Semantic and Pragmatic Language Development in Typical Acquisition, Autism Spectrum Disorders, and Williams Syndrome with Reference to Developmental Neurogenetics of the latter

by

Nadezhda N. Modyanova

B.A. in Linguistics with Cognitive Science (First Class Honours)
University College London, 2003

Submitted to the Department of Brain and Cognitive Sciences in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Neuroscience
at the
Massachusetts Institute of Technology

September 2009

© 2009 Nadezhda N. Modyanova. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature of Author: __________________________________________________________
Department of Brain and Cognitive Sciences
September 4th, 2009

Certified by: ________________________________________________________________
Kenneth N. Wexler, Ph.D.
Professor of Psychology and Linguistics
Thesis Supervisor

Accepted by: _______________________________________________________________
Earl K. Miller, Ph.D.
Picower Professor of Neuroscience
Chairman, Committee for Graduate Students
Semantic and Pragmatic Language Development in Typical Acquisition, Autism Spectrum Disorders, and Williams Syndrome with Reference to Developmental Neurogenetics of the latter

by

Nadezhda N. Modyanova

Submitted to the Department of Brain and Cognitive Sciences on September 4th, 2009 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Neuroscience

Abstract

The elucidation of the biological bases of a complex trait like human language proceeds from identification of precise behavioral phenotypes to investigation of the underlying genes.

The human behavioral parts of this dissertation focus on understanding the reasons for children’s overuse of definite article ‘the’, to refer to one of several objects in a context set, as opposed to the unique established referent. Competing theories argue the deficit is either in children’s semantic computational knowledge (of uniqueness/maximality), or in their pragmatic/social awareness/theory-of-mind development. Experiments in this dissertation focused on children’s comprehension and interpretation of the indefinite and definite determiners, as well as ‘that’, anaphors ‘another’ and ‘same’, and free relative clauses.

The results in this thesis suggest that in typically developing (TD) children the late acquisition of determiner ‘the’ is due to the late maturation of the semantic principle of maximality. Children with autism spectrum disorders (ASD) and with Williams syndrome (WS) either manifested an adult-like competence, an absence of manifestation of knowledge, or a pattern found in TD younger children (where ‘that’ is understood better than ‘the’ as referring to the salient unique referent) – indicating delay of development of the language faculty, but no deviance. This suggests that the observed deficits in ASD and WS pattern with those in TD, and hence are also semantic in nature.

The mouse neurogenetic part of this dissertation investigates whether the GTF2I family of genes, causal to WS behavioral phenotype, also contributes to WS cortical development. By overexpression of Gif2i and Gif2ird1 in the mouse neocortex via in utero electroporation, their effects on laminar patterning and cell morphogenesis during brain development are characterized.

The present results suggest that these genes can synergistically contribute to the abnormal neocortical development in WS, and thereby could contribute to language deficits in WS.

Beyond posing an explanatory challenge to linguistic theories, the research comparing typical and atypical development sheds light on the mechanisms of language development and impairment, and provides endophenotypic descriptions of ASD and WS, which are crucial for elucidating not only genetics of neurodevelopmental disorders, but also the genetic basis of the human language faculty.

Thesis Supervisor: Kenneth N Wexler
Title: Professor of Psychology and Linguistics
Thank You!

My Parents

Advisors:
Ken Wexler, Damon Page

Committee members:
Irene Heim, Mriganka Sur

Colleagues and Collaborators:
Alexandra Perovic, Christopher Hirsch, Tania Ionin, Charlotte Giessman, Vanja de Lint, Ioulia Kovelman, Pritty Patel, Lee Mavros, Margaret E. Echelbarger, and others

All the wonderful undergraduate research assistants at Wexler Lab

All participants in our language development studies - children, their parents, and their teachers
# Table of Contents

1. **Introduction** 8

2. **Background on Semantics and Pragmatics and their Development** 11
   2.1. Semantics and Pragmatics of Determiners 11
   2.2. Predictions for Acquisition 14
   2.3. Acquisition studies of definite determiners 18
      2.3.1. Spontaneous production studies 19
      2.3.2. Elicited production studies 19
      2.3.3. Theoretical explanations for patterns in production studies 21
      2.3.4. Comprehension studies 25
      2.3.5. Theoretical explanations for patterns in comprehension studies 27
      2.3.6. Summary 29
   2.4. Maximality hypothesis: semantic deficits in children’s ‘the’ 29
   2.5. Acknowledgements 31
   2.6. References 32

3. **Comprehension of Determiners in Typically Developing Children** 35
   3.1. Abstract 35
   3.2. **Experiment 1. “Puppets”** 35
      3.2.1. Method 35
         3.2.1.1. Context size variable 36
      3.2.2. Participants 37
      3.2.3. Results 38
         3.2.3.1. Random model in children’s performance 39
         3.2.3.2. Analysis of children performing well 40
         3.2.3.3. Effect of context size on interpretation of articles 41
   3.3. **Experiment 2. “Felt”** 44
      3.3.1. Method 44
      3.3.2. Participants 45
      3.3.3. Results 46
   3.4. **Experiment 3. Replication of “Felt” with Addition of ‘Same’, ‘Share’** 49
      3.4.1. Method 49
      3.4.2. Participants 49
      3.4.3. Results 49
   3.5. **Analysis. “Felt” – entire sample** 52
      3.5.1. Fine-grained analysis by half-ages 52
      3.5.2. Random model in children’s performance? 55
      3.5.3. Gender effects? 59
      3.5.4. Context effects? 61
      3.5.5. Context effects vs patterns of performance 63
   3.6. **Discussion of TD Results** 70
3.7. Acknowledgements

4. Further Investigations of Typical Acquisition
   4.1. Experiment 4. Theory of Mind (Not) in Determiner Acquisition
      4.1.1. Abstract / Motivation
      4.1.2. Language in theory of mind?
      4.1.3. Determiners and theory of mind?
      4.1.4. Method
      4.1.5. Participants
      4.1.6. Results
      4.1.7. Discussion
      4.1.8. Acknowledgements
      4.1.9. Appendix: details of procedure
      4.1.10. References

4.2. Experiment 5. Comprehension of Maximality in Free Relatives
   4.2.1. Abstract
   4.2.2. Background on Free Relatives
   4.2.3. Method
   4.2.4. Predictions
   4.2.5. Participants
   4.2.6. Results
   4.2.7. Discussion
   4.2.8. Acknowledgements
   4.2.9. References

4.3. Discussion

5. Comprehension of Determiners in Autism Spectrum Disorders
   5.1. Abstract
   5.2. Introduction to ASD
   5.3. Semantics and Pragmatics in ASD
      5.3.1. Semantics in ASD
      5.3.2. Pragmatics in ASD
      5.3.3. Theory of mind in ASD
   5.4. Why study determiners in ASD?
   5.5. Participants
   5.6. Methods
   5.7. Results
      5.7.1. Some issues on grouping participants
      5.7.2. Results of All ASD participants
      5.7.3. Grouping by Diagnosis & Endophenotypes within clinical groups
      5.7.4. Grouping by Determiners
      5.7.5. Grouping by IQ
5.7.6. Grouping by Vocabulary 115
5.7.7. Grouping by Grammar 116
5.8. Discussion 118
5.9. Acknowledgements 120
5.10. References 121

6. Williams syndrome 124

6.1. Comprehension of Determiners in Williams Syndrome 124
   6.1.1. Abstract 124
   6.1.2. Background on WS 124
   6.1.3. Semantics and Pragmatics in WS 125
      6.1.3.1. Semantics in WS 125
      6.1.3.2. Pragmatics in WS 126
      6.1.3.3. Theory of mind in WS 127
      6.1.3.4. ASD vs WS 127
   6.1.4. Predictions 128
   6.1.5. Participants 129
   6.1.6. Results 130
   6.1.7. Discussion 133
   6.1.8. Acknowledgements 133
   6.1.9. References 134

6.2. Initial Characterization of GTF2I family of Gene Homologues in Mouse Neocortical Development 136
   6.2.1. Abstract 136
   6.2.2. Background on WS region genotypes and neurological phenotypes 136
   6.2.3. Background on the GTF2I family of genes 139
   6.2.4. Aim of the present study 141
   6.2.5. Methods 142
   6.2.6. Results 144
   6.2.7. Discussion 148
   6.2.8. Acknowledgements 149
   6.2.9. References 149

7. Global Discussion and Conclusion 155
   References 160
1. Introduction

When it comes to development of definite determiners in children, what is it that is responsible for driving the developmental trajectory, what is it that makes children use ‘the’ instead of ‘a’ for a salient referent in the context set? Across multiple theories explaining the phenomenon, two camps are evident. The more predominant camp argues it is a problem of the correct use of the ‘the’, i.e. children have adult semantic knowledge, however children do not yet have correct pragmatic interpretations, which primarily have to do with distinguishing between speaker/listener knowledge. There is certainly evidence that pragmatics, e.g. correct interpretation of personal pronouns (Chien, Wexler, 1985), and of scalar implicatures (e.g. Guasti, Chierchia, Crain, Foppolo, Gualmini, Meroni (2005)) are delayed in children. The smaller camp argues it is a problem in children’s knowledge of the semantic principles that guide the correct use of ‘the’, specifically that of Maximality/Uniqueness (Wexler 2003), i.e. it is just another stage that maturation of linguistic principles goes through, on par with e.g. development of A-chains (phases) (Hirsch, Wexler 2006; Wexler 2004; Borer and Wexler 1987).

This thesis takes on the challenge of answering this question and suggesting that specific deficiency in ‘the’ in children is a manifestation of a specific semantic deficit. Supporting evidence comes not only from determiner comprehension experiments from hundreds of typically developing children, but also from children and adolescents with neurodevelopmental disorders and language impairments: autism spectrum disorders and Williams syndrome.

The next question posed herein is, what is the underlying genetic basis of human language faculty, and to what extent can investigating functions of genes in animal models shed light on the biology of human language. Current work by Chomsky and colleagues (Hauser, Chomsky, Fitch 2002 and Fitch, Hauser, Chomsky 2005) formulates hypotheses for investigations. The human language faculty certainly involves sensory-motor (perception of and reaction to external environment) and conceptual-intentional (planning, attention, control of behavior) systems which we share with other species. The question is whether recursion (ability to splice together lexical items and phrases to generate sentences) is uniquely human, unique to language, unique to language in humans, or a specific organization of otherwise existing capacities.

This thesis investigates the influence of two Williams syndrome genes on brain development using mouse models. Brain morphology in Williams syndrome is known to be abnormal, with altered white and grey matter volumes, which some have argued to stem from deficits in laminar patterning. This thesis is the first to directly investigate synergistic effects of Williams syndrome genes on brain development. It is presently impossible to make a direct and transparent connection between language deficits (not only in Williams syndrome but in other disorders as well) and abnormalities in laminar patterning, but it is certain that such abnormalities can underlie linguistic and cognitive deficits.
**Thesis Outline**

Chapter 2 discusses semantic and pragmatic properties of determiners, goes over logical predictions based on those properties, then reviews some of the child language acquisition studies of determiners, and discusses the theoretical explanations for children’s deficits.

Chapter 3 presents acquisition data on comprehension of determiners and anaphors in typically developing children.

Chapter 4 presents lack of correlation between acquisition of determiners and development of theory of mind, and presents further evidence for delayed development of Maximality in domains outside of definite determiners: in free relative clauses.

Chapter 5 presents acquisition data on comprehension of determiners in children and adolescents with Autism Spectrum Disorders in comparison to typically developing children.

Chapter 6 in part one presents acquisition data on comprehension of determiners in children and adolescents with Williams syndrome in comparison to those with Autism Spectrum Disorders and to typically developing children.

Chapter 6 in part two presents data on the effects of two Williams syndrome genes on embryonic brain development - on laminar patterning and neurite morphology - in mice.

Chapter 7 concludes.

**Publications**

The work presented in section 3.2 was previously published in conference proceedings of GALANA2 - Generative Approaches to Language Acquisition - North America in 2006. The results in 4.2 were previously published in the Proceedings of the 32nd Annual Boston University Conference on Language Development in 2008. The results in chapter 5 was presented in posters at the 28th Annual Symposium on Research in Child Language Disorders at University of Wisconsin-Madison, and also at the 2nd Annual Autism Consortium Retreat, Broad Institute of MIT and Harvard, both in 2007.

Chapters 2-3, and sections 4.1, 5, 6.1, and 6.2 are in process of submission as individual journal articles.

**Advising/Collaborations**

Work in chapters 2-6.1 ensued under guidance from Ken Wexler. Experimental details in 4.1 were worked out with Charlotte Giessman. The idea for 4.2 was suggested by Tania Ionin and was elaborated with significant help from Irene Heim. The participants in 5 and 6.1 were recruited in collaboration with Alexandra Perovic. Collection of all human behavioral data was performed by the author in collaboration with over 40 undergraduate research assistants. Collection of data from participants with autism spectrum disorders was especially made possible by assistance from Lee Mavros.

Work in chapter 6.2 ensued under extensive guidance from Damon Page.
References
2. **Background on Semantics and Pragmatics and their Development**

We first discuss semantic and pragmatic properties of determiners, go over logical predictions based on those properties, then review some of the child language acquisition studies of determiners, and discuss the theoretical explanations for children’s deficits that have been proposed.

### 2.1. Semantics and Pragmatics of Determiners

We start with semantic definitions – namely those of ‘the’ and ‘a’, following Heim (1991).

The formal semantic definition of the definite article ‘the’ (1) contains a uniqueness presupposition – that there is exactly one referent in a given context set (for example one where there are several identical Xs that may be distinguished by various means, in order to make one X unique/salient). In plural cases, uniqueness generalizes to maximality, where ‘the objects’ must refer to the entire, maximum set of objects in the context and not a subset. Maximality is uniqueness, in a sense: a singleton set is the maximal element in a set of atoms – a unique set of atoms.

The formal definition for the indefinite article ‘a’ (2) only suggests presence of at least one referent, without a context set, and may in fact be used in the absence of a referent.

To differentiate between uses of ‘a’ and uses of ‘the’, Maximize Presupposition Implicature forces use of ‘the’ instead of ‘a’ in case of singleton set. This is what makes it incorrect to say “a biological mother of mine…”.

1. Adult’s lexical entry for ‘the’ (based on Heim 1991):
   - Regardless of the utterance context (i), 
     \[ [\text{the } x] \ P \] expresses that proposition that is:
     - true at an index i, if there is **exactly one** \( x \) at i, and it is \( P \) at i
     - false at an index i, if there is exactly one \( x \) at i, and it is not \( P \) at i
     - truth-valueless at an index i, if there isn’t exactly one \( x \) at i.

2. Adult’s lexical entry of ‘a’ (based on Heim 1991):
   - A sentence of the form \([a \ x] \ P\) expresses that proposition which is true if there is at least one individual which is both \( x \) and \( P \), and false otherwise.

It turns out that the English language has a word that presents an excellent minimal contrast, at least intuitively, to the definite determiner ‘the’ - the demonstrative determiner ‘that’. Consider three examples to illustrate the differences between ‘a’ (3), ‘the’ (4), and ‘that’ (5).

3. *A guy in my class is a genius* \( \rightarrow \) Implies there is a smart male person in the set ‘my class’.

4. *The guy in my class is a genius* \( \rightarrow \) Implies there is a unique male in the context set ‘my class’, i.e. there is only one male person in ‘my class’, and is infelicitous if there are several or no males in ‘my class’.

5. *That guy in my class is a genius* \( \rightarrow \) Implies that the listener knows about a particular male person from previous discussions with the speaker. There is not necessarily a unique male in the physical world, but there is **unique, familiar** referent in **shared world-knowledge** (from prior conversations) between speaker and listener.
Thus the use of ‘that’ not only implicates uniqueness of the referent (just like ‘the’) but also requires the speaker to be aware of the listener’s state of mind when it comes the identity of the referent of ‘that’. A number of theories have been proposed to account for demonstratives and definites, their similarities and their differences.

Maclaran (1982) discusses the differences between ‘the’ and ‘that’, the key one being that while it is enough that the referent of the definite article is inferred, the referent referred to by ‘that’ must be given in the (extra)-linguistic context, it is not enough for it to be inferred. It seems that such nondemonstrative use of ‘that’ carries a number of presuppositions on the referent, which are minimally different from the definite article. ‘That’, according to Maclaran, carries a familiarity presupposition, such that both the speaker and the listener must be familiar with the referent, presumably from previous conversations. This familiarity is what makes the referent especially salient. ‘That’ also has a uniqueness presupposition, although effective not in the physical world, but in the metalinguistic knowledge space (context) shared by the speaker and the listener. Thus, while ‘the’ does not necessitate one to rely on awareness of other minds, but only on their knowledge of uniqueness and of how uniqueness applies to the context set of the utterance, ‘that’ necessitates knowledge of uniqueness and requires awareness of the context set knowledge shared by the speaker and the listener.

More recent accounts try to deal with the definite-demonstrative difference more concretely. Thus King (2001) suggests that demonstratives have an extra argument which is saturated with a speaker intention to refer. Roberts (2002) suggests that demonstratives come with an additional presupposition, namely that of accompanying speaker demonstration.

Wolter (2006 a,b) suggests reuniting semantics of definites and demonstratives, and pushing the differences into the land of domain restriction - the interpretation of nominal predicates relative to a situation (actual or possible world situation). Wolter assumes the situation-based domain restriction, where each nominal predicate enters compositional semantics with an extra argument position which is satisfied by the value of the situation variable for a particular possible world interpretation. ¹ Both ‘the’ and ‘that’ have uniqueness presupposition on the referent. Wolter argues that the condition of familiarity in itself is not necessary. The difference lies in the relative domain of interpretation that the listener must take to identify the referent – a pragmatic distinction. Wolter argues ‘the’ (definite descriptions) is interpreted relative to default situations, i.e. those associated with main predicates and corresponding to the world of discourse context, while ‘that’ (demonstrative descriptions) are interpreted relative to other salient situations, which correspond to a speaker demonstration of other salient eventuality and/or situation, or a state of experiencing the knowledge or emotion shared by discourse participants, but not necessarily identical to the discourse context (formalized in 6). Thus demonstrative, unlike definite, restricts the “value of the situation argument of their nominal complement” (Wolter 2006a:62). Wolter argues that definite ‘the’ is less marked, in the sense that it only has a uniqueness presupposition, whereas demonstrative ‘that’ is more marked – it has uniqueness plus triggers non-default situation interpretation. Thus ‘the’ tends to be preferred, unless ‘that’ is required to enhance interpretation.

¹ This is situation-based approach. The effects of domain restriction are tied to the modal parameter of the nominal predicate. Different predicates in one clause may be interpreted relative to different worlds. So each predicate enters the compositional semantics with its own modal parameter which determines which possible world the predicate is interpreted at. Predicate – has extra argument position saturated with a situation variable. Each predicate has its own situation index. Possible world = maximal situation.
Given a sentence A, a situation variable s is a default situation just in case it is bound in A. Otherwise s is a non-default situation.

Both definite and demonstrative can be used anaphorically – referring to linguistic antecedent (established by e.g. an indefinite description), also known as bridging. It is relevant to review here Avrutin (1999) and Heim’s (1982) File Change Semantics: “as a model of the discourse representation of NPs. An indefinite NP in this theory is represented in the discourse with a new file card, while a definite NP does not introduce a new file card (normally) but is incorporated into an existing one”. Avrutin (1999:37) suggests that “Indeed, the smaller the file (that is, the fewer cards it contains), the easier it is to maintain it (that is, keep record). It is plausible, therefore, that the economy (information-processing) considerations result in the following constraint on file change: Avoid introduction of new file cards, unless it is required by syntax and/or discourse.” It follows that definite NPs always have an antecedent. The use of definite NP without a previously linguistically established antecedent is ony natural in cases where the uniqueness of the referent is common knowledge. “Interestingly, this use depends on some shared knowledge, or at least on the assumption that some relevant knowledge is shared both by the speaker, and the listener.” Avrutin (1999:51) further argues that “speakers cannot felicitously use [such] definite NPs without making certain inferences about other speakers' knowledge.”

In summary, incorporation of a definite NP into discourse happens via explicit bridging or accommodation (inferential bridging), i.e. via bridging definite NP file card with some existing card. Notably, children as young as 4 years old (in Dutch, English, Russian) have adult-like interpretation of inferential bridging (argued to be implicit domain restriction) (Avrutin & Coopmans (2000)).

Consider the following examples.

(7) A car stopped in front of a house. The door was open.
(8) (on a busy street) A car stopped at the red light. That car then took a left turn.
(9) (on a busy street) A car came to screeching stop. That car came so close to running the red light!

In inferential bridging (associative anaphora (7)), ‘the door’ is taken to refer to that of the house, since a car as at least two doors, and hence has no unique door. In explicit bridging via an anaphoric definite (relational anaphora (8)), ‘a car’ is used to establish a unique referent, with

---

2 AVRUTIN (1999:45) Rules of NP Representation in Discourse
1. Instantiate the variable index of an Indefinite NP with a number of a new file card.
2. Instantiate the variable index of a definite NP with a number of an old file card.
3. Instantiate two identical variable indices with the same number, and two different indices with different numbers.
4. Instantiate the variable index of a definite NP with a number of a new file card only if this card can be bridged to another one.
‘the car’ referring back to the unique referent. Anaphoric demonstratives (9) (as Wolter argues, 2006a:74) “can be understood as zooming in on a salient subpart of the discourse context”, hence the necessity of emotive or contrastive use. “Anaphoric demonstratives can also refer to an individual who has been previously mentioned or encountered by the interlocutors but who might have been subsequently forgotten”, and (2006a:75) “anaphoric demonstrative description can be understood as referring uniquely relative to a salient situation that does not correspond to the entire unadjusted discourse context. … anaphoric definite descriptions are relativized only to the situation corresponding to the entire discourse context. Definite descriptions may not zoom out on the discourse context to include potentially forgotten individuals.”

It is also necessary to discuss referential elements which are intrinsically anaphoric – elements whose interpretation is intrinsically dependent on a linguistic antecedent. These include personal and reflexive pronouns (not the topic of current investigations), and anaphors ‘another’ and ‘same’. Referents of ‘another’ and ‘same’, just like anaphoric definite and demonstrative referents, crucially depend on existence of a unique salient referent in the discourse situation. ‘Same’ refers to that unique referent, essentially is an equal sign to the existing file card. ‘Another’ refers to a referent that is NOT the unique salient referent – something that is like that salient referent in property/kind, but is not that unique referent, but is any other one. Both of these are best used when there is a set of several similar referents, and one feature distinguishes the salient referent from non salient referents. Thus both ‘same’ and ‘another’, in a sense, can serve the role of definite and demonstrative descriptions – both require interpretation relative to the salient entity in the discourse set.

(10) Another guy in my class is a genius → somewhat weird without antecedent. A logical interpretation – there was one smart annoying person, and now there is a second one.

(11) The same guy in my class is a genius → somewhat weird without an antecedent. A logical interpretation – at some point in time there was a male genius, and now, contrary to expectations, that same male person continues to be a genius.

A similar anaphoric notion is demonstrated by this example from Russian military folklore. The speaker of the sentence is clearly a utility pole.

(12) Soldiers! You must march from me to the next utility pole!

2.2. Predictions for Acquisition

At this point, a good exercise would be to consider what it could mean to have difficulty with interpreting indefinites, definites, demonstratives, or anaphors, before looking at the acquisition data. It is possible to not have yet acquired the semantic definitions of all those determiners. Thus if children do not know that uniqueness is part of definition of ‘the’ or ‘that’ – we expect trouble. If children do not know that there must be a salient referent in the discourse set – we expect trouble. If children assume ‘a’ means something more than it does – we expect trouble.

Another alternative is that it is possible to not have yet realized the precise interaction of determiners with context, and thus difficulty with setting the appropriate situation for an interpretation, which could be due to domain restriction deficits or deficits in realizing that shared speaker/hearer knowledge is relevant to interpretation. If children do not know that ‘that’ triggers
interpretation relative to a non-default situation, or do not know about the possibility of interpretation relative to a non-default situation, children may not differentiate ‘the’ from ‘that’.

It is also possible that if children do not know about uniqueness, but are more sensitive to non-default situations of interpretations, that pragmatic knowledge may help children interpret ‘that’ even if they lack semantic uniqueness.

Note, that if Maximize presupposition implicature (i.e. there is a scale where The>A) is similar to other implicatures (e.g. ‘some’ implies ‘not all’, with a scale where All>Some), then we can expect, given studies of children’s difficulties with implicatures (Guasti, Chierchia, Crain, Foppolo, Gualmini, Meroni (2005) who show that from age 7 kids can do scalar implicatures, i.e. pragmatic inferencing (as opposed to logical)), that children will have problems with using ‘a’ correctly, and that children will think that, in some sense, ‘a’ means ‘the’.

Thus deficits in determiner interpretation could be in semantics, or pragmatics, or both.

Another good exercise is to think about what it means to have a ‘pragmatic’ deficit, or rather a ‘pragmatic processing’ deficit (as opposed to a semantic computational one). One of the few articulated proposals is that based on Grodzinsky and Reinhart (1993), and it certainly can work in accounting for children’s deficits in scalar implicatures – which people have argued is a pragmatic phenomenon. In scalar implicatures, and perhaps in referential descriptions, a reference set computation is necessary, which requires temporarily holding two or more alternative interpretations / derivations in working memory at once. Children with underdeveloped linguistic working memory may fail to hold two interpretations at once, and guess. Reinhart (to appear:10) argues that intra-subject 50% (chance) performance (as seen on apparent delay of principle B errors in children) is indication of such working memory difficulties.3 In other words, difficulty in pragmatics in children may simply be the issue of finding the right match between the context set and the right proposition. In other words, children’s pragmatic difficulties in scalar implicatures have nothing to do with children’s lack of awareness of their listener’s states of minds, but have everything to do with children’s own underdeveloped capacity for finding the right interpretation for a given context. Note that all (as far as I know) scalar implicature experiments test children’s comprehension and children’s preferred interpretations. Why do people not study production of scalar implicatures? Because in production it is hard to distinguish between the two possibilities of pragmatic deficits – those due to limited working memory / referential set comparisons, and those due to limited awareness/calculation of listener knowledge of the discourse. It is only by turning the child into a listener and focusing the child’s attention on their own interpretations that it becomes possible to distinguish the two. If pragmatic difficulties are due to lack of knowledge of listener’s state of mind – those should only be present in production. If pragmatic difficulties are due to having to juggle multiple interpretations at once – those will affect comprehension, and may as well affect production.

For such working memory issues, children may prefer to be economical, and may limit the number of cards in the discourse (according to Avrutin’s interpretations of File Card Semantics). It is more economical to have the same referent, rather than add a new one – otherwise it is necessary to hold multiple cards in mind. This view would expect for children to pick salient referents in context sets across the board, e.g. picking reflexive referent and not pronoun referent, and doing so also for ‘a’ and ‘another’ (which would lead to incorrect

——

3 Footnote on scalar implicatures: If marking on a predicate encodes neutral information that the predicate is underspecified… Build a set of alternatives, conjunct sentence with negation of alternatives that entail the sentence. AND>OR. ALL>some.
interpretations), and also for ‘the’ and ‘that’ (which would lead to correct interpretations) – but this is not what we see in acquisition (chapter 2.3). Wexler (2003) argues that this view of ‘pragmatic’ deficits as working memory limitations does not seem to make correct predictions in the acquisition of determiners – children should make mistakes with ‘a’ and not with ‘the’ – which is not the case in acquisition.

A very relevant question, which so far was not addressed, concerns the psychological reality of determiners, not just their linguistic reality. When we are discussing non/default situations, uniqueness, saliency, familiarity, the obvious question (Heim, p.c.) is “who are the relevant minds to whom the entities in these contexts are supposed to be salient or familiar. The speaker? The listener? Both? This question is crucial for an assessment of whether the production or comprehension of ‘the’ and ‘that’ involves reasoning about another mind, and if so, what kind of reasoning.”

Many authors assume that the burden of communication lies on the speaker, and not the listener. The listeners are merely attentive processors. Thus in English, which bins article space by definiteness (speaker and listener knowledge) and indefiniteness (no listener knowledge) (as opposed to binning article space by specificity (speaker knowledge)), producing an indefinite implies no listener knowledge of referent (in naming or introducing referents), and a producing definite implies listener knowledge of referent (of already-established salient referent, physical contextual information, or it implies that listener can infer the referent given their knowledge). Listener’s interpretation should be guided solely by their own attentional state. Thus while producing determiners can be a complicated task – one always has to worry about one’s listener’s point of view of the context and their knowledge of salient and unsalient referents, comprehension of determiners should be more straightforward, that is if one has good knowledge of semantic presuppositions of determiners. If you know that ‘the’ means ‘unique referent’, but you don’t know about your listener’s state of mind or you don’t know that you have to be aware of your listeners’ state of mind, you’ll be making mistakes in producing utterances that point out salient-to-you but unsalient-to-the-listener referents. In comprehension however, you’ll know automatically as soon as ‘the’ is uttered – that’s the salient referent in the context set – no second thoughts. I.e. it seems that difficulties with understanding other minds should not cause errors in comprehension. But if you don’t know semantic presuppositions of determiners, or have some non-adult form of such presuppositions, you will clearly show this lack in comprehension, unless you have some canny strategy of interpretation. Similarly in production, you may show the non-adult semantics by referring to referents in a non-adult way, whether or not you are or are not aware of necessity to be aware of other minds, unless you have some canny strategies.

If, on the other hand, the listener is not passive and only follows their own attentional state, but is actively monitoring the speaker and their shared conversational history, then, as in, e.g. Relevance theory (Wilson and Sperber 2002), listeners are actively interpreting and guessing

4 An utterance (U) is relevant if and only if it carries enough cognitive effects to balance the processing effort. (Cognitive effects occur when new information combines with existing assumptions to produce new conclusions, or to strengthen or weaken and eliminate existing assumptions).

It follows from the Communicative Principle that every U carries an expectation of optimal relevance, that is U will create enough cognitive effects to offset the processing effort required of H (Hearer) to process it, and that S (Speaker) will produce U which best reflects S’s abilities and preferences. This presumption of optimal relevance has two important
at what the speaker meant. The listeners are active interpreters. The listener would reason out (unconsciously of course), that since the speaker said ‘a’, they meant a new or any referent, because if they were referring to an old or a unique referent, they would have said ‘the’. Active listeners thus are applying just as much semantic and pragmatic and other-mind knowledge to comprehension of determiners as are speakers in producing them. Active listener’s working memory may also be continually taxed with reasoning out the possible interpretations of the speaker’s utterance. Again, if listeners are deficient in either semantic or pragmatic aspects of determiner interpretation, they may incorrectly infer the speaker’s intentions. If listeners have non-adult semantics, they may assume their speaker also has non-adult semantics, and happily follow their own (mis)interpretations. If listeners lack in awareness of other minds, they may not even think about communicational relevancy of the speaker’s utterance, ignore it altogether, or just go with what they as a listener would mean if they had produced such an utterance – from their own point of view, which may be drastically different from the intended point of view. If listeners have working memory issues, they may go with the first interpretation that comes to mind, which may or may not be the right one. The question is, do such active listeners, when they take on the role of speakers, take on the burden of producing the more relevant utterance or would they be sloppy since they know that the other person would figure it out if they cared to? For Economy principles, and that is what Relevance theory argues in a way, it is shameful to waste resources needlessly, unless it necessary to make a point. It is not economical for speakers to be sloppy since listeners would be expanding extra resources needlessly. It is not economical for speakers to take on all the communicational burden since listeners would be spending energy making sense of things any way. Is it possible that people speak in a way that makes sense for them (a result of interaction between their world knowledge and the semantics and pragmatic rules of their language) and as a result they are understood by listeners (as long as they have the same world knowledge and the same semantic and pragmatic rules of the same language as the speaker)?

The catch in all this theorizing, is that I don’t know of a way to test whether listeners are just attentive processors or active interpreters.

The main point is that if one has deficient semantic knowledge, one should make comprehension mistakes (unless one’s pragmatic knowledge (and awareness of speaker’s intentions, mediated e.g. in ‘that’ by interpretation relative to non-default situations) helps to overcome those deficits). The other main point is, if one has deficient pragmatic/other mind knowledge only, one should make production mistakes and speak as if from one’s own point of view, but no comprehension mistakes. If one is deficient in pragmatics/working memory (but not in semantic/pragmatic definitions), and if indeed referential set computations are involved in determiners as they are involved in scalar implicatures, one should make comprehension mistakes and show chance performance.

consequences: H can assume that S has not forced H into spending gratuitous effort and that every extra bit of effort will be balanced by extra cognitive effects; the second is that the first interpretation that H is satisfied with will be the only one H will come up with because H will assume that whatever change in H’s cognitive environment was intended by S to occur as a result of U will occur with least effort possible. (Modyanova 2002, unpublished term paper on Relevance-Theoretic interpretation of Tautologies).
Can we distinguish then working memory deficit from semantic knowledge deficit in comprehension? I would like to argue that yes, it is possible. Just like in apparent delay of principle B effect, children misinterpret personal pronouns when the utterance can have two interpretations, but children perform perfectly when only one interpretation is possible – namely when there is a quantifier present. This result shows that children do know principle B, but for pragmatic reasons (working memory (Reinhart, to appear) or principle P (Chien, Wexler, 1990) are being confused when having to chose one of TWO possible interpretations.

Similarly, I would like to argue, that when it comes to determiners and descriptions of unique entities, working memory/pragmatic and semantic deficits can be disambiguated. Consider two determiners – ‘same’ and ‘the’. Both refer to the salient entity in the context set, and hence may require reference set computation to figure out which is the salient entity in the context set. Semantically, ‘same’ is an equal sign to the already-established salient referent, ‘same’ is anaphoric to the salient antecedent. Semantically ‘the’ is not an equal sign, it presupposes uniqueness and existence of the referent in discourse set, but the referent may be inferred, and ‘the’ can be used anaphorically. If children are missing uniqueness (Wexler 2003), semantically ‘child-the’ indicates existence of a referent in discourse set, but the referent now no longer has to refer to the unique salient one. Just one of the salient referents will be good enough. Continuing this logic, if children’s difficulties are with working memory/pragmatic interpretations (but not semantics) – children will have problem with comprehending both ‘same’ and ‘the’ (since both involve reference set computations, and paying attention to the salient referent in the context set) in a similar way. If children’s difficulties are specifically with semantics of ‘the’ (but not with working memory/pragmatics), children will do fine on ‘same’, but will continue to have problems with ‘the’ until that time when their semantic knowledge becomes adult-like.

Let’s consider another contrast – between ‘the’ and ‘that’. Both determiners involve uniqueness semantically and both require reference set computations (maybe). The difference is that, according to Wolter (2006a,b), ‘that’ has an extra feature which calls for interpretation relative to non-default discourse situation. If children are missing uniqueness, but are sensitive to non-default interpretations (for example shared speaker/hearer knowledge), those children may do better on ‘that’ than on ‘the’. If children are not missing uniqueness, but have working memory issues, will they do better on ‘that’ than on ‘the’ since ‘that’ is more marked and draws attention to the salient subpart of the context? Well, maybe, but maybe not, since still, children would have to figure out which is the salient subpart of the context – it’s almost an extra step in computation according to Wolter (2006a,b) – figuring out which is the relevant subpart of the context. If anything, is it possible to predict that children who know uniqueness but have working memory issues will perform better on ‘the’ than on ‘that’? Answer to this question is beyond the score of the present work, but may be answered by investigating comprehension of determiners in children with Dyslexia who are argued to have working memory deficits (e.g. in Fiorin 2009). Another question beyond the scope of present work - Do children learn and apply more marked things better than less marked things?

**2.3. Acquisition studies of definite determiners and evolutions of theoretical explanations**

Children’s acquisition of determiners has been addressed in many studies. A comprehensive review is beyond the scope of the present section, but a few notable studies that illustrate acquisition are reviewed (see also Hickman 2003 for extensive reviews and alternative arguments). While children produce determiners in spontaneous speech by 3 years, adult-like
competence and performance does not come in until later – as measured by production and comprehension tasks.

