the onset of the target word, almost all infants shifted to the other picture before the end of the trial, and a large proportion (91%) shifted before the end of the latency window. By comparison when 24-month-olds were fixating the target, a moderately large proportion (40%) did not shift their gaze to the other picture before the end of the trial, and the proportion of infants who had shifted by the end of the latency window was much smaller in this direction (71%). As was observed in the latency distribution

for 18-month-olds, this pattern of more and faster distracter rejections is consistent with lexical comprehension.

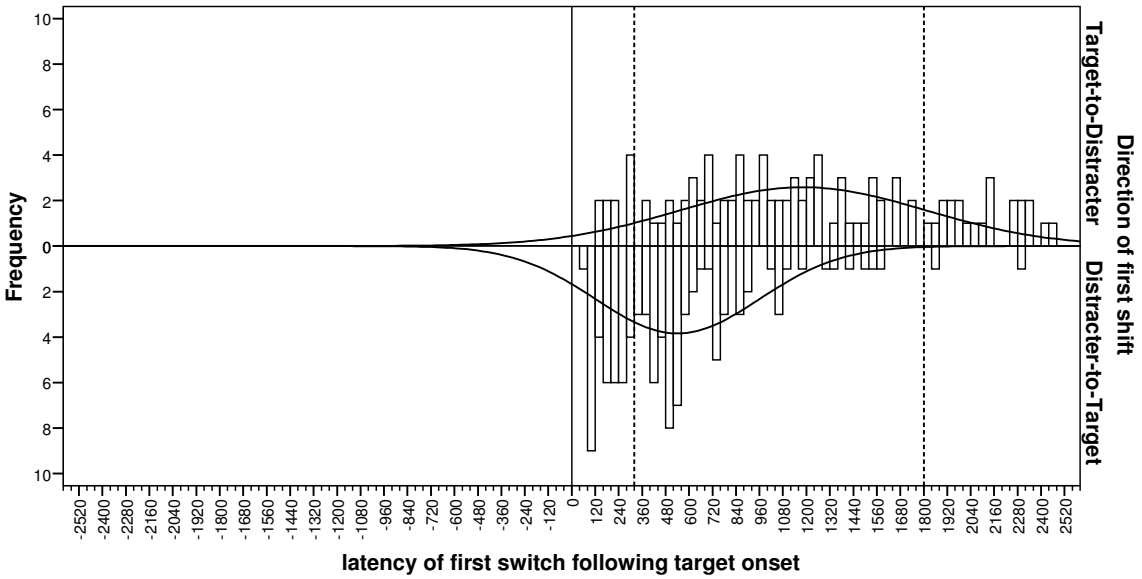
Table 3.14. Number of shifts following target onset (24m)

	Early shift <300ms	Latency Window 300-1800ms	Late shift >1800ms	No shift >2500ms
Target-to-distracter	8 (5%)	63 (40%)	24 (15%)	64 (40%)
Distracter-to-target	38 (35%)	60 (56%)	2 (2%)	7 (7%)

Figure 3.16 shows the distribution of the latency of first fixation (time taken to land a fixation on the alternative picture, measured from the onset of the target word), with the latency window of 300ms to 1800ms marked. The data in Table 3.14, and visual inspection of latency distribution in Figure 3.16, show that saccades from the distracter to the target tend to be landed faster than saccades from the target to the distracter. This appears to be the case within the 300-1800ms latency window as well as overall.

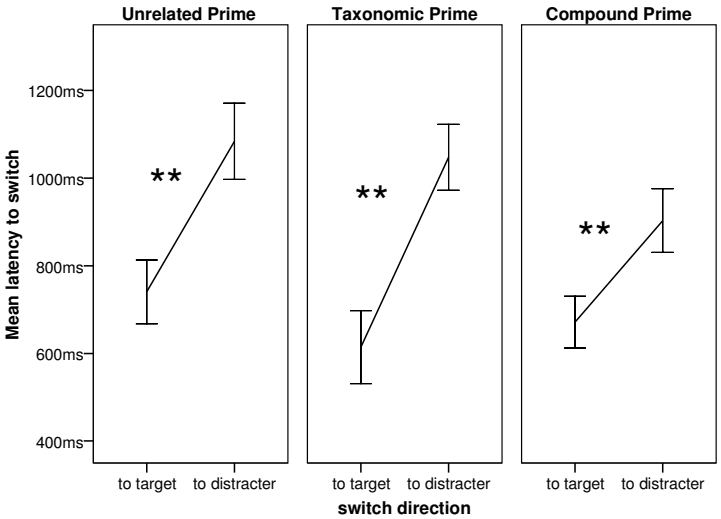
Mean latencies within the window of 300ms to 1800ms are plotted in Figure 3.17, where it is clear that latency to switch from a distracter to a target is quicker than from a target to a distracter. This is clearly the case in all priming conditions. The slope of the line appears to differ slightly between priming conditions, suggesting a possible interaction between priming condition and switch direction.

Figure 3.16. Histogram of latency over the course of the trial (24m)



Notes. Time scale depicts entire picture-presentation period. 0ms indicates onset of target word. Dashed vertical lines at 300ms and 1800ms depict the Latency window of Fernald, Perfors & Marchman (2006). Normal curves plotted over whole data set.

Figure 3.17. Latency to switch in three priming conditions (24m)



Notes. Latency window: 300ms to 1800ms. Wilcoxon's tests comparing latency of switches in each direction are marked above each comparison. **p<.01

Within the latency window, the data did not achieve Gaussian distribution in the distracter direction, even with the removal of a number of upper outliers. However, as the univariate ANOVA is considered to be robust to deviations from

normality (SPSS, "Spss 16.0 for windows", 2006), a three-way ANOVA was conducted comparing the effect of switch direction (target-to-distracter, distracter-to-target), priming condition (unrelated, taxonomic, compound) and stimulus list (List A, List B, List C), on the latency to land a saccade following the onset of naming. The data were sufficiently symmetrical to consider by ANOVA. The results of the ANOVA are summarised in Table 3.15, where there is a main effect of look direction, indicating that switches from distracter-to-target were landed faster then switches from target-to-distracter (distracter-to-target: M = 685ms, SD = 326ms; target-to-distracter: M = 1008ms, SD = 383ms; $t(135) = 5.29$, $p < .001$, $d = 0.91$). The ANOVA also revealed a strong trend towards a main effect of prime type – while the statistic did not strictly achieve significance, the p-value of .053 was extremely close to the alpha-level of .05.

Table 3.15. Three-way ANOVA for Latency (24m)

Source	*df*	*F*	*p*	partial η^2
switch direction (dir)	1	666.01**	<.001	.167
prime type	2	23.81~	.05	.048
stimulus list	2	3.00	.80	.004
dir * prime	2	0.22	.35	.018
dir * list	2	1.07	.26	.022
prime * list	4	1.35	.80	.014
dir * prime * list	4	0.42	.22	.047
S w/in group error	119	(129209)		

Notes. Mean square error given in brackets. **p<.01

When data were analysed separately according to direction of switch, no significant effects of prime condition were observed. The difference between the latency of target and distracter switches was larger in the unrelated prime condition (344ms) and the taxonomic prime condition (433ms) than in the compound prime

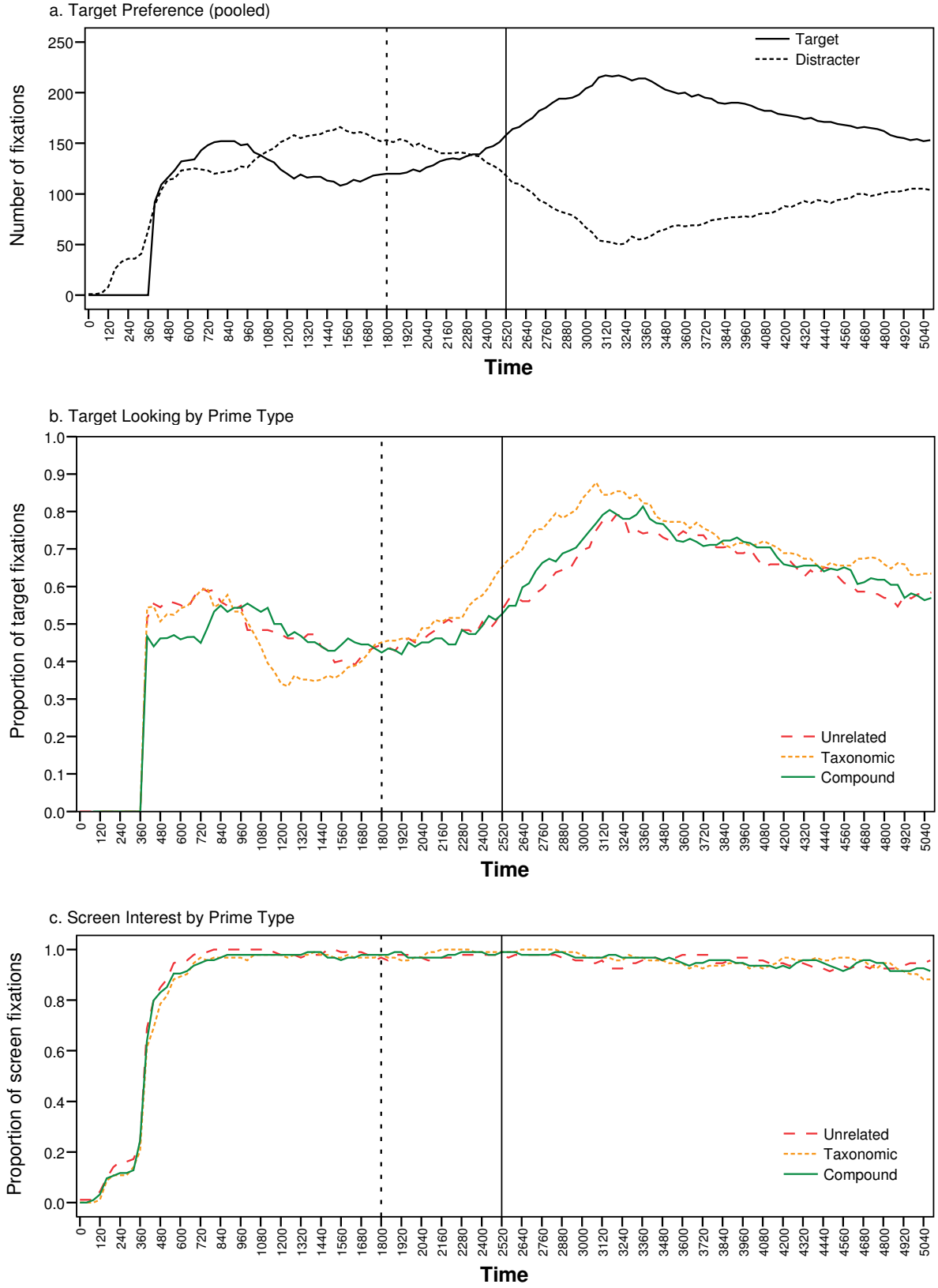
condition (232ms). Note also that switches from the target to the distracter appear to be slightly faster in the compound prime condition than switches in the other two priming conditions, and switches from the distracter to the target appear to be slightly slower in the unrelated priming condition than switches in the other two conditions. Simple comparisons between conditions revealed no further trends or effects.

As more than 30 percent of the distracter-to-target switches occurred before the beginning of the latency window (cf. Table 3.14), it is important to note that the same pattern of results was found when all valid trials were included in analysis. The large number of switches falling before the latency window will be addressed in the discussion.

Figure 3.18 depicts the time course of trials for 24-month-olds. In Figure 3.18a, fixations to targets and distracters are plotted. The main increase in target fixations begins slightly before the onset of the target word. The difference between target and distracter values at target onset is no larger than the variance observed earlier in the pre-naming phase. It is thus reasonable to interpret the naming effect as time-locked to the onset of the target word. Comparing time course data across the age groups (cf., Figure 3.12a), the 'naming bloom' following target onset is larger and more consistent for the 24-month-olds. The gradual reduction in target preference towards the end of the trial is also more pronounced.

In Figure 3.18b, the proportion of target fixations in each prime condition is shown over the course of the trial. Unlike the time courses of the 18-month-olds, no large differentiation of the priming conditions is apparent prior to naming. Target fixations in the taxonomic prime condition appear to be slightly higher than the other two conditions, but this could be due to the slightly lower trend earlier in the trial (before the onset of the target carrier).

Figure 3.18. Detailed trial time-course (24m)



Notes. Origin (0ms) indicates onset of pictures. Solid vertical line indicates onset of target word at 2500ms. Dashed vertical line indicates mean onset of target carrier phrase at 1808ms.

Figure 3.18c depicts the general screen interest over the course of the trial – total picture fixations as a proportion of the trials in each condition. Very little difference is apparent between the priming conditions. Both age groups show the same pattern of screen interest, with a sharp increase following picture onset, then consistently high interest throughout the remainder of the trial. Older infants have a much steeper increase in picture fixations at the onset of pictures.

Summary: 24-month-olds

24-month-olds demonstrated the expected pattern of lexical comprehension in the two macro-measures of target preference. The relationship of the auditory prime to the target did not influence the change in looking behaviour following naming. A weak trend was observed for sex to interact with prime type in PTL and LLK – such that in the compound prime condition, male 24-month-olds showed a moderately smaller increase in target preference than female 24-month-olds. This trend was different to the sex-by-prime interaction observed in the 18-month-olds – in which 18-month-old females showed a change in target preference in the taxonomic condition. The differences between the age groups suggest that the sex-by-prime interactions are most likely an artefact of unbalanced samples: The number of males and females differed in the general sample for each age group; In addition sex was not controlled in the allocation of stimulus individuals to stimulus lists, resulting in different numbers of boys and girls encountering different stimulus combinations.

The micro-level latency measure confirmed the reliable effect of target discrimination, with faster switches away from targets (mismatch rejections) than

switches away from targets. No effect of prime was observed in this measure, or in the detailed time courses of looking behaviour over the trial.

General discussion.

The results described in the first half of the chapter demonstrated that a priming phase mentioning an absent referent did not disrupt 18-month-olds' normal ability to show preference for a named target in a primed adaptation of a standard IPL task. The combination of behavioural data with parental report allowed systematic investigation of the item level accuracy of two measures of lexical comprehension for infants – IPL and CDI. Both tools are typically used to assess performance averaged over a number of infants, and a number of items, and there has been little previous investigation into whether either tool can deliver accuracy at the level of individual words known by individual infants. When Houston-Price, Mather and Sakkalou (2007) demonstrated British parents' underestimation of the number of words infants could identify, the authors acknowledged that one reason may have been a mismatch between the 'level' of comprehension assessed by parents, and the difficulty of the task. For example, understanding the word 'sock' could be interpreted as either 'can your infant pick out a sock from her trousers?' or 'does your infant know enough about 'sock' to know that it is not a lorry?'

In the current experiment, a match between the tools was found, providing information about item-level accuracy and about the 'level' of comprehension assessed by British parents: British parents correctly identified which words would attract behaviour consistent with lexical comprehension, and which words would not; British parents interpret 'understanding' as the ability to pick out a named picture

from a pair of semantically related pictures, with only a single presentation of unfamiliar exemplars, in an unfamiliar environment, even though a potentially distracting auditory prime preceded the basic task. That is to say, British parents set a high threshold for comprehension, marking only those words which are sufficiently 'well known' to succeed in a difficult task. Infants may well be able to identify (or infer) the referents of less-well-known words in easier tasks (for example, tasks including features such as multiple stimulus repetition, dissimilar picture pairs, longer trials, or feedback). However, the match found between CDI report and IPL performance in the current experiment adds a degree of external validity to *each* of the measurement tools, and the 'level of comprehension' at which they are most sensitive.

It is also possible that item familiarity played a role in which items were identified by infants. In the Houston-Price et al. task, all stimuli were selected from a pool of items which had been reported to be encountered at least weekly. In the current experiment, comprehension rates of above 50% provided an approximation of item familiarity at this age (on the logic that frequently encountered items were more likely to be understood). However, as no direct measure of object familiarity was included in the current design, a more systematic investigation of these elements would be needed to clarify the precise contribution of each (for example, a one-shot IPL design pitting object familiarity against distracter-type).

In the second half of the chapter, analyses focused on the three priming conditions built into the design. At the time this experiment was conducted, it was the first reported attempt to incorporate lexical priming into the body of an infant IPL task. The adaptation was thus 'minimal', with a priming phrase added to the beginning of a standard IPL trial. Given the differences between 'automatic' and

'strategic' priming effects observed over different timescales in adults (Neely, 1991), it was unclear whether priming would be evident 2500ms after the offset of the prime word. It was also unclear whether any effect of the prime would survive the picture presentation which began between the offset of the prime and the onset of the target word.

The relationship of the auditory prime to the named target picture did not create a systematic influence on looking behaviour in this minimally-primed experiment. Moderate priming-by-sex interactions were observed in both age groups. For younger infants, a priming effect appeared to lie in the pattern of looking immediately preceding target naming (more looks to targets in the taxonomic condition). The pre-naming increase in target looking prior to naming is difficult to interpret, as it is unclear how infants could infer the correct target prior to the onset of the target word in the taxonomic condition. Recall that until the onset of the target word, the prime (e.g., *elephant*) is equally taxonomically related to both the target (e.g., *dog*) and the distracter (e.g., *monkey*). Three possible explanations for early target looks in this condition are outlined below.

Lexical status & familiarity: Within picture pairs, targets tended to have higher lexical comprehension rates (cf. Chapter 2, Table 2.7) than distracters (target mean: 78%; distracter mean: 66%). If lexical comprehension rate approximates familiarity in the population, it is possible that following the onset of a target carrier, some infants correctly guessed that the more familiar item was about to be named. However, if this were the case, it is unclear why infants would not have shown the same predictive behaviour in the unrelated priming condition (where both pictures were *equally unrelated* to the prime), or why the tendency was in the *other direction* for the

compound prime (which shared an equal taxonomic relationship with both pictures, but had an additional associative relationship with the target).

Co-articulation: As the target carrier phrases were recorded whole (not spliced together), it is possible that phonological cues discriminating the onset of the target from the onset of the distracter may have been co-articulated during the target carrier phrase. Given that target carriers were counterbalanced across priming conditions, there should be no prime-dependent acoustic cues available to the infants. It is possible that viewing pictures related to a recently heard prime enhances an infant's ability to generate an 'internal phonological representation' (Mani, 2008) of a picture label, but it is again unclear why this pattern would be expected to differ between the taxonomic and the compound prime condition, as priming might be expected to enhance internal representations in both related conditions.

Data loss: Given that there were less females than males in the 18-month age group, and there were less trials available in the taxonomic priming condition, it is possible that the observed prime-by-sex interaction was the result of random noise generated by the small number of infants, or the pattern of data loss generated by the items individuals were reported not to understand. Data loss may have been gender specific, but it could also be the result of random chance.

The trend towards a sex-by-prime interaction in the 24-month-old data was much more straightforward. All infants demonstrated target recognition in all priming conditions, but there was a subtle suggestion that the compound prime generated a greater change in target naming for girls than for boys. While it is tempting to

interpret this as evidence that girls show sensitivity to associative primes in a task where boys do not, there was no systematic effect of priming for boys or for girls. The trend was also weak and unreliable, making it difficult to consider as anything more than a subtle variation in lexical comprehension rates.

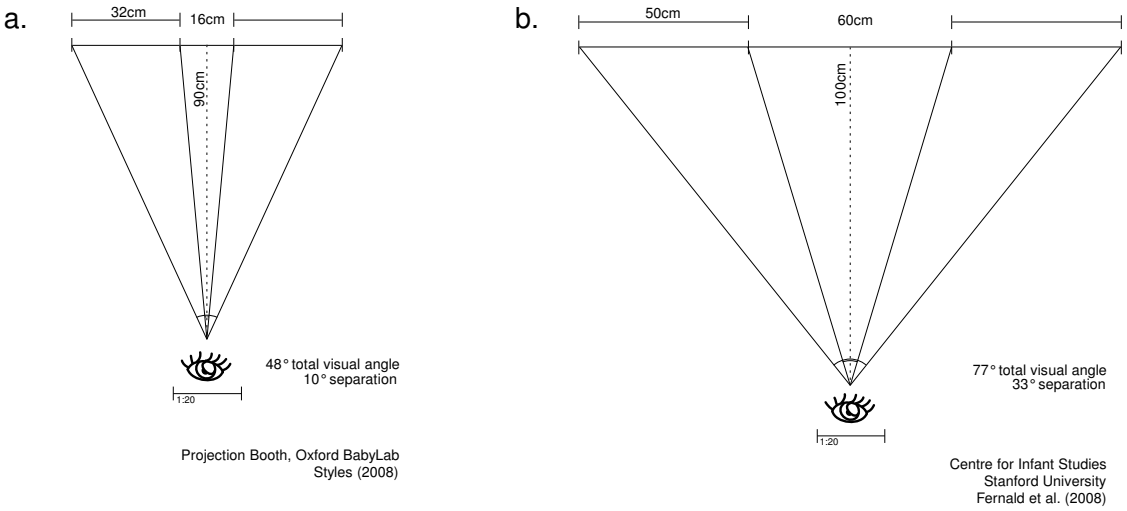
Reaction time measures

During the planning phases of the current study, it was assumed that if priming effects were not observed at the macro-level (accumulated looks across large-scale trial phases), that priming effects may be evident in micro-level (reaction-time data indexing speed of linguistic processing). Using age-appropriate analysis windows which are well established in the literature, infants in both age groups switched to targets faster than to distracters, although the trend was unreliable for the younger infants. The priming condition did not systematically influence latency of switch for either age group. It was interesting to note that these standard analysis windows did not appear to fit the distribution of the latency data in either age group. A surprisingly large proportion of the latency switches in the distracter condition fell before the before the minimum latency cut-off (18-months: 25%; 24-months: 38%), suggesting that an earlier cut-off may have been more appropriate here. An earlier latency cut-off may also have improved the reliability of the naming effect for younger infants.

There are two possible reasons why an earlier latency window may have been better suited to the current task. Firstly, there some feature of the stimulus design which may have caused infants to shift their eyes more quickly following the onset of the target word. For example, the task might have been perceived as faster paced, or the natural co-articulation of target words in their carrier phrases may have allowed faster lexical access to target words. Secondly, there are large differences in the

screen configurations of the laboratory where the latency windows were established, and the laboratory where the current experiment was carried out. Figure 3.16 illustrates the different screen configurations. As depicted in Figure 3.16b, the pictures presented at the Center for Infant Studies at Stanford University were separated by a large gap, which occupied approximately 33° of the infant's visual angle, with the outer edges of the displays extending to create a total visual angle of 77°. By contrast, in the Projection Booth at the Oxford BabyLab (Figure 3.16a), the presentation areas of the projection screen were separated by only 10°, and extended to 48°. Overlaying figures 3.16a and 3.16b would show the Oxford BabyLab picture *midlines* falling within the inner edge of the Stanford monitors.

Figure 3.19. IPL screen configuration in two laboratories



Notes. Figures a. and b. depict picture size and separation, as well as observer distance from the screen in two laboratories. Both diagrams are drawn to the same scale, and show the visual angle to the outer edge of the pictures and the visual angle of the gap.

In the current study, infants could easily initiate an eye movement from one picture to the other, many without also moving their head. In the Stanford configuration, on the other hand, images inner edges are 33° apart, a separation which commonly produces head movement (J Stephen Reznick *et al.*, 2000). The wider

visual separation of the pictures and the wider angle required for a first fixation may well have generated slower latencies in the IPL tasks used to generate the standard latency windows. Given the complicated pattern of pre-naming target switches apparent in the current experiment, it was decided not to conduct detailed further latency analyses for this experiment. In future priming studies, however, RT windows will be manually established on the basis of the data produced by the particular screen and stimulus configuration.

Interim conclusions

Parental report accurately assesses which words infants know well enough to identify in a difficult task. Parental report can thus be used as an item-level exclusion criterion for whether a trial should be included in priming analyses, on the logic that if a word is not known to an infant, then no valid lexical relationship can exist between that word and any other. Although the comprehension 'threshold' identified by parents may be unnecessarily strict for identifying a target in a less-difficult IPL task, the threshold acts as a conservative assessment of words which are well-known, and perhaps more likely to be integrated with others into an interconnected semantic system.

Although priming effects were observed in some subsets of the data, they were weak and difficult to interpret in this minimally primed adaptation if IPL. The difficulty of interpretation lies in two principal confounds caused by presentation of the pictures in the inter-stimulus interval (ISI) between the prime and the target word: the long ISI, and the difficulty of interpreting priming effects if looking differs from chance in the pre-naming phase. Further studies aimed to develop a primed IPL

method with a shorter prime-to-target ISI, to increase the likelihood that the potentially short-lived effects of spreading activation in the infant lexicon might be observed, and to develop an IPL method in which the prime and the target can be presented without any intervening stimuli which might confound direct spreading activation in the lexicon.

Chapter 4

Words-before-pictures: Picture duration



In the previous chapter, no priming effects were found when the priming phase was added to the beginning of a typical intermodal preferential looking (IPL) task. It was not possible to interpret whether the lack of priming was due to lack of connectivity in the infant lexicon, or lack of sensitivity in the testing method. As discussed previously, the pre-naming phase of the minimally-primed IPL trial introduces a complication to the lexical priming paradigm, as presenting pictures prior to the onset of the target word disrupts the standard sequence of stimulus presentation. The confound is two-fold: Firstly, semantic information generated by the pictures (including the distracter picture) may contaminate direct prime-to-target spreading activation. Secondly, and perhaps most importantly, the long prime-to-target latency vastly reduces the likelihood of automatic priming, as this form of spreading activation disperses rapidly (Neely, 1991). What is needed, therefore, is an IPL task in which the prime-to-target inter-stimulus interval (ISI) is very short (to preserve any short-lived effects of spreading activation), and in which the ISI is not interrupted by the presentation of other stimuli. In such an arrangement, the target word must therefore be presented before the onset of the test pictures: An experiment with no pre-naming phase. This chapter describes an experiment designed to establish whether infants are able to demonstrate lexical comprehension in an IPL task with no pre-naming phase, and to establish appropriate timing intervals for stimulus presentation.

Introduction

The IPL task used by Golinkoff et al. (1987) presented auditory stimuli before the onset of image pairs. Following Reznick's introduction of a 'baseline' for picture preference prior to test (J Stephen Reznick, 1990), subsequent studies (e.g., Fernald et

al., 1998; Schafer, 1998; Schafer & Plunkett, 1998; Schafer et al., 1999; Swingley et al., 1999) have incorporated a visual baseline into the body of the IPL trial, by presenting the pictures before the onset of the target word. In these contexts, it is also referred to as a pre-naming phase (e.g., Houston-Price et al., 2007; Mani & Plunkett, 2007; e.g., Meints et al., 1999), as relative preference before and after naming can be compared. Within-trial 'baselining' offers two statistical advantages over the original trial design of Golinkoff et al. Firstly, if target and distracter looking are equal prior to the onset of naming, above chance performance following the onset of naming must be due to the effect of the name. Alternatively, if pre-naming values deviate from chance, but a *change* in behaviour is observed following naming, the change provides evidence of the effect of naming.

An additional line of reasoning treats the pre-naming phase as picture familiarisation time. It is often assumed (e.g., Fernald et al., 1998; Swingley et al., 1999) that infants respond in an 'informed' way when they hear a label following familiarisation. Although it is not explicitly stated, such a view supposes infants are able to a) use the familiarisation time to explore both pictures, b) accurately form a short term association between each picture and its location and c) access this memory trace when naming occurs. Further, this account implies that infants *need to know* the location of the to-be-named picture, in order to respond correctly. However, in a paired-picture study by Swingley and Fernald (2002), 24-month-olds given more than 3000ms of familiarisation time did not show evidence of memory for both items during test. Rather, following the onset of naming, infants demonstrated high-speed judgements concerning the label's 'match' or 'mismatch' to the currently fixated image, without recourse to their memory of other elements in the display. Infants demonstrated lexical comprehension by maintaining their fixation longer when the

word labelled the fixated object, than when the word labelled another name-known object. However, the speed of their gaze shift following mismatching labels was equally rapid regardless of whether the target object was depicted elsewhere in the display. The finding that infants will shift their gaze to *any alternative picture* in the presence of a mismatching label suggests that 24-month-olds do not use memories formed in the familiarisation period when making high-speed lexical judgements. It follows that familiarisation time *may not be necessary* for infants to demonstrate high-speed online judgements during lexical processing.

The current experiment aims to establish whether lexical comprehension can be observed in measures at both the micro- and the macro- level of measurement in a study where no familiarisation period was incorporated into the body of the trial, and no priming phrase is included. To quantify the effect of un-baselined picture presentation, two additional stimulus controls were included. Firstly, the visual 'interest' of all pictures was monitored in stimulus design, such that relative interest levels could be counterbalanced, and visual salience could be assessed following test. Secondly, each picture occurred equally often as a named target and as an unnamed distracter. As an experimental pilot for a priming study, the words-before-pictures experiment was a 'one-shot' design (see Chapter 2), hence counterbalancing of target status was achieved between subjects.

The second aim of the words-before-pictures study was to identify timing intervals for use in later priming studies. Two timing scenarios were created. If the previously used pre-naming period (2500ms) was *necessary* picture exploration time, then a long picture duration might be required to observe target discrimination. That is to say, infants may require the first 2500ms for picture exploration before beginning to show target recognition. A long timing condition was thus created in which the

duration of picture presentation was 5000ms long, equivalent to the entire looking period in the previous experiment. It is also possible that infants need only a short amount of time to identify the referents of known words using their high-speed 'matching' skill. A short timing condition was created, in which picture presentation lasted for 1500ms, which was substantially shorter than the post-naming test-phase in the previous experiment. Note that the pace of the short timing condition is very fast, with only 2500ms per trial. The timing conditions varied between subjects.

Given previous reports of infants' online 'matching' skills (Swingley & Fernald, 2002), it was predicted that infants would be able show target discrimination in a 'one-shot' IPL task with no familiarisation time. However, it was unclear whether infants would perform differently in the two timing conditions. For example, it was possible that the short timing condition might be too short for target recognition to be reliable. It was also possible the timing conditions might affect different age groups in different ways, in particular, younger infants may be able to cope less well with the time pressure imposed by the short timing condition. Older infants, on the other hand, might be somewhat bored by the long picture presentation, and fail to show systematic behaviour.

It was also unclear which measures of target looking would be the most sensitive to target discrimination in the new stimulus arrangement. As an exploratory study, five measures of target looking were assessed, covering both micro- and macro- levels of description. It was predicted that the micro-level measures of reaction time might be more appropriate in the short timing condition, while the macro-level measures (concerning relative picture preference) might be more systematic in the long timing condition. In this regard, item variability was assessed, particularly with regard to which measures of target discrimination might be most

affected by the difference between items. Finally, to monitor whether developmental differences between the age groups might be due age-related changes or to language-specific development, individual vocabulary size was monitored in relation to target discrimination.

Method

Participants were recruited as in the previous experiment, and primary caregivers were asked to fill out the British CDI (Hamilton et al., 2000) in the week before their visit to the laboratory. Infants were allocated to a timing condition and stimulus list on arrival. Forty 18-month-olds were tested (19 males, 21 females; mean age: 18.1 months, range: 17.2 to 18.8). Two additional infants were removed and replaced at the time of test, one for equipment failure, and one for the parent filling out the CDI in an uninterpretable manner. Three further 18-month-olds were removed without replacement for failure to complete the task. Forty 24-month-olds were tested (15 males, 25 females; mean age: 24.0 months, range: 23.0 to 24.7). Two further 24-month-olds were removed and replaced at the time of test, following equipment failure. Six participants were later removed without replacement, one for failure to complete, and five for parental failure to return appropriately completed CDI forms. Following exclusions, there were 37 18-month-olds and 34 24-month-olds for whom both behavioural data and vocabulary information were available.

Total receptive CDI scores were checked against previously collated norms (see Chapter 1), and infants' comprehension of test items was assessed. Following preliminary analysis, two further eighteen-month-olds were removed from analysis for extremely low receptive CDI scores (in the 5th percentile). In addition, one 18-

month-old was removed from analysis when it was observed that she knew only one of the target words in her list. Exclusions were unevenly distributed across subject groups, resulting in a different number of participants in each timing condition. As shown in Table 4.1, there were 34 final participants in each age group, and across ages, 32 and 36 in each timing condition. It was decided to continue with the full sample, as creating a sub-sample with matched group sizes sample would have necessitated removal of an additional 14 infants (20%). The uneven distribution was expected to result in unequal variance between the participant conditions, making direct comparison somewhat complex. However, the value of retaining all analysable data was deemed more important in this experimental pilot.

Table 4.1. Number of participants in each subject group

	Presentation Time		
	Short (1500ms)	Long (5000ms)	Total
18-month-olds	19	15	34
24-month-olds	13	21	34
Total	32	36	68

Materials

Sixteen words were selected for use as test items, all of which were known by more than 50% of 18-month-olds (cf., Table 2.2 and Appendix B). Two stimulus lists were created in which eight 'target' items were depicted on-screen and named. The other eight items were unnamed 'distracters' which appeared alongside the target picture during the test phase of the trial. Across lists, picture pairs were yoked, and each item was used equally often as a named target and as an unnamed distracter. The picture pairs were: *frog-chair*, *butterfly-key*, *rabbit-bread*, *flower-telephone*, *duck-tree*, *bee-spoon*, *bird-clock*, *fish-television*.

Targets shared no taxonomic or associative relationships with distracters. Picture-pairs also shared no phonological onsets or rhymes, with the exception of the pair *flower-phone*[6]. Each item occurred only once in a list. Target side was counterbalanced within and across lists. Trial order was randomised on presentation. Stimulus lists are given in Appendix G.

Visual stimuli were high quality digital photographs, selected as typical exemplars of target and distracter words by a native speaker of English, and confirmed by two native speakers of British English. Pictures were presented on a 10% grey background. Pictures used as stimuli are given in Appendix H. In each picture pair, one image was judged to be "more visually interesting to babies" by two experienced infant language researchers. These images were classified as 'high interest' relative to the other picture in their pair. The relative interest of target pictures was counterbalanced within and across lists. Seven of the eight high interest targets were animals (with the exception of *flower*). Seven of the eight low interest targets were man-made artefacts (with the exception of *tree*). Pictures could thus be classified according to relative within-pair interest (high, low), animacy (animate, inanimate), kind (natural, artefact).

Audio stimuli were created in a single recording session, in a sound-attenuating booth on DAT tape sampling at 44.1 kHz. A minimum of three tokens of each auditory stimulus were produced by a female native speaker of British English, using high-affect child directed speech. The single best token of each stimulus was manually selected for clarity, typicality and affect, and edited using visual waveform editing software to remove head and tail clicks. The naming phrase was spliced together to contain an attention-getter and the target word, separated by 80ms (e.g.,

[6] The onsets have partial overlap (/f/ versus /fl/). An error occurred during stimulus recording, where the speaker accidentally recorded three tokens of the abbreviation *phone* for the planned stimulus *telephone*. The effect of individual items on the final result is quantified in the results section.

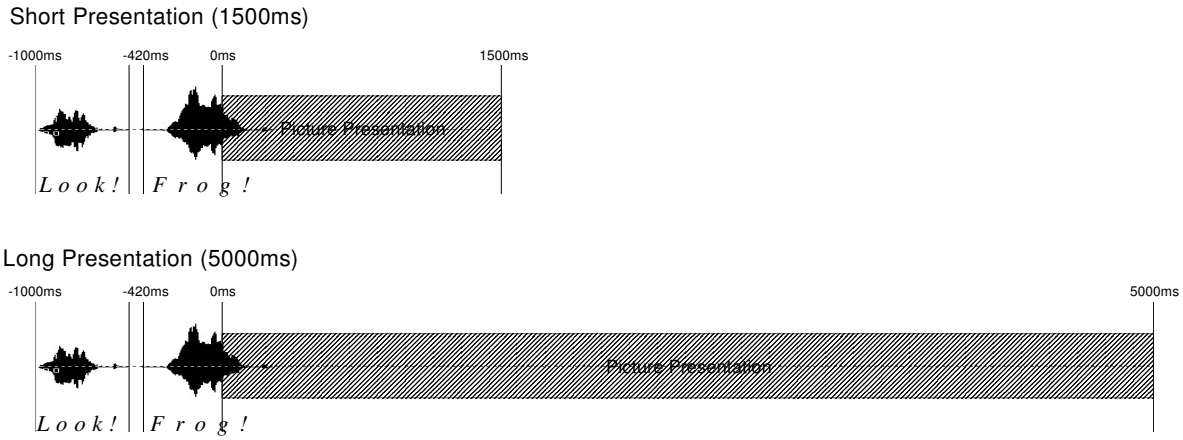
'Look! Frog!'). The duration of 'Look!' was 472ms. The duration of target words ranged from 393ms for the shortest word, 'bee,' to 901ms for 'television'. Target words had a mean duration of 603ms (*SD* = 163). The duration of individual test words is given in Table 4.2.

Table 4.2. Duration of auditory stimuli

Stimulus Type	Word	Duration (ms)
Attention Phrase	look	472
Test Item	fish	524
"	flower	708
"	frog	697
"	key	454
"	phone	614
"	rabbit	677
"	spoon	842
"	television	901
"	tree	530

Notes. Onsets and offsets of tokens were established using visual waveform editing software.

Figure 4.01. Schematic of trial timing



Notes. Picture presentation is treated as the origin of the test phase (time = 0ms). Attention-phrase and target word are separated by 80ms. Wave-forms from actual stimulus tokens, as produced by GoldWave visual waveform editing software (1995, v5.10).

Procedure

The experimental procedure was the same described in the previous experiment, with the exception of trial timing, which is illustrated in Figure 4.01. Each trial was manually initiated when the infant's attention was centred on the screen. While the screen was blank, the target phrase began (e.g., *Look! Frog!*). 420ms after the onset of the target word, picture presentation began. Note that all words other than the shortest (*Bee!*) overlapped with the onset of picture presentation. Picture-pairs remained onscreen for 1,500ms, or 5,000ms, depending on subject group. The entire procedure lasted less than one minute for the short presentation group, and less than two minutes for the long presentation group. Each infant saw eight trials from a single stimulus list. As in the previous experiment, infants sat approximately 90cm from a screen with a display area 79cm wide. Each picture was 32cm wide. The outer edges of the pictures described a visual angle of approximately 48º. The inner edges were separated by a gap of 15cm (10º). Appendix E contains a photograph of the testing booth, and a schematic of the screen configuration used in this experiment.

Scoring & measures

Recording and off-line scoring were conducted as in the previous experiment. Both large- and small-scale timing measures were again calculated: two macro-level measures assessed the relative preference for the named target picture during the course of the whole trial, and two micro-level measures of reaction time (RT) assessed the speed of infant eye-movements. The calculation of new micro-level measures is informed by statistical techniques from the visual expectation (VEx) paradigm developed by Haith and colleagues (e.g., Haith *et al.*, 1988).

Macro-measures:

As in the previous chapter, the proportion of target looking (PTL) is the total amount of time spent looking at the target (T) as a proportion of the total amount of time spent looking at both pictures (T+D). It can be represented as T/(T+D). The difference in longest look (LLK) is the difference, in milliseconds, between the single longest look to the target (t) and the single longest look to the distracter (d). It can be represented as t-d.

Micro-measures:

The time-to-fixate is a reaction time (RT) measure, which describes the time from the onset of the pictures to the beginning of the first fixation in a trial. As each trial begins with a central fixation, time-to-fixate is a measure of how long it takes infants to land a fixation on the left or the right picture (whichever was fixated first), following the onset of visual stimuli. As trials began 1000ms before the presentation of pictures, it was possible for infants to launch a predictive saccade prior to picture onset. Trials in which infants had already fixated the left or the right picture area of the screen at picture onset were excluded from RT analysis.

Time-to-disengage is a second RT measure, which describes the amount of time taken to initiate a *saccade away from* the first fixated picture. While time-to-fixate is a static measure, taken from the onset of picture presentation, time-to-disengage is a *dynamic* measure of reaction time, with the measurement beginning at the fixation of the first picture. This measure is equivalent to the *duration* of the first fixation. The 'linking hypothesis' for micro-level measures is that if infants are unable to 'map' or 'match' the spoken word to a subsequently presented picture, the speed of

their reactions to both targets and distracters will not be expected to differ. However, if infants are sensitive to a match between the target word and the target picture, RTs to targets would be predicted to differ systematically from RTs to distracters.

RT measures of different kinds are common in IPL research. Different uses of terminology, however, can cause some degree of confusion. For example, 'latency' in work from the Centre for Infant Studies at Stanford University (e.g., Fernald et al., 1998; Swingley et al., 1999), refers to the amount of time taken to *launch a saccade* away from the picture fixated at the onset of naming, while 'latency' in the Oxford BabyLab (e.g., Mani & Plunkett, 2007) refers to time taken to *land a fixation* on the alternative picture. Meanwhile, in a task with no familiarisation phase, 'latency' in work from the Infant Language Project at the University of Delaware (Golinkoff et al., 1987) refers to the amount of time taken to fixate either picture in the display, regardless of whether the fixation was the first or the second in the trial. In the current study, the terminology 'time-to-fixate' and 'time-to-disengage' was chosen to remove potential ambiguity. Both measures refer to the first fixation in a trial (time-to-fixate is equivalent to speed of first fixation; time-to-disengage is equivalent to duration of first fixation), where the direction of the look (target, distracter) is considered as an independent variable.

In addition to assessing target discrimination in the four measures outlined above, comparisons were conducted at each level (micro-, macro-), to assess relationships between the measures, and to assess whether the developmental variable of vocabulary size was correlated with performance. The discrete variance predicted by individual participants and items was also quantified for each measure, in order to assess which measures might be most robust to stimulus effects in future studies.

Results

Eighteen-month-olds were reported to know a mean of 197 of the 416 words in the British CDI, with a range of 79 words to 381 words. Twenty four-month-olds were reported to known a mean of 314 words, ranging from 170 to 413. Older infants thus knew significantly more words than younger infants ($z(34, 34) = 5.46$, $p < .001$). Within each age group, the infants allocated to the two timing conditions did not differ in CDI score. 18-month-olds were reported to know a mean of 12 of the 16 words used as stimuli (range: 3 to 16), with a mean of 6 of the 8 words named as targets in their stimulus list (range: 2 to 8). 24 month-olds were reported to known a mean of 16 stimulus words (range: 11 to 16), including a mean of 8 of the 8 words used as targets (range: 6 to 8). Infants allocated to different timing conditions did not differ in the number of stimulus words known.

Of the 544 trials available for analysis (18m: 272, 24m: 292), only those trials in which the target was reported as known were included, resulting in removal of 69 trials from 18-month-olds (25%), and 12 trials from 24-month-olds (4%).

Macro-level analysis

Participant means were created for analysis of the macro measures.

Proportion of target looking (PTL)

For analysis of PTL, two outliers (more than two standard deviations below the mean) were excluded from analysis. The outliers were means from two 18-month-olds in the short timing condition, and represented the means of two trials from one participant and three trials from another. Mean values for PTL are depicted in Figure

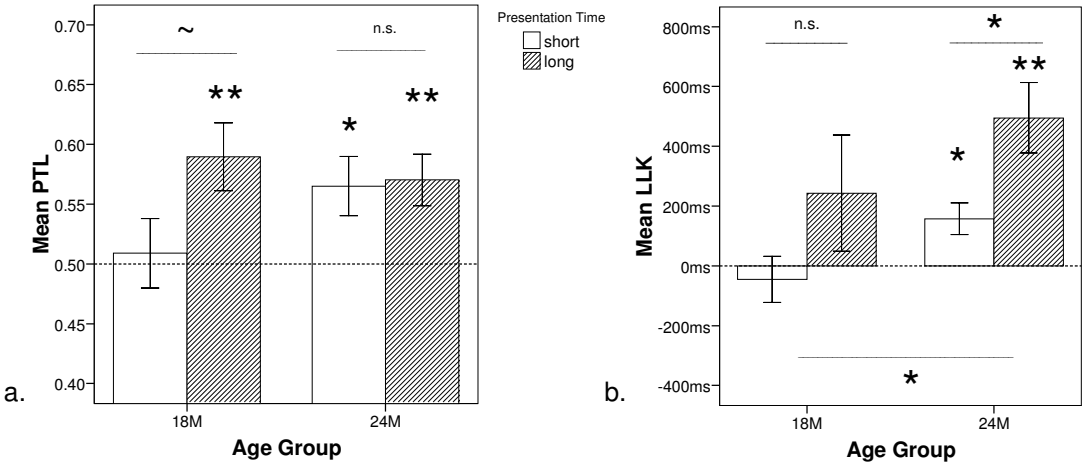
4.02a, where PTL is generally above the chance value of 0.5, with target preference evident in all conditions except one (18-month-olds in the short presentation condition). A 2x2 ANOVA was conducted comparing the effect of presentation time (short, long) and age group (18m, 24m) on PTL. No main effects or interactions were found. PTL values were thus pooled across age and timing groups. The pooled mean PTL was .56 (*SD* = .11), and was significantly above chance ($t(65) = 4.38$, $p < .001$, $d = 0.54$). This result indicates that infants reliably discriminated named targets from unnamed distracters, and the direction of the effect was consistent with that predicted by lexical comprehension.

As shown in Figure 4.02, however, when PTL values for the subgroups were assessed separately, target discrimination was reliable in only three of the four subgroups. 18-month-olds in the short timing condition did not demonstrate target discrimination. Although no interaction between age and timing was observed in the ANOVA, for 18-month-olds PTL was larger in the long timing group than in the short timing group (short: *M* = .51, *SD* = .12; long: *M* = .59, *SD* = .11). This comparison did not achieve significance ($t(30) = 1.97$, n.s., $d = 0.72$), however it approached the alpha level of .05, with a *p*-value of .058. Cohen's Standard for *d* also rates this comparison as a medium sized effect. This finding suggests that younger infants behaved moderately differently in the two timing conditions. Their failure to demonstrate target discrimination in the short timing condition is especially relevant given the larger number of subjects in this group.

These findings demonstrate that while infants *overall* were able to demonstrate target discrimination in a one-shot IPL task with no pre-naming phase, in a standard measure of IPL, younger infants tended not toshow reliable target discrimination when they were under the time-pressure imposed by the short presentation time. They

were, however, able to perform marginally better in the long timing condition, where they were able to discriminate targets from distracters. Older infants, on the other hand, appeared to be at ceiling in this measure, showing no enhancement in performance in the long timing condition.

Figure 4.02. PTL and LLK, according to age and timing condition



Notes. a. depicts the proportion of target looking (PTL). b. depicts the difference in the length of longest look (LLK). Horizontal lines indicate chance value in each measure. Single sample t-tests compare values in each subject condition to the chance value. Independent samples t-tests compare values in different subject conditions, as indicated by horizontal bars. * $p<.05$, ** $p<.01$ ~ $p<.10$ Whiskers indicate +/- one standard error.

Difference in longest look (LLK)

Following visual inspection of the distribution of LLK, three outlying participant means were removed, all of which were one standard deviation above the mean and were separated by a gap of 400ms from the next nearest value. They represented the means of two 18-month-olds (five trials each), and one 24-month-old (eight trials), all of whom were in the long timing condition, and produced mean LLK values above 2000ms. Following the removal of these three means, the remainder were normally distributed.

Results of LLK are plotted in Figure 4.02b, where target recognition is generally above the chance value of 0ms, although it was higher in the older age

group and higher in the long timing condition, and did not differ from chance for 18-month-olds in the short timing condition. Data were not sufficiently homogeneous across the subgroups to conduct comparison by ANOVA, due to generally smaller variance in the short timing condition, and generally larger variance in the younger age group. LLK data were instead subjected to simple comparisons. Mean LLK was 41ms for infants in the short presentation condition (*SD* = 300ms), and 391ms for infants in the long presentation condition (*SD* = 599ms). The difference between timing conditions was significant ($t(45.3) = 3.01$, $p = .005$, $d = 0.76$). With a mean LLK of 72ms for younger infants (*SD* = 526ms), and 352ms for older infants (*SD* = 439ms), LLK was smaller for 18-month-olds than for 24-month-olds ($t(63) = 2.33$, $p = .02$, $d = 0.59$). Within age group, 18-month-olds did not differ according to timing condition. For 24 month-olds, LLK was larger in the long timing group than in the short timing group ($t(31) = 2.33$, $p = .03$, $d = 0.84$).

Thus only older infants demonstrated reliable target identification in LLK, and were able to demonstrate this in both timing conditions. Unlike in PTL, 24-month-olds' LLK values were significantly larger in the long timing condition than in the short timing condition. While it would be typical to interpret this difference as enhanced performance, if LLK is considered as a proportion of the total presentation time, the apparent difference is no longer evident: According to look direction, longest looks differed by 158ms in the short timing interval, equivalent to 11% of the presentation time, and by 495ms in the long timing interval, equivalent to 10% of the presentation time. Thus, although LLK values were larger, it would not be appropriate to interpret this finding as enhanced performance.

Summarising target discrimination in the macro-level measures, PTL and LLK produced the same pattern of results for older infants: Target discrimination was

observed in both measures, with no meaningful difference in performance between timing conditions (the apparent LLK difference for older infants being an artefact of the different trial durations). However, PTL and LLK did differ for younger infants, who showed reliable target discrimination in PTL in the long timing condition but not in the short timing condition. This finding suggests that target discrimination was somewhat easier for young infants in the long timing condition, and that PTL was the more reliable measure of target discrimination in the current experimental procedure. 18-month-olds' failure to show target discrimination in LLK in the long timing condition also suggests the possibility that the long condition may be *too long* for optimal LLK discrimination at this age, as the 5000ms presentation period may contain more so-called 'blank stares' (Aslin, 2007). This possibility is supported by the pattern of exclusions, where the three extreme means removed from LLK were from the long timing condition, and were in excess of 2000ms in duration.

Relationship between vocabulary size and macro-measures

Vocabulary size, as assessed by CDI, is general considered to be a valuable index of linguistic and cognitive development in infants (Phillip S Dale et al., 1989; Pan *et al.*, 2004; Reese & Read, 2000)[7]. It is thus important to assess whether precise age or linguistic development *per se* is responsible for the enhanced performance of the older infants. If a positive linear relationship is found between CDI scores and macro-level measures, it suggests that vocabulary size, not age, is responsible for the differences in performance. However, if no linear relationship is evident, it demonstrates that younger infants are performing in the same manner as older infants,

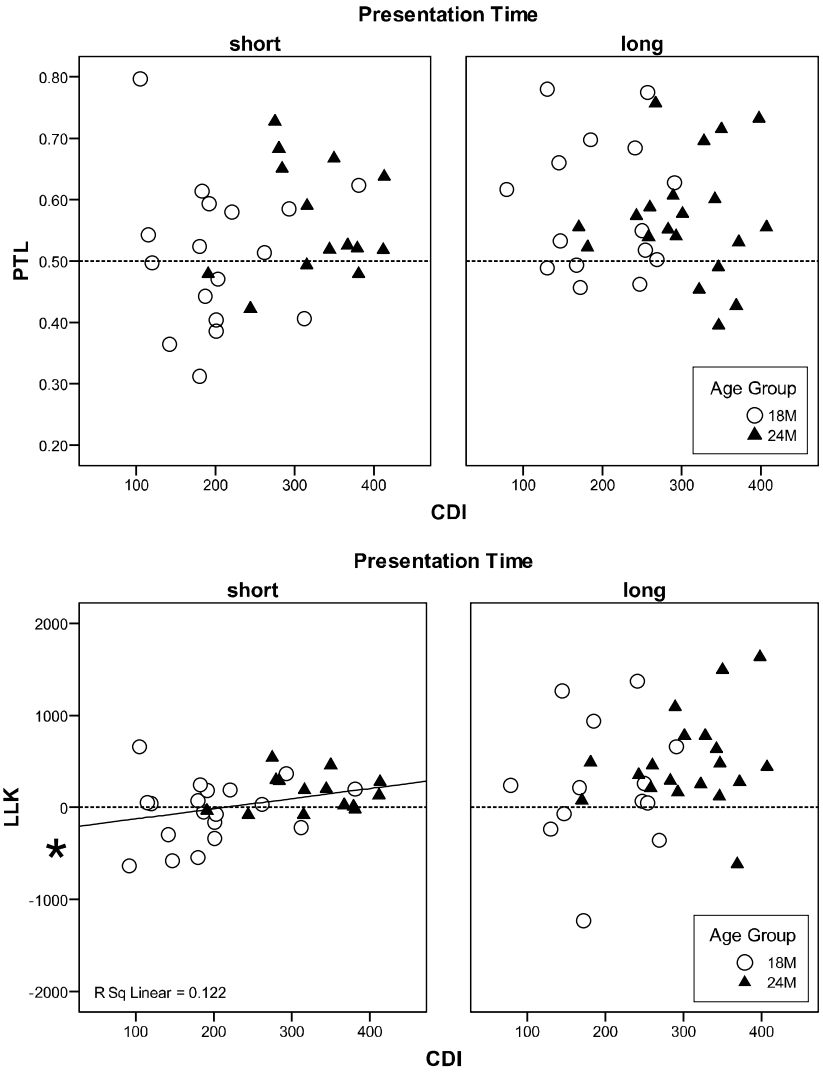
[7] Note that these studies use different CDIs in their metric, and vary according to whether lexical comprehension or production was included in the tool, and whether syntax was also a component.

but that the larger vocabularies of older infants may be driving the difference in performance due to the 'cleaner signal' created by including more trials in analysis.

Age groups were pooled for assessment of vocabulary size, as infants in the two age groups were reported to have comprehension vocabularies spread across the scale. Vocabulary is plotted against the two macro-measures in Figure 4.03, where it is clear from the scattered distribution in the upper plots that there is no linear relationship between vocabulary size and PTL. Indeed, infants' CDI comprehension scores were uncorrelated to their mean PTL ($\rho(66) = .08$, n.s.).

In the lower plots, representing LLK, the tighter clustering of values in the short timing condition suggests a possible relationship between the two measures. Individuals' CDI comprehension scores were correlated to their mean LLK at the overall level ($\rho(65) = .29$, $p = .02$). Further investigation of the timing conditions demonstrated that the correlation was lead by values in the short timing condition, where a trend towards correlation was evident ($\rho(33) = .33$, n.s.) with a *p*-value of .06 approaching the alpha-level of .05. No correlation was evident in the long timing condition ($\rho(32) = .28$, n.s.). This finding suggests that it is only under some degree of time pressure that general linguistic development (as indexed by vocabulary size) co-varies with performance in this measure.

Figure 4.03. Macro- measures plotted against comprehension vocabulary



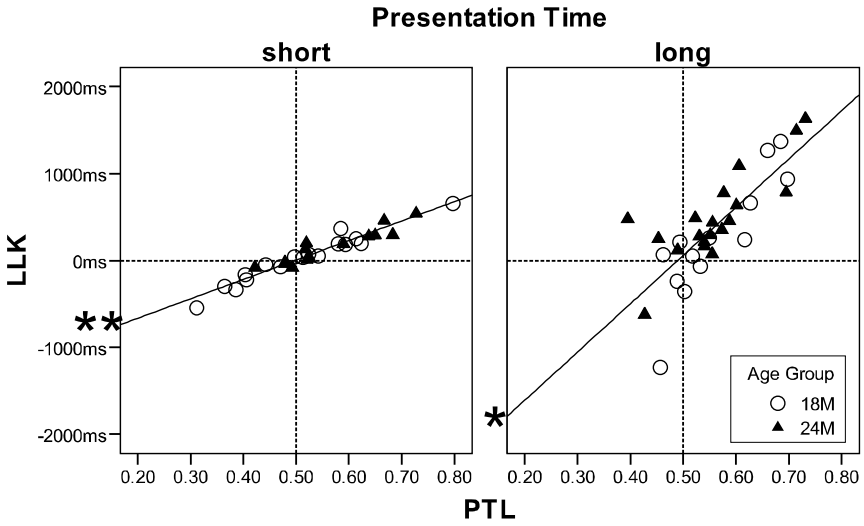
Notes. Upper plots compare CDI score and proportion of target looking (PTL). Lower plots compare CDI score and difference in longest look (LLK). CDI score is given as number of words marked as 'understood' or 'understood and also said'. Horizontal lines indicate chance value for each measure. Solid lines show linear relationship between variables. ~p<.10.

Relationship between macro-measures

Performance for each infant was compared across the two macro-level measures to establish whether the two measures indexed different performance styles, or whether infants who performed well in one measure also performed well in the other. PTL is plotted against LLK in Figure 4.04, where the means for both ages overlap through the distribution, and it is apparent that infants with a higher mean

PTL also had a higher mean LLK. Both measures exhibited more variability in the longer timing condition. Performance in PTL for each infant was correlated with their performance in LLK ($r(63) = .75$, $p < .001$). The correlation was highly significant in both timing conditions, although it was tighter in the short timing condition than in the long timing condition (short: $r(31) = .96$, $p < .001$; long: $r(31) = .80$, $p < .001$).

Figure 4.04. PTL plotted against LLK, according to age and timing condition



Notes. Dashed lines indicate chance value in each measure. Solid lines indicate linear fit for total. *p<.05, **p<.01.

Correlation between measures is sometimes taken to indicate 'converging evidence' from two measures. However, note that in the short timing condition, if infants spend more than 750ms in a single fixation then they have less than half of their time remaining to fixate the other picture. Thus, both the direction and the magnitude of PTL and LLK may be determined, to a large extent, by the duration of a single fixation. In the current experiment, the extremely tight correlation between PTL and LLK in the short timing condition suggests that the relationship between these measures may well be the outcome of statistical dependence within each trial rather than 'converging evidence'. This linear dependence may be responsible for

younger infants' failure to demonstrate reliable macro-level target discrimination in the short timing condition: Fixation speed is known to increase with age (Canfield *et al.*, 1997; Frick & Colombo, 1996; J Stephen Reznick et al., 2000), meaning that younger infants are more likely to be 'caught short' by the end of the trial. Future studies will attempt to remove this tight dependency by avoiding extremely short presentation durations.