### 2.3.1. Spontaneous production studies

Brown’s natural production data (1973:355) suggests that “children somewhere between the ages of 32 and 41 months, roughly three years, do control the specific/non-specific distinction as coded by the articles”. However, as Karmiloff-Smith writes “Brown nonetheless adds the proviso that this early productive control of the article contrast does not yet cover instances where the child is obliged to take into account his listener’s knowledge” (Karmiloff-Smith 1977:378).

Analysis by Gundel and colleagues of Brown’s corpus similarly suggests that 3 year old children use all possible determiners, and assuming that speakers take on the entire burden of communicative relevance, Gundel argue that any successful production of determiners is clear evidence toward children’s full semantic and pragmatic competence and full awareness of other’s minds. (Gundel 2009). A key assumption made by this work is that full competence is necessary for production, and that spontaneous production is evidence for full competence – that may not be the case.

Abu-Akel & Bailey (2000) analyze spontaneous speech of 17 children ages 1.5-4.8 from the Bristol corpus on CHILDES, by classifying determiners into exophoric (identifiable from situation, extralinguistic grounds, but not prior discourse) and endophoric (identifiable from linguistic discourse, anaphoric) uses, into definite and indefinite, and into specific vs nonspecific. In other words, Abu-Akel and Bailey make an effort to explicitly investigate the contexts in which children produce determiners – something that e.g. Gundel did not do so explicitly (as far as I know). 50-60% of produced DPs were exophoric, i.e. identifiable from immediate context. Over time, Abu-Akel and Bailey found increase in definite-exophoric DPs and a decrease in indefinite-endophoric (i.e. pragmatically inappropriate) DPs. The reverse pattern remained constant.

Thus while children may spontaneously PRODUCE determiners early on, they may not produce them in correct contexts. It is not enough to investigate spontaneous production – it is necessary to perform experiments where contexts of utterances can be explicitly and unambiguously investigated and controlled for.

### 2.3.2. Elicited production studies

The logic of elicited production studies is to set up a context where the children have opportunity to refer to either unsalient referents (one that are just being introduced into discourse set, ‘a’ use expected) or salient referents (already established referents or unique salient entities, ‘the’ use expected). This can be done with stories or pictures, or various physical sets of identical or merely similar items (e.g. toys) which form “item contexts”. The goal seems to be testing children’s awareness of their listener’s knowledge of un/established referents. The assumption in many of these studies is that children’s semantics is adult-like.

Nearly all studies find that children use ‘a’ for naming objects consistently, and as early as can be tested. For example Karmiloff-Smith (1979, experiment 2) found that when children were shown a bag of toys and asked “What’s in the bag?” (a situation where there is no presupposed existence of anything, no context set), children (3-12 years) always used ‘a’ correctly, at least >80% (the remaining <20% were omissions and only in 5s and younger). Similar results were found in Bresson et al (1970), Warden (1976, 1981), Schafer and de Villiers (2000).
When it comes to children’s using ‘a’ for introducing referents (i.e. mentioning entities that listener knows nothing about, or entities that have not yet been mentioned in the discourse), there is some variability across studies.

Some studies show early acquisition of ‘a’ and show that children do not overuse ‘the’ for first-mention nonsalient referents. For example 3-4 year olds produced ‘a’ 70-80% of the time (in Emslie and Stevenson (1981) when describing pictures to same-age children sitting behind a screen (2s were at 50% (and only 10-15% ‘the’)); in Maratos (1974) when describing item contexts; in Zehler & Brewer (1982) in sentence completion within narratives by 2-3 year olds; in Schaeffer and Matthewson (2005) by 2-3 year olds at 75% in contexts believed by speaker only and at 100% in context believed by nobody). I.e. children in some studies successfully use ‘a’ to refer to unestablished referents, and do not overuse ‘the’.

Other studies show a later acquisition of ‘a’, and in fact show overuse of ‘the’ in first-mention contexts. It was found that 3-8 year old children use ‘the’ (incorrectly) for referring to a nonsalient entity in a salient context about 40-70% of the time. For example 4 and 5 year olds in Bresson et al 1970 in response to question ‘who left’ in item context; 3 year olds in Garton (1983) in describing item context to a seeing/blindfolded observer; 3-8 year olds in Warden 1976,1981 in describing stories to other children behind a screen, or in another room; the ‘low’ half of 4 year olds in Maratos 1974 when describing item contexts who produced 42% ‘a’ (58% ‘the’); 3-7 year olds in Karmiloff-Smith 1979 exp 1 & 5 when referring to one of several identical objects, in exp 6 when referring to unestablished referents in a verbal story; 3-5 year olds in Schafer and de Villiers. I.e. children in some studies fail to use ‘a’ to refer to unestablished referents, and overuse ‘the’ in those cases.

Is there a coherent difference between those studies that show lack of ‘the’ overuse for unestablished referents and those that show ‘the’ overuse? One logical (semantic) possibility is that children in the former studies are assuming the listener has no context set – in which case ‘a’ is possible, but ‘the’ is not (since children’s knowledge includes the fact that referent of ‘the’ is in the context set, while referent of ‘a’ does not require a context set). In the latter studies, children could be assuming that the listener has a context set, which enables them to overuse ‘the’ (if they do not know uniqueness/Maximality in ‘the’). Unfortunately it is hard to disambiguate participants’ estimations of listeners’ knowledge regarding the presence or absence of context set – this is something that multiple studies addressed explicitly through social manipulations (below), but the results of such manipulations were not consistent.

One point of contrast across the elicited production studies is the method of elicited production. In some, various item contexts are set up (e.g. a group of identical objects, a group of similar objects (differing by color, some unique objects) and the experimenter points to various items or acts out scenarios with puppets which the children have to comment on (a more artificial set up) (Bresson et al 1970, Garton 1983, Karmiloff-Smith 1979). In others, children are given series of pictures forming a story and they have to retell the story (a more naturalistic set up) (Warden 1976, 1981, Emslie & Stevenson 1981). In both manipulations, social contexts can vary – the listener is either absent (Warden 1981), or is the experimenter (an all-knowing listener (Maratos 1974,1976)), or is a blindfolded person (Garton 1983, Warden 1976), or is another child (naïve listener) who is sitting across the table behind a screen or in another room altogether (Warden 1976, 1981; Emslie and Stevenson 1981).

However these social manipulations did not consistently affect the quality of determiners. In Garton (1983), blindfolding the observer reduced the omission of determiners (from 50% for seeing to 22% for blindfolded observer) but did not affect the patterns of production. Explaining
to the child why the screen was between them and their listener and motivating them to be 
understood helped participants in Emslie and Stevenson, relative to Warden 1976 (exp. 3) who 
simply put up the screen. However in an explicit set–up, telling a story by themselves vs to a 
visible or invisible audience did not make any difference in children’s performance (Warden 

When it comes to children’s using ‘the’ for referring to singleton referents (unique objects 
in context set) or salient referents (e.g. second mention referents known to both listener and 
speaker), there is also variation, but children are consistent in using primarily ‘the’ in these 
contexts from about 4 years of age. (Gardon (1983) finds 3s use ‘the’ 29% and ‘this’ and ‘that’ 
over 60%; Emslie and Stevenson find 3-4 year olds use definite descriptions perfectly (‘the’ and 
pronouns) (the 2 year olds were at 60% ‘the’); Warden (1976) finds 3 and 5 year olds at 90% 
‘the’; Maratsos (1974) finds 3 year olds at 55% and 4s at over 90% on ‘the’; Zehler & Brewer 
1982 find only their 3s (and not 2s) close to 80% on ‘the’; Karmiloff-Smith (1979) finds 80-90% 
‘the’ for singletons in 3-9 year olds in item contexts, and 90% ‘the’ by age 5 in verbal story 
contexts; Schafer and de Villiers find 50-70% ‘the’ for children 3-5 (with 20-30% omission and 
10-30% ‘a’); Schaeffer and Matthewson (2005) find 2-3 year olds at perfect performance).

In Summary, if children do make mistakes in using determiners, it is an “a→the” 
response, where children overuse ‘the’ in context where they ought to be using ‘a’ – for an 
established referent among a set of several similar or identical referents. It not a random 
response – children know to use ‘the’ for a singleton or a second-mention salient referent; 
children know to use ‘a’ when there is no context set. I.e. children make the mistake of using 
‘the’ for an unestablished referent within a context of several potential referents, but children do 
not make the mistake of using ‘the’ in a naming task.

2.3.3. Explanations for patterns in production studies

The most favorite explanation of overuse of ‘the’ has been ‘pragmatic’, since semantically 
children seem to know the difference between ‘a’ and ‘the’, but seem to fail to take into account 
listener knowledge.

Bresson et al (1970) link difficulties with ‘a’ in the use of the non-specific referent to the 
child’s ongoing cognitive problems in class and relation concepts (Piaget and Inhelder 1959).

Maratsos (1974) explicitly links those difficulties to Piaget’s (1951, 1955) stage of 
Egocentrism, even though only a subset of children in his data support this theory (the Low 4s) – 
most children show very good knowledge. When egocentric, children fail to estimate their 
listeners’ referential knowledge – their listener’s point of view, thus the saliency of referents is 
evaluated from the child’s point of view and not relative to that of the speaker. Children establish 
a salient referent for themselves and stop there, without introducing the referent to the speaker 
with an indefinite. Children fail to tell their conversation partner explicitly about their chosen 
referent, i.e. they fail to realize that the salience of the referent is not shared by their interlocutor5.

5 (Maratsos 1974:477): “Piaget (1955) had children tell fairy tales and myths to other children, 
and explain to other children the workings of a tap or a syringe. Piaget noted: 'The explainer 
always gave us the impression of talking to himself, without bothering about the other child. 
Very rarely did he succeed in placing himself at the latter's point of view" (Piaget 1955, p. 115). 
On logical grounds this ability might be expected to develop after the primary referential 
competence, since it seems unlikely one could estimate referential knowledge in others without
However it has been noted that the stories in Maratsos experiments were difficult, and children could choose to concentrate on the task and not on the needs of the listener (who is also the experimenter).

Warden (1976) similarly, based on children’s apparent difficulties with ‘a’ (underuse of ‘a’ in first-mention nonsalient nonunique contexts), argues that children younger than 9 years old do not take into account the social context of their referring expressions, i.e., they fail to take into account their audience’s lack of knowledge of referent. Children are either confused about the use of determiners, or are only capable of their own egocentric point of view. Warden (1981) argues that since children never use ‘a’ for a salient referent, they know its meaning. The only possibility is that children fail to implement their knowledge and that Warden’s context manipulations “failed to simplify the context sufficiently to enable children to surmount their egocentricity.” (Warden 1981:98).

More recent studies argue along the similar lines, but reinterpret egocentrism as pragmatic or theory of mind difficulties. E.g. Schafer and de Villiers (2000) suggest that use of ‘a’ vs ‘the’ entails (via a scalar implicature) that the speaker implied ‘a’, and not the stronger ‘the’. Given that computing this implicature involves other people’s beliefs, they argue that it requires Theory of Mind – a modern reinterpretation of egocentrism. Schaeffer and Matthewson (2005) argue that children, when they are trying to figure out whether their language works by definiteness or specificity, “initially lack a pragmatic concept requiring them to distinguish systematically between their own beliefs and the belief state of their interlocutor” (2005:53), it is the “Concept

having developed the required distinctions in one’s own usage first. Proper use of the two articles ‘a’ and ‘the’ clearly demands a high degree of semantic and conceptual competence.”

6 (Warden 1976:110-111): “They [children] fail to recognize the need for an indefinite expression when introducing a referent for the first time in a discourse; consequently, they also fail to recognize the constraints on the use of the definite article, namely that its use indicates an already-identified referent. The most obvious explanation for a child’s failure to identify referents is that he is unable to adopt his audience’s point of view. From his own egocentric viewpoint, a referent is specified as soon as he (the speaker) is familiar with it; he fails to realize that his audience will only become familiar with his referent after he has identified it for them verbally… however … why did nearly every child from four year upwards produce at least some identifying expressions? Were they only partially egocentric? Furthermore, why did the children who used context-bound referring expressions in expt. II only do so in appropriate contexts, namely, when their audience could see the referents. It may be argued that five-year-old children can be non-egocentric in their use of referring expressions, for example, when using demonstratives [!!!!]; but that they are still grappling with the implications of non-egocentricity for the use of the articles. It seems likely that children’s difficulty with the articles stems from the dual function of the indefinite article, namely, to indicate either an indefinite referent or a specific, but previously unidentified, referent [!!!!]. In the former case, a speaker need only consult his own knowledge of a referent, whereas in the latter case he must take account of his listener’s knowledge. Children may be forced to rely on the definite article until they have mastered the identifying function of the indefinite article; and this mastery will depend on an awareness of their audience’s point of view.” (Warden 1976:110-111) [emphasis added]. Note here the observation that children can use demonstrative determiners better than other determiners, and that children’s apparent deficit with ‘a’ (which is really overuse of ‘the’), as described by Warden, occurs in context sets where there are several potential referents. For our purposes (and anticipating Chapter 3) these observations are notable.

22
of Non-Shared Assumptions... Speaker and hearer assumptions are always independent” (2005:69). They suggest that this idea is reminiscent of Theory of Mind⁷. They argue that this accounts for children’s overuse of ‘the’ for first-mention referents (believed only by the speaker).

The parallel between Egocentrism and Theory of Mind is evident to modern cognitive scientists. Theory of Mind is a capacity to attribute and to interpret behavior in terms of mental states of conspecifics (e.g. Baron-Cohen, Tager-Flusberg, Cohen, 1993 & 2000; Astington, Harris, Olson, 1988). Theory of Mind is a more recent proposal along the same cognitive conceptual development lines as Egocentrism that concerns a module of human mind that is responsible for one’s awareness of mental states of other people. Studies show that various aspects of TOM are acquired at different ages, from infant to kindergarten. Children younger than about 4;0 display behavior that has been taken to show that they do not understand that another person will have a false belief if the child him/herself knows the true state of affairs. That is, the child takes another’s understanding of the situation to be the same as his/her own, even though the child has good reason (in the experiment) to understand that the other person should have a different belief. Approximately 50% of 4-year-old children already have this so-called first order theory of mind, as measured by a number of false belief tasks (e.g. Perner, Leekam, Wimmer, 1987; Wimmer & Perner, 1983; Call & Tomasello 1999). When children are about 5-7 years old (depending on the study and the relative complexity of the task), they acquire the ‘second order’ theory of mind, where they are now able to represent one person’s beliefs about another person’s beliefs (e.g. Sullivan, Zaitchik, & Tager-Flusberg (1994); Muris et al (1999); Astington, Pelletier, & Homer (2003)). For example, in a story used by Sullivan et al 1994, a little boy is led to believe by his mother that he would not be getting a puppy for his birthday only to find one in the basement. The child listening to this story must understand that the mother, since she still believes her son to be oblivious to the true nature of the present, believes that her son does not think he is getting a puppy, which is false given the state of the world in the story, yet true given the mother’s beliefs.

It is possible to argue that theory of mind is required for use and comprehension of definite determiners. Note however, that any proposal attributing the misuse of ‘the’ to children’s cognitive immaturity must assume that children know the adult semantic definition of ‘the’, but children do not use ‘the’ in the adult way. In short, the Egocentrism/Theory of Mind account for children’s misuse of ‘the’ suggests that children’s mistakes are not linguistic, but are due to underdeveloped cognition or pragmatics.

One must however keep in mind that it has been shown that children’s linguistic abilities do influence their awareness of other minds. Harris, de Rosnay and Pons (2005) in their review discuss that children with advanced language skills are better at the understanding of mental states, and children with deficient language abilities, such as deaf children born into non-signing families, lag in TOM, leading the researchers to conclude that TOM abilities are enhanced by children’s conversations involving multiple perspectives on a given topic and references to

⁷ Schaeffer and Matthewson (2005), footnote: “This concept might remind the reader of ‘Theory of Mind’. Although we do not exclude the possibility that the phenomenon described and analysed in the present study is related to some of Theory of Mind, we choose not to explain it in terms of Theory of Mind for several reasons: (a) from the literature on Theory of Mind it is not clear what the exact age of the acquisition of Theory of Mind is (for example, some scholars mention the age of 3;0, others 4;0, and again others 4;6); (b) there is no one precise definition of what Theory of Mind is, for example, whether it includes ‘Point of view’, whether it involves the attribution of just false beliefs to others or also other beliefs, etc.”
other’s mental states. In other words, Harris et al argue that it is the children’s pragmatic enrichment that enhances their theory of mind. However work by de Villiers and de Villiers (2003) suggests that children’s ability to embed propositions structurally, i.e. children’s syntactic knowledge, is a precursor for TOM. Indeed, Hale & Tager-Flusberg (2003) demonstrate that training children on sentential complements, sentences of the kind “John said that Mary…”, enhances TOM. 60 preschoolers, who failed a 1st order false belief and a sentential complement pretest, were randomly assigned to three groups, where over the course of two weeks they were trained on one of the following tasks: false belief, sentential complements, noun-describing relative clauses (e.g. “the dish that is broken”). Upon post tests, the group trained on relative clauses only improved on their knowledge of relative clauses, and not on TOM or complements. The group that was trained on TOM only improved on tests of TOM, but the group trained on sentential complements, in a sense a task of embedding and recursion, acquired the linguistic knowledge and significantly increased scores on a number of tasks of theory of mind.

Thus it would not be entirely correct to say that theory of mind is independent from one’s linguistic abilities, however there may exist children that may do well on measures of language, yet do poorly on measures of TOM, e.g. some children on the autism spectrum disorder.

Only a few elicited production studies argue against some version of Egocentrism / Pragmatic / Theory of Mind deficits.

Emslie and Stevenson (1981), who replicated Warden’s experiments with different results (and seem to show very good knowledge of determiners in children when the lack of context awareness in the listener is made very explicit for the children), argued against egocentrism. They predicted that if egocentric responses are inherent in children’s referential expressions, then children should make mistakes with ‘a’ on second-mention referents too (early in their development), then children should turn to making mistakes on first mention referent, i.e. developmentally children should show the following patterns: first a/a (‘a’ for first mention and ‘a’ for second mention), then the/the, then a/the. So if children know ‘a’, they would use a/the, but if children know ‘a’ only for naming things, they would use a/a. This is the logic that Emsli and Stevenson seem to argue for. That children in their experiments were mostly showing a/the pattern as early as 2 years argues against egocentrism. They further argue that egocentric performance in Warden (1976) is likely due to cognitive task demands.

Zehler & Brewer (1982) also argue against egocentrism, and as far as I can tell, they were the first to suggest that something is wrong with children’s knowledge of ‘the’ per se. They find 2-3 year old children overused ‘the’ in nonspecific contexts (only 30% of the time, but these authors finally paid attention the fact that children did use ‘the’ in nonspecific contexts), and sometimes underused ‘the’ in specific contexts. Therefore children must be overgeneralizing some principle that guides the usage of ‘the’, e.g. ‘quasi-known’, or ‘one of a few like-items’, where specification of the referent is not crucial for discourse cohesion.

In summary, children’s overuse of ‘the’ in production, in contexts where there are several potential but unestablished referents, may be due to several factors – either their deficient knowledge of semantic principles (the less favorite explanation), or due to their deficient ability to estimate their listener’s knowledge of the context set (the more favorite explanation). Simplifying the task by making the lack of listener knowledge explicit did not consistently contribute to improvement in children’s performance across studies, which suggests that children are deficient in something more than or other than listener knowledge estimation. Comprehension
experiments may provide a way to solve this issue, since the child is now placed in the role of the listener, and the child ought to know his/her own point of view of the context set and should not be able to go wrong with it. After all, a key assumption of the egocentric/pragmatic deficit hypothesis is that children’s semantic knowledge is intact – it is the application of this knowledge that is deficient. If children are now listening to utterances containing determiners, they ought to get the right interpretation – they are the listener, and they know what they see, and they know semantic principles – they should understand ‘the’ and ‘a’ well.

### 2.3.4. Comprehension studies

One repeated criticism of comprehension studies is due to Brown (1973), who argues that comprehension experiments “place a somewhat unusual communication burden” on the indefinite article, and definite article too. Karmiloff-Smith (1979) argues that such ‘afunctional contexts and unnatural uses tell us relatively little about the way in which children decode the articles”. It is much more natural to say not ‘a X’ when implying another referent, but to say ‘another X’ or ‘another one’; similarly it is more natural to say ‘the same X’, not ‘the X’ when implying the same referent.

Nonetheless, we believe that putting such burden on determiners can clearly tell us whether children are certain of their semantic knowledge, since the burden of pragmatic awareness is minimized in comprehension tasks – the child is the listener! In comprehension tasks, the child does not have to worry about the mental state of the person they are telling a story to. They are the ones listening to the story / instructions. If the communicative burden lies on the speaker, i.e. it is up to the speaker to make sure that the listener can correctly interpret their utterance – as all production studies of determiners assume – then children should not pause and wonder – why was I told this and that – they should perfectly rely on the speaker and let their actions be guided by their semantic knowledge. It can be argued that children are still forced to rely on their pragmatic knowledge of implicatures and non/default situations. This could really be problematic in use of ‘a’ – if the speaker said ‘a’, the implication is that the referent is not necessarily the salient item in context set. Children however may still go for the salient item in context set with ‘a’ if they fail to apply this implicature. But then the role of the child listener can be argued to be an active one, not a passive one.

In Maratsos (1976) comprehension experiments 3-4 year old children had to act out stories involving un/salient referents, and children correctly switched to a new referent when hearing ‘a’ X 76% of the time, and only 6% of the time when hearing ‘the’ X – these are similar stories to the ones used in his production experiments, with children now acting them out. The dichotomy between less successful production and more successful comprehension was taken by Maratsos to support the Egocentrism explanation – it predicts problems only with production, and has nothing to say about comprehension, since whatever state of mind a speaker aims to produce in a child listener would (should, given the assumption that speakers carry the communicative burden) match the child’s state of mind exactly. It is notable that this dichotomy between production and comprehension is exactly what is predicted by egocentrism. However it possible to argue that successful comprehension found by Maratsos was accidental – there were relatively few subjects, and those subjects came primarily from Harvard University area. It is an interesting question to explore whether children’s socioeconomic backgrounds (which are known be highly predictive of verbal (vocabulary) ability have any impact on knowledge of determiners.

In Karmiloff-Smith (1979) contexts of singletons and several identical items were set up for the children – exactly the same set up as in many production studies, except that now children
had to act-out the experimenter’s utterances. In experiment 12, children as young as 3 years when hearing ‘the’ went for the context which contained the singleton item >85%, and children went for the context which had several identical items <15% of the time. When hearing ‘a’, 3 year olds picked one of several items, but 4-7 year olds did so only half of the time – picking the salient singleton a lot. Only 9 year olds were above 80% correct.

In experiment 15, children themselves established a salient referent (by acting on one of the available referents, following instructions in the first clause), and then had to pick that same salient referent (upon hearing ‘the’ in a second clause) or pick another, nonsalient referent (upon hearing ‘a’ in the second clause). On average, children performed twice as many actions on two different x’s than singular x (65% vs 35%) – this is across ‘the’ and ‘a’. Children only took ‘the’ to reliably indicate the unique/salient referent at age 9, whereas 7s and 8s were essentially at chance on ‘the’, 6s were exceptionally good, and 4s and 5s overwhelmingly preferred to pick different objects for ‘the’. For ‘a’, children were more likely to pick two different objects from earliest ages.

Next, it is necessary to review acquisition (comprehension) studies of anaphors ‘same’ and ‘another’ – performed by Karmiloff-Smith (1979 experiment 16, and 1977). Recall (from section 2.1 above) that anaphors are elements whose interpretation is intrinsically dependent on a linguistic antecedent, thus referents of ‘same’ and ‘another’ are dependent on the salient established referent, and if children cannot identify which referent is salient, they will have problems with ‘same’ and ‘another’. The set up here was similar to Karmiloff-Smith’s exp. 15 (1979) where children created their own salient referents: upon hearing ‘a’ X they chose one of several referents (thus establishing a unique salient referent), and then upon hearing ‘another’ X or ‘same’ X children had the option of choosing a different referent or the same referent. In understanding ‘same’ 3 and 4 year olds picked the same KIND of item as the salient referent, but NOT the salient referent. By 5 years, children picked the salient referent. In understanding ‘another’, only 3 year olds showed deficits where 40% of the time children refused to pick ‘another’ object when all items were identical, suggesting that 3 year old children thought ‘another’ meant another KIND, not another referent. From 4 years onwards, children pick two distinct referents with ‘another’. Karmiloff-Smith also tested these same young children on comprehension of the anaphoric use of ‘the’, i.e. using ‘the’ anaphorically to refer to the salient, second mention referent – children were around chance level (in a set up similar to her experiment 15).

Comprehension of ‘this’ and ‘that’ was also investigated (e.g. Webb and Abrahamson (1976)) although from a spatial perspective. Children either sat next to the experimenter (and had the same perspective on near/far objects) or across from the experimenter (and had a different perspective – what was far for the experimenter was near for the child). 4 year olds were correct on same perspective on ‘this’ 82% but only 52% on ‘that’, and on different perspective they performed at 34% on ‘this’ and 82% on ‘that’ (i.e picked the one closer to them most of the time); 7 year olds were at 80% for same perspective ‘this’ and ‘that’, but on different perspective were at 62% on ‘this’ and performed at 74% on ‘that’. Children learned first the spatial contrast (‘this’ vs ‘that’ implies near vs far), and then learned to apply the contrast from speaker’s point of view (when those are different). This is taken to support Piagetian stages (by 4;6 – polarity for own perspective in left vs right, but not for others; by 5;8 polarity for both perspectives, but egocentric; by 7;9 – polarity for both perspectives, non-egocentric – both own and others). This study is important in the present discussion because it establishes that children (from their own point of view) can understand demonstratives.
2.3.5. *Explanations for patterns in Comprehension studies*

It seems that children in Karmiloff-Smith’s exp 12 and exp 15 are doing different things. In one case (exp 12) they seem to know ‘the’ refers to unique referent (and not one of several referents), in another (exp 15) they think ‘the’ refers to either a unique or a non-unique referent. In exp 12 they seem to be confused by ‘a’, in exp 15 they understand ‘a’ to refer to nonsalient objects. The point is that no children are doing what the adults would do until age 8-9. The key difference between experiments is the context set. In exp 15, children are establishing a salient referent for themselves in the first clause, and they know with certainty that ‘a’ in a second clause cannot refer to a salient referent again, indicating that they are keeping track of which referent is the salient one, but they seem to be ok with ‘the’ referring to a non-salient referent. In exp 12, the salient unique (singleton) and the non-unique referents are already set up for the children, and children have to infer that ‘a’ must refer to a nonsalient, nonsingleton entity, but ‘a’ really means ‘any one’ and could be also used for singleton item. What does it mean that children seemed to know ‘the’ in exp 12, but not in exp 15? If children are missing the knowledge of Maximality, they should be performing on ‘the’ just like they are performing on ‘a’ – sometimes going for the singleton and sometimes going for one of the several items, but in experiment 12 they primarily go for the singleton referent. One possible explanation is that, in cases where there is a presupposition that a context set exits (e.g. when ‘the’ is used), children have a preference for a context set which dictates the action, where they don’t have to make a choice – there is only one possible action. In case of the singleton – there is only one referent to pick, otherwise the choice (if it matters to the children) is having to pick one of three. In other words, good performance in exp 12 on ‘the’ may indicate children’s sensitivity of semantics of ‘the’ which call for a context set and the semantics of ‘a’ which do not. When ‘a’ is used, there is no context set, so children pick either response. ‘The’, such that it involves a context set, may indicate to children to narrow the context of interpretation and help them identify the singleton referents. Thus it seems possible to argue that in exp 12 children may interpret ‘the’ correctly without knowledge of uniqueness/Maximality. It will take further investigation to determine whether this explanation might be correct. One question is why the same explanation does not apply to exp 15 and to our experiment in Chapter 3.

There is variation in children’s competence and performance in elicited production and in comprehension studies, but notably there is consensus! Production results suggested children’s deficits are mostly in underuse of ‘a’ for nonsalient referents (i.e. overuse of ‘the’ for non-unique nonsalient referents). Comprehension data suggested that children’s deficits are sometimes with ‘the’ (overuse of ‘the’ to refer to nonsalient referents) and sometimes with ‘a’ (using ‘a’ to refer to salient singletons, although technically this is consistent with semantics of ‘a’). *In other words, whenever studies find deficits, they are deficits, both in production and in comprehension, with overuse of ‘the’ to refer to nonsalient nonunique referents.* Depending on the study, those deficits continue until children are 9 years old (or 4 years old if the children come from Harvard University area, in case of Maratsos8).

One thing that we would like to argue is certain, is that poor comprehension data is unexplainable by pragmatic/egocentric deficits on the part of the child. Recall that Maratsos argued that dichotomy between poor production and good comprehension is evidence for

---

8 The issue of environmental effects on development of determiners, i.e. the possibility of better educated, more verbal parents inducing knowledge of determiners earlier in their children is being investigated (Modyanova, Hirsch, Perovic, Wexler, in preparation).
Egocentrism. What about poor comprehension? Children seem to ‘forget’ which is the salient referent between clause one and clause two of instructions (e.g in Karmiloff-Smith exp 15), something that is implausible. Can children really switch their attention that quickly between clauses, and have issues with maintaining referents and focus of attention? Most production and comprehension studies above assume that in a conversation it is only the speaker who is primarily charged with representing the mind of their interlocutor, and not the hearer, who implicitly trusts the speaker to guide their attention to the salient entities in the discourse set. If the speaker is successful in directing the hearer’s attention to the salient entities in the discourse set, this is not because the hearer recognizes the speaker’s intention to do so. Rather, it is possible that there are innate mechanisms of joint attention at work. (Heim, p.c.).

If however the listener is not passive, but is active (something that none of the studies reviewed so far seem to assume), Egocentric/pragmatic deficits may result in the listener is misconstruing the salient set as relative to a context. I.e. children may make up their own salient referents which would differ from the intended salient referents. This however seems not to be the case, at least in children who understand ‘same’ and ‘another’ correctly (in Karmiloff-Smith 1977) earlier than they understand ‘the’ correctly. Taking File Change Semantics, as an example, the Egocentric/Pragmatic deficit in comprehension will result in children making up (imagine) their own referential file cards that are not mentioned in the discourse, i.e. an ungrounded index card. This ungrounded index card creation is very uneconomical. Thus the only way to interpret the Egocentricism’s hypothesis for children’s miscomprehension of ‘the’, is to say that children are active listeners and are really bad at paying attention to the context of interpretation of the utterance or are really bad at matching the context of the utterance with the salient context of interpretation. This however was shown not to be the case with ‘same’ and ‘another’ above.

Could Zehler & Brewer (1982) be right in arguing that children must be overgeneralizing some principle that guides the usage of ‘the’? Although in explaining the comprehension results, Karmiloff-Smith argued that the unnatural uses and afunctional contexts prevented younger children from showing their knowledge, she also suggests that only older 8-9 children “can make a more ‘abstract’ analysis of the utterance, no longer requiring functional clues... It may therefore be reasonable to hypothesize that something very important takes place in linguistic development from 8 to 9 years, i.e. that children are then capable of a more abstract analysis of an utterance, and no longer require stress on intralinguistic functional clues or extralinguistic situational clues.” [emphasis added] Inadvertently, Karmiloff-Smith in the previous sentence makes a very important observation – a linguistic change, a change in knowledge, may after all be the better explanation for children’s deficits, rather than changes in awareness of other’s minds.

The drastic contrast between comprehension of ‘same’ and ‘another’ on one hand, and ‘a’ and ‘the’ on the other, is also deeply troubling for pragmatic and Piagetian theories, as Karmiloff-Smith notes. “This led to a critical evaluation of Piaget’s implicit contention that young children are using determiners anaphorically”, and that “language is not only the tool of intelligence for representing ongoing cognitive development, but that it is also a problem area for children within its own right. It is suggested that the importance of young children’s processing procedures in the linguistic environment has hitherto been underestimated in Piaget’s interactive epistemology” (Karmiloff-Smith 1977:377). She asks “is there something essentially linguistic at work?” (1977:386).

One frequent observation in comprehension studies is that children seem to prefer to pick two distinct referents rather than one unique referent. A study by Foley et al (2000) illustrates a
similar point. The participants were presented with three characters, e.g. Elmo, Kermit, and Snowman, each of which had his own sandwich, and instructed to make “Elmo touches a/the/his sandwich and Snowman does too”. In ellipsis, two interpretations are possible. The Bound (sloppy) interpretation would have subjects making each character touch their own sandwich, i.e. one to one mapping of characters and sandwiches. The Referential (strict) interpretation would have subjects picking only one sandwich, namely, in the above example, Elmo’s. i.e. there’s a two to one mapping of characters to sandwiches. Thornton and Wexler (1999) show that children know sloppy and strict readings for pronouns. Further, given addition of ‘too’ in stimuli, we expect action on the same object – since ‘too’ implies parallel action. Children ages 4-7 prefer the one to one (bound/sloppy, action on two objects) interpretation where each character gets their own sandwich, regardless of whether the sentences used indefinite (40-55% ‘action on different objects’) or definite (50-60%) articles or pronouns (70%). Other interpretations (action by two characters on same object, action by all three characters on all three objects) take up 10-20% of children’s answers. Adults on the other hand, take the referential/strict/two-to-one reading for indefinite and definite referents, but only slightly prefer the bound reading for pronouns. The relevant point for us from this study is that children below age 8 do not have adult’s interpretations, and are often treating ‘the’ as ‘a’, and preferring to act on two different objects regardless of whether they are hearing a pronoun, a definite or an indefinite determiner.

2.3.6. Summary

It was noted that the indefinite article is correctly produced by children in the absence of a context set, and the definite article is correctly used in presence of a single salient referent. However many of studies show that children incorrectly produce (and understand) the definite article to refer to a nonsalient, not-previously-identified referent from a group of several identical entities (i.e. a partitive context), regardless of whether they are visible objects or imagined characters from a verbally presented story. On the other hand, some studies show that children’s overuse of the definite article in production is to some extent an artifact of the experimental design (e.g. Emslie & Stevenson 1981). Children also seem to know that ‘the’, unlike ‘a’, needs a context set, but children seem to think ‘the’ can refer to one of the nonsalient referents. At the same time, children understand ‘same’ and ‘another’ well. While egocentric/pragmatic explanations work for most elicited production studies, comprehension errors with ‘the’ (and relatively much less comprehension errors with ‘same’) are unexpected, since egocentric/pragmatic view makes the assumption that children’s semantic knowledge is intact and pragmatics is not really involved in comprehension as it is in production. Perhaps the right explanation is not in children’s deficits with pragmatics, but in children’s deficits in semantics.

2.4. Maximality hypothesis: semantic deficits in children’s ‘the’

The key issue for the pragmatic class of theories is children’s mistakes in comprehension. While production of determiners necessitates awareness of the state of the listener’s knowledge about the ongoing information exchange, and consequently a lack of such awareness may well inhibit correct use of determiners, no such things goes on during comprehension. The child is the listener, the child is evaluating own knowledge about the information in a conversational exchange. NO! Let me rephrase that. The child does not need to evaluate own knowledge. The child simply knows! The child simply knows the salient referents in the context set, because the child is the one who defines the salient referents. So why then does the majority of children under the age of 7-8 make comprehension mistakes?
One possibility is that children are not passive listeners, they are active listeners, they are applying their own semantic knowledge, and that semantic knowledge may be deficient.

Wexler (2003) argues that children’s overuse of the definite article is a deficit in children’s computational system of language, arising specifically due to the lack of a component of the semantic definition of ‘the’ – Maximality (Uniqueness). This proposal is based on Heim’s (1991) semantic definitions of articles and her observation that Egocentrism cannot explain all the patterns of children’s use of ‘the’ without further stipulations. It cannot explain children’s overuse of ‘the’ in contexts such as the ‘one of the children is laughing’ Making Noise story (e.g. Karmiloff-Smith 1979). There, children are told about ‘lots of boys and girls’ but there is no information on the referents beyond that – they are not shown, they are not enumerated. Thus there are no unique referents for children to focus on and make salient for themselves, short of inventing such referent out of thin air. Egocentrism also cannot explain children’s lack of overuse of ‘a’ with singletons (unique referents), which is guided by the maximize presupposition implicature, which children are supposed to lack if children have pragmatic problems.