Micro-measures

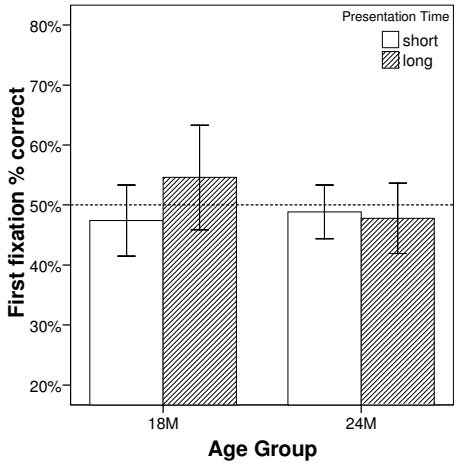
As in the analysis of macro-level measures, trials in which infants were not reported to know the target word were again excluded. For reaction time analysis, also excluded were trials in which infants were fixating the left or right picture area at onset (18m: 32 trials (16%), 24m: 75 trials (29%)).

Percent of correct first fixations

Before beginning analysis of RT measures, a check was conducted concerning the direction of the first fixation. Participant means were calculated to indicate the percentage of trials in which infants' first fixation was (correctly) directed towards the named target picture. The data were normally distributed overall, but were unsuitable for analysis by ANOVA (due to larger variance in the long timing condition, and in the younger age group). The percent of first fixations correctly landing on the target picture is depicted in Figure 4.05, where there is very little difference between sub-groups, and the mean values are very close to the chance value of 50%. As simple comparisons revealed no reliable differences between age groups or timing conditions, data were pooled. The first fixation was correct in 49% of all trials

(SD = 26%), a value which did not differ significantly from chance 50%. This finding suggests that the first fixation is not systematically directed.

Figure 4.05. Percent of first fixations correct, according to age and timing condition



Notes. Dashed line indicates chance value (50%). Whiskers show +/- one standard error.

Time-to-fixate

Following the VEx method described in detail by Canfield, Smith, Brezsnyak and Snow (1997), visual inspection of histograms was used to identify the time period in which time-to-fixate was normally distributed. Identifying the beginning of a Gaussian distribution allows exclusion of those fixations which were 'pre-emptive' or 'expectancy' based (Haith et al., 1988), rather than responsive. In addition, identifying the end of a normal distribution in the right-hand tail allows removal of trials in which infants were unusually slow in responding. Following removal of anticipatory fixations, Canfield et al. (1997) employed a cutoff for the right hand tail of two standard deviations above the mean. In the current study, the upper cutoff was established visually instead. One reason for this was the relatively long presentation period during which late fixations could be initiated. In Canfield et al's study, trials were 1000ms in duration, thus automatically excluding fixations slower than 1000ms.

In the current study, by comparison, even the short presentation condition allowed fixations to be initiated up to 50% later than in the former study. Even a small number of such late fixations would dramatically skew the mean in the current task, making it a poor measure of central tendency for selecting further outliers. Thus the visual method was employed for both the lower and the upper cutoffs.

Note that unlike the 'pre-screening' process described by Fernald et al. (2008), or the 'traditional cutoffs' described by Rose, Feldman, Jankowski & Caro (2002) and Reznick, Chawarska & Betts (2000), no a-priori patterns of fixation were deemed 'extreme' prior to manual inspection. In eye tracking and VEX research it is commonly understood that apparently small differences in experimental setup can create differences in RT – a point highlighted by Canfield et al. (1997). Factors affecting eye-movement RT include the degree of peripherality (Rayner, 1978), and the continuity of central stimuli (Rose et al., 2002), as well as the level of ambient light during test (J Stephen Reznick et al., 2000). In addition, within a given experimental configuration, trials may differ systematically in their RT distribution across participant conditions. In particular, it was predicted that time pressure may induce differences in the speed of first fixations, and that the speed of first fixations may also vary according to the direction of the first look. Data were visually inspected separately for each look direction and timing condition. However, as VEX researchers working with younger infants have found a decline in age-related changes in purely visual RTs toward the end of the first year (Canfield et al., 1997; J Stephen Reznick et al., 2000; Rose et al., 2002), data were pooled across age group for the visual inspection of data.

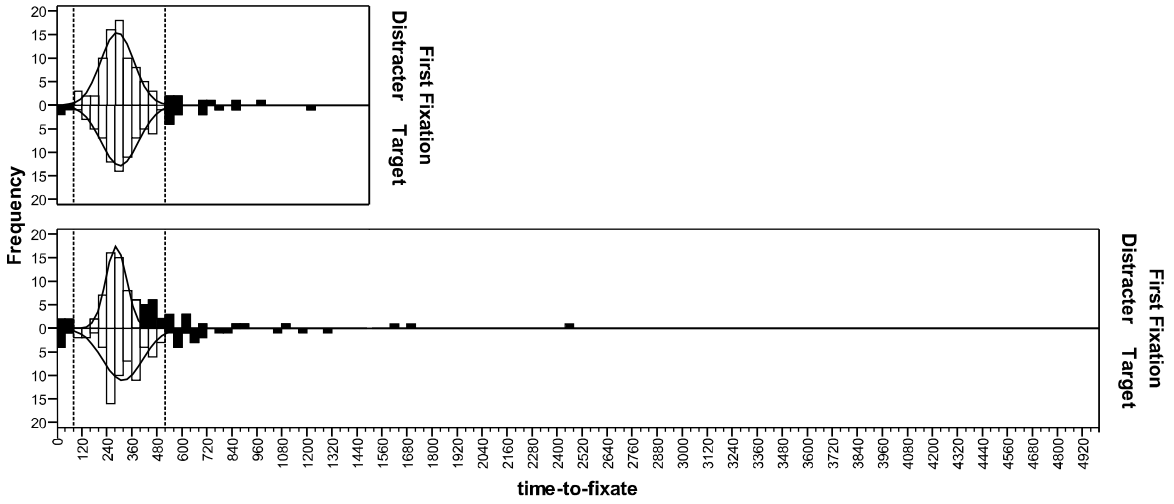
Histograms showing the distribution of time-to-fixate over the trial are given in Figure 4.06. Following Canfield et al.'s method (1997) histogram bins are

equivalent to the duration of a single frame of video (40ms), as this represents the accuracy of the frame-by-frame coding system (+/- 20ms scientific error). A window of 80ms to 520ms was selected to exclude pre-emptive and late fixations. 15 of the 345 valid trials were pre-emptive (4%), and 48 were 520ms or slower (14%). Following exclusion these trials, data from three of the four subsets were normally distributed. Distracter fixations in the long timing condition remained abnormally distributed, and further inspection of this subset revealed a clear bimodal distribution with a small late peak beginning after 400ms (note three dark bars inside the analysis window for distracter fixations in the long timing condition). Removal of fixations slower than 400ms in this subset (13 trials) resulted in normal distribution in all subsets of the data.

Time to fixate in the two age groups is depicted in Figure 4.07a, where target fixation takes approximately the same amount of time for both ages, but older infants exhibit a faster time-to-fixate for distracters. Time-to-fixate was not suitable for analysis by ANOVA due to differences in variance between the subsets of the data. Simple effects were calculated instead. The mean time-to-fixate when the picture first fixated was the target was 309ms (SD = 92ms), and when the first fixated picture was the distracter, 288ms (SD = 68ms). The difference between these trials was significant (adjusted $t(252.5) = 2.11$, $p = .04$, $d = 0.24$)[8], indicating general discrimination of named targets from unnamed distracters. Note that fixations to targets were *slower* than fixations to distracters.

[8] Where marked, t-value and degrees of freedom are adjusted for unequal variances between groups.

Figure 4.06. Histogram of time-to-fixate over the course of a trial



Notes. Distribution of all valid trials. Vertical dashed lines indicate manually selected window (80ms to 520ms). Dark bars represent excluded trials. White bars represent the trials included in analysis, with normal curves plotted.

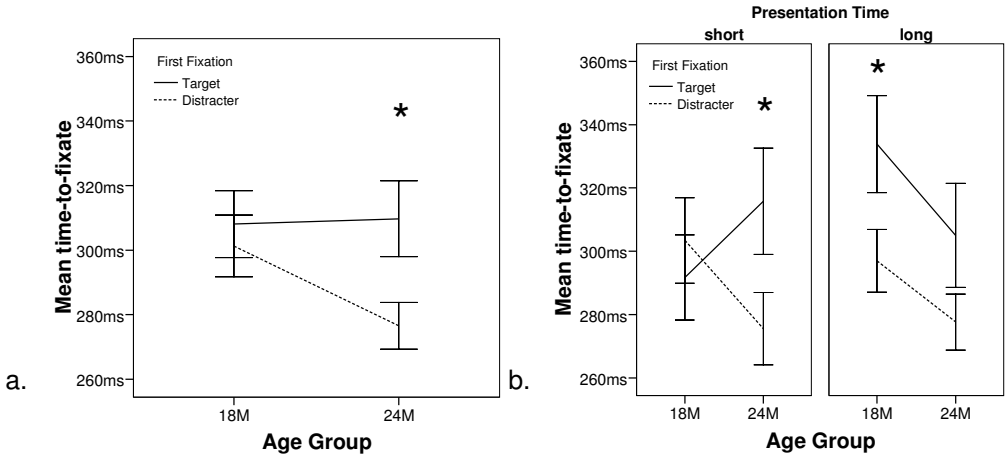
According to age group, the older infants demonstrated reliable target discrimination (adjusted $t(166.0) = 2.14$, $p = .02$, $d = 0.36$), while the younger infants did not ($t(126) = 0.47$, n.s., $d = 0.08$). According to timing condition, the long timing condition produced reliable target discrimination (adjusted $t(101.4) = 2.23$, $p < .05$, $d = 0.38$), while the short timing condition did not ($t(147) = 0.84$, n.s. , $d = 0.14$).

Figure 4.07b shows time-to-fixate separated into the four subject groups, where the interaction between timing and age is clear. Younger infants in the long timing condition perform quite like older infants, and exhibit target discrimination. When analysed separately, 24-month-olds showed the same pattern in both timing conditions, but target discrimination was only reliable in the short timing condition.

Thus, even though the percentage of correct first fixations suggested unsystematic saccadic planning, the speed of the first fixation suggests some degree of target sensitivity prior to the first fixation. Time-to-fixate was *slower* in the target direction than in the distracter direction in half of the experimental subgroups, and this finding indicates that infants who were slower to initiate their first fixation in a

trial were more likely to correctly fixate the target. This implies a processing cost in target identification, which results in a speed/accuracy trade-off. This result is somewhat surprising, as infants would need to use peripheral visual information *prior to fixation* in order demonstrate this effect.

Figure 4.07. Time-to-fixate, according to age and timing condition



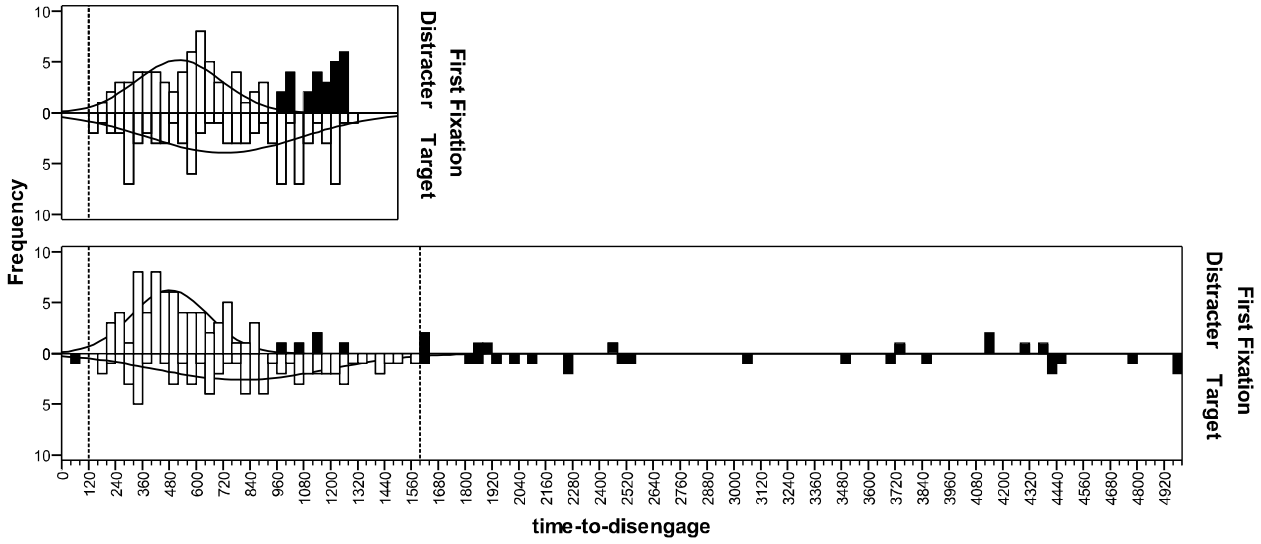
Notes. a. depicts time-to-fixate according to age only. b. depicts time-to-fixate according to age and presentation time. Independent samples t-tests compare RT in target and distracter direction. *p<.05. **p<.01. Whiskers show +/- one standard error.

Time-to-disengage

Histograms of the distribution of time-to-disengage over the trial are given in figure 4.08. Visual inspection of histograms suggested that the majority of values for time-to-disengage fell between 120ms and 1600ms, and that this dense distribution was likely to be responsive, rather than predictive. Of the 338 valid trials for which time-to-disengage data was available (i.e., the target word was known, the infant was not looking at a screen area at the onset of the images and the infant made more than one look in the course of the trial), time-to-disengage was less than 120ms in one trial, and was 1600ms or greater in 61 trials (18%). Within this window, data for time-to-disengage were abnormally distributed. When exclusions for time-to-fixate were carried over (only normally distributed first fixations were included), distribution was

of time-to-disengage was closer to normal distribution. This exclusion constituted a further 31 trials (9%).

Figure 4.08. Histogram of time-to-disengage over the course of the trial



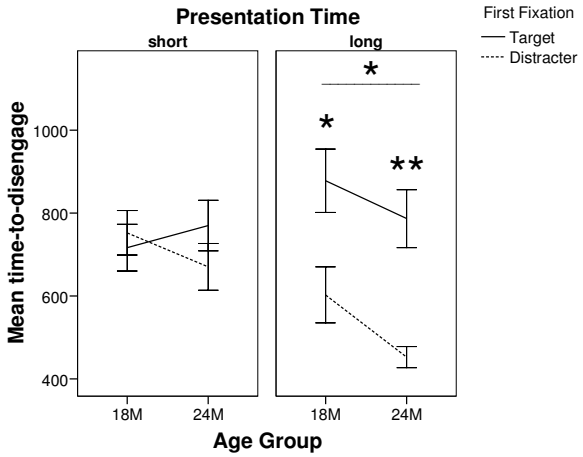
Notes. Bars show original distribution of all valid trials. Dashed lines indicate the manually selected window (120ms to 1600ms). Dark bars represent the trials excluded from analysis. White bars represent trials included in analysis, with normal curves overlaid. Note that this distribution was not normal.

Further inspection revealed that time-to-disengage remained abnormally distributed in the short timing condition. Here, data were bi-modally distributed, with a late peak beginning at 960ms. However, as this late peak constituted 22% of the remaining sample, the late values were retained, and non-parametric tests were conducted.

The means for disengaging from targets and distracters in each participant condition are depicted in Figure 4.09, where the mean time-to-disengage from targets was 774ms (*SD* = 356), and from distracters, 633ms (*SD* = 312ms). Distracters were disengaged from faster than targets ($z(124, 121) = 3.02$, $p = .003$). In Figure 4.09, the difference between timing conditions is also quite clear: In the short timing condition, no difference was evident between age groups ($z(82, 66) = 0.47$, n.s.), nor between

look directions ($z(77, 71) = 0.32$, n.s.). However, in long timing condition saccades away from distracters were launched faster than targets ($z(47, 50) = 4.51$, $p < .001$). This was true of both 18-month-olds ($z(15, 20) = 2.35$, $p = .02$) and 24-month-olds ($z(32, 30) = 3.61$, $p < .001$). In addition to the general effect of target discrimination, older infants were reliably faster than younger infants in the long timing condition ($z(35, 62) = 2.26$, $p = .02$) as is evident in Figure 4.09.

Figure 4.09. Time-to-disengage, according to age and direction of first fixation



Notes. Wilcoxon's rank sums tests compare RT in target and distracter direction. *p<.05. **p<.01. Whiskers show +/- one standard error.

Taken together, the results of the micro-level measures demonstrate that infants do not direct more of their first fixations to targets than they do to distracters, as demonstrated by the chance-values for percent correct first fixations. However the speed and duration of the first fixation (time-to-fixate and time-to-disengage) demonstrated sensitivity to the direction of the first look, with slower fixations to targets and faster disengagement from distracters. This general effect was mediated however, by age and time pressure: Only young infants in the slow condition and old infants in the fast condition reliably demonstrated target discrimination in time-to-

fixate. In time-to-disengage, however, both age groups demonstrated reliable target discrimination, but only in the long timing condition.

These findings suggest that the short timing condition is too short to allow consistent response patterns. Interestingly, for the older infants, the long timing condition may well have been *too long*, as older infants failed to show consistent target discrimination in the long timing condition, even though were able to in the short timing condition. In a long presentation experiment, older infants may exhibit less motivation to correctly land the very first fixation, in response to the low time pressure.

Comparison of performance in two micro-level measures

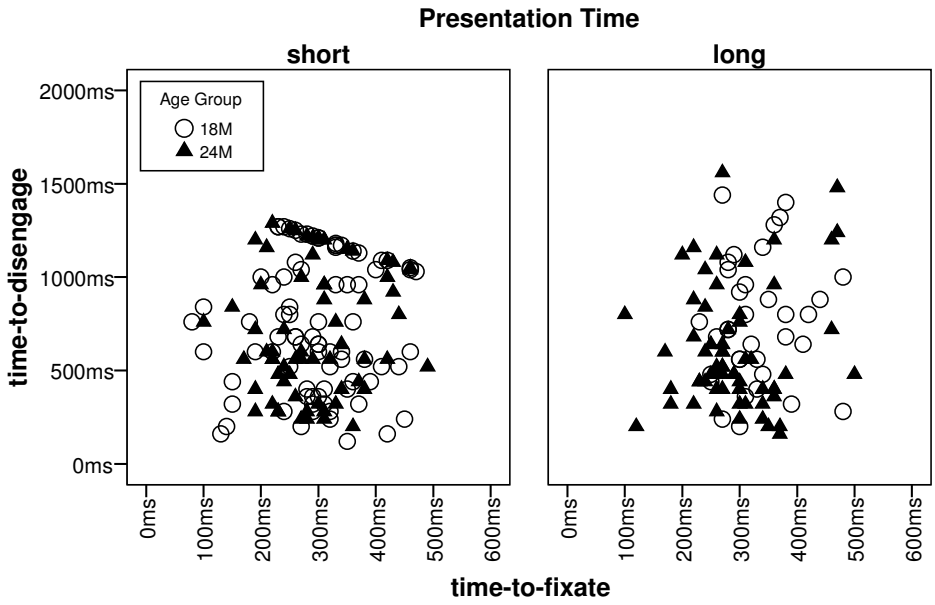
Some of the earliest reaction time experiments suggested that slower RTs in a standard psychometric task were also coupled with participants' slower gait when walking, and a slower disposition *overall* (van Essen, 1937). The micro-level measures in the current experiment (time-to-fixate and time-to-disengage) represent sequential RTs. As such, there is a possibility that the second measure is simply a reflexive outcome of the first (e.g., fast-generates-fast, or slow-generates-slow), or that both measures index the same underlying performance style (e.g., a generally fast responder), or short-term response mode (e.g., a fast trial). A correlation between sequential RTs in the current experiment, would imply that time-to-disengage was a reflexive outcome of the time-to-fixate, and therefore, uninformative. Time-to-fixate is plotted against time-to-disengage in Figure 4.10, where the values for 18-month-olds overlap substantially with the values for 24-month-olds. In the long timing condition, the dark triangles are distributed further to the left, indicating the 24-month-olds' faster time-to-disengage. It is interesting to note the linear clustering of

values at the upper right of the distribution in the short timing condition. The origin of this clustering is quite clear: In the long timing condition, where looks in both directions were relatively unconstrained by time pressure, the upper limit of normally distributed time-to-fixate was 520ms, and time-to-disengage, 1560ms. Thus, a slow fixation followed by a long look-duration would sum to 2080ms. That is, 580ms longer than the picture presentation in the short timing condition. This scatter-plot demonstrates the extent to which values for time-to-disengage were constrained by the short timing interval, and also explains the emergence of the large late peak in the distribution of time-to-disengage for the short timing condition. This timing constraint is also likely to account for the lack of target discrimination in time-to-disengage in the short timing condition. The comparison of onset and offset of the first fixation demonstrates even more clearly how the short timing condition is dominated by a single fixation, as was observed in the comparison of macro-measures. This finding again suggests that the short timing condition is too short to be a reliable measure of target discrimination in future studies.

No correlation between micro-level measures was observed at the overall level ($\rho(245) = .06$, n.s.), within the different stimulus conditions (short: $\rho(148) = .05$, n.s.; long: $\rho(97) = .07$, n.s.), nor within the different age groups (18m: $\rho(117) = .02$, n.s.; 24m: $\rho(128) = .05$, n.s.). This finding suggests that time-to-disengage is independent of time-to-fixate, not merely a reflexive outcome of the participants' general response speed. The lack of correlation is consistent with the findings of Frick, Colombo and Saxon (1999), who found that infants' general look duration (or 'pace') did not correlate with the speed of fixation shifts to peripheral stimuli, so long as they were

not fixating a central stimulus[9]. The independence of the two measures makes time-to-disengage a valuable additional measure of target discrimination.

Figure 4.10. Time-to-fixate plotted against time-to-disengage for each trial



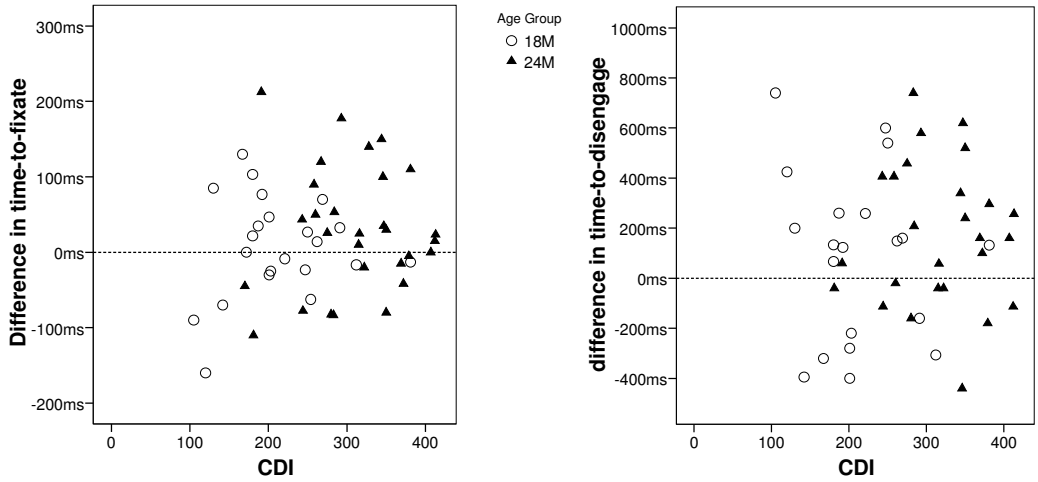
Notes. No correlation was observed between micro-level RT measures.

Comprehension vocabulary

For each infant, mean time-to-fixate and time-to-disengage was calculated in each look direction. Difference measures were calculated (t-d) for comparison to the developmental variable of vocabulary size. CDI score is plotted against micro-measures in Figure 4.11, where the older infants (dark triangles) are distributed further to the right, given their generally larger vocabularies.

[9] This is contrasted with a "competition" condition in which a central stimulus remained onscreen. Average look duration (established in a separate task) *was* correlated with fixation speed in the competition condition, indicating that those who are typically long lookers are also slow disengagers. Note that in Frick et al. (1999), only a single peripheral stimulus was presented at a time. Common to both the non-competition condition and the current experiment is that infants were not fixating a salient central stimulus at the onset of test stimuli.

Figure 4.11. Two RT measures plotted against CDI comprehension score



Notes. Difference measures calculated as mean for target direction minus mean for distracter direction. Dashed lines indicate chance value (0ms).

The difference values are widely dispersed, with more values falling above chance than below, indicative of the general trend for target discrimination. No correlation was found between CDI comprehension score and time-to-fixate, nor between CDI score and time-to-disengage.

Accounting for item variability

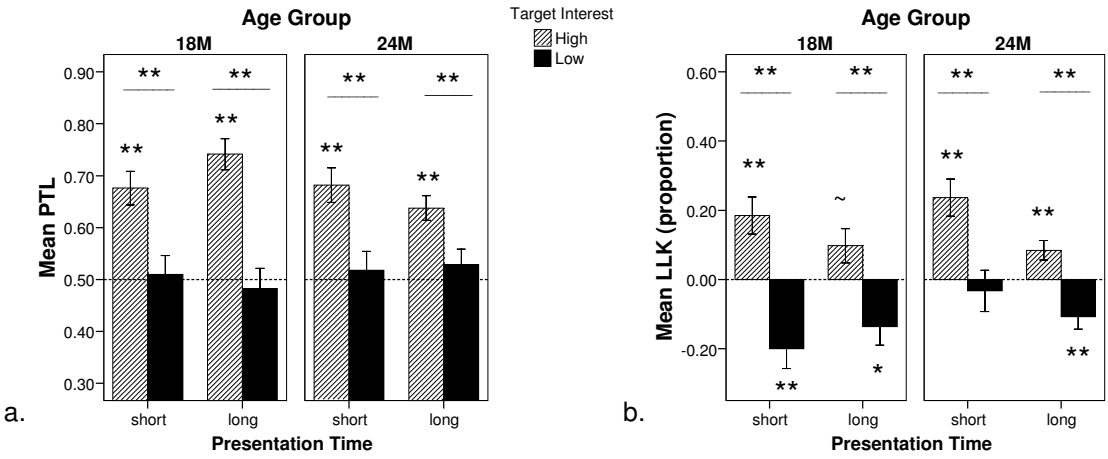
In this one-shot design, all infants saw the same sixteen items, half of which were named, and they saw each item only once. Although target status was counterbalanced between lists, the individual items included in the test were predicted to influence performance in the task. For future studies, it is important to establish a) what features of items are most likely to influence performance, and b) which measures are most susceptible to item variability in the current task. For this reason, three analyses were conducted across each of the measures. First, the impact of the within-pair relative image ranking was assessed. Second, target discrimination was assessed for each item. Thirdly, the variance introduced by individual items was quantified using regression techniques. The former two analyses allow particular item

attributes to be assessed, while the latter allows the measures themselves to be compared for their robustness.

Relative within-pair interest

Figure 4.12 depicts the two macro-measures of target discrimination (PTL and LLK) according to whether the named target was ranked as high or low interest within its picture pair. The striped bars (representing looking to high interest pictures) are consistently higher than the dark bars (representing looking to low interest pictures), suggesting that high interest targets attract more and longer looks than low interest targets. This plot also summarises the results of single sample t-tests comparing each subset of the data to chance, and independent samples t-tests comparing high and low interest items for each participant group. In both measures, performance was significantly higher for high interest targets, in all subject conditions.

Figure 4.12. Relative within-pair picture interest in two macro-level measures.



Notes. a. depicts the proportion of target looking (PTL). b. depicts the longest look difference (LLK), expressed as a proportion of the picture presentation window. Horizontal lines indicate chance value in each measure. Single sample t-tests compare each subset of the data to the chance value in that measure. Independent samples *t*-tests compare high and low interest levels. *p<.05, **p<.01, ~p<.10. Whiskers show +/- one standard error.

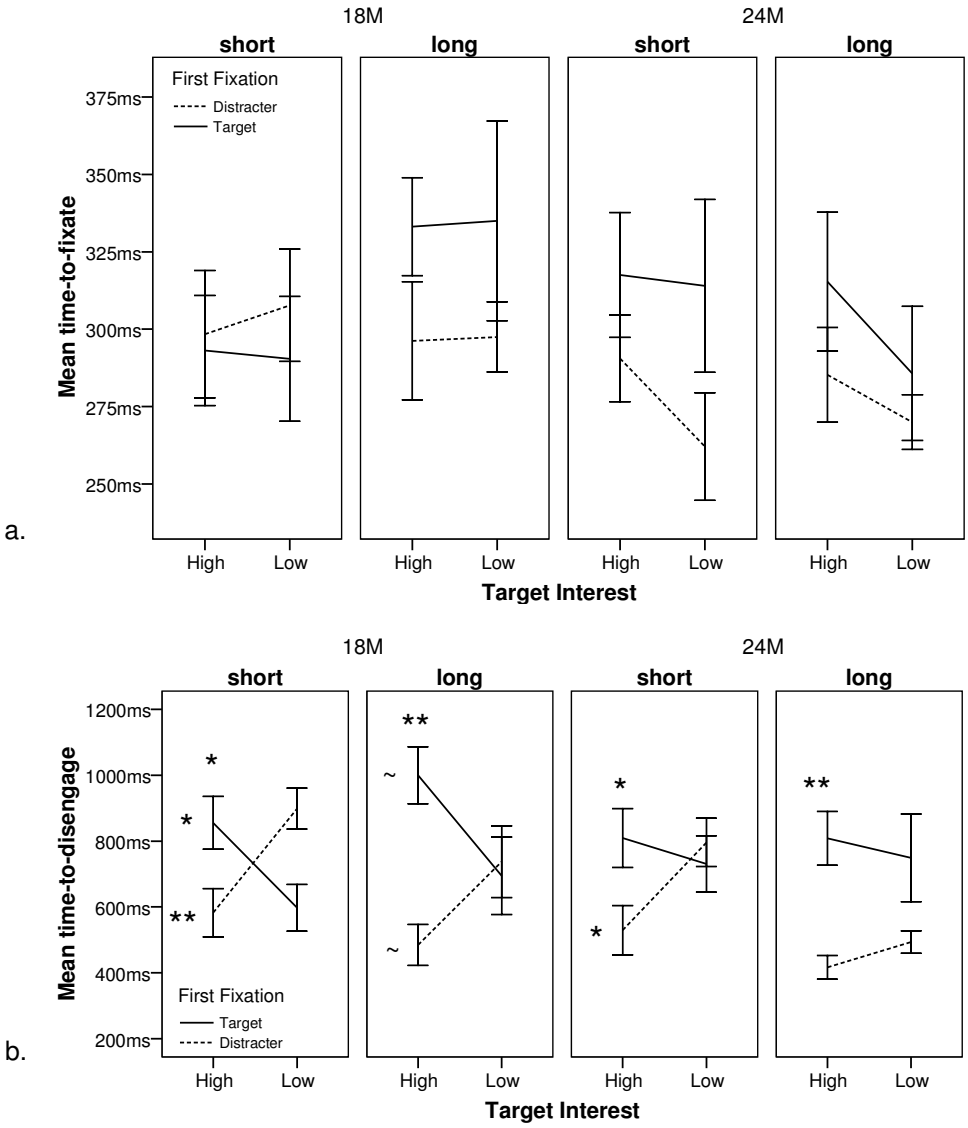
PTL, depicted in Figure 4.12a, was above chance for high interest targets, indicating robust target discrimination for these items. However, note that if the interest level of the picture were the only factor driving looking behaviour (without any influence from the name), then low interest targets would be expected to attract equivalent *below chance* looking. This was not the case in PTL, a finding which suggests that picture naming influenced behaviour even on low interest targets.

In LLK on the other hand, as illustrated in Figure 4.12b, while high interest targets generally generated above chance target looking (with the exception of 18-month-olds in the long timing condition, where the comparison failed to reach significance), low interest targets in three of the four conditions attracted below chance performance in this measure. This suggests that the overall results for LLK were led, to some extent by the relative interest of the pictures in each pair.

Figure 4.13 illustrates RTs to high and low interest targets as measured by time-to-fixate and time-to-disengage. Here the solid lines (representing RTs to and from targets) are generally higher than the dashed lines (representing RTs to and from distracters), indicating the general pattern of slower fixations to targets, and faster disengagements from distracters. This plot also summarises the results of comparisons across interest level, and compares look direction as a measure of target discrimination. In time-to-fixate (Figure 4.13a), within-pair interest did not affect RT in any condition, and the pooled data show target discrimination (targets fixated slower than distracters) in two of the four conditions, as demonstrated in the analysis of participant means (18m short: $t(80) = 0.62$, n.s., $d = 0.14$; 18m long: adjusted $t(40.9) = 2.02$, $p = 0.05$, $d = 0.63$; 24m short: adjusted $t(54.2) = 1.99$, $p = 0.05$, $d = 0.54$; 24m long: adjusted $t(58.9) = 1.47$, n.s., $d = 0.38$)[10].

[10] Where marked, adjusted *t*-value and degrees of freedom given, as between group variances unequal.

Figure 4.13. Relative within-pair picture interest in two micro-level measures.



Notes. In time-to-fixate, independent samples t-tests found no difference between high and low interest targets. In time-to-disengage, Wilcoxon's rank sums tests compare high and low interest targets when marked *beside* the line for target or distracter. Wilcoxon's rank sums tests compare direction of first look when marked *above* the comparison. **p*<.05, ***p*<.01. Whiskers show +/- one standard error.

In time-to-disengage (Figure 4.13b), however, within-pair picture interest affected looking behaviour to some extent for 18-month-olds in both timing conditions, and in both look directions: When the target was less interesting, 18-month-olds in the short timing condition showed distracter preference. In the long timing condition, within-pair interest only had a marginal effect on their behaviour in

each look direction. Older infants were less affected by the within-pair interest of the target as only distracter fixations in the short timing condition were affected by interest level. In the long timing condition, within-pair interest level did not have a reliable effect on time-to-disengage in either age group, and the pooled data show clear target discrimination (18m long: $z(15, 20) = 2.35$, $p = .02$; 24m long: $z(32, 30) = 3.61$, $p < .001$).

Although within-pair interest did have some effect on time-to-disengage, it is worth bearing in mind that if relative picture interest were the only factor influencing time-to-disengage, we would expect fast disengagement from low interest targets, coupled with slow disengagement from high interest distracters. This pattern was observed for 18-month-olds in the short timing condition, but not in the other three conditions. While target discrimination was not reliable for low interest targets in time-to-disengage, the target name interfered with responding to interest alone in three of the four participant conditions.

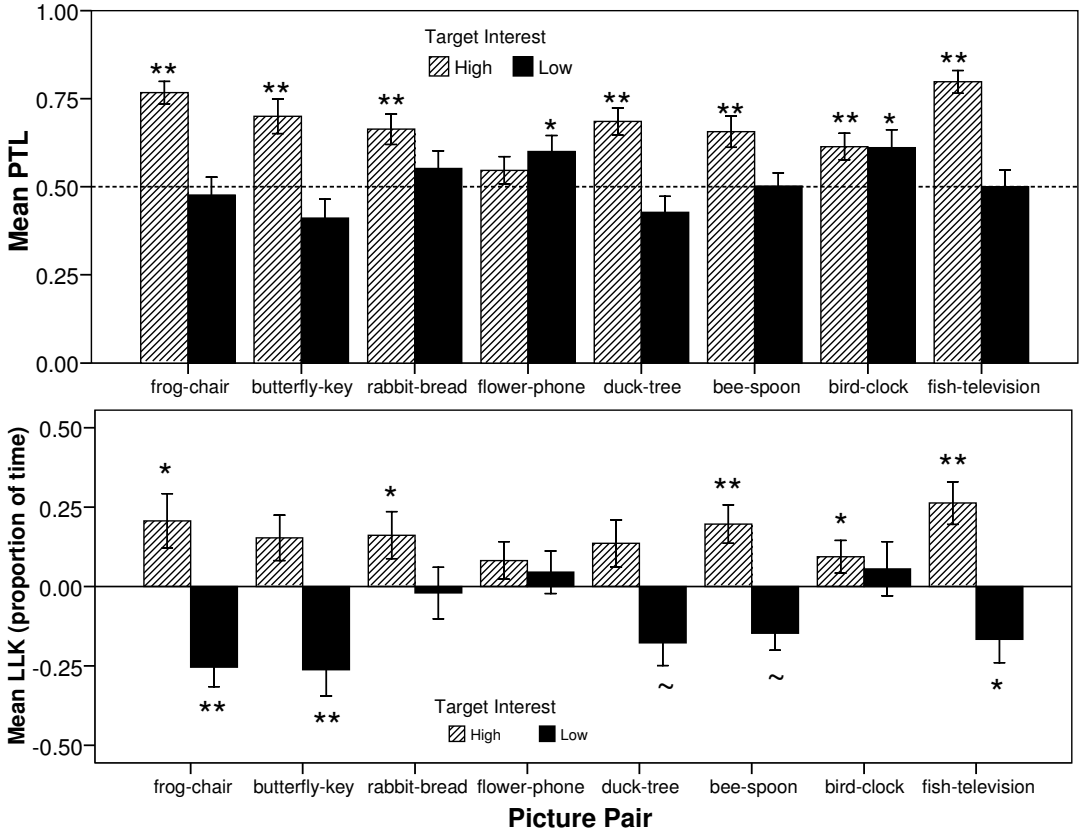
The findings from the comparison of within-pair interest indicate that the macro-level measures were heavily influenced by the relative interest of pictures within a pair. The finding is not especially surprising as seven of the eight picture pairs contained one animal and one man-made artefact. LLK was more heavily affected than PTL, where *distracter* preference was observed for low interest targets in most conditions. The micro-level measures were somewhat less susceptible to the relative interest of the target, with time-to-fixate showing no effect of interest level, and reverse preference only observed in a single condition for time-to-disengage. Clearly, when items from the same semantic category must be avoided (as in the planned series of priming studies), the monitoring and counterbalancing of interest levels will therefore be of critical importance.

Individual items.

Across all subject conditions, following the onset of naming, more than half of the items (9 of 16) were looked at significantly more than would be predicted by chance (PTL of .50) – the pattern expected following naming. The items which did not show the expected preference were *chair*, *key*, *bread*, *flower*, *tree*, *spoon* and *television*. All of these items were inanimate and, as is clear in Figure 4.14, all-bar-one were ranked as 'less interesting' in their picture pair. The only pair in which preference was inverted according to the approximate rank was the pair *flower-phone* (the only pair in which both items were inanimate, and which shared a similar onset). No items, however, received a mean PTL which was below chance. This suggests that while the 'more interesting' item in a pair may exert some 'pull' on infants' attention, their performance was not completely biased by the more interesting items, and thus, the effect of naming was still evident overall.

By contrast, only five of the 16 items consistently attracted a difference in longest look which was more than would be predicted by chance. Longest look differences were above chance for *frog*, *rabbit*, *bee*, *bird* and *fish*, did not differ from chance for *bread*, *phone*, *duck*, *spoon* and *clock*, and were below chance for *chair*, *key* and *television* (*tree* and *spoon* were not significantly below chance, but approached the alpha level of .05, with *p*-values of .062 and .059, respectively). Note that the five lowest performing items were the same in both PTL and LLK.

Figure 4.14. Means for each picture in two macro-measures, pooled across subject conditions



Notes. Grouped by picture pair and relative 'interest' ranking within each pair. Upper plot depicts proportion of target looking (PTL). Lower plot depicts difference in longest look (LLK) expressed as a proportion of the total looking time in each condition. Single-sample t-tests compare value for each item to the chance value for that measure. *p<.05. **p<.01 ~p<.10. Whiskers show +/- one standard error.

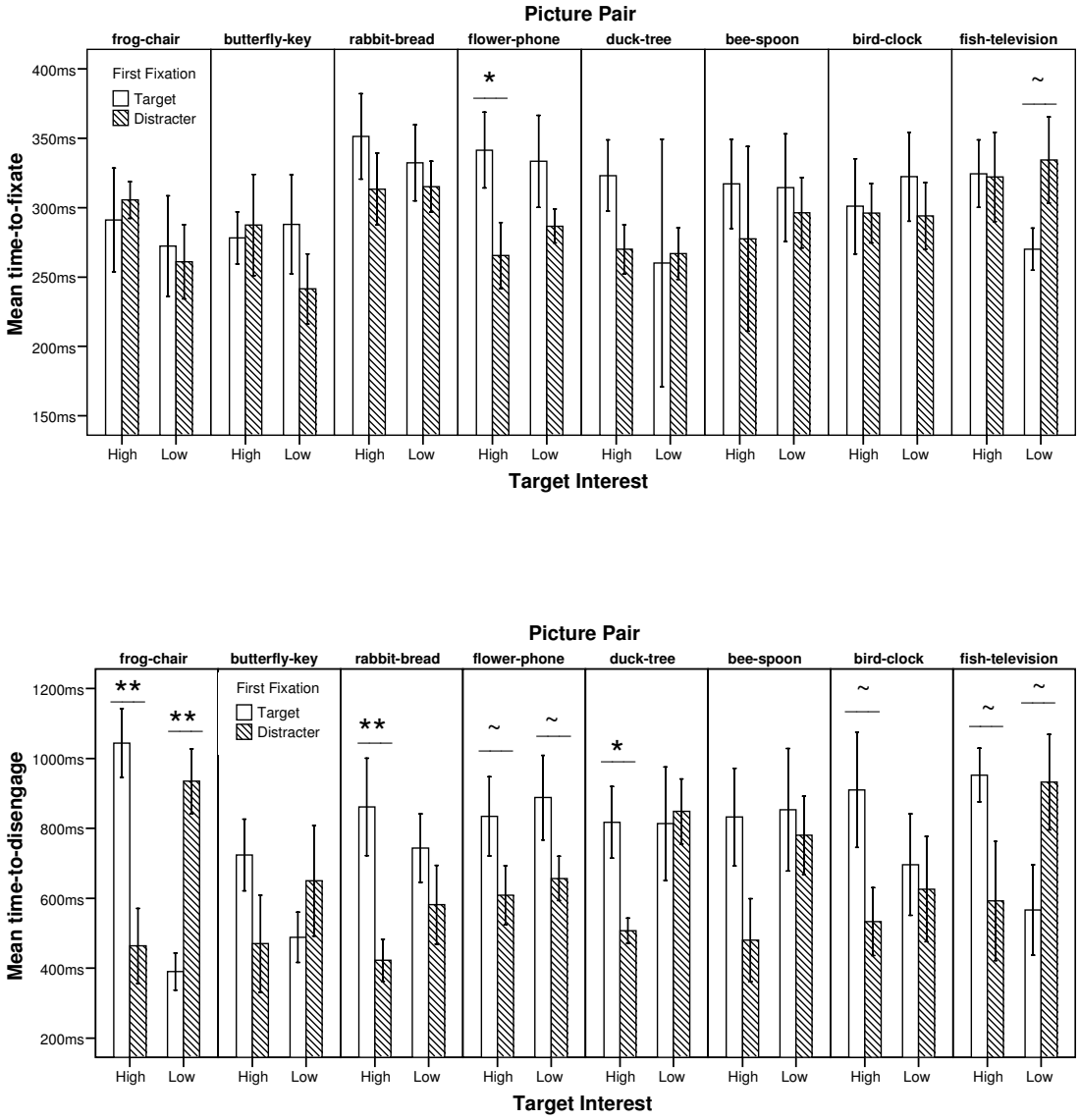
Figure 4.15 plots the micro-level measures of looking for each item. In these plots, the white bars (representing first fixations to a target) tend to be higher than the striped bars (representing first fixations to distracters), as would be expected given the general pattern of slower fixations to targets and faster disengagement from distracters. It is clear that the variance in these measures is quite large. In time-to-fixate, the only item which independently demonstrated target discrimination in time-to-fixate was *flower*, which was fixated slower when it appeared as the named target, than when it appeared as an unnamed distracter. The item *television*, however, tended in the opposite direction, with faster looks to the named *television* (and slower looks to the *fish*). In time-to-disengage, three items demonstrated reliable target

discrimination: *frog*, *rabbit* and *duck*, with slower disengagement from these items when they appeared as targets than when they appeared as distracters. In addition, *flower*, *phone*, *bird* and *fish* were disengaged from moderately slower when they appeared as targets. These comparisons were not significant, but approached the alpha level of .05, with *p*-values of .07, .11, .06 and .09, respectively. In time-to-disengage, two items tended to attract looking opposite to the general pattern of behaviour: the low interest target *chair* attracted significantly shorter first fixations than its distracter *frog*, and the low interest target *television* tended to attract shorter fixations than its distracter *fish*, although the comparison did not achieve significance in this case (*p*-value of .06). It thus appears that the items *frog* and *fish* were highly salient to infants, exhibiting quite a strong pull on infants' looking behaviour.

The results for individual items show that only a small number of items attracted reliable target discrimination when the data are reduced to this level (only one presentation per item, per infant), and were pooled across four conditions which did not always produce reliable target-discrimination (for example, 18-month-olds in the short timing condition failed to show target discrimination in any measure). This analysis is therefore an assessment of which items are *best recognised* given the current range of experimental conditions. The analyses also give an indication of which items might most influence looking behaviour over and above the effect of target naming. The picture-pairs *frog-chair*, *butterfly-key* and *fish-television* appeared to have the most dramatic distracter effects. These three pairs shared the property of having very colourful animals (yellow and black, bright blue and black, and bright orange, respectively). By comparison, *rabbit-bread*, *duck-tree*, *bee-spoon*, and *bird-clock* all contained duller coloured animals, and do not appear to have as dramatic a

pull towards the animal when it was the distracter. This finding suggests that the particular animals used in a task should be monitored carefully in future studies.

Figure 4.15 Means for each picture in two micro-measures, pooled across subject condition



Notes. Grouped by picture pair and relative 'interest' ranking for each item. Upper plot depicts time-to-fixate. Independent samples t-tests compare time-to-fixate in each look direction. Lower plot depicts time-to-disengage. Wilcoxon's rank sums tests compare time-to-disengage in each look direction *$p \leq .05$. **$p < .01$ ~ $p \leq .10$. Whiskers show +/- one standard error.

Quantification of individual and item effects in each measure.

In order to establish which experimental variables were responsible for the greatest amount of variance in each measure, categorical regressions of the four measures of target discrimination (PTL, LLK, time-to-fixate and time-to-disengage) were conducted in a stepwise manner. This analysis also allows assessment of which measures are the most robust to known sources of variance, and might be of most value in future studies. Categorical regression allows each category of an independent variable to be assigned a numerical value via optimal scaling techniques, which generate the best possible model for predicting a continuous dependent variable.

Independent variables of interest were tested one at a time in optimally scaled regressions. Those variables which returned significant fits to the data were then combined into multivariate models, in order to find the model which accounted for the largest proportion of the variance. The variables entered independently were: participant (each individual participant), direction of first fixation (target, distracter), target word (*bee*, *bird*, *bread*, *butterfly*, *chair*, *clock*, *duck*, *fish*, *flower*, *frog*, *key*, *phone*, *rabbit*, *spoon*, *tree*), relative within-pair interest (high, low), target animacy (animate, inanimate), target kind (natural, artefact), picture-pair (categorical, one to eight). In addition, two experimental variables were included: trial speed (fast, slow), which was calculated according to the median split of time-to-fixate for each age group[11]; and trial number (sequential, one to eight), which was included to check on whether learning effects over the course of the trial were a significant source of variance. Best fitting models for each subset of the data are reported in Tables 4.4 to 4.7, and are summarised graphically in Figure 4.16.

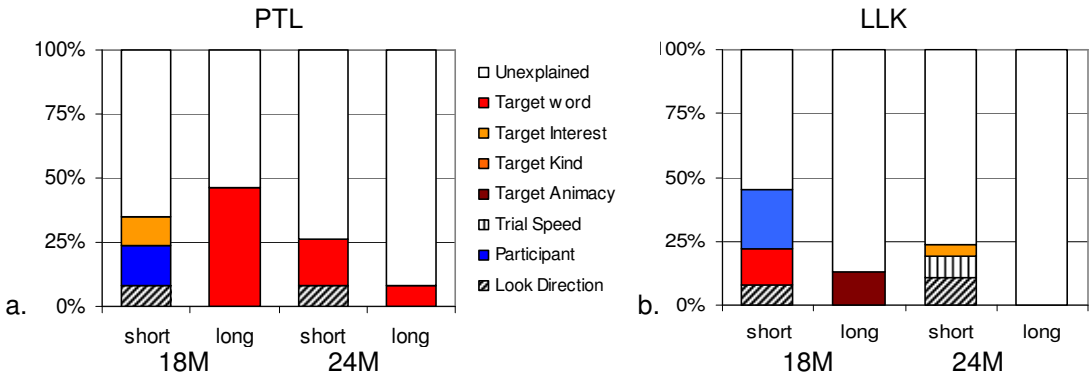
[11] Trial speed was not included in the models of time-to-fixate.

In the models of macro-measures, three features are worth noting. Firstly, both macro-measures show a significant amount of variance due to the influence of individual items, as indicated by the red, orange and brown portions of the Figure 4.16. In general, older infants show less variance due to individual items, particularly in the long timing condition, where LLK appears to be unaffected by individual items.

Secondly, in the short timing condition, both age groups and both measures demonstrate a significant amount of variance due to the direction of the first fixation, as is clear in the hashed portion at the base of columns in Figure 4.16. Given that the macro-level measures are designed to assess the accumulation of relative preference (rather than online processing) over a period of time, this finding again suggests that the short picture presentation is too short for the diagnostic characteristics of macro-measures to be appropriate. Recall that younger infants failed to show target discrimination in both macro-measures in the short timing condition. This may have been caused, in part by the variance due to the direction of the first look, which was shown to be unsystematically directed (according to the percentage of correct fixations). Finally, younger infants show a significant amount of variance according to individual participants in the short timing condition.

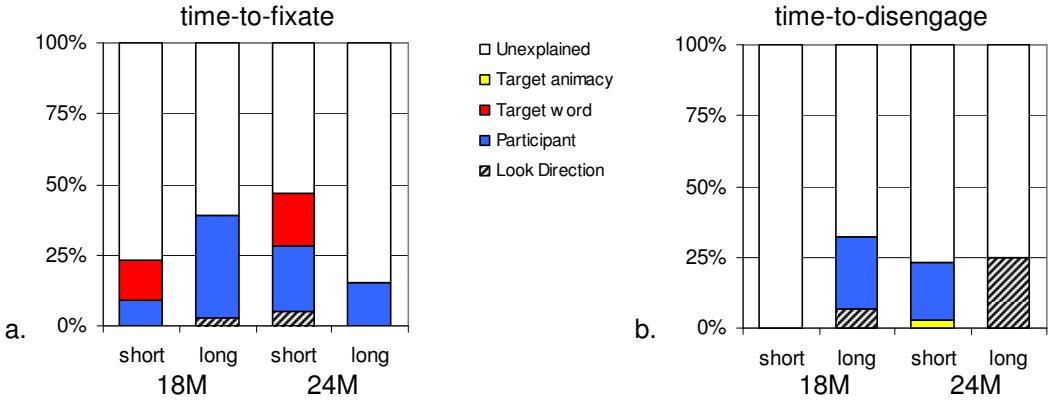
Models of variance for micro-level measures are plotted in Figure 4.17. Here, the most interesting feature is the dramatic reduction in the predictive nature of item variables.

Figure 4.16. Best fitting models of variance for macro-level measures



Notes. a. depicts the percentage of variance in proportion of target looking (PTL) accounted for by models d, e, l and m outlined in Table 4.6. b. depicts the percentage of variance in the longest look difference (LLK) accounted for by models d, e and i outlined in Table 4.7.

Figure 4.17. Best fitting models of variance for micro-level measures



Notes. a. depicts the percentage of variance in time-to-fixate accounted for by models d, e, l and m outlined in Table 4.8. b. depicts the percentage of variance in time-to-disengage accounted for by models d, e and i outlined in Table 4.9.

Although time-to-fixate is influenced by individual items in the short timing condition, no effect of item variability was found in the long timing condition, and very little was evident in the measure time-to-disengage. This suggests that online 'match' or 'mismatch' judgements are more automatic than the macro-level measures, which show much greater item variability. By contrast, the micro-level measures show much more variability due to the individual participants.

Taken together, these finding demonstrate that the micro-level measures appear to be more robust to item effects, while the macro-level measures show more consistency across participants. The short timing condition was found to be problematic (with the macro-level measures affected by the direction of the first look). Despite the target discrimination demonstrated by the older infants in the short timing condition, the confound of a single look dominating the macro-measures should be avoided in future studies. Removal of this timing condition would also reduce the influence of individual items from the variance of micro-measures. With large amounts of variance predicted by individual participants, the micro-level measures would clearly benefit from more trials per infant, and from a larger subject pool.

Discussion

The current experiment was designed to test whether infants would be able to demonstrate target discrimination in a cut-down version of a standard intermodal preferential looking (IPL) task. The words-before-pictures experiment was designed to pilot features required for a short-SOA priming task: Trials contained no picture familiarisation time, and each item was only encountered once. Picture presentation time was manipulated between participant groups to establish how long pictures would be required to remain on-screen. The most important finding of the current study is that infants at 18 and 24 months of age can demonstrate general target discrimination in the current IPL task with no familiarisation period or repetition of trials. This validates the basic task for use in future priming studies. The experimental manipulation also adds to our understanding of different measures used to assess target discrimination in IPL.

The effect of naming a picture varied across subject groups, and across the different measures, as is summarised in Table 4.3. The short timing condition (1,500ms picture presentation) was found to be too short for a number of reasons. Firstly 18-month-olds in this timing condition failed to show target discrimination in any measure. However 18-month-olds *were* able to demonstrate target recognition in three of four measures when the images remained on screen for a longer duration. Secondly, even 24-month-olds failed to show target discrimination in the short timing condition in time-to-disengage, despite demonstrating target discrimination in time-to-fixate. This is most likely due to the trial ending before the infant was able to naturally shift away from a long fixation, as graphically illustrated in the scatter-plot of Figure 4.10. Finally, a statistical dependency between PTL and LLK in the short timing condition was demonstrated by the tight linear relationship between measures. This feature of the short timing condition is undesirable as it confounds whether the two measures index the same *type* of behaviour, or the same two fixations.

Table 4.3. Summary of target discrimination in four measures

		PTL	LLK	t-t-fixate	t-t-disengage
18m	short	-	-	-	-
	long	Y	-	Y	Y
24m	short	Y	Y	Y	-
	long	Y	Y	-	Y

In the long timing condition, both age groups demonstrated target discrimination in three of four measures. 18-month-olds failed to discriminate targets from distracters in LLK. This may have been due to features of particular stimulus items attracting long stares during the 5,000ms time window, as LLK in this condition did show significant variability due to items. 24-month-olds failed to show target discrimination in time-to-disengage – a surprising finding considering that they *did*

show target discrimination in the *short* timing condition in this measure. One possible interpretation for this finding is that older infants do not perceive time pressure in the 5,000ms condition, and are less motivated to perform 'correctly' at the very start of the trial. This suggests that picture durations somewhere between 1,500ms and 5,000ms might be beneficial for both 18- and 24-month-olds in future studies.

Micro-level measures

Infants heard the start of the target word before they knew which side the target would appear. A single measurement of time-to-fixate from the onset of picture presentation was predicted to index a 'blind' motor reflex, randomly directed to the left or right. As such, time-to-fixate was expected to be a useful baseline against which infants' subsequent saccades could be measured. Contrary to expectation, when combined with the direction of the first look, this measure *was* sensitive to target discrimination in two subsets of the data (young infants in the slow condition and older infants in the fast condition). The direction of the effect was also somewhat surprising, with *slower* looks to targets than to distracters. This finding suggested a processing cost in accurate performance (first looks to targets were launched slower).

It might seem surprising that infants as young as 18 months-of-age were able to process suddenly appearing peripheral stimuli before the first saccade had been launched. However, in the current experiment, pictures were presented only 5° from the centre of the screen, with their outer edges extending to 23° from the centre. Adult studies have shown the human visual system has a high degree of accuracy at such small eccentricities, including the ability to extract semantic information from very briefly presented photographs prior to saccadic movement (Thorpe *et al.*, 2001). Furthermore, recent eye-tracking studies have demonstrated that when an auditory

word *precedes* the presentation of a picture pair, adults rapidly fixate the named target in their first fixation almost 100% of the time (Vucetic, 2008). It is therefore possible that the infant visual system is sufficiently advanced to identify visual features of pictures prior to launching their first eye-movement.

An alternative explanation for target discrimination in time-to-fixate, is that infants begin to process an online target 'match' after planning or even launching the saccade. In this account, the processing cost of 'matching' causes a saccadic slow down, leading to a slower fixation. This second account might fit better with the finding that the percentage of correct first fixation was at chance. Both accounts, however, demonstrate that early visual processing is engaged in the first few hundred milliseconds of a trial, even when the pictures have never been seen before, and the infant does not know which location the target will occur in.

Item effects

Three methods of monitoring individual item effects were used in the current experiment: The ranking of relative within-pair interest was assessed for its effect on target identification; target discrimination for individual items was reported; and item variability was quantified in each of the four measures of target discrimination. From these analyses it is clear that different items attract different levels of visual attention, and this has an impact on target discrimination. The relative assessment of within-pair interest clearly demonstrated more and longer looks to high-interest targets – a finding which strengthens the necessity for counterbalancing interest levels within future stimulus sets. In the analysis of individual items, animates generally attracted more looking than inanimates, which could result in above chance *distracter looking* in LLK, but did not bias the other measures as much. The micro-level measures were

less likely to show target discrimination at this level of analysis, but also appeared to show less variability between items. This suggestion was confirmed in the categorical regressions of each measure, where the micro-level measures showed very little variability according to item descriptors, and almost none outside the short timing condition.

Comparison to previous IPL findings concerning latency

IPL studies have rarely been conducted without a familiarisation phase. In one study, Golinkoff, Hirsch-Pasek, Cauley and Gordon (1987) report mean latencies of 3.43 seconds to targets, and 4.23 seconds to distracters, with infants making significantly faster fixations to targets than to distracters. Two major discrepancies are evident between their findings and those reported here. The first is the direction of the effect, with faster latencies to targets than to distracters, and the second is the scale of the measurement: Golinkoff et al's 17-month-olds produced latencies six to eight times slower than the time-to-fixate demonstrated by 18- and 24-month-olds reported here.

It is worth noting that Golinkoff et al.'s latency measure does not take into account the *sequence of first fixations*. Instead, it reports the time taken to fixate each picture in a single trial. For example, for the pair *cookie – sock*, the authors report a mean latency to *cookie* of 2.93 seconds, and to *sock*, 4.79 seconds. In this example, it can be assumed that the majority of infants looked first at the cookie, then shifted their gaze to the sock some seconds later. However, as this measure contains pooled data from both first and second fixations, it is somewhat difficult to interpret the results directly. For example, if first fixations were equally fast in both look directions, then the *duration* of the first fixation (time-to-disengage) would determine

the latency to the second picture. If, on the other hand, the speed of the first fixation (time-to-fixate) were variable and the duration, static, the *speed* of the first fixation would determine the latency to the second. However, if these two measures (time-to-fixate and time-to-disengage) varied in opposite directions, the pooled data might show no difference between target and distracter direction, even though a difference originally existed in the raw data. As an additional complication, if the latency of both fixations were completely static, yet *more infants* fixated the target first, then the pooled data would show faster mean latency to targets, as there would be more first fixations in that pool. In sum, this particular 'latency' measure conflates two measurements from a single trial which are not independent of each other, and confounds the measure with a possible distributional bias in look direction. While legitimate behavioural differences were observed, it is impossible to pinpoint the locus of the effect, and some important effects may well have been masked.

It then follows that the Golinkoff et al., 'latency' measure is very different to the two RT measures used in the current study. However, the direction of the effect is not incompatible with findings reported here. Taking the example of 18-month-olds in the long timing condition (who were most similar to Golinkoff et al.'s participants, both in age and in trial timing), if we assume that it takes approximately 200ms to launch a saccade away from the first fixated picture, and begin fixating the second, it is possible to reconstruct a 'latency' measure from the means generated in the current study: Target-first trials would yield a *target latency* of 334ms and *distracter latency* of 1078ms (calculated as 878ms + 200ms). Distracter-first trials would yield a *target latency* of 803ms (calculated as 603ms + 200ms), and a *distracter latency* of 297ms. Pooling these means across trials from each look direction (as in Golinkoff et al.) would produce an approximate mean target latency of 569ms and an approximate

mean distracter latency of 688ms. Note that now, target latency is smaller than distracter latency – the same direction as that reported in Golinkoff et al. (1987).

However, these artificially created latency values are still much smaller than the latencies of 2.93 or 4.79 seconds reported by Golinkoff et al. (1987). Four key differences between the experiments have been identified, which most likely account for the discrepancy. Two differences concern the experimental procedure (time pressure, and laboratory configuration), and two concern data handling (online coding and treatment of missing values). Each of these points is discussed in detail below.

Time pressure

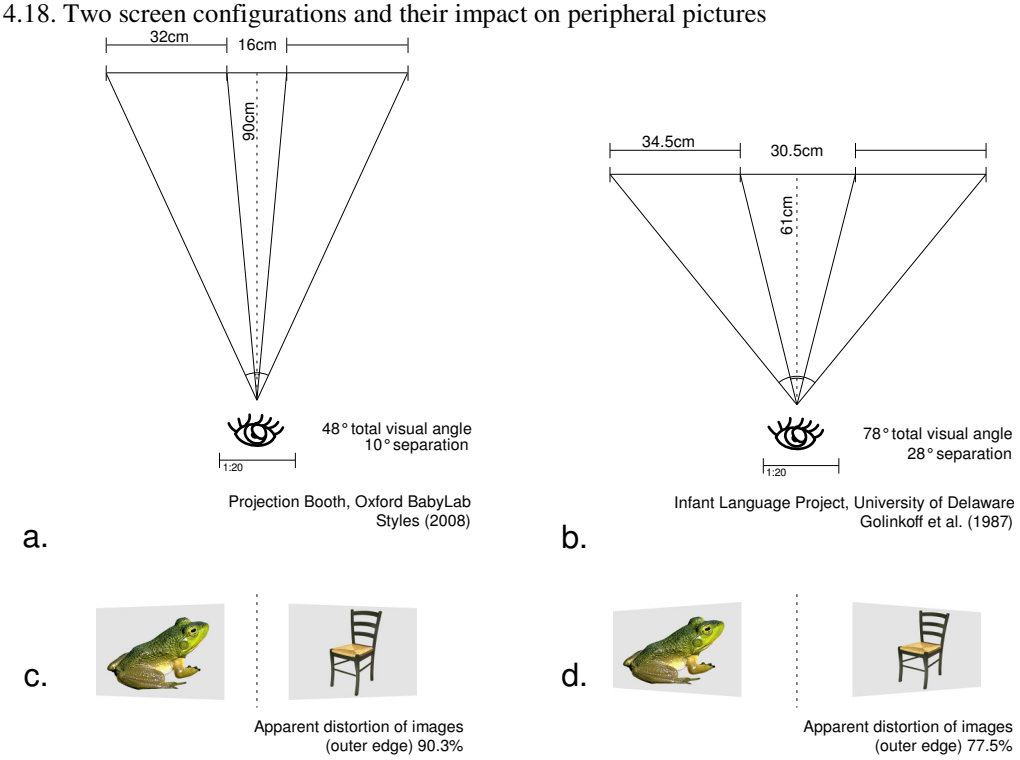
The Golinkoff et al. task had a much slower pace than the current task, with 3,000ms auditory presentation followed by 8,000ms visual presentation. Thus trials were almost twice as long as those in long timing condition presented here. The slower pace of the Golinkoff task may have lead to less perception of 'time pressure' by infants, and a more leisurely rate of response.

Laboratory configuration

The configuration of the screens reported by Golinkoff et al. is notably different to that reported here. As depicted in Figure 4.18b, the monitors of the Infant Language Project at Delaware University were separated by a large gap, and the infant sat much closer to the display. The gap between the screens occupied approximately 28° of the infant's visual angle, with the outer edges of the displays extending to create a total visual angle of 78°. By contrast, in Projection Booth at the Oxford BabyLab (Figure 4.20a), the presentation areas of the projection screen were separated by only 10°, and extended to 48°. Overlaying figures 4.18a and 4.18b would

show the Oxford BabyLab picture *midlines* falling on the inner edge of the Delaware monitors.

Figure 4.18. Two screen configurations and their impact on peripheral pictures



Notes. Figures a. and b. depict picture size and separation, as well as observer distance from the screen in two laboratories. Both diagrams are drawn to the same scale, and show the visual angle to the outer edge of the pictures and the visual angle of the gap. Figures c. and d. depict the perceptual distortion of pictures in configuration a, and b, respectively. Distortions were calculated according to the formula in Appendix F. For illustrative purposes, pictures in c. and d. were identical in size prior to distortion, and were separated according to their respective screen configurations.

In the current study, most infants could easily initiate an eye movement from the centre of the screen to a picture, without also moving their head. Some infants were even able to launch saccades between left and right pictures without head movement. In the Delaware configuration, on the other hand, the centre of each image is almost 30° from the centre, a separation which commonly produces head movement (J Stephen Reznick et al., 2000). The wider visual separation of the pictures and the wider angle required for a first fixation may well have slowed responses in the Golinkoff et al. task.

A related outcome of the screen configuration is the angular distortion of images viewed at different eccentricities. Figure 4.18c and 4.18d show the apparent distortion pictures incur when presented in the different screen configurations. Although the distortion is not dramatic in either case, for infants who have had no prior experience with the exemplars used at test, the more pronounced distortion produced by wider separation may also contribute to slower first responses.

Online coding.

In the current method, offline coding of taped looking behaviour allows temporal accuracy down to the 40ms duration of a single video frame. Fixations as fast as 120ms were included in the normally distributed values for time-to-disengage. Coding of looks in the Golinkoff et al. task was conducted on-line, using a manual button-press system. As adults typically take a few hundred milliseconds to initiate a button-press, it is thus possible that an online coder might not have been able to respond accurately to such short fixations, biasing the data towards longer latency values.

Missing values.

Perhaps the strongest contender for the source of the difference, is the treatment of missing values. Missing values for the second fixation (which occurred when infants only made a single fixation in the course of a trial) were replaced with the maximum possible value of 8 seconds (the full duration of picture presentation). Inclusion of these high values may have inflated the latency means. As the authors do not report the proportion of trials in which infants fixated only a single picture, it is difficult to gauge the impact of missing values on their analysis. Separating the values

for first fixation (time-to-fixate) and first saccade away (time-to-disengage) removes this complication, as only trials with a valid value are included in analysis of each measure.

In summary, all four differences may have contributed to the large difference in RTs reported in this study and in the study by Golinkoff and colleagues.

Interim conclusions

Infants can demonstrate lexical comprehension in both macro- and micro-level measures in an IPL task with no pre-naming picture familiarisation period, and no repetition of items. Younger infants are less reliable at demonstrating lexical comprehension, and were particularly affected by time pressure in the short timing condition. The influence of timing condition was particularly pronounced in the macro-level measures. In particular, 18-month-olds did not show target discrimination in LLK, nor in PTL in the short timing condition.

The counterbalancing of target items and interest levels within and across subjects was sufficient to 'balance out' a-priori image preferences which may have otherwise skewed the results. However, detailed analyses demonstrated that picture interest had a large effect on looking behaviour in this un-baselined task. The micro-level measures (time-to-fixate and time-to-disengage) were less affected by systematic differences between items, suggesting that online language processing tasks override general picture preference, but that these effects may be short lived (that is, limited only to the first fixation or two).

It appears that target discrimination by 18- and 24-month-olds infants is detectable in the words-before-pictures version of the IPL task, which features a one-

shot stimulus presentation, no pre-naming picture familiarisation time, and a rapid presentation of target words. The short presentation duration was clearly too short to be valuable in the future, especially for the younger infants, and there was some suggestion that the long presentation duration was possibly too long for some RT measures to remain diagnostic for older infants. The following recommendations are proposed: a) picture presentation longer than 1,500ms; b) picture presentation shorter than 5,000ms; c) careful counterbalancing of picture interest; d) larger number of infants in each subject group. Chapter 5 introduces these adaptations.

Table 4.4. Summary of best-fitting categorical regression models of PTL

Group	Model	Type	Coefficient(s)	*F*	*df*	*p*	Imp.	Adj. r^2
18m short	a	single	relative interest	14.17**	1, 92	<.001	1.00	.13
	b	single	look direction (dir)	11.11**	1, 92	.001	1.00	.10
	c	single	participant (S)	1.58~	16, 92	.094	1.00	.09
	d	multiple	interest * dir * S	3.81**	18, 92	<.001	-	.36
			relative interest	24.54**	1	<.001	(.32)	(.11)
			direction	14.42**	1	<.001	(.22)	(.08)
			S	28.38**	16	<.001	(.46)	(.16)
Group	Model	Type	Coefficient(s)	*F*	*df*	*p*	Imp.	Adj. r^2
18m long	e	single	target word	5.75**	15, 83	<.001	1.00	.46
	f	single	dir	6.02*	1, 83	.016	1.00	.06
	g	multiple	target word * dir	5.31**	16, 83	<.001	1.00	.45
			target word	74.56*	15	<.001	(1.0)	(.45)
			dir	0.01	1	.928	(.00)	(.00)
Group	Model	Type	Coefficient(s)	*F*	*df*	*p*	Imp.	Adj. r^2
24m short	h	single	target word	2.061*	15, 97	.020	1.00	.14
	i	single	dir	10.69**	1, 97	.001	1.00	.09
	j	single	trial speed	5.74*	1, 97	.019	1.00	.05
	k	single	S	1.61~	13, 97	.098	1.00	.08
	l	multiple	target word * dir	3.09**	16, 97	<.001	1.00	.26
			target word	36.38**	15	<.001	(.71)	(.18)
			dir	15.36**	1	<.001	(.29)	(.08)
Group	Model	Type	Coefficient(s)	*F*	*df*	*p*	Imp.	Adj. r^2
24m long	m	single	target word	1.86*	15, 134	.033	1.00	.08

Notes. Each model is given a different index letter. Single-factor models of variables of interest were run first, to establish relevant predictors for composite categorical regressions. Model fits of up to *p*-value of .15 are reported. Only the single best fitting predictor of item variance is reported and was included in composite regressions due to the large correlation between indicator variables (e.g., all animates had high relative interest). Only the single best-fitting composite model is reported. Figures in brackets sum to the statistic for the composite model. 'Imp' refers to the relative 'importance' of each coefficient in a model, and is used to calculate the adjusted r^2 for each coefficient in the model. *p<.05, **p<.01, ~p<.10.

Table 4.5. Summary of best fitting categorical regressions models of LLK

Group	Model	Type	Coefficient(s)	*F*	*df*	*p*	Imp.	Adj. r^2
18m short	a	single	S	2.01*	16, 53	.040	1.00	.23
	b	single	target word	1.61~	14, 53	.120	1.00	.14
	c	single	dir	8.65**	1, 53	.005	1.00	.13
	d	multiple	S * dir * target	2.39*	31, 53	.018	1.00	.45
			S	39.29**	16	<.001	(0.51)	(.23)
			look direction	13.07**	1	.002	(0.18)	(.08)
			target word	22.49**	14	<.001	(0.31)	(.14)
Group	Model	Type	Coefficient(s)	*F*	*df*	*p*	Imp.	Adj. r^2
18m long	e	single	animacy	6.80*	1, 38	.013	1.00	.13
Group	Model	Type	Coefficient(s)	*F*	*df*	*p*	Imp.	Adj. r^2
24m short	f		dir	8.62**	2, 66	<.001	1.00	.19
	g		trial speed	9.77**	1, 65	.003	1.00	.12
	h		relative interest	3.28~	1, 66	.075	1.00	.03
	i	multiple	dir * speed * interest	7.75**	3, 65	<.001	1.00	.24
			dir	8.86**	1	.004	0.45	(.11)
			trial speed	4.76*	1	.033	0.33	(.08)
			relative interest	3.63	1	.062	0.22	(.05)
Group	Model	Type	Coefficient(s)	*F*	*df*	*p*	Imp.	Adj. r^2
24m long	-	-	-	-	-	-	-	-

Notes. Each model is given a different index letter. Single-factor models of variables of interest were run first, to establish relevant predictors for composite categorical regressions. Model fits of up to *p*-value of .15 are reported. Only the single best fitting predictor of item variance is reported and was included in composite regressions due to the large correlation between indicator variables (e.g., all animates had high relative interest). Only the single best-fitting composite model is reported. Figures in brackets sum to the statistic for the composite model. 'Imp' refers to the relative 'importance' of each coefficient in a model, and is used to calculate the adjusted r^2 for each coefficient in the model. *p<.05, **p<.01, ~p<.10.

Table 4.6. Summary of best fitting categorical regression models of time-to-fixate

Group	Model	Type	Coefficient(s)	*F*	*df*	*p*	Imp.	Adj. r^2
18m short	a	single	target word	2.69**	15, 81	.003	1.00	.24
	b	single	S	7.97*	18, 81	.026	1.00	.18
	c	multiple	target * S	1.75*	33, 81	.038	1.00	.23
			target word	33.95**	15	<.001	(.62)	(.14)
			S	20.06**	18	<.001	(.38)	(.09)
Group	Model	Type	Coefficient(s)	*F*	*df*	*p*	Imp.	Adj. r^2
18m long	d	single	S	3.03**	14, 45	.005	1.00	.39
	e	single	dir	3.54~	1, 45	.067	1.00	.05
	f	multiple	S * dir	2.92**	15, 45	.006	1.00	.39
			S	38.28**	14	<.001	(.93)	(.36)
			dir	1.80	1	.190	(.07)	(.03)
Group	Model	Type	Coefficient(s)	*F*	*df*	*p*	Imp.	Adj. r^2
24m short	g	single	S	3.92**	12, 66	<.001	1.00	.35
	h	single	target word	1.52	15, 66	.134	1.00	.11
	i	single	dir	4.13*	1, 66	.046	1.00	.05
	j	multiple	S * target * dir	3.12**	28, 66	.001	1.00	.47
			S	42.45**	12	<.001	(.50)	(.23)
			target word	39.00**	15	<.001	(.41)	(.19)
			dir	9.62**	1	.004	(.10)	(.05)
Group	Model	Type	Coefficient(s)	*F*	*df*	*p*	Imp.	Adj. r^2
24m long	k	single	S	1.74	18, 55	.061	1.00	.15

Notes. Each model is given a different index letter. Single-factor models of variables of interest were run first, to establish relevant predictors for composite categorical regressions. Model fits of up to *p*-value of .15 are reported. Only the single best fitting predictor of item variance is reported and was included in composite regressions due to the large correlation between indicator variables (e.g., all animates had high relative interest). Only the single best-fitting composite model is reported. Figures in brackets sum to the statistic for the composite model. 'Imp' refers to the relative 'importance' of each coefficient in a model, and is used to calculate the adjusted r^2 for each coefficient in the model. *p<.05, **p<.01, ~p<.10.

Table 4.7. Summary of best fitting categorical regression models of time-to-disengage

Group	Model	Type	Coefficients	*F*	*df*	*p*	Imp.	Adj. r^2
18m short	-	-	-	-	-	-	-	-
Group	Model	Type	Coefficients	*F*	*df*	*p*	Imp.	Adj. r^2
18m long	a	single	dir	6.80*	1, 34	.014	1.00	.15
	b	single	S	2.14~	12, 34	.059	1.00	.29
	c	multiple	S * dir	2.14~	14, 34	.059	1.00	.32
			S	21.39**	13	<.001	(.77)	(.25)
			dir	5.52*	1	.029	(.23)	(.07)
Group	Model	Type	Coefficients	*F*	*df*	*p*	Imp.	Adj. r^2
24m short	d	single	S	2.22*	12, 53	.023	1.00	.18
	e	single	animacy	2.28~	1, 65	.097	1.00	.03
	f	multiple	S * animacy	2.47*	13, 65	.011	1.00	.23
			S	28.56**	12	<.001	(.87)	(.20)
			animacy	4.61*	1	.036	(.13)	(.03)
Group	Model	Type	Coefficients	*F*	*df*	*p*	Imp.	Adj. r^2
24m long	g	single	dir	21.21**	1, 61	<.001	1.00	.25

Notes. Each model is given a different index letter. Single-factor models of variables of interest were run first, to establish relevant predictors for composite categorical regressions. Model fits of up to *p*-value of .15 are reported. Only the single best fitting predictor of item variance is reported and was included in composite regressions due to the large correlation between indicator variables (e.g., all animates had high relative interest). Only the single best-fitting composite model is reported. Figures in brackets sum to the statistic for the composite model. 'Imp' refers to the relative 'importance' of each coefficient in a model, and is used to calculate the adjusted r^2 for each coefficient in the model. *p<.05, **p<.01, ~p<.10.

Chapter 5

Quick-Primed IPL



The previous chapter demonstrated that it is possible to observe lexical comprehension in a short IPL task with no pre-naming phase. The main advantage of this design is that it allows a priming phrase to be introduced directly before the onset of the target word, without the inclusion of intervening stimuli. The current chapter describes a 'quick-primed' experiment for 18- and 24-month-olds, where an auditory prime preceded the target word, and the target word began before the onset of the pictures. Primes were either unrelated to targets, or related by taxonomy and association. The experiment was replicated across two timing conditions and two age groups. Significant priming was observed for the older age group, in macro- and micro-level measures of target discrimination. Priming was not reliable in the younger age group. The pattern of behaviour was very similar across the age range, suggesting a tendency towards primed responding across the age and language development range. An additional advantage of this new experimental design was that the short gap between the prime and the target made it more likely to generate automatic language processing effects rather than strategic response patterns. Possible mechanisms for the observed priming effect are discussed.

Introduction

In adult priming studies, a distinction is often made between 'automatic' and 'strategic' priming. Automatic priming is understood to reflect cognitive processes such as spreading activation in the lexicon. The speed of these effects, and their replicability across a wide range of tasks and stimulus sets, is understood to reflect short-term processing effects which are beyond conscious control, and are impervious to the broader task context. Strategic priming, on the other hand, is understood to be

influenced by task effects, such as expectation, semantic matching, or the generation of a short-term response set, against which potential responses may be checked (for overview, see McNamara, 2005). Experimental manipulations used to reduce strategic priming include a) disguising the prime-to-target relationship by using a large number of filler items (e.g., Neely, 1977), b) masking the pairing of primes and targets by requiring responses to all items (e.g., Shelton & Martin, 1992), and c) shortening the inter-stimulus interval (ISI) between the presentation of the prime and the target (Moss et al., 1995; Neely, 1977). While the first two approaches reduce the *accuracy of strategy formation* by masking the task design, the third approach precludes the *use of response strategies* by speeding stimulus presentation, and isolating the high-speed effects of preliminary processing. To make the most of this third possibility, a short prime-to-target ISI was employed, to capture the early stages of any possible spreading activation. Moss, Ostrin, Tyler & Marslen-Wilson's prime-to-target ISI of 200ms (1995) was selected, on the grounds that infants' phonological processing speed is similar to adults' (Swingley, Pinto & Fernald, 1999).

As in the words-before-pictures experiment, there was a short stimulus onset asynchrony (SOA) between the onset of the target word and the onset of picture presentation. The SOA was included to allow infants to begin processing the target word before the test phase of the trial began. This delay was included to prevent the sudden onset of visual information from confounding or blocking any rapid spread of activation from prime to target, and also to ensure that semantic information generated by the pictures was not active in the early stages of target word processing.

Previous IPL studies have demonstrated that infants as young as 18 months-of-age can extract sufficient acoustic information from the first 300ms of a word to correctly identify its referent in a free-looking task (Fernald et al., 2001). For the

current experiment, a 'long' SOA of 400ms was selected to allow sufficient processing of the target word prior to picture presentation. A SOA of 420ms was piloted in the words-before pictures experiment (Chapter 4), demonstrating that the new SOA would likely be sufficient for lexical comprehension to be observed in the sequence of first fixations, and over longer periods of time. The current experiment was then replicated in a second group of participants, with a shorter SOA, of 200ms. The short SOA was selected to increase the task difficulty, as pictures would appear earlier in target word processing. Recall that in Chapter 3, a match between parental CDI report and IPL performance was observed when the task had a particular level of difficulty. In the current experiment, the diagnosticity of the task was expected to increase when the task was more difficult. For example, if infants in the long SOA condition performed at 'ceiling', it might not be possible to observe differences between the priming conditions. However, if infants performed below their usual standard in the short SOA condition, facilitatory (or inhibitory) effects of the auditory prime might become more apparent. As all experimental parameters in the two SOA conditions were identical, the participant groups are considered together in a single analysis.

Method

Participants were recruited as described in previous chapters, and sent the Oxford CDI by post. Most brought their completed CDIs with them on the day of their experimental visit. In a small number of cases, they completed the form on the day of testing, either during their visit, or returned by post shortly afterward.

Seventy two 18-month-olds were tested (40 males, 32 females; mean age: 18.2 months; range: 17.1 to 18.8). Eleven additional infants were removed and replaced at the time of test, 4 for failure to complete, 1 for fussiness, 1 for experimenter error, and 5 for parental failure to return the CDI. One infant was later removed without replacement due to data loss. Seventy two 24-month-olds were tested (34 males, 38 females; mean age: 24.1 months; range: 23.4 to 25.0). Fourteen additional infants were removed and replaced at the time of test, 2 for failure to complete, 6 for parental failure to return CDI forms, 1 for parental failure to fill out a CDI form appropriately, and 5 for experimenter error. One infant was later removed without replacement due to parental interference during test. There were thus 71 18-month-olds and 71 24-month-olds for whom both behavioural and CDI data were available for analysis.

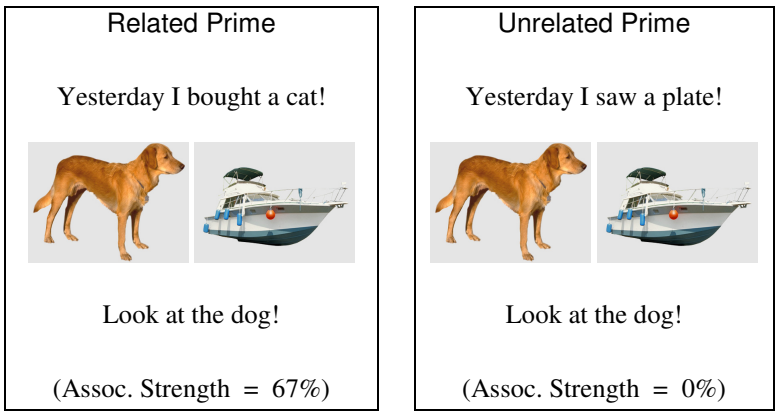
Total receptive CDI scores were checked against previously collated norms, and infants' comprehension of test items was assessed. Following preliminary analysis, three eighteen-month-olds were removed from analysis for extremely low receptive CDI scores (in the 5th percentile of previously collected CDIs for 18-month-olds). To ensure that individual differences between participants did not outweigh differences between priming conditions, four further eighteen-month-olds were removed from analysis when it was observed that they contributed trials to only one priming condition.

Materials

As in the previous chapters, stimuli for the current experiment were selected from the pool of words understood by more than 50% of 18-month-olds, established in Chapter 2 (Table 2.7). Thirty six words were selected for use as primes, targets and distracters, arranged into two stimulus lists. As in the minimally primed IPL

experiment (Chapter 3), picture pairs were yoked across two stimulus lists, such that each target occurred in each priming condition. 'Related' word-pairs were basic-level taxonomic sisters (e.g., prime: *cat*; target: *dog*) with an attested forward association in adult British English (Moss & Older, 1996 and cf. Appenndix D). The related word-pairs were equivalent to the 'compound' priming condition in the minimally primed experiment. Unrelated word-pairs shared no semantic or associative relationship, and no phonological onset or rhyme (e.g., prime: *plate*; target: *dog*). The priming conditions are illustrated in Figure 5.01.

Figure 5.01. A yoked picture-pair in two priming conditions



Notes. Association strength from Moss and Older (1996), c.f., Appendix D.

With only two priming conditions in the current experiment, it was possible for each prime to function as both a related prime and as an unrelated prime, across lists. The counterbalancing of both primes and targets across lists makes interpretation of item- or list-effects more straightforward than in the minimally primed experiment. As in previous experiments, distracters shared no phonological, semantic or associative relationship with the prime or target (e.g., prime: *cat*, target: *dog*, distracter: *boat*). Within a stimulus list, half of the primes were related and half

unrelated, and no stimulus was repeated. Each infant saw 12 trials from a single stimulus list. The stimulus lists are given in Appendix G.

In the current stimuli configuration, all infants experienced the same 12 items labelled as targets, consistently paired with the same 12 unnamed distracters. This differs from the words-before-pictures experiment (Chapter 4), in which all items acted equally often as targets and as distracters across participants. The careful counterbalancing of interest levels within the stimulus list was conducted to balance out a-priori differences in the interest of individual targets and distracters.

In the construction of stimulus lists, stimuli were ranked according to three approximate interest levels. In the 'high' interest level, *animals* (all mammals) were expected to be visually interesting to infants because of the generally high interest of faces, and the property of autonomous motion. In the medium interest level, *vehicles* were expected to be moderately interesting due to their visual complexity, and the property of motion, and *foods* were expected to be moderately interesting due to the salient property of taste. In the 'low' interest level, *furniture*, *clothing* and *tableware* were expected to be less visually interesting than items in the other levels. Target and distracter pairs were counterbalanced across the interest levels such that each list contained four targets and four distracters from each interest level, and for each high interest target paired with a low interest distracter, there was also a low interest target paired with a high interest distracter.

Visual stimuli were selected and edited as in previous chapters, with the exception that pictures appeared on a 10% grey background. Visual stimuli are given in Appendix H. Audio stimuli were recorded and edited as in previous chapters. As in the minimally primed experiment, primes were recorded at the end of a short sentence (e.g., *Yesterday, I saw a cat*!). The durations of priming phases are given in Table 5.1.

Table 5.1. Duration of prime phrases

Prime phrase	Duration (ms)
Yesterday, I ate an apple	2225
Yesterday, I ate a biscuit	2125
Yesterday, I bought a boot	2365
Yesterday, I saw a cat	2140
Yesterday, I bought a cot	2135
Yesterday, I saw an elephant	2593
Yesterday, I bought a hat	2120
Yesterday, I saw a lorry	2064
Yesterday, I saw a pig	1931
Yesterday, I bought a plate	1898
Yesterday, I saw a sheep	2171
Yesterday, I saw a train	2030
M (*SD*)	2150 (187)

Notes. Onsets and offsets of tokens established using visual waveform editing software.

Given the short ISI between the prime and the target, targets were recorded as single-word utterances. Targets were spoken in high-affect citation form, as though the speaker were excited, or surprised to see the target. The pragmatics of the auditory stimuli were designed to mimic a plausible conversation shift, as might be encountered if a parent happened to be talking about a non-present object when a new object suddenly came into view. Note that the single-word delivery of the target word differs from the presentation of targets in both previous experiments. The duration of target words is given in Table 5.2.

Table 5.2. Duration of target words

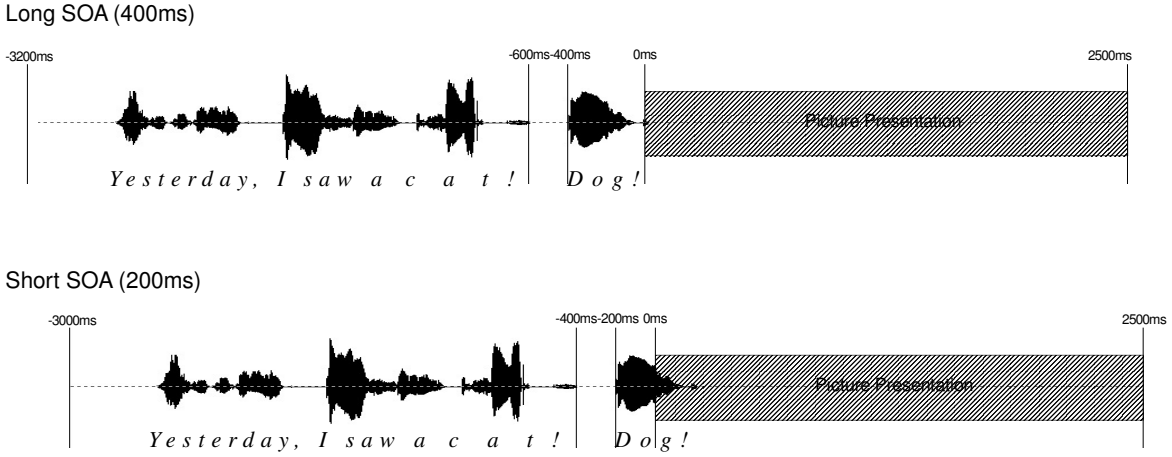
Target word	Duration (ms)
banana	650
bed	380
bus	496
cake	498
car	452
coat	569
cow	695
cup	484
dog	411
horse	742
mouse	697
shoe	535
M (*SD*)	551 (120)

Notes. Onsets and offsets of tokens established in visual waveform editing software.

Following a short 'settling in' period in the play area, the 'infant timed' procedure was conducted as described as in previous chapters, with the exception of trial timing. While the screen was blank, the priming phrase began (e.g., *Yesterday, I saw a cat!*), followed by an inter-stimulus interval (ISI) of 200ms, then the target word in isolation (e.g., *Dog!*). Presentation of pictures began at a stimulus onset asynchrony (SOA) of 400ms or 200ms from the onset of the target word, depending on participant group. Picture pairs remained onscreen for 2,500ms. A schematic of trial timing in the two SOA conditions is given in Figure 5.02, where the onset of picture presentation is treated as the origin (0ms). Note that all target words overlap with the onset of picture presentation, with the exception of the shortest target word (*bed*), in the long SOA condition. Trial order was randomised on presentation. Target side was counterbalanced. Infants were tested in the same testing booth and with the

same screen configuration as in previous experiments. A photograph of the testing booth and a diagram of the screen configuration are given in Appendix E.

Figure 5.02. Schematic of trial timing



Origin (0ms) indicates the onset of the test-phase of the trial, when pictures are presented on the screen. Wave-forms from actual stimulus tokens, as produced by GoldWave visual waveform editing software (1995, v5.10).

Scoring & measures

Scoring was conducted as in previous experiments, with the exception that frame-by-frame coding was conducted using custom software. Macro- and micro-level timing measures introduced in the previous chapters were calculated: Proportion of target looking (PTL) and the difference in the length of the longest look duration (LLK) assessed the relative preference for the named target picture during the course of the whole trial; and two micro-level measures of reaction time (RT), time-to-fixate, and time-to-disengage, assessed the speed of infant eye-movements.

In the priming context, the primary concern of the analysis is not whether target discrimination is evident, but whether there is a difference in the pattern of behaviour when the same target is preceded by two different primes. The macro-level

measures (PTL and LLK) inform our understanding of how auditory primes might influence the overall interest in named targets over the two and a half seconds of picture presentation. It is predicted that if related primes increase the activation of target words via spreading activation, then when the target word is heard, the activation will be reinforced, leading to more interest in the target picture. It is also possible, however, that increased activation in the target word might lead to enhanced processing efficiency, and therefore *less* interest in targets over the full duration of the trial. For this reason, the micro-structure of the trial and the sequence of first fixations is also valuable, as high-speed RT measures are understood to index the ease and speed of linguistic processing under different lexical conditions. Thus, if primed targets are identified faster, or primed distracters rejected faster, it would indicate that the auditory prime has caused a difference in the real-time processing of the match between the target word and the first fixated picture.

Results

Age group and vocabulary size

CDI scores were collated for two purposes – to establish which test words each individual knew at the time of test, and to generate an estimate of total vocabulary size for each infant. The mean receptive CDI score for 18-month-olds was 197 words (SD = 65) out of a possible 416. Twenty-four-month-olds had larger vocabularies than 18-month-olds ($z(71, 63) = 7.91$, $p < .0005$), with a mean of 318 words (SD = 67). At 18 months-of-age, participants in the long SOA group had smaller vocabularies than those in the short SOA group ($z(33, 30) = 1.95$, $p = .05$). This between-group difference in vocabulary size did not create any difference in 18-

month-olds' looking behaviour[12], nor did it affect the number of test items known by the two SOA groups. At 24 months-of-age there was no difference in the vocabulary sizes of the two SOA groups.

Eighteen-month-olds were reported to understand a mean of 9 of the 12 primes used as test stimuli (*SD* = 2.5) and 10 of the 12 target words (*SD* = 2.2). 24-month-olds were reported to understand a mean of 11 primes (*SD* = 1.4) and 11 targets (*SD* = 1.2). Twenty four-month-olds knew more test items than 18-month-olds (primes: $z(71, 63) = 5.77$, $p < .001$; targets: $z(71,63) = 5.92$, $p < .001$).

Lexical comprehension and trial exclusions

Of the original 756 trials available for analysis (18m: 756 trials, 24m: 852 trials) only those trials in which both the prime and the target were reported as known were included in analysis, resulting in the removal of 253 trials (33%) from 18-month-olds, and 78 trials (9%) from 24-month-olds. Trials in which infants did not fixate the picture areas of the screen were also removed from the analysis, accounting for a further 3% of remaining trials (38 of 1277 trials). The final number of valid trials in each condition is given in Table 5.3.

[12] A vocabulary-matched sub-sample of 18-month-olds was selected to check whether differences in mean vocabulary size predicted differences in looking behaviour. Infants reported to understand less than 100 words or more than 300 words were removed from this reduced analysis, resulting in the removal of four infants from each SOA group. Vocabulary differences were no longer evident in this sub-sample ($z(55) = 354$, n.s.). The sub-sample produced the same pattern of results as reported in the main analysis, and is omitted for brevity.

Table 5.3. Number of trials available in each subject group

	Timing condition		
	Long SOA (400ms)	Short SOA (200ms)	Total
18-month-olds	230 (47%)	256 (53%)	486 (100%)
24-month-olds	363 (48%)	390 (52%)	753 (100%)

The smaller number of trials available for 18-month-olds was predicted to generate larger variance in this age group, complicating a straightforward analysis of age. For this reason, analyses were conducted separately for each age group. Finally, for the analysis of RT measures, 283 of the remaining trials (23%) were excluded because the infant looked to one side of the screen prior to the onset of pictures.

Participant means were generated for each of the macro-measures, proportion of target looking (PTL) and longest look difference (LLK). As in the previous chapter, the direction of the first fixation was also used to calculate the percentage of correct (target) first fixations, and was treated as a independent variable in the RT measures, time-to-fixate and time-to-disengage.

Preliminary analysis of stimulus list

In this experiment, the same target pictures were shown in both priming conditions. Therefore within each list, six targets shared a priming condition; *dog, cow, banana, shoe, cup* and *bed* were all preceded by related primes in List A, and unrelated primes in List B, while *horse, mouse, cake, bus, car* and *coat* were preceded by unrelated primes in List A, and by related primes in List B. If targets in one sub-list happen (by chance) to be more visually interesting than targets in the other sub-list, then the high-interest pictures could attract more looking overall. Such an effect might generate discrepancies in the pattern of behaviour for participants who saw the

different stimulus lists, as related primes would be from the more interesting sub-list for one group of infants, and from the less interesting sub-list for the other group. Preliminary inspection of the data indicated that infants exposed to the two lists performed somewhat differently, suggesting that this was indeed the case. To account for a-priori biases between the items, sub-list is therefore included as a factor in all following analysis. This approach is consistent with the most rigorous standards of statistical analysis for counterbalanced conditions across the same set of items in language-based tasks (Raajmakers, 2003; Raajmakers *et al.*, 1999).

The following sections present data from the two age groups separately, followed by a comparison of the developmental and individual variables across the entire subject pool. Changes in looking behaviour over the course of the trial are considered in detailed time-course plots, and these analyses are followed by a detailed investigation of stimulus effects.

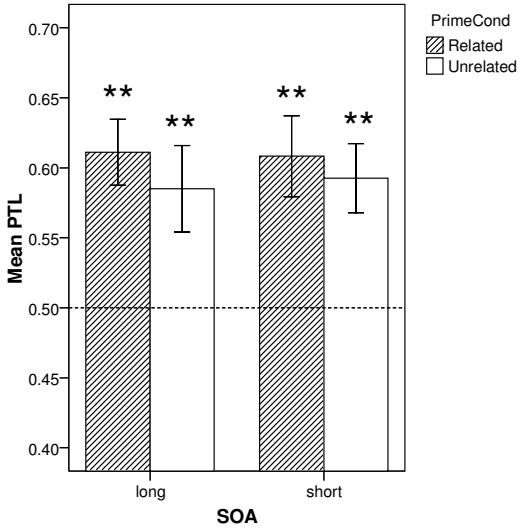
18-month-olds

PTL (18-month-olds)

Two outlying participant means were removed from analysis. Both means were more than two standard deviations below the mean, and represented a single trial with a PTL of 0, indicating that the infant had not looked at the target once during the course of the trial. Both means represented unrelated trials from infants in the long SOA group. Figure 5.03 depicts mean PTL for each SOA group, with the priming conditions plotted separately. In this plot, mean PTL appears to be above the chance level of .50 in all conditions – consistent with the previously reported finding of lexical comprehension for words marked as understood. Mean PTL appears to be slightly higher for targets in the related prime condition than in the unrelated prime

condition in both SOA groups, although the difference is very small. No difference is apparent between the two SOA conditions.

Figure 5.03. PTL according to SOA group and prime condition (18m)



Notes. Proportion of target looking (PTL). Dashed horizontal line indicates chance value of .5. One sample t-tests comparing means to chance value are marked above each bar. Independent samples t-tests comparing means for each priming condition are marked above pair. **p<.01. Whiskers show +/- one standard error.

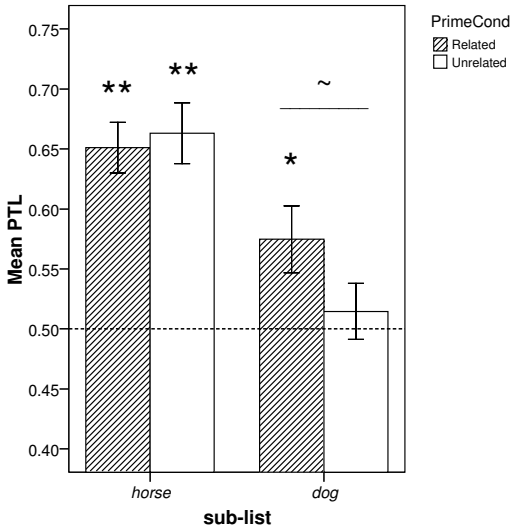
A three-level ANOVA compared the effect of SOA group (long, short), prime condition (related, unrelated) and stimulus sub-list (*horse*, *dog*), on PTL. The results are given in Table 5.4, where there was a highly significant main effect of stimulus sub-list. This finding demonstrates that stimuli in the two sub-lists reliably generated different looking patterns, an effect most likely due to a-priori differences in picture salience.

Table 5.4. Three-way ANOVA of PTL (18m)

Source (between S)	*df*	*F*	*p*	partial η^2
SOA	1	0.02	.879	<.001
prime	1	1.02	.314	.009
sub-list	1	20.24**	<.001	.152
SOA * prime	1	0.12	.734	.001
SOA * sub-list	1	0.98	.325	.009
prime * sub-list	1	2.07~	.153	.018
SOA * prime * sub-list	1	0.62	.434	.005
S w/in group error	113	(0.02)		

Data were pooled across SOA group for simple comparisons. Figure 5.04 shows PTL for each of the two stimulus sub-lists, separated across priming conditions. Concerning the main effect of stimulus sub-list, mean PTL for targets in the *horse* sub-list was .66 (*SD* = .13), and for targets in the *dog* sub-list, .55 (*SD* = .15). Thus, target preference was significantly above the chance level of .50 in both sub-lists (*horse*: $t(57) = 9.55$, $p < .001$, $d = 2.53$; *dog*: $t(62) = 2.48$, $p = .02$, $d = 0.63$), indicating that infants spent more time looking at targets that at distracters in both stimulus sub-lists. However, PTL in the *horse* sub-list was significantly higher than in the *dog* sub-list ($t(119) = 4.44$, $p < .001$, $d = 0.82$), confirming that targets in one sub-list were more visually interesting than the other.

Figure 5.04. PTL according to sub-list and prime condition (18m)

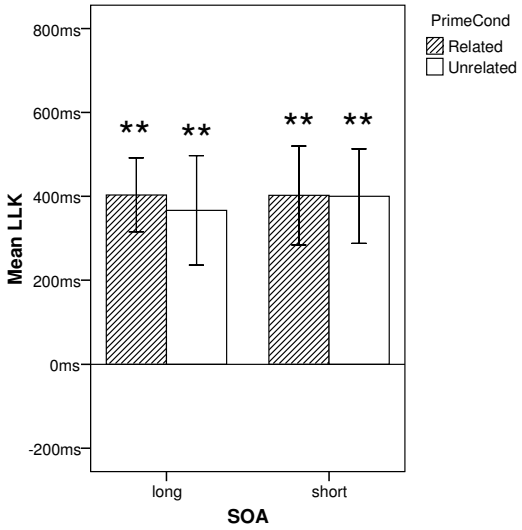


Notes. Proportion of target looking (PTL). Dashed horizontal line indicates chance value of .50. One sample t-tests comparing means to chance value are marked above each bar. Independent samples t-tests comparing means for each priming condition are marked above pair. **p<.01, *p<.05, ~p<.10. Whiskers show +/- one standard error.

LLK (18-month-olds)

One mean was removed from the 18-month-old data to satisfy Gaussian distribution in all levels of the LLK distribution. This mean represented a single trial in which the difference in longest look was -2500ms, indicating that the infant was looking in the distracter direction at the onset of pictures, and did not shift her gaze throughout the trial. Mean LLK in the different SOA groups is plotted in Figure 5.05, where a similar pattern of results is visible to that observed in PTL. That is, LLK was above the chance level of 0ms, with little difference between timing conditions and little difference between priming conditions.

Figure 5.05. LLK according to SOA group and prime condition (18m)



Notes. Longest Look Difference (LLK). Horizontal line indicates chance value of 0ms. One sample t-tests comparing means to chance value are marked above each bar. Independent samples t-tests comparing means for each priming condition are marked above pair of bars where significant. $^{**}p<.01$. Whiskers show +/- one standard error.

A three-way ANOVA compared the effect of SOA group (long, short), prime type (related, unrelated) and sub-list (*horse, dog*) on PTL. The results of the ANOVA are given in Table 5.5, where a highly significant main effect of stimulus sub-list was again observed. Figure 5.06 depicts mean LLK according to target sub-list and priming condition. There were higher LLK values for targets in the *horse* sub-list than in the *dog* sub-list (*horse*: $M = 608$ms, $SD = 580$ms; *dog*: $M = 192$ms, $SD = 588$ms; $t(120) = 3.92$, $p < .001$, $d = 0.72$). LLK in both sub-lists was above the 0ms level predicted by chance (*horse*: $t(58) = 8.05$, $p < .001$, $d = 2.12$; *dog*: $t(62) = 2.59$, $p = .01$, $d = 0.66$), indicating that infants showed greater target preference for the named picture in both sub-lists.

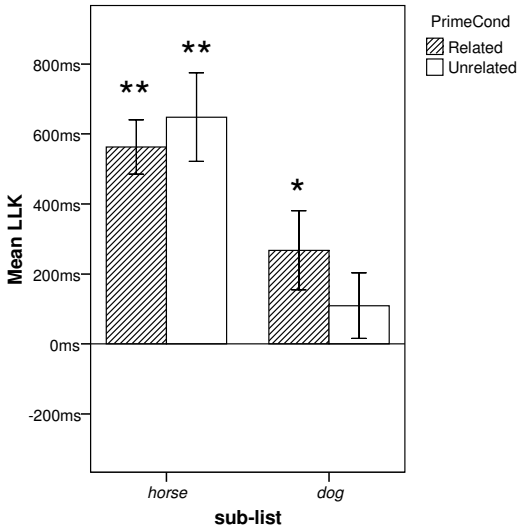
Table 5.5. Three-way ANOVA of LLK (18m)

Source (between S)	*df*	*F*	*p*	partial η^2
SOA	1	<0.01	.990	<.001
prime	1	0.18	.672	.002
list	1	15.32**	<.001	.118
SOA * prime	1	0.07	.789	.001
SOA * list	1	1.34	.249	.012
prime * list	1	1.28	.260	.011
SOA * prime * list	1	0.15	.703	.001
Error	114	(350129.13)		

Notes. Mean square error given in brackets. **p<.01.

Simple effects confirmed that there was no significant difference between the priming conditions for targets in the *horse* sub-list (related: M = 562ms, SD = 410ms; unrelated: M = 648ms, SD = 703ms; $t(57) = 0.56$, n.s., $d = 0.15$), nor for targets in the *dog* sub-list (related: M = 268ms, SD = 649ms; unrelated: M = 110ms, SD = 512ms; $t(61) = 1.07$, n.s., $d = 0.27$). Although the comparison did not achieve significance, it is worth noting that the difference between priming conditions in the *dog* sub-list is in the expected direction – with higher LLK for related primes than for unrelated primes. Furthermore, unrelated targets in the *dog* sub-list did not attract target looking above the chance level of 0ms, when assessed individually. The same pattern of results was shown in the analysis of PTL, where the difference resulted in a near-significant interaction of prime and item sub-list.

Figure 5.06. LLK according to sub-list and prime condition (18m)



Notes. Difference in longest look (LLK). Horizontal line indicates 'chance' value of 0ms. Single sample t-tests comparing LLK to chance value are marked above bar. Independent samples t-tests comparing priming conditions are marked above pair of bars where significant. $^{**}p<.01$, $^{*}p<.05$. Whiskers show +/- one standard error.

Taken together, the results of the two macro-level measures demonstrate that 18-month-olds show significant target preference for words listed as known by their parents (as demonstrated in previous experiments), although in both measures, they failed to show above-chance target preference for unrelated targets in the (less interesting) *dog* sub-list. There was a significant difference in performance between the two stimulus sub-lists, suggesting that a-priori target salience generated more looking to the targets in the *horse* sub-list than to targets in the *dog* sub-list. Although the priming condition did not interact significantly with the general effect of stimulus salience, there was a trend toward higher values for related primes in the *dog* sub-list. This trend was apparent in both measures, but only approached statistical reliability in the PTL measure.

Direction of first look (18-month-olds)

For investigation of the reaction time measures, it is necessary to consider the direction of the first fixation in each trial. Of the 375 trials available for latency analysis (in which the infant did not fixate the left or right area of the screen prior to the onset of images), 18-month-olds fixated the distracter first in 163 trials (43%), and the target first in 212 trials (57%). A one-way chi-square test confirmed that there were significantly more first fixations to the target than to the distracter (adjusted[13] $\chi^2(1) = 6.40$, $p = .01$). This finding differs from the results of the words-before-pictures experiment (Chapter 4), in which first fixations were equally distributed between targets and distracters.

Table 5.6 gives the number of first fixations in each direction, according to the priming condition of the trial. In this table it is clear that the number of first fixations to targets is higher in both priming conditions, but no difference in distribution is evident between the priming conditions.

Table 5.6. Number of first fixations to target and distracter according to prime type

	Look Direction	
	Target	Distracter
Related Prime Condition	103	80
Unrelated Prime Condition	109	83

No relationship between direction of first fixation and priming condition was evident in a chi-square test of direction (target, distracter) and prime (related, unrelated), (adjusted $\chi^2(1) < 0.01$, n.s.). This was also the case when data were considered separately according to SOA group (long: adjusted χ^2 (1) = 0.74, n.s.; short: adjusted

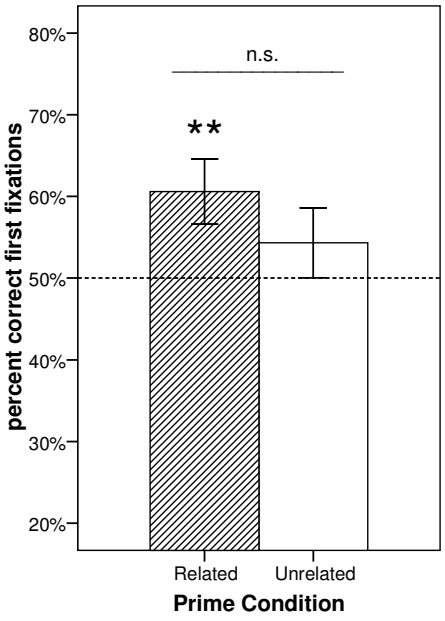
[13] Yates' continuity corrected statistic is given for Pearson's χ^2 when degrees of freedom equal 1.

$\chi^2(1) = 0.43$, n.s.), and also when sub-list was considered (*horse* sub-list: adjusted $\chi^2(1) = 0.06$, n.s.; *dog* sub-list: adjusted $\chi^2(1) < 0.01$, n.s.).

The number of trials in which infants' first fixation was to the target was taken as a percentage of the number of valid trials contributed by the infant, to generate a percentage of 'correct' first fixations. Figure 5.07 depicts the percent of correct first fixations for 18-month-olds, separated by prime condition. In this plot, the percentage of target fixations is slightly higher for trials in the related prime condition ($M = 61\%$, $SD = 31\%$), than in the unrelated condition ($M = 54\%$, $SD = 33\%$) a comparison which did not achieve significance ($t(119) = 1.07$, n.s., $d = 0.20$). However, note that it is only in the related prime condition that the number of target fixations is significantly above chance.

As the number of looks distributed between targets and distracters is asymmetrical, fixation direction is considered as an independent variable for further RT analysis.

Figure 5.07. Percent correct first fixations (18m)



Notes. One sample t-tests comparing mean to chance value of 50%, are marked above bar. Independent samples t-tests comparing priming conditions are marked across pair-of-bars. **p<.01.

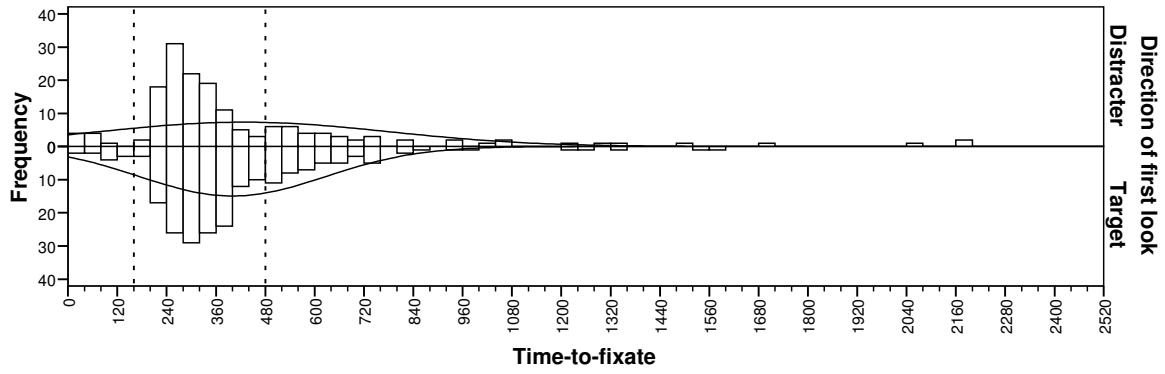
Time-to-fixate (18-month-olds)

Figure 5.08 shows the distribution of first fixations over the course of the trial, according to the direction of first fixation. The visual inspection method introduced in the previous chapter was used to establish an appropriate 'window' for analysis of the RT measure. In this distribution, a small spike of early fixations is evident in the first few hundred milliseconds, prior to the bulk of the trials. Note that the window of 160ms to 480ms is similar to the window selected in the previous, words-before-pictures chapter (80ms to 520ms).

Following selection of the analysis window, individual trial data were considered, without the generation of participant means. Recall that each infant saw six trials in each of the sub-lists, half of which were in each priming condition. With the inclusion of look direction as an independent variable in the analysis, there were

insufficient trials in each of the conditions to generate valid 'means' (half of the 'mean' values would be generated from a single trial). Instead, data from each trial were considered independently.

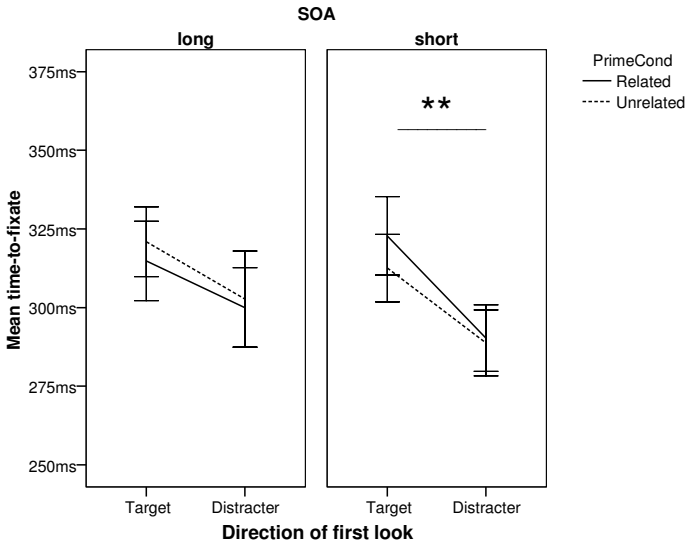
Figure 5.08. Histogram of time-to-fixate over the course of the trial (18m)



Notes. Analysis window marked from 160ms to 480ms. Normal curves plotted over whole data set (prior to exclusion).

Figure 5.09 depicts time-to-fixate in the two SOA groups, according to the priming condition and the direction of first fixation. In this plot, fixations to targets appear to be slower than fixations to distracters in both SOA groups, and little difference is apparent between priming conditions. The finding of slower fixations to targets is consistent with the results of the words-before-pictures experiment (Chapter 4), where the finding was interpreted as a processing cost in the presence of a match between the target picture and the spoken word. This finding replicates the previous result, with new items and a larger population of subjects. To further investigate the possible influence of SOA, prime condition and stimulus effects on the speed of the first fixation, parametric analyses were conducted.

Figure 5.09. Time-to-fixate according to look direction and SOA group (18m)



Notes. Independent samples t-tests comparing look direction are shown above comparison. **p<.01. Whiskers show +/- one standard error.

A four-level ANOVA was conducted on time-to-fixate, comparing the effects of fixation direction (target, distracter), SOA group (long, short), priming condition (related, unrelated), and stimulus sub-list (*horse*, *dog*). The results of the ANOVA are given in table 5.7, where a highly significant main effect of look direction was observed, along with a highly significant interaction between look direction and stimulus sub-list. The same pattern of results was obtained when separate ANOVAs were conducted for looks in each direction.