Now suppose children are missing the Uniqueness presupposition from their definition of ‘the’: for them, the definite article simply means that there exists a referent in a context (formalized in 13). Essentially, children’s ‘the’ means ‘one of’ the things in a context set, instead of ‘the unique’ thing. In plural cases, ‘the things’, according to children, may refer to a plural subset of the entities in question, rather than the entire set. However even in lacking uniqueness/maximality, children are still able to differentiate between the definite and the indefinite article, as the definition of ‘a’ has no context set, but ‘the’ does. Observe also that children may be inconsistent with their use of the definite article, which would stem from what children chose the context set to be. Children may decide that the salient referent is the entire context set, as there is no reason, semantically, to rule out the bigger set. Children may also pick the restricted, smaller set that only contains the salient referent. Furthermore, children’s pragmatic implicature of maximize presupposition is intact. Given the results of many acquisition studies, it does look as if this is what children are using in producing and comprehending the definite and the indefinite articles. They only produce ‘a’ in the absence of a context set, and produce ‘the’ in contexts containing a singular referent, contexts requiring reference to one of a set of objects. Children understand ‘the’ to refer to one of a set of equally salient things (incorrectly). This can also apply to plural sets. Plural ‘the’ picks out a maximal whole set of items, and not a subset. Children lacking Maximality may interpret ‘the’ to pick out a plural but not necessarily the maximal set. This is what they do. In Karmiloff-Smith’s (1979) experiment 14, children were presented with toy parking lots containing a variety of cars and trucks and were instructed to e.g. “put the closed cars into the garage”. The four and five year olds’ responses showed violations of maximality with children only picking the entire set of referents on average 12% and 55%, respectively. Children 6 years of age and older showed performance above 80%.

(13) Children’s lexical entry for ‘the’ (Wexler 2003)

Regardless of the utterance context (i), [the* x] P expresses that proposition that is:
true at an index i, if there is an x at i, and it is P at i
false at an index i, if (i) there is an x at i, and there is no x such that x is P at i
truth-valueless at an index i, if there is no x at i

Thus, if children lack the uniqueness/maximality presupposition from the definite article, but have intact pragmatics (at least as much as is relevant for interpretation of determiners), the
main pattern in acquisition, the overuse of ‘the’ instead of ‘a’, seems to be fully accounted for, and key predictions are made. Definite determiners and definite plural DPs are not the only aspects of language that involve Maximality/uniqueness. So do questions, and free relatives, and demonstratives.

Herein, we investigate the Maximality hypothesis by investigating comprehension of anaphoric use of ‘the’ and ‘that’, ‘a’, ‘another’ and ‘same’.

We predict that if children’s deficit is semantic (and not pragmatic/theory of mind) in nature, we expect a difference in acquisition between ‘same’ and ‘the’ – children will do better on ‘same’ that does not involve Maximality. We expect difference between ‘the’ and ‘that’ – children may do better on ‘that’ if they are aware of its pragmatic restriction to interpretation in non-default contexts. We expect that children can figure out the relevant salient context set early on, and will show good knowledge for determiners that do not involve Maximality, but do involve knowledge of the salient context – ‘same’ and ‘another’, i.e. we aim to replicate the findings of Karmiloff-Smith of ‘same’ and ‘another’ in English.

If children’s deficits are not semantic, and children do know Maximality as well as definitions of other determiners, but are deficient in speaker/hearer knowledge distinctions, we expect children to do very well across the board if children are passive listeners letting the speaker direct their attention. If children are active listeners on the other hand, and they are deficient in pragmatics, they may have problems with those items that rely more on the speaker/hearer knowledge on their interpretation, e.g. in ‘that’ (according to theory of ‘that’ by Wolter 2006a,b). Except that we will not see that in children’s performance – semantic knowledge, knowledge of uniqueness should be enough in choosing the salient referent for anaphoric use of ‘the’ and ‘that’. Thus if children’s deficit is not semantic, there should be no difference in performance between ‘same’, ‘the’, and ‘that’. Children, as active interpreters, could still have difficulty establishing the salient context set, and that may cause problems with all determiners, including ‘another’ where we should see lots of incorrect actions on salient referent.

2.5. Acknowledgements

We would like to thank Irene Heim for enlightening theoretical discussions. Any errors are ours.


2.6. References


3. Comprehension of Determiners in Typically Developing Children

To elucidate the situation in the development of referential abilities, we study children’s comprehension of articles, allowing children to demonstrate their preference for the interpretation of the anaphoric use of definite and indefinite and demonstrative determiners.

3.1. Abstract

Is children’s overuse of ‘the’ due to their deficient knowledge of semantics or due to their deficient use of such knowledge due to pragmatic or egocentric factors? We tested 203 children ages 3 to 9 years on comprehension of the indefinite determiner ‘a’, anaphor ‘another’, and definite determiners ‘the’ and ‘that’. Of these children, 74 were also tested on comprehension of anaphor ‘same’ and verb ‘share’. Children at three years of age were able to pick the correct referent for ‘same’ 65% of the time, and 83% of the time at age four, indicating their knowledge of the salient referent. Children were also able to pick the correct referent for ‘another’ at all ages, again indicating their knowledge of the salient referent in the context set by picking not that referent. At the same time, only 40% of 4-5 year olds consistently showed correct interpretation of ‘the’. The majority of children (70%) consistently picked the salient referent for ‘the’ only from the age of seven years. Additionally, between 5% to 40% of children across ages were able to pick the correct referent of ‘that’ better than they were able to pick the correct referent of ‘the’, with this pattern peaking around 5 years. The opposite pattern was only observed in two six-year olds. These results are taken to support the Maximality hypothesis, that children are deficient in their knowledge of uniqueness/Maximality in the definite article, and is taken to argue against theories where children’s knowledge is intact, but performance suffers from pragmatic/cognitive limitations.

3.2. Experiment 1. “Puppets”

3.2.1. Method

Figure 3.2.1. Act-out task set up
An act-out paradigm was used, closely following Karmiloff-Smith’s (1979) experiment 15. This particular paradigm was chosen because children can establish a salient referent for themselves as part of the task, and do not have to infer the saliency of a referent from the context of singleton vs several identical nonsalient items. This kind of set up worked well for Karmiloff-Smith when investigating ‘same’ and ‘another’, and we aim to use a consistent paradigm for the new investigation of anaphoric use of ‘that’.

Children’s comprehension of different articles, as illustrated by children’s actions is studied. Several objects are laid out in front of a child, here six fences, three balloons, three spoons, and three logs. The significance of three vs six context sets will be explained later. A child is given two puppet actors, here Kanga and Froggy, and asked to follow the investigator’s instructions: “Kanga, push a balloon” [break to allow for action] “and then Froggy, push a/the/that balloon.” The first clause always contains the indefinite article, and is used to establish a salient referent out of a group of identical objects. The second clause asks subjects to act on the same object that was acted on in the first clause if ‘the’ and ‘that’ are used and interpreted in adult manner, or to act on any object if ‘a’ is used and interpreted in adult manner. Children were presented with exemplars of each condition 4 times in a randomized order. Two verbs were used for the act-out task: push and kiss.

It may be argued that the instruction phrase is not exactly a natural set up, since adults are more likely to say ‘the same X’, rather than simply ‘the X’. However, our goal is investigate children’s knowledge of ‘the’, ‘that’, and ‘a’ in a ‘pure’ setting – without additional disambiguating terms.

The dependent variable is the number of actions by children on the same object in both clauses. Observe that children are placed in the position of the listener, thus they have to evaluate their own knowledge about the context and the referents, and what the speaker’s words are meant to imply.

### 3.2.1.1. Context size variable

The idea for investigating the possible effects of the size of the context set arose following a closer examination of Karmiloff-Smith’s (1979) data from her experiment 15. 43 French children aged 4-9 years were tested on two conditions with ‘a’ or ‘the’ in the second clause. There was only one presentation of each condition, but additional items were added ad hoc. The context set consisted of four objects.

Briefly going over her results (Table 3.2.2, where percentage of actions on the same object are shown), the indefinite article is mostly correctly interpreted to mean ‘any’ or ‘a different one’ depending on the age of the subjects. The definite article is differentiated from the indefinite article by all ages except 4, 5 and 7 year olds. Perfect knowledge of ‘the’, which would consist of 100% actions on the same object is not attained by any subjects, but is approximated by 6 and 9 year olds. Children are clearly presenting a deficit in comprehension and correct interpretation of the definite article.

<table>
<thead>
<tr>
<th></th>
<th>4;0-4;11</th>
<th>5;0-5;11</th>
<th>6;0-6;11</th>
<th>7;0-7;11</th>
<th>8;0-8;11</th>
<th>9;0-9;11</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>19%</td>
<td>10%</td>
<td>14%</td>
<td>42%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>the</td>
<td>26%</td>
<td>30%</td>
<td>72%</td>
<td>42%</td>
<td>47%</td>
<td>81%</td>
</tr>
</tbody>
</table>

Table 3.2.2. Percentage ‘same’ responses by French children from Karmiloff-Smith, exp 15

If we consider the probabilities for children’s performance given a context size of 4 objects and a single presentation of each condition, it becomes evident that younger children are
performing at chance levels. Once one of 4 objects is made salient, there is a 75% chance of picking one of the other three nonsalient objects, if the child picks a randomly – i.e. as if ignoring the determiner altogether. The chance for picking the same, salient object randomly is only 25%. It is important to note that these chance levels are only applicable in case there is only one presentation of each condition. Well, this is indeed the performance that French 4 and 5 year-olds are giving. They are picking about 80% different objects for ‘a’, and about 70% different object for ‘the’.

Thus our goal was to pick two context sizes such that chance levels would be sufficiently different, yet the experimental set-up would still be manageable. There are two items for every determiner condition that have 3 elements of a set of objects, and two other items for every conditions that have 6 elements of a set of objects, giving us 33% vs 16% chance levels, respectively. Note that average chance levels over both determiner context conditions is 25%. Additionally, we can argue that there are two ways to define a salient set. For children who are missing uniqueness it could be that the salient set of referents is the entire context set in front of them, or it could be the unique salient referent. Thus a smaller context size of 3 items may be easier for children to apprehend at once, but a larger set of 6 items may be too much to take in, leading the child to select the unique salient referent. On the other hand, we may see evidence of random model in children’s performance – children, by chance, may choose the salient referent more often in the 3-item context than in the 6-item context. Using two different context set sizes, we can reliably investigate whether children are responding uniformly or randomly.

3.2.2. Participants

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean Age</th>
<th>N</th>
<th>Minimum Age</th>
<th>Maximum Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>3s</td>
<td>3;7</td>
<td>12</td>
<td>3;2</td>
<td>3;11</td>
</tr>
<tr>
<td>4s</td>
<td>4;6</td>
<td>18</td>
<td>4;0</td>
<td>4;11</td>
</tr>
<tr>
<td>5s</td>
<td>5;5</td>
<td>13</td>
<td>5;0</td>
<td>5;10</td>
</tr>
<tr>
<td>6s</td>
<td>6;6</td>
<td>12</td>
<td>6;2</td>
<td>6;11</td>
</tr>
<tr>
<td>7s</td>
<td>7;8</td>
<td>12</td>
<td>7;1</td>
<td>7;11</td>
</tr>
<tr>
<td>8s</td>
<td>8;4</td>
<td>11</td>
<td>8;0</td>
<td>8;9</td>
</tr>
<tr>
<td>9s</td>
<td>9;6</td>
<td>11</td>
<td>9;1</td>
<td>9;11</td>
</tr>
<tr>
<td>10s</td>
<td>10;4</td>
<td>4</td>
<td>10;0</td>
<td>10;10</td>
</tr>
<tr>
<td>Adult</td>
<td>19</td>
<td>2</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>Total</td>
<td>6;10</td>
<td>95</td>
<td>3;2</td>
<td>19;0</td>
</tr>
</tbody>
</table>

Table 3.2.3. Participants in ‘Puppets’ Task
3.2.3. Results

Our results, summarized in Figure 3.2.4, replicate those found previously for ‘a’ and ‘the’. The indefinite article is correctly interpreted as ‘any’ by all age groups, remaining around 30% - very close to the average 25% chance levels. The definite article is differentiated from the indefinite article, but proficiency is not yet fully attained even by the 10 year-old group. Our new result is that the determiner ‘that’ is differentiated from the definite article, with more children picking more same objects upon hearing ‘that’, at almost all ages. Proficiency with ‘that’ is attained by 9 year olds – earlier than with ‘the’. Adults are excluded from further analyses.

A mixed repeated measures ANOVA was performed (Determiner (3 levels) by Context (2 levels) as repeated measures by Age group (8 levels) as between subject measure). Mauchly’s test of Sphericity for within subjects effects of determiner (χ²(2) = 10.05, p = .007) and the interaction between determiner and context (χ²(2) = 7.03, p = .03) come out significant, hence we must accept the hypothesis that the variances of the differences between the levels of variables are significantly different, and the assumption of sphericity has been violated. Therefore, degrees of freedom were corrected using the Huynh-Feldt estimates of sphericity (ε = .993 for determiner and ε = 1 for the interaction). For main effect of context, the assumption of sphericity is met. Results revealed a significant effect for determiner (F(1.98, 168.73) = 47.72, p<.0001) and context (F(1,85) = 11.52, p = .001). The effect of age group was significant (F(7,85)=2.5, p=.023). The age by determiner type is the only interaction that remotely approached significance (F(13.89, 168.73) = 1.47, p = .128).

<table>
<thead>
<tr>
<th>Age</th>
<th>3s</th>
<th>4s</th>
<th>5s</th>
<th>6s</th>
<th>7s</th>
<th>8s</th>
<th>9s</th>
<th>10s</th>
<th>Adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>A%</td>
<td>21%</td>
<td>33%</td>
<td>23%</td>
<td>33%</td>
<td>42%</td>
<td>34%</td>
<td>34%</td>
<td>37%</td>
<td>62%</td>
</tr>
<tr>
<td>The</td>
<td>39%</td>
<td>47%</td>
<td>46%</td>
<td>67%</td>
<td>54%</td>
<td>64%</td>
<td>75%</td>
<td>75%</td>
<td>100%</td>
</tr>
<tr>
<td>That</td>
<td>42%</td>
<td>57%</td>
<td>67%</td>
<td>60%</td>
<td>60%</td>
<td>70%</td>
<td>95%</td>
<td>94%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Figure and Table 3.2.4. Percentage ‘same’ responses by English-speaking children
(error bars represent +/-1 standard error)
Post-Hoc pairwise comparisons (adjusted for multiple comparisons with Bonferroni correction) revealed significant difference between the 3 year old and the 9 year old group (p=.018) (no other comparisons of age groups were significant), significant difference between all determiners (‘a’ vs ‘the’ (p<.001), ‘a’ vs ‘that’ (p<.001), ‘the’ vs ‘that’ (p=.006)), and significant difference between 3-item and 6-item contexts (p=.001).

It can be noted that all children (age groups) differentiate between the indefinite and the definite articles. This in itself however cannot be taken as evidence for knowledge of Uniqueness, as the in/definite articles can be distinguished by existence of context. The indefinite implies no context consideration. The definite determiners imply consideration of the context, and in case of “child’s the” there are two context possibilities: the entire set of objects can be considered, or alternatively only the object that was previously made salient can be taken as the context set.

Thus the results so far strongly suggest that English-speaking children in an act-out task are able to interpret ‘that’ to refer to an established salient referent better than they can interpret ‘the’. This difference in interpretation is not predicted by Egocentrism (pragmatic/cognitive deficit), but is predicted by Uniqueness (computational system of language deficit).

3.2.3.1. Random model in children’s performance

Figures 3.2.6-3.2.8 summarize the proportion of subjects performing in particular ways on determiners, i.e. picking the ‘same’ reference 0, 1, 2, 3, or 4 times, out of the 4 possible presentations. A random model of responses is derived by considering probabilities given a context size and number of presentations (Figure 3.2.5). Given that children are presented with each conditions 4 times (2 times with 3-item context and 2 times with 6-item context), it is possible to derive the proportion of subjects that will produce from 0 to 4 out of 4 possible ‘same’ responses. Thus by pure chance, more subjects will show 0 and 1 ‘same’ responses. If such a pattern of responses is seen in the data, then subjects are guessing.

Individual subject’s performance on ‘a’ shows that the older participants are eerily reminiscent of the random response model, which is in fact the correct interpretation for ‘a’. The younger participants show a preference for acting on two different objects. Perhaps this is evidence of an interpretation strategy. Individual subject’s performance on ‘the’ shows that the younger participants display a random model pattern of responses. The older participants show less random model and more knowledge of ‘the’, however performance is not perfect. Individual subjects’ performance on ‘that’ shows that a lot of younger children are now moving away from the random model and showing some knowledge of ‘that’. Older participants show firm knowledge – there is no room for random model. Observe, that all proportions generally fit a unimodal distribution, suggesting children to be on a continuum of knowledge. If we saw bimodal distributions, it would suggest distinct populations within subjects, such that some children knew exactly what was going on, and the rest failing the task.

---

1 Average of 3-item and 6-item contexts: e.g. P(same=4) = (1/4)^4; e.g. P(same=0) = (3/4)^4.
3.2.3.2. Analysis of children performing ‘well’

We will define children as “performing well” on ‘the’ and ‘that’ when they perform the action by the two actors on the ‘same’ element of a set of objects 3 or 4 times per condition (i.e. at least 75% ‘same’ performance). Participants were subdivided by their pattern of performance between ‘the’ and ‘that’, with “A” pattern denoting good performance on both ‘the’ and ‘that’, “B” pattern denoting bad performance on ‘the’ with a relatively better performance on ‘that’, “C” pattern denoting a bad performance on ‘that’ with a relatively better performance on ‘the’, and finally “D” pattern denoting a bad performance in both ‘the’ and ‘that’. Figure and Table 3.2.9 shows for each age the proportion of children showing a given pattern on ‘the’ and ‘that’. The children who are performing well on both ‘the’ and ‘that’ tend to be older, while children performing poorly on both tend to be younger. It should be noted that children performing well on ‘that’ only (B pattern) are well distributed throughout the age ranges, while only 5 children perform better on ‘the’ (C pattern).
3.2.3.3. Effect of context size on interpretation of articles

Recall that children were presented with either 3 or 6 elements of a set of identical objects. The ANOVA (above) showed a significant main effect of context, therefore a closer look is warranted. The first clause of instructions established a salient referent, and the second clause asked the child to act on either the same or a different referent. Thus the definite determiner phrases here are endophoric, i.e. the referent is identifiable from immediate linguistic context: if the first clause had ‘a balloon’, and second clause had ‘the balloon’, an adult does not need to consider a physical context set in order to interpret the referent. At the same time, the phrase may also be interpreted exophorically, meaning the salient context set and the salient unique referent is identifiable on extralinguistic grounds or from the specific physical situation of the utterance. Since children are guessing referents due to their inconsistent knowledge of ‘the’, children may show different performance depending on the context set.

The Uniqueness account suggests that children will take either the endophoric unique discourse-determined context set of size one i.e. the established referent, in which case we may see the ‘same’ response. Alternatively, children may take the exophoric situation determined context set of size 3 or 6, depending on the trial, in which case they may respond according to a random model, with probabilities of responses determined by size of that set. Thus the uniqueness account predicts that more same responses will be given for a context size of 3 than context size 6 (Figure and Table 3.2.10).

It is difficult to see what the Egocentrism account predicts. It seems that it does not predict a differential or random treatment of articles at all, so context size should not matter. Thus
we expect to see little random responses, as children will be focusing on one object, and producing the same performance in both contexts. If kids pick a unique element for themselves, then the context selected is unique, and children may produce more ‘same’ responses than uniqueness account predicts.

We derive the random model of responses for each context size via probability performance levels (Figure and Table 3.2.10). Chance levels for action on the same object given two presentations for the 3 and for the 6 contexts are given. When looking at the ‘visible values’, i.e. one and two responses out of two, it follows that by pure chance, more children are likely to act on the same referent when there are 3 elements of a set in front of them, than when there are 6 elements.

<table>
<thead>
<tr>
<th>‘SAME’</th>
<th>3-ITEM</th>
<th>6-ITEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero</td>
<td>0.44</td>
<td>0.69</td>
</tr>
<tr>
<td>One</td>
<td>0.44</td>
<td>0.28</td>
</tr>
<tr>
<td>Two</td>
<td>0.11</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Figure and Table 3.2.10. Probability ‘same’ responses: random model for 3 vs 6 object context, given two presentations

Figure 3.2.11 shows the difference in percent ‘same’ responses between the 3-item and the 6 item context for each age and for each determiner. As expected, context size significantly affects performance on the indefinite article for all age groups. The indefinite indeed means ‘any one’ for children, as it does for adults. While context size affects performance in younger children on ‘that’, indicating imperfect knowledge, there is no context influence in the older children, suggesting lack of doubt in referent choice guided by familiarity, and singular choice of context set. Context size significantly affects performance on the definite article ‘the’ for the younger children and overall, suggesting a choice of context set from two possibilities: the entire set of objects or the already-established referent. The above analyses show that children are sensitive to context set size, something that is not predicted by Egocentrism, but is expected on the Uniqueness account.

The way such pattern can be predicted by egocentrism/pragmatic difficulties is if an assumption is added: the larger the context size, the more likely a child is to let his/her mind wonder, and to pick a different (own) salient referent from the true salient referent. This however does not really seem plausible to us.
Figure 3.2.11. Proportion difference ‘same’ responses between 3-item and 6-item context set
3.3. **Experiment 2. “Felt”**

3.3.1. **Method**

Another version of the above experiment was developed, eliminating any potential issues affecting children’s performance. One potential confound is memory and attention limitations of young children. There were many things a subject had to attend to in the experimental set-up, so it was possible that children forgot which object was affected in the first clause. There was extensive set up of stimuli between conditions, and at times it was possible for the participants to forget the salient referent – if for example the puppet that made the initial action was then returned to the starting position. Reliance on participants’ remembering of the initial salient referent may place unnecessary cognitive demands and may obscure participants’ knowledge, and may perhaps lead to the context-size effects observed. In the new task, the goal was to make the initial salient referent visually obvious so as to eliminate memory involvement.

A second problem was small number of sentences per condition. Thus the new version has 6 sentences per condition. A condition is added to balance out ‘same’ versus ‘different’ actions - ‘a/another’. This ‘another’ condition is also a test of anaphoric interpretations, as its correct interpretation means ‘a different entity from the salient one in the context set’, which requires one to be aware of the salient entity in the context set, but does not require knowledge of Maximality for correct interpretation.

The experiment is in a form of a book made of felt cloth, with arrays of 3 or 6 identical objects permanently attached to the pages. Objects included apple, watermelon slice, car, carrot, Christmas tree, star, flower, ice-cream cone, baseball, heart, witch’s hat. Every page contains an array of target items and an array of distractor objects. Subjects are given two Velcro-backed ‘actors’, Fishy and Turtle. Thus the characters can be physically stuck to the page by children who can see what they were/are doing.

There were two ways of making sure the subjects are paying attention: whether they are affecting the right kind of object, and whether they are doing that with the correct actor (the order in which the actors affect the objects is randomized).

Firstly, subjects are introduced to the two actors and different objects that will be used. A few practice trials follow where children become comfortable with using actors to touch either same or different objects. Subjects are told that in this ‘game’, they have to decide whether Fishy and Turtle go to same or different objects based on what the experimenter says. Article use is avoided during introduction to avoid biasing subjects’ responses.

Comprehension of items is evaluated when the experimenter gives the child instructions on how to manipulate the actors given the context of the current page. The first clause of instructions always contains the indefinite article ‘a’ and serves to establish a unique, salient, visually distinct referent within the context set – the object that has Fishy or Turtle on it. The second clause contains one of the four articles – ‘a’, ‘another’, ‘the’, or ‘that’. Note that simple present tense is used in all sentences in experiment, unlike imperative which was used in ‘Puppets’.

Instructions go like this: “Fishy touches an apple, [pause to allow for action of Fishy touching one of the apples] and Turtle touches a/another/the/that apple.” An adult would respond like this: pick two different objects upon hearing ‘another’, pick any object upon hearing ‘a’, and
pick the same (salient) object upon hearing ‘the’ and ‘that’. The number of times the same object was subsequently acted upon served as the dependent variable.

The modulation of the context size (three-item vs six item) was retained in this experiment, to confirm the effects observed previously.

Thus the comprehension experiment tests how children interpret ‘the’ versus ‘a’, specifically whether they know Uniqueness or whether they take account of the pragmatic conditions for establishing context sets – that both the speaker and listener must have means for determining the context sets. It also tests whether children use pragmatic properties (old information, known to listener) of ‘that’ to decide how to use the correct article.

### 3.3.2. Participants

<table>
<thead>
<tr>
<th>Age group</th>
<th>Mean age</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>3.47</td>
<td>19</td>
<td>3.00</td>
<td>3.96</td>
</tr>
<tr>
<td>4</td>
<td>4.62</td>
<td>18</td>
<td>4.08</td>
<td>4.97</td>
</tr>
<tr>
<td>5</td>
<td>5.45</td>
<td>21</td>
<td>5.00</td>
<td>5.91</td>
</tr>
<tr>
<td>6</td>
<td>6.55</td>
<td>21</td>
<td>6.10</td>
<td>6.98</td>
</tr>
<tr>
<td>7</td>
<td>7.54</td>
<td>21</td>
<td>7.02</td>
<td>7.96</td>
</tr>
<tr>
<td>8</td>
<td>8.62</td>
<td>15</td>
<td>8.00</td>
<td>8.99</td>
</tr>
<tr>
<td>9</td>
<td>9.42</td>
<td>14</td>
<td>9.00</td>
<td>9.92</td>
</tr>
<tr>
<td>Total</td>
<td>6.36</td>
<td>129</td>
<td>3.00</td>
<td>9.92</td>
</tr>
</tbody>
</table>

Table 3.3.2. Participants in ‘Felt’ task

The participants included 129 children from Boston and Cambridge daycares and afterschool programs. The results of this study are more clear than those using ‘puppets’ version of the experiment, since children no longer have to rely on remembering the salient referent but they are able to clearly read it off the context. The clear-cut trends in the data were already evident with about ten children per age group, however more children were recruited to verify the patterns as well as for purposes of the ongoing autism study.
### Results

![Developmental Trajectory of Determiners](chart.png)

**Figure and Table 3.3.3. Percentage ‘same’ responses by children**
(percent choice of salient entity in the context set (action on the same object by both actors))
(error bars represent ±1 standard error)

<table>
<thead>
<tr>
<th>Age</th>
<th>3;0-3;11 (19)</th>
<th>4;0-4;11 (18)</th>
<th>5;0-5;11 (21)</th>
<th>6;1-6;11 (21)</th>
<th>7;0-7;11 (21)</th>
<th>8;0-8;11 (15)</th>
<th>9;0-9;11 (14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A%</td>
<td>2.7%</td>
<td>5.5%</td>
<td>15.0%</td>
<td>10.3%</td>
<td>13.5%</td>
<td>6.7%</td>
<td>13.2%</td>
</tr>
<tr>
<td>Another%</td>
<td>0.8%</td>
<td>5.5%</td>
<td>7.2%</td>
<td>4.8%</td>
<td>1.7%</td>
<td>0.0%</td>
<td>1.2%</td>
</tr>
<tr>
<td>The%</td>
<td>2.7%</td>
<td>44.5%</td>
<td>46.8%</td>
<td>54.0%</td>
<td>70.7%</td>
<td>74.5%</td>
<td>85.7%</td>
</tr>
<tr>
<td>That%</td>
<td>8.8%</td>
<td>49.0%</td>
<td>62.7%</td>
<td>50.8%</td>
<td>81.0%</td>
<td>80.0%</td>
<td>97.7%</td>
</tr>
</tbody>
</table>

A mixed repeated measures ANOVA was performed (Determiner (4 levels) by Context (2 levels) as repeated measures by Age group (7 levels) as between subject measure). Mauchly’s test of Sphericity for within subjects effects of determiner ($\chi^2(5) = 184.5$, $p < .001$) and the interaction between determiner and context ($\chi^2(5) = 11.1$, $p = .048$) come out significant, hence we must accept the hypothesis that the variances of the differences between the levels of variables are significantly different, and the assumption of sphericity has been violated. Therefore, degrees of freedom were corrected using the Greenhouse-Geiser estimates of sphericity for determiner ($\varepsilon = .508$ for determiner) and Huynh-Feldt estimates of sphericity for the interaction ($\varepsilon = 1$ for the interaction). For main effect of context, the assumption of sphericity is met. Results revealed a significant effect for determiner ($F(1.52, 185.8) = 232.9$, $p<.001$), with determiner type by age interaction reaching high significance ($F(9.13, 185.5) = 8.9$, $p < .001$). Neither context, nor interaction between context and age reached significance. Interaction between determiner and context was significant ($F(3,366)=5.4$, $p=.001$). The three way
interaction (determiner by context by age) was not significant. The effect of age group was significant (F(6,122)=10.3, p<.001).

Post-Hoc pairwise comparisons (adjusted for multiple comparisons with Bonferroni correction) revealed significant difference between the 3 year old group and all other age groups (all ps<=.001, except for 4 year olds (p=.012)), 4 year olds also differed from the 9 year olds (p=.02). No other age group comparisons showed significance. There was also significant difference among all determiners (all ps <.001, i.e. each determiner was significantly different from all other determiners for all ages). There was no difference between contexts.

Comprehension of the indefinite article ‘a’ remained between 3% and 15% actions on the same object for all ages studied. This is lower that what would be expected for chance performance (25%) if children took ‘a’ to mean ‘any one’.

Comprehension of ‘another’ was surprisingly good (i.e. practically no actions on the same object) for even the youngest participants and indicated a number of things. Firstly, ‘another’ is an anaphor, i.e. has a distinct mechanism of interpretation from the indefinite and definite articles. Secondly, interpretation of ‘another’ involves being aware of the salient entity in the context set, and picking not the salient one! In other words, as early as 4 years, when children begin to distinguish between indefinite and definite articles, the notion of ‘salient entity in the context set’ is known to children. Had this not been known by children, they would be performing around 25% actions on the same object, which, obviously, they are not. Children are differentiating ‘a’ from ‘another’ significantly (see post-hoc tests above).

Comprehension of the definite determiner ‘the’ showed an S-shaped curve found in the development of many biological systems, with an initial plateau around 50% and lack of change in performance for 4, 5, and 6 year olds, with the following rapid increase between ages 6 and 8, with 8 year olds already reaching adult-like levels of performance.

Comprehension of ‘that’ was better than comprehension of ‘the’ for all ages studied (see post-hoc tests above). Ignoring, for the moment, the dip in the 6 years, interpreting ‘that’ as referring to the salient entity in the context set gives children a whole two years worth of advantage over interpreting ‘the’. Using ‘that’ helps 5 year old children interpret 63% of utterances correctly, whereas on ‘the’ only 7 year olds reach the same level.

Participants were subdivided by their pattern of performance between ‘the’ and ‘that’ (Figure and Table 3.3.4), with “A” pattern denoting good performance on both ‘the’ and ‘that’, “B” pattern denoting bad performance on ‘the’ with a relatively better performance on ‘that’, “C” pattern denoting a bad performance on ‘that’ with a relatively better performance on ‘the’, and finally “D” pattern denoting a bad performance in both ‘the’ and ‘that’. 3

We find that no 3s show adult-like interpretation, only 45% of 4-6s show adult-like interpretation. This analysis makes more evident the steep increase between ages 6 and 8, with

---

2 Chance performance (see section on random model) suggests that children, by chance will produce many 1/6 or 2/6 responses, which would result in about 25% ‘same’ responses.

3 Pattern the vs that: If the difference (absolute value of ‘the’-‘that’)<2, then (if the<4 assign “D”, else assign “A”); else (if ‘the’<‘that’ then (if that<3, then “D”, else “B”), else (if the<3, then “D”, else “C”)). Random model of responses (section 7) shows that it is extremely unlikely to get 4/6 or better by chance, hence for ‘the’ and ‘that’, 4-6/6 is ‘good’ performance. The difference between ‘the’ and ‘that’, to be significant, is decided to be at least 2. In case of such differential performance, it is necessary to show knowledge of at least 3/6 on ‘the’ or ‘that’.
80% of 8 year olds showing adult like interpretations. Apparent increase in knowledge of ‘that’ relative to ‘the’ in 5s and 7s is due to a subset of these children interpreting ‘that’ better than ‘the’. The apparent dip in knowledge of ‘that’ in 6s is due to 2 children interpreting ‘the’ better than ‘that’.

Figure and Table 3.3.4. Proportion of children performing well by age

<table>
<thead>
<tr>
<th></th>
<th>3(19)</th>
<th>4(18)</th>
<th>5(21)</th>
<th>6(21)</th>
<th>7(21)</th>
<th>8(15)</th>
<th>9(14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>0.95</td>
<td>0.56</td>
<td>0.33</td>
<td>0.43</td>
<td>0.19</td>
<td>0.20</td>
<td>0.00</td>
</tr>
<tr>
<td>C, that&lt;the</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.10</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>B, that&gt;the</td>
<td>0.05</td>
<td>0.00</td>
<td>0.24</td>
<td>0.00</td>
<td>0.14</td>
<td>0.00</td>
<td>0.14</td>
</tr>
<tr>
<td>A, the=that</td>
<td>0.00</td>
<td>0.44</td>
<td>0.43</td>
<td>0.48</td>
<td>0.67</td>
<td>0.80</td>
<td>0.86</td>
</tr>
</tbody>
</table>
3.4. Experiment 3. Replication of “Felt” with addition of ‘same’ and ‘share’

3.4.1. Method
Additional control conditions were built into the “Felt” task to verify that children have the concept of ‘sameness’, know which is the salient referent in the context set independent of Maximality in ‘the’, and to assure that children do not slip into a strategy of simply picking two different objects. These unambiguous sentences involved ‘same’ and ‘share’, and replaced the second clauses of utterances: “Fishy touches an apple, and Turtle touches the same apple” OR “…and Turtle shares the apple with Fishy.”

3.4.2. Participants
Another group of participants, which included 104 children and teenagers from Boston, Cambridge, Wellesley, and Michigan, was tested on the version of ‘felt’ task which included the ‘same’ and ‘share’ control conditions.

<table>
<thead>
<tr>
<th>Age</th>
<th>Mean</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>3.59</td>
<td>15</td>
<td>2.92</td>
<td>3.98</td>
</tr>
<tr>
<td>4</td>
<td>4.46</td>
<td>19</td>
<td>4.00</td>
<td>4.96</td>
</tr>
<tr>
<td>5</td>
<td>5.43</td>
<td>7</td>
<td>5.09</td>
<td>5.97</td>
</tr>
<tr>
<td>6</td>
<td>6.41</td>
<td>12</td>
<td>6.00</td>
<td>6.93</td>
</tr>
<tr>
<td>7</td>
<td>7.34</td>
<td>9</td>
<td>7.06</td>
<td>7.71</td>
</tr>
<tr>
<td>8</td>
<td>8.61</td>
<td>8</td>
<td>8.03</td>
<td>8.89</td>
</tr>
<tr>
<td>9</td>
<td>9.46</td>
<td>4</td>
<td>9.06</td>
<td>9.93</td>
</tr>
<tr>
<td>Total</td>
<td>5.76</td>
<td>74</td>
<td>2.92</td>
<td>9.93</td>
</tr>
</tbody>
</table>

Table 3.4.1. Child participants

<table>
<thead>
<tr>
<th>Age</th>
<th>Mean</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>“10”</td>
<td>11.44</td>
<td>19</td>
<td>10.00</td>
<td>13.14</td>
</tr>
<tr>
<td>“20”</td>
<td>16.74</td>
<td>11</td>
<td>14.59</td>
<td>19.66</td>
</tr>
</tbody>
</table>

Table 3.4.2. Adolescent and Adult participants

3.4.3. Results

<table>
<thead>
<tr>
<th>OLDER GROUP</th>
<th>A</th>
<th>ANO</th>
<th>THE</th>
<th>THAT</th>
<th>SAME</th>
<th>SHARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>10;0-13;1 (19)</td>
<td>9.7%</td>
<td>0.0%</td>
<td>97.3%</td>
<td>98.2%</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>14;6-19;8 (11)</td>
<td>10.7%</td>
<td>0.0%</td>
<td>94.0%</td>
<td>94.0%</td>
<td>97.8%</td>
<td>97.8%</td>
</tr>
</tbody>
</table>

Table 3.4.3. Percent ‘same’ actions by adolescent and adult participants
The trends in performance on the test conditions were the same in this sample as in the previous sample (3.3.). Participants in the younger ages showed a much better performance on ‘same’ and ‘share’ than on the definite articles, indicating that they have concepts of sameness, and awareness of the salient entity in the context set. This additionally shows that their lack of adult-like performance on the definite determiners is not due to their lack knowledge of the fact that toy characters can appear in the same space (i.e. on a same apple) together.