To clarify the main effect of look direction, simple comparisons were conducted. The overall mean time-to-fixate was 310ms (SD = 60ms), with slower first fixations in the target direction (M = 318ms, SD = 70ms) than in the distracter direction (M = 294ms, SD = 62ms). The difference between look directions was significant ($t(256) = 2.80$, $p = .005$, $d = 0.35$), demonstrating that 18-month-olds' looking pattern was consistent with target discrimination prior to landing the first fixation. This finding parallels the results of infants in the words-before-pictures

experiment (Chapter 4), and provides further evidence supporting the finding that infants are able to extract semantic information from para-foveal stimuli very quickly, and that 'matching' stimuli to visual information has a processing cost for infants, as demonstrated by their slower fixation times.

Table 5.7. Four-way ANOVA of time-to-fixate (18m)

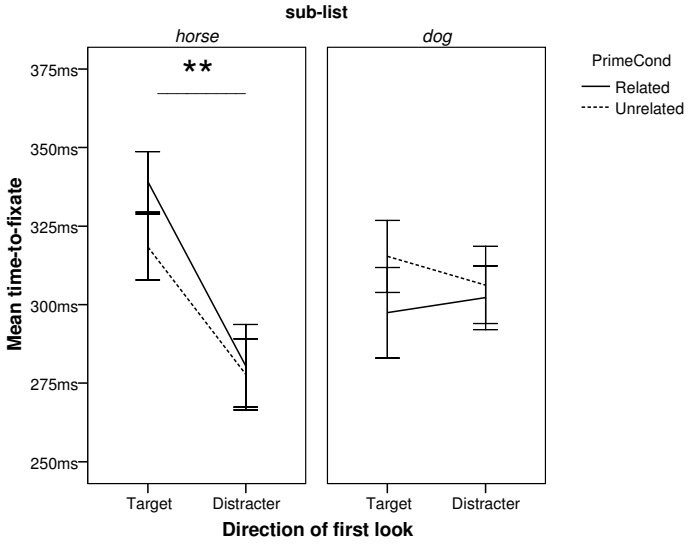
Source	*df*	*F*	*p*	partial η^2
direction (dir)	1	7.12**	.008	.029
SOA	1	0.34	.559	.001
prime	1	0.02	.901	.000
sub-list	1	0.04	.852	.000
dir * SOA	1	0.35	.555	.001
dir * prime	1	0.01	.935	.000
SOA * prime	1	0.53	.468	.002
dir * SOA * prime	1	0.09	.769	.000
dir * sub-list	1	7.29**	.007	.029
SOA * sub-list	1	1.43	.233	.006
dir * SOA * sub-list	1	0.04	.847	.000
prime * sub-list	1	1.01	.315	.004
dir * prime * sub-list	1	1.14	.287	.005
SOA * prime * sub-list	1	3.05~	.082	.012
dir * SOA * prime * sub-list	1	0.47	.496	.002
S w/in group error	242	(4427.53)		

Notes. Mean square error given in brackets. ~p<.10, **p<.01.

To clarify the significant interaction of look direction with stimulus sub-list, Figure 5.10 depicts the mean time-to-fixate targets and distracters in each of the two stimulus sub-lists, with the priming conditions also shown. In this plot, it is clear that target discrimination (slower first fixations to named targets) was only evident for targets in the *horse* sub-list (target: M = 328ms, SD = 64ms; distracter: M = 279ms, SD = 56ms; $t(118) = 4.16$, $p < .001$, $d = 0.77$), while no clear pattern of target

discrimination was apparent in the *dog* sub-list (target: *M* = 307ms, *SD* = 75ms; distracter: *M* = 304ms, *SD* = 64ms; adjusted[14] *t*(133.8) = 0.28, n.s., *d* = 0.05). It is interesting to note, however, that there was a tendency toward priming in the horse sub-list, with slower looks to related targets than to unrelated targets. The comparison did not achieve significance, with a *p*-value of .15 barely approaching the alpha-level of .05, suggesting that the interaction is not reliable. However, this weak trend appears in the *horse* sub-list, while the trend toward priming observed in the macro-level measures (PTL and LLK) was for items in the *dog* sub-list.

Figure 5.10. Time-to-fixate according to sub-list (18m)



Notes. Independent sample t-tests comparing look direction are shown below comparison. Independent samples t-test comparing priming condition are shown above the pair of values. **p<.01. Whiskers show +/- one standard error.

Concerning the near-significant three-way interaction of prime with SOA and sub-list, Table 5.8 gives the mean time-to-fixate for participants in each SOA group, according to the stimulus sub-list, with the results of independent samples *t*-tests comparing priming conditions. For 18-month-olds in the short SOA group, stimuli in

[14] The adjusted *t*-value and adjusted degrees of freedom are reported, due to unequal variances identified in a Levene's test ($F(136) = 4.52$, $p < .05$).

the *horse* sub-list (targets and distracters pooled) were fixated significantly slower in the related prime condition than in the unrelated prime condition. No other subsets of the data were affected.

Table 5.8. Mean time-to-fixate according to SOA group and sub-list (18m)

		Priming condition					
		Related	Unrelated				
SOA	sub-list	*M (SD)*	*M (SD)*	*df*	*t*	*p*	*d*
long	*horse*	301 (50)	314 (74)	54	0.75	.458	0.204
	dog	315 (78)	315 (72)	60	0.14	.989	0.004
short	*horse*	333 (72)	296 (61)	62	2.22*	.030	0.564
	dog	290 (63)	307 (69)	74	1.10	.275	0.235

Notes. Independent samples t-tests compare priming conditions. *p<.05

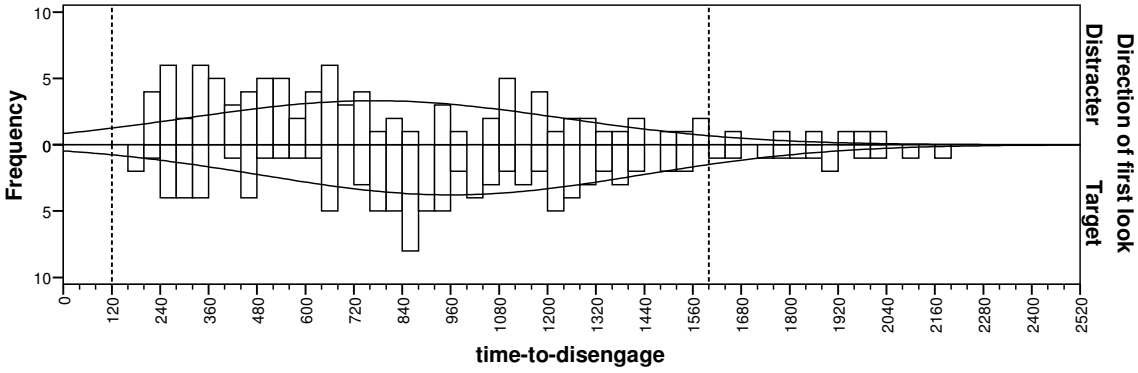
These findings demonstrate that the speed of the first fixation was affected by the name of the picture. Where the first fixation was to the *target*, the time-to-fixate was *slower* than when the first fixation was towards the *distracter*. Thus, when 18-month-olds take more time to initiate their first eye movement, they are more likely to fixate the named target in their first look. Fixation speed was also affected by particular stimulus combinations, but the prime condition did not account for a significant portion of the variance in this age group. The finding is consistent with the pattern of behaviour demonstrated in the macro-level measures.

Time to disengage (18-month-olds)

For the analysis of time-to-disengage, only those trials which were normally distributed in the time-to-fixate measure were included in analysis. Figure 5.11 shows the distribution of those trials, separated by look condition. Following manual inspection of histograms, trials with a time-to-disengage of over 1600ms were

excluded from analysis. The data were not normally distributed at the overall level, nor within subsets of the data. However, given that the univariate ANOVA is considered to be robust to deviations from normality so long as homogeneity of variance is present between subsets, analysis by ANOVA was conducted. Non-parametric tests were used in place of simple comparisons.

Figure 5.11. Histogram of time-to-disengage over the course of a trial (18m)

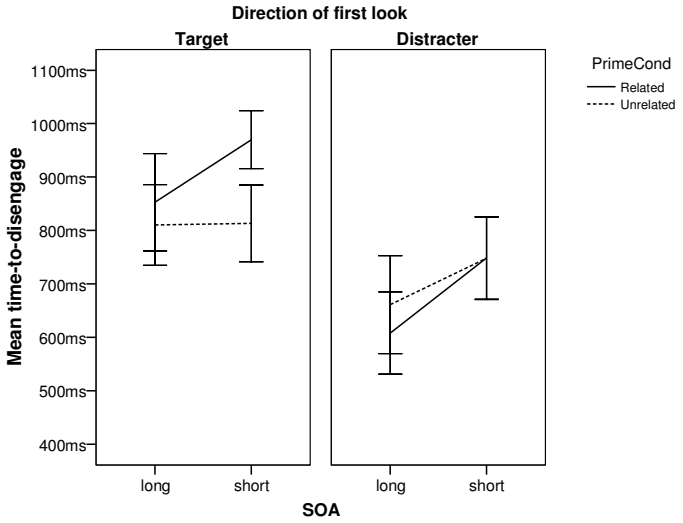


Notes. Analysis window marked from 120ms to 1600ms. Normal curves plotted over whole data set (prior to exclusion).

Figure 5.12 depicts time-to-disengage in each fixation direction, separated according to SOA group and priming condition. In this plot, when the first fixation is to the target, time-to-disengage is slower than when the first fixation is to the distracter. Recall that time-to-disengage is measured from the onset of the first fixation, and therefore shows that the duration of the first fixation is longer when the infant fixates the target first. This pattern of behaviour is consistent with distracter 'rejection'. In addition, time-to-disengage appears to be slightly slower in the short timing condition – as might be expected if infants needed more time to process pictures which began so soon after the onset of the target word. In this plot there also appears to be a moderate difference between the priming conditions, such that target

looks are *even longer* and distracter looks, *even shorter* when the prime was related to the target.

Figure 5.12. Time-to-disengage according to SOA and prime condition (18m)



Notes. Whiskers show +/- one standard error

To quantify these apparent effects, a four-level ANOVA of time-to-disengage was conducted, comparing the influence of direction of fixation (target, distracter), SOA group (long, short), prime condition (related, unrelated) and stimulus sub-list (*horse*, *dog*). The results of the ANOVA are given in Table 5.9 where a highly significant main effect of direction is apparent, along with a highly significant interaction between fixation direction and stimulus sub-list. The main effect of SOA also approached significance, with a *p*-value of .08 approaching the alpha-level of .05.

Concerning the main effect of look direction, first looks to targets were longer (M = 851ms, SD = 363ms) than first looks to distracters (M = 699ms, SD = 383ms), and non-parametric tests confirmed that this comparison was highly significant ($z(99, 91) = 2.91$, $p = .004$). Concerning the near-significant main effect of SOA group, first fixations in the long SOA group (M = 733ms, SD = 390ms) were

moderately shorter than first fixations in the short SOA group (M = 817ms, SD = 369ms), however the comparison did not achieve significance ($z(87, 103) = 1.58$, n.s.), with a p-value of .11. The direction of the trend is still of interest, however. Given the notion of 'time pressure' introduced in the previous chapter, 18-month-olds might have been expected to exhibit shorter fixations in the short SOA condition due to the faster 'pace' of the task. Here, they demonstrate the opposite pattern, with shorter fixations in the long SOA condition. Although the difference was unreliable, it suggests that infants who have heard more of the word prior to picture onset need less time before rejecting an unnamed distracter picture.

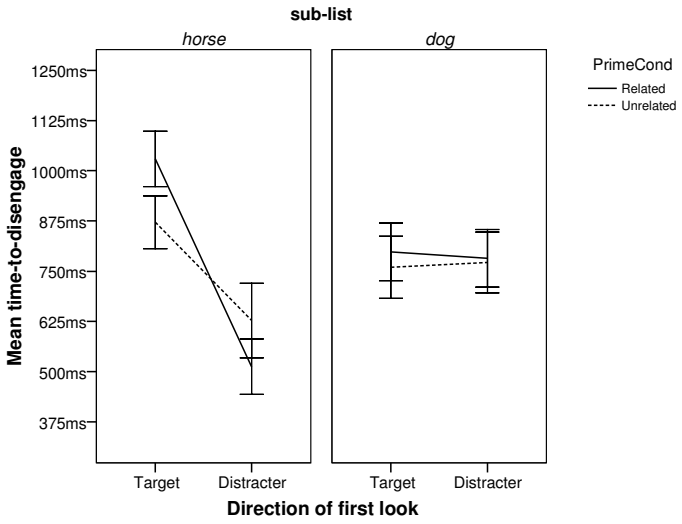
Table 5.9. Four-way ANOVA of time-to-disengage (18m)

Source	*df*	*F*	*p*	partial η^2
direction (dir)	1	11.30**	.001	.061
SOA	1	3.08~	.081	.017
prime	1	0.09	.762	.001
sub-list	1	0.02	.884	<.001
dir * SOA	1	0.09	.766	.001
dir * prime	1	1.50	.223	.008
SOA * prime	1	0.76	.385	.004
dir * SOA * prime	1	0.43	.513	.002
dir * sub-list	1	12.75**	<.001	.068
SOA * sub-list	1	0.29	.592	.002
dir * SOA * sub-list	1	0.06	.806	<.001
prime * sub-list	1	0.04	.837	<.001
dir * prime * sub-list	1	1.43	.233	.008
SOA * prime * sub-list	1	0.55	.461	.003
dir * SOA * prime * sub-list	1	1.92	.168	.011
Error	175	(135458.92)		

Notes. Mean square error given in brackets. ~p<.10, **p<.01.

Concerning the interaction of look direction with sub-list, Figure 5.13 plots time-to-disengage separately for each stimulus sub-list, according to look direction. In this plot it is clear that items in the *horse* sub-list attracted strong target discrimination, while those in the *dog* sub-list did not. Non parametric tests confirmed that longer fixations to the target were only evident in the *horse* sub-list (target: M = 934ms, SD = 325ms; distracter: M = 575ms, SD = 350ms; $z(35, 47) = 4.21$, $p < .001$), while no difference was evident in the *dog* sub-list (target: M = 777ms, SD = 383ms; distracter: M = 777ms, SD = 386ms; $z(56, 25) = 0.03$, n.s.).

Figure 5.13. Time-to-disengage according to stimulus sub-list and priming condition (18m)



Notes. Whiskers show +/- one standard error

Furthermore, in figure 5.13 the degree of target discrimination demonstrated for items in the *horse* sub-list, as indicated by the slope of the line connecting values for targets and distracters, differed according to prime condition. Table 5.10 summarises Wilcoxon's tests comparing look direction within each stimulus sub-list, separated by priming condition. This table also includes a calculation of the mean difference between the look directions (Δ-direction). In this table it is clear that target

discrimination is significant in both priming conditions, but that the effect is larger and more reliable in the related prime condition, where looks to *horse* targets are 491ms longer than looks to *horse* distracters. This can be compared to the unrelated prime condition, where looks to the target are only 245ms longer than looks to the distracter. However, bearing in mind that prime condition did not generate a significant interaction with look direction and stimulus group, this finding should be treated as a trend.

Table 5.10. Mean time-to-disengage according to SOA group and sub-list (18m)

		Direction of first fixation					
		Target	Distracter				Δ direction
sub-list		*M (SD)*	*M (SD)*	*N*	*U*	*p*	*M*
horse	Related	1004 (314)	513 (275)	56	44.5**	<.001	491
	Unrelated	872 (328)	627 (403)	52	152.5*	.044	245
dog	Related	798 (343)	782 (386)	44	384.5	.908	16
	Unrelated	760 (417)	771 (393)	38	325.5	.883	-11

Notes. Wilcoxon's tests compare look direction. **p<.01 *p<.05

Taken together, the RT measures showed that 18-month-olds demonstrate significant target discrimination in the speed of the first fixation and the time taken to shift away from the first fixated picture. Targets were fixated slower than distracters, indicating a processing cost in 'matching' an auditory string to a picture prior to fixation. First fixations to targets were maintained longer, indicating high-speed rejection of incorrectly fixated distracters. This general effect was mediated by an interaction with the stimulus sub-list in both measures, such that target recognition was only evident for items in the *horse* sub-list, and not for items in the *dog* sub-list. In both RT measures there was a moderate trend toward a priming effect for items in the *horse* sub-list.

Summary: 18-month-olds

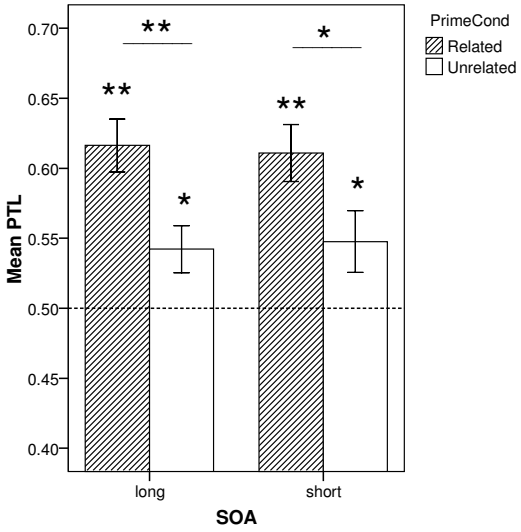
18-month-olds showed target recognition in all four measures, PTL, LLK, time-to-disengage, and time-to-fixate, for words reported by their parents as known. 18-month-olds percentage of correct first fixations was also above chance. 18-month-olds' performance was not reliably affected by the SOA condition in which they saw the stimuli, except for a minor trend in their time-to-disengage from pictures, which suggested a processing cost in the short SOA group. 18-month-olds exhibited a large difference in the salience of items in the two stimulus sub-lists, with items in the *horse* sub-list attracting more visual preference overall, and attracting more reliable target recognition in RT measures. The sub-lists also interacted with a weak tendency toward priming for 18-month-olds. Interestingly where moderate priming trends were observed for items in the *dog* sub-list in the macro-level measures, in RT measures, priming trends were evident for items in the *horse* sub-list.

24-month-olds

PTL (24-month-olds)

Figure 5.14 depicts the mean proportion of target looking (PTL) for 24-month-olds according to SOA group and priming condition. In this plot mean PTL is above the chance level of .50, indicating preference for the named target over the unnamed distracter. Mean PTL is higher in the related condition than in the unrelated condition, suggesting greater preference for named targets when they were preceded by related primes than when they were preceded by unrelated primes. Little difference is apparent between the SOA groups.

Figure 5.14. PTL according to SOA group and prime condition (24m)



Notes. Proportion of target looking (PTL). Horizontal dashed line indicates 'chance' value of .50. Single sample t-tests comparing PTL to chance are marked above each bar. Independent samples t-tests comparing priming conditions are marked above pair of bars. **p<.01, *p<.05. Whiskers show +/- one standard error.

A three-level ANOVA was conducted comparing the influence of SOA group (long, short), prime condition (related, unrelated), and stimulus sub-list (*horse*, *dog*) on PTL. The results of the ANOVA are given in Table 5.11. As in the analysis of 18-month-old PTL, a highly significant main effect of sub-list was evident. Unlike in the analysis of 18-month-olds a highly significant effect of priming condition was also evident, with no interaction between the two factors. Data were pooled across SOA group for simple comparison.

Table 5.11. Three-way ANOVA of PTL (24m)

Source (between S)	*df*	*F*	*p*	partial η^2
SOA	1	<0.01	.982	<.001
prime	1	15.66**	<.001	.105
sub-list	1	31.13**	<.001	.189
SOA * prime	1	0.65	.422	.005
SOA * sub-list	1	1.25	.266	.009
prime * sub-list	1	0.05	.825	<.001
SOA * prime * sub-list	1	0.96	.330	.007
Error	134	(0.01)		

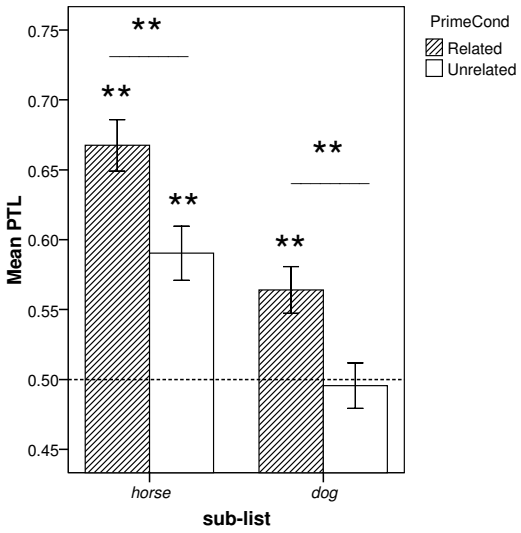
Notes. Mean square error in brackets. **p<.01

Target preference was generally above the level of .50 predicted by chance (M = .58, SD = .12; t(141) = 7.81, p < .001, d = 1.32), indicating that 24-month-olds showed significant preference for the named target over the unnamed distracter. Concerning the main effect of prime condition, PTL in the related prime condition (M = .61, SD=.12) was higher than in the unrelated priming condition (M = .54, SD = .12), and simple effects confirmed that the difference was significant (t(140) = 3.51, p < .001, d = 0.59). This finding indicates that target preference was *greater* when the named target was preceded by a related word than when it was preceded by an unrelated word.

To investigate the main effect of stimulus sub-list, Figure 5.15 illustrates PTL for items in the two stimulus sub-lists. PTL for items in the *horse* sub-list attracted higher PTL (M = .63, SD = .12), than items in the *dog* sub-list (M = .53, SD = .10), and simple effects confirmed that this difference was significant (t(140) = 5.14, p < .001, d = 0.87). This finding indicates that older infants, like younger infants, showed an effect of stimulus set. The direction of the effect again suggests that the

targets in the *horse* sub-list were more interesting to look at than the targets in the *dog* sub-list. Critically, the effect of image interest did not interact with the general effect of the prime condition; 24-month-olds showed a similar increase in primed target preference when looking at the interesting targets in the *horse* sub-list (related: $M = .67$, $SD = .11$; unrelated: $M = .59$, $SD = .12$; $t(69) = 2.87$, $p = .005$, $d = 0.69$), as they did when looking at targets in the less interesting *dog* sub-list (related: $M = .56$, $SD = .10$; unrelated: $M = .50$, $SD = .10$; $t(69) = 2.94$, $p = .005$, $d = 0.71$). This finding is important because it demonstrates that the effect of the priming relationship is evident *even when* the experimental stimuli contain a-priori differences in image interest, *and* when the trial structure precludes within-trial or within-experiment baselining.

Figure 5.15. PTL according to stimulus sub-list and prime condition (24m)

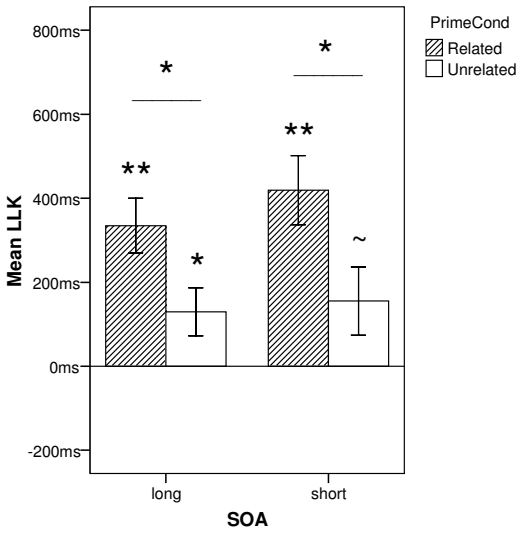


Notes. Horizontal dashed line indicates 'chance' value of .05. Single sample t-tests comparing PTL to chance are marked above each bar. Independent samples t-tests comparing priming conditions are marked above pair of bars. **p<.01. Whiskers show +/- one standard error.

LLK (24-month-olds)

Figure 5.16 depicts the longest look difference (LLK) in each of the SOA groups, with the priming conditions shown separately. In this plot, LLK is above the level of 0ms predicted by chance, indicating that 24-month-olds look longer at targets than distracters in their single longest look of the trial. It is also clear in this plot that targets preceded by related primes attract greater target preference than targets preceded by unrelated primes. Little difference is apparent between the priming conditions.

Figure 5.16. LLK according to SOA group and prime condition (24m)



Notes. Longest look difference (LLK). One-sample t-tests comparing LLK to chance value of 0ms are marked above each bar. Independent samples t-tests comparing priming conditions are marked across pair of bars. $^{**}p<.01$, $^{*}p<.05$. Whiskers show +/- one standard error.

A three-way ANOVA was conducted on LLK to investigate the effects of SOA group (short, long), priming condition (related, unrelated) and stimulus sub-list (*horse*, *dog*). The results of the ANOVA are given in Table 5.12, where there a highly significant main effect of prime condition, was observed along with a highly

significant main effect of stimulus sub-list. As in the analysis of PTL, no interaction was evident between the two effects. Data were pooled across SOA group for comparison of simple effects. In general, LLK was above the level which would be predicted by chance (M = 258ms, SD = 442ms; $t(138) = 6.89$, $p < .001$, $d = 0.58$), indicating that 24-month-olds made longer looks to named targets than they did to unnamed distracters. The longest look difference was larger in the related prime condition (M = 403ms, SD = 567ms), than in the unrelated prime condition (M = 383ms, SD = 669ms), and this difference was significant ($t(137) = 3.26$, $p = .001$, $d = 0.56$).

Table 5.12. Three-way ANOVA of LLK (24m)

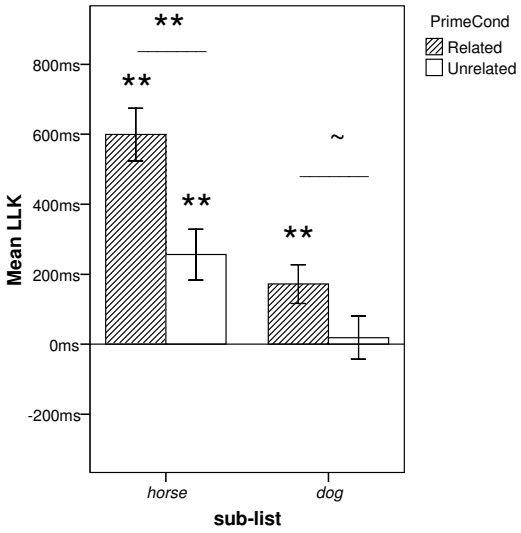
Source	*df*	*F*	*p*	partial η^2
SOA	1	0.57	.452	.004
prime	1	12.77**	<.001	.089
sub-list	1	23.55**	<.001	.152
SOA * prime	1	<0.01	.991	<.001
SOA * sub-list	1	0.77	.383	.006
prime * sub-list	1	1.61	.207	.012
SOA * prime * sub-list	1	1.33	.250	.010
Error	131	(156800.8)		

Notes. Mean square error in brackets. **p<.01

Figure 5.17 shows LLK separated according to stimulus sub-list and priming condition. Simple effects confirmed the same pattern of results as PTL; target preference was greater in the *horse* sub-list (M = 418ms, SD = 468ms) than in the *dog* sub-list (M = 97ms, SD=349ms), and this comparison was significant ($t(137) = 4.58$, $p < .001$, $d = 0.78$). Although the effect of sub-list did not interact with priming

condition, the effect of the prime was more reliable in the *horse* sub-list (related: M = 599ms, SD = 435ms; unrelated: M = 256ms, SD = 441ms; $t(68) = 3.27$, $p = .002$, $d = 0.79$), than it was in the *dog* sub-list (related: M = 172ms, SD = 327ms; unrelated: M = 19ms, SD = 359ms; $t(67) = 1.85$, n.s., $d = 0.45$), where the comparison failed to achieve significance, with a p-value of .07 approaching the alpha level of .05. This pattern is consistent with that found in PTL for the 24-month-olds; although there was a large and reliable stimulus effect, the effect of the priming condition was also very reliable, and did not significantly differ according to stimulus sub-list.

Figure 5.17. LLK according to sub-list and prime condition (24m)



Notes. Difference in longest look (LLK), measured in milliseconds. One-sample t-tests comparing LLK to chance value of 0ms are marked above bar. Independent samples t-tests comparing priming conditions are marked above pair of bars. **p<.01, ~p<.10.

Direction of first look (24-month-olds)

Of the 581 trials in which valid reaction time data was available, 24-month-olds fixated the distracter first in 245 trials (42%), and the target first in 336 trials (58%). A one-way chi-square test confirmed that there were significantly more first fixations to the target than to the distracter (adjusted $\chi^2(1) = 14.25$, p <.001). Table

5.13 gives the number of first fixations in each direction, according to the priming condition of the trial. As was found for 18-month-olds, the number of first fixations to targets was higher in both priming conditions, but the frequency distribution did not differ between the priming conditions, as confirmed by a chi-square test comparing direction (target, distracter) and prime condition (related, unrelated) (adjusted $\chi^2(1) <.001$, n.s.)[15]. This was also the case when considered separately according to SOA group (long SOA: adjusted $\chi^2(1) = 0.35$, n.s.; short SOA: adjusted $\chi^2(1) = 0.31$, n.s.;), and also sub-list (*horse* sub-list: adjusted $\chi^2(1) = 1.54$, n.s.; *dog* sub-list: adjusted $\chi^2(1) = 1.69$, n.s.).

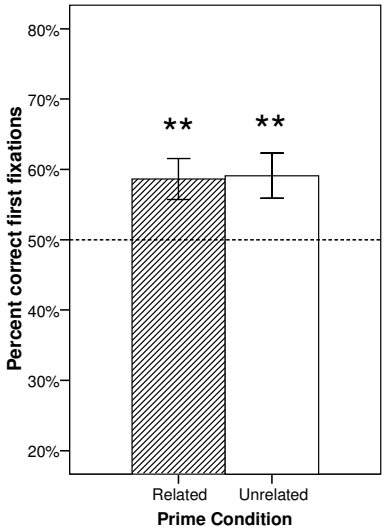
Table 5.13. Number of first fixations to target and distracter according to prime type (24m)

Priming condition	Look Direction	
	Target	Distracter
Related	162	118
Unrelated	174	127

The number of target fixations was used to generate a percentage score for each individual, of 'correct' first fixations, out of all valid trials. Figure 5.18 depicts the individual mean for the percentage of first fixations correct. As was observed for 18-month-olds, there was no difference between the two priming conditions. This finding suggests that although the first fixation was reliably directed to the named target more often than the unnamed distracter, the influence of the prime had no effect on processing at this level.

[15] Yates' continuity corrected statistic is given for Pearson's χ^2 when degrees of freedom equal 1.

Figure 5.18. Percent correct first fixations according to prime condition (24m).

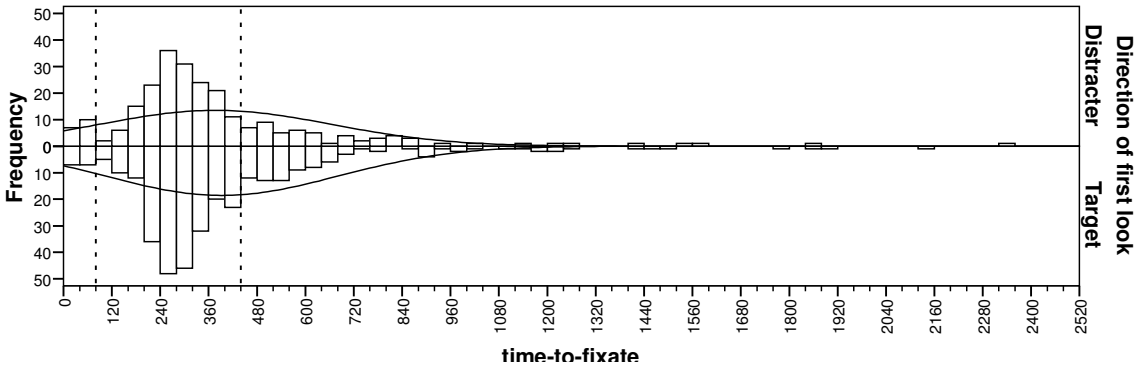


Notes. One sample t-tests comparing mean to chance value of 50%, are marked above bar. Independent samples t-tests comparing priming conditions are marked across pair-of-bars where significant. **p<.01.

Time-to-fixate (24-month-olds)

Figure 5.19 shows the distribution of first screen fixations over the course of the trial, separated by the direction of the first fixation. Visual inspection was again used to establish a 'window' for analysis, this time from 80ms to 440ms. Note that this window begins and ends slightly earlier than the window for 18-month-olds (160ms to 480ms). Again, in the RT measures, data from individual trials are considered, as it is not possible to create valid participant means in each of the experimental conditions. One further low value was removed from analysis when it was observed that it created a skew in the data from the short SOA group, in the target direction.

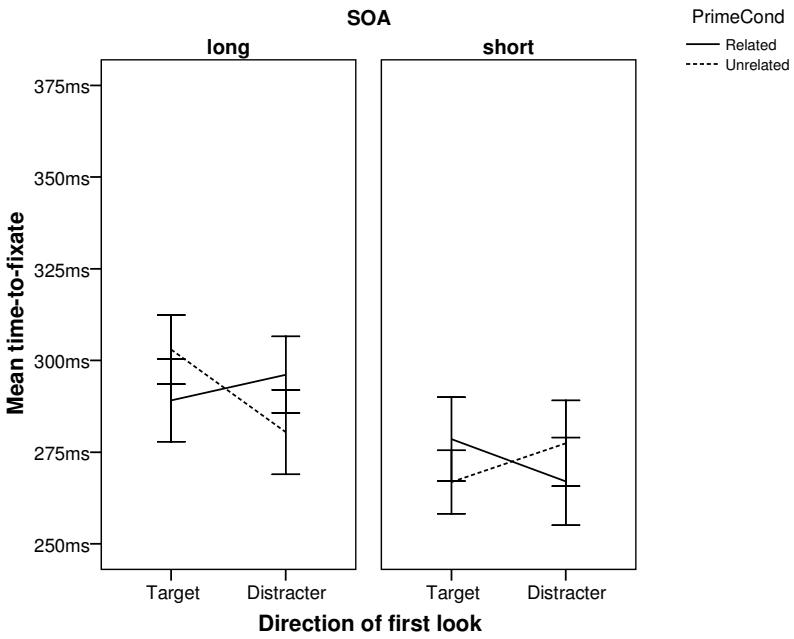
Figure 5.19. Histogram of time-to-fixate over the course of the trial (24m)



Notes. Analysis window marked from 80ms to 440ms. Normal curves plotted over whole data set (prior to exclusions).

Figure 5.20 depicts the mean time-to-fixate according to SOA group and priming condition. In this plot, a moderate difference is apparent between the SOA group, with infants in the short SOA group making slightly faster first fixations than infants in the long SOA group. No systematic difference is apparent between the priming conditions, which overlap substantially in both SOA groups.

Figure 5.20. Time-to-fixate according to SOA group and prime condition (24m)



Notes. +/- one standard error.

A four-way ANOVA was conducted on time-to-fixate, to investigate the influence of look direction (target, distracter), SOA group (long, short), prime condition (related, unrelated) and stimulus sub-list (*horse*, *dog*). The results of the ANOVA are given in Table 5.14, where a significant main effect of SOA group was apparent. Simple effects clarified that 24-month-olds in the short SOA group made faster fixations (M = 272ms, SD = 76ms) than 24-month-olds in the long SOA group (M = 293, SD = 75), and this difference was significant ($t(398) = 2.74$, $p = .006$, $d = 0.28$).

Table 5.14. Four-way ANOVA of time-to-fixate distracters (24m)

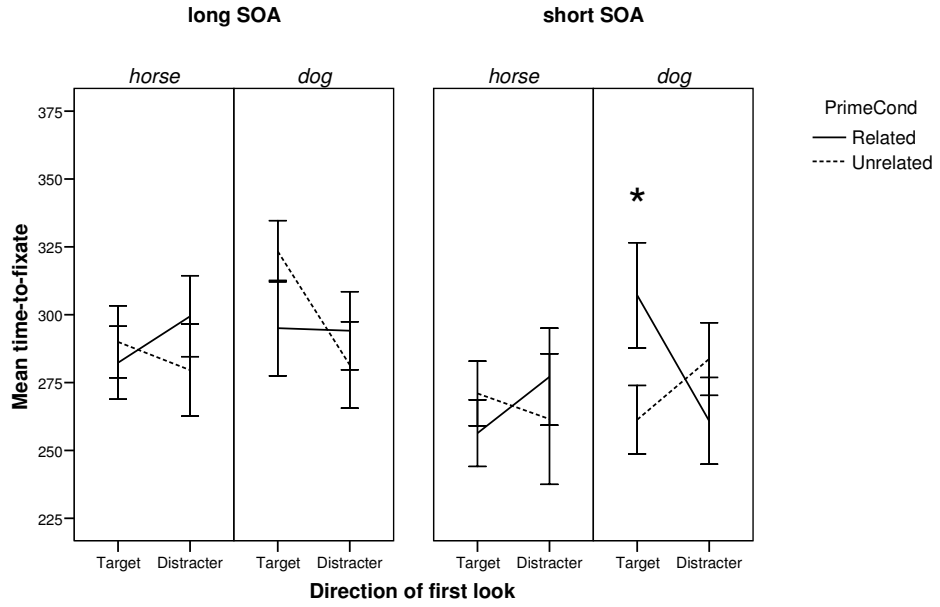
Source	*df*	*F*	*p*	partial η^2
direction (dir)	1	0.59	.442	.002
SOA	1	6.91**	.009	.018
prime	1	0.11	.745	<.001
list	1	2.01	.157	.005
dir * SOA	1	0.14	.708	<.001
dir * prime	1	0.23	.633	.001
SOA * prime	1	0.19	.664	<.001
dir * SOA * prime	1	2.86~	.092	.007
dir * list	1	1.80	.181	.005
SOA * list	1	0.01	.944	<.001
dir * SOA * list	1	0.05	.824	<.001
prime * list	1	0.01	.928	<.001
dir * prime * list	1	1.83	.177	.005
SOA * prime * list	1	0.63	.429	.002
dir * SOA * prime * list	1	3.15~	.077	.008
S w/in group error	384	(5682.15)		

Notes. Mean square error given in brackets. ~p<.10, **p<.01.

The ANOVA also revealed a near significant three-way interaction between direction of first fixation, SOA group and prime condition, and a near significant four-

way interaction between direction of first look, SOA group, prime condition and stimulus sub-list. To investigate the pattern of interactions Figure 5.21 depicts time-to-fixate according to all four factors. In this plot it is apparent that the only reliable difference between priming conditions occurs in the target direction for infants in the short SOA group, who were fixating items in the *dog* sub-list. It is interesting to note that the primed fixation delay occurs in the most 'difficult' condition: the fastest timing conditions and least interesting stimulus items. In the *horse* sub-list on the other hand, related targets are fixated slightly faster. Given the difficulty of interpreting these results, and the non-significant nature of the interactions, they may well be artefactual, and should only be treated as weak trends in the data.

Figure 5.21. Time-to-fixate according to stimulus sub-list and prime condition (18m)



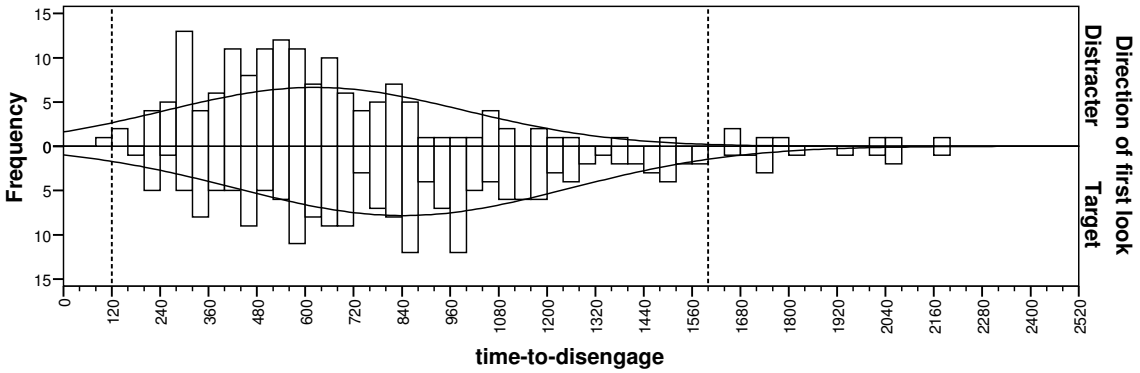
Notes. +/- one standard error

Time-to-disengage (24-month-olds)

The distribution of time-to-disengage is plotted in Figure 5.22, where only those trials where the speed of the *first* fixation was normally distributed are included

in analysis. The manually selected analysis window excluded trials with a time-to-disengage of less than 120ms, and trials with a time-to-disengage of 1600ms or more. As in the 18-month-old data, time-to-disengage was not normally distributed within the different subsets of the data. Due to differences in the variance between the two look directions, each direction is analysed in separate univariate ANOVAs, which are considered robust to deviations from normality. Non parametric tests are employed instead of simple comparisons.

Figure 5.22. Histogram of time-to-disengage over the course of the trial (24m)

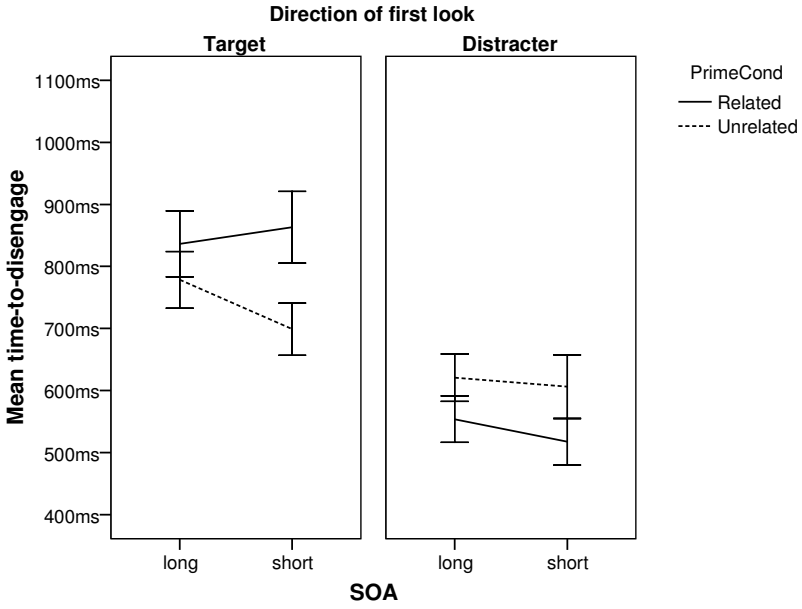


Notes. Manually selected analysis window: 120-1600ms. Normal curves plotted over all valid trials.

Figure 5.23 plots mean time-to-disengage according to look direction, with SOA group and priming condition also shown. In this plot, time-to-disengage is generally slower in the target direction. Little difference is observable between the SOA groups. The priming conditions appear to differ somewhat in both look directions, with faster distracter rejections and slower target disengagement in the related prime condition. Time-to-disengage from targets (M = 784ms, SD = 341ms) was slower than time to disengage from distracters (M = 570ms, SD = 259ms), and the difference between look directions was significant ($z(192, 147) = 5.96$, $p < .001$). This indicates that infants rejected a picture faster when the picture mismatched the

unfolding label. This finding is consistent with the pattern observed in younger infants and in the previous experiment.

Figure 5.23 Time-to-disengage according to SOA group and prime condition (24m)



Notes. Whiskers show +/- one standard error.

Time-to-disengage from targets was subjected to a three-way ANOVA comparing the influence of SOA group (long, short), priming condition (related, unrelated) and stimulus sub-list (*horse*, *dog*). The results of the ANOVA are given in Table 5.15, where there is a significant main effect of prime condition. The main effect of sub-list did not achieve significance, with a *p*-value of .14. Time-to-disengage from targets was slower in the related prime condition (M = 849ms, SD = 354ms), than it was in the unrelated prime condition (M = 735ms, SD = 325ms). A Wilcoxon's test confirmed that the difference between the rank of the scores was significant ($z(83, 109) = 2.07$, $p = .04$).

Table 5.15. Three-way ANOVA of time-to-disengage from targets (24m)

Source	*df*	*F*	*p*	partial η^2
SOA	1	0.28	.599	.002
prime	1	5.13*	.025	.027
sub-list	1	2.19	.141	.012
SOA * prime	1	0.95	.332	.005
SOA * sub-list	1	0.35	.556	.002
prime * sub-list	1	0.28	.596	.002
SOA * prime * sub-list	1	0.98	.324	.005
Error	184	(114765.9)		

Notes. Mean square error in brackets. *p<.05

Time-to-disengage from distracters was subjected to a three-level ANOVA comparing the influence of SOA group (long, short), prime condition (related, unrelated) and stimulus sub-list (*horse*, *dog*). In this comparison, a main effect of prime approached significance, with a *p*-value of .08 close to the alpha-level of .05. The main effect of sub-list did not achieve significance, with a *p*-value of .16. In this look direction there was *faster* rejection of distracters in the related prime condition (M = 532ms, SD = 228ms), than in the unrelated condition (M = 606ms, SD = 282ms). This difference did not achieve significance ($z(72, 75) = 1.52$, n.s.), with a *p*-value of .13.

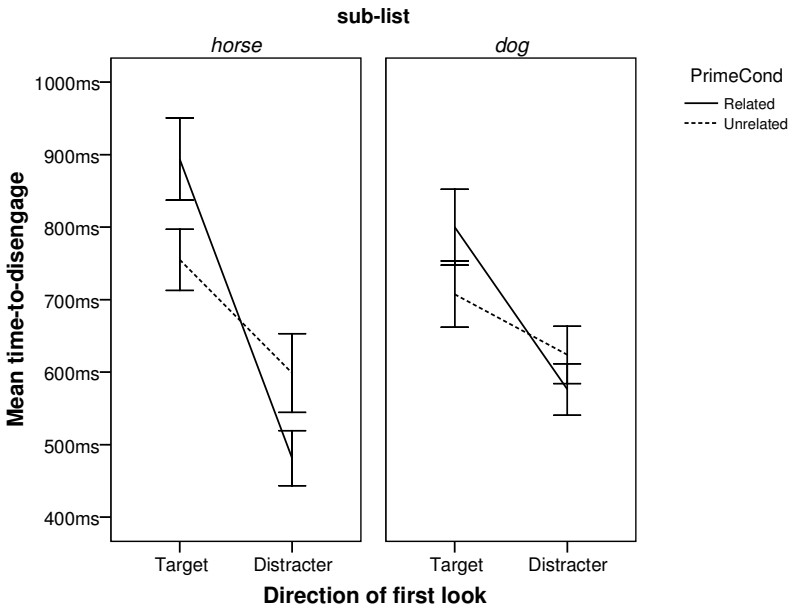
Table 5.16. Three-way ANOVA of time-to-disengage from distracters (24m)

Source	*df*	*F*	*p*	partial η^2
SOA	1	0.51	.476	.004
prime	1	3.17~	.077	.023
sub-list	1	1.99	.161	.014
SOA * prime	1	<0.01	.962	<.001
SOA * sub-list	1	0.31	.582	.002
prime * sub-list	1	0.54	.466	.004
SOA * prime * sub-list	1	1.00	.319	.007
Error	137	(64716.8)		

Notes. Mean square error in brackets. ~p<.10

Given that both look directions showed a moderate trend toward an effect of sub-list in the macro measures, sub-list is considered along with the other independent variables. Figure 5.24 shows time-to-disengage according to sub-list, where an interaction is apparent between prime condition, look direction and sub-list: The overall pattern of data is the same in both sub-lists, but the effects of the look direction (faster rejection of distracters) and priming condition (larger difference between values for target and distracter) are moderately more pronounced for items in the *horse* sub-list.

Figure 5.24. Time-to-disengage according to sub-list and prime condition (24m)



Notes. Whiskers show +/- one standard error.

To further clarify this interaction, Table 5.17 gives the mean time-to-disengage according to SOA group and sub-list. Table 5.17 also gives a difference measure (Δ-direction) which illustrates the magnitude of the difference in time to disengage from targets, as compared to distracters. In this table, Δ-direction is larger for the related prime condition, where the effect of fixation direction is highly significant. Within each priming condition, Δ-direction is larger for items in the *horse* sub-list than for items in the *dog* sub-list – a pattern which results in no evidence for target recognition for targets in the *dog* sub-list when they were preceded by an unrelated prime. This interaction follows the pattern observed for infants at 18 months-of-age, who showed target discrimination only for items in the *horse* sub-list, and in this sub-list the effect of target discrimination was larger for items in the related prime condition.

Table 5.17. Mean time-to-disengage according to SOA group and sub-list (24m)

		Direction of first fixation					
		Target	Distracter				Δ direction
sub-list		*M (SD)*	*M (SD)*	*N*	*U*	*p*	*M*
horse	Related	894 (373)	469 (211)	73	203.0**	<.001	425
	Unrelated	755 (337)	599 (307)	96	739.0*	.027	156
dog	Related	800 (331)	576 (230)	82	489.0**	.001	224
	Unrelated	708 (307)	612 (265)	88	793.5	.146	96

Notes. Wilcoxon's tests compare look direction. $^{**}p < .01$, $^{*}p < .05$

In summary, 24-month-olds were significantly slower to disengage from targets than they were to disengage from distracters, an effect which was larger in the related priming condition than in the unrelated priming condition, and which was moderately influenced by stimulus sub-list.

Summary: 24-month-olds

In the micro-level measures of PTL and LLK, 24-month-olds showed target preference which was significantly above chance. Their performance was strongly affected by the stimulus sub-list; items in the *horse* sub-list attracted greater target preference than items in the *dog* sub-list. The effect of list did not interfere with a strong, general effect of prime condition, in which targets preceded by related primes attracted greater target preference than targets preceded by unrelated primes. No effects of SOA group were observed.

When the sequence of first fixations were considered, 24-month-olds reliably made their first fixation to the target more often than to the distracter – an effect which suggests target discrimination prior to planning their first eye movement, but which may have other interpretations. The speed of the first fixation was primarily influenced by the SOA group – with faster first fixations from infants in the short

SOA group. Note that this effect is in the opposite direction to the effect observed for 18-month-olds, who were slower to fixate targets in the short SOA condition. One possible interpretation for this difference is that older infants have more advanced linguistic processing skills, and rather than being delayed by the early onset of the pictures, older infants are able to respond to the faster pace of stimulus presentation, with moderately quicker first fixations.

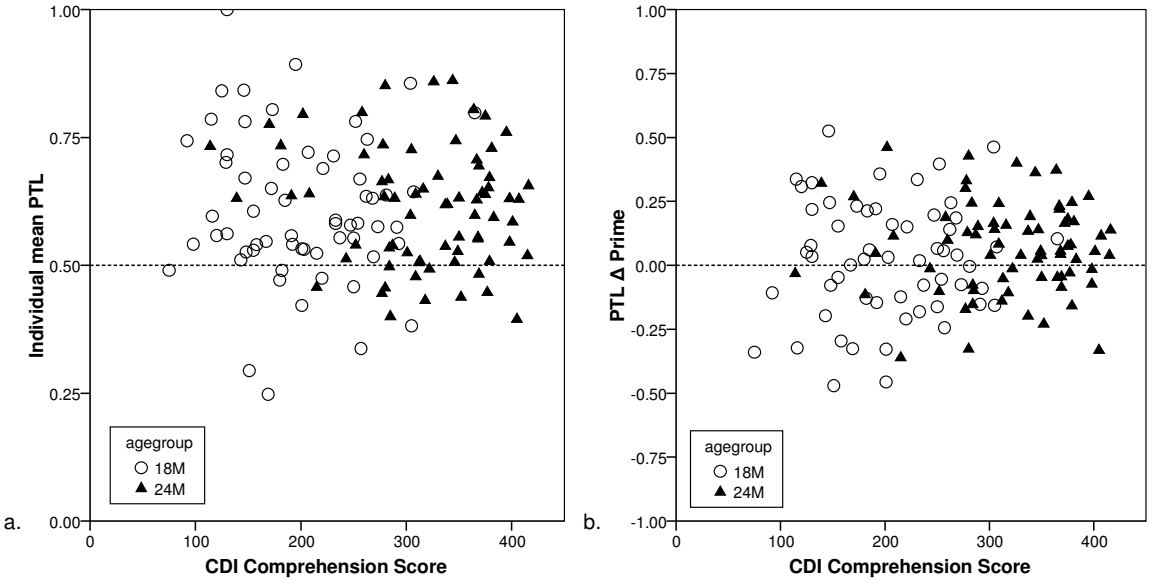
In micro-level measures, prime and SOA did not significantly interact with time-to-fixate, although there was a trend toward a sub-list interaction, suggesting that the influence of a-priori visual salience is active during the earliest stages of the referent identification task. Finally, 24-month-olds were slower to disengage from named targets than from unnamed distracters, and this effect was larger in the related priming condition than in the unrelated condition, and larger for the items which attracted greater visual interest. Taken together, these findings demonstrate that targets preceded by related primes attract greater visual interest than targets preceded by unrelated primes, and that this difference in interest is reflected in the duration of the very first fixation. The faster speed of distracter rejection indicates the influence of the priming condition on the high-speed 'matching' skill infants demonstrated in the previous experiment. This skill was larger and more reliable for one group of items than for the other.

Age and vocabulary development

Mean PTL was calculated for each infant, to facilitate comparison of individuals across the language development scale. PTL is plotted against CDI comprehension score for infants in both age groups in Figure 5.25a. In this plot it is clear that there is substantial overlap between the two age groups in terms of their

total CDI score, and PTL scores do not appear to differ greatly between age groups The spread of the PTL scores is broad (ranging from .25 to 1.0), with majority of infants showing above-chance mean PTL. As infants with smaller vocabularies are likely to lose more trials to the lexical exclusion criteria, it is unsurprising that infants at the lower end of the vocabulary scale show greater variance in their mean PTL. No linear trend is evident between total vocabulary size and individual mean PTL.

Figure 5.25. PTL and Δ-Prime for PTL according to age and lexicon size



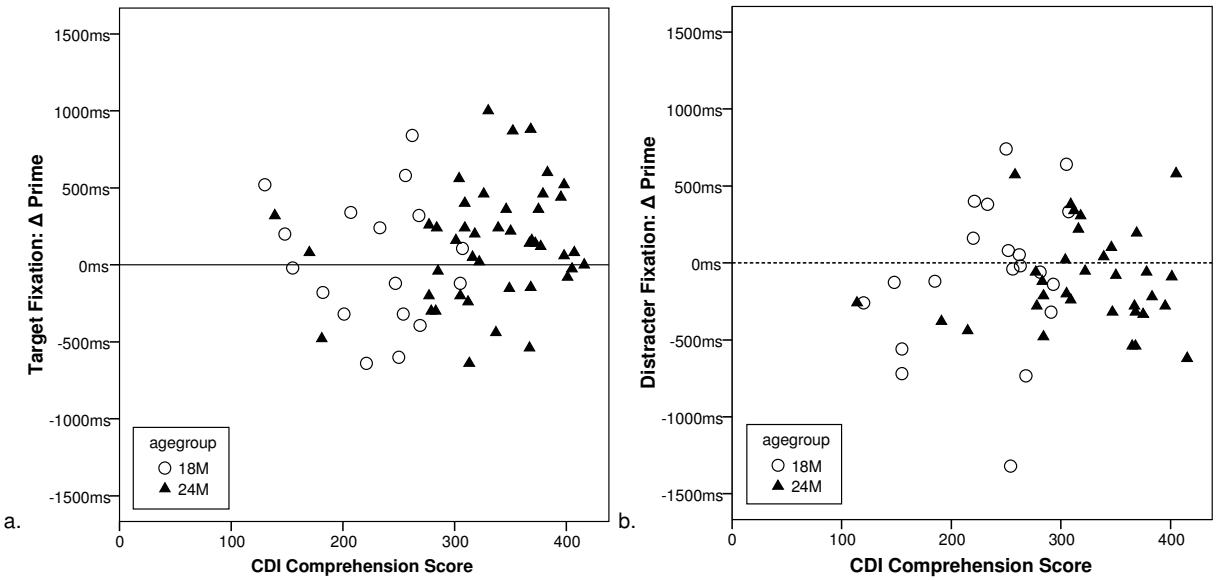
Notes. Each point represents an individual participant. Chance value in each measure marked by horizontal dashed line.

In Figure 5.25b, Δ-prime was calculated as the difference between the priming conditions for each individual, and this plot also shows performance across the language development scale. Δ-prime is quite dispersed, with wider dispersal for infants with smaller vocabularies. Again no difference is apparent between the behaviour of infants in the two age groups, and no linear trend is evident.

The mean time-to-disengage for each participant was also calculated, according to priming condition. Δ-prime was calculated as the difference between the related prime condition and the unrelated prime condition, and was generated

separately for each look direction, where data from both priming conditions were available for an individual. Figure 5.26 plots Δ-prime for each fixation direction, against individual infants' CDI score. In the target direction (Figure 5.26a) the majority of values are above chance, indicating slower disengagement from targets in the related prime condition. In the distracter direction (Figure 5.26b) the majority of values fall below chance, indicating faster disengagement from distracters in the related prime condition. In both plots there is substantial overall in the range of values generated by each age group, and no linear trend is evident over the CDI scale.

Figure 5.26. Δ-Prime for time-to-disengage, split by fixation direction



Notes. Each point represents an individual participant. Chance value in each measure marked by horizontal dashed line.

Taken together, the findings from both PTL and time-to-disengage across age and vocabulary development suggest that there is little difference between the performance of infants in the two age groups, other than a general reduction in variance. The reduction in variance is due to the larger vocabularies of 24-month-olds providing more trials in which both the target and the distracter were name-known. Recall that 18-month-olds tended to show a moderate priming effect (particularly in

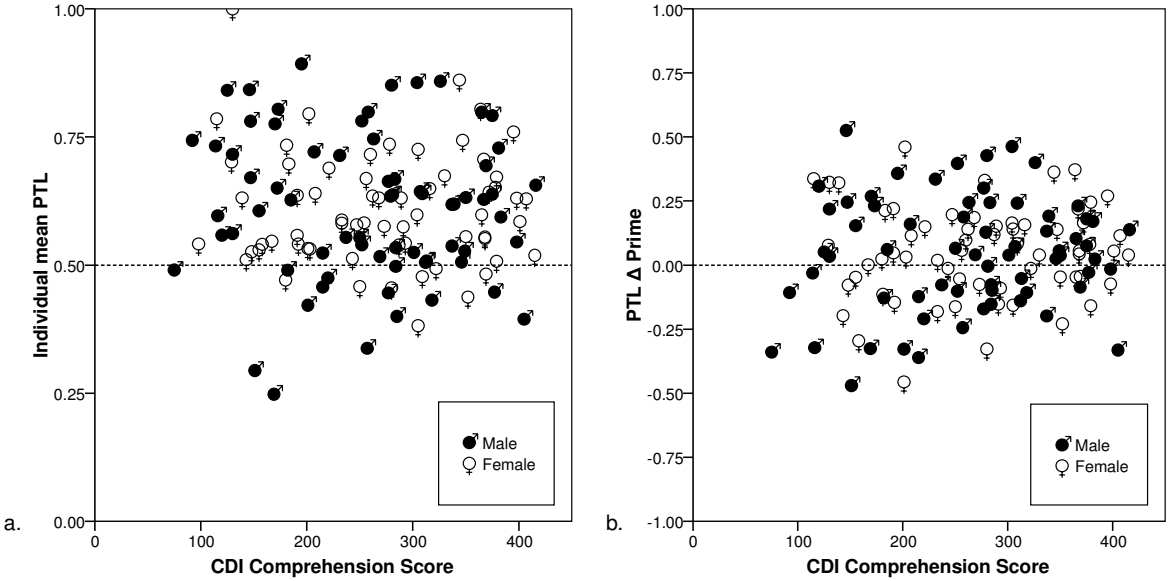
the stimulus sub-list with the more interesting pictures), but the effect was not significant when their data was analysed independently. In addition, these findings suggest that there is no monotonic increase in the magnitude of the priming effect with an increase in vocabulary development. This finding is especially interesting, as it suggests that the priming effect is not dependent on some 'critical mass' of vocabulary items, but may be present from quite early stages in lexicon development.

Comparison of sexes

In this section, the distributional data are reconsidered according to the sex of the participant, to interrogate possible sex effects which may be evident at different stages in lexicon development. Individual mean PTL and Δ-prime for PTL are plotted against CDI score with the sexes marked separately. In these plots it is clear that male and female infants overlap substantially across the CDI scale and across each of the measure. There is also no apparent linear trend according to the sex of participants.

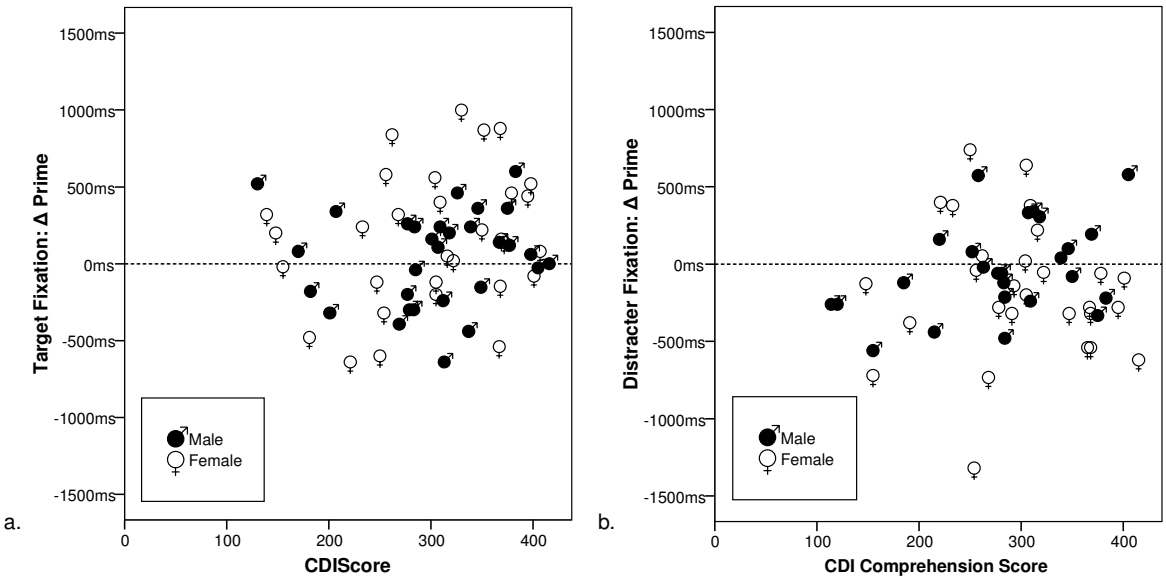
Figure 5.27 depicts Δ-prime for time-to-disengage in each of the look directions, with the sexes marked separately. In this plot too, values for males and females are fairly evenly distributed throughout the data set, with slightly more extreme values for females. While this may be taken to indicate that female infants show stronger priming effects in RT measure, it should be noted that female infants generated both the highest and the lowest values in each measure. The extreme values are likely to be generated by chance.

Figure 5.27. PTL and Δ-Prime for PTL in relation to lexicon size and sex



Notes. Each point represents an individual participant. Chance value in each measure marked by horizontal dashed line.

Figure 5.28. Δ-Prime for time-to-disengage, compared to lexicon size and sex



Notes. Each point represents an individual participant. Chance value in each measure marked by horizontal dashed line.

Trial time course

To investigate changes in the pattern of looking over the course of the whole trial, time courses for each age group were plotted. Figure 5.29 depicts the time course of looking for 18-month-olds. In Figure 5.29a, the number of trials in which infants were fixating the target is plotted against the number of trials in which infants were fixating the distracter. Over the first 200ms, the number of fixations is very low. This represents the portion of the trial in which pictures have only just appeared on the screen, and infants have yet to make their first fixation. Over the following 200ms, there is approximately equal interest in target and distracter pictures, suggesting that fixations landed prior to 400ms have not incorporated peripheral visual information about the identity of the target picture. From 400ms to the end of the trial targets consistently attract more fixations than distracters. The naming 'bloom' occurs from about 400ms to about 1600ms. There is a moderate decrease in target looking in the second half of the trial, which gradually recovers towards the end of the trial.

In Figure 5.29b, the proportion of fixations directed to the target is plotted separately for each priming condition. In this plot, in the first 200m there is a sharp peak in the proportion of target looking in the unrelated condition, and an early low peak in the related condition. Given that screen fixations are only present in eleven trials at this stage, these early peaks should be considered as artefacts. Following this early stage, interest in targets and distracters is very similar for the majority of the naming 'bloom'. In the unrelated prime condition, there is an increase in target fixations towards the end of the trial, a pattern which contributes to the late rise in target interest observed in Figure 5.29a.

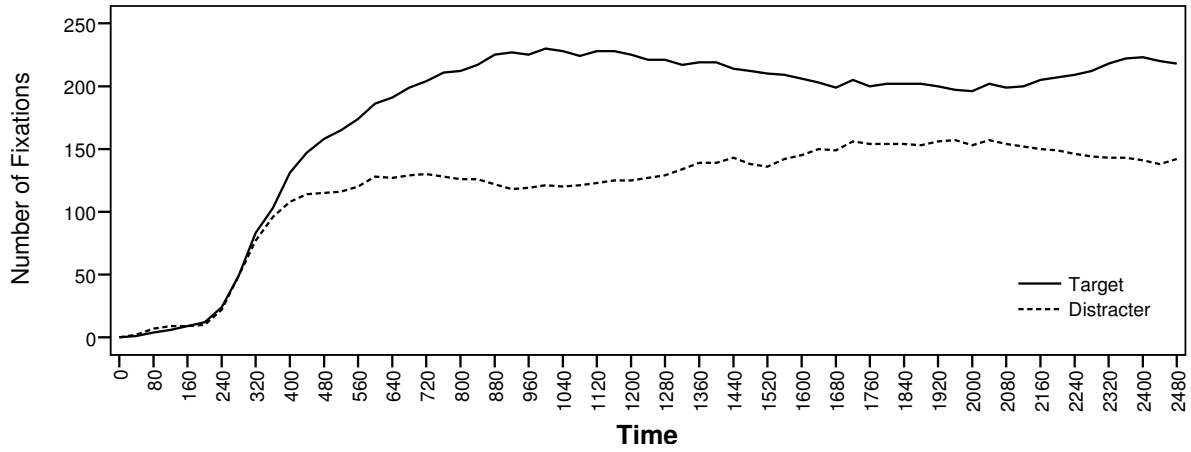
In figure 5.29c, the total number of trials in which infants are fixating either picture is plotted as a proportion of all valid trials, in each priming condition. This

plot summarises general screen interest, and can show differences in the onset of first fixations and differences in continued interest over the course of the trial. In this plot, screen looking increases slightly more quickly in unrelated prime trials than it does in related prime trials. After the first 1000ms of the trial, screen interest is approximately equal in both conditions, and remains almost at ceiling for the remainder of the trial. The temporary early gap in screen interest between the related and unrelated conditions is interesting as it indicates longer delays in launching fixations in related trials, combined with blinks and switches. This pattern was not clear in the analysis of time-to-fixate, but the current analysis also includes infants who were fixating the screen at picture onset, and infants whose first fixation speed was not normally distributed.

In Figure 5.30, time courses for 24-month-olds are depicted for comparison. In figure 5.30a, the main increase in fixations to both targets and distracters begins from about 200ms. The increase in target fixations is slightly steeper than the increase in distracter fixations, and a clear naming 'bloom' shows more fixations to targets for the majority of the trial. There is a gradual decrease in target preference, leading to almost equal picture looking for the last 500ms of the trial. 18-month-olds and 24-month-olds show very similar patterns of target preference over the course of the trial.

Figure 5.29. Detailed trial time course (18m)

a. Target Preference



b. Target looking in different prime conditions



c. Screen interest in different prime conditions

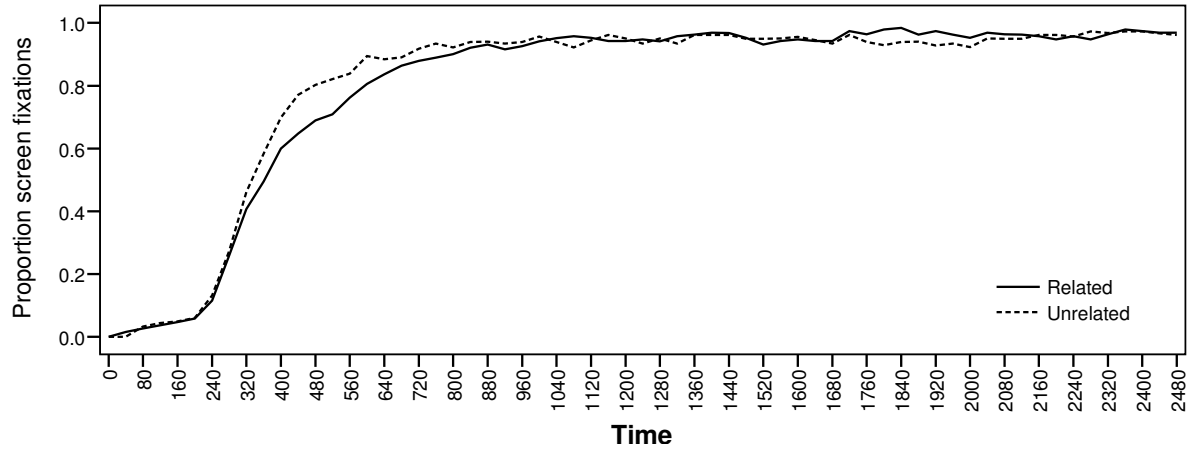
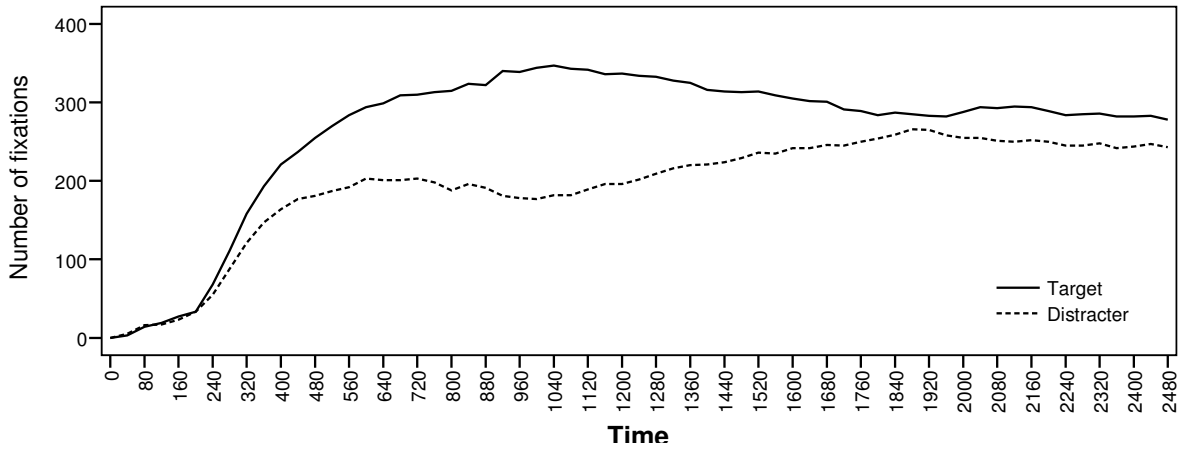


Figure 5.30. Detailed trial time course (24m)

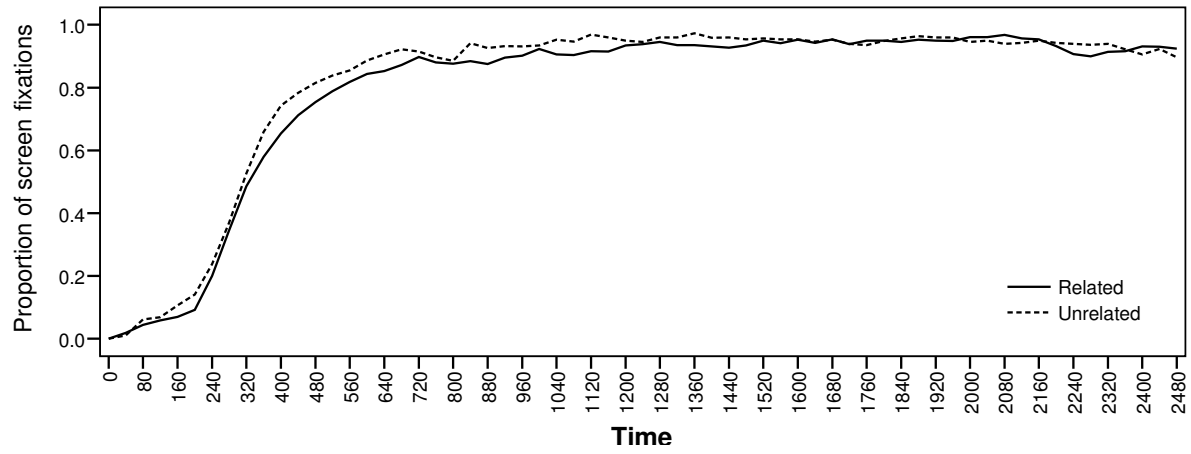
a. Target Preference



b. Target looking in different prime conditions



c. Screen interest in different prime conditions



According to prime condition (Figure 5.30b), the advantage for targets in the related condition over targets in the unrelated condition is clear and long lasting, from 500ms to the end of the trial. This differs from the pattern observed for 18-month-olds, where there was very little difference between the priming conditions until late in the trial. This finding is consistent with the general pattern of results in which priming was unreliable for 18-month-olds, and only evident for some items, while priming for 24-month-olds was clear and reliable in three of the four measures of target preference (PTL, LLK and time-to-disengage).

Finally, Figure 5.30c depicts screen interest in the different priming conditions, where the same pattern was observed for both age groups: There was a rapid increase in screen fixations from about 200ms to 800ms, during which time, there were more early fixations in the unrelated prime condition than in the related prime condition. For the remainder of the trial, both priming conditions continued almost at ceiling.

Again, the influence of the prime in this early period was not evident when time-to-fixate was considered. Note that this effect may well have been complicated by differences between test items, which are discussed in detail in the following section.

Item effects

Throughout the analyses conducted in this chapter, clear item effects were evident between the two stimulus sub-lists. In the *horse* sub-list, the target-distracter pairs were *horse-sock, mouse-table, cake-trousers, bus-monkey, car-bowl,* and *coat-bear.* The targets in this sub-list consistently attracted more visual interest, as measured overtly by PTL and LLK, and implicitly via the interaction of sub-list with

RT variables. Conversely, items in the *dog* sub-list attracted less looking overall, and, in some conditions, did not attract behaviour consistent with target discrimination. In the *dog* sub-list, the target distracter pairs were *dog-boat, cow-toast, banana-lion, shoe-cheese, cup-pushchair,* and *bed-chicken*. Given the reliability of the sub-list effect, it is worth investigating this effect in more detail.

One possibility is that target words in the different sub-lists have different levels of familiarity for infants. If this were the case, infants might be expected to show more reliable target discrimination for words which are more familiar. To investigate this possibility, lexical comprehension rate can be used as a proxy for population-level familiarity. According to the norming study carried out in Chapter 2, target words in the *horse* sub-list had a mean comprehension rate of 72% (*SD* = 16%), while target words in the *dog* sub-list had a mean comprehension rate of 89% (*SD* = 8%). The difference in ranked comprehension rates was significant ($U(12) = 4.0$, $p = .025$). The effect was in the *unexpected direction*, however, with higher comprehension rates for the words that attracted the least target looking. This might suggest that the visual interest in the *horse* sub-list was a type of novelty preference. However, it does not explain why the more familiar words were those which attracted less target discrimination – sometimes failing to exceed chance. Furthermore, the results of the minimally primed experiment (Chapter 3) suggest that for target discrimination, individual target status (known, unknown) would be more important than population-level assessments of familiarity. Given the difficulty of interpreting this finding, further sources of variation between the sub-lists are also investigated.

It was also possible that the familiarity of the distracter pictures could be involved in an effect of sub-list. The mean lexical comprehension of words in the

horse sub-list was 64% (*SD* = 14%), and the in *dog* sub-list, 67% (*SD* = 6%). No difference was evident ($U(12) = 9$, n.s.).

An additional source of variability came to light after considering the verbs used in prime 'carrier phrases'. Verbs were selected if they were acquired early, and were semantically plausible in the sentential context. Ten of the twelve carrier phrases contained the verb *saw*, as in 'Yesterday I saw a cat,' or the verb *bought*, as in 'Yesterday I bought a plate'. The verbs *bought* and *saw* were semantically unrestricted, given the nouns used in the current experiment. The carrier phrases for the primes *apple* and *biscuit*, on the other hand, contained the verb *ate*. Unlike *saw* and *bought*, *ate* is a semantically restrictive verb, which only takes food as its direct object. This effectively means that *apple* and *biscuit* trials contain more than one biasing relationship between the prime phrase and the target. Recent research has demonstrated that infants at the end of their second year can use the semantics of verbs to guide predictive looks to restricted object categories (Fernald, 2005).

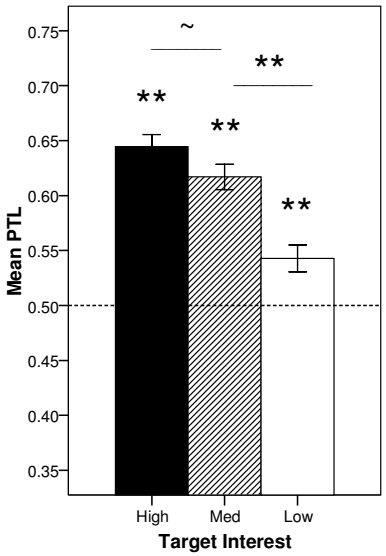
However, the two restrictive verbs were balanced across the sub-lists: In *horse* sub-list *ate a biscuit* congruously preceded *cake* in the related condition, and incongruously preceded *coat* in the unrelated condition; in the *dog* sub-list, *ate an apple* congruously preceded *banana* in the related condition, and incongrouously preceded *bed* in the unrelated condition. Given this tight balance, the restrictive verbs are unlikely to be a source of the generally higher interest in targets from the *horse* sub-list, nor for the minor interactions of sub-list with priming condition.

It was also possible that the interest ranking of the items differed across the sub-lists, and if the interest-level ranking produced a consistent effect on the looking pattern, it could account for the item effect. In the preparation of stimuli for the current experiment, targets were allocated to three approximate interest ranks: high

interest for animals, medium interest for vehicles and food, and low interest for inanimate objects, clothing, furniture and tableware. The interest level of the target was found to influence the magnitude of target preference. For example, Figure 5.31 depicts PTL pooled across age group, SOA and priming condition. In this plot PTL for targets ranked as 'high' interest was higher than for targets ranked as 'medium' interest, which was in turn higher than for targets ranked as 'low' interest. Target preference for items from all three interest levels was above the .50 level predicted by chance.

A difference was observed between the interest levels of items in the two sub-lists. The *horse* sub-list, which attracted more target looking overall, had two high interest targets, three medium interest targets and a single low interest target. The *dog* sub-list had two high interest targets, one medium interest target and three low interest targets. This means that targets in the *horse* sub-list were indeed more interesting on average than targets in the *dog* sub-list.

Figure 5.31. PTL for target pictures in three interest levels



Notes. Proportion of target looking (PTL). One sample t-tests comparing PTL to chance value of .50 are marked above each bar. Independent samples t-tests comparing interest levels are marked above pair of bars. Whiskers show +/- one standard error.

In addition, when distracter interest was also considered, the six targets in the *dog* sub-list were paired with two high-interest distracters and four low-interest distracters, while targets in the *horse* sub-list were paired with two high-interest distracters and four medium-interest distracters. This means that as well as being slightly less interesting, targets in the *dog* sub-list had distracters which were, on average, *more distracting*. It is clear that this asymmetry in the stimulus design is the most likely source of the stimulus effects observed throughout the analysis.

However, the primary concern in a priming study is not whether target discrimination is evident, nor the magnitude of target discrimination in different sub-sets of the items. The primary concern is whether measures *typically used to determine target discrimination* differ when the same item is encountered in two different priming conditions. Despite the strong influence of the sub-lists, differences between priming conditions were also observed for the older infants, where the effect of prime condition was highly significant.

General discussion

The primed IPL task was designed to incorporate the stimulus order of the words-before pictures task into a fast-paced, primed IPL scenario, in which the organisation of the stimuli mimicked adult sequential priming tasks. The task replaces the 'lexical decision' task of adult priming studies with an infant-friendly lexical comprehension task. Picture pairs were displayed following the auditory presentation of two words, whose relationship and timing could be manipulated. Given the high speed of stimulus presentation, and the potentially confusing priming phrase, it is noteworthy that infants at both ages demonstrated reliable target identification in macro- and micro-level measurements. While there was some variation in the degree of comprehension demonstrated across items, the general finding of target discrimination indicates that the current task effectively taps into the same comprehension skill as demonstrated in previous experiments.

In addition, differences between the two priming conditions were observed in the older age group, in both micro- and macro- level measures of looking behaviour. 24-month-olds fixated targets faster and longer, and showed more overall target interest when the target was preceded by a related word. This constitutes a clear 'priming effect', indexing the effect of the verbal context on looking behaviour. These priming effects were robust to the general effect of stimulus interest.

While the younger infants did not reliably demonstrate overall priming effects, tendencies toward priming were evident for some items, and their overall pattern of looking behaviour was very similar to the performance of older infants, with the exception that their performance was more variable. This finding demonstrates that

the current design was sufficiently sensitive to capture differences in picture interest which were brought about not by the act of labelling, but by varying the penultimate word prior to picture presentation. The priming condition not only influenced overall interest in the named target, but also the ease and speed of target identification (and distracter rejection).

Fast pace

The fast pace of the design included a short prime-to-target inter-stimulus interval (ISI) of 200ms, short SOAs of 200ms and 400ms, and a relatively short period of free looking time (2,500ms). The SOA of 400ms had been piloted in the words-before-pictures experiment (Chapter 4). The picture presentation time in the current experiment fell between the two picture durations previously piloted (1500ms, 5000ms). The fast pace of stimulus presentation and the relatively short picture presentation period were selected to isolate the early stages of linguistic processing, and hopefully to reduce any 'strategic' priming effects. Although the results of the minimally primed IPL experiment (Chapter 3) suggested that priming phrases prior to picture onset do not disrupt infants' general ability to identify a named target, it was unclear whether the very fast timing employed in the current design would generate a different pattern of results. From this perspective, it is reassuring to see that even when the prime precedes the target at such a short ISI, target recognition was clear (although for younger infants it was only robust for some of the items).

The stimulus onset asynchronies of 200ms and 400ms were selected to fall either side of the previously established auditory decision point (300ms), at which infants are able to identify a named picture regardless of whether the rest of the string was presented (Fernald et al., 2001) – or in this case, potentially disrupted due to the

appearance of the pictures. Very little difference was found between the SOA conditions, in terms of target recognition or priming. In the macro-level measures, no difference was apparent between SOA conditions, suggesting that the sudden onset of pictures did not differentially disrupt overall interest in named targets, nor did it change the degree of priming observed. In the reaction time measures, younger infants showed some sensitivity to SOA in the duration of their first fixation, suggesting that processing was more efficient (i.e., distracter rejection was faster) when they had heard more of the target word prior to picture onset. For older infants, the *speed* of the first fixation was influenced by the pace of the task, with faster fixations in the short SOA condition. There was also a weak tendency for the shorter SOA condition to generate slightly more pronounced priming effects in some subsets of the data. This could be due to enhanced automaticity in the task, but given the weakness of the trend, it is possibly artefactual.

Stimulus controls

Two stimulus controls were considered critical for the quick-primed IPL experiment. First, in order to avoid inadvertent memory effects which might interfere with priming, no stimuli were repeated within a testing session. This approach is advocated by McNamara (2005) in situations where testing periods are relatively short. Secondly, and perhaps most importantly, only those trials in which infants were reported to understand both the prime and the target were included in the analysis. The minimally primed IPL study (Chapter 2) established that parental report was a reliable index of which words 18-month-olds would be able to comprehend in a relatively difficult task. Although the picture pairs here presented a lower level of difficulty (as all items were drawn from different categories, and were thus less likely

to be confusable), the exclusion of trials containing words reported as unknown was still considered to be important. From a theoretical perspective, it makes little sense to investigate possible 'links' between words whose status in the lexicon is unknown. Furthermore, at the experimental level, the inclusion of unknown words generates a pattern of noise which is likely to mask genuine priming effects: If a prime or a target is not understood in a 'related' trial, that trial effectively becomes 'unrelated,' as only one word (the target or the prime) shares a relationship with the named picture. Yet if a prime in an 'unrelated' trial is not understood, the trial remains 'unrelated' (as only the target word shares a relationship with the target picture). This skew would reduce the difference between the two trial types, potentially masking legitimate priming effects. Eliminating this confound is particularly relevant for infants with small vocabularies, for whom this noise would be greatest. Pre-testing of items (e.g., in a standard IPL task) could have provided additional data about each infant's comprehension, but would have confounded the stipulation of non-repetition. Instead, lexical comprehension of test items was assessed via parental report.

Lexical Organisation and Processing. Prime-target pairs were selected which shared both semantic and associative relationships, in order to maximised the likelihood of an associative 'boost' (Moss et al., 1995). It is interesting to note that where adult word associations are present between common infant words, the word-pairs tend also to be category co-ordinates. Given the small size of the infant lexicon in the second year, co-occurrence of word association and taxonomic relationship may be a salient property of the developing lexicon. Further investigations would be needed to establish which, of taxonomy or association, is the greater source of organisation in infants' knowledge structures.

What causes the infant priming effect?

Although the current experiment has been framed as an infant-friendly adaptation of a standard adult priming method, it is unclear whether the current task actually taps into the same level of linguistic processing as the priming effects observed in adult studies. The complication arises out of the combination of auditory word presentation, with pictures in the test phrase. In this scenario, infants respond to the pictures, *not to the target word itself*. This means that any effect of priming on infants lexical processing can only be assessed via the visual task. The complication is that there may be semantic or visual processing occurring in parallel to the lexical processing. And the observed priming effect may derive from either source.

Thus the primed IPL effect could be produced by 'lexical facilitation' or by 'visual mediation'. In lexical facilitation, the prime enhances activation of the target word via spreading activation in the lexicon. This 'word-to-word' priming would cause greater interest in the target picture due to enhanced lexical activation of the target word. This kind of priming (which is variously referred to as 'semantic priming', or 'lexical priming'), is the kind typically identified in the adult priming literature, where processing is assessed by response times to actual test words. It is this kind of prime which provides the most direct evidence of lexical connectivity or organization in the lexicon.

In visual mediation, on the other hand, the prime may trigger a conceptual representation of the prime concept. That concept may share visual features with the target (such as 'four legs' for *cat* and *dog*, or 'smooth surface' for *plate* and *cup*). This 'word-to-picture' priming would thus enhance visual interest in the target picture without affecting lexical processing of the prime. If the priming effect observed here is due to word-to-picture priming, it would provide evidence about semantic feature

mapping across different known objects, but it would not provide information about the impact of hearing a prime word on processing a target word. And therefore, does not provide information about the word-level organization of the lexicon.

One reason for preferring a lexical account for the current findings is that the lexicon may have limited resources available for activation. For example, hearing *plate* might spread activity to related areas of the lexicon (including *cup*, *bowl*, *food*, etc.), resulting in reducing the resting activation level of unrelated areas (e.g., *dog*, *boat*, *shoe*). This account would lead to delayed or inhibited comprehension of targets primed by unrelated words. Although inhibition was not clearly evident in the current experiment, target discrimination was less reliable in the unrelated priming condition, especially as measured by time-to-disengage (c.f., Figure 5.24). These findings may indicate some inhibition of normal target discrimination.

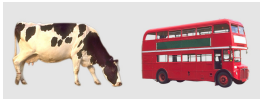
One interpretation of Styvers & Tenenbaum's (2005) lexical modelling implies that there would be no substantial change in the nature of lexical connectivity over the course of lexicon development. Hence, where a priming effect is observed in infancy, it is not unreasonable to propose that it derives from the same mechanisms as adult priming effects (i.e., spreading activation in the lexicon). However, the current data do not rule out accounts which operate at different levels of processing. As the observed effect may be due to lexical facilitation, visual mediation or to some combination of both, further studies are needed to clarify the precise mechanisms underlying the priming effect observed in the current experiment. The following chapter describes an experiment which attempts to tease apart these two possibilities.

Interim conclusions

Infants in their second year discriminated named target pictures from unnamed distracter pictures in a quick-primed IPL task which included sequential word presentation. For older infants, target discrimination was influenced by whether the penultimate word was also related to the named target picture. With clear, enhanced target recognition in three of four measures of target discrimination. It remains unclear whether the priming effects observed in the older age group are due to spreading activation at a lexical or at a conceptual level of representation. It also remains to be seen whether one type of relationship (associative or taxonomic) provides organisational structure in the infant lexicon, or whether an adult-like mix of structures is encoded. Nonetheless, these findings demonstrate that priming effects can be observed in the online language processing of 24-month-old infants.

Chapter 6

Half Target – Quick Prime



In the previous chapter, a quick-primed IPL experiment was conducted for infants at 18- and 24 months-of-age. In the quick-primed design, the organization of the stimuli allowed the prime word to be acoustically presented directly before the spoken target word. In order to isolate *automatic* priming effects in the early stages of lexical processing, stimuli were presented quickly, and the relationship of the prime to the target was systematically manipulated. The onset of the target word was followed by a referent identification task tapping into infants' lexical comprehension. Different priming conditions were shown to influence looking patterns for the older infants, with more and longer target looking, as well as faster distracter rejections, for those words which were preceded by a related prime.

The locus of the priming effect was, however, ambiguous, as it could be accounted for by two possible processing models: 'lexical facilitation' or 'visual mediation'. The current chapter reports an experiment for 18-, 24- and 30-month-olds which attempts to tease apart these two possible mechanisms by investigating the role of target word processing in primed looking behaviour. In this two-by-two design, half of all primes were related to the target word, and half of all targets were named.

Introduction

The motivation for the body of research described in this thesis is to investigate relationships in the developing lexicon by observing the effect of hearing one word (the prime) on the linguistic processing of another word (the target). In the adult lexicon, the spreading activation account of lexical processing (introduced in Chapter 1) proposes that changes in the ease and speed of lexical processing can be used to infer the psychological reality of relationships between words in the lexicon.

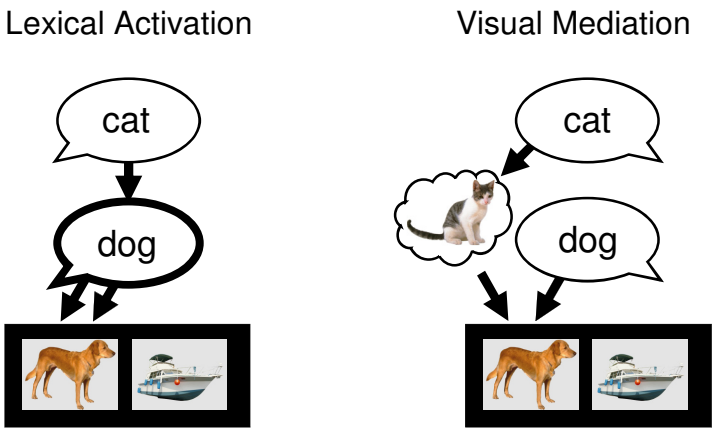
In the quick-primed experiment (Chapter 5), the cross-modal nature of the IPL task meant that while infants processed the acoustic target label, it was responses to *pictures* which were monitored. Because the test item was not the spoken word itself, the locus of the observed priming effect was ambiguous.

The previous chapter introduced two processing models which were able to account for the observed priming effect. 'Lexical facilitation' and 'visual mediation' are illustrated in Figure 6.01. In the lexical facilitation account, some activity flows from the prime to the target (e.g., *cat* to *dog*), prior to the presentation of the target word. When the target word is heard shortly after the prime, the summed activity (from spreading activation and from the incoming phonological information) results in activation of the word *dog* which is enhanced relative to the unrelated priming condition. The spreading activation model implies that the target word (e.g., *dog*) would be processing faster when preceded by a related prime, and may also result in greater interest in the referent of the target, due to the word's higher level of activation. Enhanced processing is considered as a difference between the two priming conditions: The spreading activation model is compatible with either facilitation of normal target processing (e.g., in the related condition), or with inhibition of normal target processing (e.g., in the unrelated condition).

According to the visual mediation account, on the other hand, the prime word activates a mental representation of the prime concept, which includes information about the prime's visual form. As all primes and targets in the previous experiment were taxonomic sisters, there was some overlap between the visual features of certain primes and targets. For example, *cat* and *dog* are visually similar, due to shared features such as faces, legs and tails (although note that some prime-target pairs such as *cheese* and *bread* are less visually similar). In the visual mediation account, when

the pictures appear, the visual similarity between the internal representation of the prime and the observable target picture enhances interest in the target picture. In this processing account, *cat* and *dog* need not be linked in the lexicon, and the effect of visual similarity (prime-to-picture) may be extra-linguistic.

Figure 6.01. Two possible priming mechanisms



One way of teasing apart these two mechanisms is to investigate the role of the target word in the producing the priming effect. In the lexical activation account, the prime exerts its influence on the activation level of the target word itself (word-word priming), while in the visual mediation account, the prime word and the target word independently generate interest in the target picture (word-picture priming). If target labelling is manipulated along with the priming relationship, then the effects of the prime and the target may be observed independently. For example, if the prime is related to the target, but the target remains unnamed, spreading activation would predict no difference in target word processing, while the visual mediation account would predict enhanced target interest. If both processes (spreading activation and visual mediation) act in parallel, then more complex interactions between naming and priming may emerge. The detailed time course of the trial is predicted to be

informative in this regard. This chapter presents a two-by-two design, in which the priming conditions of the previous chapter (related and unrelated) are counterbalanced across two target labelling conditions (named and unnamed). In this scenario, all infants hear an auditory prime on all trials, followed by a target label, or a neutral 'attention phrase' equivalent in duration (e.g., Look!).

In the previous chapter, significant priming effects were only found in the older age group. This finding suggested that priming effects might only be observable in children with large enough vocabularies that few trials would be lost to lexical exclusions. It is also possible that priming effects develop with age and/or linguistic development, and that *even larger* priming effects might be observable in older infants. For this reason, a third age group was added. The current experiment includes 18-month-olds, 24-month-olds and 30-month-olds.

Controlling for *a priori* preference

In previous chapters, three methods were employed to control for differences in the *a priori* interest of test stimuli: within-trial baselines, counterbalanced labelling, and counterbalancing relative interest levels. Each method has its advantages and disadvantages in IPL, particularly as they relate to the priming context. In the minimally-primed IPL experiment (Chapter 3), the 'gold-standard' of a within-trial baseline was used, to discretely quantify each individual's interest in the target at the time of test, *before* assessing change following the onset of a label. While this method provides the most stringent test of the influence of a label (over and above general interest), it is incompatible with fast-paced priming, as it requires visual stimuli to be presented during the inter-stimulus interval (ISI) between the offset of the prime and

the onset of the target. The dual complications of long ISI, and semantic interference from the pictures, meant that alternative solutions were required.

In the words-before-pictures experiment (Chapter 4), instead of a pre-naming baseline, all pictures acted equally often as named targets and unnamed distracters across the sample population. While this did not allow baselines for each individual, the approximate interest of targets and distracters was counterbalanced across the population, and the effect of labelling remained significant across the pooled items. However, as the purpose of lexical priming is to observe differences in behavioural responses to a *single item* under different linguistic contexts, a third method of controlling for item differences was sought.

In the quick-primed experiment (Chapter 5), where the same items were named as targets for all infants, the interest level of pictures was counterbalanced within the stimulus list, according to a three-level scale of picture interest (animals, vehicles and food, simple inanimates). In this method, picture preference varied systematically according to interest level, with more looking to the 'high' interest items. However, due to strict lexical constraints, such as avoiding shared phonology, it proved to be difficult to counterbalance the relative interest of picture-pairs across priming conditions and within stimulus sub-lists. This resulted in an unbalanced stimulus set, where differences in the general interest of the sub-lists were large and pervasive. Over and above the stimulus group effect, the effect of the priming condition *was* evident for older infants, but legitimate priming effects may have been masked (e.g., for the younger infants) by the 'noise' created by asymmetrical interest levels.

Of the two interest controls appropriate in a quick-primed scenario (counterbalancing labelling across targets and distracters, counterbalancing within-

pair interest across conditions), the first was selected for the current experiment, with the expectation that 'noise' due to unbalanced interest levels would be reduced. To achieve this end, a second stimulus set was created so that each picture could appear as a named target for some infants and as an unnamed distracter for others, allowing counterbalancing across the population.

Having piloted two different stimulus onset asynchronies (SOAs) and found little reliable difference between them, the longer SOA (400ms) was selected for the current experiment.

Method

Participants were recruited as described in previous chapters, and sent the Oxford CDI by post. Most brought their completed CDIs with them on the day of their experimental visit. In a small number of cases, they completed the form on the day of testing, either during their visit, or returned by post shortly afterward.

Forty two 18-month-olds participated in the half-target experiment. Eight were excluded from analysis (seven for parental failure to return an appropriately filled out CDI, and one for experimenter error). Forty seven 24-month-olds participated in the experiment, with seven exclusions (four for parental failure to return a completed CDI, two for experimenter error and one for equipment failure). Thirty five 30-month-olds participated, with four exclusions (one for parental failure to return a completed CDI form, two for equipment failure and one for interruption by a family member during test). The experimental sample thus comprised of 34 18-month-olds (19 male, 15 female; M = 18.3, range = 17.0 to 19.3), 40 24-month-olds (25 male, 15 female; M = 24.1, range = 22.9 to 24.9) and 32 30-month-olds (16 male, 16 female; M = 30.5,

range = 29.6 to 31.6) for whom both behavioural and CDI data were available for analysis.

Total receptive CDI scores were checked against previously collated norms, and infants' comprehension of test items was checked. Following preliminary checks, one 18-month-old was removed from analysis for an extremely low receptive CDI score (in the 5th percentile of previously collected CDIs for 18-month-olds). To ensure that individual differences between participants did not outweigh differences between priming conditions, eleven further eighteen-month-olds were removed from analysis when it was observed that they did not contribute trials to each of the four priming conditions. For three 18-month-olds, CDIs were anonymised prior to recording a full CDI comprehension score. As these individuals had passed preliminary checks for vocabulary size and stimulus comprehension, mean replacement was used, giving the three individuals the mean score of 18-month-olds included in analysis.

Materials

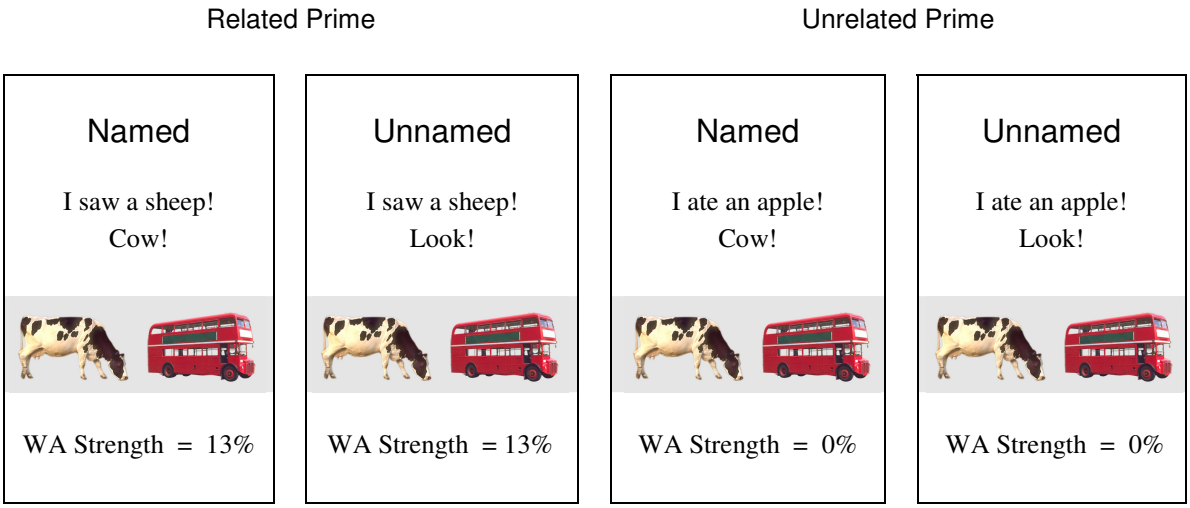
Stimuli for the current experiment were selected from the pool of words understood by more than 50% of 18-month-olds, established in Chapter 2 (Table 2.7). It was necessary to include a small number of additional words which did not reach the original criteria for inclusion in the stimulus list. The body parts *hand* and *foot* had comprehension rates at 18 months of 78% and 71%, respectively, but were not included in the original list of stimulus words because they were not considered visually 'isolatable' objects. Given the larger number of prime-target pairs required in the current experiment (24 instead of 12), the visual criterion was relaxed for this experiment. Similarly, *door* (87%), *window* (62%), *ear* (83%) and *eye* (87%) were added to the stimulus pool. Finally, the words *sun* (40%) and *moon* (43%) did not

quite achieve the 50% comprehension rate cutoff, but as they were known by almost half of the infants at 18-months, they were included to complete the stimulus set. Forty eight words were selected for use as primes, targets and distracters, and arranged into four stimulus lists, in each of two stimulus sets.

Within a stimulus set, picture pairs were yoked across the four lists, and each infant only saw each picture pair once. Across lists, each target occurred in each of the two priming conditions: 'Related' word-pairs were basic-level taxonomic sisters (e.g., prime: *sheep*; target: *cow*) with an attested forward association in adult British English (Moss & Older, 1996 and cf. Appendix D). Unrelated word-pairs shared no semantic or associative relationship, and no phonological onset or rhyme (e.g., prime: *apple*; target: *cow*). In half of the related trials and half of the unrelated trials, the target remained unnamed, and was replaced by a neutral attention phrase (e.g., *Look!*). Distracters shared no phonological, semantic or associative relationship with the prime or target (e.g., prime: *sheep*, target: *cow*, distracter: *bus*). In this two-by-two design, naming and priming were counterbalanced across picture pairs and across prime types. The four stimulus conditions are illustrated in Figure 6.02.

The second stimulus set was constructed from the same items in order that each picture could occur equally often as a named target as an unnamed distracter. Primes were counterbalanced across stimulus sets, meaning that a prime preceded a related target in one stimulus set, and an unrelated target in the other. Picture pairs were yoked within a stimulus set, and re-paired across stimulus sets. Each infant saw 12 trials from a single stimulus list. The eight stimulus organisations are given in Appendix G.

Figure 6.02. A yoked picture pair in four stimulus conditions



Notes. Word Association (WA) Strength from (Moss & Older, 1996).

Visual stimuli were selected and edited as in previous chapters, with the exception that pictures appeared on a 10% grey background. Visual stimuli are given in Appendix H. Audio stimuli were recorded and edited as in previous chapters. As in the minimally primed experiment, primes were recorded at the end of a short sentence. To reduce the likelihood of lost trials to infant 'boredom', the prime phrase in the current experiment was shorter, (e.g., *I saw a cat*!). The durations of priming phases are given in Table 6.1.

Table 6.1. Duration of prime phrases

Prime phrase	Duration (ms)
I ate an apple	1053
I bought a bib	935
I ate a biscuit	1103
I saw a boot	1099
I bought a bowl	1302
I saw a car	1373
I saw a cat	1118
I ate some cheese	1290
I bought a cot	1135
I bought a cup	1141
I saw a door	1514
I saw an ear	1094
I saw an elephant	1336
I saw a hand	1299
I bought a hat	1134
I saw a lion	1278
I saw a lorry	1174
I saw a pig	1200
I saw a rabbit	1396
I saw a sheep	1059
I saw the sun	1345
I saw a table	1134
I saw a train	1372
I bought a watch	1231
M (*SD*)	1213 (137)

Notes. Onsets and offsets of tokens established using visual waveform editing software.

As in the previous experiment, targets were recorded as single-word utterances. The durations of target words are given in Table 6.2. As half of the trials were to contain no target label, there was a possibility that infants might become distracted or dissatisfied if they heard the same neutral ‘dummy target’ on numerous

occasions. To maintain infants' attention, the three attention phrases were used, so that each infant heard each only twice. The durations of dummy targets are given in Table 6.3. The mean duration of dummy targets is (720ms) was extremely close to the mean duration of the target labels (718ms). Dummy targets were substituted into unnamed trials using a Latin Square order.

Table 6.2. Duration of target words

Target word	Duration (ms)
banana	826
bed	478
bicycle	799
boat	682
bread	588
bus	577
cake	658
chair	683
clock	759
coat	739
cow	669
dog	591
duck	584
eye	655
foot	811
horse	884
moon	765
mouse	892
nappy	691
plate	692
shoe	801
spoon	867
tiger	792
window	743
M (*SD*)	718 (108)

Notes. Onsets and offsets of tokens established in visual waveform editing software.

Table 6.3. Duration of dummy targets

Target word	Duration (ms)
look	665
hey	769
wow	728
M (*SD*)	720 (52)

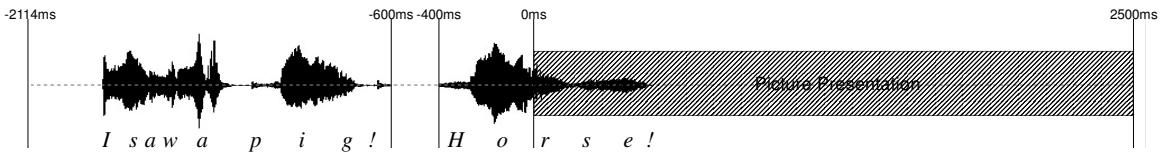
Notes. Onsets and offsets of tokens established in visual waveform editing software.

Procedure

Following a short 'settling in' period in the play area, the 'infant timed' procedure was conducted as described as in the previous chapters, with the exception that only a single SOA was used (400ms), and testing was conducted in a different booth. While the screen was blank, the priming phrase began (e.g., *I saw a pig!*), followed by an inter-stimulus interval (ISI) of 200 ms, then the target word in isolation (e.g., *Horse!*). Pictures began at a stimulus onset asynchrony (SOA) of 400ms from the onset of the target word, and remained onscreen for 2,500ms.

A schematic of trial timing is given in Figure 6.03, where the onset of picture presentation is treated as the origin (0ms). All target words overlapped the onset of picture presentation.

Figure 6.03. Schematic of trial timing



Origin (0ms) indicates the onset of the test-phase of the trial, when pictures are presented on the screen. Wave-forms from actual stimulus tokens, as produced by GoldWave visual waveform editing software (1995, v5.10).

In the previous experiment, it was noted that infants tended to look up at the centrally mounted speaker during the auditory presentation of the priming phrase, while the screen remained dark. It was also noticed that some infants looked away from the screen completely, and did not return their attention to the screen until late in the trial. In the current experiment, shorter priming phrases were used to reduce the likelihood of lost trials. In addition, a small attention-device was displayed onscreen during the priming phrase. Previous pilots using a standard fixation cross had

suggested that the static shape was unappealing to infants. The attention-device used in the current experiment was a circle with a blue spiral pattern, which rotated slowly and pulsed in size. The device was animated from a still gif in Flash. It appeared onscreen from the onset of the trial until 1,000ms. It acted as a fixation point at the beginning of the trial, but concluded prior to the onset of most prime words. The attention-device is illustrated in Fig 6.04, and its location is illustrated in Appendix E.

Target side was counterbalanced. Each stimulus list was presented in two random orders. Infants were tested in a booth adjacent to the one used in previous experiments. A diagram of the screen configuration are given in Appendix E (Booth 2).

Figure 6.04. Onscreen attention device



Scoring & measures

Scoring was conducted as in previous experiments, with the exception that frame-by-frame coding of behavioural states was conducted using custom coding software. Macro- and micro-level timing measures were calculated. The proportion of target looking (PTL) assessed relative preference for the named target picture during the course of the whole trial. The direction of the first fixation was used to calculate the percentage of first fixations to the target, as a measure of accuracy prior to fixation. Time-to-fixate, and time-to-disengage, micro-level measures of reaction time

(RT), assessed the speed of infant eye-movements. Detailed time courses for each condition were also produced for each age group.

Results

Age group and vocabulary size

The mean receptive CDI score for 18-month-olds was 201 words (*SD* = 41) out of a possible 416. 24-month-olds knew 324 words (*SD* = 72), and 30-month-olds, 357 words (*SD* = 62). Thirty-month-olds had larger receptive vocabularies than 24-month-olds ($z(32, 40) = 2.27$, $p = .02$), who, in turn, had larger receptive vocabularies than 18-month-olds ($z(40, 22) = 5.26$, $p < .001$).

Eighteen-month-olds were reported to understand a mean of 10 of the 12 primes (*SD* = 1.2) and 10 of the 12 targets (*SD* = 1.6). Twenty four-month-olds were reported to understand a mean of 11 primes (*SD* = 1.2) and 11 targets (*SD* = 1.2). Thirty-month-olds were reported to understand a mean of 12 primes (*SD* = 0.7) and 12 targets (*SD* = 1.0). Twenty four-month-olds knew significantly more of the test items than 18-month-olds ($z(40, 22) = 4.35$, $p < .001$). Infants in the two older age groups did not differ significantly in the number of stimulus words known ($z(32, 40) = 0.12$, n.s.).

Lexical comprehension and trial exclusions

Of the original 1,116 trials available for analysis (18m: 264 trials, 24m: 468 trials; 30m: 384 trials) only those trials in which both the prime and the target were reported as known were included in analysis. This resulted in the removal of 72 trials

(27%) from 18-month-olds, 48 trials (10%) from 24-month-olds, and 24 trials (6%) from 30-month-olds. Trials in which infants did not fixate the picture areas of the screen for more than 80ms were also removed from the analysis, accounting for a further 9 trials in total.

The smaller number of trials available for 18-month-olds was predicted to generate larger variance in this age group, complicating a straightforward analysis of age. For this reason, analyses were conducted separately for the younger infants, while data from the two older age groups were pooled. Finally, for the analysis of RT measures, 67 trials were excluded because the infant looked to one of the picture areas prior to the onset of pictures.

As in the previous chapter, the direction of the first fixation was also used to calculate the percentage of correct (target) first fixations, and was treated as a dependent variable in the RT measures, time-to-fixate and time-to-disengage.

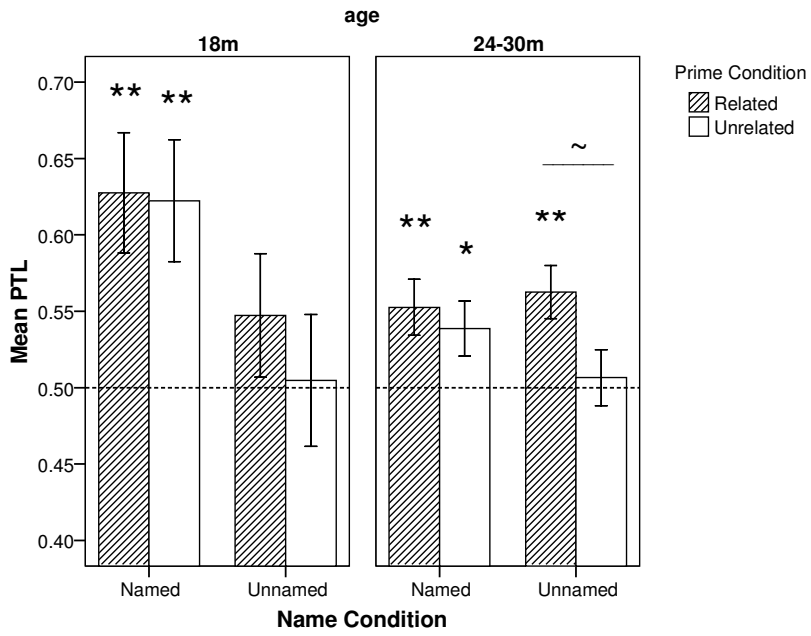
Proportion of target looking (PTL)

When participant means for PTL were created in each of the four conditions, it introduced substantial asymmetry across the eight stimulus sub-lists, complicating straightforward parametric analysis. Instead, raw PTL scores for each trial were considered, and the data here were sufficiently symmetrical for analysis. As preliminary analysis revealed no difference between the PTL of 24- and 30-month-olds, data were pooled across the older age groups.

Figure 6.05 depicts the proportion of target looking (PTL) for younger and older infants, according to whether the target was named, with the priming conditions shown separately. For 18-month-olds, target preference is high in the two named

conditions, and appears to be above the .50 value predicted by chance. In the unnamed conditions, however, PTL is lower, and close to the chance value of .50. In the older age groups, PTL is generally quite close to chance in all conditions, with PTL in the three test conditions (named related, named unrelated, unnamed related) tending to be a little higher than PTL in the control condition (unnamed unrelated). For both younger and older infants, mean PTL in the control condition is at chance – as would be expected given that neither of the auditory stimuli share a relationship (phonological, semantic or associative) with either picture.

Figure 6.05. PTL for younger and older infants according to condition



Notes. Horizontal dashed line indicates chance value of .50. One sample t-tests comparing PTL to chance are marked above each bar, where significant. Wilcoxon's rank sums tests comparing values between priming conditions are marked above pair of bars. $^{**}p<.01$, $^{*}p<.05$, $\sim p<.10$. Whiskers show +/- one standard error.

For younger infants, a three-way ANOVA compared the effect of name condition (named, unnamed), prime condition (related, unrelated) and stimulus sub-list (one to eight), on PTL. The results of the ANOVA are given in Table 6.4, where a

main effect of target condition is evident, with no other significant effects or interactions.

18-month-olds showed greater preference for named targets (M = .63, SD = .27) than for unnamed targets (M = .53, SD = .29), and Wilcoxon's rank sums tests confirmed that the difference was significant ($z(97, 94) = 2.60$, $p = .01$). Single sample t-tests demonstrated that it was only in the named trials that target preference was above the chance value of .50 (named: $t(93) = 4.49$, $p < .001$, $d = 0.46$; unnamed: $t(96) = 0.89$, n.s., $d = 0.09$). This finding is consistent with comprehension of named items which were reported as known, as reported in previous chapters. The effect of target label did not differ significantly according to priming condition or stimulus list.

Table 6.4. Three-way ANOVA of PTL (18m)

Source	*df*	*F*	*p*	partial η^2
name	1	5.18*	.02	.032
prime	1	0.13	.72	.001
sub-list	7	0.33	.94	.014
name * prime	1	0.17	.68	.001
name * sub-list	7	0.62	.74	.027
prime * sub-list	7	0.92	.49	.039
name * prime * sub-list	7	0.19	.99	.008
Error	159	(0.08)		

Notes. Mean square error given in brackets. *p<.05

The older age groups showed a different pattern of response. A three-way ANOVA compared the effect of target condition (named, unnamed), prime condition (related, unrelated) and stimulus sub-list (one to eight) on PTL. The results of the ANOVA are given in Table 6.5, where a significant main effect of prime condition is evident, along with a significant main effect of stimulus sub-list. In general, PTL was higher in trials which contained a related prime (M = .56, SD = .25) than for trials

which contained an unrelated prime (M = .52, SD = .25). This difference did not approach significance in a Wilcoxon's rank sums test (z(381, 391) = 1.63, n.s.), with a p-value of .10.