This time older children and teenagers also participated. This older group showed perfect knowledge of all determiners. They gave 10% actions on the same object for the indefinite determiner – a bit lower than expected by chance levels, but differing ‘a’ from ‘another’ which got 0% actions on the same object. We believe that a larger sample size as well as a clearer experimental set up showed the true state of adult knowledge, unlike the knowledge shown by only two adults tested on “Puppets” who showed unusually high actions on the same object for indefinite ‘a’.

A mixed repeated measures ANOVA was performed (Determiner (6 levels) as repeated measures by Age group (7 levels) as between subject measure). Mauchly’s test of Sphericity for within subjects effects of determiner ($\chi^2(14) = 117.4, p < .001$) come out significant, hence we
must accept the hypothesis that the variances of the differences between the levels of variables are significantly different, and the assumption of sphericity has been violated. Therefore, degrees of freedom were corrected using the Greenhouse-Geiser estimates of sphericity for determiner ($\epsilon = .572$ for determiner). Results revealed a significant effect for determiner ($F(2.86, 191.8) = 232.8$, $p<.001$), with determiner type by age interaction reaching high significance ($F(17.2, 191.8) = 2.9$, $p < .001$). The effect of age group was significant ($F(6,67)=3.9$, $p=.002$).

Post-Hoc pairwise comparisons (adjusted for multiple comparisons with Bonferroni correction) revealed significant difference between the 3 year old group and the 6 year olds ($p=.034$), the 3s and the 8s ($p=.005$), the 3s and the 9s ($p=.042$). There was also significant difference between all determiners (all $ps <.001$), except for ‘a’ vs ‘another’ ($p=.001$), for ‘that’ vs ‘share’ ($p=.003$), and for ‘that’ vs ‘same’ ($p=.01$). Only ‘share’ vs ‘same’ was not significant.

In dividing participants by their patterns of performance on ‘the’ and ‘that’ (Figure and Table 3.4.5), we find similar trends as in the previous sample. “A” pattern is only shown by 30-40% of 4-6 year olds, but by close to 90% of 8 year olds. In this sample, we find a larger proportion of participants showing the “B” pattern, and not a single child showing the “C” pattern. We also find that 20% of 3 year olds show ‘A’ pattern unlike 0% of 3 year olds in the previous sample. This could be due to a sampling error. It could be because the saliency of the referents exemplified by additional control conditions (‘same’ and ‘share’) influenced children in their identification of the salient referent of ‘that’, but still was powerless to help children identify the salient referent of ‘the’. Why was this powerless to help children identify the salient referent of ‘the’? Because children are missing Maximality/uniqueness.

![Figure and Table 3.4.5. Proportion children performing well](image)

<table>
<thead>
<tr>
<th>age (number of participants)</th>
<th>3(15)</th>
<th>4(19)</th>
<th>5(7)</th>
<th>6(12)</th>
<th>7(9)</th>
<th>8(8)</th>
<th>9(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>0.60</td>
<td>0.42</td>
<td>0.14</td>
<td>0.17</td>
<td>0.22</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>B, that&gt;the</td>
<td>0.20</td>
<td>0.21</td>
<td>0.57</td>
<td>0.42</td>
<td>0.11</td>
<td>0.13</td>
<td>0.00</td>
</tr>
<tr>
<td>A, the=that</td>
<td>0.20</td>
<td>0.37</td>
<td>0.29</td>
<td>0.42</td>
<td>0.67</td>
<td>0.88</td>
<td>1.00</td>
</tr>
</tbody>
</table>
3.5. **Analysis. “Felt” – entire sample**

Participants in the analyses in this section included the combined sample of 203 participants from both ‘Felt’ experiments. This combination is possible because, while there were some differences in the data between the two ‘Felt’ experiments, these are not significant. ANOVA (dependent variable = the pattern of performance that each participant is showing; between subjects factor = experiment (Felt part 1 (experiment 2) vs Felt part 2 (experiment 3))) showed no significant effect of experiment (p=.444). When Age group was added as another between subjects factor, the effect of experiment approached but did not reach significance (F(1,188)=3.4, p=.066), the effect of age was highly significant (F(6,188)=10.7, p<.001), the interaction was not significant (p=.8). Thus we can take advantage of the 203 participant strong sample size and analyze the patterns of performance in fine detail.

3.5.1. **Fine-grained analysis by half-ages**

Given the large number of participants in the combined ‘Felt’ experiment, it is worthwhile to look at children’s development of determiners on a finer scale than a year. For this reason, children are divided into ‘half-age’ groups, i.e. 3-3.49, 3.5-3.9, 4-4.49, 4.5-4.9, etc. A special age of interest here is in the 6-7 year old range, since this is where children experience a change in their knowledge trends in ‘the’ and ‘that’ – that is where the inflection point seems to be in the maturation of uniqueness/Maximality. Children at 6 years show the ‘A’ pattern (good knowledge of ‘the’ and ‘that’) more like the 4 and 5 year olds, children 7 years show the ‘A’ pattern more like the older children. Can we map this inflection point precisely?

### Participants

<table>
<thead>
<tr>
<th>Group</th>
<th>3</th>
<th>3.5</th>
<th>4</th>
<th>4.5</th>
<th>5</th>
<th>5.5</th>
<th>6</th>
<th>6.5</th>
<th>7</th>
<th>7.5</th>
<th>8</th>
<th>8.5</th>
<th>9</th>
<th>9.5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>12</td>
<td>22</td>
<td>15</td>
<td>22</td>
<td>15</td>
<td>13</td>
<td>18</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>5</td>
<td>18</td>
<td>10</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 3.5.1 All participants in ‘Felt’ divided by half-ages
Results

A mixed repeated measures ANOVA was performed (Determiner (4 levels) by Context (2 levels) as repeated measures by half-age group (14 levels) as between subject measure). Mauchly’s test of Sphericity for within subjects effects of determiner ($\chi^2(5) = 247.8$, $p < .001$) and the interaction between determiner and context ($\chi^2(5) = 41.1$, $p < .001$) come out significant, hence we must accept the hypothesis that the variances of the differences between the levels of variables are significantly different, and the assumption of sphericity has been violated. Therefore, degrees of freedom were corrected using the Greenhouse-Geiser estimates of sphericity for determiner ($\varepsilon = .544$ for determiner) and Huynh-Feldt estimates of sphericity for the interaction ($\varepsilon = .969$ for the interaction). For main effect of context, the assumption of sphericity is met. Results revealed a significant effect for determiner ($F(1.63, 308.3) = 371.7$, $p < .001$), with determiner type by age interaction reaching high significance ($F(21.2, 308.3) = 5.1$, $p < .001$). Neither context, nor interaction between context and age reached significance. Interaction between determiner and context was significant ($F(2.9, 549.6) = 9.2$, $p < .001$). The three way interaction (determiner by context by age) was finally significant ($F(37.8, 549.6) = 1.5$, $p = .025$). The effect of age group was significant ($F(13, 189) = 5.7$, $p < .001$).

Post-Hoc pairwise comparisons (adjusted for multiple comparisons with Bonferroni correction) revealed significant difference between the 3s and 4s ($p = .039$), 3s and 5s ($p = .009$), and 3s and 6s-9.5s ($p$s range from $p = .036$ to $p < .001$). 3.5s differ from 5s ($p = .04$), and from 6.5s-9.5s ($p$s range from .04 to <.001), but not from 8s (likely because there are only 5 children in the 8.0-8.49 age range). There was also significant difference between all determiners (all $p$s <.001). There was no difference between contexts.
Analyzing the entire sample and binning the children by half-year groups, we can see the distinct jump in knowledge of ‘the’ in the six year olds. Younger 6s perform similarly to the 4 and 5 year old children. Older 6s perform similarly to 7-8 year old children. This categorical jump in performance is significant. We borrow the logic from Hirsch and Wexler (2006)’s studies of development of syntactic phases via studies of nonactional passives. Question is what drives children’s performance. If it is internal, maturing knowledge, then we expect little variance in the age of attainment of the semantic principles. If it is language-external factors, such as paying attention to the speaker-hearer knowledge distinction, we would expect great variance with respect to the age when individual children acquire definite articles. Clearly, we see that the former is the case for ‘the’ – 40% of children 4-6 show adult-like knowledge, and almost 70% of 6.5 year olds show that knowledge. For children younger than 6.5 the latter seems to be the case for ‘that’. Before knowledge of semantic principles is in place (before 6.5), a subset of children can use the pragmatics in ‘that’ better than lack of semantics in ‘the’ in choosing the salient referent. Once the semantic principle is in place (6.5 and beyond), the majority of children have no need of relying on pragmatics – hence performance on ‘the’ and ‘that’ tracks closely together.

![Figure 3.5.3](chart.png)

**Figure and Table 3.5.3.** Proportion children performing well, by half-ages

<table>
<thead>
<tr>
<th></th>
<th>3 (12)</th>
<th>3.5 (22)</th>
<th>4 (15)</th>
<th>4.5 (22)</th>
<th>5 (15)</th>
<th>5.5 (13)</th>
<th>6 (18)</th>
<th>6.5 (15)</th>
<th>7 (15)</th>
<th>7.5 (15)</th>
<th>8 (5)</th>
<th>8.5 (18)</th>
<th>9 (10)</th>
<th>9.5 (8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>D</td>
<td>C</td>
<td>B</td>
<td>A</td>
<td>0.17</td>
<td>0.05</td>
<td>0.40</td>
<td>0.41</td>
<td>0.40</td>
<td>0.38</td>
<td>0.28</td>
<td>0.67</td>
<td>0.73</td>
<td>0.80</td>
</tr>
<tr>
<td>B</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.06</td>
<td>0.07</td>
<td>0.07</td>
<td>0.20</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.06</td>
</tr>
<tr>
<td>C</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>D</td>
<td>0.83</td>
<td>0.77</td>
<td>0.40</td>
<td>0.55</td>
<td>0.20</td>
<td>0.38</td>
<td>0.44</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
<td>0.11</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

54
3.5.2. Random model in children’s performance?

Next, we investigate patterns of performance relative to a random model of responses for each determiner. The random model here was calculated for 3 presentations of 3-item contexts and 3 presentations of 6-item contexts, i.e. we calculated the probability that children will give us 0/6, 1/6, etc ‘same’ responses (Figure and Table 3.5.4). Such performance might be expected for ‘a’ (which means ‘any one’), but not for ‘another’. If younger children’s knowledge of ‘the’ is deficient, we might see some random responses in ‘the’ too. If children are picking their referents by chance, we expect a great deal of 1/6 and 2/6 ‘same’ responses, some 0/6 and 3/6 responses, and practically no 4/6 or greater ‘same’ responses as those are unlikely by chance.

We can see that with ‘Felt’ set-up, as opposed to ‘Puppets’ set-up, we observe practically no random patterns of responses in groups of subjects. (Figures 3.5.5-3.5.8).

It follows from this data, that there are almost no ages when children show such random-model profile of answers in any of the determiners in ‘Felt’ task.

<table>
<thead>
<tr>
<th>Number actions on the ‘same’ object</th>
<th>Random model estimate for 3<em>3 and 6</em>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.17147</td>
</tr>
<tr>
<td>1</td>
<td>0.36008</td>
</tr>
<tr>
<td>2</td>
<td>0.30350</td>
</tr>
<tr>
<td>3</td>
<td>0.13083</td>
</tr>
<tr>
<td>4</td>
<td>0.03035</td>
</tr>
<tr>
<td>5</td>
<td>0.00360</td>
</tr>
<tr>
<td>6</td>
<td>0.00017</td>
</tr>
</tbody>
</table>

Figure and Table 3.5.4. probability of performance on random model.

![Bar graph showing probability of performance on random model](image)

Figure 3.5.5. Proportion of children giving 0, 1, 2, 3, 4, 5, or 6 ‘same’ responses for ‘another’
Figure 3.5.6. Proportion of children giving 0, 1, 2, 3, 4, 5, or 6 ‘same’ responses for ‘a’

Figure 3.5.7. Proportion of children giving 0, 1, 2, 3, 4, 5, or 6 ‘same’ responses for ‘the’
Proportion of children who get ‘the’ (at 4/6 or better) plateaus from 4 to 6 years. Proportion of children who get ‘that’ varies within that same plateau. Once children are 6.5 years old or older, there is no wide variation between ‘the’ and ‘that’. This pattern suggests that before 6.5, children who are getting ‘that’ are doing so via some interpretational strategy which some children get and some children don’t – hence the variation. Note that the same young fives that get almost as many correct referents with ‘a’, and ‘another’ as they do with ‘that’, only 40% of these same young fives get correct referent with ‘the’. If children do not know uniqueness in ‘the’ and ‘that’, they can think ‘the’ means ‘one of salient things in context set’, and they can think that ‘that’ also means ‘one of salient things in the context set’. A reasonable strategy is that children take ‘that’ to indicate interpretation relative to non-default context, which for children may focus the interpretation on the true salient entity in the context. Once children know uniqueness (by 6.5 years), that knowledge is sufficient for them to interpret the salient entity in the discourse set.
Three year olds primarily show 0/6 responses for all determiners, although notably there are few ‘smart’ three-year olds who show 4/6 and better on ‘the’ and ‘that’, but those are very unique individuals. For ‘another’, 86% of younger children and 94% of older children show 0/6 ‘same’ responses, with other children showing some 1/6 and 2/6 responses. For ‘a’, 60% of children give 0/6, 25% give 1/6 responses, and about 20% of children give 2/6 and 3/6 responses – this pattern does not change with age. For ‘the’ and ‘that’ three year olds show primarily 0/6 responses with some 1/6 responses. In all other ages, children give many more 6/6 responses for ‘that’ than for ‘the’, almost twice as many (19% vs 38% in 4-6s, 51% vs 73% in 6.5-9s). At younger ages for ‘the’ about 10% of children show each of 1/6, 2/6, 3/6, 4/6, 29% show 0/6, and 14% show 5/6. For the same younger ages using ‘that’ to indicate the referent pushes many these children toward 6/6 responses. At older ages, there are still some children showing poor performance on ‘the and ‘that’, but what is notable is that many more children are sure of the referent of ‘that’ than of ‘the’: older children show 6/6 by a third more for ‘that’ than ‘the’, and show 5/6 more for ‘the’ than for ‘that’ (26% vs 12%).

In other words, children’s average ‘50%’ ‘same’ responses performance on ‘the’ at younger ages, is not that. There is almost bimodal distribution of participants. There are around half of participants who show adult-like knowledge (4/6 or higher). With the rest of participants showing poor performance (3/6 or lower), which is consistent with chance performance according to our model of random responses.
3.5.3. Gender effects?

Participants

<table>
<thead>
<tr>
<th>Age</th>
<th>Gender</th>
<th>Mean</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>f</td>
<td>3.47</td>
<td>17</td>
<td>3.00</td>
<td>3.95</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>3.59</td>
<td>17</td>
<td>2.92</td>
<td>3.98</td>
</tr>
<tr>
<td>4</td>
<td>f</td>
<td>4.57</td>
<td>18</td>
<td>4.08</td>
<td>4.97</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>4.51</td>
<td>19</td>
<td>4.00</td>
<td>4.87</td>
</tr>
<tr>
<td>5</td>
<td>f</td>
<td>5.42</td>
<td>15</td>
<td>5.00</td>
<td>5.91</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>5.42</td>
<td>12</td>
<td>5.01</td>
<td>5.64</td>
</tr>
<tr>
<td>6</td>
<td>f</td>
<td>6.60</td>
<td>22</td>
<td>6.10</td>
<td>6.98</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>6.27</td>
<td>12</td>
<td>6.00</td>
<td>6.58</td>
</tr>
<tr>
<td>7</td>
<td>f</td>
<td>7.47</td>
<td>19</td>
<td>7.05</td>
<td>7.96</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>7.48</td>
<td>11</td>
<td>7.02</td>
<td>7.90</td>
</tr>
<tr>
<td>8</td>
<td>f</td>
<td>8.67</td>
<td>13</td>
<td>8.10</td>
<td>8.94</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>8.55</td>
<td>10</td>
<td>8.00</td>
<td>8.99</td>
</tr>
<tr>
<td>9</td>
<td>f</td>
<td>9.53</td>
<td>10</td>
<td>9.00</td>
<td>9.93</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>9.30</td>
<td>8</td>
<td>9.06</td>
<td>9.80</td>
</tr>
<tr>
<td>Total</td>
<td>F</td>
<td>5.89</td>
<td>114</td>
<td>3.00</td>
<td>9.93</td>
</tr>
<tr>
<td>Total</td>
<td>M</td>
<td>5.82</td>
<td>89</td>
<td>2.92</td>
<td>9.80</td>
</tr>
</tbody>
</table>

Table 3.5.10. Participants in “Felt” divided by age and by gender

Results

A mixed repeated measures ANOVA was performed (Determiner (4 levels) by Context (2 levels) as repeated measures by Age group (7 levels) and gender (2 levels) as between subject measures). Mauchly’s test of Sphericity for within subjects effects of determiner ($\chi^2(5) = 250.1$, $p < .001$) and the interaction between determiner and context ($\chi^2(5) = 41.9$, $p < .001$) come out significant, hence we must accept the hypothesis that the variances of the differences between the levels of variables are significantly different, and the assumption of sphericity has been violated. Therefore, degrees of freedom were corrected using the Greenhouse-Geiser estimates of sphericity for determiner ($\epsilon = .543$ for determiner) and Huynh-Feldt estimates of sphericity for the interaction ($\epsilon = .97$ for the interaction). For main effect of context, the assumption of sphericity is met. Results revealed a significant effect for determiner ($F(1.63, 307.9) = 395.5$, $p<.001$), with determiner type by age interaction reaching high significance ($F(9.8, 307.9) = 9.9$, $p < .001$). Neither determiner by gender interaction, nor the determiner by age by gender interaction, nor context, nor interaction between context and age, nor the interaction between context and gender, nor the interaction between context age and gender reached significance. Interaction between determiner and context was significant ($F(2.9,550.7)=9.6$, $p<.001$). The three way interaction between determiner, context and age approached significance ($F(17.5,550.7)=1.6$, $p=.065$). The three way interaction between determiner, context, and gender and the four way interaction were not significant. The between subjects effect of age group was
significant \((F(6,189)=10.6, \ p<.001)\), but there was no effect of gender, and no interaction between age and gender.

**Figure 3.5.11. % actions on the ‘same’object in females and males**

**Figure 3.5.12. Proportion of females and males performing well**

While there is no significant effect of gender and no interactions between determiner and gender, there is noticeable difference in graphs in performance between females and males ‘the’ at ages 6 and 7. Females’ development takes off by 6 years and by 7 years over 70% of them show good performance. Males’ development lags behind, and takes off only by 7 years and reaches good performance in majority of children only by 8 years. Furthermore, more males show “B” pattern at age 6 than females. Thus males are a year behind females in developing semantic knowledge.
3.5.4. Context effects?

Participants

<table>
<thead>
<tr>
<th>Age</th>
<th>Mean</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>3.53</td>
<td>34</td>
<td>2.92</td>
<td>3.98</td>
</tr>
<tr>
<td>4</td>
<td>4.54</td>
<td>37</td>
<td>4.00</td>
<td>4.97</td>
</tr>
<tr>
<td>5</td>
<td>5.42</td>
<td>27</td>
<td>5.00</td>
<td>5.91</td>
</tr>
<tr>
<td>6</td>
<td>6.48</td>
<td>34</td>
<td>6.00</td>
<td>6.98</td>
</tr>
<tr>
<td>7</td>
<td>7.48</td>
<td>30</td>
<td>7.02</td>
<td>7.96</td>
</tr>
<tr>
<td>8</td>
<td>8.62</td>
<td>23</td>
<td>8.00</td>
<td>8.99</td>
</tr>
<tr>
<td>9</td>
<td>9.43</td>
<td>18</td>
<td>9.00</td>
<td>9.93</td>
</tr>
</tbody>
</table>

Table 3.5.13. Participants in “Felt” by age

Results

Below are Random models for potential probabilistic replies for picking the same entity zero, once, twice or three times, in a 3-item or a 6-item context given 3 presentations of each. As in the ‘puppets’ experiment, subjects should be more likely, by chance, to pick the same objects in the smaller sized context (see discussion above in the ‘effect of context size’ section of ‘puppets’ experiment). In other words, chance performance in the 3-item context should be 33%, and chance performance in the 6-item context should be 16%. Thus the expected difference between 3-item and 6-item contexts should be around 16%.

<table>
<thead>
<tr>
<th>Number 'same' actions</th>
<th>3-ITEM context</th>
<th>6-ITEM context</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.296</td>
<td>0.579</td>
</tr>
<tr>
<td>1</td>
<td>0.444</td>
<td>0.347</td>
</tr>
<tr>
<td>2</td>
<td>0.222</td>
<td>0.069</td>
</tr>
<tr>
<td>3</td>
<td>0.037</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Figure and Table 3.5.14 Random model context
While in the repeated measures ANOVA (above in half-age analysis) there was no significant effect of context, interaction between determiner and context was significant ($F(2.9,549.6)=9.2, p<.001$). The three way interaction (determiner by context by half-ages) was significant ($F(37.8,549.6)=1.5, p=.025$).

This time around, participants have a tendency to perform slightly more actions on the same object as expected for 3-item rather than 6-item contexts only for ‘a’; for ‘the’, participants perform slightly more action on the same object for 6-item contexts rather than for 3-item contexts – contrary to random model effects. There are no context effects for ‘another’ and hardly any ‘that’ (Figure 3.5.15).

Notably, the highest differences for ‘the-3’ vs ‘the-6’ occur in children of ages 4, 5, and 6 (9%, 10%, and 13% respectively) – the time where they are at the developmental plateau. From age 7 onwards the effect of context almost disappears (less than 3% with the exception of 8 year olds (7%)). For ‘a-3’ and ‘a-6’ the largest differences are in the 4 and 5 year old children (12% and 9% respectively), appearing somewhat less in the 6 (4%), 7 (8%) and 9 (8%) year olds, and practically absent in 8 year olds. For ‘that-3’ and ‘that-6’ the differences are less than 3% for all ages except for 6 year olds (7%). For ‘another-3’ and ‘another-6’ the differences are less than 3% for all ages.

So, as before, for children’s interpretation of ‘that’, there is no room for probabilities. For ‘another’ we find no variation either. However for ‘the’ and ‘a’ context is playing a role. For the indefinite article ‘a’ the context effect is as predicted by the random model – since ‘a’ means ‘any one’, children show that they know this definition by picking slightly more ‘same’ objects when there are fewer referents in the context set. For the definite article ‘the’ the context effect is the opposite of what is predicted by the random model. What seems to be going on in the younger children is that the larger context set is forcing them to pay more attention to the salient referent. Perhaps while young children are able to hold three items in their attention at once, and thereby assigning them equal saliency, young children are not able to hold six items in their attention at once, and therefore focus on the truly salient referent.
3.5.5. Context effects vs patterns of performance

To explore the effect of context further, we divide participants into those who show differential performance in 3 vs 6 contexts (group 1) and those who do not (group 0), and evaluate their patterns of knowledge. Manipulation of the context size (3 items vs 6 items) showed that around 30% of children are sensitive to it. Recall that while in case of ‘a’ children show semi-random performance and chose the salient referent slightly more often in the 3-item context than in the 6-item context, in case ‘the’ children chose the salient referent more in the 6-item context – contrary to random-effects models.

Table 3.5.16. Proportion children (not) showing sensitivity to context size by age.

<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>average</th>
</tr>
</thead>
<tbody>
<tr>
<td>the</td>
<td>76%</td>
<td>82%</td>
<td>54%</td>
<td>68%</td>
<td>57%</td>
<td>71%</td>
<td>64%</td>
<td>67%</td>
</tr>
<tr>
<td>the</td>
<td>82%</td>
<td>76%</td>
<td>68%</td>
<td>54%</td>
<td>71%</td>
<td>64%</td>
<td>67%</td>
<td>57%</td>
</tr>
<tr>
<td>a</td>
<td>54%</td>
<td>68%</td>
<td>71%</td>
<td>64%</td>
<td>73%</td>
<td>73%</td>
<td>61%</td>
<td>67%</td>
</tr>
<tr>
<td>a</td>
<td>68%</td>
<td>54%</td>
<td>71%</td>
<td>64%</td>
<td>73%</td>
<td>73%</td>
<td>61%</td>
<td>67%</td>
</tr>
<tr>
<td>the</td>
<td>64%</td>
<td>67%</td>
<td>73%</td>
<td>73%</td>
<td>61%</td>
<td>67%</td>
<td>67%</td>
<td>65%</td>
</tr>
<tr>
<td>the</td>
<td>67%</td>
<td>67%</td>
<td>65%</td>
<td>67%</td>
<td>67%</td>
<td>67%</td>
<td>65%</td>
<td>72%</td>
</tr>
<tr>
<td>a</td>
<td>67%</td>
<td>67%</td>
<td>65%</td>
<td>67%</td>
<td>67%</td>
<td>67%</td>
<td>65%</td>
<td>72%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>average</th>
</tr>
</thead>
<tbody>
<tr>
<td>the</td>
<td>24%</td>
<td>18%</td>
<td>46%</td>
<td>32%</td>
<td>43%</td>
<td>29%</td>
<td>36%</td>
<td>33%</td>
</tr>
<tr>
<td>the</td>
<td>18%</td>
<td>24%</td>
<td>46%</td>
<td>32%</td>
<td>43%</td>
<td>29%</td>
<td>36%</td>
<td>33%</td>
</tr>
<tr>
<td>a</td>
<td>46%</td>
<td>32%</td>
<td>43%</td>
<td>29%</td>
<td>36%</td>
<td>33%</td>
<td>27%</td>
<td>27%</td>
</tr>
<tr>
<td>a</td>
<td>32%</td>
<td>46%</td>
<td>43%</td>
<td>29%</td>
<td>36%</td>
<td>33%</td>
<td>27%</td>
<td>27%</td>
</tr>
<tr>
<td>the</td>
<td>36%</td>
<td>33%</td>
<td>27%</td>
<td>27%</td>
<td>39%</td>
<td>26%</td>
<td>33%</td>
<td>33%</td>
</tr>
<tr>
<td>the</td>
<td>33%</td>
<td>33%</td>
<td>27%</td>
<td>27%</td>
<td>39%</td>
<td>26%</td>
<td>33%</td>
<td>33%</td>
</tr>
<tr>
<td>a</td>
<td>33%</td>
<td>33%</td>
<td>27%</td>
<td>27%</td>
<td>39%</td>
<td>26%</td>
<td>33%</td>
<td>33%</td>
</tr>
<tr>
<td>a</td>
<td>33%</td>
<td>33%</td>
<td>27%</td>
<td>27%</td>
<td>39%</td>
<td>26%</td>
<td>33%</td>
<td>33%</td>
</tr>
<tr>
<td></td>
<td>34%</td>
<td>38%</td>
<td>34%</td>
<td>38%</td>
<td>34%</td>
<td>38%</td>
<td>34%</td>
<td>38%</td>
</tr>
</tbody>
</table>
For a3 vs a6: No difference participants

Figure 3.5.17. % actions on the same object by context

Figure 3.5.18. Proportion children giving particular performances

Table 3.5.19. Paired t-tests for those not differentiating between ‘a’ contexts

<table>
<thead>
<tr>
<th>a context = 0</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANO3 - ANO6</td>
<td>-0.576</td>
<td>145</td>
<td>0.565</td>
</tr>
<tr>
<td>THE3 - THE6</td>
<td>-4.168</td>
<td>145</td>
<td>5.26E-05</td>
</tr>
<tr>
<td>THAT3 - THAT6</td>
<td>-0.686</td>
<td>145</td>
<td>0.494</td>
</tr>
<tr>
<td>AA3 - THE3</td>
<td>-12.126</td>
<td>145</td>
<td>0</td>
</tr>
<tr>
<td>AA6 - THE6</td>
<td>-13.933</td>
<td>145</td>
<td>0</td>
</tr>
<tr>
<td>THE3 - THAT3</td>
<td>-6.22</td>
<td>145</td>
<td>0</td>
</tr>
<tr>
<td>THE3 - THAT6</td>
<td>-6.324</td>
<td>145</td>
<td>0</td>
</tr>
<tr>
<td>THE6 - THAT3</td>
<td>-2.746</td>
<td>145</td>
<td>0.006795</td>
</tr>
<tr>
<td>THE6 - THAT6</td>
<td>-3.238</td>
<td>145</td>
<td>0.001491</td>
</tr>
<tr>
<td>AA3 - ANO3</td>
<td>1.814</td>
<td>145</td>
<td>0.072</td>
</tr>
<tr>
<td>AA3 - ANO6</td>
<td>1.351</td>
<td>145</td>
<td>0.179</td>
</tr>
<tr>
<td>AA6 - ANO3</td>
<td>1.814</td>
<td>145</td>
<td>0.072</td>
</tr>
<tr>
<td>AA6 - ANO6</td>
<td>1.351</td>
<td>145</td>
<td>0.179</td>
</tr>
</tbody>
</table>

Those who do not differentiate ‘a’ contexts (72% of participants), show context differences on ‘the’. (reach 4/6 by 7 years of age). These children seem more categorical in their reasoning, e.g. ‘a’ = another one (indeed ‘a’ performance tracks with ‘another’ closely). They also show context differences in ‘the’. It seems that their semantics does not provide a clear definition of a given determiner, they try to guess (use a strategy). Relative saliency of an item relative to its context seems to be what’s going on in ‘the’. It seems that what is going on is in The3 children can take in all three items, so technically, any of the three can be salient – look, they are right in front of the child, whether or not fishy/turtle is on them. But in The6 – there are too many items for children to take in at once, so ‘salient’ is more likely to be the on with fishy/turtle on it.
For a3 vs a6: with difference participants

![Graphs showing% actions on the same object by context and proportion children giving particular performances.](Image)

**Figure 3.5.20. % actions on the same object by context**  
**Figure 3.5.21. Proportion children giving particular performances**

<table>
<thead>
<tr>
<th>a context</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA6 - AA3</td>
<td>-3.711</td>
<td>56</td>
<td>0.0004767</td>
</tr>
<tr>
<td>ANO3 - ANO6</td>
<td>1.694</td>
<td>56</td>
<td>0.096</td>
</tr>
<tr>
<td>THE3 - THE6</td>
<td>-1.154</td>
<td>56</td>
<td>0.253</td>
</tr>
<tr>
<td>THAT3 - THAT6</td>
<td>0.704</td>
<td>56</td>
<td>0.484</td>
</tr>
<tr>
<td>AA3 - THE3</td>
<td>-4.596</td>
<td>56</td>
<td>0</td>
</tr>
<tr>
<td>AA6 - THE6</td>
<td>-8.115</td>
<td>56</td>
<td>0</td>
</tr>
<tr>
<td>THE3 - THAT3</td>
<td>-3.647</td>
<td>56</td>
<td>0.001</td>
</tr>
<tr>
<td>THE3 - THAT6</td>
<td>-2.32</td>
<td>56</td>
<td>0.024</td>
</tr>
<tr>
<td>THE6 - THAT3</td>
<td>-1.847</td>
<td>56</td>
<td>0.070</td>
</tr>
<tr>
<td>THE6 - THAT6</td>
<td>-1.53</td>
<td>56</td>
<td>0.132</td>
</tr>
<tr>
<td>AA3 - ANO3</td>
<td>8.204</td>
<td>56</td>
<td>0</td>
</tr>
<tr>
<td>AA3 - ANO6</td>
<td>8.3</td>
<td>56</td>
<td>0</td>
</tr>
<tr>
<td>AA6 - ANO3</td>
<td>3.188</td>
<td>56</td>
<td>0.002</td>
</tr>
<tr>
<td>AA6 - ANO6</td>
<td>4.014</td>
<td>56</td>
<td>0</td>
</tr>
</tbody>
</table>

Those who do differentiate ‘a’ contexts (28% of participants), show NO context differences on ‘the’. (reach 4/6 by 6 years of age). These children seem more sensitive (sure of) to semantics, e.g. ‘a’ = “any one” (adult meaning), in fact in 4 year olds, A3 = The (!), and random model context effects are evident with this interpretation. These children are also somewhat more likely to show ‘adult’ interpretation of ‘the’ earlier.
For the3 vs the6: No difference participants

![Graph showing actions on the same object by context](image1)

![Graph showing proportion of children giving particular performances](image2)

**Figure 3.5.23. % actions on the same object by context**

**Figure 3.5.24. Proportion children giving particular performances**

**Table 3.5.25. Paired t-tests for those not differentiating between ‘the’ contexts**

<table>
<thead>
<tr>
<th>the context =0</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA6 - AA3</td>
<td>-1.377</td>
<td>130</td>
<td>0.171</td>
</tr>
<tr>
<td>ANO3 - ANO6</td>
<td>2.023</td>
<td>130</td>
<td>0.045</td>
</tr>
<tr>
<td>THAT3 - THAT6</td>
<td>0.294</td>
<td>130</td>
<td>0.769</td>
</tr>
<tr>
<td>AA3 - THE3</td>
<td>-10.955</td>
<td>130</td>
<td>0</td>
</tr>
<tr>
<td>AA6 - THE6</td>
<td>-11.248</td>
<td>130</td>
<td>0</td>
</tr>
<tr>
<td>THE3 - THAT3</td>
<td>-3.486</td>
<td>130</td>
<td>0.001</td>
</tr>
<tr>
<td>THE3 - THAT6</td>
<td>-3.077</td>
<td>130</td>
<td>0.003</td>
</tr>
<tr>
<td>THE6 - THAT3</td>
<td>-3.486</td>
<td>130</td>
<td>0.001</td>
</tr>
<tr>
<td>THE6 - THAT6</td>
<td>-3.077</td>
<td>130</td>
<td>0.003</td>
</tr>
<tr>
<td>AA3 - ANO3</td>
<td>3.193</td>
<td>130</td>
<td>0.002</td>
</tr>
<tr>
<td>AA3 - ANO6</td>
<td>3.707</td>
<td>130</td>
<td>0</td>
</tr>
<tr>
<td>AA6 - ANO3</td>
<td>2.446</td>
<td>130</td>
<td>0.016</td>
</tr>
<tr>
<td>AA6 - ANO6</td>
<td>3.239</td>
<td>130</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Those who do not differentiate ‘the’ contexts (65% of participants), also show no context differences on ‘a’. (reach 4/6 by 7 years of age).
For the3 vs the6: with difference participants

![Graph showing number of responses vs age]  ![Graph showing proportion of participants]  

**Figure 3.5.26. % actions on the same object by context**  
**Figure 3.5.27. Proportion children giving particular performances**

**Table 3.5.28. Paired t-tests for those differentiating between ‘the’ contexts**

<table>
<thead>
<tr>
<th>the context</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA6 - AA3</td>
<td>-3.753</td>
<td>71</td>
<td>0.0003541</td>
</tr>
<tr>
<td>ANO3 - ANO6</td>
<td>-0.241</td>
<td>71</td>
<td>0.81</td>
</tr>
<tr>
<td>THE3 - THE6</td>
<td>-4.349</td>
<td>71</td>
<td>4.49E-05</td>
</tr>
<tr>
<td>THAT3 - THAT6</td>
<td>-0.478</td>
<td>71</td>
<td>0.634</td>
</tr>
<tr>
<td>AA3 - THE3</td>
<td>-6.457</td>
<td>71</td>
<td>0</td>
</tr>
<tr>
<td>AA6 - THE6</td>
<td>-12.804</td>
<td>71</td>
<td>0</td>
</tr>
<tr>
<td>THE3 - THAT3</td>
<td>-7.577</td>
<td>71</td>
<td>0</td>
</tr>
<tr>
<td>THE3 - THAT6</td>
<td>-6.775</td>
<td>71</td>
<td>0</td>
</tr>
<tr>
<td>THE6 - THAT3</td>
<td>-1.276</td>
<td>71</td>
<td>0.206</td>
</tr>
<tr>
<td>THE6 - THAT6</td>
<td>-1.905</td>
<td>71</td>
<td>0.061</td>
</tr>
<tr>
<td>AA3 - ANO3</td>
<td>5.958</td>
<td>71</td>
<td>0</td>
</tr>
<tr>
<td>AA3 - ANO6</td>
<td>4.916</td>
<td>71</td>
<td>0</td>
</tr>
<tr>
<td>AA6 - ANO3</td>
<td>2.559</td>
<td>71</td>
<td>0.013</td>
</tr>
<tr>
<td>AA6 - ANO6</td>
<td>2.218</td>
<td>71</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Those who do differentiate ‘the’ contexts (35% of participants), also show context differences on ‘a’. They also reach good performance only by 8 years. Note these participants are evenly distributed across the two ‘felt’ experiments. ‘Felt’ part one had 45 of 129. ‘Felt’ part two had 27 of 74. These participants show context effect on ‘a’ that goes in the direction predicted by the random model, so it is not the fact they are not subject to chance performance. However in case of ‘the’ something is going on. Performance on The6 tracks closely with performance on ‘That’. Performance on ‘the3’ looks exactly like the random model of responses. At the same time, participants are differentiating ‘the3’ from ‘a3’. The fact that only 50% of children are showing adult-like knowledge till they are 8 years old suggests that their semantics does not
provide a clear definition of a given determiner, and that they try to guess (use a strategy). Relative saliency of an item relative to its context seems to be what’s going on in ‘the’. It seems that what is going on is in ‘The3’ children can take in all three items, so technically, any of the three can be salient – look, they are right in front of the child, whether or not Fishy/Turtle is on them. But in ‘The6’ – there are too many items for children to take in at once, so ‘salient’ is more likely to be the one with Fishy/Turtle on it. This is evidence supporting the idea that children who are lacking uniqueness are variable in their choice of the context of the utterance – which could be the whole context set or the unique salient referent.
Figure 3.5.29. Proportion of participants giving 0, 1, 2 or 3 ‘same’ answers.
3.6. Discussion of TD Results

We presented data on comprehension of determiners in children as tested on act-out tasks. The classical act-out task in which children physically pushed around actors and referents (“Puppets”) already showed the developmental trends in definite and indefinite determiners later replicated by the more constrained act-out task made of felt (“Felt”), which were in turn replicated in another sample of participants on ‘Felt’ with additional conditions, however the results from ‘Puppets’ were far noisier. The investigated effects of random models of performance as well as effects of differences in context size were very much pronounced in ‘Puppets’, and following the data from ‘Felt’, it can be safely concluded that majority, but notably not all, of those effects were due to straying children’s attention due to experimental set-up issues.