Table 6.5. Three-way ANOVA of PTL (24-30m)

Source	*df*	*F*	*p*	partial η^2
name	1	0.57	.45	.001
prime	1	3.92*	.05	.005
sub-list	7	2.13*	.04	.020
name * prime	1	1.74	.19	.002
name * sub-list	7	0.32	.94	.003
prime * sub-list	7	0.69	.68	.006
name * prime * sub-list	7	0.64	.72	.006
Error	740	(0.06)		

Notes. Mean square error given in brackets. *p<.05

When the trial conditions were investigated separately, older infants showed significant target preference in all three test conditions: in both named conditions (related: M = .55, SD = .26, t(195) = 2.85, p = .005, d = 0.20; unrelated: M = .54, SD = .25, t(194) = 3.60, p < .001, d = 0.16), and in the unnamed related condition (M = .56, SD = .24, t(191) = 2.15, p = .03, d = 0.25). No preference for the target was shown in the control condition, where the target was neither named nor primed (M = .51, SD = .25, t(188) = 0.36, n.s., d = 0.03), suggesting that *a priori* interest in the different pictures was adequately balanced across the population.

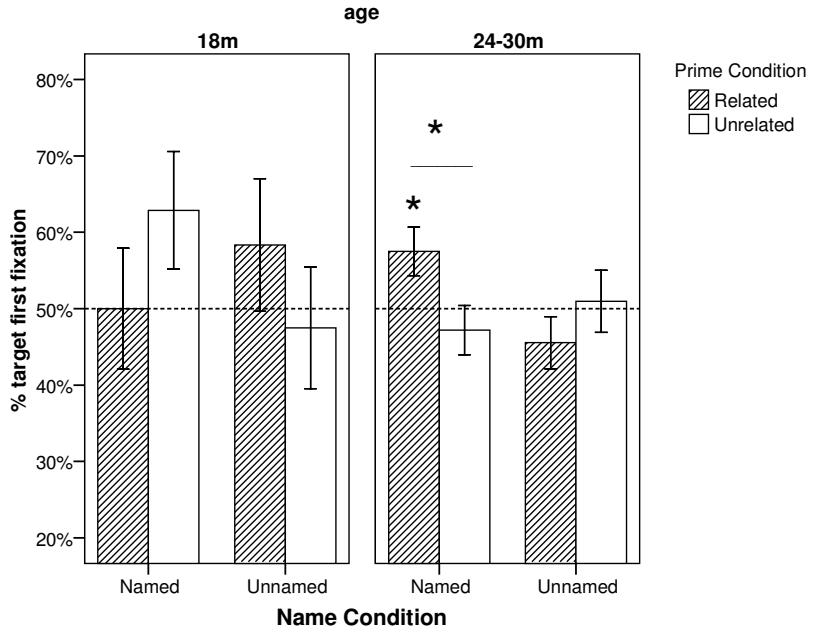
This pattern of behaviour suggests that both the target word and the prime word generated above chance target preference for older infants. However, the lack of a reliable difference between the test conditions and the control (as would have been demonstrated by an interaction between the prime condition and the name condition), suggests that this trend is not reliable. Furthermore, the lack of any difference

between the three test conditions does not provide clarification of whether lexical facilitation or visual mediation might have been at work.

Percentage of target first fixations

For each infant, the number of trials in which their first screen fixation was to the target picture was calculated, as a proportion of all valid trials in each stimulus condition (trials were not included if the infant was fixating a picture area at image onset). The percentage of target first fixations is plotted in Figure 6.06, with data pooled across the two older age groups. In this plot, the variance is large for the younger infants, remaining close to the chance value of 50%. The large amount of overlap in values means that little difference is evident between the four stimulus conditions. Older infants exhibit less variance in the percentage of first fixations to target. Like younger infants, their values remain close to chance. The two named conditions appear to differ somewhat from each other, with a greater percentage of correct first fixations in the named related condition than in the named unrelated condition.

Figure 6.06. Percentage of first fixations to target



Notes. Horizontal dashed line indicates chance value of 50%. One sample t-tests comparing each value chance are marked above each bar, where significant. Wilcoxon's rank sums tests comparing values in different priming conditions are marked above pairs of bars, where significant. *p<.05. Whiskers show +/- one standard error.

To investigate the pattern in more detail, younger infants' percentage of first fixations to targets were subjected to a three-way ANOVA comparing naming (named, unnamed), priming (related, unrelated) and stimulus sub-list (one to eight). The results of the ANOVA are given in Table 6.6, where a significant main effect of stimulus sub-list is evident, along with highly significant two-way interaction between priming condition and stimulus sub-list, and a significant three-way interaction between naming, priming and stimulus sub-list. As no significant main effect of prime condition is evident, this finding suggests that any apparent priming effects are an outcome of stimulus organisation, combined with the discrete variance introduced by the different populations of infants who saw items in the different conditions. The three-way interaction is difficult to interpret, given the number of sub-lists. The finding suggests, however, that behaviour was heavily influenced by particular stimulus combinations in the eight sub-lists.

Table 6.6. Three-way ANOVA of percentage of first fixations to target (18m)

Source	*df*	*F*	*p*	partial η^2
name	1	0.35	.56	.006
prime	1	0.62	.44	.011
sub-list	7	2.12~	.06	.216
name * prime	1	2.54	.12	.045
name * sub-list	7	1.18	.33	.133
prime * sub-list	7	3.60**	<.01	.318
name * prime * sub-list	7	2.35*	.04	.234
Error	54	(958.51)		

Notes. Mean square error given in brackets. **p<.01, *p<.05, ~p<.10

For older infants, data were analysed in a four-way ANOVA comparing the influence of the age group (18m, 24m), target name (named, unnamed), prime (related, unrelated) and stimulus sub-list (one to eight) on the percentage of first fixations to the target. The results of the ANOVA are given in Table 6.7, where a highly significant main effect of stimulus sub-list is evident, along with a significant interaction between naming and priming. Two further interactions approached significance: a two-way interaction between age group and prime, and a four-way interaction between age, naming, priming and sub-list. The main effect of stimulus sub-list demonstrated that targets in different sub-lists attracted significantly different percentages of first fixations, suggesting that some items were more visually interesting than others.

Table 6.7. Four-way ANOVA of percentage of first fixations to target (24-30m)

Source	*df*	*F*	*p*	partial η^2
age	1	0.02	.89	<.001
name	1	1.15	.28	.005
prime	1	0.57	.45	.003
sub-list	7	3.26**	<.005	.094
age * name	1	0.23	.63	.001
age * prime	1	3.39~	.07	.015
name * prime	1	4.28*	.04	.019
age * name * prime	1	1.27	.26	.006
age * sub-list	7	0.25	.97	.008
name * sub-list	7	0.42	.89	.013
age * name * sub-list	7	0.35	.93	.011
prime * sub-list	7	1.55	.15	.047
age * prime * sub-list	7	0.80	.58	.025
name* prime * sub-list	7	0.66	.70	.021
age * name * prime * sub-list	7	1.81~	.09	.055
Error	219	(838.17)		

Notes. Mean square error given in brackets. **p<.01, *p<.05, ~p<.10

To investigate the significant name-by-prime interaction, data in each naming condition were assessed separately. For named targets, there were significantly more first fixations to the target in the related prime condition (M = 58%, SD = 27%) than there were in the unrelated prime condition (M = 47%, SD = 27%), and Wilcoxon's signed ranks tests confirmed that the difference between ranked scores was significant (z(71, 71) = 2.19, p = .03). For unnamed targets, by comparison, no difference was observed in the percentage of first fixations which were directed to the target (related: M = 47%, SD = 27%; unrelated: M = 51%, SD = 34%; z(70, 71)= 1.07, n.s.). This pattern of behaviour demonstrates that infants only showed greater accuracy in directing their first look to the target, when it was both named and primed by a related word. This finding suggests that the infants are better able to recognise the target *prior to fixation* when the target word had been primed by a related word. This

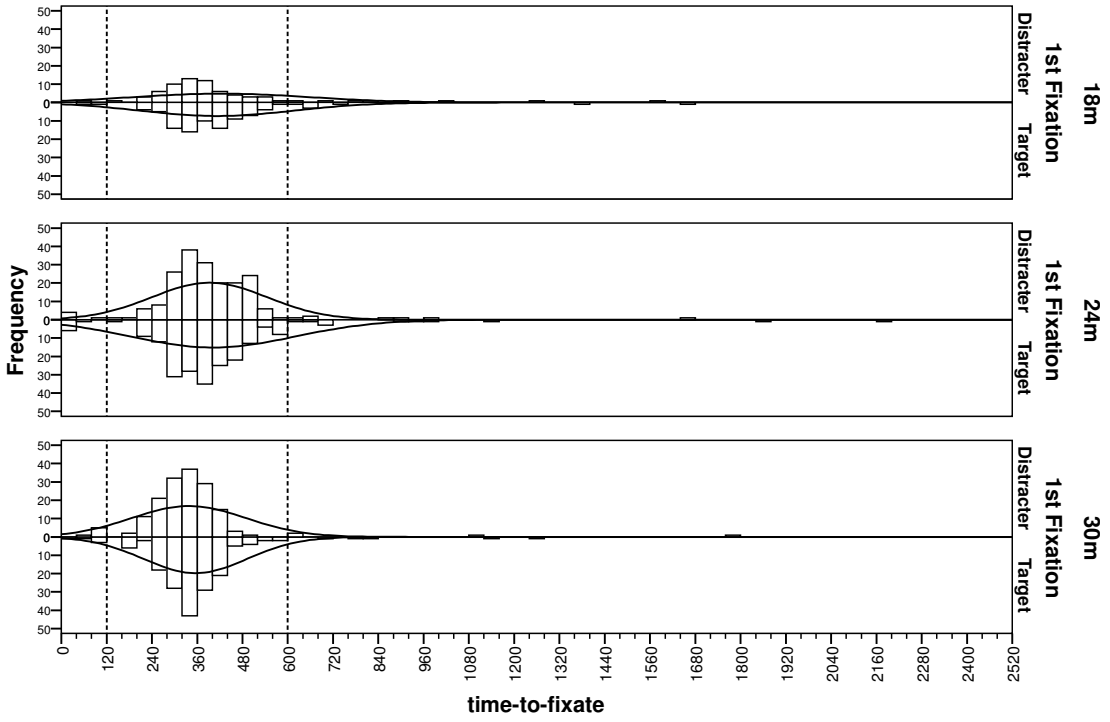
finding is consistent with the lexical facilitation account, whereby the effect of the prime is only evident when the target word is also heard.

Concerning the near-significant age-by-prime interaction, stimulus condition only significantly affected performance for 24-month-olds ($\chi^2(3) = 9.29$, $p = .03$), and not for 30-month-olds ($\chi^2(3) = 2.44$, n.s.). The near-significant four-way interaction is difficult to interpret in detail, but indicates that the significant target-by-prime interaction varied across age, and was inconsistent across the stimulus sub-lists.

Time-to-fixate

The amount of time taken from the onset of the pictures to the onset of the first fixation is plotted for each trial in Figure 6.07. In this plot, there are very few 'predictive' screen fixations, and in all three age groups, predictive fixations have ceased by 120ms. The majority of fixations fall between 120ms and 600ms, for all age groups, and the distribution of these fixations suggests that they are responsive, and close to Gaussian distribution. Little difference is evident between the speed of fixations to the target and to the distracter. Preliminary analysis confirmed that there was no difference between the speed of fixations to the target and to the distracter (18m: $z(62, 84) = 1.11$, n.s.; 24-30m: $z(333, 348) = 0.76$, n.s.). Data were therefore pooled across direction for further analyses.

Figure 6.07. Histogram of time-to-fixate over the course of the trial

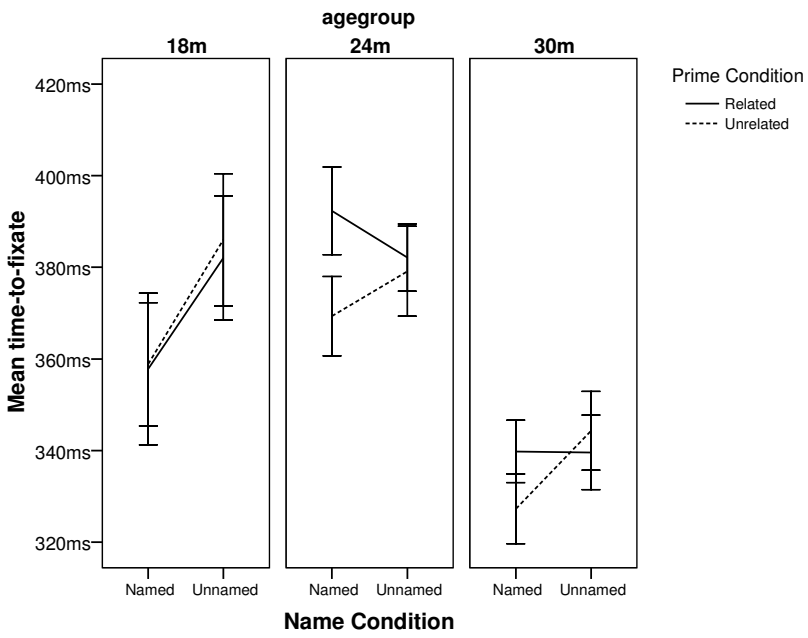


Notes. Origin (0ms) indicates onset of pictures. Analysis window marked from 120ms to 600ms. Normal curves plotted over whole data set.

Figure 6.08 depicts time-to-fixate in each age group according to name and prime condition. For younger infants, fixations to targets appear to be faster in the named conditions than in the unnamed conditions. Note that this contrast is different to those presented previously. Previous chapters have focussed on different reaction times to a (named) target as opposed to an (unnamed) distracter. This chart is pooled across look direction, and presents the speed of fixations to either picture, according to the naming condition – that is, whether the second word presented in the trial was a known label (e.g., *dog*) or a dummy attention phrase (e.g., *Look!*). The trend for younger infants to look faster in the named conditions suggests that they are making faster fixations in response to known labels – regardless of which direction they look. Little difference is apparent between the priming conditions for younger infants.

For older infants (24- and 30-month-olds), in Figure 6.08, when two content words are presented in sequence (the named conditions) first fixations appear to be slightly slower when the two words are related, than when they are unrelated. This trend fits with the finding reported in previous chapters that processing a semantic match between stimuli (label and object, or related prime and target) increases infants' processing time, thus slowing a fixation. By contrast, when the second word is an attention phrase, no difference between priming conditions is apparent, confirming that primes were indeed unrelated to the dummy target words.

Figure 6.08. Time-to-fixate according to age group (pooled across look direction)



Notes. Whiskers show +/- one standard error.

18-month-olds

Time-to-fixate for younger infants was subjected to a three-way ANOVA comparing the influence of naming (named, unnamed), priming (related, unrelated) and stimulus

sub-list (one to eight). The results of the ANOVA are given in Table 6.8, where no significant main effects or interactions are evident. There is a near-significant main effect of stimulus sub-list (with a *p*-value of .08 approaching the alpha-level of .05), suggesting that the apparent trend for faster responses in the naming condition (observed in Figure 6.08) may have been an artefact of the stimulus combinations.

Table 6.8. Three-way ANOVA of time-to-fixate (18m)

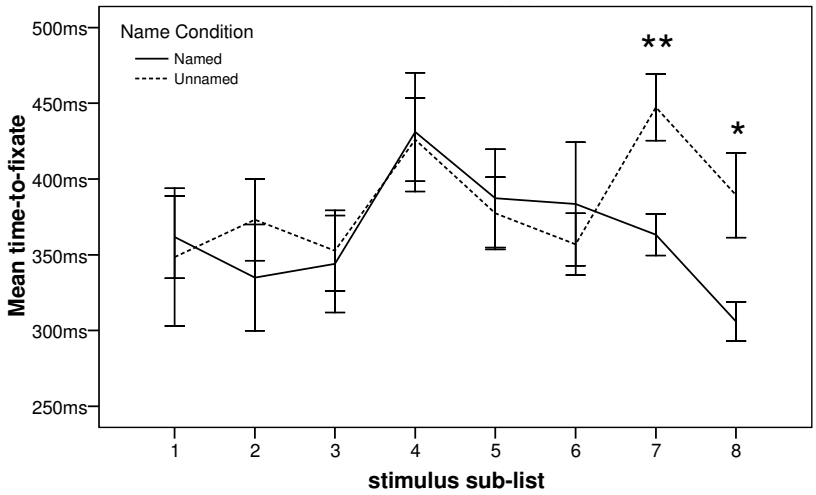
Source	*df*	*F*	*P*	partial η^2
name	1	2.01	.16	.017
prime	1	0.02	.89	<.001
sub-list	7	1.88~	.08	.104
name * prime	1	0.56	.46	.005
name * sub-list	7	0.76	.62	.045
prime * sub-list	7	1.36	.23	.077
name * prime * sub-list	7	0.83	.57	.048
Error	144	(7128.48)		

Notes. Mean square error given in brackets. ~p<.10

To explore this possibility, the Figure 6.09 plots time-to-fixate for 18-month-olds according to naming condition and to stimulus sub-list. In this plot it is clear that the observed trend was lead by only two of the eight sub-lists (seven and eight). These sub-lists contained the targets *mouse, bread, clock*, and *cow, cake, bed*, respectively. According to the interest levels established in the previous chapter, both lists contained one 'high' interest target (an animal), one 'medium' interest target (food) and one 'low' interest target (a household object). There is no clear reason for naming to be speed responses more in these sub-lists, than in other sub-lists. It is thus likely that the differences in the speed of responses to targets in these stimulus sub-lists were caused by the different individuals who saw the stimuli in the different conditions.

The high level of lexical exclusion for 18-month-olds increases this possibility, as few infants saw all four conditions.

Figure 6.09 Time-to-fixate according to stimulus sub-list (18m)



Notes. Wilcoxon's ranked sums tests comparing naming conditions are marked above values where significant. **p<.01, *p<.05. Whiskers show +/- one standard error.

To summarise the results for 18-month-olds, no reliable effects of test condition were evident, other than irregular variance between stimulus sub-lists.

Older infants

The two older age groups were analysed separately in the micro-level measures, due to the large difference in reaction times evident in Figure 6.08 and 6.12. For 24-month-olds, there was insufficient homogeneity of variance in time-to-fixate for analysis by ANOVA. To check for main effects of naming and priming, data were first assessed at the overall level. No difference was observed in time-to-fixate pictures according to naming condition (named: M = 382ms, SD = 89ms; unnamed: M = 381ms, SD = 82ms, z(185, 185) = 0.21, n.s.), nor priming condition (related: M = 387ms, SD = 85ms; unrelated: M = 374ms, SD = 87ms, z(185,

185) = 1.14, n.s.). To check for an interaction between naming and priming, the data were analysed separately in each of the naming conditions. In named trials, the related priming condition generated slightly slower fixations (*M* = 392ms, *SD* = 95ms) than the unrelated trials (*M* = 369ms, *SD* = 81ms), although this difference was not significant ($z(98, 87) = 1.4$, n.s.). In the unnamed trials, no difference was apparent between priming conditions (related: *M* = 382ms, *SD* = 72ms; unrelated: *M* = 379ms, *SD* = 93ms; $z(96, 89) = 0.13$, n.s.).

Concerning the role of stimulus sub-list in 24-month-olds time-to-fixate, according to a one way ANOVA, there was no general effect of sub-list on time-to-fixate ($F(7, 362) = 1.57$, n.s., partial $\eta^2 = .03$). Within the four trial conditions, there was also no systematic effect of sub-list, according to Kruskal Wallis tests (named related: $\chi^2(7) = 5.40$, n.s.; named unrelated: $\chi^2(7) = 7.94$, n.s.; unnamed related: $\chi^2(7) = 10.81$, n.s.; unnamed unrelated: $\chi^2(7) = 2.86$, n.s.). Thus no reliable effects of trial condition nor of stimulus organisation were evident in time-to-fixate.

Overall time-to-fixate was faster for 30-month-olds (*M* = 337ms, *SD* = 68ms) than for 24-month-olds (*M* = 338ms, *SD* = 68ms), and this difference was significant ($z(311, 370) = 6.78$, $p<.001$). Data from 30-month-olds was sufficiently homogeneous for analysis in a three-way univariate ANOVA comparing the influence of name (named, unnamed), prime (related, unrelated) and stimulus sub-list (one to eight). The results of the ANOVA are given in Table 6.9, where no significant main effects or interactions are evident. There is a near-significant three-way interaction between target, prime and stimulus sub-list, with a *p*-value of .06 approaching the alpha-level of .05.

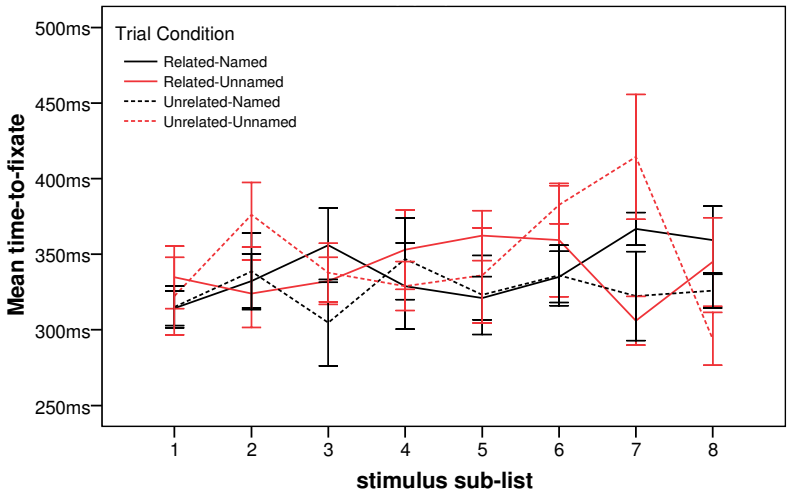
Table 6.9. Three-way ANOVA of time-to-fixate (30m)

Source	*df*	*F*	*p*	partial η^2
target	1	2.11	.15	.007
prime	1	0.04	.84	<.001
sub-list	7	0.96	.46	.024
target * prime	1	1.98	.16	.007
target * sub-list	7	0.64	.72	.016
prime * sub-list	7	1.31	.24	.032
target * prime * sub-list	7	1.94~	.06	.046
Error	279	(4611.32)		

Notes. Mean square error given in brackets. ~p<.10

To explore the three-way interaction further, Figure 6.10 depicts time-to-fixate in each of the stimulus sub-lists, according to the trial condition. In this plot no systematic pattern is evident between the trial conditions, suggesting that the source of the variance is due to different individuals encountering groups of stimuli in different conditions.

Figure 6.10 Time-to-fixate according to trial condition (30m)



Notes. Whiskers show +/- one standard error.

Taken together, the findings from all three age groups suggest that the large number of stimulus groups complicates analysis of time-to-fixate. The dual source of variance

(stimulus combinations coupled with different participants experiencing stimuli in different conditions) makes it difficult to observe systematic patterns in the looking behaviour of infants in this age range.

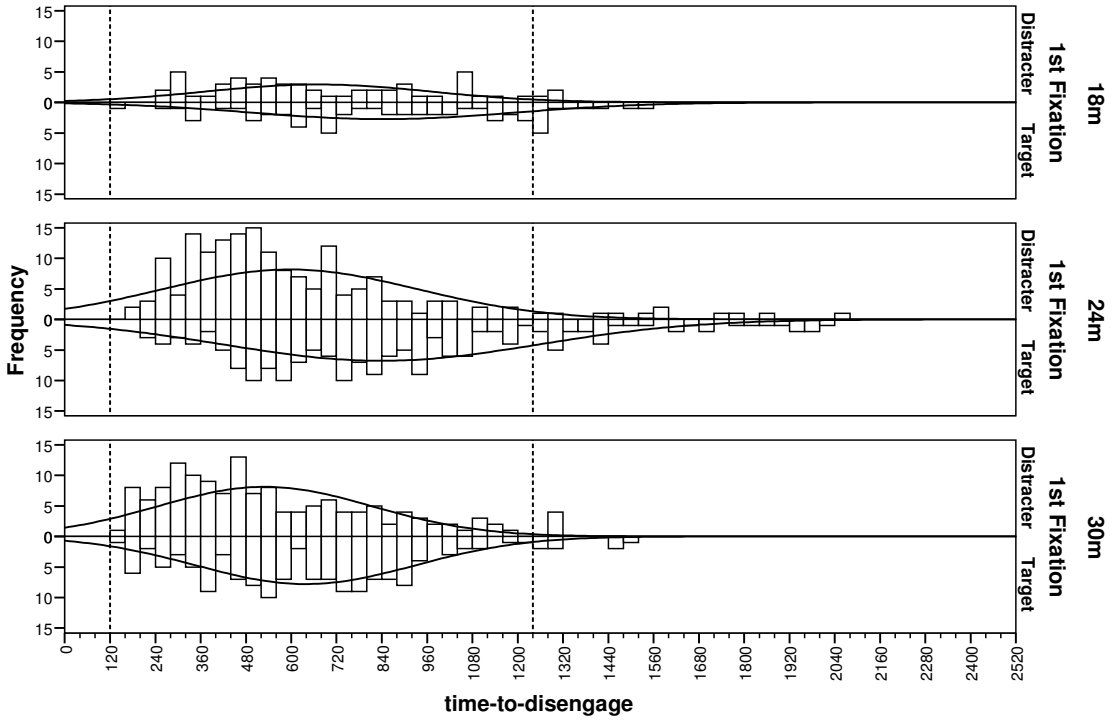
To summarise the results of time-to-fixate for the older infants, there were significant differences between the age groups, with faster responses from older infants. There were also no significant differences between the look directions, and no reliable effects of test conditions (name, prime).

Time-to-disengage

For those trials in which the first fixation fell within the fixation window (120ms to 600ms), time-to-disengage was calculated as the time from the onset of the first fixation to a picture to the onset of the first saccade away. Figure 6.11 plots the distribution of time-to-disengage over the remainder of the trial. In all age groups, time-to-disengage begins after 120ms, and the majority of fixations have concluded by 1240ms. These values are taken as the window for analysis. In this figure, time-to-disengage appears to be slower when the first fixated picture is the target. The distribution was non-Gaussian, and non-parametric analyses were therefore conducted.

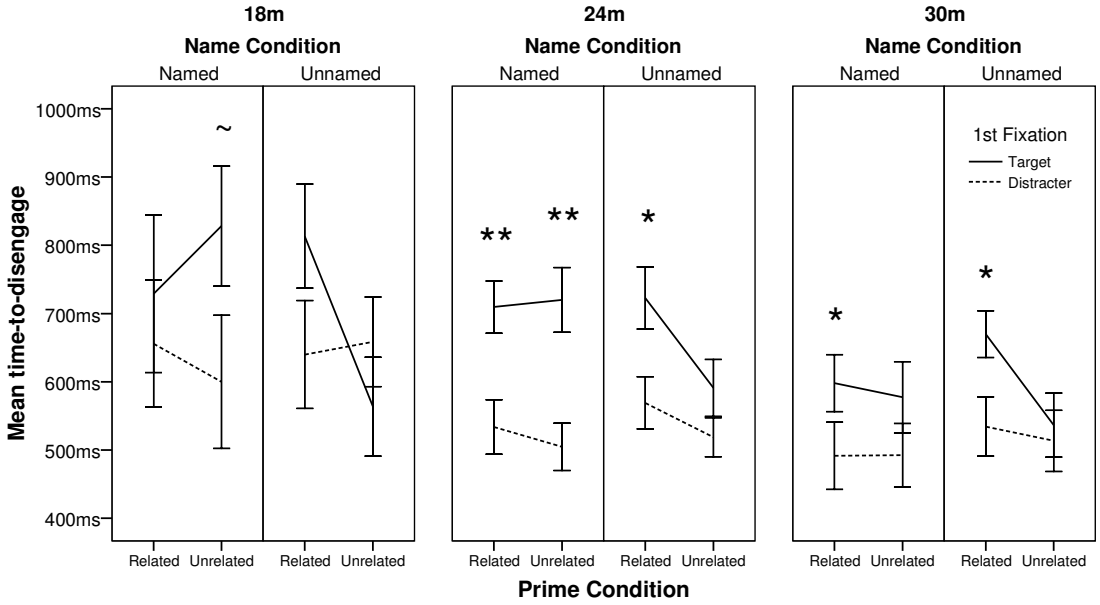
Figure 6.12 depicts time-to-disengage in each of the age groups, according to naming condition, priming condition and direction of first fixation. The general trend for time-to-disengage from distracters to be faster than time-to-disengage from targets is apparent in all three age groups. The trend is more pronounced in the older infants, and is not evident in the control condition (unnamed unrelated).

Figure 6.11. Histogram of time-to-disengage over the course of the trial



Notes. Origin (0ms) indicates the onset of the first fixation. Vertical dashed lines indicate window from 120ms to 1240ms. Normal curves plotted over the whole data set.

Figure 6.12. Time-to-disengage in three age groups



Notes. Wilcoxon's rank sums tests comparing time-to-disengage in two look directions are marked above pair of values where significant. **p<.01, *p<.05, ~p<.10. Whiskers show +/- one standard error.

18-month-olds

For 18-month-olds, data were insufficiently homogeneous to analyse according to a univariate ANOVA incorporating stimulus sub-list. Time-to-disengage was analysed in a three-way ANOVA comparing the influence of direction of first fixation (target, distracter), name condition (named, unnamed) and prime condition (related, unrelated). The results of the ANOVA are given in Table 6.10, where no significant main effects or interactions are evident. The three-way interaction between name, prime and direction barely approaches significance with a *p*-value of .09. No influence of stimulus sub-list was evident in Kruskall Wallis tests at the overall level ($\chi^2(7) = 0.81$, n.s.), according to look direction (target: $\chi^2(7) = 0.22$, n.s.; distracter: $\chi^2(7) = 0.27$, n.s.), or within trial condition (named related: $\chi^2(6) = 0.14$, n.s.; named unrelated: $\chi^2(6) = 0.49$, n.s.; unnamed related $\chi^2(7) = 0.12$, n.s.; unnamed unrelated: $\chi^2(7) = 0.44$, n.s.). The pattern of time-to-disengage was thus unsystematic for 18-month-olds.

Table 6.10. Three-way ANOVA of time-to-fixate (18m)

Source	*df*	*F*	*p*	partial η^2
direction (dir)	1	2.41	.13	.029
name	1	0.32	.58	.004
prime	1	0.59	.44	.007
dir * name	1	0.83	.37	.010
dir * prime	1	0.21	.65	.003
name * prime	1	1.26	.27	.015
dir * name * prime	1	3.00~	.09	.035
Error	82	(81981.05)		

Notes. Mean square error given in brackets. ~$p<.10$

Older infants

Time-to-disengage was insufficiently homogeneous for analysis by univariate ANOVA including stimulus sub-list. The data were instead subjected to a four-way ANOVA assessing the influence of age (18m, 24m), direction of first fixation (target, distracter), name condition (named, unnamed) and prime condition (related, unrelated). The results of the ANOVA are given in Table 6.11, where a highly significant main effect of age is evident, along with a highly significant main effect of look direction, and a significant main effect of prime. The interaction of name with prime approaches significance, with a *p*-value of .08.

Table 6.11. Three-way ANOVA of time-to-fixate (24-30m)

Source	*df*	*F*	*p*	partial η^2
age	1	7.22**	<.01	.013
direction (dir)	1	32.42**	<.001	.058
name	1	0.03	.87	<.001
prime	1	4.86*	.03	.009
age * dir	1	2.49	.12	.005
age * name	1	0.90	.34	.002
dir * name	1	1.36	.24	.003
age * dir * name	1	0.61	.44	.001
age * prime	1	0.02	.88	<.001
dir * prime	1	1.07	.30	.002
age * dir * prime	1	0.29	.59	.001
name * prime	1	3.08~	.08	.006
age * name * prime	1	0.03	.86	<.001
dir * name * prime	1	1.56	.21	.003
age * dir * name * prime	1	0.03	.86	<.001
Error	531	(60619.05)		

Notes. Mean square error given in brackets. ~*p*<.10

Concerning the main effect of age, time-to-disengage was slower for 24-month-olds (*M* = 604ms, *SD* = 248ms), than for 30-month-olds (*M* = 552ms,

SD = 260ms), and this difference was significant ($z(286, 261) = 2.43$, $p = .02$). This finding demonstrates that older infants are faster to disengage from pictures in general (both distracters and targets). Concerning the main effect of look direction, time-to-disengage was slower when the first fixation was to the target (*M* = 641ms, *SD* = 255ms), than when the first fixation was to the distracter (*M* = 521ms, *SD* = 242ms), and this difference was significant ($z(263, 284) = 5.69$, $p < .001$). This finding is consistent with the results of the words-before-pictures experiment (Chapter 4) and the quick primed experiment (Chapter 5), in which targets were disengaged from slower than distracters. The finding is consistent with the notion of rapid online distracter rejection.

Concerning the main effect of prime, time-to-disengage was slower in trials where the auditory prime was related to the target picture (*M* = 602ms, *SD* = 257ms), than when the prime was unrelated (*M* = 554ms, *SD* = 251ms), and this difference was significant ($z(284, 263) = 2.14$, p=.03). As this effect did not interact with look direction, it suggests that related primes generate slower time-to-disengage to all pictures (both targets and distracters).

This main effect was moderated by the near significant interaction between name and prime, such that the difference between priming conditions was larger in the unnamed condition. When assessed separately in the two naming conditions, no difference between priming conditions was evident in trials where the target was named (related: *M* = 588ms, *SD* = 263ms; unrelated: *M* = 566ms, *SD* = 268ms; $z(142, 134) = 0.81$, n.s.). When the target was unnamed, time-to-disengage was slower in related trials (*M* = 616ms, *SD* = 252ms) than in unrelated trials (*M* = 542ms, *SD* = 231ms), and this difference was significant ($z(142, 129) = 2.30$, p =.02). This finding is difficult to interpret as it means infants spend longer looking at *either*

picture following a related prime, but only when they *had not* also heard the target label. While it is possible that this RT difference indexes the cost of processing a cross-modal match, it is difficult to explain why the influence of the prime would be as large when the infant is fixating the target compared to when they are fixating the distracter. It is possible that an interaction between prime, target and fixation direction is masked by the complex stimulus sub-list design.

To check for the influence of stimulus sub-list, the two age groups were assessed separately. For 24-month-olds, sub-list had a near significant effect on the time-to-disengage from pictures ($\chi^2(7) = 13.34$, n.s.), with a *p*-value of .06 approaching the alpha level of .05. For 30-month-olds, sub-list had no observable effect ($\chi^2(7) = 9.11$, n.s.). When assessed according to look direction, sub-list had a near significant effect on the time-to-disengage from targets ($\chi^2(7) = 13.33$, n.s.), with a *p*-value of .07. Sub-list also had a near significant effect on the time-to-disengage from distracters ($\chi^2(7) = 13.23$, n.s.), with a *p*-value of.07.

Time series

Detailed time series were generated from the pool of trials in which the first fixation was within the time-to-fixate window (120ms to 600ms). Figure 6.13 to 6.15 depict the number of fixations to target and distracters in each frame of analysis over the course of the trial, with each of the four stimulus conditions shown separately. Figure 6.13 depicts 18-month-olds, 6.14 depicts 24-month-olds, and 6.15 depicts 30-month-olds.

In Figure 6.13, in all four conditions, there is an increase in fixations to the picture areas over the first 500ms of the trial, indicating the onset of the first fixations.

After this initial increase, the number of target and distracter fixations varies about a consistent mean, which appears not to decline over the course of the trial. This indicates that infants who made their first fixation during the fixation window (120ms to 600ms) did not look away from the screen before the end of the trial (as would have been evident if the mean value had declined). The small number of trials available for 18-month-olds in each condition produces an inconsistent, jagged profile in each of the plots. The large variance makes it difficult to draw strong conclusions from the profiles.

In the two named conditions, on the left of the figure, the number of target fixations is larger than the number of distracter fixations for almost the whole duration of the trial: In the named related condition the difference is very small, but lasts until the end of the trial; In the named unrelated condition, the difference is much more pronounced, showing a 'naming bloom' from about 360ms to 1600ms, following which, the number of fixations to targets and distracters is similar. It is difficult to reconcile the small number of target fixations in the named related condition with the PTL finding of highly significant target preference. It is likely that PTL was boosted in the original analysis by trials in which the infant was fixating an image area prior to picture onset (nine trials), or by trials in which the first fixation was initiated outside the window for first fixations (seven trials). These exclusions constitute 33% of the valid trials available for analysis in the named related condition.

In the unnamed related condition, the number of target fixation is greater than the number of distracter fixations in the first half of the response profile, generating something of a 'naming bloom' between 360ms and 1600ms. The peak is followed by a trough in the opposite direction, however, suggesting that if the related prime generated early target looking, the effect was short-lived. The equal amounts of target

preference and distracter preference are consistent with the finding of no overall target preference in PTL.

In the unnamed unrelated condition, the number of target and distracter fixations is very similar throughout the trial, consistent with the finding of no preference for the target in PTL. Note, however, that overall shape and variance of the plot for the unnamed unrelated condition is very similar to the plot for the named related condition, again, making it difficult to draw conclusions from the shapes of the plots for 18-month-olds.

In Figure 6.14, time courses for the 24-month-olds are plotted. More trials were available for older infants, and this is evident in the relative smoothness of the curves. The rapid increase in the number of fixations from 160ms to 600ms indicates the onset of the first fixation in each trial. No decline in picture looking is evident over the course of the trial, indicating that 24-month-olds did not tend to look away from the screen.

The profiles in all three test conditions (named related, named unrelated, unnamed related) show a 'naming bloom' during the first 1600ms of the trial. In all three conditions this is followed by a switch to distracter preference. This is consistent with the finding that the three test conditions attracted greater than chance target preference in PTL, but that the magnitude of the preference was very small. It is only in the named related condition that the naming bloom follows a higher number of target fixations in the onset stage. This is consistent with the finding that older infants showed an above chance percentage of fixations to the target only in the named related condition. Compared to the three test conditions, the control condition (unnamed unrelated) shows little difference between the number of fixations to targets and to distracters, and little change over the course of the trial.

Figure 6.15 plots the time-courses for 30-month-olds. All four profiles show a rapid onset in picture fixations, with no decline in mean values over the course of the trial, indicating that the oldest age group did not lose interest in the screen before the end of the trial. In the two related prime conditions, high amplitude 'naming blooms' can be seen between 480ms and 1280ms. In the named related condition, this is followed by equal numbers of fixations to targets and distracters for the duration of the trial. In the unnamed related condition, target interest is somewhat more prolonged, lasting until 1760ms. In the named unrelated condition, target interest is smaller in magnitude, but lasts longer, from about 600ms to 1680ms.

In the unnamed unrelated condition, the number of target and distracter fixations is similar throughout the trial, with a slight advantage for targets. Note however that the variance in the unnamed unrelated condition is very similar to the named related condition. Given the magnitude of the variance in these plots, it is difficult to draw firm conclusions about whether looking behaviour in these two conditions is substantially different.

Finally, to investigate the onset-contingent pattern of looking behaviour, Figure 6.16 and Figure 6.17 depict the proportion of target fixations in trials where the first fixation was to the target, and to the distracter, respectively. In both plots and all three age groups, the most salient feature of these plots is the degree of overlap between the four conditions. No point in the time series suggests that it would be a more sensitive window for analysis of the different conditions.

Figure 6.13. Detailed trial time course in four conditions (18m)

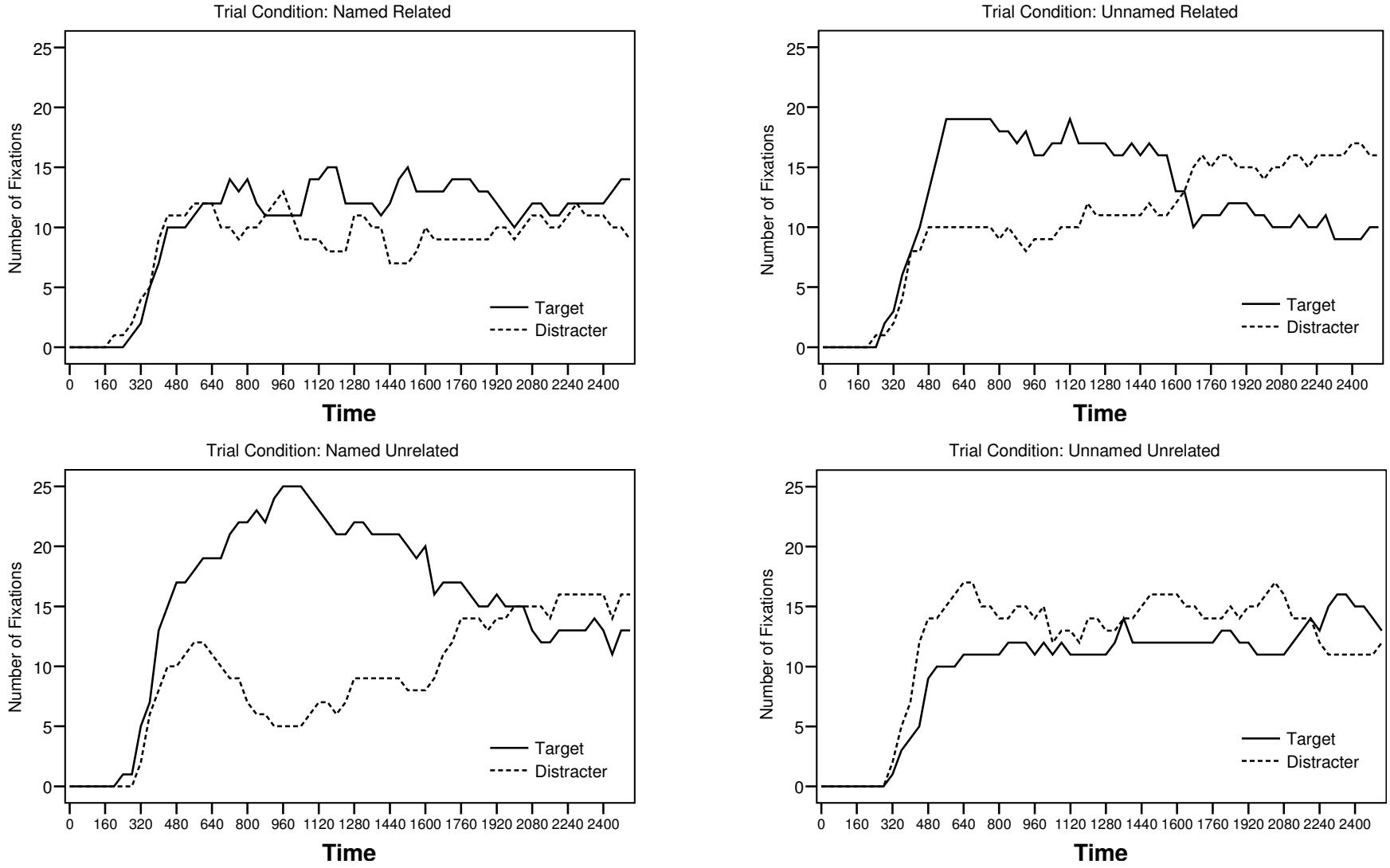


Figure 6.14. Detailed trial time course in four conditions (24m)

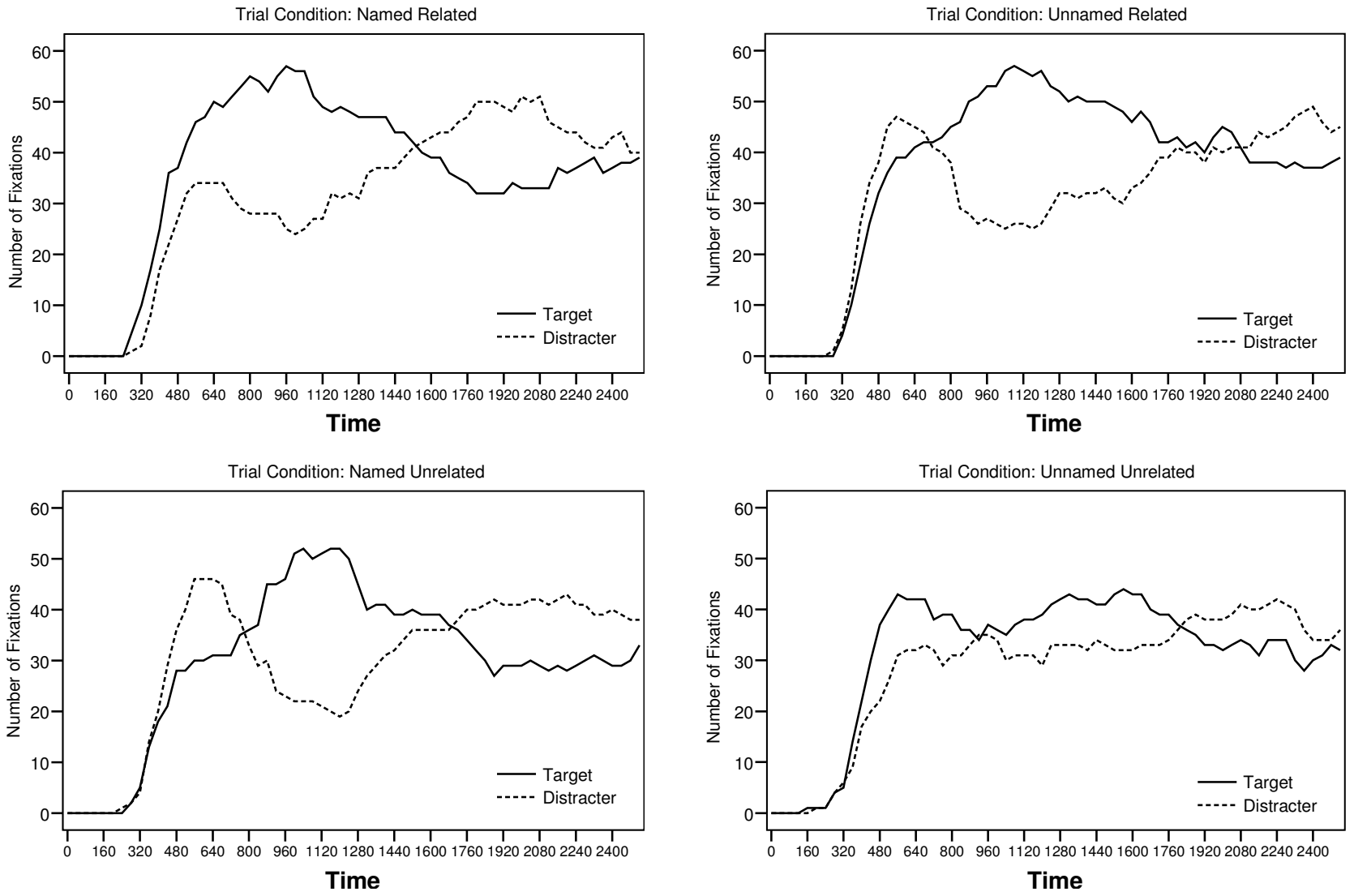


Figure 6.15. Detailed trial time course in four conditions (30m)

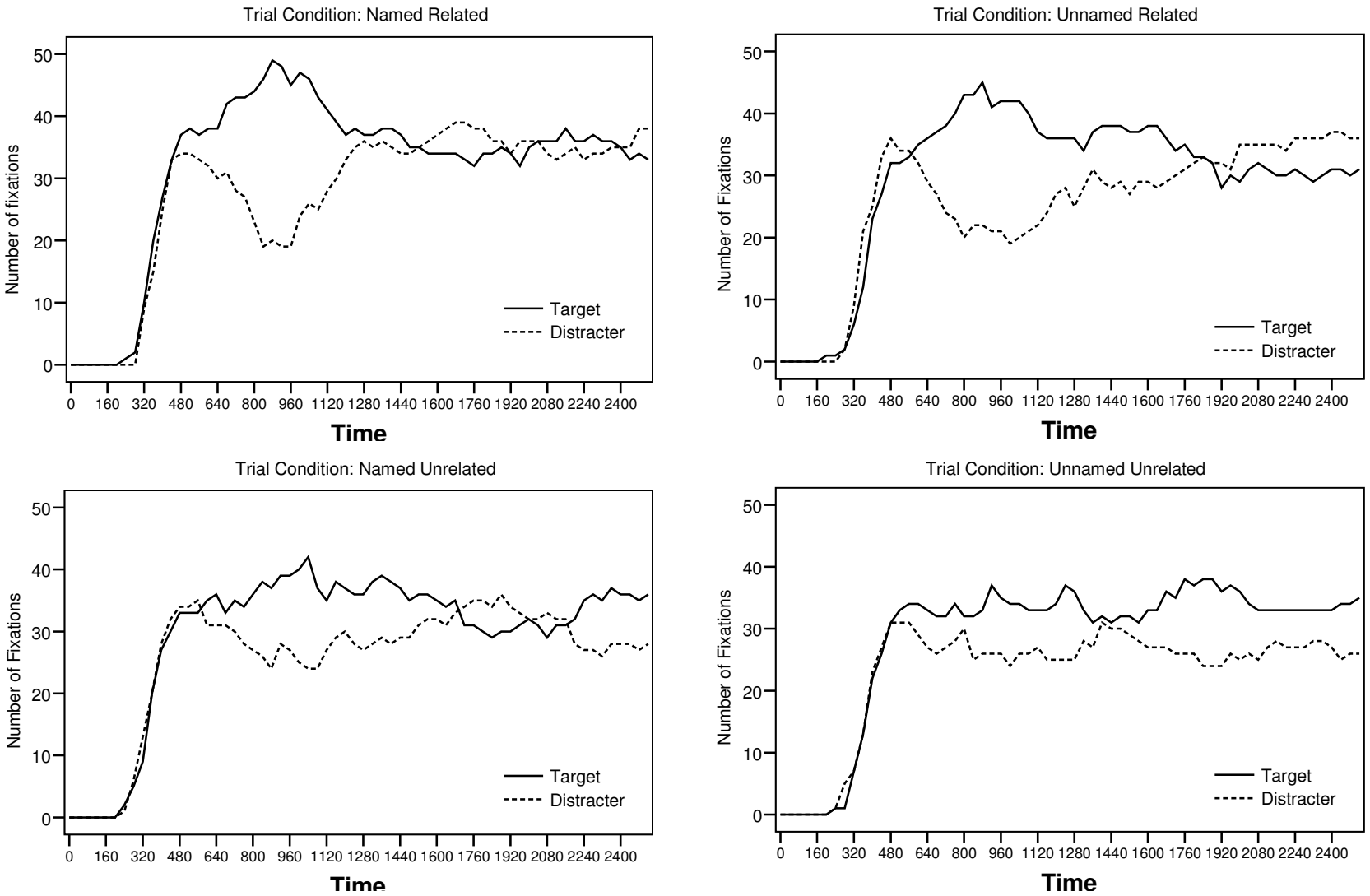


Figure 6.16. Target-First trials: Proportion of target fixations

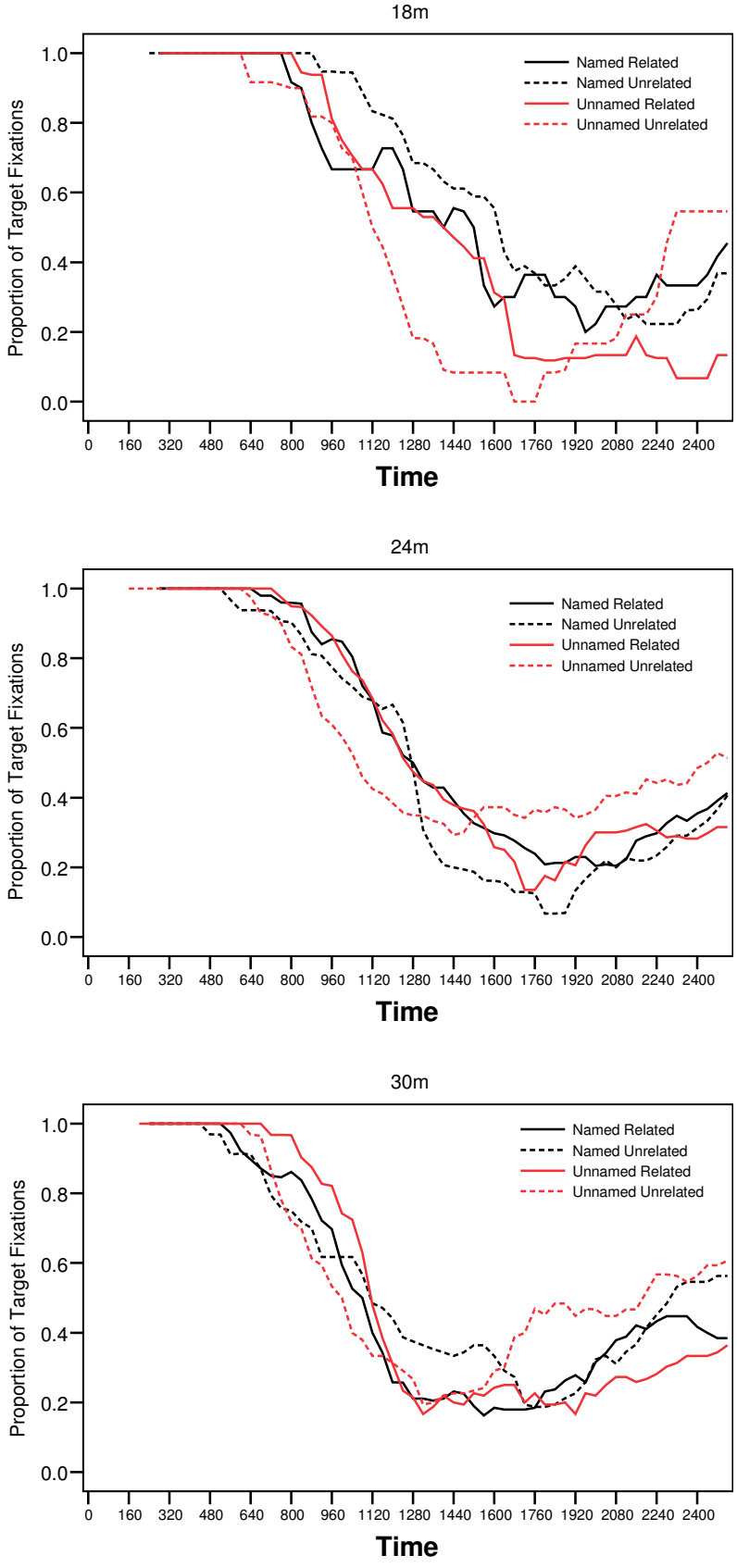
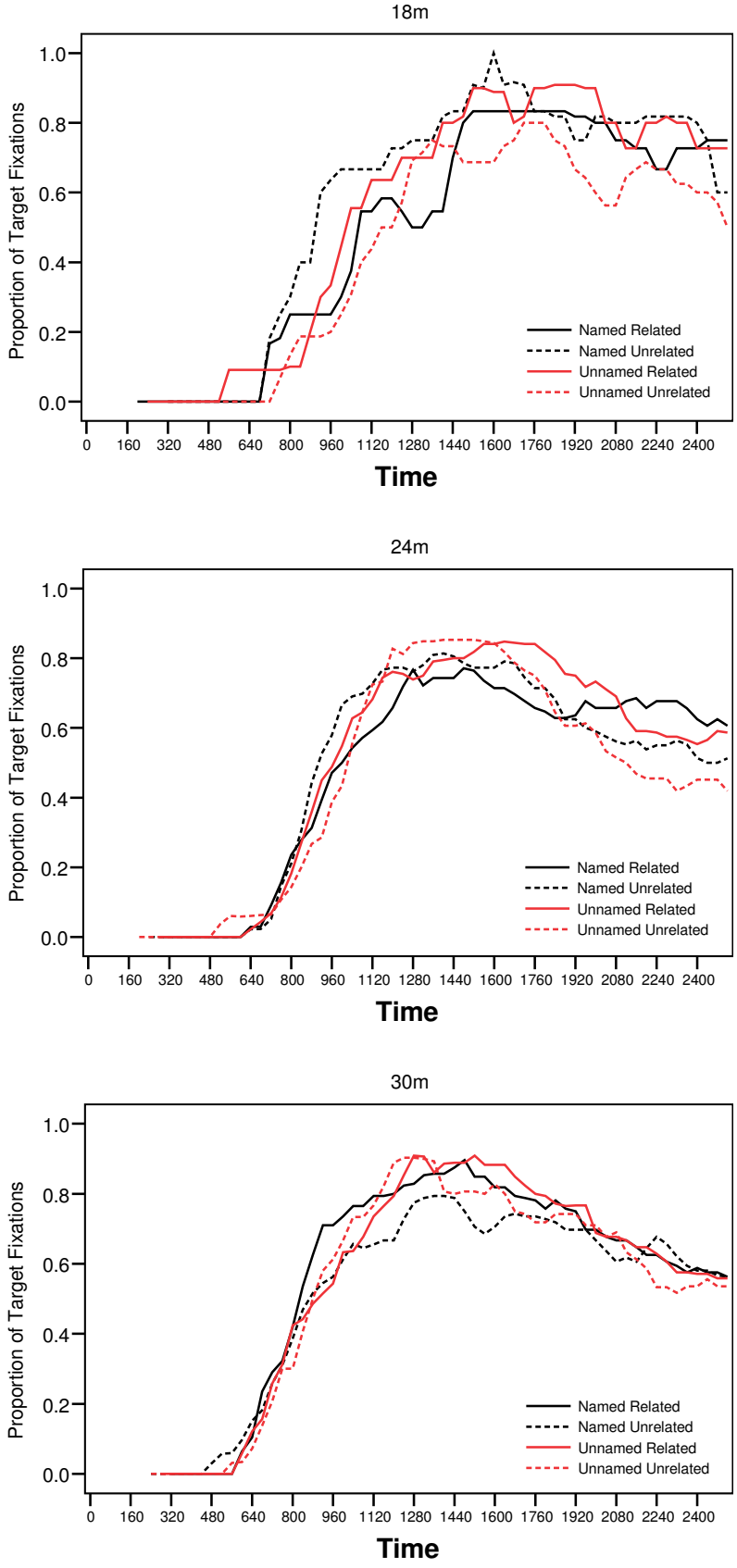


Figure 6.17. Distracter-First Trials: proportion of target fixations



General discussion

The purpose of the current experiment was to tease apart two processing mechanisms which might account for the priming effects observed in the quick-primed experiment (Chapter 5). The 'lexical facilitation' account (standard lexical priming) relied on linguistic processing of the target word for its effects to be seen, while the 'visual mediation' account could generate interest in the target without presentation of the target word. The current experiment introduced a naming manipulation, repeating the priming conditions presented in the original quick-primed experiment, with the target named in only half of the trials. This was intended to act as a test of the influence of the label, independent from the prime. It was predicted that the two named conditions would replicate the related and unrelated priming conditions from the previous experiment, while the third test condition (unnamed related) would add additional information about the mechanisms involved.