MANOVA (dependent variables = percent ‘same’ responses for ‘a’, ‘the’ and ‘that’; independent variable = task (‘Felt’ or ‘Puppets’) did not show a significant difference between ‘Felt’ and ‘Puppets’ for ‘the’ and ‘that’, but did show it for ‘a’ (F(1,290)=72.1, p<.001). When age group was added as another between subjects factor, the effect of task did not change, the effect of age was highly significant for ‘the’ (F(6,278)=8.5, p<.001) and ‘that’ (F(6,278)=10.2, p<.001), not significant for ‘a’. The interaction between age and task was not significant for any determiner.

Figure 3.6.1. Development of Determiners on ‘Puppets’ and ‘Felt’ tasks

One notable difference was in the 3 year old group on ‘the’ and ‘that’ who seemed to perform relatively better on ‘the’ and ‘that’ in ‘Puppets’ than on ‘Felt’ (about 20% difference). Are they are avoiding acting with Fishy and Turtle on the same object in ‘Felt’ – is it a strategy to always pick a different one? Perhaps it is. Perhaps, the difference here is in the memory demands of ‘Puppets’. If young children have trouble remembering in ‘Puppets’ what just happened, we expect to see random model (chance) performance on all determiners – which is close to what we observe. In ‘Felt’ however, there are no memory demands – the salient unique referent is clearly
visible, and children are less likely to show random model performance where by chance they may land on the salient referent some of the time. In this regard, ‘Felt’ is a clearer test than ‘Puppets’ of children’s knowledge – children’s competence, since in ‘Felt’ performance issues (e.g. memory demands) are minimized. The same notable difference between ‘Puppets’ and ‘Felt’, although going in the opposite direction, occurs in the 7-9 year old participants. It seems their semantic competence is obscured by the cognitive demands of ‘Puppets’. In the 4-6 year old children however, the nature of the task did not seem to influence their performance.

Another notable difference, across all participants, is in ‘a’. In ‘Puppets’ participants picked the salient referent for ‘a’ around 30%, while it was around 10% in ‘Felt’. This could be explained by random model performance in ‘Puppets’, but a very much reduced random model performance in ‘Felt’, which perhaps allowed participants to clearly invoke maximize presupposition (if ‘a’ is used and not ‘the’, that indicates a nonsalient referent).

Overall, the data produced by testing children on ‘Felt’, a more constrained version of an act out task, produced clear results. Three year olds showed almost no knowledge of ‘the’ or ‘that’ in comprehension, but were able to pick out the salient referent as referred to by ‘same’ or ‘share’, and were able to pick out the unsalient referent as referred to by ‘another’. From age 4 till age 6 children plateaued in their knowledge. 40% of children showed adult-like knowledge of ‘the’ and ‘that’, between 5% and 40% of these children understood ‘that’ to refer to the salient referent in the context set better than ‘the’. The opposite pattern was found only in two children of age six. A big jump in knowledge is observed in seven year olds – now 70% of them show adult-like knowledge, as do 80% of 8 year olds.

About 30% of participants at all ages showed sensitivity to the size of the context set – by chance levels children should be more likely to pick the salient referent in a smaller context set size (3-item) than in a larger context set size (6-item). In ‘Puppets’, we found such effects of context set on ‘a’ and ‘the’, but not on ‘that’. In ‘Felt’, we confirmed these effects on interpretation of the indefinite article ‘a’ and a lack of these effects on interpretation of ‘that’. It seems when it comes to ‘that’ children are certain of their interpretation and are not susceptible to context effects. With ‘the’ we found something new in ‘Felt’. Children were going against chance
levels, and were picking more same objects when there were 6 items in the context set. The only logical explanation is that children are able to make all three objects in a context set salient, but they are unable to make all six objects in the context set salient, perhaps due to attentional/working memory issues, and hence focus their attention on the truly salient item in the context set – the referent of ‘the’.

Thus we show that children can identify the salient entity in the discourse set as early as 3-4 years (they know ‘another’, and ‘same’, and sometimes ‘that’), but at the same time children do not know the salient entity with ‘the’. For the logic of pragmatic deficit theories – whether egocentrism or others – this is a paradox. For egocentric/pragmatic view to work in comprehension, children should be able to pick their own salient referents – ones that are different from the salient referents implied by the speaker. But children can and do pick correct salient referent with ‘same’. For the logic of semantic deficit theories – this is supportive evidence. If children are deficient in their knowledge of speaker-listener distinctions, and/or in their ability to pay attention to the truly salient entity in the context set, it should not make much difference which determiner is used to refer to the salient referent. If children are deficient in their knowledge of the semantic principle of uniqueness / maximality, which article is used to refer to the salient referent makes all the difference: if children are forced to rely on uniqueness, as in case of ‘the’, when they lack uniqueness, which is at least through age 6, their interpretation breaks down; however if children realize they do not have to rely on uniqueness but instead can rely on pragmatic features of familiarity or the feature signifying non-default context interpretation, as in the case of ‘that’, they show more successful comprehension.

In our future publications we will address the question – what factors can explain the presence or the absense of knowledge of ‘the’ and ‘that’ in children – and we will investigate children’s nonverbal reasoning, vocabulary levels, overall grammatical comprehension levels, as well as children’s syntactic knowledge of actional and psychological passives, and children’s syntactic and pragmatic interpretation of personal and reflexive pronouns (Modyanova, Perovic, Hirsch, Wexler, in preparation).

3.7. Acknowledgements

Our heartfelt thanks go to all the participants, their families, their teachers, and the child care centers that allowed us to carry out these studies on their premises.

The biggest thanks goes to Tess (Tara) Deduch, who helped me design the Felt book stimuli. The next biggest thanks goes to the undergraduate research assistants who have worked with me, helped me design stimuli, and who helped me run the subjects (in alphabetical order by first name): Alexandra Huston-Carico, Allegra Cornaglia, Amy LeMessurier, Avril Kenney, Charles Agoos, Charlotte Yang, Christopher Watson, Clare Dean, David Mathieson, Dolapo (Nikki) Longe, Elisabeth Lex, Elisse Lockhart, Elizabeth Boehm, Elizabeth Roman, Frances Choi, Isadora Nahmanson, Issac Buenrostro, Katherine Boothe, Kathryn Germer, Leah Bostead, Margaret E Echelbarger, Matthew McKinley, Orit Giguinsky, Rafael Raya, Ritu Tandon, Rob Wells, Sang-Hee Min, Ted Kang, Tiffany Ho, Tracey Ragsdale, Veronica Cole, Yuki Yukung.
4. Further Investigations of Typical Acquisition

In Chapter 3, we presented results suggesting that children’s deficits in ‘the’ are semantic (due to immature uniqueness/Maximality), and not pragmatic/egocentric in nature. To fully support such a view, it is necessary to further investigate children’s semantic and pragmatic (other mind awareness) development, in way that distinctly separates them.

This chapter presents results from two experiments from the opposite ends of theoretical continuum. Experiment 4 investigated potential correlations between comprehension of determiners and development of theory of mind, as way to directly test children’s awareness of other’s points of view and whether that is involved in comprehension of determiners. Experiment 5 investigated children’s comprehension of Free Relative clauses – constructions which unequivocally invoke Maximality in certain contexts, and which, we would like to argue, should not involve any pragmatic considerations in their interpretations and hence may present a challenge for egocentric/pragmatic accounts. Together, these experiments provide further evidence for the semantic deficits hypothesis.

4.1. Experiment 4. Theory of Mind (Not) in Determiner Acquisition

4.1.1. Abstract / Motivation

The interconnection between development of theory of mind and development of language is a hotly debated topic. Some show that TOM depends on language development, especially sentential clauses, others show that’s not the case (reviewed below). But can theory of mind really be involved in definite article acquisition? This is an implicit assumption of many elicited production studies and pragmatic theories of acquisition of determiners where children’s deficit in ‘the’ is attributed to children’s inability to consider the point of view of their listener (reviewed in section 2.3). Theory of mind in determiners was explicitly considered by Gundel (2009) (where the fact that children use determiners early in spontaneous speech is suggestive of knowledge of TOM), Schaeffer and Matthewson (2005), and Shaefer and de Villiers (2000). To reiterate, many studies investigate whether language contributes to theory of mind, but those studies that investigate determiners suggest that theory of mind contributes to language development (at least in determiners) – the two fields have opposing points of view!

For all the talk about egocentrism and theory of mind in determiners, nobody before us has actually investigated this in acquisition.

Thus, an experiment developed in collaboration with Charlotte Giessman is testing in parallel the development of articles (using the ‘felt’ task) and a nonverbal theory of mind test (adapted from work by Baron-Cohen and Tomasello) which excludes any potential linguistic confounds. Preliminary results show similarity in patterns of acquisition on group levels, however there is no correlation on an individual level, thus some children are excellent at theory of mind, yet poor on articles, and the opposite pattern is also found. Thus it seems that the hypothesis of egocentrism/theory of mind for acquisition of articles, while plausible, does not stand up to the test.
4.1.2. Language in Theory of mind?

The topic of the role of language in theory of mind has been already raised in Chapter 2. An important question is whether language (grammar/syntax, vocabulary) or nonverbal intelligence affects Theory of Mind development (or vise versa), and to what extent can a test of TOM measure development of social reasoning independently of other factors. Harris, de Rosnay, Pons (2005) argue that children with advanced language skills (i.e. those who are exposed to lexical and syntactic enrichment through maternal conversations) are better at mental state understanding. Studies by de Villiers demonstrate that typical children need to have syntactic structure for sentential complements prior to successfully performing on false-belief tasks, suggesting that first order theory of mind inferencing builds on structural linguistic knowledge (e.g. De Villiers & Pyers 2002). Hale & Tager-Flusberg (2003) support this in a training study where typical children trained on TOM task improved only on the TOM task, but children trained on a task of sentential complements improved both on the syntactic structure comprehension and on a TOM task. A recent study however fails to show such exclusive dependence on syntax, instead suggesting a dependence of false-belief understanding on general language ability: both syntactic and semantic proficiencies, but not working memory (Slade & Ruffman 2005). Perner et al (2003) also show that in German, where complement structures are required in sentences not involving mental states, complement structure is in place before TOM: children ages 2;6-4;6 could understand ‘want that’ complement clauses before ‘say that’, ‘think that’, and false belief.

Perhaps a behavioral genetic study of TOM in over a thousand 5-year-old twin pairs can help elucidate what is going on (Hughes, Jaffee, Happe, Taylor, Caspi, Moffitt 2005). It was found that individual differences in theory of mind correlated with verbal ability (as measured by vocabulary knowledge), and that knowledge of TOM was primarily predicted by environmental and not genetic factors. This finding is in contrast to behavioral genetic studies of syntactic acquisition that find strong genetic factors. Behavioral genetic studies with TD twins show that development of syntactic constructions, such as passives and verbal inflections, has substantial heritability (genetics) but little shared environment effects, with higher correlations between abilities of monozygotic twins than dizygotic twins (Ganger et al 2005, Ganger 1998). Furthermore, development of syntactic verbal inflections has no correlation with children’s verbal levels and maternal education (Rice, Wexler, Hershberger 1998). But vocabulary levels are known to be driven by environmental effects in twins (e.g. Bishop et al 2005). Therefore, if there is a genetic component in syntactic development (and not much environment), but not much genetics and plenty of environment in theory of mind development, the development of the two correlates because it happens in parallel, but not because one influences the other. The fact that children in Hale & Tager-Flusberg (2003) performed better on theory of mind after being trained on sentential complements is a direct illustration of environmental influences on theory of mind development, as is the fact that deaf children, who lack ‘normal’ levels of input due to having hearing parents, are delayed in theory of mind, but not deaf children who are exposed to sign language from birth (Schick, de Villiers, de Villiers, Hoffmeister 2007). Thus it seems that whereas computational aspects of language (syntax) are driven by genetics, cognitive strategies – such as vocabulary levels and theory of mind – are driven by environmental contributions – as evidenced by the fact that richness of input to the children enhances theory of mind.

As far as we know, there were no behavioral genetic studies of semantic and/or pragmatic knowledge in twins. Given that Maximality is part of the semantic computational language faculty, we can predict that it can pattern alongside syntactic development. If on the other hand,
children’s deficits with determiners are due to theory of mind, we expect to see the majority of variance in performance being accounted for by environmental factors. A twin behavioral study of determiners would be the perfect answer to these questions, but given the absence of such a study, we can investigate acquisition of determiners in parallel to theory of mind.

4.1.3. Determiners and Theory of Mind – independent of each other?!

Let us try to incorporate theory of mind with semantics. Let us assume that the listeners DO actively try to figure out what the speaker means (although it is usually assumed that the burden of effective communication is on the speaker). Determiners help listeners pick out referents, and may play key distinguishing roles in interpretations. The semantic definition of ‘the’ tells a person to pick the uniquely salient referent in the context set. ‘That’ on the other hand tells a person that the familiar uniquely salient referent is in the context set located in the information space shared by both the speaker and the listener. Observe that while the definition of ‘that’ explicitly refers to minds other than the listener, ‘the’ only needs a context set – if the context set is not clear, the listener may think about the hearer’s intentions, but that is optional.

Let us illustrate what we mean. Figure 4.1.1 shows the path of reasoning involved in 1st order vs 2nd order standard TOM tasks (reviewed in 2.2.3.). For the 1st order TOM, the Ignorant person refers to the one who leaves the room and does not witness the Hider moving the object of interest from one place to another. For the 2nd order TOM, the Informed person is one who believes that the Knowledgeable person is not aware of the object of interest, although the Knowledgeable person actually is. (e.g. Perner & Wimmer 1985)

![Figure 4.1.1. Schematic representation of reasoning involved in Theory of Mind](image)

![Figure 4.1.2. Schematic representation of hypothesized reasoning in interpretation of determiners](image)
Figure 4.1.2 illustrates in the same manner the reasoning involved in interpretation of the definite determiners ‘the’ and ‘that’. It is evident that while understanding ‘that’ involves the speaker and the listener sharing a mental space representing the salient unique referent and interfacing with the context, understanding of ‘the’ depends on the context set. In other words, children do not need to know anything about theory of mind to interpret ‘the’ in comprehension – only their knowledge of the salient object in the context set and knowledge of maximality, no false belief is involved, and no necessary consideration of what the speaker has in mind. But children do need awareness of others’ minds for interpretation of ‘that’, because they have to consider the saliency of the referent relative to the shared knowledge between the speaker and the listener. Thus if children’s performance on determiners depended on theory of mind, then we would expect, if anything, children to do worse on ‘that’, that requires pragmatics. Yet results in sections 3.2-3.5 above make a clear case that just the reverse is true.

4.1.4. Method

The nonverbal theory of mind test was used because it excludes any potential linguistic confounds (adapted from work by Call & Tomasello 1999; Colle, Baron-Cohen & Hill 2007). This involved two experimenters – a hider and a communicator. The hider hid a reward (a poker chip which the kids traded for stickers at the end of the game) under one of two identical Styrofoam cups/plates, the action visible to the communicator only. The communicator tried to help the participant to locate the reward (by placing a brightly-colored ring on the location of reward from the communicator’s point of view), but occasionally left the room. During the absence of the communicator, the hider would sometimes either visibly displace the reward from one location to another (in control conditions), or switch the containers without revealing the location of the reward (in control and test). Thus, the communicator occasionally ended up with a false belief and eventually indicated an incorrect location for the reward, which the kids were supposed to pick up on. The only verbal stimuli was the question “Where is the chip?” uttered by the hider to communicator to initiate indication of communicator’s point of view, and to the participant to initiate the participant’s response. The participant had to point to the final location of the chip.

Several conditions were devised (see appendix for complete descriptions). In the False Belief condition, the containers switched locations while communicator was away (thus Communicator had a false belief about the location of the reward). In the True Belief the communicator was present during the switching of the containers. In Control Belief, the communicator was away, but there was no container switch.

Typically developing children successfully passed the nonverbal task in a similar way as the verbal task (Call & Tomasello 1999): children ages 4;5-5 were successful in passing both. Children with autism (verbal/language levels at 2 year old level) failed only the false belief conditions (less than 20% correct), but passed the control conditions (80% correct), whereas typically developing 4 year old participants got false belief (almost 60% correct). (Colle, Baron-Cohen & Hill 2007).
Children were also tested on ‘Felt’ (as described in 3.3). The order of administration of TOM task and Felt task was randomized.

4.1.5. Participants

Participants included 51 children (3;1-11;4) tested on both the ‘felt’ version of determiner act-out task and the nonverbal theory of mind. The majority of these children came from prior samples ran on ‘Felt’. 8 children were excluded from the present analysis for failing control conditions on TOM task. Given the small number of subjects, for analysis data from three and four year olds was combined (17 children, mean age 4.2), data from five and six year olds was combined (13 children, mean age 5.9), data from seven-ten year olds was combined (13 children, mean age 9.0).

<table>
<thead>
<tr>
<th>Age (number of participants)</th>
<th>3(5)</th>
<th>4(12)</th>
<th>5(6)</th>
<th>6(7)</th>
<th>7(4)</th>
<th>8(1)</th>
<th>9(4)</th>
<th>10(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONT--visible</td>
<td>92.00%</td>
<td>66.67%</td>
<td>66.67%</td>
<td>70.67%</td>
<td>66.67%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CONT-invisible</td>
<td>82.00%</td>
<td>77.00%</td>
<td>89.67%</td>
<td>71.67%</td>
<td>51.33%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CONT-ignore</td>
<td>92.33%</td>
<td>82.00%</td>
<td>97.33%</td>
<td>74.33%</td>
<td>74.33%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1.5. Percent correct responses on control conditions on TOM task

4.1.6. Results

Firstly, following other studies of nonverbal theory of mind, we confirmed children’s ability to understand the nonverbal theory of mind method using Control and Belief control conditions (all groups perform well (66% or better, except 5-6s on control Belief condition)).

Looking next at performance on False Belief and on “Felt”, we find, intriguingly, that levels of false belief performance match those of determiner performance on group levels. 3-4 year olds perform around 30-40% correct on False belief and ‘the’ and ‘that’; 5-6 year olds are around 60% on false belief, 60% on ‘the’ and 77% on ‘that’; and 7-10 year olds are around 70-80% on all tasks. Note that at the same time, 3-4s were closer to 50-60% on anaphor ‘same’, and
5-6s were close to ceiling performance on these items – suggesting children know what is the salient referent in the context set and can pay attention very well to the tasks.

Figure and Table 4.1.6. Percent correct responses (TOM) and percent actions on same object (FELT)

<table>
<thead>
<tr>
<th>Age Group</th>
<th>BELIEF-false</th>
<th>THE</th>
<th>THAT</th>
<th>Share</th>
<th>Same</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-4s(17)</td>
<td>43.00%</td>
<td>32.33%</td>
<td>34.33%</td>
<td>58.75%</td>
<td>54.50%</td>
</tr>
<tr>
<td>5-6s(13)</td>
<td>61.67%</td>
<td>62.83%</td>
<td>77.00%</td>
<td>97.75%</td>
<td>95.50%</td>
</tr>
<tr>
<td>7+s(13)</td>
<td>71.67%</td>
<td>84.67%</td>
<td>88.50%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

However there is no correlation between false belief and determiner performance within each of these groups. Correlations across ages simply reveal an overall development, rather than causation. If we were to measure height of children, that would also correlate with theory of mind and determiner knowledge – simply because children are growing beings. If we investigate children’s performances for a given False Belief Score – there is only correlation between scoring zero on FB and scoring 26% on ‘the’ and ‘that’. With other scores on False Belief there is no correlation with determiner knowledge.

Table 4.1.7. Determiner performance for a given False Belief score

<table>
<thead>
<tr>
<th>FB</th>
<th>THE</th>
<th>THAT</th>
<th>AGE</th>
<th>#</th>
<th>Age Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>25%</td>
<td>27%</td>
<td>4.5</td>
<td>10</td>
<td>3.5-6.92</td>
</tr>
<tr>
<td>1</td>
<td>77%</td>
<td>73%</td>
<td>6.3</td>
<td>5</td>
<td>4.1-10.25</td>
</tr>
<tr>
<td>2</td>
<td>68%</td>
<td>80%</td>
<td>7.1</td>
<td>15</td>
<td>3.5-11.33</td>
</tr>
<tr>
<td>3</td>
<td>63%</td>
<td>69%</td>
<td>6.1</td>
<td>13</td>
<td>4.42-9.17</td>
</tr>
</tbody>
</table>

Next we characterized children’s performance as either ‘good’ or ‘bad’. ‘Good performance’ counts getting 2/3 or more correct on the TOM task, and getting 4/6 or more correct on the Felt task. This is similar analysis to what was done for ‘Felt’. In ‘Felt’ Participants were subdivided by their pattern of performance between ‘the’ and ‘that’, with “A” pattern denoting
good performance on both ‘the’ and ‘that’, “B” pattern denoting bad performance on ‘the’ with a relatively better performance on ‘that’, “C” pattern denoting a bad performance on ‘that’ with a relatively better performance on ‘the’, and finally “D” pattern denoting a bad performance in both ‘the’ and ‘that’. The trend in this data was similar to what was found in the larger sample tested on ‘Felt’.

In TOM, participants were subdivided by their pattern of performance between ‘the’ and ‘false belief’, with “A” pattern denoting good performance on both ‘the’ (4/6 or better) and FB (2/3 or better), “B” pattern denoting bad performance on FB with a relatively better performance on ‘the’, “C” pattern denoting a bad performance on ‘the’ with a relatively better performance on FB, and finally “D” pattern denoting a bad performance in both ‘the’ and FB.

On individual level, almost 50% of 3-4s were bad at both FB and ‘the’, 29% were doing better at FB than at ‘the’, with a minority doing well on both or doing better on ‘the’ than on FB. In 5-6 year olds, the majority of children (38%) showed better knowledge of FB than ‘the’, 31% showed good knowledge on both, and 23% showed better knowledge on ‘the’ than on FB. Comparison of False Belief relative to ‘that’ yielded similar results in this sample (data not shown).

Figure and Table 4.1.8. Proportions of children performing ‘well’ and otherwise on theory of mind relative to ‘the’.

<table>
<thead>
<tr>
<th></th>
<th>3-4s (17)</th>
<th>5-6s (13)</th>
<th>7-10s (13)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>0.47</td>
<td>0.08</td>
<td>0.00</td>
</tr>
<tr>
<td>C the&lt;tom</td>
<td>0.29</td>
<td>0.38</td>
<td>0.08</td>
</tr>
<tr>
<td>B the&gt;tom</td>
<td>0.06</td>
<td>0.23</td>
<td>0.15</td>
</tr>
<tr>
<td>A</td>
<td>0.18</td>
<td>0.31</td>
<td>0.77</td>
</tr>
</tbody>
</table>

4.1.7. Discussion

We investigated whether Theory of Mind may play a role in acquisition of determiners. We find that in 3-6 year old children, around 30-40% of participants know false belief better, but do not know ‘the’. In the 5-6 year olds, there were 23% of participants who knew ‘the’ better than false belief. These results suggest that development of false belief leads the development of ‘the’ and ‘that’, however false belief is not sufficient for the development of ‘the’ and ‘that’. Based on the study of ‘the’ vs ‘that’ in chapter 3, and this study of theory of mind, there is a two year difference between attainment of semantic knowledge vs other mind/pragmatic/theory of mind knowledge: ‘the’ develops around age 7, whereas first order theory of mind is in place by age 5.
If theory of mind, specifically the ability to estimate other people’s knowledge of the (discourse) context, is causal to the development of determiner ‘the’, why do children spend two years digesting their ability to know other people minds before finally applying uniqueness correctly?

Given this data, it is still possible to argue that theory of mind is necessary for determiner acquisition, but we clearly show that theory of mind is not sufficient for determiner acquisition, and it cannot be the sole cause of children’s deficits in ‘the’.

4.1.8. Acknowledgements

Many thanks are due to Charlotte Giessman for collaboration on this study. We would like to thank Wexler Lab undergraduate researchers for assistance in data collection. Our heartfelt thanks go to all the participants, their families, their teachers, and the child care centers that allowed us to carry out this study on their premises.

4.1.9. Appendix: Details of Procedure

<table>
<thead>
<tr>
<th>Task (adapted from Baron-Cohen et al. 2006 and Call &amp; Tomasello 1999)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Material</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Instruction</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Pretest</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Screening / Control test</strong></td>
</tr>
<tr>
<td><strong>Visible displacement</strong></td>
</tr>
<tr>
<td>Scenario</td>
</tr>
<tr>
<td>-------------------------------</td>
</tr>
<tr>
<td>Invisible displacement</td>
</tr>
<tr>
<td>Ignore communicator (hardest)</td>
</tr>
<tr>
<td>Belief task</td>
</tr>
<tr>
<td>False belief</td>
</tr>
<tr>
<td>True belief</td>
</tr>
<tr>
<td>Control belief</td>
</tr>
</tbody>
</table>
4.1.10. References


4.2. **Experiment 5. Comprehension of Maximality in Free Relative Clauses**

4.2.1. Abstract

The present section adds to the debate around ‘the’ by experimentally testing another linguistic phenomenon that involves maximality, namely Free Relatives (e.g. “[What is on the table] is red”). Unlike definite DPs, Free Relatives do not seem to exhibit salience effects, and thus are not salience-dependent, and therefore should not present comprehension difficulties for non-semantic reasons (Heim, p.c.). We found that children lack awareness of the maximal interpretation in free relatives at least until 6-8 years of age, the age at which it was shown (e.g. Karmiloff-Smith 1979, Modyanova & Wexler 2007, Chapter 3 herein) that children begin to reliably interpret the definite article in the adult way. These results are taken to support the Maximality hypothesis.

4.2.2. Background on Free Relatives

**Semantics of Free Relatives**

A Free Relative (1) is a headless (or nonlexically headed) relative clause found in many languages, taking singular agreement on the matrix verb, and replaceable by a truth-conditionally equivalent definite plural DP (2) (e.g. Caponigro 2002, 2004). Narrow scope indefinite free relatives (not found in English) will not be addressed herein. Two presuppositions play a role in free relatives. The maximality presupposition (that the clause must refer to the maximal entity) and the homogeneity presupposition (that the predicate must be true of every subpart of the maximal entity, i.e. the maximal set must be internally homogenous with respect to the predicate (Lobner 1985, Gajewski 2005)). These presuppositions in free relatives become evident in consideration of several contexts: if there are one or more red things on the table, the sentence is true (a). If there are some blue things in addition to some red things (b), the sentence is neither true nor false, but constitutes a presupposition failure. The failure is of the presupposition of homogeneity - the maximal entity of things has some red and some blue elements, and thus does not constitute a homogenous set. As a result, the Maximality presupposition fails because the only way to prevent a failure of homogeneity is to select a nonmaximal subset. I.e. in a context like (b) the predicate ‘red’ is expected to hold true of the entire referent, denoted by the free relative, not a part of it (homogeneity presupposition), but then the clause cannot refer to the maximal entity (Maximality presupposition). The sentence fails in context (b). Similar judgments are observed in the case of plural DPs. In absence of a referent in context (c), DPs and free relatives seem to differ: an empty set seems unacceptable for DPs, but is sometimes possible for free relatives.

(1) [FreeRel What is on the table] is red
(2) [DP The things on the table] are red

(3) contexts:

a. Any number of red things
b. *?Some red things, some blue things
c. ?No things {null set is a plural entity}
For the purposes of the current investigation, we are going to make the assumption that children know (have) the property of homogeneity (though this remains to be experimentally investigated). If adults are forced to make a judgment on a free relative referring to a subset, they cannot preserve homogeneity without violating Maximality – they would have to select the homogenous nonmaximal set, or a nonhomogenous maximal set – neither of which is an acceptable alternative. The sentence fails for adults. Children however, if they are missing Maximality in free relatives (but they have homogeneity), are happy to satisfy homogeneity by selecting the nonmaximal subset, something that adults cannot do.

A notable point (Heim, p.c.) is that there seems to be a subtle but important difference between plural DPs and free relatives. Free relative clauses, unlike plural DPs, do not seem to show salience effects – they do not require the referent to be salient in the context set. Let us consider some examples that illustrate salience effects (all due to Heim, p.c., originating from McCawley). For definite description in (4), it is clear that ‘the dog’ is same dog as ‘my dog’ – it is the salient dog in the context set. The same holds true of plural definites (5). In free relative clauses (6), however, this notion of saliency fails to apply and the result is bizarre. There seems to be no way to separate out the set of things under the table that were just put there (the salient set) from the set of things under the table that were already there. Thus, the only possible interpretation of (6) is that the set of things under the table ends up sitting on itself, and even that does not make sense. This is in stark contrast to the salient dog(s) in (4) and (5) and the distinct nonsalient another dog. Free relatives do not do saliency.

(4) I took my cat and my dog to the park. There, the dog got into a fight with another
(5) I took my cat and my two dogs to the park. There, the dogs got into a fight with another dog.
(6) I put some of my stuff under the bed and some under the table. What was under the table sat on top of some other stuff already piled up there.

Thus in (1) the free relative clause seems to refer to “the set of everything in the universe that is on the table”. In (2) the plural DP seems to refer to “the set of everything that is on the table and that is salient”. In other words the semantics of free relatives can be argued to be not saliency dependent, in the same way that the semantics of ‘the’ is saliency dependent. If this is true (and it certainly merits further study), then investigating comprehension of maximality in free relative clauses is a ‘pure’ test of semantics – the saliency of the context plays no role, hence no pragmatic deficits should influence comprehension of free relatives – only semantic deficits may do so.

Sources of Maximality in Free Relatives

An account of the semantics of free relatives must explain the source of the maximality presupposition. As illustrated above, free relatives may be substituted by truth-conditionally equivalent definite DPs (as long as salience effects are not invoked). And herein lies a potential problem. DPs are usually expressions of type e (denoting individuals, like ‘Bill’), but depending on the context may shift into type <e,t> (denoting characteristic functions of a set of individuals), or type <<e,t>,t> (the characteristic function of a set of properties, e.g. ’every X’) (Partee & Rooth 1983). Wh-constituents, such as relative clauses, denote sets of individuals (Cooper 1983).
How can one make a relative clause ‘shift down’ to denote a single individual in case of the free relative? How to solve this type mismatch?

Two recent proposals (Jacobson 1995; Caponigro 2002, 2004) find the answer in type-shifting rules of Partee (1987). Notably, the Iota type shifting rule is a way for a property (type <e,t>) to map into its maximal individual (one that contains all other individuals of the set, and is of type e), and is only defined for properties that characterize one and only one individual. Thus a predicative expression, e.g. a wh-clause, may shift into an individual-denoting expression, a DP, but only when the set that the predicate characterizes is a single entity.

Where the two accounts differ is in the source of maximality. Jacobson proposes maximality to be encoded in the wh-words, implying that wh-clauses always display maximality, regardless of context. Caponigro argues against that, since depending on the context, not all wh-questions have to absolutely be answered in a completely exhaustive manner: there is room for the involvement of pragmatic, contextual restrictions on interpretation of wh-clauses. Caponigro, in unifying the account over both definite and indefinite free relatives (not found in English), claims that maximality results directly from the type mismatch itself, between the set denoted by the free relative (type <e,t>) and the fact that the rest of the sentence requires an individual (type e). Thus, unlike lexically headed relative clauses, which describe sets of objects, the free relative denotes a maximal plural entity, which explains the maximality presupposition.

Prior Studies of Acquisition of Free Relatives

Although children’s interpretation of free relatives was never previously investigated, their syntactic acquisition was. Flynn & Lust (1981) tested 96 children (aged 3;6-7;7) on elicited imitation and act-out comprehension tasks of several kinds of object relative clauses: lexically headed (with determinate (7) or non-determinate (8) heads) and nonlexically headed (free relatives(9)).

(7) Big Bird pushes the balloon which bumps Ernie
(8) Ernie pushes the thing which touches Big Bird
(9) Cookie Monster hits what pushes Big Bird

Free relatives were the easiest for children to imitate, and children spontaneously turned other forms of relative clauses into free relatives. The Determinate Head (7) was the best comprehended. In all conditions, children showed 75% correct performance only by age 6;6; younger children performed at or below 50%. (Flynn & Lust 1981). Since then, a similar pattern of acquisition, with free relatives emerging before or alongside headed relative clauses, was observed in first and second language acquisition crosslinguistically (Flynn & Foley 2004). It is argued that there is an easier semantic-syntactic mapping for free relatives than for lexically headed clauses.

This study illustrates that free relatives clauses can be understood by children (from a syntactic point of view), therefore it is possible to investigate children’s comprehension of semantics in free relative clauses, without worrying too much about any syntactic confounds.

4.2.3. Method

Maximality in free relatives can be studied by presenting participants with contexts that do not contain a maximal plural entity uniquely described by the predicate and result in a truth-valueless utterance and a presupposition failure. Investigating presupposition failures is known to
be a challenge, thus we wanted to set up a relatively simple but informative behavioral task. We investigated the interpretation of free relatives with a match/mismatch (yes/no) task presented on a computer, and, in addition to participants’ responses, their reaction times were also recorded. Neither a ‘yes’ nor a ‘no’ response is an entirely correct answer in case of a presupposition failure, but we wanted to keep the experiment simple and therefore we did not provide a third answer option (‘neither true nor false’). Participants’ reaction times are expected to be informative as to any uncertainties in their answers – if participants experience a presupposition failure, it is reasonable to expect that will take them longer to process. If participants do not experience a presupposition failure, they will answer faster.

Participants were presented with a single picture at a time on the computer screen, and were asked a yes/no question. Free relatives were formulated in form of a question because in pilot testing with declarative free relatives children pointed to the referent(s), instead of evaluating the truth-value of the utterance and replying with ‘yes’ or ‘no’. Figure 4.2.1 shows one of the illustrations. The elephant is hiding some apples under a blanket, and some other apples are lying around the elephant. All pictures involved colors and animals, and children were tested on their color words and animal names in the introduction to the task.

![Figure 4.2.1. Example of stimulus picture](image)

There were three conditions, each with eight items. In the “yes” control condition, a picture denoted a complete set of items, and was correctly described by the test sentence. Thus if all the apples under the blanket were red, the sentence here would be: “Is what is under the blanket red?”. In the “no” control condition, the test sentence failed to question the pictured set: e.g. all the apples under the blanket were green, but the test sentence would still be “Is what is under the blanket red?”. Here we made sure there was still a potential referent in the picture, e.g. a red apple, but not under the blanket, to ensure that children had an opportunity to say ‘yes’. The key test condition was ‘max’ (‘maximal’), which involved a set consisting of two kinds of objects, e.g. some green and some red apples under the blanket, but the test sentence only asked about a part of the set, “Is what is under the blanket red?”, which constituted a violation of the maximality presupposition. The right answer is ‘neither true nor false’ (which we do not provide), and we expect those who know maximality to say ‘no’ (however it is also entirely possible for them to say ‘yes’), and to take longer to make their judgment relative to the control “no” and “yes” conditions.
Potential Phases in Free Relatives

When free relatives appear in copular sentences like above, there may be a concern that any children’s deficits in comprehension of free relatives may be not due to problems with maximality or domain restriction, but due to children’s universal phase requirement (UPR), which causes children to lack the ability to define defective phases and be unable to interpret passives, unaccusatives and raised constructions (Wexler 2004). Because of the structure of the investigated free relatives, this confound is unlikely. Copular sentences are distinguished in two ways (e.g. Higgins 1973). In predicational sentences, the referential DP comes before the copula with the predicate following it (10). In specificational sentences, the referential DP is post-verbal, and the predicative constituent is in the first position (11). The latter, but not the former, requires raising of the predicational constituent to specTP, necessitating crossing a phasal boundary. As predicted by UPR, children do have deficits with comprehension of inverse copulas, performing at 75% correct only by age 7, whereas non-inverted copulas are understood at near 100% rates as early as 3 years of age. (Hirsch & Wexler, 2007). Free relatives in the current study are predicational (12), not involving raising, and hence should not pose problems for subjects for syntactic reasons.

(10) The pig is the animal who hugs [non-inverted copula, predicational]
(11) The animal who hugs is the pig [inverse copula, specificational]
(12) [RefDP What is under the blanket] is [AdjPred red]

4.2.4. Predictions

As stated earlier, for the purposes of the current investigation, we are going to make the assumption that children know (have) the property of homogeneity (though this remains to be experimentally investigated). If adults are forced to make a judgment on a free relative referring to a subset, they cannot preserve homogeneity without violating Maximality – they would have to select the homogenous nonmaximal set, or a nonhomogenous maximal set – neither of which is an acceptable alternative. The sentence fails for adults. Children however, if they are missing Maximality in free relatives (but they have homogeneity), are happy to satisfy homogeneity by selecting the nonmaximal subset, something that adults cannot do.

If children do not know the maximality presupposition, we would see variable verbal answers to conditions violating maximality, and perhaps the reaction times would be the same across the test and the control conditions. If subjects do know Maximality (e.g. adult participants), we would see longer reaction times relative to control conditions, and these longer-than-expected reaction times would illustrate subjects’ awareness of the Maximality presupposition, the failure of which they end up experiencing.

The key prediction is if Maximality is involved in interpretations of free relatives, and if Maximality is involved in definite determiners (which we know is the case), then we should see similar patterns in acquisition of free relatives and the definite determiner if children are lacking Maximality. Thus we expect children younger than 6-7 years to not show us knowledge of Maximality within the Free Relative task. If on the other hand, Maximality is not what is causing children’s definites in ‘the’, then we expect children to perform similarly to adults very early on, on Free Relatives.

Predictions of the pragmatic / Theory of Mind / Egocentric view are harder to formulate because, as argued in other chapters, it does not readily apply to comprehension without additional assumptions, and because children would need to be hypothesized to be unpredictably
selecting their own salient context sets which do not match the actual salient context sets, which to us does not seem plausible. It is feasible to imagine that children could focus on the subset described by the free relative, as opposed to the entire set, but that should clash with their adult-like knowledge of maximality. Thus it would seem on this view that children may take a bit longer with the maximal condition early on, since they are expected to have adult-like competence. Furthermore, if children are focusing on a wrong subset, we would expect a great deal of ‘no’ responses.

Furthermore, it is worth it here to repeat a point made by Heim (p.c.) that free relative clauses do not involve saliency at all. Hence the potential reasoning that we hypothesize could be taken by the pragmatic/egocentric view in the above paragraph cannot hold. “Indeed, if the semantics of ‘what's on the table’ is simply ‘the set of everything in the universe that is on the table’, then … egocentrism cannot possibly be an explanation for non-adult-like comprehension. If salience plays no role in determining reference in the first place, the child's comprehension can't depend on what's salient to them.” (Heim, p.c).

4.2.5. Participants

Participants were recruited from Boston, Cambridge, and Wellesley, MA daycare and after school programs, and from higher education institutions. There were sixteen younger children (3;0-5;11, mean age 4;0), thirteen children aged 6;0-8;11 years (mean age 7;0) termed ‘middle’ herein, nine older children (9;0-12;4, mean age 10;2), and twenty two college students (18-25 years old). An older group of eleven adults (aged 43-56) was also tested, but their results were not qualitatively different from the college students (older adults take much longer to react) and are not included.

4.2.6. Results

We find (Table 4.2.2) that all children performed well on control conditions involving free relatives but containing no maximality violations. Everybody correctly said ‘yes’ to ‘yes’ conditions where the sentence matched the picture, and everybody correctly rejected ‘no’ conditions where there was a clear mismatch between the picture and the sentence. Most children said ‘yes’ (i.e. seemed to agree with) violations of maximality in the ‘max’ conditions. As expected, the adult group also showed similar acceptance of maximality violations as children (remember, the right answer there is neither ‘yes’ nor ‘no). Although the rates of performance on ‘max’ condition do not differ across groups, individual subject analysis shows that all children younger than 6 years accepted Maximal violations, but four of the nine children in the old group did primarily reject them.

<table>
<thead>
<tr>
<th></th>
<th>Young (3-5)</th>
<th>Middle (6-8)</th>
<th>Old (9-12)</th>
<th>College (18-24)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAX</td>
<td>17%</td>
<td>22%</td>
<td>33%</td>
<td>30%</td>
</tr>
<tr>
<td>NO</td>
<td>83%</td>
<td>95%</td>
<td>97%</td>
<td>98%</td>
</tr>
<tr>
<td>YES</td>
<td>95%</td>
<td>96%</td>
<td>96%</td>
<td>98%</td>
</tr>
</tbody>
</table>

Table 4.2.2. % ‘Correct’ Responses

The reaction time data elucidates what is going on (Figure 4.2.3). The younger children take an equal amount of time to reply to all conditions, indicating their lack of awareness of maximality. The older children and adults take longer to reply to the maximal condition, than to
the control conditions, indicating their awareness that something unusual is going on in sentences violating the maximality presupposition. This awareness begins to emerge as early as 6-8 years of age, with the middle group taking as much time to say ‘no’ correctly to the ‘no’ condition, as they did to say ‘yes’ to the ‘max’ condition. When children are 9 years old, their performance RTs mirror those of adults.

Younger children’s reaction times were not significantly different across conditions. The middle group’s (6-8 year olds) RTs for ‘yes’ vs ‘max’ conditions were approaching significance (t(12)=2.07, p=.061), and the older groups’ (9-12 year olds) differences for ‘yes’ vs ‘max’ conditions were significant (t(8)=2.5, p=.038). College students’ (18-24 year olds) RTs significantly differentiated between ‘no’ and ‘max’ (t(21)=3.1, p=.005) and ‘yes’ and ‘max’ conditions (t(21)=3.58, p=.002). Crucially, those few participants in the 9-12 year olds and the college groups, who rejected violations of maximality, took as much time as those who failed to reject the violations (with no significant differences). A difference plot graph (Figure 4.2.4) shows the same trends.
4.2.7. Discussion

The results of the first-ever study of children’s maximal interpretation in free relatives appear consistent with predictions made by the maximality hypothesis (Wexler 2003): very young children would show variable acceptance rates of conditions involving violations of the maximality presupposition, but they do not show any slower reaction times relative to control conditions. The age at which awareness of Maximality appears in Free relatives (6-8 years old) matches the age at which awareness of Maximality appears in the definite determiner ‘the’.

The younger children, ages 3-5, took an equal amount of time to reply to all conditions, indicating that they had no particular doubts as to any of their responses across conditions. It is likely that younger children do not know the maximality presupposition, and thus do not think twice about their answers to the ‘max’ condition, but merely pick the subset described by the picture. The middle group, ages 6-8, began to be aware of the failure of the maximality presupposition, and took as much time to correctly reject the ‘no’ condition, as to accept the ‘max’ conditions. The older children and the adults took much longer to react to the maximal condition, than on both the control conditions. The majority of these participants responded affirmatively, but a minority did reject the ‘max’ conditions. All, however, took about the same time to provide either response. That "no" answers take longer than "yes" answers is an extremely reliable and well-known result in psychological experimentation. Thus the fact that adults took longer giving "yes" answers for presupposition failures provides convincing evidence that they are not sure what to say, and that an answer in the ‘max’ condition is problematic for them.

Our results can be explained if the older children and the adults do know maximality, but do not know how to react to a presupposition failure, given a constrained yes or no answer choice, and are likely to accept it, while taking a while to think about their answer.

The results do not appear to be consistent with the pragmatic/egocentric hypotheses. Firstly, these would predict that children know Maximality and hence should show longer (adult-like) reaction times to the maximality presupposition failure than we found, and would predict higher rates of rejection responses than we found. And Secondly, it is possible to argue that free relatives do not involve salience, and hence there is no way for their interpretation to be deficient for pragmatic/egocentric reasons.

The present findings are also corroborated by Yatsushiro (2007), who investigated children’s knowledge of the anti-uniqueness presupposition of ‘every’. If children are missing or just beginning to develop uniqueness, they will also be missing antiuniqueness. ‘Every’ was argued to have a number of presuppositions: the existence (the argument of ‘every’ cannot be an empty set), the anti-duality (the argument of ‘every’ cannot be a two-membered set), and the anti-uniqueness (the argument of ‘every’ cannot be a singleton set) (Sauerland 2007). Yatsushiro tested 120 monolingual German children ages 6-9 years using a presupposition judgment task, where the subject’s job was to say whether an observer of the situation depicted in a picture would say a particular sentence. For example, in one illustration there was a woman sitting, and a man and three children standing. Somebody with full knowledge of presuppositions of ‘every’ would reject sentences like (13) (violating anti-uniqueness) and like (14) (violating truth, and violating anti-duality). All participants showed near 100% knowledge of the existence presupposition, but the anti-uniqueness for ‘every’ was only achieved by adults (90%). The 6-9 year old children progressed from 35% to 63% correct.

(13) Every mother of mine is sitting on a chair
(14) Every grandfather of mine is also sitting on a chair
Testing presupposition failure is known to be a challenge, and as our percent-correct response results show, even adult participants are likely to accept presupposition failures. However measuring participants’ reaction times clearly indicates the levels of (un)certainty of their replies, and can be successfully used to estimate the participants’ competence. Needless to say, further experiments are necessary, ones that perhaps provide participants with a third potential response, in addition to ‘yes’ and ‘no’.

Our further studies will include testing comprehension of free relatives in parallel to plural DPs, and perhaps using a picture-matching and/or a presupposition judgment task (as in Yatsushiro) in addition to the current match/mismatch reaction-time task. We would also like to investigate the homogeneity presupposition, as well as the potential differences in interpretations of free relatives and plural DPs.

That maximality (in determiners and free relatives) and anti-uniqueness (in ‘every’) problems seem to pattern in analogous ways adds further proof to the hypothesis that children’s semantic competence, like their syntactic knowledge, takes time to develop. Studies also show that children’s pragmatics takes time to mature too. Perhaps we should aim at investigating semantics independently of pragmatics, which will be difficult for they work hand in hand, but is certainly possible given the present results.

4.2.8. Acknowledgements

We would like to thank Tania Ionin, Kai von Fintel, Vina Tsakali, Alexandra Perovic, Suzanne Flynn, and especially Irene Heim for discussions and suggestions, the audience of BUCLD 32 for comments and suggestions, and MIT and Wellesley undergraduates for assistance in testing and illustrating, especially Margaret Echelbarger, Ritu Tandon, Christopher Watson, and Alexandra Huston-Carico. Our heartfelt thanks go to all the participants, their families, their teachers, and the child care centers that allowed us to carry out this study on their premises. All remaining errors are ours.
4.2.9. References


4.3. **Discussion of Further Investigations into the Typical Development of Semantics and Pragmatics**

In this chapter, further investigation of semantic and pragmatic abilities in typically developing children provides support to the semantic deficit theory in determiner acquisition.

We presented data on comprehension of determiners in children in parallel to their comprehension of false belief conditions. According to our interpretation of pragmatic deficit in determiners theories, children’s deficits with ‘the’ are due to inability to consider the difference between speaker/hearer knowledge, and hence false belief should track closely with development of ‘the’, and no children should show knowledge of ‘the’ ahead of knowledge of false belief. On semantic deficit account, there should be no such correlation.

It was found here that while on group level, the knowledge of false belief tracks with the knowledge of determiners, on individual level there was wide variation, with almost 40% of children in the 5-6 year range showing understanding of other’s minds and of False Belief, but showing no understanding of ‘the’, and with 30% showing good knowledge on both, and with 23% showing better knowledge on ‘the’ than on false belief. The same 5-6 year olds showed close to perfect comprehension of anaphoric ‘same’, suggesting that they can tell what is the salient referent in the context set. Development of first order theory of mind certainly leads development of definite determiners, but children lacking theory of mind can and do show knowledge of determiners. Furthermore, if theory of mind, and hence the ability to clearly estimate other’s knowledge of context is known by 5 years, why is it that majority of children only reach good performance levels on ‘the’ only when they are seven year old? The answer is clear – children’s deficits in determiners stem from something other than theory of mind deficits — they stem from underdeveloped system of semantics.

To further support the semantic deficits, we investigated comprehension of maximality presupposition in free relative clauses using a match/mismatch version of a truth value judgement task. The maximality presupposition is argued to be part of free relative clauses, such that the proposition denoted by the free relative clause must refer to a whole context set (with plural or singleton members), as specified by the predicate. The homogeneity presupposition (that the predicate must be true of every subpart of the maximal entity) is also part of free relative clauses, however for the purposes of the present study we make the assumption that this is known to children (although this remains to be investigated). Note that it is possible to argue (Heim, p.c.) that unlike definite descriptions, free relatives can be interpreted without salience, and hence constitute a ‘pure’ test of semantic competence of maximality. If deficits in maximality underly children’s deficits in ‘the’, children should show similar deficits in comprehension of maximality in free relative clauses. If however children’s knowledge of maximality is intact and their deficits in ‘the’ occur for non-semantic reasons (egocentric/pragmatic), children should show adult-like comprehension of maximality in free relative clauses, unless certain extra assumptions are made by the egocentric/pragmatic view.

We find that children can evaluate whether or not free relative clauses match context descriptions that contain whole context sets. Evaluation of partial set contexts that violate maximality presented a difficulty for children, or rather a lack of difficulty – violation of maximality presupposition caused adults and older children through puzzlement and extensive
pauses in reaction times relative to conditions without maximality violations. Younger children, at least through 6 years of age, showed no such puzzlement, indicating that they were happy to apply the free relative description to the part of the set described by the predicate. This performance in young children violates maximality, but children, being deficient in maximality, do not notice this. 6-8 year old children began to show puzzlement, and only 9 year olds showed adult-like patterns – indicating that they are aware of the resulting failure of the maximality presupposition. The advantage of investigating free relative clauses is that their interpretation does not require reliance on anybody’s else’s state of knowledge, not does it require selecting a salient context set and salient referent – participants are simply evaluating whether a context matches the set denoted by the free relative sentence. Deficits in comprehension of maximality in free relative clauses can only stem from semantic deficits in maximality.

In future, we will investigate comprehension of definite descriptions in parallel with free relatives (items that involve maximality in their interpretation) and in parallel with items involving scalar implicatures, to investigate further the nature of semantic and pragmatic deficits in children.
5. Comprehension of Determiners in Autism Spectrum Disorders

In previous chapters we provided extensive evidence that delay in acquisition of the definite determiner is likely semantic in nature, and is not due to children’s pragmatic or theory of mind deficits. This chapter investigates the acquisition of determiners in people with autism spectrum disorders, a range of disorders with the defining characteristic of missing or abnormal pragmatic, social, and theory of mind abilities, and as such presenting the potential for ‘minimal pair’ comparison to typically developing children and the theories of acquisition of ‘the’. Given the prevalence of these disorders (e.g. Yeargin-Allsopp et al 2003), understanding the range of their language dis/abilities is important not only for scientific reasons but also for potential diagnostic and therapeutic reasons.

5.1. Abstract

We investigate the comprehension of definite (‘the’ and ‘that’) and indefinite (‘a’ and ‘another’) articles in Autism Spectrum Disorders (ASD). Interpretation of these elements, denoting salient and unique or simply existing referents, relies on the knowledge of both pragmatic and semantic aspects of language. While pragmatics is known to be impaired, little is known about the mastery of semantics, i.e. compositional meanings of sentences, in ASD. Thirty one children and adolescents clinically diagnosed with ASD (Chronological Age: 5;11-18;6) and 93 controls (CA: 3;0-17;1), individually matched on non-verbal reasoning (KBIT), vocabulary (PPVT), or grammar (TROG), were tested. Those participants with Asperger syndrome and PDD-NOS, who had the highest grammar levels, showed adult knowledge levels of articles. Another subset of participants with Asperger syndrome and PDD-NOS performed at levels below their mental age, but showed patterns found in typically developing younger controls. All participants with the diagnosis of autism showed an extremely poor knowledge of articles. Across all participants knowledge of articles was most well predicted by scores on the test of reception of grammar, as opposed to scores on the test of nonverbal reasoning or the test of vocabulary. These results indicate a gradation of knowledge of articles across the spectrum, which may have useful implications in the diagnostic process as well as in the elucidation of genetic endophenotypes of ASD.

5.2. Introduction to ASD

The precise nature of languages impairments in autism spectrum disorders is still not a hugely investigated topic, and few of the findings from the rich literature on the typical language acquisition have made it over. This is a first study of investigating semantic and pragmatic knowledge of articles/determiners in an unselected group of children with autism spectrum disorders, with the ultimate goal of increasing the data about the endophenotypes in the heterogeneous disorder of autism.

Deficits in language and communication are known to be one of the defining characteristics and diagnostic criteria of autism (APA, 1994). Language abilities in ASD are in process of being characterized, but already it looks like different subgroups show distinct patterns of language development ranging from nonexistent to delayed to normal. Autism spectrum disorders are associated with communicative difficulties (Downs & Smith, 2004), especially
issues with pragmatics – the use of language in conversational situations, given a particular context, and a particular state of knowledge of listeners and speakers. Some subtypes of autism have also been associated with deficiencies in the computational system of language – sentence structure and function words such as auxiliaries and articles. Rapin and Dunn (2003), who looked at an unselected sample of children with autism, suggest existence of at least two subgroups in ASD, those with language problems in phonology and syntax, and those with language problems in semantics and pragmatics, with no overlap.

A large number of studies has focused on individuals with autism who are high-functioning, i.e. whose non-verbal IQ is within normal range, and who can pass the first-order (involving reasoning about a person’s thoughts) and second-order (involving reasoning about what one person thinks about another person’s thoughts) false belief tasks (Bauminger & Kasari, 1999; Kaland et al 2002). However they still have difficulty recognizing faux pas (Baron-Cohen et al, 1999), and they have to reason out other’s behavior in social situations - something that non-autistic people understand instinctively (Luckett et al 2002).

However, difficulties with grammatical morphemes experienced by some children with autism in Roberts at al (2004) do not seem to be associated with IQ scores recorded for the participants in this study. Despite their normal range IQ, the performance of some of the participants on measures of grammar was still poor. Conversely, some of the children with lower IQ showed unimpaired performance on the same measures. In the same vein, asynchronies in the development of computational aspects of language as opposed to general cognitive abilities (the latter reflected in the level of receptive vocabulary) have also been noted in Kjelgaard & Tager-Flusberg (2001). These authors report that at least two thirds of their subject population of 89 children with autism, aged 6-16, showed poorer scores on standardized measures of grammar than on measures of vocabulary.

In the following pages, I summarize what is already known about semantic and pragmatic knowledge in people with autism spectrum disorders (ASD) and characterize what is known about theory of mind in these disorders. While some things are known about language in ASD, there are not many studies of precise aspects of semantics, and a dearth of studies of articles. Given heterogeneous language and cognitive abilities within ASD, it is important to study abilities of subgroups in detail, otherwise precise delineations will be impossible, preventing further genetic advances.

5.3. Semantics and Pragmatics in ASD

Although definite and indefinite articles were not studied in detail in ASD, other semantic and pragmatics aspects of language were investigated, and this literature can be used to formulate hypotheses regarding knowledge of articles in ASD.

Note that by semantics here we do not mean vocabulary levels, nor do we mean ability to classify words by concepts (colors, etc). There is a profound, but often missed difference, between knowing the meaning of e.g. ‘dog’, and of knowing the meaning of ‘not’, ‘the’, ‘every’. The latter are both more conceptually abstract and more linguistic – these are functional closed class items which typical children often take time to acquire. These referential terms are used as pointers to referents in conversations, and their successful use requires not only pragmatic awareness, but also knowledge of their semantic meanings which are part of the computational system of language, and are distinct form conceptual meanings of open class words.

The definition of pragmatics also differs between studies of autism and studies of typical development. In case of autism studies – pragmatics is about speech acts, i.e. using language to
request and to question, and about discourse cohesion – sticking to the topic of a conversation, and about understanding non-literal language use in ironies and metaphors. In typical development, pragmatic studies address, among other things, more subtle aspects of children’s knowledge, e.g. children’s comprehension of scalar implicatures (logical entailments) – limiting or extending semantic meanings of words such as ‘some’ and ‘every’ based on the contexts.

Tager-Flusberg (1981) reviewed studies to that date and concluded that semantic and pragmatic knowledge is especially deficient in autism. Evidence for pragmatic deficits was established – several studies found the now classical deficits in speech acts, speaker-hearer relationships, turn-taking, differentiating old and new information. Semantic deficits – there defined as knowing conceptual relations between words, e.g. that needle and thread go together – were shown to be impaired. Notably, Tager-Flusberg argues that “The evidence for semantic abnormalities in autism is only indirect, but the main weakness of the hypothesis is the absence of good research” (1981:50).

Groen, Zwiers, van der Gaag, and Buitelaar (2008) are the most recent to review a wide range of studies of language in ASD. Those studies investigating semantics focused on vocabulary – either receptive or expressive – which shows high correlation of nonverbal reasoning IQ levels. Subtle semantic problems however were shown in conceptualization of abstract terms and emotional states. Groen et al conclude that “the exact nature of the semantic deficits in autism remains to be established.” (2008:1418).

5.3.1. Semantics in ASD

Specifically, several studies investigated spontaneous speech in children with autism for presence or absence of function words, such as articles, and for overall discourse cohesion and reference which is often mediated through articles.

Simmons and Baltaxe (1975) in spontaneous speech of 4 of 7 teenagers with autism of average IQ found that semantic violations e.g. incongruency in word use, problems with pronouns, and inconsistent referential abilities in linking referents across sentences.

Bartolucci, Pierce & Streiner (1980) studied spontaneous speech in ten year old children with autism (mental age of 6 years) who showed 15% omission of articles in obligatory contexts.

Baltaxe and D’Angiola (1992) investigated discourse cohesion, especially reference (e.g. pronominal bridging across sentences (e.g. “John is out. He’ll be back at five”), as well as demonstrative and comparative reference), ellipsis (leaving out words or clauses which contain old information), and conjunction (using ‘and’, ‘because’). They find that compared to typical children (aged 3.5 years), children with specific language impairment (SLI) (aged 7-8 year and matched on language levels) were half as successful in discourse cohesion, and that children with autism (aged 7-8 years also matched on language levels to typical kids) were only 30% as successful as typical kids. Autistic participants were especially impaired in reference and ellipsis – using these less than 60% correct – compared to 85% correct in SLI kids and above 95% in TD kids. Note that typically developing children as early as 4 years are capable of referential bridging (Avrutin and Coopmans 2000).

Baltaxe and D’Angiola (1996) find that compared to typical children (aged 3.5 years) and children with specific language impairment (aged 7-8 years) (matched on language levels), children with autism (aged 7-8 years) were less successful in using pronouns and demonstratives to tie together discourse. Children with autism produced these at only half the rate of SLI children, and only third of rate of TD children. Additionally children with autism made four times as many mistakes as TD children, and twice as many mistakes as SLI.
More recently, Colle, Baron-Cohen, Wheelwright and van der Lely (2008) investigated story-telling discourse in adults with high-functioning autism (HFA) or with Asperger syndrome. Referential (anaphoric) use of pronouns and definite and indefinite articles with noun phrases to successfully introduce, reintroduce and maintain referents in a story by ASD participants did not drastically differ from a control group. However subtle deficits showed that ASD participants used less pronouns to maintain reference, and preferred to use full noun phrases. Arnold, Bennetto, Diehl (2009) found results similar to Colle et al (2008) but in children and adolescents with high-functioning autism ages 9-17.

Dahlgren and Sandberg (2008) found correlation between referential ability and theory of mind in 7-14 year old children with autism of normal intelligence. The referential ability task used involved producing verbal descriptions of a single card among a range of cards so that another person could successfully pick it out. This ability correlated with first-order theory of mind.

In summary, people with autism spectrum disorders who are higher functioning seem to show less deficits than those who are low functioning. Precise studies of comprehension of articles were not performed.

5.3.2. Pragmatics in ASD

Pragmatic impairments in autism (in Rice, Warren & Betz 2005) include a narrow range of speech acts (Loveland, Landry, Hughes, Hall & McEvoy, 1988; Wetherby, 1986); deficits in narratives and in conversations (Loveland & Tunali, 1993; Tager-Flusberg & Anderson 1991; Tager-Flusberg & Sullivan 1995); and a lack of consideration of the listener’s perspective (Paul & Cohen, 1984). More studies support earlier findings.

Sabbagh (1999) discusses that conversation difficulties in people with autism and those with right cerebral hemisphere damage stem from inability to understand the importance of communicative intentions of their conversation partner. This is supported by a recent study of narratives in high functioning autism: children’s narratives were unusual and less coherent than those of control subjects, suggesting deficit in awareness of listener’s state of knowledge (Diehl, Bennetto & Young 2006). Individuals with Asperger syndrome similarly show problems during conversations involving emotional and social topics (Adams et al 2002). Jolliffe & Baron-Cohen (1999) investigated comprehension of ‘strange stories’ by adults with HFA and Asperger syndrome and found that participants could provide answers to mental state questions but had difficulty in providing contextually appropriate mental state answers. For example in a “Sarcasm story”, a lady says that it was a lovely day for a picnic when it was raining. A person with HFA commented that she was “pretending that everything was OK in order to make Tom feel happier” – a clearly incorrect interpretation given the context.

On the other hand, Kremer-Sadlik (2004) shows that children with high functioning autism and Asperger syndrome are able to identify listener’s conversational needs. In conversational exchanges at subjects’ homes, subjects produced relevant responses accepted by their conversation partners. It may be the case that conversation partners of these subjects may themselves have abnormal pragmatic needs given the existence of broad autism phenotype in first degree relatives of people with ASD (e.g. Bishop et al 2004).

Lopez & Leekam (2003) investigated context integration – a key ability for successful discourse – in 15 children and adolescents with HFA, who showed good performance in both visual and verbal context enhanced identification of objects and homograph disambiguation.
Specific pragmatic deficits of the kind found in typical development were only studied recently.

Dennis, Lazenby, Lockyer (2001) investigated high-functioning children with autism and their understanding and use of pragmatic inferences in mental state words, especially comprehension of the presupposition and implicature of complements. Participants listened to sentences and judged their truth value in metaphor comprehension (matching metaphor with situation) and script inferencing (completing stories). For example, <Karen thinks that the door is shut> does not entail <the door is shut>; “I have butterflies in my stomach” goes with a > first-day-at-school picture. Dennis et al found that HFA participants know what mental state words mean, and what mental state words imply regarding given or presupposed knowledge (70-80%). However, HFA were not able to infer what mental state words implied in context, showing only a third level of performance of controls.

Noveck, Guelminger, Georgieff, Labruyere (2007) studied a famous TD phenomenon where children misinterpret the scope of ‘every’ relative to ‘no’ in sentences such as “every horse did not jump over the fence” in a context where two of three horses jumped over the fence, and judge this sentence to be false, taking it to mean “no horse jumped over”. Participants with autism (mean age 16 years, mental age 8 years, just under 70% passing first order false belief task) showed close to 50-50 child-like and adult-like interpretations. TD 4-year-old controls matched teenagers with autism in their performance. These results indicate difficulties with pragmatic scalar implicatures in autism.

Pijnacker, Hagoort, Buitelaar, Teunisse, Geurts (2009) studied scalar implicatures in adults with HFA and Asperger syndrome. These pragmatic inferences included inferring that “some” entails “not all”, in underinformative ‘some’, e.g. “some sparrows are birds”. While this is logically true, it is pragmatically false because it entails that “not all sparrows are birds”. ASD participants were as good as TD controls in judging such sentences. However HFA were almost twice as likely to show a logical interpretation for underinformative condition than Asperger participants, but there was no significant difference between HFA and typical controls. Pijnacker et al also looked at participants’ reaction times: people with HFA took about a second longer to make their judgments but those participants that showed logical interpretations were faster by a second than those who correctly gave pragmatic interpretations.

In summary, the severity of pragmatic deficits (whether speech-acts or scalar implicatures) tracks closely with severity of autistic symptoms.

5.3.3. Theory of mind in ASD

One proposal is that issues with pragmatics in autism are likely stemming from problems with theory of mind (TOM). Hale & Tager-Flusberg (2005) studied natural language samples of children with autism, looking for use of topic-related contingent utterances; subjects’ vocabulary and TOM was also measured. The researchers conclude that deficits in pragmatic language use are linked to deficits in TOM. Joseph & Tager-Flusberg (2004) investigated children with autism on several tests of understanding mental states and executive control skill such as working memory and planning. Children’s performance on these tests explained the variance in deficiencies in communication, but not in social interaction nor in repetitive behaviors. These findings suggest a deficit in social-perceptual processes, thus preventing children with autism interpreting mental states through information expressed in eye gazes and voices.

Steele, Joseph, Tager-Flusberg (2003) show that knowledge of Theory of Mind is differentially deficient in children autism (57 children, 4-14 years old (mean age 7), verbal
mental age of 5-6 years). Only 25% of children with Autism passed the first order theory of mind task and only 12% passed a second-order task. Children with PDD-NOS diagnosis (verbal mental age 8 years, chronological age 12 years) pass the first order false-belief more often than those with autism, at 36% (Sicotte & Stemberger 1999). Bauminger and Ksari (1999) show that children with HFA (8-14 years old, all within normal levels of intelligence) pass second order false belief task at 68% (vs 89% in matched TD controls) and that HFA’s performance correlates with full and verbal IQ, with those passing the task having higher standard scores by a standard deviation. Once study investigated the difference between PDD-NOS and HFA abilities. Begeer, Rieffe, Terwogt, Stockmann (2003) had an ingenious task where participants with ASD of average intelligence were given a task to perform but an experimenters sabotaged their efforts by removing a key object necessary for successful performance. Another experimenter, oblivious to the sabotage and thus with a false belief, directed participants to proceed with the task. 90% of TD controls informed the experimenter with the false belief of the sabotage, children with PDD-NOS (aged 7-11 years) also corrected the false belief, but only when they knew they would be reward to completing the initial task; children with HFA (aged 7-11 years) tended to not respond. Children with Asperger fail only higher-order metaphor and faux-pas tasks (Baron-Cohen et al, 1999). Thus the three ASD groups provide not only differential levels of language abilities, but also of pragmatic and theory of mind abilities, allowing for comparison and contrast. Excluding any of ASD group from a study may fail to provide a complete gradation of abilities, and may paint a misleading picture.

An important question is whether language (grammar/syntax, vocabulary) or nonverbal intelligence affects Theory of Mind development (or vise versa), and to what extent can a test of TOM measure development of social reasoning independently of other factors. This issues was raised in chapters 2 and 4 as well. Harris, de Rosnay, Pons (2005) argue that children with advanced language skills (i.e. those who are exposed to lexical and syntactic enrichment through maternal conversations) are better at mental state understanding. Studies by de Villiers demonstrate that typical children need to have syntactic structure for sentential complements prior to successfully performing on false-belief tasks, suggesting that first order theory of mind inferencing builds on structural linguistic knowledge (e.g. De Villiers & Pyers 2002). Hale & Tager-Flusberg (2003) support this in a training study where typical children trained on TOM task improved only on the TOM task, but children trained on a task of sentential complements improved both on the syntactic structure comprehension and on a TOM task. A recent study however fails to show such exclusive dependence on syntax, instead suggesting a dependence of false-belief understanding on general language ability: both syntactic and semantic proficiencies, but not working memory (Slade & Ruffman 2005). Perner et al (2003) also show that in German, where complement structures are required in sentences not involving mental states, complement structure is in place before TOM: children ages 2;6-4;6 could understand ‘want that’ complement clauses before ‘say that’, ‘think that’, and false belief.

For children with autism, one can also observe correlations between linguistic and cognitive abilities and TOM. One study comparing children with autism, PDD-NOS and typical children on multiple tests found social cognitive ability as measured by first and second order TOM tasks are best predicted by verbal memory, performance IQ, age and gender (Buitelaar et al 1999). Another study, following assessment of children with autism, Asperger syndrome, specific language impairment and typical controls, links first order theory of mind to comprehension and production of belief terms such as think, know and guess, showing links between development of TOM and development of communicative competence (Ziatas & Durkin 1998). Happe (1995) in
reviewing a multitude of studies concludes that TOM is especially impaired in autism, subjects needing to be of greater mental age to pass TOM tasks than typically developing children.