18-month-olds replicated the pattern of the quick-primed experiment in the macro-level measure of PTL, by responding to target naming only, with no effect of prime type. Older infants did not replicate the pattern for named trials in previous experiment. Instead, they showed more looking to the target than the distracter when either the prime, or the target, *or both words* were related to the target picture, with no difference in performance between the two named conditions. Their pattern of behaviour was not systematic when stimulus sub-list was included in analysis, and the generally low level of their target preferences meant that the difference between the control condition and the test conditions was not reliable.

The micro-level measure of percentage of first fixations to target gave the strongest evidence for one processing mechanism over another. While 18-month-olds showed unsystematic behaviour, older infants (24- and 30-month-olds), performed

above chance in the named related condition only. They demonstrated the ability to integrate para-foveal semantic information into high-speed decisions about which direction of fixate, but only when the target had been both named and primed by a related word. This finding is consistent with the lexical facilitation account of priming, in which the prime enhances the ease and speed of target processing via spreading activation prior to the onset of the target word. In the visual mediation account, feature matching from either the prime or the target word to the target picture would be delayed until the onset of the pictures. Thus any additive effect of prime and target visual matching would not be predicted to occur earlier than prime matching or target matching. This finding of speeded target recognition prior to initiating the first saccade strongly supports a lexical processing account of priming effects for two-year-olds.

Considering the two findings for older infants together, the micro-level measure (percentage of correct first fixations) shows a high-speed, on-line language processing effect, which results in greater accuracy in the very first fixation. The macro-level measure of PTL, on the other hand, represents a slower behavioural outcome, measured over two and a half seconds. In this measure, target preference is weak, and occurs in *any* condition where a relationship exists between a spoken word and a picture. These findings are consistent with automatic spreading activation in the first few hundred milliseconds of the trial, followed by a strategic response over the remainder of the trial. Possible reasons for the particular pattern of strategic response are discussed in the section which follows.

Discrepancies between experiments

The effect of the prime enhancing the percentage of correct first fixations was not observed in the quick-primed IPL experiment. The discrepancy may be due to the onscreen attention device, which increased the likelihood that infants were fixating a central portion of the screen at image onset. This also increases the likelihood of a small visual angle from the point of fixation to the target picture, improving the quality of the peripheral visual information. Review of videos suggests that a number of infants blinked at about the time the fixation device disappeared from the screen, and the duration of their blink may have effectively increased the stimulus onset asynchrony from the target word to the target picture. Such a delay may have improved general performance in the related prime condition. A third possibility is that the onscreen fixation device may have been moderately confused infants, thus generating a novel response pattern.

Older infants (24-month-olds in particular) failed to replicate the robust differences between priming conditions observed in the previous experiment. The differences were observed in the macro measures of PTL and LLK, as well as the micro-level measure of time-to-disengage. The two named conditions in the current experiment (named related, named unrelated) were directly equivalent to the two priming conditions in the previous experiment (related, unrelated). There are several possible reasons why the general effect of priming may not have emerged.

Loss of power. As there were less instances of each condition in the current experiment (three compared to six), the difference could have been due to a lack of power to observe differences between groups. However, the loss of power would predict similar means with larger variances. While the previous experiment reported a proportion of target looking of .61 in the related prime condition and .54 in the

unrelated condition, in the current experiment, both values were low (.55 in the related named condition, .54 in the unrelated named condition). While loss of power may have contributed to the failure to replicate, it may not be the primary source of the discrepancy.

Less strict visual stimulus controls. In the current experiment, a small number of words included as stimuli had originally been excluded from the stimuli set established in Chapter 2. These words had originally been judged as inappropriate stimulus items, as they were difficult to isolate from other features in their environment (e.g., eye from face, or moon from sky). The inclusion of these pictures, divorced from their usual context, and surrounded by fuzzy boundaries, may have generated additional noise in the looking pattern.

Onscreen attention device. In the previous quick-primed IPL experiment, some infants looked away from the screen during the priming phrase, and 3% of trials were lost to inattention. In the current experiment, with the onscreen attention device, less than 1% of trials were lost to inattention. However, review of videos indicated that while some infants appeared enjoy watching the attention device during the first 1,000ms of the trial, others appeared to be annoyed by it. Infants who were distracted by the animated shape may have been less attentive to the auditory stimuli, or less attentive during the test phase of the trial.

Time Pressure. In the current experiment, prime phrases were reduced to simple sentences (e.g., I saw a pig.), lasting on average 1213ms. The prime phrases in the previous experiment were longer (e.g., Yesterday, I saw a cat.), lasting an average of 2150ms. The shorter prime phrase reduced the amount of time infants had between test phases by almost one second. This effectively speeds up the pace of the task with more stimuli presented in a shorter time. Infants may have had greater difficulty

coping with the reduced time between trials, especially as the fast pace was coupled with the onscreen attention device.

Cue Validity. One consequence of the design of the current experiment is the variation in auditory cue validity. In the previous quick-primed IPL experiment, the target word was 100% predictive of the target image. In addition, the prime was related 50% of the time. That is to say, there was always one word with a relationship to the target, sometimes two. Auditory stimuli were thus highly reliable cues to the content of one of the pictures. By listening to the auditory cues, infants may have been able to infer the structure of the 'game' (experimental task) and could enjoy identifying 'matches' between named targets and unnamed distracters (many infants smiled, clapped, or pointed when they fixated the named target).

By contrast, in the current experiment, primes and targets had cue validities of only 50% each. This means that there were some trials in which no words were related to the target, and that the reliability of both cues was questionable. The unreliable nature of the cues makes it difficult to infer the purpose of the game, or to derive consistent satisfaction from 'playing along'. In short, this game was less fun to play. Post hoc review of video recordings supports this interpretation (compared to the quick-primed IPL experiment, there were more gestures and facial expressions were consistent with annoyance, e.g., shaking head at screen, looking away early, closing eyes). It is possible that older infants, aware of the unreliable nature of the auditory cues, employed a form of strategic responding which may have overshadowed systematic responses to the trial conditions.

Variation across stimulus sub-lists. The method of counterbalancing for differences between stimuli in the current experiment was to present all pictures as targets to one group of infants, and as distracters to another group. While this method

has been employed successfully previously, in combination with the trial types of the current experiment, it resulted in eight discrete stimulus sub-lists. Each sub-list contained only three targets, which differed systematically according to the trial conditions different infants observed them in.

Because all counterbalanced designs contain a confound between variance caused by items and variance caused by participants, it is important to establish whether the test conditions are artefacts of this mixed variance, or whether they are systematic effects which remain in spite of other sources of variance. Sub-list variation can derive from differences in the effectiveness of particular stimulus combinations (e.g., is spoon as good a prime for bowl as sheep is for cow?) or from differences in the individual populations who observed the items in different conditions (e.g., does Jane like to look at bananas as much as John does?). The inclusion of the sub-list in the analysis of a priming design is a fundamental control which allows the independent observation of the 'treatment effect' (e.g., prime condition), while acknowledging other sources of systematic variation (Raajmakers, 2003; Raajmakers et al., 1999). When the number of sub-lists is small and the items are well controlled, there is no reason for systematic effects of sub-list to interfere with the effect under assessment. Recall that in the quick-primed experiment (Chapter 5), the effect of sub-list was compatible with a robust priming effect for older infants. However, in the current experiment, the large number of sub-lists generated a large amount of unsystematic noise and complicated straightforward analysis.

Given the large number of minor manipulations between the two experiments, it is difficult to determine the precise cause of the failure to replicate the main priming effect from the previous experiment. The most likely causes of the difference between

experiments is due to low cue validity, in conjunction with the large amount of noise generated by the numerous stimulus combinations.

Further complications

In the Latin Square ordering of dummy targets across stimulus conditions, two disruptions of the standard stimulus arrangement occurred. In List B3, in the unrelated unnamed trial where the target picture was *eye* and the distracter *dog*, the dummy target 'Look!' was used, introducing a semantic confound between the acoustic dummy target and the target picture. This accounted for approximately 1% of trials, but may have influenced the looking pattern in stimulus sub-list 6, where it accounted for 25% of trials. In List B2, in the related unnamed condition where the target picture was *cow*, the dummy target was allocated as '*Wow!*' The phonological overlap of the dummy target with the target picture introduces a confound in the naming condition. This trial accounts for approximately 1% of all trials, but 25% of the trials in sub-list 8.

As in the previous chapter, three of the primes were preceded by the semantically restrictive verb *ate*. Unlike in the previous chapter, the verb was not counterbalanced across all conditions and sub-lists. The verb (which may have generated a boost to primes in the related condition) occurred in sub-lists two, three, four and seven, and twice in list eight. These trials account for 13% of all trials, and the effect of the semantically restrictive verb on the different sub-lists is unknown.

Interim conclusions

The current experiment sought to distinguish between lexical facilitation and visual mediation as possible accounts of the processing underlying the previously observed priming effect. In one measure, percentage of first fixations to target, older infants demonstrated a high speed ability to fixate named targets in their first look only when the target was both named and primed by a related word. As the effect of the prime was not evident without the target word, this finding provides evidence for the lexical facilitation account of primed language processing.

However, the same experiment failed to replicate the main findings of the original, that is, greater preference for named targets, and faster rejection of unnamed distracters, when preceded by a related prime. Given the lack of replication of the basic effect, and the pervasive complications of stimulus combinations, it is difficult to draw firm conclusions from the findings of the current experiment. Further research is needed to establish whether the effect observed here is reliable without the current stimulus combinations.

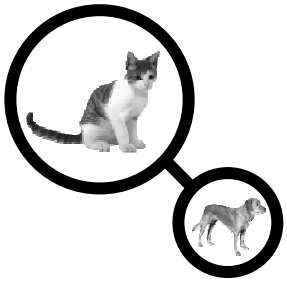
Given the large number of complications encountered in the current experiment, some recommendations for future quick-primed IPL research are proposed:

1. The effect of stimulus variations are analysed most effectively when the number of stimulus lists is kept small.
2. Counterbalancing interest levels within lists is effective, and preferable to using all items as targets and distracters.

3. Filler trials in which infants can perform a standard IPL task may help to raise the cue validity in a complex priming task, and reduce strategic responding.
4. An onscreen attention device during the priming phrase will lose less trials to inattention, but may increase task difficulty.
5. Short priming phrases are not necessarily better, as they may increase the task's time pressure.

Chapter 7

Summary & conclusions



The quick-primed IPL method established over the course of this thesis is a combination of the adult lexical priming paradigm, and the infant inter-modal preferential looking (IPL) task. It is a flexible method, able to assess the speed and accuracy of lexical access during spoken language comprehension. This method combines fine-grained temporal sensitivity with an intuitive, infant-friendly task. Following preliminary investigations into the size and composition of the lexicon in the second year (Chapter 2), this thesis documents the development of the experimental methodology. Over a series of experiments (Chapters 3-5), methodological parameters such as stimulus sequence, picture duration, prime-to-target inter-stimulus interval (ISI), and target-to-pictures stimulus onset asynchrony (SOA) were established. Using the new method, priming effects were observed in the language comprehension of two-year-olds (Chapters 5-6). While there were some complications in the interpretations of the findings, the processing mechanism which best described the observed effects, was a lexical account relying in spreading activation between connected words in the lexicon (Chapter 6).

This research was motivated by questions about semantic organisation in the lexicon during infancy, and the development of adult-like structures. Adult-like semantic organisation in the early stages of lexical development has been predicted by statistical models of lexicon growth (Hills et al., 2008b; Steyvers & Tenenbaum, 2005). The preferential attachment account arising out of these models describes how early acquired words become deeply embedded in the lexicon, offering a possible explanation for the relationship between age-of-acquisition and frequency effects in adult language processing (Morrison & Ellis, 1995; Morrison et al., 1992). However, network models based on the words acquired by 16- to 30-month-olds suggest more than one organisational structure is possible during infancy (Hills et al., 2008a; Hills

et al., 2008b), leading to the possibility that re-organisation occurs with age. Experimental data were therefore required to establish *when relationships develop* in the infant lexicon, and *which kinds of relationship* provide organisation in the developing semantic system.

Behavioural research into the semantic systems of children (Betjemann & Keenan, 2008; Duncan & Kellas, 1983; Hartfield & Conture, 2006; Hashimoto et al., 2007; McCauley et al., 1976; Nation & Snowling, 1999; Radeau, 1983; Schvaneveldt et al., 1977) and toddlers (Krackow & Gordon, 1998) were shown to have limited applicability in infancy, due to the demands of the tasks employed: Lexical priming techniques used with adults typically rely on higher-order meta-linguistic knowledge (e.g., Meyer & Schvaneveldt, 1976), while simplified versions of the task (e.g., primed picture naming, object decisions, and verbal memory tasks) require verbal production skills at a minimum. Given these methodological limitations, relationships within the infant lexicon remained untested until the current programme of research.

The robust priming effects observed in the quick-primed experiment (Chapter 5) were the first to suggest that adult-like structure could be observed in the lexicons of children under the age of three. However, a complication remained in the interpretation of the priming effect – given that the design included both auditory primes and visual representations of objects, it was unclear whether the mechanism was truly 'lexical' or whether it might rely on some other (non-linguistic) mechanism. The follow up study (Chapter 6) attempted to clarify whether the two-year olds' priming effect was lexical in origin, or relied on visual matching from the prime to the target picture. The results of the study suggested a lexical interpretation, but these results were complicated by a failure to replicate the general pattern of priming

observed in the original experiment. Strategic responding by older infants was thought to have played a role in the confusing pattern of results.

Additional support for the lexical account of the priming effect comes from a recent study, reported by Styles, Arias-Trejo and Plunkett (2008). The study (Experiment 3, conducted by the second author of the paper) was almost identical in design to the experiment reported in Chapter 6 (including that targets were only named in half of the trials), but employed three methodological manipulations. First, an intermediate age group (21 months) was targeted, as strategic responding may not yet have developed in this group. Second, to increase the automaticity of responding, target and distracter images were selected to share phonological onsets, making them cohort-competitors during the early stages of lexical access. Third, to maximize the likelihood of phonological interference, the shorter target-to-picture SOA was selected from the two SOAs (400ms, 200ms) established in the quick-primed experiment (Chapter 5).

As in experiments reported in this thesis, 18-month-olds in Styles, Arias-Trejo and Plunkett's Experiment 3 (2008) responded to target naming, but were unaffected by lexical primes presented at short latencies. By contrast, 21-month-olds showed robust target recognition *only* in the condition where the target was both named and primed by a related word. Unlike in the half-target experiment reported here (Chapter 6), the magnitude of the effect was clear in standard, macro-level measures, and the difference between conditions was not complicated by stimulus combinations or strategic responding. The combination of short SOA with a phonological onset-competitor meant that infants had only begun to process the target word when two equally valid images appeared. To identify the correct target, infants needed to *continue processing* the target word, while inhibiting responses to a cohort

competitor. In this case, 21-month-olds only recognised the target in the related-named condition. Thus phonological interference from the distracter picture was only overcome with support from a related prime. Moreover, 21-month-olds did not show preference for primed, but unnamed targets (as would be predicted by a visual mediation account). This pattern of general inhibition coupled with primed processing facilitation is only compatible with a lexical, or word-word priming account (recall that the 'visual mediation' account predicted target preference even when the prime was related and the target unnamed).

Thus, when the results of this recent study are combined with the experimental groundwork laid out in this thesis, the quick-primed IPL method can be taken to demonstrate priming effects at the lexical level of representation, making it a valuable new method for investigating word-level semantic relationships in the developing lexicon.

Principle experimental findings

Preliminary studies

The composition of the lexicon at 18 months-of-age was investigated in Chapter 1, using 178 Oxford CDIs collected from infants a four week age range. This is the largest reported study of such a narrow age window: This study incorporates data from more than twice the number of 18-month-olds previously reported in large-scale CDI studies (Fenson et al., 1994; Hamilton et al., 2000). In addition, it is the first large-scale CDI study to go beyond total comprehension and production scores, and to investigate the components making up the infant lexicon, along with their relations to other lexical variables.

Three sets of norms were created: norms for the total comprehension and production scores of individuals (Table 2.2); norms for the comprehension rate of each word in the Oxford CDI (Appendices B & C); and a collation of adult word associations between nouns in the 18-month-old lexicon (Appendix D). These norms were used as a baseline for all other studies in the thesis: Low vocabulary infants were identified and removed from analysis; comprehension rates were used to establish a pool of nouns appropriate for use as experimental stimuli (Table 2.9); and adult word associations were manipulated in the primed experimental designs (Chapters 3, 5 and 6). The clear numeric dominance of nouns in the 18-month-old vocabulary was taken as good reason to limit the current series of investigations to nouns.

In addition, the norming process allowed investigation of the internal detail of British infants' lexicons at 18 months-of-age. Individuals' comprehension and production scores were consistent with previously reported norms, varying by only a few percentage points at each percentile (Fenson et al., 1994; Hamilton et al., 2000). As demonstrated elsewhere (Fenson et al., 1994; Hamilton et al., 2000), individuals' comprehension rates correlated with their production rates, and comprehension was consistently ahead of production. The previously reported trend for North American production scores to be higher than British production scores (Hamilton et al., 2000) was replicated, however the production advantage for North American infants was restricted to certain word types (e.g., action words, adjectives, pronouns). This suggested that some of the North American advantage may have been generated by the population-specific development of the North American tool.

The norms for the comprehension rate of individual words are the first reported lexical norms for British English. The British comprehension rates for individual words correlated with the North American comprehension rates (Philip S

Dale & Fenson, 1996). Comprehension rates for nouns were also compared with other lexical variables known to influence adult lexical access: noun comprehension rate was shown to correlate with both age-of-acquisition and adult word frequency. These findings suggest that where other measures are unavailable (i.e., untested) comprehension rate could act as a proxy measure for either age-of-acquisition or lexical frequency in infant language studies.

A recent paper by Goodman, Dale and Li (2008) adds an interesting counterpoint to these data. Their study compared age-of-acquisition derived from production and comprehension rates in the LEX database (Philip S Dale & Fenson, 1996) with child directed frequency counts in the CHILDES data set (MacWhinney, 2000). Across all words in the MacArthur-Bates CDI (Fenson et al., 1994), they found that *only nouns* showed a correlation between age-of-acquisition in comprehension and the typical frequency in parental speech. For age-of-acquisition in production, all word classes showed a within-class frequency effect. Their findings corroborate the noun pattern reported here, with spoken frequency counts more relevant to the population of interest.

The first experimental study in this series (Chapter 2) was initially planned as a long-range priming study. While no effects of priming were observed, the main contribution of this study was the finding that parental report of CDI comprehension accurately predicted the words for which 18-month-olds demonstrated lexical comprehension – that is, *untrained parents know which words their infants understand*. This finding provides external validation of the item-level accuracy of both the CDI tool and the IPL method, and counters the claims of previous researchers that parental report is inaccurate at the item level (Houston-Price et al., 2007; Tomasello & Mervis, 1994). Furthermore, this experiment provides evidence

about the *level of comprehension* at which parents make their CDI assessments: British parents interpret 'understanding' as the ability to pick out a named item from a member of the same category when the task is fairly difficult (unfamiliar items, presented only once in an unfamiliar environment).

One critical application of this finding is that it validates the use of comprehension scores in the CDI for item-level control of individuals' word knowledge at the time of test. Furthermore, given that British parents set a high threshold for lexical comprehension, the use of this data for lexical exclusion can be seen as conservative (excluding some words which may be known, but ensuring that only known words are assessed in the priming context).

Another important finding of this minimally primed experiment was that infants were able to demonstrate age appropriate lexical comprehension abilities even when the regular task was preceded by a priming phrase, presented in darkness. This was an important finding, as it indicated that infants were not confused by the non sequitur, and that priming phrases could be employed in other adaptations of the IPL paradigm without the need for masking or visual accompaniment.

As no priming effects were observed in the minimally primed adaptation of the standard IPL design, the words-before-pictures experiment (Chapter 4) was used to pilot a truncated version of the IPL task, in which the confounding pre-naming baseline was removed. This experiment was conducted to ensure that lexical comprehension was observable in the 'un-baselined' version of the task, and to pilot stimulus order, timing intervals, and appropriate measures for analysis. Short and long picture-presentation durations were trialled. The target-to-test stimulus SOA of approximately 400ms was found to be appropriate, and lexical comprehension was observed in the looking behaviour of both age groups.

The short picture presentation duration (1,500ms) was shown to be too short for 18-month-olds to demonstrate reliable comprehension in some measures, and was also too short for appropriate analysis of the time-to-disengage measure. The long picture presentation duration (5,000ms) appeared to reduce the efficacy of some measures of target discrimination in older infants. An intermediate picture duration of 2,500ms was proposed for future studies. In addition to these primary findings, investigation of the looking pattern across different items led to the observation that macro-level measures of looking behaviour appeared to be more heavily influenced by the salience of different pictures, while the micro-level measures were less affected. This suggested that the micro-level measures (time-to-fixate and time-to-disengage) might be more reliable in un-baselined tasks.

An interesting side product of the words-before-pictures experiment (Chapter 4), was the finding that infants indexed the match between a spoken word and a picture *before* they had landed their first fixation. Although the accuracy of directing the first fixation was at chance, differences in the speed of the first fixation could only be attributed to semantic information about the identity of the peripheral images being integrated with the incoming auditory string. The cost of processing a match between the picture and the word was a slowdown in time-to-fixate. This high-speed peripheral visual ability has not been reported previously for infants, although it is consistent with findings from the adult literature (Morgan & Meyer, 2005; Thorpe et al., 2001; Vucetic, 2008). This effect was likely to have been facilitated by the small angular gap between pictures in the current screen configuration.

Primed studies

Following the methodological parameters developed in the preliminary studies, the primed studies (Chapters 5 & 6) introduced a lexical priming phrase to the words-before-pictures stimulus organisation. In the tightly controlled quick-primed experiment (Chapter 5), age-appropriate lexical comprehension was demonstrated for named items, along with robust priming effects for older infants. These effects were evident both for micro- and macro-levels of measurement, indicating that the auditory prime influenced the ease and speed of target recognition as well as the overall interest in the target picture. Systematic stimulus effects were also uncovered in this experiment, but they did not interfere with the basic effect of the priming condition on looking behaviour.

In the half-target experiment (Chapter 6), the complexity of the design and the un-intuitive nature of the 'game' led to a failure to replicate the main priming effect of the previous study. Eighteen-month-olds demonstrated their usual ability to identify named targets, but older infants showed a more complex pattern of results. Surprisingly, the measure of percentage of correct first fixations showed a priming effect, which, although weak, could only be explained via a lexical facilitation account: Infants were only able to direct their first look towards the target with greater-than-chance accuracy in those cases where the target was named and the prime was related. This finding was especially salient given that infants had *not* shown above-chance performance in the direction of first fixations in *any of the preceding experiments*. This finding therefore suggests facilitation above the level of accuracy normally available to two-year-olds in the very earliest stages of the trial.

Developmental trajectory

The developmental trajectory suggested by this series of studies (with additional evidence provided by Styles Arias-Trejo and Plunkett (2008)) shows that priming effects emerge between 18 and 21 months-of-age. This could be taken to indicate that lexical relationships develop between these ages. However it is also possible that experimental features of the current task are unsuitable for observing priming effects in 18-month-olds. For example, prime-to-target spreading activation may take longer for younger infants (thus younger infants may require longer ISIs), or memory capacity may limit younger infants processing of phonological information to only the most recent item (hence, their performing as though no prime was presented). Further manipulations of the method would be required to ascertain whether a set of experimental parameters more appropriate to 18-month-olds can be developed.

Limitations of the current work

The current body of work is primarily methodological. Although it was inspired by questions about the detailed organisation of the infant lexicon, the bulk of the work reported here outlines the step-by-step process of developing a new experimental method. Although a promising new method has been developed, the main theoretical contributions of the method have yet to be seen. While an experimental effect consistent with spreading activation in the lexicon has been observed in the looking behaviour of two-year-olds, and the lexical nature of the effect gains further support from external evidence (Styles et al., 2008), but further investigations will be required in order to answer the theoretical questions posed in the introduction: What kinds of relationships are evident in the infant lexicon? Which

ones are most strongly integrated into infant semantic systems? And does the organisation of the developing lexicon change over time?

One complication encountered during the experimental work was the feature of stimulus variability. It is standard practice in lexical priming studies to ensure that individuals do not encounter a test item on more than one occasion, to avoid confounding lexical priming with other memory effects or short term dependencies (McNamara, 2005). When working in infancy, however, the small vocabularies of participants means that only a small number of unique trials can be constructed. Given the lack of individual baselines in the primed task, stringent stimulus controls must be employed to balance out stimulus effects. In the quick-prime experiment, the chosen approach produced a stimulus pool where the effects of stimulus variation were large, but well controlled, and did not generate noise in the experimental design. In the half-target experiment (Chapter 6) on the other hand, a different method of stimulus control was selected, and the results were much harder to quantify. Taken together, the designs of these experiments provide guidelines for future research of this kind, which can be summarised as follows: Approximate rankings of picture salience based on animacy and other salient features systematically predict looking behaviour in the second year; Counterbalancing approximate interest levels within and across stimulus lists is preferable to counterbalancing target and distracter status in separate lists; Counterbalancing approximate interest levels within test conditions (as well as stimulus sub-lists) is likely to produce more balanced stimulus sub-lists, and to simplify analysis further. Too many stimulus sub-conditions not only complicates analysis but introduces unnecessary noise to already noisy infant data. With these guidelines in mind, simple and effective stimulus combinations are no doubt

achievable, so long as they are incorporated into the earliest stages of experimental design.

An additional limitation of the current body of work is that it deals primarily with nouns. The decision to focus on nouns was driven by dramatic noun bias observed in the CDI norming study (Chapter 1). As the experimental work was designed to be a proof-of-concept, it was considered necessary to select the largest word class in the infant lexicon from which to draw stimuli. Although this approach is consistent with the majority of adult priming studies, it also means that the finding of lexical level relationships can only be extrapolated to other items of the same word class.

Future directions

Having established an infant friendly priming method with high temporal sensitivity, straightforward manipulations in the selection of stimuli could be used to examine the detailed structure of the infant lexicon.

Examining different types of lexical relationship in infancy

The priming effects observed in the current series of experiments were evident when prime-target pairs were related both associatively and taxonomically. This kind of relationship was selected with the intention of maximising possible priming effects, given the 'associative boost' observed in adult priming studies (Moss et al., 1997). As it was initially unclear whether *any* priming effects would be observed in the novel method, the selection of the strongest possible priming relationship was well founded. The original intention, once the basic method had been established, was to use primed

IPL to investigate different types of lexical relationship. However, given the large number of subjects and time consuming nature of the methodological development, the methodological approach has yet to be applied to different kinds of relationship which may exist early in the developing lexicon. This means that a variety of interesting questions have yet to be addressed.

Association versus taxonomy. The modelling of Styvers and Tenebaum (2005) predicts that associative relationships will be evident early in lexicon development. The selection of the strongest possible priming relationship in the current series of experiments meant that any discrete effects of taxonomy or association were confounded. Thus it is impossible to tell whether the observed priming effects reflect taxonomic organisation, associative organisation, or both, in the lexicons of two-year-olds. Further investigation could also reveal whether association and taxonomy have different time-courses in the spread of activation through the infant lexicon, as has been demonstrated for adults (Moss et al., 1997).

Taxonomy versus theme. As discussed in Chapter 2, the word associations typically produced by children tend to differ from those of adults, with taxonomic responding increasing with age (Brown & Berko, 1960; D. Entwisle, 1966; Ervin, 1961; Francis, 1972; MacNeil, 1966; Woodrow & Lowell, 1916). It is therefore possible that non-taxonomic relationships play a greater role than classical semantic knowledge in the semantic systems of infants. Further supporting evidence comes from the fact that school-aged children are better able to explain non-taxonomic relationships (Hashimoto et al., 2007), and from processing differences between

taxonomic and thematic relationships in school-aged children (Nation & Snowling, 1999), albeit in a population with a developmental deficit.

A logical extension of the quick-primed IPL method would therefore employ primes with different kinds of semantic relationship to the target. A primed design comparing thematic primes with taxonomic primes could readily be achieved with or without association, so long as the trial-by-trial cue validity remained high. One limitation to studies of this kind is that there are very few attested adult associations between thematically related nouns in the infant lexicon. One might expect that child word association norms would provide more associations from which to select stimulus pairs, but preliminary reviews conducted in the course of this research suggested that it is typically the *proportion of responses which are non-taxonomic* which varies between childhood and adulthood, rather than the *number of non-taxonomic relationships* apparent between words. This consideration would not be problematic if association were excluded from the design.

In an alternative approach, thematic verb-noun relationships could be employed in the quick-primed IPL design, exploiting the fact that the acoustic prime need not be depicted visually. However, given that a verb appearing shortly before a noun (e.g., *kick – ball*) resembles a grammatical sentence fragment, primes of this type may envoke quite different processing mechanisms. Thus direct comparison of verb-noun primes with noun-noun primes should be undertaken with caution.

Semantic feature norms. Shared semantic features may also be a source of semantic organisation in the infant lexicon. Given McRae, de Sa and Seidenberg's (1997) demonstration of semantic feature priming in adults, it would be reasonable to predict that semantic features might also show priming effects in infancy. Hills et al.'s

(2008b) comparison of two infant lexical networks modelled on semantic feature norms or associative norms makes the comparison of semantic features with association especially interesting. As raised in the introduction (Chapter 1), it is possible that both systems exist in tandem, but differences in the time-course of spreading activation may be evident. Primed investigations of semantic features might also allow comparisons of whether infants show organisation according to visual feature overlap (e.g., *bird – aeroplane*) or to functional feature overlap (e.g. *bus – aeroplane*).

Phonology. An adaptation of the primed IPL method has recently been used to investigate how activation flows between words at the phonological level of representation in the infant lexicon. In this work by Mani and Plunkett (2008) infants saw a picture of a familiar object, prior to the presentation of a pair of pictures. The target label was presented shortly after the onset of the picture pair. 18-month-olds and 24-month-olds recognised target words faster when the prime's (unspoken) label was phonologically related to the target, than they did when the prime was unrelated. In addition, they found that the size of the phonological neighbourhood influenced reaction times for 24-month-olds. Their finding demonstrates adult-like phonological organisation of words in the lexicons of children by two years of age, providing further evidence that infant lexicons contain adult-like features by age two. Further research into the different levels of representation could employ the auditory quick-primed IPL method established here to make comparisons between the time course of phonological and semantic activation in the lexicon.

Onomatopeoia. The CDI norming study (Chapter 2) identified an unexpected feature of the lexicons of 18-month-olds – the tight relationship between the comprehension rate and the production rate for words in the 'animal sounds' section of the Oxford CDI. The special status of onomatopoeia in the infant lexicon makes it an interesting avenue for investigation, especially as these semantic associates do not rely on taxonomy. Onomatopoeic primes could be compared across related (e.g., *woof – dog*) and unrelated word pairs (e.g., *quack – dog*). While certain controls might be needed to ensure that lexicalised relationships (e.g., *moo-cow*) were not included in the design, these kinds of relationships might be embedded in the semantic structures of infants at even younger ages. Recall that these words had the highest median comprehension rate of all of the word types in the O-CDI (all onomatopoeia words were understood by at more than 50% of 18-month-olds), meaning that infant-by-infant lexical exclusions could be minimised.

Further methodological adaptations

The adaptations of the method described above all rely on the simple substitution of different word types into the basic trial structure of the quick-primed IPL task. More broad-ranging future directions for infant priming techniques could include tasks which mask the explicit pairing of primes and targets (e.g., by incorporating a feedback phase at the end of the trial (Killing & Bishop, 2008), thus making every target a prime for the following trial); using priming in a train-test paradigm to investigate how quickly newly-learned words become integrated into the infant lexicon (as per Dumay & Gaskell, 2006); or incorporating pupil dilation as a measure of violation of expectation (as per Sirois & Jackson, 2007) during primed language comprehension.

Broader applications

Two further applications of the quick-primed IPL method may also have value in the future: clinical assessment of individual differences in infancy, and clinical assessment of adult populations with special needs. Firstly, quick-primed IPL may be able to identify children at risk of reading or learning difficulties prior to the start of formal education. Two populations of school-aged children have been shown to exhibit atypical responses to semantic priming tasks: Children with poor reading comprehension (Nation & Snowling, 1999), and children with reading disability (Betjemann & Keenan, 2008). If lexicon organisation is indeed continuous throughout development, as predicted by Steyvers & Tenenbaum (2005), it is possible that abnormal performance in a priming task during infancy could be a marker for later difficulties in reading development. If this were the case, interventions based of supporting semantic *relationships* may be particularly advantageous for at-risk children while the lexicon is developing. In addition, the implicit nature of the quick-primed IPL task (a task without complex instructions, requiring only eye movements as a response), might be particularly useful in clinical assessments of adult patients with special needs, for example, left hemisphere stroke patients with language deficits. Clearly, the novel method established in the current thesis has wide-spread applicability to psycholinguistic inquiry.

References

Amso, D., & Johnson, S. P. (2005). Selection and inhibition in infancy: Evidence from the spatial negative priming paradigm. *Cognition, 95*, B27-B36.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22*, 261-295.

Antos, S. J. (1979). Processing facilitation in a lexical decision task. *Journal of Experimental Psychology: Human Perception and Performance, 5*(3), 527-545.

Aslin, R. (2007). What's in a look? *Developmental Science, 10*(1), 48-53.

Bailey, T. M., & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development, 17*(2), 1265–1282.

Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of Child Language, 6*, 183-200.

Betjemann, R. S., & Keenan, J. M. (2008). Phonological and semantic priming in children with reading disability. *Child Development, 79*(4), 1086-1102.

Bloom, L. (1973). *One word at a time: The use of single word utterances before syntax*. The Hague: Mouton.

Bloom, P. (2002). *How children learn the meanings of words*. Cambridge, MA & London, UK: MIT Press.

Boulenger, V., Décoppet, N., Roy, A. C., Paulignan, Y., & Nazir, T. (2007). Differential effects of age-of-acquisition for concrete nouns and action verbs: Evidence for partly discrete representations? *Cognition, 10*, 131-146.

Bremner, A. J., Mareschal, D., Destrebecqz, A., & Cleermans, A. (2007). Cognitive control of sequential knowledge at 2 years of age: Evidence from an incidental sequence learning and generation task. *Psychological Science, 18*, 261-266.

Bretherton, I., McNew, S., Snyder, L., & Bates, E. (1983). Individual differences at 20 months: Analytic and holistic strategies in language acquisition. *Journal of Child Language, 10*, 293-320.

The british national corpus, version 3. (2007). Distributed by Oxford University Computing Services on behalf of the BNC Consortium.

Brown, R., & Berko, J. (1960). Word association and the acquisition of grammar. *Child Development, 31*, 1-14.

Brown, R., & McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior, 5*(4), 325-337.

Callaghan, T. C. (2000). Factors affecting children's graphic symbol use in the third year: Language, similarity, and iconicity. *Cognitive Development, 15*, 185.

Canfield, R. L., Smith, E. G., Brezsnyak, M. P., & Snow, K. L. (1997). Information processing through the first year of life: A longitudinal study using the visual expectation paradigm. *Monographs of the Society for Research in Child Development, 62*(2), 1-170.

Catling, J. C., & Johnston, R. A. (2006). Effects of age of acquisition and priming on picture naming. *The Quarterly Journal of Experimental Psychology, 59*(8), 1443-1453.

Chamberlain, A. F., & Chamberlain, I. C. (1904). Studies of a child. *Pedagogical Seminary, 11*, 264-291, 452-483.

Chapman, L. J., Chapman, J. P., Curran, T. E., & Miller, M. B. (1994). Do children and the elderly show heightened semantic priming? How to answer the question. *Developmental Review*, 159-185.

Clark, E. (1973). What's in a word? On the child's acquisition of semantics in his first language. In T. E. Moore (Ed.), *Cognitive development, and the acquisition of language*. New York: Academic Press.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407-428.

Dale, P. S., Bates, E., Reznick, J. S., & Morisset, C. (1989). The validity of a parent report instrument of child language at twenty months. *Journal of Child Language, 16*, 239-249.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers, 28*(1), 125-127.

Darwin, C. (1877). A bibliographical sketch of an infant. *Mind: A Quarterly Review of Psychology and Philosophy, 7*(3), 285-294.

Deese, J. (1962). Form class and the determinants of word association. *Journal of Verbal Learning and Verbal Behavior, 1*, 79-84.

Dumay, N., & Gaskell, G. (2006). Sleep-associated changes in the mental representation of spoken words. *Psychological Science, 18*(1), 35-39.

Duncan, E. M., & Kellas, G. (1983). Developmental changes in the internal structure of semantic categories. *Journal of Experimental Child Psychology, 26*, 328-340.

Entwisle, D. (1966). *Word associations of young children*. Baltimore, Maryland: The Johns Hopkins University Press.

Entwisle, D. R., Forsyth, D. F., & Muuss, R. (1964). The syntactic-paradigmatic shift in children's word associations. *Journal of Verbal Learning and Verbal Behavior, 3*(19-27), 19-29.

Ervin, S. M. (1961). Changes with age in the verbal determinants of word-association. *American Journal of Psychology, 74*, 361-372.

Fantz, R. L. (1963). Pattern vison in the newborn infant. *Science, 140*(3564), 296-297.

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language, 41*, 469-495.

Feldman, H. M., Dollaghan, C. A., Campbell, T. F., Kurs-Lasky, M., Janosky, J. E., & Paradise, J. L. (2000). Measurement properties of the macarthur communicative development inventories at ages one and two years. *Child Development, 71*(2), 310-322.

Fenson, L., Bates, E., Dale, P. S., Goodman, J., Reznick, S. J., & Thal, D. J. (2000). Measuring variability in early child language: Don't shoot the messenger. *Child Development, 71*(2), 232-328.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5), 1-173.

Fernald, A. (2005). Young children's inferential use of verb information in learning new objects, *Society for Research in Child Development Biennial Meeting*. Atlanta, Georgia: Kevin C Rose.

Fernald, A., & Hurtado, N. (2006). Names in frames: Infants interpret words in sentence frames faster than words in isolation. *Developmental Science, 9*(3), F33-F40(31).

Fernald, A., Perfors, A., & Marchman, V. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the second year. *Developmental Psychology, 42*(1), 98-116.

Fernald, A., Pinto, J. P., Weinberg, A., & McRoberts, G. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science, 9*(3), 228-231.

Fernald, A., Swingley, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development, 72*(4), 1003-1015.

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. M. Fernandez & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 97-135). Amsterdam: John Benjamins.

Francis, H. (1972). Toward an explanation of the syntagmatic-paradigmatic shift. *Child Development, 43*(949-958).

Frick, J. E., & Colombo, J. (1996). Individual differences in infant visual attention: Recognition of degraded visual forms by four-month-olds. *Child Development, 67*, 188-204.

Frick, J. E., Colombo, J., & Saxon, T. F. (1999). Individual and developmental differences in disengagement of fixation in early infancy. *Child Development, 70*(3), 537-548.

Friedrich, M., & Friederici, A. D. (2004). N400-like semantic incongruity effect in 19-month-olds: Processing known words in picture contexts. *Journal of Cognitive Neuroscience, 16*(8), 1465-1477.

Friedrich, M., & Friederici, A. D. (2005a). Lexical priming and semantic integration reflected in the event-related potential of 14-month-olds. *Cognitive Neuroscience and Neuropsychology, 16*(6), 653-656.

Friedrich, M., & Friederici, A. D. (2005b). Semantic sentence processing reflected in the event-related potentials of one- and two-year-old children. *Neuroreport, 16*(16), 1801-1804.

Funnell, E., Hughes, D., & Woodcock, J. (2006). Age of acquisition for naming and knowing: A new hypothesis. *The Quarterly Journal of Experimental Psychology, 59*(2), 268-295.

Ganger, J., & Brent, M. R. (2004). Reexamining the vocabulary spurt. *Developmental Psychology, 40*(4), 621-632.

Goldfield, B., & Reznick, J. S. (1990). Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of Child Language, 17*, 171-183.

Goldin-Meadow, S., Seligman, M. P., & Gelman, R. (1976). Language in the two-year old. *Cognition, 4*, 189-202.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language, 14*, 23-45.

Golinkoff, R. M., Mervis, C. B., & Hirsch-Pasek, K. (1994). Early object labels: The case for a developmental lexicals principles framework. *Journal of Child Language, 21*(1), 125-155.

Goodman, J., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language, 35*, 515-531.

Haith, M. M., Hazan, C., & Goodman, G. S. (1988). Expectation and anticipation of dynamic visual events by 3.5-month-old babies. *Child Development, 59*(2), 457-479.

Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a british communicative development inventory. *Journal of Child Language, 27*, 689-705.

Harris, M., Yeeles, C., Chasin, J., & Oakley, Y. (1995). Symmetries and asymmetries in early lexical comprehension. *Journal of Child Language, 22*, 1-18.

Hartfield, K. N., & Conture, E. G. (2006). Effects of perceptual and conceptual similarity in lexical priming of young children who stutter. *Journal of Fluency Disorders, 31*, 303-324.

Hashimoto, N., McGregor, K. K., & Graham, A. (2007). Conceptual organisation at 6 and 8 years of age: Evidence from the semantic priming of object decisions. *Journal of Speech, Language, and Hearing Research, 50*, 161-176.

Heinze, H.-J., Muente, T.-F., & Kutas, M. (1998). Context effects in a category verification task as assessed by event-related brain potential (erp) measures. *Biological Psychology, 47*(2), 121-135.

Hills, T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. B. (2008a). Categorical structure in early semantic networks of nouns. In B. C. Love, K. McRae & V. M. Sloutsky (Eds.), *Proceedings of the 30th annual conference of the cognitive science society* (pp. 161-166). Washington, DC: Cognitive Science Society.

Hills, T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. B. (2008b). Is there preferential attachment in the growth of early semantic networks. In B. C. Love, K. McRae & V. M. Sloutsky (Eds.), *Proceedings of the 30th annual conference of the cognitive science society* (pp. 155-160). Washington, DC: Cognitive Science Society.

Houston-Price, C., Mather, E., & Sakkalou, E. (2007). Discrepancy between parental report of infants' receptive vocabulary and infants' behaviour in a preferential looking task. *Journal of Child Language, 34*(4), 701-724.

Houston-Price, C., Plunkett, K., & Harris, P. (2005). Word-learning wizardry at 1;6. *Journal of Child Language, 32*, 175-189.

Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language, 57*, 460-482.

Huttenlocher, J. (1974). The origins of language comprehension. In R. L. Solso (Ed.), *Theories in cognitive psychology: The loyla symposium*. New York: Earlbaum.

Jenkins, J. J., & Russell, W. A. (1960). Systematic changes in word association norms. *Journal of Abnormal and Social Psychology, 60*(3), 293-304.

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2001). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 40*, 133-156.

Killing, S. E. A., & Bishop, D. V. M. (2008). Move it! Visual feedback enhances validity of preferential looking as a measure of individual differences in vocabulary in toddlers. *Developmental Science, 11*(4), 525-530.

Koff, R. (1963). Systematic changes in children's word-association norms 1916-63. *Child Development, 36*(1), 299-305.

Krackow, E., & Gordon, P. (1998). Are lions and tigers substitutes or associates? Evidence against the slot filler accounts of children's early categorization. *Child Development, 69*(2), 347-354.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day american english*. Providence, RI: Brown University Press.

Kutas, M., & Hillyard, S. A. (1980). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology, 11*(2), 99-116.

Leech, G., Rayson, P., & Wilson, A. (2001). *Word frequencies in written and spoken english; based on the british national corpus* (1 ed.). London: Longman.

Leopold, W. (1948). Semantic learning in infant language. *Word, 4*, 173-180.

Lin, E. L., & Murphy, G. L. (2001). Thematic relations in adults' concepts. *Journal of Experimental Psychology: General, 130*(1), 3-28.

Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review, 7*(4), 618-630.

MacNeil, D. (1966). A study of word association. *Journal of Verbal Learning and Verbal Behavior, 5*(548-557).

MacWhinney, B. (2000). *The childes project: Tools for analyzing talk. Third edition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Mani, N. (2008). Infants automatically generate a phonological representation of fixated images, *New Directions in Word Learning*. University of York, UK.

Mani, N., Coleman, J., & Plunkett, K. (2008). Phonological specificity of vowel contrasts at 18-months. *Language and Speech, 51*(1&2), 3-21.

Mani, N., & Plunkett, K. (2007). Phonological specificity of consonants and vowels in early lexical representations. *Journal of Memory and Language, 57*, 252-272.

Mani, N., & Plunkett, K. (2008). Phonological priming in infancy. In B. C. Love, K. McRae & V. M. Sloutsky (Eds.), *Proceedings of the 30th annual conference of the cognitive science society*. Washington, DC: Cognitive Science Society.

Markman, E. M. (1989). *Categorization and naming in children*. Cambridge: MA: MIT Press.

Mather, E., & Plunkett, K. (In Press). Learning words over time: The role of stimulus repetition in mutual exclusivity. *Infancy*.

Matsukawa, J., Snodgrass, J., & Doniger, G. (2005). Conceptual versus perceptual priming in incomplete picture identification. *Journal of Psycholinguistic Research, 34*(6), 515-540.

McCauley, C., Weil, C. M., & Sperber, R. D. (1976). The development of memory structure as reflected by semantic-priming effects. *Journal of Experimental Child Psychology, 22*, 511-518.

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. London: Psychology Press.

McRae, K., & Boisvert, S. (1998). Automatic semantic similarity priming. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24*(3), 558-572.

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods, Instruments, & Computers, 37*(4), 547-559.

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General, 126*(2), 99-130.

Meints, K., Plunkett, K., & Harris, P. (1999). When does an ostrich become a bird? The role of typicality in early word comprehension. *Developmental Psychology, 35*(4), 1072-1078.

Meints, K., Plunkett, K., Harris, P. L., & Dimmock, D. (2002). What is 'on' and 'under' for 15-, 18- and 24-month olds? Typicality effects in early comprehension of spatial prepositions. *British Journal of Developmental Psychology, 20*, 113-130.

Merriman, W. E., & Bowman, L. L. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development, 54*(3&4, Serial No. 220), 1-132.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognising pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*, 227-234.

Meyer, D. E., & Schvaneveldt, R. W. (1976). Meaning, memory structure, and mental processes: People's rapid reactions to words reveal how stored semantic information is retrieved. *Science, 192*, 27-33.

Miller, G. A. (1995). Wordnet:An on-line lexical database. *International Journal of Lexicography, 3*((4)).

Moore, K. C. (1896). The mental development of the child. *Psychological Review Monograph Supplements, 1*(3), 1-153.

Morgan, J. L., & Meyer, A. S. (2005). Processing extrafoveal objects during multiple object naming. *Journal of Experimental Psychology: Learning, Memory and Cognition, 31*(3), 428-442.

Morrison, C. M., Chappell, T. D., & Ellis, A. W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *Quarterly Journal of Experimental Psychology, 50A*(3), 528-559.

Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*(1), 116-133.

Morrison, C. M., Ellis, A. W., & Quinlan, P. T. (1992). Age of acquisition, not word frequency, affects object naming, not object recognition. *Memory and Cognition, 20*(6), 705-714.

Moss, H. E., McCormick, S. F., & Tyler, L. K. (1997). The time course of activation of semantic information during spoken word recognition. *Language and Cognitive Processes, 12*(5/6), 695-731.

Moss, H. E., & Older, L. (1996). *Birkbeck word association norms*. Hove, UK: Psychology Press.

Moss, H. E., Ostrin, R. K., Tyler, L. K., & Marslen-Wilson, W. D. (1995). Accessing different types of lexical semantic information: Evidence from priming. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*(4), 863-883.

Murphy, G. L. (2001). Causes of taxonomic sorting by adults: A test of the thematic-to-taxonomic shift. *Psychonomic Bulletin & Review, 8*(4), 834-839.

Nation, K., & Snowling, M. J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. *Cognition, 70*, B1-B13.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106*, 226–254.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition*. Hillsdale NJ: Earlbaum.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The university of south florida word association, rhyme, and word fragment norms. http://w3.usf.edu/FreeAssociation/.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development, 38*(1-2), 1-135.

Nelson, K. (1977). The syntagmatic-paradigmatic shift revisited: A review of research and theory. *Psychonomic Bulletin & Review, 84*(1), 93-116.

Palermo, D. S. (1971). Characteristics of word association responses obtained from children in grades one through four. *Developmental Psychology, 5*(1), 118-123.

Pan, B. A., Rowe, M. L., Spier, E., & Tamis-Lemonda, C. (2004). Measuring productive vocabulary of toddlers in low income families: Concurrent and predictive validity of three sources of data. *Journal of Child Language, 31*, 587-608.

Postman, L., & Keppel, G. (Eds.). (1970). *Norms of word association*. New York: Academic Press.

Raajmakers, J. G. W. (2003). A further look at the "language-as-a-fixed-effect fallacy". *Canadian Journal of Psychology, 57*(3), 141-151.

Raajmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with the "language-as-a-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language, 41*, 416-426.

Radeau, M. (1983). Semantic priming between spoken words in adults and children. *Canadian Journal of Psychology, 4*, 547-556.

Radeau, M., Morais, J., & Dewier, A. (1989). Phonological priming in spoken word recognition: Task effects. *Memory and Cognition, 17*, 525-535.

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin, 85*(3), 618-660.

Reese, G., & Read, S. (2000). Predictive validity of the new zealand macarthur communicative development inventory: Words and sentences. *Journal of Child Language, 27*, 255-266.

Reznick, J. S. (1990). Visual preference as a test of infant word comprehension. *Applied Psycholinguistics, 11*(2), 145-166.

Reznick, J. S., Chawarska, K., & Betts, S. (2000). The development of visual expectations in the first year. *Child Development, 71*(5), 1191-1204.

Reznick, J. S., & Goldfield, B. A. (1994). Diary vs. Representative checklist assessment of productive vocabulary. *Journal of Child Language, 21*, 465-472.

Ring, E. D., & Fenson, L. (2000). The correspondence between parent report and child performance for receptive and expressive vocabulary beyond infancy. *First Language, 20*, 141-159.

Robinson, B. F., & Mervis, C. B. (1999). Comparing productive vocabulary measures from the cdi and a systematic diary study. *Journal of Child Language, 26*(1), 177-185.

Roget, P. M. (1991). *Roget's thesaurus of english words and phrases*: Project Guttenberg, MICRA Inc.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104*(3), 192-233.

Rose, S. A., Feldman, J. F., Jankowski, J. J., & Caro, D., M. (2002). A longitudinal study of visual expectation and reaction time in the first year of life. *Child Development, 73*(1), 47-61.

Rosenzweig, M. R. (1970). International kent rosanoff word association norms, emphasizing those of french male and female students and french workmen. In L. Postman & G. Keppel (Eds.), *Norms of word association*. New York & London: Academinc Press.

Rossell, S. L., & Nobre, A. C. (2004). Semantic priming of different affective categories. *Emotion, 4*(4), 354-363.

Savin, H. B. (1963). Word-frequency effect and errors in the perception of speech. *Journal of the Acoustical Society of America, 36*(2), 200-206.

Schafer, G. (1998). *Word learning in infancy.* Unpublished D.Phil, University of Oxford, Oxford.

Schafer, G., & Plunkett, K. (1998). Rapid word learning under tightly controlled conditions. *Child Development, 69*(2), 309-320.

Schafer, G., Plunkett, K., & Harris, P. L. (1999). What's in a name? Lexical knowledge drives infants' visual preferences in the absence of referential input. *Developmental Science, 2*(2), 188-195.

Schvaneveldt, R. W., Ackerman, B. P., & Semlear, T. (1977). The effect of semantic context on children's word recognition. *Child Development, 48*, 612-616.

Shelton, J. R., & Martin, R. C. (1992). How semantic is automatic semantic priming? *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*(6), 1191-1290.

Simcock, G., & DeLoache, J. (2006). Get the picture? The effects of iconicity on toddlers' reenactment from picture books. *Developmental Psychology, 42*(6), 1352–1357.

Sirois, S., & Jackson, I. (2007). *Pupil dilation and infant cognition.* Paper presented at the IEEE 6th International Conference on Development and Learning 2007, London, UK.

Smiley, S. S., & Brown, A. L. (1979). Conceptual preference for thematic or taxonomic relations: A nonmonotonic age-trend from preschool to old age. *Journal of Experimental Child Psychology, 28*, 249-257.

Snyder, L., Bates, E., & Bretherton, I. (1981). Context and content in early lexical development. *Journal of Child Language, 8*, 565-582.

Spelke, E. (1976). Infants' intermodal perception of events. *Cognitive Psychology, 8*, 553-560.

Spruyt, A., Hermans, D., De Houwer, J., & Eelen, P. (2004). Automatic non-associative semantic priming: Episodic affective priming of naming responses. *Acta Psychologica, 116*(2004), 39-54.

Spss 16.0 for windows. (2006). (2006). Chicago, IL: SPSS Inc.

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science, 29*, 41-78.

Stiles, J. (1994). On the nature of informant judgements in inventory measures. And so what is it you want to know? *Monographs of the Society for Research in Child Development, 59*(5), 180-185.

Styles, S. J., Arias-Trejo, N., & Plunkett, K. (2008). Priming and lexical interference in infancy. In B. C. Love, K. McRae & V. M. Sloutsky (Eds.), *Proceedings of the 30th annual conference of the cognitive science society* (pp. 651-656). Washington, DC: Cognitive Science Society.

Swingley, D., & Aslin, R. (2007). Lexical competition in young children's word learning. *Cognitive Psychology, 54*, 99-132.

Swingley, D., & Fernald, A. (2002). Recognition of words referring to present and absent objects by 24-month-olds. *Journal of Memory and Language, 46*(1), 39-56.

Swingley, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition, 71*(2), 73-108.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (re)consideration of context effects. *Journal of Verbal Learning and Verbal Behaviour, 18*, 645-659.

Taine, M. (1877). On the acquisition of language by children. *Mind: A Quarterly Review of Psychology and Philosophy, 2*(6), 252-259.

Thomas, D. G., Campos, J. J., Shucard, D. W., Ramsay, D. S., & Shucard, J. (1981). Semantic comprehension in infancy: A signal detection analysis. *Child Development, 52*(3), 798-803.

Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York: Teachers' College, Columbia University.

Thorpe, S. J., Gegenfurtner, K. R., Fabre-Thorpe, M., & Bülthoff, H. H. (2001). Detection of animals in natural images using far peripheral vision. *European Journal of Neuroscience, 14*, 869-876.

Tomasello, M., & Mervis, C. B. (1994). The instrument is great, but measuring comprehension is still a problem. *Monographs of the Society for Research in Child Development, 59*(5), 174-179.

Torkildsen, J. v. K., Syversen, G., Simonsen, H. G., Moen, I., & Lindgren, M. (2007). Electrophysiological correlates of auditory semantic priming in 24-month-olds. *Journal of Neurolinguistics, 20*, 332-351.

van Essen, J. (1937). Time and reaction. *The American Journal of Psychology, 50*(1), 414-428.

Vucetic, V. (2008). *Words and pictures: The effect of labelling on visual object processing.* Unpublished D.Phil, University of Oxford, Oxford, UK.

Waxman, S. R., & Braun, I. (2005). Consistent (but not variable) names as invitations to form object categories: New evidence from 12-month-old infants. *Cognition, 95*, B59-B68.

White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representation. *Journal of Memory and Language, 59*(1), 114-132.

Woodrow, H., & Lowell, F. (1916). Children's association frequency tables. *Psychological Monographs, 22*(97).

REFERENCES

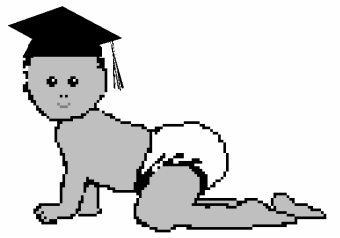
Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*: Touchstone Applied Science Associates (T A S A), Incorporate.

Zipf, G. K. (1935). *The psychobiology of human language*. Boston, MA: Houghton-Mifflin.

Appendix A: Oxford CDI

OXFORD UNIVERSITY
BABYLAB

Communicative Development Inventory
- A UK adaptation of the MacArthur CDI * -



Dear parent,

The following is a list of words that are typical in children's vocabularies.

For words that your child **understands but does not yet say**, place a mark in the first column, labelled "**U**".

	U	U/S
crocodile	●	○

For words that your child **understands and also says,** place a mark in the second column, labelled "**U/S**".