Thus, there is no clear answer as to causation, a nonverbal first-order false belief task (Colle, Baron-Cohen & Hill, 2007 and Call & Tomasello, 1999) may remove any potential language confounds and provide an estimate of social/pragmatic reasoning independent of language. Call & Tomasello tested typical children on the nonverbal and a verbal false-belief tasks and found nearly identical performance, confirming the validity of the nonverbal task. Colle et al tested autistic and specific-language-impaired children of similar very poor linguistic abilities (at typical 2 year-old level) and similar chronological age (8 years), but differing in mental age (5 years for autism group, 7 years for SLI group). They replicated the patterns found with standard verbal false-belief task: children with SLI, despite having same poor language skills as children with autism, perform well on the false-belief tasks, unlike children with autism. Thus the nonverbal false-belief task may be used as an estimate of social/pragmatic reasoning independent of language.

5.4. Why study articles in ASD?

The study of articles in TD acquisition has important implications for the study of the organization of the linguistic system in the population with ASD, in the face of reported difficulties with semantic and pragmatic aspects of linguistic knowledge.

The theories around typical acquisition of articles are pertinent to ASD. The egocentricity/Pragmatic deficits hypotheses were argued above (in Chapters 2 and 4.1) to be analogous to the conclusion in the study of cognitive development that young children have difficulties in understanding and using the fundamental concept that other people have minds and that these minds access information in particular ways, i.e. difficulties in Theory of Mind (e.g. Baron-Cohen, Tager-Flusberg, Cohen, 1993 & 2000; Astington, Harris, Olson, 1988). In other words on this view children have adult-like knowledge of semantic principles, but fail to use them correctly. While Egocentricity looks like it is quite analogous to TOM considerations, TD children already do well on first-order false belief tasks by age 4-5 (Chapter 4.1 and elsewhere), hence the overuse of the definite article in older, 6-7 year-old TD children (Chapter 3) is likely to have another explanation. The alternative theory was proposed in which children are hypothesized to lack (or have difficulty computing) a crucial semantic property of ‘the’, namely the semantic principle of the presupposition of Uniqueness/Maximality (Wexler 2003).

Thus, while the Egocentrism idea predicts correlation between presence of first order TOM in an individual and the individual’s competence in the article system (demonstrated by non-overuse of ‘the’), difficulties with Uniqueness predict that there should not be such a correlation. While the ideas of egocentrism and uniqueness may seem very similar, they differ in a crucial way. Egocentrism and lack of Theory of Mind (as possibly the same developmental cognitive stage) are limitations on the child’s pragmatic system, the system of a child’s understanding of what other minds know about a given linguistic context, while the lack of Uniqueness in a child would be a developmental linguistic stage in the computational system of language – the language learning mechanism, one in which the child did not know complete semantics of ‘the’, and not a limitation on understanding what other minds know.

The article ‘that’ is especially interesting because it has particular pragmatic features that condition usage, in addition to the property of Uniqueness. Referents of ‘that’, if not pointed to physically, must be invoked in memory, in the metalinguistic common knowledge of both conversation partners, i.e. in the nondefault context of interpretation. If TD children have
outgrown Egocentric/Pragmatic difficulties (or first-order false belief difficulties), but have difficulties with Uniqueness/Maximality, then they could use these additional pragmatic features to figure out the correct referent of ‘that’. That is, pragmatics may save the day even with the computational semantic difficulty. If Uniqueness is the correct explanation of overuse of ‘the’, we might expect TD children at a certain age to do well on sentences with ‘that’, even though they fail when it comes to correctly interpreting ‘the’. On the other hand, if overuse of ‘the’ is a pragmatic (Egocentricity/TOM) difficulty, then we do not predict a dissociation between ‘the’, on the one hand, and ‘that’ on the other – children should do well on both in comprehension. These are differential predictions for TD children (with results from TD children supporting the Uniqueness/Maximality deficit hypothesis in Chapters 3 and 4.1), and these predictions hold also of children with ASD (if the pragmatic features of familiarity or nondefault context of interpretation are known to ASD children).

Children with ASD however, given their inherent issues with theory of mind and pragmatics, may not be able to use additional pragmatic feature in ‘that’. If they are deficient in uniqueness, they will perform poorly on both ‘the’ and ‘that’. However if their knowledge of uniqueness is intact, they will be able to perform well on ‘the’, and their knowledge of uniqueness may be sufficient for successful interpretation of ‘that’ in contexts which do not strictly rely on non-default interpretation or familiarity of the referent or other people’s states of mind – as can be argued to be the case in the present comprehension task. If however ASD children are forced to rely on the pragmatic features in ‘that’, we may expect deficits in ‘that’ relative to ‘the’.

Finally, we have a chance to formulate a good hypothesis and to carry out a structured study of specifically semantic and pragmatics aspects of language. Interpretation of referents denoted by ‘that’ is guided by semantic principles and by pragmatic considerations, which could be problematic for people with ASD if they have to rely on those pragmatic considerations. Interpretation of ‘the’, on the other hand, is guided by semantic principles which form part of the computational core of the human grammatical ability, and can therefore be predicted to be either poor or good depending on the state of the grammar in individuals with ASD.

5.5. Participants

The ASD subjects for the present study were recruited in collaboration with Alexandra Perovic, through parent support groups and schools for children with disabilities. The clinical diagnosis was made by clinicians in the New England area and reported to us by participants’ parents. We included all children with a diagnosis of an autism spectrum disorder, ages 6-18, who were able to combine two-word utterances and who were able to complete all tasks in our study (tasks that included as control items comprehension of simple transitive sentences).

Participants were individually matched to TD controls on three measures, thus forming three control groups per disorder. TD controls were recruited from Boston/Cambridge area daycares and afterschool programs. One group of TD controls was matched to probands on raw non-verbal cognitive abilities, as measured by Matrices subtest of KBIT-II (usually no more than one point off). This group allows us to factor out the influence of general cognition on participants’ linguistic performance. The second control group was matched on raw scores of a standardized measure of receptive vocabulary (PPVT-III, maximum 4 points off), to tease apart the influence of general language abilities, i.e. lexical knowledge, on the more abstract aspects of linguistic knowledge of language assessed by our experimental task. The third control group was matched on raw score of standardized measure of grammar (TROG-2, usually no more than one
point off), to establish whether there are differences in the subtle aspects of linguistic knowledge between control and disordered children even if they are matched for the general level of grammatical competence. Every effort was made to match the participants on gender and to choose only those control participants whose standard scores (SS) on the standardized tests were close to the average (usually between 85 and 115). In this way the performance of control participants on the study task could not be interpreted as being due to their superior or inferior general cognitive or linguistic abilities. As participants with autism are aged between 6-18 years, it may not be desirable to match them to TD controls of the same chronological age. TD children at age 10 are expected to be at ceiling performance on the experimental probe (Chapter 3).

5.6. Methods

Standardized assessment tools are expected to provide a picture of a general level of participants’ linguistic functioning that will enable comparison of their abilities with those in other studies reported in the literature. Nonverbal reasoning were assessed with of Matrices subtest of the Kaufman Brief Intelligence Test (KBIT) (Kaufman & Kaufman, 1990). Receptive vocabulary was assessed with the aid of Peabody Picture Vocabulary Test-III (PPVT-III) (Dunn & Dunn, 1997). Receptive grammar was evaluated with Test of Reception of Grammar, 2nd edition (TROG-2) (Bishop, 2003).

However, these tools are limited in giving an exact evaluation of participants’ mastery of the particular linguistic structures that form the core of human computational knowledge of language. To investigate knowledge of semantics and pragmatics of articles novel experimental probe was developed, based on studies with typically developing children and work reported in the literature on the acquisition of targeted items.

Comprehension of definite (‘the’, ‘that’) and indefinite (‘a’, ‘another’) articles was investigated using an act-out experiment, with 6 tokens of each type of 4 conditions being used, allowing for within subject analysis of performance on particular types of stimuli. Performance on similar experiments has been investigated by the authors, by Karmiloff-Smith (1979, experiment 15) in TD French children, and others e.g. Maratsos (1976).

The experiment is in a form of a book made of felt cloth, with arrays of 3 or 6 identical objects permanently attached to the pages. Objects included apple, watermelon slice, car, carrot, Christmas tree, star, flower, icecream cone, baseball, heart, witch’s hat. Every page contains an array of target items and an array of distractor objects. Subjects are given two Velcro-backed ‘actors’, Fishy and Turtle. Thus the characters can be physically stuck to the page by children who can see what they were/are doing.

There were two ways of making sure the subjects are paying attention: whether they are affecting the right kind of object, and whether they are doing that with the correct actor (the order in which the actors affect the objects is randomized).

Firstly, subjects are introduced to the two actors and different objects that will be used. A few practice trials follow where children come comfortable with using actors to touch either same or different objects. Subjects are told that in this ‘game’, they have to decide whether to Fishy and Turtle go to same or different objects based on what the experimenter says. Article use is avoided during introduction to avoid biasing subjects’ responses.

Comprehension of items is evaluated when the experimenter gives the child instructions on how to manipulate the actors given the context of the current page. The first clause of instructions always contains the indefinite article ‘a’ and serves to establish a unique, salient,
visually distinct referent within the context set – the object that has Fishy or Turtle on it. The second clause contains one of the four articles – ‘a’, ‘another’, ‘the’, or ‘that’.

Instructions go like this: “Fishy touches an apple, [pause to allow for action of Fishy touching one of the apples] and Turtle touches a/another/the/that apple.” An adult would respond like this: pick two different objects upon hearing ‘another’, pick any object upon hearing ‘a’, and pick the same (salient) object upon hearing ‘the’ and ‘that’. The number of times the same object was subsequently acted upon served as the dependent variable.

Figure 5.1. Experimental set-up

Thus the comprehension experiment tests how children interpret ‘the’ versus ‘a’, specifically whether they know Uniqueness or whether they take account of the pragmatic conditions for establishing context sets – that both the speaker and listener must have means for determining the context sets. It also tests whether children use pragmatic properties (old information, known to listener) of ‘that’ to decide how to use the correct article.

5.7. Results

5.7.1. Some issues on grouping participants

Kjelgaard and Tager-Flusberg (2001) found receptive vocabulary (as measured by PPVT) to be a better indicator of linguistic abilities (articulation, expressive vocabulary) and of IQ than children’s performance on CELF (Clinical Evaluation of Language Fundamentals, which measures some syntax, working memory for language, morphology). This could be because they could only analyze half of their ASD participants on CELF, primarily those with near-normal IQ levels, as the rest did not pass CELF. Or this could be because CELF is an extensive test measuring much more than syntactic competence. In this regard, TROG (used here) focuses primarily on syntax comprehension and is much shorter than CELF.

Using the criteria that Kjelgaard and Tager-Flusberg (2001) adopted for defining language subgroups in autism, Roberts, Rice and Tager-Flusberg (2004) in their study of finiteness in autism spectrum disorders, divided their sample of children with ASD according to PPVT standard scores – 85 and above (normal), 70-84 (borderline), below 70 (impaired), and found overall that finiteness levels correlate with vocabulary levels.

Alternative binning may be problematic. For example binning by diagnosis (autism vs PDD-NOS vs Asperger) may be subject to parental or clinical biases – in some states, children may be more likely to receive services through schools with a diagnosis of autism, rather than a diagnosis PDD-NOS. Binning by non-verbal IQ is problematic due to potential inaccuracies of IQ testing, since people with ASD show different nonverbal reasoning skills based on the
particular tests (Dawson, Soulieres, Gernsbacher, Mottron 2007).

However dividing children by vocabulary makes a key assumption that vocabulary is indicative of syntactic/semantic/linguistic ability. While vocabulary and cognitive scores may be causative to syntactic language abilities in people with mental retardation (e.g. Facon, Facon-Bollengier, Grubar 2002), the dissociation between verbal and nonverbal abilities is known in autism, and thus a priori it may incorrect to make this assumption. In typically developing children, syntactic/grammatical abilities were shown to be independent of vocabulary levels (Rice, Wexler, Hershberger 1998).

We must keep in mind that one of the ultimate goals is endophenotyping – finding subgroups characterized in a clear-cut ways, for enhancing gene finding. (e.g. Bradford et al (2001) show that incorporating language phenotypes strengthens evidence of linkage to autism). Another goal is to enhance diagnosis (e.g. Botting, Conti-Ramsden (2003)).

Thus we will explore the optimal binning/grouping strategy, and to that end our sample will be divided five times: by diagnosis, by standard score on nonverbal reasoning (KBIT), by standard score on vocabulary (PPVT), by standard score on grammar (TROG), and by performance patterns on the test of determiners.

Recall that the task involves interpreting ‘a’, ‘another’, ‘the’ or ‘that’ as pointers to an already-established referent or to a new referent. Thus “number same responses” refers to the number of times subjects decide that an already-talked-about entity is referred to, and attach the two toy actors to a single object out of an array of several. This seems better than percentage correct, because, as discussed above, for the ‘a’ condition, the correct response depends on the interpretation. The target adult responses would be closer to 0-10% same actions for ‘a’ and ‘another’, and close to 100% same actions for ‘the’ and ‘that’.

5.7.2 Results for All ASD participants

Participants

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>KBIT SS</th>
<th>PPVT SS</th>
<th>TROG SS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ASD</strong></td>
<td>mean</td>
<td>11.17</td>
<td>87.39</td>
<td>84.65</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>0.66</td>
<td>7.35</td>
<td>4.98</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>6</td>
<td>19</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>number</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td><strong>KBIT controls</strong></td>
<td>mean</td>
<td>7.72</td>
<td>106.26</td>
<td>107.90</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>0.54</td>
<td>6.64</td>
<td>2.14</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>3</td>
<td>15</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>number</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td><strong>PPVT controls</strong></td>
<td>mean</td>
<td>8.64</td>
<td>104.80</td>
<td>100.87</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>0.70</td>
<td>7.38</td>
<td>1.89</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>3</td>
<td>17</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>number</td>
<td>31</td>
<td>30</td>
<td>31</td>
</tr>
<tr>
<td><strong>TROG controls</strong></td>
<td>mean</td>
<td>6.48</td>
<td>103.35</td>
<td>104.06</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>0.58</td>
<td>6.64</td>
<td>2.28</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>3</td>
<td>17</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>number</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
</tbody>
</table>

Table 5.2. Details on all the ASD and all the TD controls
Results

Figure 5.3. Number ‘same’ responses (out of 6) as a measure of knowledge of determiners in all ASD and their TD controls (error bars represent +/− 1 SE)

We find that knowledge of indefinite articles ‘a’ and ‘another’ does not differ among participant groups, however knowledge of ‘the’ and ‘that’ does so.

MANOVA (between subjects factor = group (ASD vs three TD controls), dependent variables = performance on ‘a’, ‘another’, ‘the’, ‘that’) showed significant effect of group for ‘the’ (F(3,120)=3.6, p=.015) and ‘that’ (F(3,120)=3.4, p=.019), but not for ‘a’ and ‘another’. Tukey Post Hoc analysis of group showed significant differences on ‘the’ between ASD group and KBIT controls (p=.012), and on ‘that’ between ASD and KBIT controls (p=.022) and between ASD and PPVT controls (p=.05). However R² for the model was less than .1 for all dependent variables.

Adding chronological age as a covariate increased R² for the model to .33 (adjusted R² .3) for ‘the’ and .24 (adjusted R² .21) for ‘that’, and showed significant effect of age of participants for ‘another’ (F(1,119)=3.9, p=.048), ‘the’ (F(1,119)=42.5, p<.001) and ‘that’ (F(1,119)=24.8, p<.001). The effect of group was significant for ‘the’ (F(3,119)=11.5, p<.001) and ‘that’ (F(3,119)=8.9, p<.001).

We further investigated the effect of participants’ standard scores on article performance, by adding KBIT SS, PPVT SS and TROG SS scores as covariants. The effect of group was now lost, the effect of age preserved at same significance levels for ‘the’ and ‘that’. KBIT SS scores showed effect on ‘the’ (F(1,115)=5.4, p=.022). PPVT SS scores showed no effect. TROG SS scores showed effect on ‘that’ only (F(1,115)=11.7, p=.001). R² for the model increased to .49 (adjusted R² .43) for ‘the’ and .49 (adjusted R² .46) for ‘that’.

We now turn the focus specifically on the definite determiners, since this is where variation lies. Participants were subdivided by their pattern of performance between ‘the’ and ‘that’, with “A” pattern denoting good performance on both ‘the’ and ‘that’, “B” pattern denoting bad performance on ‘the’ with a relatively better performance on ‘that’, “C” pattern denoting a bad performance on ‘that’ with a relatively better performance on ‘the’, and finally “D” pattern denoting a bad performance in both ‘the’ and ‘that’. ‘Good’ performance was considered if children got 4/6 or greater ‘same’ actions on the same object. ‘Bad’ performance was 3/6 or less.
For ‘B’ and ‘C’ patterns children had to get a difference of 2 ‘same’ actions between ‘the’ and ‘that’ for the difference to count. Thus difference between 3 and 4 does not count, but a difference between 3 and 5 does. This is was done to figure out whether children are showing adult-like knowledge (A pattern), showing deficits especially in semantics (but not pragmatics) (B pattern, found in TD children), showing deficits especially in pragmatics (but not in semantics) (C pattern, not found in TD children), or showing deficits across the board (D pattern).

The proportion of participants showing these patterns is graphed. It becomes immediately clear that participants are not showing a random performance. As a group, close to 60% of children with ASD show no knowledge of articles (pattern “D”), with only 25% showing good knowledge (“A” pattern), and the remaining 15% showing the “B” pattern where they do better on ‘that’ than on ‘the’. No control group has similar distribution of participants showing that many poor responses. The TROG control group (which is the youngest) is showing over 40% “A” pattern. In the KBIT and PPVT control groups almost 60% of participants are showing good proficiency with ‘the’ and ‘that’. Observe that only one participant in the TROG control group is showing the “C” pattern.

This analysis illustrates one of the main points of the data. Children with ASD are more deficient in their knowledge of articles than expected given their nonverbal reasoning abilities, more deficient than expected given their vocabulary levels, and more deficient than expected given their grammatical levels. As a group ASD children are older than their controls, but their world experience does not help them determine salient referents in context sets.

Figure and Table 5.4. Proportion of all ASD participants showing adult like (A), semantic deficit (B), pragmatic deficit (C) or null knowledge (D) patterns.

<table>
<thead>
<tr>
<th></th>
<th>ASD</th>
<th>KBIT-control</th>
<th>PPVT-control</th>
<th>TROG-control</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>0.58</td>
<td>0.29</td>
<td>0.29</td>
<td>0.29</td>
</tr>
<tr>
<td>C</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
</tr>
<tr>
<td>B</td>
<td>0.16</td>
<td>0.13</td>
<td>0.13</td>
<td>0.26</td>
</tr>
<tr>
<td>A</td>
<td>0.26</td>
<td>0.58</td>
<td>0.58</td>
<td>0.42</td>
</tr>
</tbody>
</table>

ANOVA (between group factor: ASD vs 3 TD groups, dependent variable = proportion participants showing a particular pattern) showed significant effect of group (F(3,120)=5.7, p=.025), with Tukey post hoc tests showing ASD group to be different from KBIT group and
PPVT group (p=.041 for both), with R² for the model less than .1. Adding age as covariant increased R² to .3 (adjusted R² .28). Both effects were significant: age (F(1,119)=51.8, p<.001), group (F(3,119)=14.8, p<.001). We further investigated the effect of participants’ standard scores on article knowledge patterns, by adding KBIT SS, PPVT SS and TROG SS scores as covariants. There were still significant effects of age (F(1,115)=46.3, p<.001), but not group. KBIT SS and PPVT SS scores here showed no effect. TROG SS scores showed effect (F(1,115)=5.9, p=.017). R² for the model increased to .45 (adjusted R² .42).

This analysis shows that whether we are measuring knowledge of individual articles, or patterns of definite article knowledge, the effects are similar. From here onward, we will analyze patterns of performance on the definite articles in different groupings of participants.

5.7.3. Grouping By diagnosis

The goal in this section is to make sense of the factors that drive participants knowledge of articles. First of, we are dividing all ASD participants into their clinical diagnoses groups.

Participants

<table>
<thead>
<tr>
<th>Ages</th>
<th>Nonverbal Reasoning AE/SS</th>
<th>Vocabulary AE/SS</th>
<th>Grammar AE/SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autism (13)</td>
<td>11.29 (6.47-17.23)</td>
<td>6.14 / 66.85 (40-103)</td>
<td>5.28 / 59.85 (40-87)</td>
</tr>
<tr>
<td>PDD-NOS (10)</td>
<td>9.38 (5.99-12.21)</td>
<td>9.05 / 97.5 (69-126)</td>
<td>8.75 / 94.8 (71-120)</td>
</tr>
</tbody>
</table>

Table 5.5. ASD participants divided by their clinical diagnosis

<table>
<thead>
<tr>
<th>Ages</th>
<th>KBIT controls</th>
<th>PPVT controls</th>
<th>TROG controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autism (aut) controls</td>
<td>5.73 (3-7.69)</td>
<td>5.64 (3-7.95)</td>
<td>4.28 (3.62-5.35)</td>
</tr>
<tr>
<td>PDD-nos (pdd) controls</td>
<td>7.95 (4-14.59)</td>
<td>8.51 (5.91-12.24)</td>
<td>6.98 (3-11.91)</td>
</tr>
<tr>
<td>Asperger (asp) controls</td>
<td>10.66 (7.72-15.21)</td>
<td>13.69 (9.14-17.11)</td>
<td>9.44 (5.2-17.08)</td>
</tr>
</tbody>
</table>

Table 5.6. TD controls corresponding to subdivisions by ASD clinical diagnosis

The participants included thirteen children clinically diagnosed with autism, ten children clinically diagnosed with Pervasive Developmental Disorder – Not Otherwise Specified (PDD-NOS), and eight children clinically diagnosed with Asperger syndrome (AS). Their performance on a battery of standardized tests showed varying degrees of abilities: children with autism showed significant impairments in comprehension of grammar, vocabulary, and non-verbal reasoning; children with PDD-NOS showed age-level vocabulary and non-verbal reasoning, but were delayed in grammar. Children with Asperger syndrome showed above average vocabulary, and average non-verbal reasoning abilities and grammar skills. Thus in all categories of ASD, grammar skills lag behind nonverbal IQ and vocabulary skills. Nine groups of younger typically developing controls were included, individually matched to probands on raw scores on TROG, PPVT, and KBIT: three groups of controls for each of the three proband groups, allowing us to compare probands’ performance on articles to that of typically developing children of similar cognitive profiles.
If there are no differences in performance on articles between younger controls and probands, then development of articles proceeds in typical, albeit delayed (relative chronological age) fashion. However, if probands are performing differently or considerably worse than controls than article development, than it is specifically impaired in ASD.

**Results**

Multivariate ANOVA (the four articles as dependent variables, the 12 subject groups as between subject variables) showed significant effect of group on ‘the’ (F(11,112)=8.6, p<.001) and ‘that’, (F(11,112)=8.2, p<.001) with R² of .45 (.4 adjusted). Tukey’s post hoc tests showed significant differences for ‘the’ between the Asperger and the autism groups (p<.001), autism and its KBIT control group (p=.046), and autism and all PDD-NOS and Asperger control groups (p=.012 - p<.001), and for ‘that’ between the Asperger and the autism groups (p<.001), autism and PDD-NOS (p=.021), autism and its KBIT controls (p=.004), autism and its PPVT controls (p=.016), and autism and all PDD-NOS and Asperger groups (all ps<.001). No groups differ in their performance on either ‘a’ or ‘another’. Addition of standardized test scores as covariates, showed that only TROG SS has significant effect only ‘that’ (F(11,108)=10.6, p=.001).
MANOVA of just the ASD participants binned by diagnosis revealed significant effect of such binning on ‘the’ (F(2,28)=16.3, p<.001) and ‘that’ (F(2,28)=26.5, p<.001), and overall model had R^2 of .54 (adjusted .5) for ‘the’ and .65 (adjusted .63) for ‘that’.

Binning participants by diagnosis brings out very clear patterns. Children with autism are showing a profound absence of knowledge of articles. They are more impaired at articles than all of their control groups, more impaired than their nonverbal, verbal and grammar mental ages would predict. This is suggestive of a severe semantic and pragmatic deficit in these children.

Children with PDD-NOS as a group show better performance than children with autism, although the difference is not significant due to extensive within group variation in PDD-NOS, however like children with autism, the PDD-NOS are performing worse than their controls, worse than predicted by their mental ages. Note, PDD-NOS are performing slightly better on ‘that’ than on ‘the’ – a pattern found in typical development.

Children with Asperger syndrome in start contrast show knowledge of articles similar to their controls – at almost adult-like levels. Although the differences between groups are not significant, the Asperger group performs most like their grammar (TROG) controls. Observe that these results are surprising. We predicted children with Asperger syndrome to perform more poorly on ‘that’ than on ‘the’, given their pragmatic deficits. If anything, children with Asperger seem to be doing a bit worse on ‘the’ than on ‘that’ – a pattern found in typical development.

Endophenotypes within clinically defined groups

Patterns of performance within ASD groups were also analyzed. We are calling these endophenotypes, which (1) Are internal phenotypes & measurable components unseen by unaided eye but discoverable by examination; (2) Are simpler clues to the genetics underpinning psychiatric disease than the disease syndrome itself; (3) Are latent genetically influenced traits, which may be related only indirectly to disease symptoms. (from Glannon 2004, Skuse 2001, Gottesman & Gould 2003). We explore the idea that a constrained test of semantics and pragmatics (of articles) may contribute to accurately defining the autism spectrum disorders endophenotypes allowing for a more precise genotypic characterization.

Participants with autism did not show any differences in performance within their clinical subgroup. Participants with Asperger syndrome and with PDD-NOS however did.

<table>
<thead>
<tr>
<th>Further characterization of performance</th>
<th>Age</th>
<th>Nonverbal Reasoning AE/SS</th>
<th>Vocabulary AE/SS</th>
<th>Grammar AE/SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDD-NOS: autism-like performance (5)</td>
<td>9.47</td>
<td>9.4/100</td>
<td>9.73/102.2</td>
<td>5.9/75.8</td>
</tr>
<tr>
<td></td>
<td>(5.99-12.21)</td>
<td>(86-126)</td>
<td>(73-120)</td>
<td>(55-99)</td>
</tr>
<tr>
<td>PDD-NOS: child-like performance (2)</td>
<td>9.2</td>
<td>6.96/86</td>
<td>7.5/83.5</td>
<td>6.0/77.5</td>
</tr>
<tr>
<td></td>
<td>(7.06-11.34)</td>
<td>(69-103)</td>
<td>(82-85)</td>
<td>(76-79)</td>
</tr>
<tr>
<td>PDD-NOS: adult-like performance (3)</td>
<td>9.36</td>
<td>9.86/101</td>
<td>7.95/90</td>
<td>7.42/84.67</td>
</tr>
<tr>
<td></td>
<td>(6.08-11.19)</td>
<td>(93-111)</td>
<td>(71-102)</td>
<td>(74-97)</td>
</tr>
<tr>
<td>Asperger: child-like performance (3)</td>
<td>10.62</td>
<td>15.19/126.67</td>
<td>13.28/118.33</td>
<td>8.14/91</td>
</tr>
<tr>
<td></td>
<td>(6.12-16.19)</td>
<td>(104-151)</td>
<td>(91-133)</td>
<td>(74-109)</td>
</tr>
<tr>
<td>Asperger: adult-like performance (5)</td>
<td>14.76</td>
<td>13.95/97</td>
<td>17.32/108.6</td>
<td>11.37/96.4</td>
</tr>
<tr>
<td></td>
<td>(10.82-18.56)</td>
<td>(73-109)</td>
<td>(100-121)</td>
<td>(76-111)</td>
</tr>
</tbody>
</table>

Table 5.10. Endophenotypes within clinically defined groups
Figure and Table 5.11. Article endophenotypes within clinical diagnosis subgroups and their TD controls

<table>
<thead>
<tr>
<th></th>
<th>ASPER-GER</th>
<th>asp kbit</th>
<th>asp ppvt</th>
<th>asp trop</th>
<th>AUTISM</th>
<th>aut kbit</th>
<th>aut ppvt</th>
<th>aut trop</th>
<th>PDD-NOS</th>
<th>pdd kbit</th>
<th>pdd ppvt</th>
<th>pdd trop</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.46</td>
<td>0.54</td>
<td>0.62</td>
<td>0.50</td>
<td>0.30</td>
<td>0.20</td>
<td>0.10</td>
</tr>
<tr>
<td>C</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.10</td>
</tr>
<tr>
<td>B</td>
<td>0.38</td>
<td>0.00</td>
<td>0.00</td>
<td>0.13</td>
<td>0.00</td>
<td>0.15</td>
<td>0.23</td>
<td>0.23</td>
<td>0.20</td>
<td>0.20</td>
<td>0.10</td>
<td>0.40</td>
</tr>
<tr>
<td>A</td>
<td>0.63</td>
<td>1.00</td>
<td>1.00</td>
<td>0.88</td>
<td>0.00</td>
<td>0.38</td>
<td>0.23</td>
<td>0.15</td>
<td>0.30</td>
<td>0.50</td>
<td>0.70</td>
<td>0.40</td>
</tr>
</tbody>
</table>

ANCOVA (12 participant groups by Pattern by KBIT, PPVT and TROG SS scores as covariates) showed significant effect of group (F(11,108)=4.4, p<.001), and a significant effect of TROG SS (F(1,108)=5.3, p=.023). R² for the model is .47 (adjusted .41). Tukey’s post hoc tests showed significant difference between the autism and the Asperger groups (p<.001), autism and its KBIT controls (p=.031), autism and all pdd-nos and Asperger control groups (ps<=.001).

Participants with Asperger syndrome showed two distinct patterns of performance. 63% of AS were completely adult-like - showing perfect knowledge, and 37% showed the pattern of performance found in typical development, namely poor performance on ‘the’, and good performance on ‘that’ (B pattern), but with a more pronounced difference than found typically. Such a pattern goes against our predictions: if anything children with AS should do worse on ‘that’ (more pragmatics involved) than on ‘the’. While good performance on ‘that’ continues to await explanation, poor performance on ‘the’ seems to go with lower grammar scores: the poorly performing AS subgroup has on average a lower TROG score than the adult-like AS subgroup. Thus, given good nonverbal intelligence and good vocabulary skills, overall grammatical development is predictive of knowledge of the article ‘the’ in participants with Asperger syndrome.

Children with PDD-NOS showed three patterns of performance. These patterns were the two found in the Asperger groups, child-like (B pattern, 20% of PDD-NOS subjects) and adult-like (A pattern, 30% of PDD-NOS subjects), and the pattern found in the group with autism (C pattern, 50% of PDD-NOS subjects), with participants’ scores on a standard measure of grammar correlating with their performance on ‘the’ and ‘that’. Thus, as with the Asperger subgroups, overall grammatical development is the driving force for article development in people with PDD-NOS.
5.7.4. Grouping By articles/determiners

Herein, we are subdividing ASD participants (along with their KBIT, PPVT, and TROG controls) by their patterns of performance on articles, and investigating the resulting patterns of their nonverbal reasoning, vocabulary, and grammar abilities.

### Participants

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Nonverbal Reasoning AE/SS</th>
<th>Vocabulary AE/SS</th>
<th>Grammar AE/SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null (no) performance (18)</td>
<td>10.79 (5.99-17.23)</td>
<td>7.04/76.06 (40-126)</td>
<td>6.51/71.61 (40-120)</td>
<td>4.54/61.61 (55-99)</td>
</tr>
<tr>
<td>Child-like (diff) performance (5)</td>
<td>10.05 (6.12-16.19)</td>
<td>11.9/110.4 (69-151)</td>
<td>10.97/104.4 (82-133)</td>
<td>7.28/85.6 (74-109)</td>
</tr>
<tr>
<td>Adult-like (same) performance (8)</td>
<td>12.74 (6.08-18.56)</td>
<td>12.42/98.5 (73-111)</td>
<td>13.8/101.63 (71-121)</td>
<td>9.89/92 (74-111)</td>
</tr>
</tbody>
</table>

Table 5.12. ASD participants divided by their performance on articles

<table>
<thead>
<tr>
<th>ages</th>
<th>kbit controls</th>
<th>ppvt controls</th>
<th>trog controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>no controls</td>
<td>6.26 (3-9.01)</td>
<td>6.61 (3-12.24)</td>
<td>4.79 (3-10)</td>
</tr>
<tr>
<td>dif controls</td>
<td>9.91 (6.58-15.21)</td>
<td>10.12 (5.91-16.76)</td>
<td>8.27 (5.2-17.08)</td>
</tr>
<tr>
<td>same controls</td>
<td>9.63 (4-14.59)</td>
<td>12.3 (7.28-17.11)</td>
<td>9.18 (4.89-12.19)</td>
</tr>
</tbody>
</table>

Table 5.13. TD controls corresponding to subdivisions by ASD performance on articles

### Results

ANOVA (12 participant groups by KBIT, PPVT, and TROG SS scores as dependent variables) shows significant effect of such grouping on all variables: KBIT SS (F(11,111)=4.5, p<.001, R^2=.31 (adjusted .24)), PPVT SS (F(11,111)=6.8, p<.001, R^2=.4 (adjusted .34)), TROG SS (F(11,111)=16.6, p<.001, R^2=.6 (adjusted .58)). Tukey’s post hoc test showed the following contrasts. The null ASD group was significant different from the child-like ASD group on KBIT (p=.005), PPVT (p=.003), and TROG (p=.007). The null ASD group was also different form the adult-like ASD group on PPVT (p=.001), and TROG (p<.001).

Dividing ASD participants by their performance on articles indicates that those who show no knowledge of articles (null group) have lower nonverbal reasoning, vocabulary and grammar abilities, than those who show at least some knowledge of articles (child-like (diff) and adult-like (same)). The difference between the child-like group (who shows partial knowledge) and the adult-like group lies in their grammar skills what are in the average range for adult-like group and below-average for the child-like group.
Figure and Table 5.14. Article endophenotypes within subgroups defined by article knowledge and their TD controls

<table>
<thead>
<tr>
<th></th>
<th>DIFF</th>
<th>dif kbit</th>
<th>dif ppvt</th>
<th>dif trog</th>
<th>NO</th>
<th>no kbit</th>
<th>no ppvt</th>
<th>no trog</th>
<th>SAME</th>
<th>same kbit</th>
<th>same ppvt</th>
<th>same trog</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>0</td>
<td>0.2</td>
<td>0.2</td>
<td>0</td>
<td>1</td>
<td>0.44</td>
<td>0.44</td>
<td>0.50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.00</td>
<td>0.06</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.6</td>
<td>0</td>
<td>0.22</td>
<td>0.22</td>
<td>0.28</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>0.8</td>
<td>0.8</td>
<td>0.4</td>
<td>0</td>
<td>0.33</td>
<td>0.33</td>
<td>0.17</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

ANOVA (12 participant groups by Pattern as dependent variable by KBIT, PPVT and TROG SS scores as covariates) showed significant effect of group (F(11,108)=7.0, p<.001), and the effect of TROG SS approached significance (F(1,108)=2.9, p=.088). R² for the model is .55 (adjusted .5). Tukey post hoc tests showed significant differences between the null and child-like groups (p=.005), the null and the adult-like groups (p<.001), the null and its KBIT (p=.001) and PPVT (p=.001) and TROG (p=.039) controls.

MANOVA of just the ASD participants binned by performance revealed significant effect of such binning on ‘the’ (F(2,28)=156.2, p<.001) and ‘that’ (F(2,28)=170.7, p<.001), and overall model had R² of .92 for both ‘the’ and ‘that’.