	U	U/S
crocodile	○	●

If your child uses a different pronunciation of a word (e.g., 'bickie' for biscuit, or 'telly' for television) - mark the word anyway.

Occasionally we list two alternative forms - please **underline** the one your child understands and/or produces.

	U	U/S
pool/pond	○	●

Please fill in the whole circle exactly as shown above, do not just tick or partly fill the circle.

correct marking - ● incorrect markings - Ⓥ or ⊙

This inventory is a comprehensive "catalogue" of words that are used by many different children across a wide age range, so do not worry if your child knows only a few of them at the moment!

If you have any additional comments or information that you think we should consider, please add these at the end of this inventory.

Thank you very much!

* For information and original copies of the MacArthur CDI, please contact the Developmental Psychology Lab, San Diego State University, San Diego, CA 92182, USA.

OXFORD UNIVERSITY BABYLAB

Communicative Development Inventory

Subject code

.....

Your name: ……

Child's name: ……. Male/female: …….

Birth date of child: …../……/…… Today's date: …../……./…….

Animal sounds	U	U/S		U	U/S
baa baa	O	O	ouch	O	O
choo choo	O	O	quack	O	O
cockadoodledoo	O	O	uh oh	O	O
grr	O	O	vroom	O	O
meow	O	O	woof	O	O
moo	O	O	yum	O	O

Animals	U	U/S		U	U/S
animal	O	O	horse	O	O
bear	O	O	kitten	O	O
bee	O	O	lamb	O	O
bird	O	O	lion	O	O
bunny / rabbit	O	O	monkey	O	O
butterfly	O	O	mouse	O	O
cat	O	O	owl	O	O
chicken	O	O	penguin	O	O
cow	O	O	pig	O	O
deer	O	O	pony	O	O
dog	O	O	puppy	O	O
donkey	O	O	sheep	O	O
duck	O	O	spider	O	O
elephant	O	O	squirrel	O	O
fish	O	O	tiger	O	O
frog	O	O	turkey	O	O
giraffe	O	O	turtle	O	O
goose	O	O			

Vehicles	U	U/S		U	U/S
aeroplane / plane	O	O	bus	O	O
bicycle / bike	O	O	car	O	O
boat	O	O	fire engine	O	O
lorry / truck	O	O	pushchair/buggy	O	O
motor-bike	O	O	train	O	O

Toys	U	U/S		U	U/S
ball	O	O	doll	O	O
balloon	O	O	pen	O	O
block / brick	O	O	teddy bear	O	O
book	O	O	toy	O	O
bubble	O	O			

Food and Drink	U	U/S		U	U/S
apple	O	O	food	O	O
banana	O	O	ice cream	O	O
biscuit	O	O	jam	O	O
bread	O	O	juice	O	O
butter	O	O	meat	O	O
cake	O	O	milk	O	O
carrot	O	O	orange	O	O
cereal	O	O	pasta / spaghetti	O	O
cheese	O	O	peas	O	O
chicken	O	O	pizza	O	O
chips	O	O	sweets	O	O
coffee	O	O	tea	O	O
drink	O	O	toast	O	O
egg	O	O	water	O	O
fish	O	O			

Body Parts	U	U/S		U	U/S
arm	O	O	hair	O	O
belly button / tummy button	O	O	hand	O	O
cheek	O	O	head	O	O
ear	O	O	knee	O	O
eye	O	O	leg	O	O
face	O	O	nail	O	O
finger	O	O	nose	O	O
foot	O	O	toe	O	O
tongue	O	O	tummy	O	O
tooth	O	O	mouth	O	O

Clothes	U	U/S		U	U/S
bib	O	O	dress	O	O
boot(s)	O	O	glasses / specs	O	O
button	O	O	hat	O	O
coat	O	O	jacket	O	O

Clothes	U	U/S		U	U/S
jeans	O	O	shoe	O	O
jumper / sweater	O	O	shorts	O	O
nappy	O	O	sock	O	O
necklace	O	O	trousers	O	O
pyjamas	O	O	zip	O	O
shirt	O	O			

Furniture and Rooms	U	U/S		U	U/S
bath / bathtub	O	O	living room	O	O
bathroom	O	O	play pen	O	O
bed	O	O	potty	O	O
bedroom	O	O	refrigerator / fridge	O	O
chair	O	O	rocking chair	O	O
cooker / stove / oven	O	O	settee / sofa	O	O
cot	O	O	sink	O	O
door	O	O	stairs	O	O
drawer	O	O	table	O	O
garage	O	O	TV / television	O	O
high chair	O	O	window	O	O
kitchen	O	O			

Outside	U	U/S		U	U/S
beach	O	O	outside	O	O
bucket	O	O	park	O	O
church	O	O	party	O	O
flower	O	O	pool	O	O
garden	O	O	rain	O	O
house	O	O	school	O	O
moon	O	O	shop	O	O
sky	O	O	swing	O	O
slide	O	O	tree	O	O
snow	O	O	wall	O	O
spade	O	O	water	O	O
star	O	O	work	O	O
stone	O	O	zoo	O	O
sun	O	O			

Household items	U	U/S		U	U/S
bin	O	O	bowl	O	O
blanket	O	O	box	O	O
bottle	O	O	broom	O	O

Household items	U	U/S		U	U/S
brush	O	O	paper	O	O
clock	O	O	penny	O	O
comb	O	O	picture	O	O
cup	O	O	pillow	O	O
dish	O	O	plant	O	O
dummy	O	O	plate	O	O
fork	O	O	purse	O	O
glass	O	O	radio	O	O
hammer	O	O	rubbish	O	O
hoover / vacuum	O	O	scissors	O	O
jug	O	O	soap	O	O
key	O	O	spoon	O	O
lamp	O	O	telephone	O	O
light	O	O	toothbrush	O	O
medicine	O	O	towel	O	O
money	O	O	watch	O	O
mug	O	O			

People	U	U/S		U	U/S
aunt	O	O	girl	O	O
baby	O	O	grandma	O	O
boy	O	O	grandpa	O	O
brother	O	O	lady	O	O
child	O	O	man	O	O
daddy	O	O	mummy	O	O
doctor	O	O	nanny	O	O
friend	O	O	people	O	O
person	O	O	teacher	O	O
policeman	O	O	uncle	O	O
sister	O	O			

Games and Routines	U	U/S		U	U/S
bath	O	O	no	O	O
breakfast	O	O	pat-a-cake	O	O
bye bye	O	O	peekaboo	O	O
dinner	O	O	please	O	O
don't	O	O	shh / hush / shush	O	O
hello	O	O	tea	O	O
hi	O	O	thank you	O	O
lunch	O	O	wait	O	O
nap	O	O	want to	O	O
night night	O	O	yes	O	O

Action Words	U	U/S		U	U/S
bite	O	O	know	O	O
blow	O	O	like	O	O
break	O	O	look	O	O
bring	O	O	love	O	O
bump	O	O	make	O	O
call	O	O	open	O	O
carry	O	O	play	O	O
catch	O	O	pull	O	O
clean	O	O	push	O	O
cry	O	O	put	O	O
cuddle	O	O	read	O	O
cut	O	O	ride	O	O
dance	O	O	run	O	O
draw	O	O	say	O	O
drink	O	O	scratch	O	O
drive	O	O	see	O	O
drop	O	O	show	O	O
eat	O	O	shut / close	O	O
fall	O	O	sing	O	O
feed	O	O	sleep	O	O
find	O	O	smile	O	O
finish	O	O	splash	O	O
get	O	O	stop	O	O
give	O	O	swim	O	O
go	O	O	swing	O	O
have	O	O	take	O	O
hear	O	O	tell	O	O
help	O	O	throw	O	O
hit	O	O	tickle	O	O
hug	O	O	walk	O	O
hurry	O	O	wash	O	O
jump	O	O	watch	O	O
kick	O	O	wipe	O	O
kiss	O	O	write	O	O

Descriptive Words	U	U/S		U	U/S
all gone	O	O	clean	O	O
asleep	O	O	cold	O	O
bad	O	O	dark	O	O
big	O	O	dirty	O	O
blue	O	O	dry	O	O
broken	O	O	empty	O	O
careful	O	O	fast	O	O

Descriptive Words	U	U/S		U	U/S
fine	O	O	old	O	O
gentle	O	O	pretty	O	O
good	O	O	red	O	O
green	O	O	sad	O	O
happy	O	O	scared	O	O
hard	O	O	sick	O	O
hot	O	O	sleepy	O	O
hungry	O	O	soft	O	O
hurt	O	O	thirsty	O	O
little	O	O	tired	O	O
nasty	O	O	wet	O	O
naughty	O	O	yellow	O	O
nice	O	O			

Question words	U	U/S		U	U/S
how	O	O	where	O	O
what	O	O	who	O	O
when	O	O	why	O	O

Time	U	U/S		U	U/S
day	O	O	now	O	O
later	O	O	today	O	O
morning	O	O	tomorrow	O	O
night	O	O	tonight	O	O

Pronouns	U	U/S		U	U/S
her	O	O	my	O	O
his	O	O	that	O	O
I	O	O	this	O	O
it	O	O	you	O	O
me	O	O	your	O	O
mine	O	O			

Prepositions	U	U/S		U	U/S
away	O	O	on	O	O
back	O	O	out	O	O
down	O	O	there	O	O
in	O	O	under	O	O
inside	O	O	up	O	O
off	O	O			

Quantifiers	U	U/S		U	U/S
all	O	O	not	O	O
again	O	O	other	O	O
another	O	O	same	O	O
more	O	O	some	O	O
none	O	O			

Extra words	U	U/S		U	U/S
chase (action)	O	O		O	O
smell (action)	O	O		O	O
	O	O		O	O
	O	O		O	O
	O	O		O	O
	O	O		O	O

Additional Questions:

Does anyone speak to your child in a language other than English (if so, which language)?

Has your child ever had any hearing problems, including glue ear?

Was your child born more than six weeks premature?

Thank you for your help.
If you have any further comments, please write them below.

Appendix B: CDI comprehension rate (Alphabetic)

Lexical Comprehension Rate at 18 months – Alphabetic

Percentage of 18-month-olds reported to understand words in the Oxford CDI

A

aeroplane / airplane	83.2
again	65.4
all	20.1
all gone	84.9
animal	38.5
another	23.5
apple	77.1
arm	55.3
asleep	52.0
aunt	11.2
away	22.9

B

baa baa	81.6
baby	88.3
back	22.9
bad	21.8
ball	97.8
balloon	81.6
banana	90.5
bath	96.6
bath / bathtub	93.3
bathroom	62.0
beach	11.2
bear	53.1
bed	83.2
bedroom	60.3
bee	60.9
belly button / tummy button	53.6
bib	75.4
bicycle / bike	70.9
big	20.7
bin	69.3
bird	88.8
biscuit	87.7
bite	53.1
blanket	36.3
block / brick	59.8
blow	52.5
blue	15.1
boat	62.6
book	95.5
boots	53.6
bottle	62.6
bowl	57.5
box	46.9
boy	38.0
bread	72.1
break	18.4
breakfast	78.8
bring	60.9
broken	27.4
broom	25.1
brother	12.3
brush	70.9
bubble	60.3
bucket	38.5
bump	45.3
bunny / rabbit	77.7
bus	69.3
butter	24.6
butterfly	55.3
button	41.9
bye bye	98.9

C

cake	55.3
call	16.2
car	96.6
careful	59.8
carrot	47.5
carry	36.3
cat	94.4
catch	36.3
cereal	25.1
chair	82.1
cheek	28.5
cheese	63.1
chicken	59.8
chicken	24.6
child	8.9
chips	43.6
choo choo	52.5
church	10.6
clean	41.3
clean	39.1
clock	65.4
coat	78.2
cock-a-doodle-doo	50.3
coffee	14.5
cold	41.3
comb	31.3
cooker / stove / oven	39.1
cot	72.1
cow	85.5
cry	46.9
cuddle	86.6
cup	79.3
cut	9.5

D

daddy	97.2
dance	82.7
dark	19.6
day	7.8
deer	11.7
dinner	56.4
dirty	60.9
dish	25.1
doctor	9.5
dog	98.3
doll	59.2
don’t	58.7
donkey	27.4
door	87.7
down	67.0
draw	44.1
drawer	35.8
dress	26.8
drink	83.8
drink	82.1
drive	24.0
drop	25.1
dry	14.5
duck	91.1
dummy	29.6

E

ear	82.7
eat	71.5
egg	31.8
elephant	54.2
empty	38.5
eye	86.0

F	
face	43.6
fall	33.5
fast	13.4
feed	26.8
find	45.8
fine	1.7
finger	82.1
finish	30.2
fire engine	26.8
fish	76.0
fish	44.1
flower	77.7
food	48.6
foot	70.4
fork	44.7
friend	11.2
frog	56.4

G	
garage	14.5
garden	73.7
gentle	53.6
get	63.7
giraffe	37.4
girl	27.9
give	76.0
glass	31.3
glasses / specs	48.0
go	56.4
good	63.7
goose	13.4
grandma	58.7
grandpa	59.2
green	13.4
grr	60.9

H	
hair	91.1
hammer	20.1
hand	77.1
happy	21.2
hard	7.8
hat	88.3
have	17.9
head	75.4
hear	25.7
hello	89.4
help	29.6
her	5.6
hi	36.3
high chair	67.0
his	5.6
hit	32.4
hoover / vacuum	55.9
horse	77.1
hot	75.4
house	57.0
how	8.4
hug	62.0
hungry	45.8
hurry	15.1
hurt	24.6

I	
I	11.2
ice cream	40.8
in	50.8
inside	26.8
it	13.4

J	
jacket	21.2
jam	18.4
jeans	10.6
jug	8.9
juice	72.1
jump	50.8
jumper / sweater	44.7

K	
key	73.2
kick	44.7
kiss	92.7
kitchen	64.2
kitten	24.0
knee	43.6
know	7.3

L	
lady	24.0
lamb	25.1
lamp	17.3
later	17.9
leg	57.0
light	62.0
like	24.0
lion	64.8
little	13.4
living room	21.2
look	65.4
lorry / truck	62.0
love	35.2
lunch	55.9

M	
make	10.6
man	35.2
me	28.5
meat	13.4
medicine	29.1
meow	84.9
milk	87.2
mine	33.0
money	30.7
monkey	57.0
moo	88.3
moon	42.5
more	82.7
morning	7.8
motor-bike	33.5
mouse	54.2
mouth	77.1
mug	13.4
mummy	98.9
my	16.2

N	
nail	16.8
nanny	34.6
nap	23.5
nappy	92.2
nasty	9.5
naughty	48.0
necklace	16.8
nice	27.9
night	20.7
night night	79.9
no	99.4
none	11.2
nose	92.7
not	12.8
now	12.8

O	
off	42.5
old	1.1
on	51.4
open	68.7
orange	36.9
other	13.4
ouch	57.0
out	40.2
outside	67.0
owl	41.3

P	
paper	41.3
park	36.9
party	4.5
pasta / spaghetti	35.2
pat-a-cake	37.4
peas	46.9
peekaboo	88.8
pen	52.0
penguin	33.0
penny	13.4
people	13.4
person	6.1
picture	43.0
pig	77.7
pillow	30.2
pizza	17.3
plant	29.6
plate	51.4
play	48.0
play pen	5.0
please	53.6
policeman	5.6
pony	10.1
pool	17.3
potty	35.2
pretty	20.7
pull	32.4
puppy	20.1
purse	27.4
push	43.6
pushchair	77.7
put	43.6
pyjamas	54.2

Q	
quack	85.5

R	
radio	29.1
rain	45.8
read	57.0
red	12.3
refrigerator / fridge	55.9
ride	26.8
rocking chair	7.8
rubbish	28.5
run	45.8

S	
sad	11.7
same	6.1
say	34.1
scared	3.4
school	15.6
scissors	25.1
scratch	10.1
see	33.5
settee / sofa	48.0
sheep	70.4
shh / hush / shush	65.9
shirt	19.6
shoe	98.3
shop	39.1
shorts	11.7
show	38.0
shut / close	74.3
sick	6.1
sing	52.5
sink	20.1
sister	16.2
sky	41.9
sleep	75.4
sleepy	35.8
slide	57.0
smile	41.3
snow	7.8
soap	39.7
sock	91.6
soft	14.5
some	13.4
spade	24.6
spider	47.5
splash	70.4
spoon	78.8
squirrel	26.3
stairs	80.4
star	40.2
stone	31.8
stop	54.7
sun	40.2
sweets	26.8
swim	42.5
swing	64.2
swing	52.0

T	
table	63.7
take	28.5
tea	44.1
tea	56.4
teacher	5.6
teddy bear	87.7
telephone	85.5
tell	8.9
thank you	71.5
that	19.6
there	33.5
thirsty	39.1
this	14.5
throw	62.0
tickle	65.9
tiger	51.4
tired	52.5
toast	69.3
today	5.6
toe	69.8
tomorrow	5.0
tongue	50.3
tonight	2.8
tooth	74.9
toothbrush	87.2
towel	56.4
toy	60.3
train	66.5
tree	66.5
trousers	55.3
tummy	82.1
turkey	3.9
turtle	10.1
TV	75.4

U	
uh oh	74.3
uncle	8.9
under	28.5
up	59.8

V	
vroom	59.2

W	
wait	39.1
walk	74.3
wall	15.1
want to	25.7
wash	62.0
watch	54.2
watch	26.3
water	60.9
water	57.0
wet	41.3
what	31.3
when	5.0
where	58.7
who	29.1
why	2.8
window	61.5
wipe	45.8
woof	86.6
work	22.9
write	12.3

Y	
yellow	13.4
yes	86.0
you	21.8
your	21.2
yum	62.0

Z	
zip	27.4
zoo	4.5

compiled by Suzy Styles 2004

Appendix C: CDI comprehension rate (by Category)

Lexical Comprehension Rate at 18 months – by Type

Percentage of 18-month-olds reported to understand words in the Oxford CDI

Animals

98.3	dog	54.7	elephant	4.5	turkey
94.4	cat	54.2	bear		
91.6	duck	51.4	tiger		
88.8	bird	48.0	spider		
86.0	cow	41.9	owl		
78.2	bunny / rabbit	39.7	animal		
78.2	pig	37.4	giraffe		
77.7	horse	33.0	penguin		
76.5	fish	27.9	donkey		
70.9	sheep	26.8	squirrel		
65.4	lion	25.1	lamb		
61.5	bee	24.0	kitten		
60.3	chicken	20.7	puppy		
57.5	monkey	13.4	goose		
57.0	frog	12.3	deer		
55.9	butterfly	10.6	pony		
54.7	mouse	10.1	turtle		

Animal Sounds

88.8	moo
87.2	woof
86.0	quack
85.5	meow
81.6	baa baa
74.9	uh oh
62.0	yum
61.5	grr
59.2	vroom
56.4	ouch
52.5	choo choo
50.8	cockadoodledoo

Action Words

92.7	kiss	52.5	sing	33.0	hit
86.6	cuddle	52.5	blow	32.4	pull
82.7	dance	51.4	jump	30.7	help
82.1	drink	47.5	play	30.2	finish
76.5	give	46.9	cry	27.9	take
75.4	sleep	46.4	run	27.4	feed
74.3	walk	46.4	wipe	26.8	ride
74.3	shut / close	45.8	bump	26.8	watch
71.5	eat	45.3	find	26.8	hear
70.4	splash	44.7	draw	25.7	drop
68.7	open	44.1	kick	24.6	like
66.5	tickle	44.1	put	24.6	drive
65.4	look	43.0	push	18.4	break
63.7	get	43.0	swim	17.3	have
62.6	wash	41.9	clean	16.8	call
62.6	hug	41.9	smile	15.1	hurry
62.0	throw	39.1	show	12.8	write
60.9	bring	37.4	carry	11.2	make
57.5	read	36.9	catch	10.6	scratch
56.4	go	36.3	love	10.1	cut
54.7	stop	34.6	see	9.5	tell
54.2	bite	34.6	say	7.8	know
52.5	swing	34.1	fall		

Body Parts

93.3 nose
91.1 hair
86.6 eye
83.2 ear
82.1 tummy
82.1 finger
78.2 mouth
78.2 hand
76.0 head
75.4 tooth
70.9 foot
70.4 toe
58.1 leg
56.4 arm
54.2 belly button / tummy button
50.8 tongue
44.7 face
44.1 knee
29.1 cheek
17.9 nail

Clothing

98.3 shoe
92.2 sock
92.2 nappy
88.3 hat
78.8 coat
76.0 bib
56.4 trousers
54.7 boots
54.2 pyjamas
48.6 glasses / specs
44.7 jumper / sweater
42.5 button
27.9 zip
27.4 dress
21.8 jacket
20.7 shirt
17.3 necklace
12.8 shorts
11.2 jeans

Food & Drink

90.5 banana
88.3 biscuit
87.2 milk
83.8 drink
77.7 apple
72.6 bread
71.5 juice
69.8 toast
63.7 cheese
60.9 water
55.9 cake
49.7 food
48.0 carrot
47.5 peas
44.7 chips
44.1 tea
44.1 fish
40.8 ice cream
36.9 orange
35.2 pasta / spaghetti
31.8 egg
27.4 sweets
26.3 cereal
25.1 butter
24.6 chicken
18.4 jam
17.9 pizza
14.5 coffee
14.0 meat

Games & Routines

99.4 no
98.9 bye bye
96.6 bath
89.4 hello
88.3 peekaboo
86.0 yes
79.9 night night
78.8 breakfast
70.9 thank you
65.9 shh / hush / shush
58.7 don't
57.0 tea
56.4 lunch
55.9 dinner
54.2 please
39.1 wait
37.4 hi
37.4 pat-a-cake
25.1 want to
24.0 nap

Toys

97.8 ball
96.1 book
87.7 teddy bear
82.1 balloon
60.9 toy
60.3 bubble
60.3 block / brick
58.7 doll
52.5 pen

Vehicles

96.6 car
83.2 aeroplane
77.7 pushchair
70.9 bicycle / bike
69.3 bus
67.0 train
63.1 boat
62.6 lorry / truck
33.5 motor-bike
27.9 fire engine

Furniture & Rooms

93.9	bath / bathtub	64.2	table	36.3	drawer
87.7	door	62.6	bathroom	35.2	potty
83.8	bed	61.5	window	21.8	living room
81.6	chair	60.9	bedroom	21.2	sink
80.4	stairs	56.4	refrigerator / fridge	15.1	garage
74.9	TV			8.4	rocking chair
72.6	cot	48.0	settee / sofa	5.0	play pen
67.0	high chair	39.7	cooker / stove / oven		
64.8	kitchen				

Household Objects

87.7	toothbrush	54.7	watch	30.2	medicine
85.5	telephone	52.0	plate	29.6	rubbish
79.9	cup	46.9	box	29.6	plant
79.3	spoon	45.3	fork	29.6	radio
73.2	key	43.6	picture	27.9	purse
71.5	brush	41.9	paper	25.7	dish
69.8	bin	39.7	soap	25.1	scissors
65.9	clock	37.4	blanket	24.6	broom
62.6	light	31.8	comb	20.1	hammer
62.0	bottle	31.3	glass	17.9	lamp
58.1	bowl	30.7	money	13.4	mug
57.0	towel	30.7	pillow	12.8	penny
55.3	hoover / vacuum	30.2	dummy	9.5	jug

Outside

78.2	flower	43.0	moon	23.5	work
74.3	garden	41.9	sky	17.9	pool
67.0	tree	41.3	star	16.2	school
67.0	outside	40.2	sun	15.1	wall
64.8	swing	39.1	shop	11.7	church
57.5	house	39.1	bucket	10.6	beach
57.5	slide	37.4	park	8.4	snow
56.4	water	32.4	stone	5.0	party
46.4	rain	25.1	spade	5.0	zoo

People

98.9	mummy	34.6	nanny	11.2	aunt
97.2	daddy	28.5	girl	9.5	doctor
88.8	baby	24.0	lady	8.9	child
59.8	grandpa	16.8	sister	8.9	uncle
58.7	grandma	13.4	people	6.1	person
38.0	boy	11.7	friend	5.6	policeman
35.8	man	11.7	brother	5.6	teacher

Adjectives

85.5	all gone	39.7	clean	15.1	soft
76.0	hot	38.5	empty	14.5	little
63.7	good	36.3	sleepy	14.0	yellow
61.5	dirty	27.9	nice	14.0	green
60.3	careful	27.4	broken	14.0	fast
54.2	gentle	25.7	hurt	12.8	red
53.1	tired	21.8	bad	12.3	sad
52.5	asleep	21.2	happy	10.1	nasty
48.6	naughty	21.2	big	8.4	hard
46.9	hungry	20.7	pretty	6.7	sick
41.9	cold	20.1	dark	3.9	scared
41.3	wet	15.6	blue	1.7	fine
40.2	thirsty	15.1	dry	1.7	old

Prepositions

68.2	down	27.9	inside	16.8	my
60.3	up	22.9	back	14.5	this
51.4	on	22.9	away	14.0	it
50.8	in	33.0	mine	11.2	I
43.0	off	28.5	me	5.6	her
40.8	out	22.3	you	5.6	his
34.6	there	21.8	your		
29.6	under	20.1	that		

Quantifiers

83.2	more	21.2	all	12.8	not
65.4	again	14.0	some	11.7	none
24.0	another	13.4	other	6.1	same

Question Words

58.1	where	29.1	who	5.6	when
31.8	what	8.9	how	3.4	why

Times

20.7	night	8.4	morning	5.6	tomorrow
17.9	later	7.8	day	3.4	tonight
13.4	now	6.1	today		

Compiled by Suzy Styles 2004

Appendix D: Word Associations

Word Associations 18M – Semantic type

Associative Strength (Collated from Moss et al., 1996)

Taxonomic Relationships

cat	dog	66.7
dog	cat	51.1
truck	lorry	53.3
boot(s)	shoe	30.0
plate	cup(s)	25.2
lorry	truck	23.8
mouse	cat	20.9
slide	swing	18.8
sheep	cow	13.3
hat	coat	12.5
coat	hat	12.5
toothbrush	brush	11.1
cot	bed	9.5
elephant	mouse	8.9
truck	car	8.9
bowl	spoon	7.1
pig	cow	6.2
train	bus	6.2
biscuit	cake	4.8
bowl	cup	4.8
lorry	car	4.8
car	bike	4.8
bicycle	car	4.8
swing	slide	4.4
cat	mouse	4.2
duck	bird	4.2
bus	car	4.1
bus	train	4.1
car	bicycle	3.6
bowl	plate	2.4
plate	spoon	2.4
car	truck	2.4
apple	banana	2.4
lorry	bus	2.4
mouse	elephant	2.3
bear	dog	2.3
rabbit	dog	2.3
cake	biscuit	2.2
cat	bird	2.2
dog	horse	2.2
sheep	dog	2.2
sheep	pig	2.2
trousers	shoes	2.2
watch	clock	2.2
train	boat	2.1
train	car	2.1
truck	bus	2.1
truck	train	2.1
bib	nappy	2.1
pig	horse	2.1
pig	sheep	2.1
rabbit	cat	1.2
rabbit	duck	1.2
biscuit	cheese	1.2
car	bus	1.2
cup	plate	1.1

Non-Taxonomic

boat	house	10.4
monkey	tree	10.3
bowl	fish	9.5
bowl	dog(gy)	4.8
plate	fish	4.8
plate	table	4.8
flower	bed	4.8
mouse	house	4.7
apple	tree	2.4
flower	bird	2.4
house	car	2.4
house	tree	2.4
mouse	clock	2.3
cake	cup	2.2
cake	spoon	2.2
cake	table	2.2
dog	house	2.2
light	cheese	2.2
telephone	rabbit	2.2
train	dog	2.1
slide	banana	2.1
fish	bowl	2.0
monkey	banana	1.1

Word Associations 18M - Alphabetical listing by stimulus word

Associative Strength (Collated from Moss et al., 1996)

Taxonomic Relationships

apple	banana	2.4
bear	dog	2.3
bib	nappy	2.1
bicycle	car	4.8
biscuit	cake	4.8
biscuit	cheese	1.2
boot(s)	shoe	30.0
bowl	cup	4.8
bowl	plate	2.4
bowl	spoon	7.1
bus	car	4.1
bus	train	4.1
cake	biscuit	2.2
car	bicycle	3.6
car	bike	4.8
car	bus	1.2
car	truck	2.4
cat	bird	2.2
cat	dog	66.7
cat	mouse	4.2
coat	hat	12.5
cot	bed	9.5
cup	plate	1.1
dog	cat	51.1
dog	horse	2.2
duck	bird	4.2
elephant	mouse	8.9
hat	coat	12.5
pig	cow	6.2
pig	horse	2.1
pig	sheep	2.1
plate	cup(s)	25.2
plate	spoon	2.4
lorry	bus	2.4
lorry	car	4.8
lorry	truck	23.8
mouse	cat	20.9
mouse	elephant	2.3
rabbit	cat	1.2
rabbit	dog	2.3
rabbit	duck	1.2
sheep	cow	13.3
sheep	dog	2.2
sheep	pig	2.2
slide	swing	18.8
swing	slide	4.4
toothbrush	brush	11.1
train	boat	2.1
train	bus	6.2
train	car	2.1
trousers	shoes	2.2
truck	car	8.9
truck	bus	2.1
truck	lorry	53.3
truck	train	2.1
watch	clock	2.2

Non-Taxonomic

apple	tree	2.4
boat	house	10.4
bowl	fish	9.5
bowl	dog(gy)	4.8
flower	bed	4.8
monkey	tree	10.3
mouse	house	4.7
cake	cup	2.2
cake	spoon	2.2
cake	table	2.2
dog	house	2.2
fish	bowl	2.0
flower	bird	2.4
house	car	2.4
house	tree	2.4
light	cheese	2.2
monkey	banana	1.1
mouse	clock	2.3
plate	fish	4.8
plate	table	4.8
slide	banana	2.1
telephone	rabbit	2.2
train	dog	2.1

Word Associations 18M - Alphabetical list by response word

Associative Strength (Collated from Moss et al., 1996)

Taxonomic Relationships

apple	banana	2.4
cot	bed	9.5
car	bicycle	3.6
car	bike	4.8
cat	bird	2.2
duck	bird	4.2
cake	biscuit	2.2
train	boat	2.1
toothbrush	brush	11.1
car	bus	1.2
lorry	bus	2.4
train	bus	6.2
truck	bus	2.1
biscuit	cake	4.8
bicycle	car	4.8
bus	car	4.1
lorry	car	4.8
train	car	2.1
truck	car	8.9
dog	cat	51.1
mouse	cat	20.9
rabbit	cat	1.2
biscuit	cheese	1.2
watch	clock	2.2
hat	coat	12.5
pig	cow	6.2
sheep	cow	13.3
bowl	cup	4.8
plate	cup(s)	25.2
bear	dog	2.3
cat	dog	66.7
rabbit	dog	2.3
sheep	dog	2.2
rabbit	duck	1.2
mouse	elephant	2.3
coat	hat	12.5
dog	horse	2.2
pig	horse	2.1
truck	lorry	53.3
cat	mouse	4.2
elephant	mouse	8.9
sheep	pig	2.2
bib	nappy	2.1
bowl	plate	2.4
cup	plate	1.1
pig	sheep	2.1
boot(s)	shoe	30
trousers	shoes	2.2
swing	slide	4.4
bowl	spoon	7.1
plate	spoon	2.4
slide	swing	18.8
bus	train	4.1
truck	train	2.1
car	truck	2.4
lorry	truck	23.8

Non-Taxonomic

monkey	banana	1.1
slide	banana	2.1
flower	bed	4.8
flower	bird	2.4
fish	bowl	2
house	car	2.4
light	cheese	2.2
mouse	clock	2.3
cake	cup	2.2
train	dog	2.1
bowl	dog(gy)	4.8
bowl	fish	9.5
plate	fish	4.8
boat	house	10.4
mouse	house	4.7
dog	house	2.2
telephone	rabbit	2.2
cake	spoon	2.2
cake	table	2.2
plate	table	4.8
apple	tree	2.4
monkey	tree	10.3
house	tree	2.4

Untested for Associative Strength

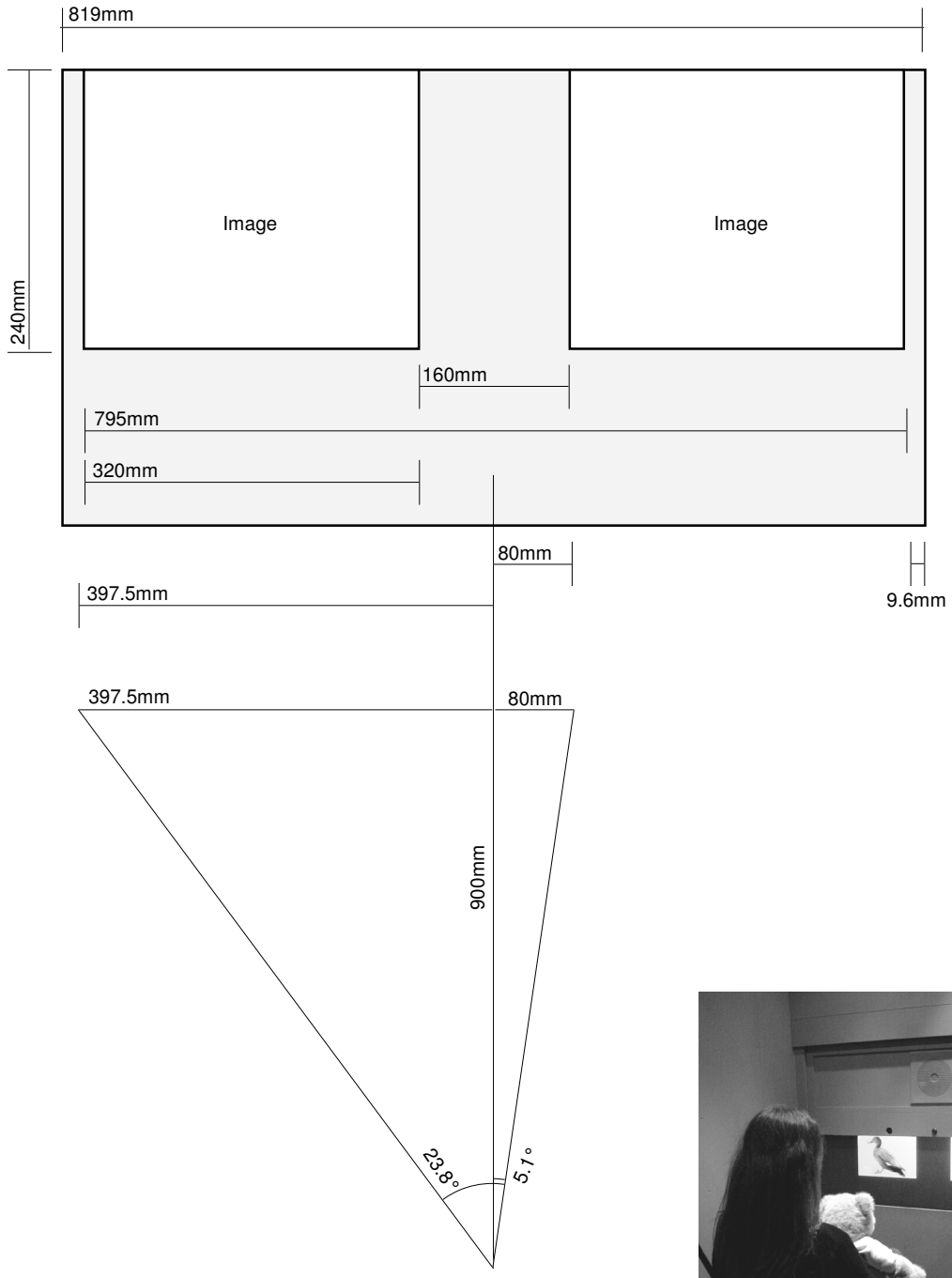
(avoid as first in priming pair) (Collated from Moss et al., 1996)

banana	all	hoover	all
bed	all	horse	all
bike	all	lion	all
bird	all	nappy	all
bottle	all	pushchair	all
bunny	all	pyjamas	all
chair	all	refrigerator	all
cheese	all	shoe	all
chicken	all	sock	all
clock	all	table	all
cow	all	tiger	all
fridge	all	toast	all
high-chair	all	towel	all

Appendix E: Laboratory Testing Booths

Booth 1

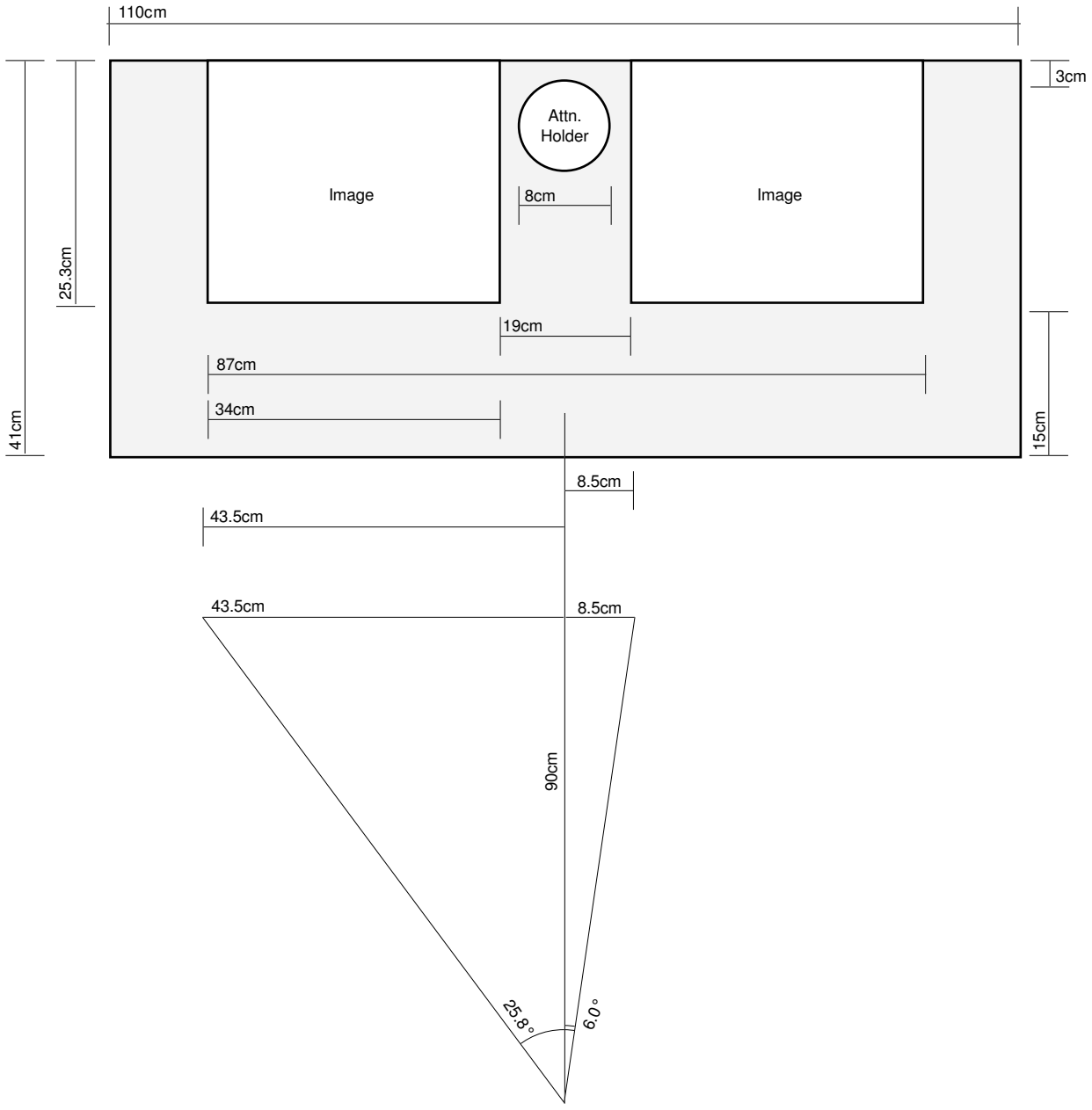
Minimally-primed IPL, words-before-pictures & quick-primed IPL





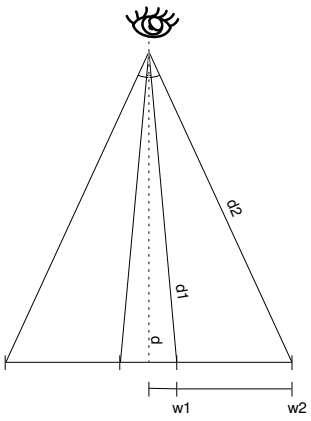
Booth 2 Screen

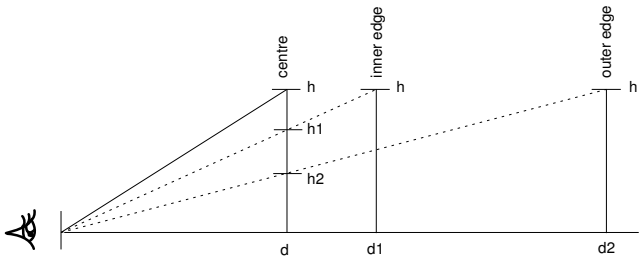
Half-target quick primed IPL



Appendix F: Screen Distortion

Perspective Distortion:
Calculations for peripheral images in IPL





Measured Values

d	distance from screen
h	height of picture
w1	width from centre to inner edge
w2	width from centre to outer edge

Apparent height

$$h1 = \frac{d * h}{\sqrt{d^2 + w1^2}}$$

Distortion = h1 / h

$$h2 = \frac{d * h}{\sqrt{d^2 + w2^2}}$$

Distortion = h2 / h

Appendix G: Stimulus Lists

Minimally Primed IPL (Chapter 3)

List A

Prime & Carrier	Attention		Target	Distracter	Prime Type	WA-Strength[a]
Yesterday, I saw a pig	Look…	Which...	Horse	Lion	Compound	
Yesterday, I saw a rabbit	Which...	Can…	Cow	Chicken	Taxonomic	
Yesterday, I saw an elephant	Can…	Where's	Dog	Monkey	Taxonomic	0
Yesterday, I bought a boot	Where's	Look…	Mouse	Bear	Unrelated	0
Yesterday, I ate[b] an apple	Look…	Which...	Banana	Toast	Compound	
Yesterday, I saw a train	Which...	Can…	Cake	Bread	Unrelated	0
Yesterday, I saw an aeroplane	Can…	Where's	Bus	Lorry	Taxonomic	0
Yesterday, I saw a car	Where's	Look…	Bicycle	Pushchair	Compound	
Yesterday, I saw a sheep	Look…	Which...	Coat	Sock	Unrelated	0
Yesterday, I bought a hat	Which...	Can…	Shoe	Pyjamas	Taxonomic	0
Yesterday, I bought a plate	Can…	Where's	Cup	Bowl	Unrelated	0
Yesterday, I saw a cat	Where's	Which...	Bed	Highchair	Unrelated	0

List B

Prime & Carrier	Attention		Target	Distracter	Prime Type	WA-Strength[a]
Yesterday, I bought a plate	Can…	Where's	Horse	Lion	Unrelated	0
Yesterday, I bought a hat	Where's	Look…	Cow	Chicken	Unrelated	0
Yesterday, I saw a cat	Look…	Which...	Dog	Monkey	Compound	
Yesterday, I saw a sheep	Which...	Can…	Mouse	Bear	Taxonomic	0
Yesterday, I ate[b] cheese	Can…	Where's	Banana	Toast	Taxonomic	0
Yesterday, I ate[b] an biscuit	Where's	Look…	Cake	Bread	Compound	
Yesterday, I ate an apple	Look…	Which...	Bus	Lorry	Unrelated	0
Yesterday, I saw a train	Which...	Can…	Bicycle	Pushchair	Taxonomic	0
Yesterday, I bought trousers	Can…	Where's	Coat	Sock	Taxonomic	0
Yesterday, I saw a boot	Where's	Which...	Shoe	Pyjamas	Compound	
Yesterday, I saw an elephant	Can…	Where's	Cup	Bowl	Unrelated	0
Yesterday, I bought a cot	Where's	Look…	Bed	Highchair	Compound	

List C

Prime & Carrier	Attention		Target	Distracter	Prime Type	WA-Strength[a]
Yesterday, I saw a cat	Look…	Which...	Horse	Lion	Taxonomic	0
Yesterday, I saw a sheep	Which...	Can…	Cow	Chicken	Compound	
Yesterday, I bought a cot	Can…	Where's	Dog	Monkey	Unrelated	0
Yesterday, I saw an elephant	Where's	Look…	Mouse	Bear	Compound	
Yesterday, I saw a car	Look…	Which...	Banana	Toast	Unrelated	0
Yesterday, I ate[b] an apple	Which...	Can…	Cake	Bread	Taxonomic	0
Yesterday, I saw a train	Can…	Where's	Bus	Truck	Compound	
Yesterday, I ate cheese	Where's	Look…	Bicycle	Pushchair	Unrelated	0
Yesterday, I bought a hat	Look…	Which...	Coat	Sock	Compound	
Yesterday, I saw a pig	Which...	Can…	Shoe	Pyjamas	Unrelated	0
Yesterday, I bought a spoon	Can…	Where's	Cup	Bowl	Taxonomic	0
Yesterday, I bought a table	Where's	Which...	Bed	Highchair	Taxonomic	0

[a] Forward Adult Word Association norms from Moss & Older (1996)

[b] Semantically restrictive verb

Words-before-pictures (Chapter 4)

List A

Target Phrase	Target	Distracter
Look! Frog!	Frog	Chair
Look! Key!	Key	Butterfly
Look! Rabbit!	Rabbit	Bread
Look!Phone!	Phone	Flower
Look!Duck!	Duck	Tree
Look!Spoon!	Spoon	Bee
Look! Bird!	Bird	Clock
Look! Television!	Television	Fish

List B

Prime & Carrier	Target	Distracter
Look! Chair!	Chair	Frog
Look! Butterfly!	Butterfly	Key
Look! Bread!	Bread	Rabbit
Look! Flower!	Flower	Phone
Look! Tree!	Tree	Duck
Look! Bee!	Bee	Spoon
Look! Clock!	Clock	Bird
Look! Fish!	Fish	Television

Quick Primed IPL (Chapter 5)

List A

Prime & Carrier	Target	Distracter	Prime Type	WA-Strength[a]
Yesterday, I saw a cat	Dog	Boat	Related	66.7
Yesterday, I saw a sheep	Cow	Toast	Related	13.3
Yesterday, I ate[b] an apple	Banana	Lion	Related	2.4
Yesterday, I bought a boot	Shoe	Cheese	Related	30.0
Yesterday, I bought a plate	Cup[b]	Pushchair	Related	25.2
Yesterday, I bought a cot	Bed	Chicken	Related	9.5
Yesterday, I saw a train	Horse	Sock	Unrelated	0
Yesterday, I saw a lorry	Mouse	Table	Unrelated	0
Yesterday, I saw an elephant	Cake	Trousers	Unrelated	0
Yesterday, I bought a hat	Bus	Monkey	Unrelated	0
Yesterday, I saw a pig	Car	Bowl	Unrelated	0
Yesterday, I ate[b] a biscuit	Coat	Bear	Unrelated	0

List B

Prime & Carrier	Target	Distracter	Prime Type	WA-Strength[a]
Yesterday, I bought a plate	Dog	Boat	Unrelated	0
Yesterday, I bought a boot	Cow	Toast	Unrelated	0
Yesterday, I bought a cot	Banana	Lion	Unrelated	0
Yesterday, I saw a cat	Shoe	Cheese	Unrelated	0
Yesterday, I bought a sheep	Cup[b]	Pushchair	Unrelated	0
Yesterday, I ate[b] an apple	Bed	Chicken	Unrelated	0
Yesterday, I saw a pig	Horse	Sock	Related	2.1
Yesterday, I saw an elephant	Mouse	Table	Related	8.9
Yesterday, I ate[b] a biscuit	Cake	Trousers	Related	4.8
Yesterday, I saw a train	Bus	Monkey	Related	6.2
Yesterday, I saw a lorry	Car	Bowl	Related	4.8
Yesterday, I bought a hat	Coat	Bear	Related	12.5

Notes. Word Association norms from (Moss & Older, 1996)

Half Target Quick Primed IPL (Chapter 6)

List A1

Prime & Carrier	Target	Distracter	Prime Cond.	Name Cond.	WA-Strength[a]
I saw a lorry	bus	cow	Named	Related	2.4
I saw a boot	shoe	window	Named	Related	30.0
I saw the sun	moon	bed	Named	Related	11.1
I bought a cot	duck	bread	Named	Unrelated	0
I bought a hat	plate	bicycle	Named	Unrelated	0
I ate a biscuit	foot	tiger	Named	Unrelated	0
I saw a pig	horse	cake	Unnamed	Related	2.1
I ate an apple	banana	mouse	Unnamed	Related	2.4
I saw a table	chair	coat	Unnamed	Related	43.0
I ate some cheese	dog	eye	Unnamed	Unrelated	0
I saw an elephant	boat	clock	Unnamed	Unrelated	0
I saw a lion	nappy	spoon	Unnamed	Unrelated	0

List A2

Prime & Carrier	Target	Distracter			WA-Strength[a]
I bought a watch	bus	cow	Named	Unrelated	0
I saw a car	shoe	window	Named	Unrelated	0
I saw a door	moon	bed	Named	Unrelated	0
I saw a rabbit	duck	bread	Named	Related	1.2
I bought a cup	plate	bicycle	Named	Related	1.1
I saw a hand	foot	tiger	Named	Related	9.5
I bought a bowl	horse	cake	Unnamed	Unrelated	0
I saw an ear	banana	mouse	Unnamed	Unrelated	0
I saw a sheep	chair	coat	Unnamed	Unrelated	0
I caw a cat	dog	eye	Unnamed	Related	2.2
I saw a train	boat	clock	Unnamed	Related	2.1
I bought a bib	nappy	spoon	Unnamed	Related	2.1

Notes. Word Association norms from (Moss & Older, 1996), with supplementary values from (Postman & Keppel, 1970).

List A3

Prime & Carrier	Target	Distracter	Prime Cond.	Name Cond.	WA-Strength[a]
I saw a lorry	bus	cow	Unnamed	Related	2.4
I saw a boot	shoe	window	Unnamed	Related	30.0
I saw the sun	moon	bed	Unnamed	Related	11.1
I bought a cot	duck	bread	Unnamed	Unrelated	0
I bought a hat	plate	bicycle	Unnamed	Unrelated	0
I ate a biscuit	foot	tiger	Unnamed	Unrelated	0
I saw a pig	horse	cake	Named	Related	2.1
I ate an apple	banana	mouse	Named	Related	2.4
I saw a table	chair	coat	Named	Related	43.0
I ate some cheese	dog	eye	Named	Unrelated	0
I saw an elephant	boat	clock	Named	Unrelated	0
I saw a lion	nappy	spoon	Named	Unrelated	0

List A4

Prime & Carrier	Target	Distracter			WA-Strength[a]
I bought a watch	bus	cow	Unnamed	Unrelated	0
I saw a car	shoe	window	Unnamed	Unrelated	0
I saw a door	moon	bed	Unnamed	Unrelated	0
I saw a rabbit	duck	bread	Unnamed	Related	1.2
I bought a cup	plate	bicycle	Unnamed	Related	1.1
I saw a hand	foot	tiger	Unnamed	Related	9.5
I bought a bowl	horse	cake	Named	Unrelated	0
I saw an ear	banana	mouse	Named	Unrelated	0
I saw a sheep	chair	coat	Named	Unrelated	0
I caw a cat	dog	eye	Named	Related	2.2
I saw a train	boat	clock	Named	Related	2.1
I bought a bib	nappy	spoon	Named	Related	2.1

Notes. Word Association norms from (Moss & Older, 1996), with supplementary values from (Postman & Keppel, 1970).

List B1

Prime & Carrier	Target	Distracter	Prime Cond.	Name Cond.	WA-Strength[a]
I saw a car	bicycle	horse	Named	Related	3.6
I bought a hat	coat	banana	Named	Related	12.5
I saw a door	window	foot	Named	Related	2.1
I bought a bib	tiger	shoe	Named	Unrelated	0
I saw a lorry	spoon	chair	Named	Unrelated	0
I bought a table	eye	dog	Named	Unrelated	0
I saw an elephant	mouse	plate	Unnamed	Related	8.9
I ate some cheese	bread	duck	Unnamed	Related	1.2
I bought a watch	clock	boat	Unnamed	Related	2.2
I ate an apple	cow	bus	Unnamed	Unrelated	0
I saw a train	cake	nappy	Unnamed	Unrelated	0
I bought a bed	bed	moon	Unnamed	Unrelated	0

List B2

Prime & Carrier	Target	Distracter			WA-Strength[a]
I bought a cup	bicycle	horse	Named	Unrelated	0
I saw a pig	coat	banana	Named	Unrelated	0
I saw the sun	window	foot	Named	Unrelated	0
I saw a lion	tiger	shoe	Named	Related	26.1
I bought a bowl	spoon	chair	Named	Related	7.1
I saw an ear	eye	dog	Named	Related	2.1
I bought a boot	mouse	plate	Unnamed	Unrelated	0
I saw a hand	bread	duck	Unnamed	Unrelated	0
I saw a rabbit	clock	boat	Unnamed	Unrelated	0
I saw a sheep	cow	bus	Unnamed	Related	13.3
I ate a biscuit	cake	nappy	Unnamed	Related	4.8
I bought a cot	bed	moon	Unnamed	Related	9.5

Notes. Word Association norms from (Moss & Older, 1996), with supplementary values from (Postman & Keppel, 1970).

List B3

Prime & Carrier	Target	Distracter	Prime Cond.	Name Cond.	WA-Strength[a]
I saw a car	bicycle	horse	Unnamed	Related	3.6
I bought a hat	coat	banana	Unnamed	Related	12.5
I saw a door	window	foot	Unnamed	Related	2.1
I bought a bib	tiger	shoe	Unnamed	Unrelated	0
I saw a lorry	spoon	chair	Unnamed	Unrelated	0
I bought a table	eye	dog	Unnamed	Unrelated	0
I saw an elephant	mouse	plate	Named	Related	8.9
I ate some cheese	bread	duck	Named	Related	1.2
I bought a watch	clock	boat	Named	Related	2.2
I ate an apple	cow	bus	Named	Unrelated	0
I saw a train	cake	nappy	Named	Unrelated	0
I bought a bed	bed	moon	Named	Unrelated	0

List B4

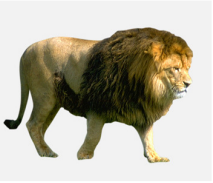
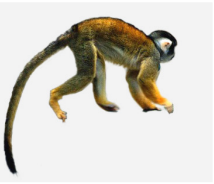
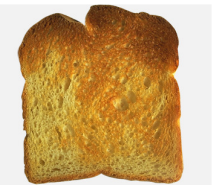
Prime & Carrier	Target	Distracter			WA-Strength[a]
I bought a cup	bicycle	horse	Unnamed	Unrelated	0
I saw a pig	coat	banana	Unnamed	Unrelated	0
I saw the sun	window	foot	Unnamed	Unrelated	0
I saw a lion	tiger	shoe	Unnamed	Related	26.1
I bought a bowl	spoon	chair	Unnamed	Related	7.1
I saw an ear	eye	dog	Unnamed	Related	2.1
I bought a boot	mouse	plate	Named	Unrelated	0
I saw a hand	bread	duck	Named	Unrelated	0
I saw a rabbit	clock	boat	Named	Unrelated	0
I saw a sheep	cow	bus	Named	Related	13.3
I ate a biscuit	cake	nappy	Named	Related	4.8
I bought a cot	bed	moon	Named	Related	9.5

Notes. Word Association norms from (Moss & Older, 1996), with supplementary values from (Postman & Keppel, 1970).

Appendix H: Visual Stimuli

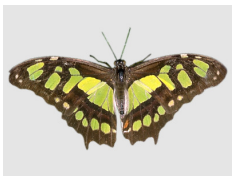
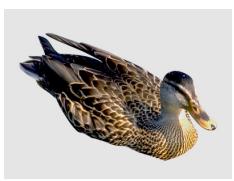
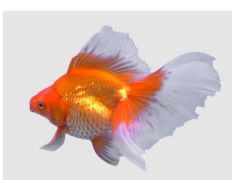
Minimally-Primed IPL (Chapter 3)

Lists A & B

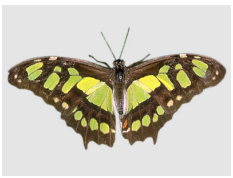
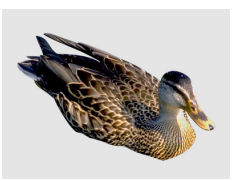
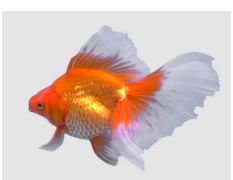
TARGET	DISTRACTER	TARGET	DISTRACTER
			
			
			
			
			
			

Words-before-pictures (Chapter 4)

List A

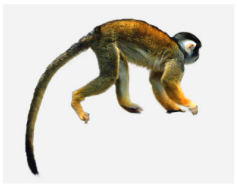
TARGET	DISTRACTER
	
	
	
	
	
	
	
	

List B

TARGET	DISTRACTER
	
	
	
	
	
	
	
	

Quick-Primed IPL (Chapter 5)

Lists A & B

TARGET	DISTRACTER	TARGET	DISTRACTER
			
			
			
			
			
			

.

Half Target Quick Primed IPL (Chapter 6)

Lists A1-A4

TARGET	DISTRACTER	TARGET	DISTRACTER
			
			
			
			
			
			

Lists B1-B4

TARGET	DISTRACTER	TARGET	DISTRACTER
			
			
			
	