5.7.5. Grouping by IQ

The goal in this section is to further make sense of the factors that drive participants knowledge of articles. Here, we are dividing ASD participants (along with their KBIT, PPVT, and TROG controls) according to their IQ levels, to investigate whether this works for reducing the variation across ASD.
Participants

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Nonverbal Reasoning AE/SS</th>
<th>Vocabulary AE/SS</th>
<th>Grammar AE/SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline IQ (5)</td>
<td>12.28 (6.91-15.69)</td>
<td>4.47/43.6 (40-54)</td>
<td>5.02/55.6 (40-71)</td>
<td>3.8/55.6 (55-58)</td>
</tr>
<tr>
<td>Low IQ (6)</td>
<td>13.13 (8.04-17.23)</td>
<td>7.29/69.5 (66-73)</td>
<td>7.65/64.5 (40-104)</td>
<td>4.89/62 (55-76)</td>
</tr>
<tr>
<td>Normal IQ (20)</td>
<td>10.31 (5.99-18.56)</td>
<td>10.98/103.7 (84-151)</td>
<td>10.58/97.95 (50-133)</td>
<td>7.44/81.15 (55-111)</td>
</tr>
</tbody>
</table>

Table 5.15. ASD participants divided by their Nonverbal Reasoning Standard Score

<table>
<thead>
<tr>
<th>ages</th>
<th>kbit controls</th>
<th>ppvt controls</th>
<th>trog controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>base controls</td>
<td>4.17 (3-5.61)</td>
<td>5.99 (4.81-7.95)</td>
<td>3.94 (3.62-4.85)</td>
</tr>
<tr>
<td>low controls</td>
<td>6.9 (4.59-8.1)</td>
<td>7.37 (3-12.1)</td>
<td>5.32 (3.74-7.06)</td>
</tr>
<tr>
<td>norm controls</td>
<td>8.85 (4-15.21)</td>
<td>9.69 (3.62-17.11)</td>
<td>7.46 (3-17.08)</td>
</tr>
</tbody>
</table>

Table 5.16. TD controls corresponding to subdivisions by ASD IQ

Results

ANCOVA (12 participant groups by Pattern by KBIT, PPVT and TROG SS scores as covariates) showed significant effect of group (F(11,108)=2.8, p=.003), and effect of TROG SS
(F(1,108)=8.8, p=.004). R² for the model is .41 (adjusted .33). Tukey post hoc tests showed no significant differences between ASD participant groups. Only baseline ASD group was significantly different from Normal group’s controls (KBIT p=.009, PPVT p=.009, TROG p=.007).

MANOVA of just the ASD participants binned by IQ revealed moderate significant effect of such binning on ‘the’ (F(2,28)=3.4, p=.05) and ‘that’ (F(2,28)=4.2, p=.025), and overall model had R² of .19 (adjusted .14) for ‘the’, and .23 (adjusted .18) for ‘that’.

Dividing ASD participants by their nonverbal reasoning abilities blurs distinctions between groups. While those with baseline IQ show no knowledge of articles, those with low and average IQ show mixed pattern of knowledge. Thus nonverbal reasoning levels are not the driving force behind comprehension of articles.

5.7.6. Grouping by Vocabulary

The goal in this section is to further make sense of the factors that drive participants knowledge of articles. Here, we are dividing ASD participants (along with their KBIT, PPVT, and TROG controls) according to their vocabulary levels – something that Kjelgaard and Tager-Flusberg (2001) argued works well for reducing the variation across ASD.

Participants

<table>
<thead>
<tr>
<th>Participants</th>
<th>Age</th>
<th>Nonverbal Reasoning AE/ SS</th>
<th>Vocabulary AE/SS</th>
<th>Grammar AE/SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline Vocab (8)</td>
<td>13.65 (10.48-17.23)</td>
<td>6.26/56.75 (40-84)</td>
<td>4.98/47.25 (40-58)</td>
<td>3.87/55 (55-55)</td>
</tr>
<tr>
<td>Low Vocab (7)</td>
<td>9.23 (6.91-11.34)</td>
<td>6.88/80.86 (54-99)</td>
<td>6.59/77 (71-82)</td>
<td>5.1/64.14 (55-83)</td>
</tr>
<tr>
<td>Normal Vocab (16)</td>
<td>10.78 (5.99-18.56)</td>
<td>11.71/105.56 (73-151)</td>
<td>12.28/106.69 (85-133)</td>
<td>8.16/86.5 (55-111)</td>
</tr>
</tbody>
</table>

Table 5.18. ASD participants divided by their Vocabulary Standard Score

<table>
<thead>
<tr>
<th>ages</th>
<th>kbit controls</th>
<th>ppvt controls</th>
<th>trog controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base controls</td>
<td>5.68 (3-7.69)</td>
<td>5.55 (3-7.95)</td>
<td>4.17 (3.62-4.85)</td>
</tr>
<tr>
<td>Low controls</td>
<td>6.49 (3.9-9.06)</td>
<td>6.93 (4.81-9.1)</td>
<td>5.86 (3.81-10.14)</td>
</tr>
<tr>
<td>Norm controls</td>
<td>9.27 (4-15.21)</td>
<td>10.94 (4.57-17.11)</td>
<td>7.91 (3-17.08)</td>
</tr>
</tbody>
</table>

Table 5.19. TD controls corresponding to subdivisions by ASD Vocabulary
Results

ANCOVA (12 participant groups binned by vocabulary, by Pattern, by KBIT, PPVT and TROG SS scores as covariates) showed significant effect of group (F(11,108)=1.9, p=.042), and effect of TROG SS (F(1,108)=6.2, p=.014). R² for the model is .36 (adjusted .28). Tukey post hoc tests showed significant differences between the baseline and the normal ASD groups (p=.026), baseline and normal controls (p<.001).

MANOVA of just the ASD participants binned by vocabulary revealed significant effect of such binning on ‘the’ (F(2,28)=8.4, p=.001) and ‘that’ (F(2,28)=12, p<.001), and overall model had R² of .37 (adjusted .33) for ‘the’ and .47 (adjusted .43) for ‘that’.

Dividing ASD participants by vocabulary levels seems a bit better than dividing children by nonverbal reasoning, but still not great. There is not much difference between the baseline and the low vocabulary groups.

5.7.7. Grouping By Grammar

The goal in this section is to further make sense of the factors that drive participants knowledge of articles. Here, we are dividing ASD participants (along with their KBIT, PPVT, and TROG controls) according to their grammar levels – something that Kjelgaard and Tager-Flusberg (2001) showed does not works well for reducing the variation across ASD. 

Figure and Table 5.20. Article endophenotypes within subgroups defined by Vocabulary and their TD controls

<table>
<thead>
<tr>
<th></th>
<th>BASE</th>
<th>base kbit</th>
<th>base ppvt</th>
<th>base trog</th>
<th>LOW</th>
<th>low kbit</th>
<th>low ppvt</th>
<th>low trog</th>
<th>NORM</th>
<th>norm kbit</th>
<th>norm ppvt</th>
<th>norm trog</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>1.00</td>
<td>0.38</td>
<td>0.63</td>
<td>0.75</td>
<td>0.71</td>
<td>0.57</td>
<td>0.14</td>
<td>0.14</td>
<td>0.31</td>
<td>0.13</td>
<td>0.19</td>
<td>0.13</td>
</tr>
<tr>
<td>C</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.06</td>
</tr>
<tr>
<td>B</td>
<td>0.00</td>
<td>0.25</td>
<td>0.25</td>
<td>0.13</td>
<td>0.14</td>
<td>0.00</td>
<td>0.14</td>
<td>0.57</td>
<td>0.25</td>
<td>0.13</td>
<td>0.06</td>
<td>0.19</td>
</tr>
<tr>
<td>A</td>
<td>0.00</td>
<td>0.38</td>
<td>0.13</td>
<td>0.13</td>
<td>0.14</td>
<td>0.43</td>
<td>0.71</td>
<td>0.29</td>
<td>0.44</td>
<td>0.75</td>
<td>0.75</td>
<td>0.63</td>
</tr>
</tbody>
</table>
Participants

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Nonverbal Reasoning AE/SS</th>
<th>Vocabulary AE/SS</th>
<th>Grammar AE/SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline Grammar (16)</td>
<td>11.21 (6.47-17.23)</td>
<td>6.61/71.19 (40-103)</td>
<td>6.18/66.63 (40-115)</td>
<td>4.27/57.56 (55-69)</td>
</tr>
<tr>
<td>Low Grammar (6)</td>
<td>9.43 (6.08-15.15)</td>
<td>8.04/93.67 (69-125)</td>
<td>8.96/95.83 (71-131)</td>
<td>5.81/77 (74-83)</td>
</tr>
</tbody>
</table>

Table 5.21. ASD participants divided by their Grammar Standard Score

<table>
<thead>
<tr>
<th>ages</th>
<th>kbit controls</th>
<th>ppvt controls</th>
<th>trog controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>base controls</td>
<td>6.1 (3-9.01)</td>
<td>6.44 (3-12.24)</td>
<td>4.57 (3.62-6.21)</td>
</tr>
<tr>
<td>low controls</td>
<td>7.09 (4-9.06)</td>
<td>8.51 (5.91-12.1)</td>
<td>6.56 (4.89-10.14)</td>
</tr>
<tr>
<td>norm controls</td>
<td>11.01 (6.42-15.21)</td>
<td>12.65 (6.2-17.11)</td>
<td>9.82 (3-17.08)</td>
</tr>
</tbody>
</table>

Table 5.22. TD controls corresponding to subdivisions by ASD Grammar

Results

ANCOVA (12 participant groups binned by vocabulary, by Pattern, by KBIT, PPVT and TROG SS scores as covariates) showed significant effect of group (F(11,108)=3.8, p<.001), and
effect of TROG SS ($F(1,108)=6.3$, $p=.014$). $R^2$ for the model is .45 (adjusted .38). Tukey post hoc
tests showed significant differences between the baseline and the low ASD groups ($p<.001$), the
baseline and the normal ASD groups ($p=.001$). Baseline ASD group differed from its KBIT
control ($p=.009$), its PPVT control ($p=.016$), and from all low and normal control groups
($p<.001$).

MANOVA of just the ASD participants binned by grammar revealed significant effect of
such binning on ‘the’ ($F(2,28)=24.1$, $p<.001$) and ‘that’ ($F(2,28)=79.9$, $p<.001$), and overall
model had $R^2$ of .63 (adjusted .61) for ‘the’ and .85 (adjusted .84) for ‘that’.

Dividing ASD participants by their grammatical abilities makes the differences in
performance on articles very evident between groups – more so than when dividing participants
by their IQ or vocabulary levels. Now there is a clear difference between the baseline group and
the low and the normal ASD groups. The difference in performances are there in the data, and
binning participants by their grammatical abilities brings out these differences.

5.8. Discussion

In summary, children with autism show no knowledge of definite articles, children with
PDD-NOS show a delay in definite article development, and lastly those with Asperger syndrome
show near-perfect adult knowledge of definite articles. Observe that all populations except autism
differentiate indefinite (a, another) and definite (the, that) articles. That the autism group only
gives 6% same responses on ‘the’ or ‘that’ is striking – only three year old typically developing
children with much lower cognitive and language skills show this kind of performance. If this
result holds up with a larger group, it will be a very important result in the study of language
disability in autism.

Investigation of differential subject groupings revealed that while no binning is perfect,
binning by vocabulary or by nonverbal IQ does not do a good job of distinguishing the really
poor performing participants (null (D) pattern) from the mildly impaired (child-like (B) pattern)
from the nonimpaired (adult-like (A) performance). On the other hand, binning by clinical
diagnosis clearly distinguished the poorest performing children from the highest performing
participants. Binning by overall grammatical skills distinguished the poorest performing from
mildly delayed, and the poorest performing from the highest performing. Such distinctions are
key for determining which group of participants would require the most language assistance.

The binning-by-different-groups analysis also shows that the driving force in knowledge
of definite articles in ASD is not one’s vocabulary levels, and it is not one’s nonverbal IQ – it is
only one’s grammar levels, here as measured by TROG. Regardless of how the ASD participants
were grouped, TROG SS always came out as a significant covariate. It is worth it to restate this
clearly: knowledge of articles in ASD is best predicted by their standard scores on a standardized
test of comprehension of grammatical structures such as transitive and passive voice sentences,
relative clauses, sentences with negation, prepositions, personal pronouns – all aspects of
language that have to do with syntactic ability. TROG does not measure definite articles in any
way. The fact that overall grammatical development is a significant predictor of article
knowledge merits further investigation and suggests that knowledge of articles is mediated by the
computational language faculty, just as syntactic ability is.

In terms of theoretical significance for semantic and pragmatic theories, our results on the
patterns of article knowledge within ASD support the idea that acquisition of definite determiner
‘the’ is mediated by maturation of knowledge of semantic properties and not solely pragmatics.
We do not find children who do worse on ‘that’ than on ‘the’ within ASD (C pattern). But we
find some ASD children who perform like typically developing children - better on ‘that’ than on ‘the’ (B pattern) – indicating a deficit in uniqueness/Maximality, with a relative strength in knowledge of pragmatic features of ‘that’ (of familiarity or interpretation relative to unusual context). This is surprising given known ‘pragmatic’ deficits in ASD. It seems that despite their social difficulties, some of the ASD participants know the pragmatic feature “information known to both speaker and hearer” or ‘interpretation relative to nondefault context’, which is what helps them interpret ‘that’ correctly, while they are interpreting ‘the’ incorrectly due to lack of uniqueness/Maximality presupposition. Good performance on both ‘the’ and ‘that’ (A pattern) in some children is evidence of intact knowledge of uniqueness, but does not tell us whether the knowledge of pragmatic features in ‘that’ is intact, since in our task, uniqueness may be enough for comprehension of ‘that’. Knowledge of the indefinite article ‘a’ and of anaphor ‘another’ seems intact, as we are not seeing random performance which can result in 25% ‘same’ performance.

Perhaps some comments from two of our Asperger participants can elucidate what is going on. Following the task, we asked how they made their decisions. One male participant (18 years of age) said: “you said ‘the apple’, ‘the’ can refer to a previously described object, therefore Turtle goes to same location as Fish”… This sort of explicit reasoning occurred with every single one of the items in the task. This participant, when questioned about ‘the’ vs ‘that’ said: “’the’ can involve a set of objects; ‘that’ can involve only one.” It is possible to interpret this statement to mean that ‘the’ can refer to a singular or a plural item. Another male participant (16 years of age) said: “‘the’ can mean any and specific… ‘that’ is single specific thing”. Observe that the definition of ‘the’ of the latter is exactly what one would get if ‘the’ was missing uniqueness – as what we think is going on in TD children (Wexler 2003) – ‘the’ can refer to a unique specific thing, or it can refer to one of the things in the context set.

In further studies, our aim is to replicate the present findings in a larger set of participants with ASD. It is also important to study acquisition of anaphor ‘same’ – just like we did in Chapter 3.4 with Typically Developing participants – since correct performance of ‘same’ does not require uniqueness/maximality, but does require awareness of the salient entity in the context set. Additionally, investigating comprehension of free relative clauses in ASD, something that we argued in Chapter 4.2 requires only maximality and not salience, may be very informative. In the present sample, we cannot firmly answer the question of whether participants with autism have this awareness of the salient entity in the discourse set.

The observed performances reveal varying degrees of impaired, delayed, and normal semantic and pragmatic knowledge in our sample of children with ASD. No participants showed any atypical patterns of article comprehension, in fact patterns found in typical development were exaggerated in some children with PDD-NOS and with Asperger syndrome. Such distinct patterns of performance could strongly contribute to the phenotypic characterization of ASD. Indeed, perhaps knowledge of articles can act as an endophenotype, a simpler clue to the genetics underlying ASD, rather than, or in addition to, the three hallmark components of autism. A constrained test of semantics and pragmatics (of articles) may contribute to accurately defining the autism spectrum disorders endophenotypes allowing for a more precise genotypic characterization and diagnostics.
5.9. Acknowledgements

This research was supported by the Anne and Paul Marcus Family Foundation to the Brain Development and Disorders Project (BDDP) at MIT, and by the Singleton Fellowship to the author. Our heartfelt thanks go to all the participants and their families, Alexandra Perovic, Lee Mavros - the coordinator of BDDP, the audiences of the September 2006 Cog Lunch at the department of Brain and Cognitive Sciences, MIT and the June 2007 Symposium on Research in Child Language Disorders at Madison, Wisconsin, and the wonderful MIT and Wellesley undergraduates of the Wexler ab/Normal Language Lab: Tess (Tara) Diduch, Leah Bogsted, Nikki (Dolapo) Longe, Frances Choi, Elisabeth Lex, Robert Wells, Christopher Watson, Alexandra Huston-Carico, Katherine Boothe, and above all Margaret Echelbarger.

5.10. References


6. Williams Syndrome

In chapter 5, we showed that acquisition of determiners in children with autism spectrum disorders does not proceed a deviant pattern from typical acquisition – it is simply severely delayed, over and beyond what would be expected given the cognitive functioning levels of those children. In the first part of this chapter, we investigate acquisition of determiners in what sometimes has been called the mirror opposite of autism – Williams syndrome – whose sufferers are known for their hypersociability and linguistic fluency. Un/Fortunately first impressions are misleading, and often those with Williams syndrome perform only slightly better, but not too differently, than those with autism. In the second part of this chapter, we take advantage of the fact that the genetics underlying Williams syndrome are known, and investigate, using mouse models, whether two of those genes contribute to the abnormal brain development in Williams syndrome, which in turn could contribute to the language impairments experienced by those with Williams syndrome.

6.1 Comprehension of Determiners in Williams Syndrome

6.1.1 Abstract

We investigate the comprehension of definite (‘the’ and ‘that’) and indefinite (‘a’), articles, as well as anaphor ‘another’, in children and adolescents with Williams syndrome (WS). Interpretation of these elements, denoting salient and unique or simply existing referents, relies on the knowledge of both pragmatic and semantic aspects of language. There is a debate whether pragmatics is impaired in WS, and there is little known about the mastery of semantics, i.e. compositional meanings of sentences, in WS. We tested 6 children and adolescents with WS (Chronological Age: 6;5-20;7) and 18 typically developing controls (CA: 4;7-16;0), individually matched on non-verbal reasoning (KBIT), vocabulary (PPVT), or grammar (TROG). We find that participants with WS show a variety of knowledge patterns – with 50% of participants showing very poor knowledge on ‘the’ and ‘that’. One WS participant showed perfect knowledge of ‘the’ and ‘that’, and the remaining participants showed a pattern found in typical acquisition where ‘that’ is understood better than ‘that’. WS group performed most like its typically developing grammar controls of mean age 6 years. In comparison, children and adolescents with autism spectrum disorders (group matched to WS on age, IQ, vocabulary and grammar) all showed no knowledge of articles. These preliminary results suggest that Williams syndrome impacts the semantic and pragmatic language systems, but not as much as autism does.

6.1.2 Background on Williams Syndrome

Williams syndrome is caused by a deletion of 25 genes on chromosome 7q11.23. The incidence is 1 in 7,500 births (on the basis of FISH test). Physical characteristics include unusual facies -‘elfin faces’; hypercalcemia, heart and kidney abnormalities, hypersensitivity to sound, auditory and visual problems, growth retardation. Neuropsychologically, people with WS are hypersociable and easily approach strangers (Doyle, Bellugi, Korenberg, Graham 2004), yet they have Theory of Mind deficits and pragmatic deficits similar to those found in autism (Porter, Coltheart, Langdon 2007; Laws Bishop 2004; Philofsky, Fidler, Hepburn 2007). WS also have severe to mild cognitive impairment (mean IQ 55, range 40-90); deficits in visuospatial
perception, spatial memory and processing, but preserved face recognition, all unlike people with autism. Language abilities in WS were believed to be delayed but intact, or at least preserved relative to their cognitive deficits, however some recent studies (in Rice, Warren & Betz 2005, and others) show that WS have problems with complex aspects of grammar involving dislocation of sentential elements, e.g. passive sentences with psychological verbs and raising sentences with ‘seem’ that involve similar syntax (Perovic & Wexler 2006, 2007). From a syntactic point of view, even adult individuals with WS linguistically resemble typically developing 5-6 year old children. Other skills in adults with WS remain at 6-8 year age equivalent (Howlin, Davies, Udwin, 1998).

Given the unique profile of deficits and abilities in WS, it is important to study their knowledge of definite and indefinite articles. These small words are a very specific aspect of language that bridges semantics, pragmatics and maybe even theory of mind. Knowledge of semantics and pragmatics and TOM has been studied in WS, but never that of articles. Studying knowledge of articles in WS is important not only for the sake understanding WS further, but also for the sake of making sense of theoretical explanations of articles in typical development (TD).

6.1.3. Semantics and Pragmatics in Williams syndrome

Although definite and indefinite articles were not studied in detail in WS, other semantic and pragmatics aspects of language were investigated, and this literature can be used to formulate hypotheses regarding knowledge of articles in WS.

Note that by semantics here we do not mean vocabulary/lexicon levels, nor do we mean ability to classify words by concepts (colors, etc). There is profound, but often missed difference, between knowing the meaning of e.g. ‘dog’, and of knowing the meaning of ‘not’, ‘the’, ‘every’. The latter are both more conceptually abstract and more linguistic – these are functional closed class items which typical children often take time to acquire. These referential terms are used as pointers to referents in conversations, and their successful use requires not only pragmatic awareness, but also knowledge of their semantic meanings which are part of the computational system of language, and are distinct from conceptual meanings of open class words.

The definition of pragmatics also differs between studies of WS and studies of typical development. In case of WS studies – pragmatics is often about speech acts, i.e. using language to request and to question, and about discourse cohesion – sticking to the topic of a conversation, and about understanding non-literal language use in ironies and metaphors. In typical development, pragmatic studies address, among other things, more subtle aspects of children’s knowledge, e.g. children’s comprehension of scalar implicatures (logical entailments) – limiting or extending semantic meanings of words such as ‘some’ and ‘every’ based on the contexts.

6.1.3.1. Semantics in Williams Syndrome

Inevitably, we review lexical knowledge studies. WS experience a late lexical onset, attributed to a deficit in early phonological processing (Nazzi, Paterson & Karmiloff-Smith 2003). Production of words on a naming task in WS has been shown to be similar to that of TD children matched on mental age, but WS produced more gestures indicating word finding difficulties (Bello, Capirci, Volterra (2004)). People with WS were initially argued to have ‘bizarre’ semantics producing e.g. unusual animal names, however semantic priming – which indicates organization of mental dictionary and semantic memory – was shown to be normal in an online task in WS (Tyler, Karmiloff-Smith, Voice, Stevens, Grant, Udwin, Davies, Howlin
They however suggest that integration of semantic information may still be abnormal in sentence comprehension. Reading in WS is argued to impaired due to their weaker semantic knowledge (Laing, Hulme, Grant, Karmiloff-Smith, 2001). Few studies address specific semantic terms. Comprehension of spatial terms - prepositions ‘in’, ‘on’, ‘above’, and ‘below’ but not nonspatial terms (including negation ‘neither’...’nor’) was impaired in participants with WS, thus suggesting that their visuospatial difficulties also concern linguistic spatial terms (Phillips, Jarrold, Baddeley, Grant and Karmiloff-Smith (2004)). Laing and Jarrold (2007) find in two picture matching tasks that WS can perform on conditions that require a basis of existing semantic knowledge better than on conditions that require a mental representation of spatial relations. Pictures of animals were used, but names of animals were not – the animals were colored in different colors and were referred to based on their color. In the former condition animals on pictures were drawn the same size, e.g. penguins and giraffes, therefore WS had to rely on lexical/conceptual knowledge to answer which animal; in the latter condition animals were drawn either the in correct size relation or in the opposite (e.g. penguin bigger than giraffe), so participants had to suppress lexical knowledge and pay attention to spatial relation in pictures. Technically, WS and TD performed no better than chance at semantic conditions. WS performed better at spatial than semantic (slightly), but TD were way better at spatial. So the authors seem to be overinterpreting their results. WS were at chance on both conditions.

To our knowledge, only one study addressed knowledge of articles in WS. Elicited production of articles and nouns in French WS participants showed that they produce more errors than do TD children, with more errors of omission of articles (20%) and also more lack of gender concord between the article and the noun (50%) (as required in French) (Karmiloff-Smith, Grant, Berthoud, Davies, Howlin, Udwin 1997). This deficit in syntactic gender marking was attributed to WS’s morphosyntactic difficulties, and not due to a lack of knowledge of gender or articles independently. Finally, Mervis & John (2008) caution against generalizations by showing that depending how participants with WS are matched, the relative difficulties with spatial vocabulary disappeared. In summary, those with Williams syndrome may have semantic deficits, especially where semantics/lexical abilities intersect with morphosyntax, as in e.g. gender marking.

6.1.3.2. Pragmatics in Williams Syndrome

Those with Williams syndrome were thought to not have issues with social communication given their friendly personality (Bellugi, Lichtenberger et al 1999; Bellugi, Adolphs et al 1999). Rice et al in their 2005 review of language disorders say that pragmatics has not been studied extensively in WS. Since then Stojanovic (2006) showed that people with WS have a hard time interpreting information requests from their conversation partners and providing enough information of them. Thus people with WS do have difficulties with social conversational interaction. Laws and Bishop (2004) used Children’s Communication Checklist (a questionnaire that parents fill out) to show that participants with WS show some degree of pragmatic language impairment as well as more difficulties with social relationships and restricted interests, suggesting a parallel between WS and ASD. WS scored lower on inappropriate initiation of conversations, stereotyped conversations and use of contexts than participants with Down syndrome and language impairments; on discourse coherence (e.g. maintaining referents with in/definite articles) WS scored higher than language impaired groups but lower than TD controls. But in direct comparison, WS outscore ASD participants. Philofsky, Fidler, Hepburn (2007) used the same test and found that while both groups showed significant deficits relative to TD controls, WS performed better than ASD on pragmatic language, especially on discourse
coherence, stereotyped language (although not by a lot); nonverbal communication and social relations subscales showed ASD to have twice as many deficits as WS.

Comprehension of personal pronouns (which in TD is delayed due to children’s inability to set pragmatic constraints on the context of interpretation or due to children’s inability to consider multiple referential representations due to their limited working memory) in 50% of people with Williams syndrome is adult-like, and in 30% of WS is TD child-like suggesting a “delay in the acquisition of constraints regulating coreferential interpretation of pronouns.” (Perovic, Modyanova, Wexler, reanalysis; quote from Perovic, Wexler 2007).

Verbal working memory was argued to be a relative strength in WS, but unlike in TD, it correlates with grammatical abilities, suggesting that children with WS rely on working memory to learn grammar (Robinson, Mervis, Robinson 2003).

In summary, participants with WS may have pragmatic difficulties in maintaining referents using articles due to their issues with discourse cohesion.

6.1.3.3. Theory of Mind in Williams Syndrome

It has been shown that Theory of Mind is not a strong area in individuals with Williams syndrome. While 94% of WS pass first order false belief tasks, only 31% pass a harder 2nd order false belief task (usually passed by TD 7 year olds), but 88% pass an easier 2nd order task (that us usually passed by TD 5 years – with clearer and easier-to-remember stories); 50% of WS knew metaphor and sarcasm well (unlike ASD people who interpret these literally), and the rest found sarcastic statements easier than metaphors (i.e. socially constrained is easier than cognitively constrained) (Karmiloff-Smith, Klima, Bellugi, Grant, Baron-Cohen 1995). In contrast, other studies show deficits in first-order and second order theory of mind abilities in people with WS that are similar to those in ASD: only 24% of WS pass first order TOM, and only 45% pass second order TOM (Tager-Flusberg & Sullivan 2000; Sullivan & Tager-Flusberg 1999). Additionally, people with WS were unable to identify ironic jokes, instead judging them to be lies, just as younger typically developing children do (Sullivan, Winner & Tager-Flusberg 2003). A recent study (Skwerer, Verbalis, Schofield, Faja, Tager-Flusberg 2006) used a version of Eyes Test developed by Baron-Cohen for diagnosing of theory of mind deficits in those with Asperger syndrome and high-functioning autism (Baron-Cohen, Wheelwright, Hill, Raste & Plumb 2001) where participants have to correctly pick an adjective describing a gaze depicted in a picture of a person’s eyes. WS individuals are known to be good at face processing, however they showed worse performance than age matched normal controls (and similar performance to age-, IQ-, and language- matched people with learning and intellectual disabilities). These results clearly support WS’s deficits in interpretation of social communication. Santos and Deruelle (2009) suggest that WS’s TOM deficits (measure by attribution of intentions) are modality specific – with WS doing much better at verbal story tasks, and much worse at visual tasks (comic strips depicting scenarios).

6.1.3.4. Autism spectrum disorders vs (in) Williams syndrome

While on the surface ASD and WS seem mirror opposites, especially in visuospatial domain and social domain, more accurate studies suggest common social/theory of mind/pragmatic deficits. Klein-Tasman, Mervis, Lord and Phillips (2007) find that up to 50%-70% of participants with WS when tested on ADOS module 1 (autism diagnostic observation schedule for those with no or beginning language skills) score as being autistic or on the autism spectrum. WS are impaired in eye contact, joint attention, interacting normally with other people
it seemed that other human beings were their favorite toy, not another sentient human being
with independent wishes and thoughts.

Thus, we are seeing similar pragmatic difficulties in people with ASD and WS. Despite
their different communication styles, both groups experience difficulties with social
communication. It is essential then to characterize the precise nature of this deficit. Previous
studies often use communication checklists and general tests of language to access people’s
difficulties with semantics and pragmatics (e.g. Farmer and Oliver 2005). The results are often
telling, but specific study of certain aspects of language will provide a clearer picture.

6.1.4. Predictions

It could be argued people with WS and people with autism and PDD-NOS have similar,
poor, profiles when it comes to their linguistic, pragmatic and theory of mind abilities, although
WS fair slightly better than ASD. Thus it is possible, then, that knowledge of semantics and
pragmatics, as measured by a test of definite determiners, will be similar in ASD and WS, if they
are matched on their overall cognitive and linguistic functioning. Please see section 5.4 for
discussion of predictions for determiners in ASD which is also relevant here, since similar
predictions can be made for WS participants.

Relative to typically developing children, it is expected that WS participants could
perform below TD children that would be matched on to WS on cognitive levels and grammatical
levels, especially given Laws and Bishop (2004) findings on WS’s problems on discourse
cohesion.

WS participants also present a wonderful comparison group for elucidating the theoretical
accounts for definite articles. On one hand, WS are gregarious and hypersocial, so if it weren’t
for studies explicitly showing social, pragmatic and theory of mind deficits in WS, it would be
easy to suggest that WS would have no problems with articles for pragmatic reasons. It is also
possible that WS have semantic deficits, but based on existing lexical studies, it is hard to argue
for them. Since we are assuming that semantic knowledge of uniqueness/maximality is part of the
computational language faculty, then we expect some correlation between grammatical/syntactic
development and semantic development. WS participants are known to show a deficit in syntactic
movement (passive and raising comprehension, acquired by 7-8 year in TD (Perovic, Wexler,
2006, 2007), but no deficit in syntactic antecedents (reflexive pronouns comprehension, acquired
by 4-5 years in TD (Perovic, Wexler, 2007), suggesting that WS are (even when they are adults)
at a linguistic level of a typically developing 5-6 year old child. Based on the typical
development study, some 5-6 year olds would know definite and indefinite articles perfectly, and
some would show no knowledge, and yet others may lack semantic knowledge in ‘the’ but
overcome this lack in ‘that’ using its the extra saliency and requirements on context of
interpretation.

Given this reasoning, it is possible to make the following prediction. If deficient
pragmatics is what causes difficulties in comprehension of determiners in all populations,
children with WS should perform just like ASD children matched on age and cognitive abilities.
If deficient semantics is what causes difficulties in comprehension of determiners, participants
with WS should perform just like TD 5-6 year olds.
6.1.5. Participants

The WS subjects, with confirmed genetic diagnosis of WS, were recruited by Alexandra Perovic in 2005 with help of Williams Syndrome Association (WSA) in the USA. Two participants were tested in 2006, and four participants were recontacted in 2009 to ask for their participation in the present study.

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>KBIT SS</th>
<th>PPVT SS</th>
<th>TROG SS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WS</strong></td>
<td>mean</td>
<td>13.87</td>
<td>85.17</td>
<td>79.00</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>2.27</td>
<td>15.01</td>
<td>2.13</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>6.4</td>
<td>20.6</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>number</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td><strong>Autism/PDD-NOS (ASD)</strong></td>
<td>mean</td>
<td>11.49</td>
<td>88.00</td>
<td>77.83</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>1.76</td>
<td>15.56</td>
<td>13.19</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>6.5</td>
<td>17.2</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>number</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td><strong>KBIT controls</strong></td>
<td>mean</td>
<td>9.48</td>
<td>100.17</td>
<td>102.67</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>1.43</td>
<td>11.08</td>
<td>3.19</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>5.6</td>
<td>15.2</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>number</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td><strong>PPVT controls</strong></td>
<td>mean</td>
<td>9.65</td>
<td>109.67</td>
<td>98.33</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>1.71</td>
<td>15.01</td>
<td>1.45</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>4.6</td>
<td>16.0</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td>number</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td><strong>TROG controls</strong></td>
<td>mean</td>
<td>6.02</td>
<td>97.67</td>
<td>99.67</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>0.44</td>
<td>3.87</td>
<td>3.93</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>4.6</td>
<td>7.1</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>number</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 6.1.1. Details of WS participants and their TD and ASD controls

The ASD subjects for the present study were recruited in collaboration with Alexandra Perovic, through parent support groups and schools for children with disabilities. The clinical diagnosis was made by clinicians in the New England area and reported to us by parents. For the present comparison with Williams syndrome, only those with autism or PDD-NOS diagnosis were included, since those with Asperger are expected to significantly differ from WS participants on nonverbal reasoning and grammatical abilities, which may invalidate comparisons.

WS participants were individually matched to TD controls on three measures, thus forming three control groups. TD controls were recruited from Boston/Cambridge area daycares and afterschool programs. One group of TD controls was matched to probands on raw non-verbal cognitive abilities, as measured by Matrices subtest of KBIT-II (usually no more than one point off). This group allows us to factor out the influence of general cognition on participants' linguistic performance. The second control group was matched on raw scores of a standardized measure of receptive vocabulary (PPVT-III, maximum 4 points off), to tease apart the influence...
of general language abilities, i.e. lexical knowledge, on the more abstract aspects of linguistic knowledge of language assessed by our experimental task. The third control group was matched on raw score of standardized measure of grammar (TROG-2, usually no more than one point off), to establish whether there are differences in the subtle aspects of linguistic knowledge between control and disordered participants even if they are matched for the general level of grammatical competence. Every effort was made to match the participants on gender and to choose only those control participants whose scores on the standardized tests were close to the average (between 85 and 115). In this way the performance of control participants on the study task could not be interpreted as being due to their superior or inferior general cognitive or linguistic abilities.

WS participants were also group matched to a subset ASD participants from Chapter 5 (especially those with autism or with PDD-NOS) on age, raw and standard scores of the three standardized tests. Herein, this subgroup will be referred to ASD.

The testing procedure was the same as for ASD participants in Chapter 5.6.

6.1.6. Results

We find that knowledge of indefinite articles ‘a’ and ‘another’ does not differ among participant groups, however knowledge of ‘the’ and ‘that’ does so.

MANOVA (number ‘same’ responses on ‘a’, ‘another’, ‘the’, and ‘that’ as dependent variables, 5 groups (WS, ASD, and three WS control groups) as between subject factors) showed significant effect of group for ‘the’ (F(4, 25)=3.72, p=.017) and ‘that’ (F(4, 25)=3.95, p=.013), but not for ‘a’ and ‘another’. Fisher’s LSD posthoc tests (selected to bring out the subtle effects
in the small number of subjects) showed differences on ‘the’ approaching significance between ASD and WS group (p=.086), and significant differences on ‘the’ between ASD group and WS KBIT and PPVT control groups (p=.004 and p=.003 respectively), but WS group did not differ significantly from its controls; significant differences on ‘that’ were observed between WS and ASD groups (p=.015), and between ASD and all WS controls (p=.001 for KBIT controls, p=.004 for PPVT controls, p=.04 for TROG controls). R² for the model is .373 (adjusted .27) for ‘the’ and .387 (adjusted .289) for ‘that’.

We now turn the focus specifically on the definite determiners, since this is where variation lies. Participants were subdivided by their pattern of performance between ‘the’ and ‘that’, with “A” pattern denoting good performance on both ‘the’ and ‘that’, “B” pattern denoting bad performance on ‘the’ with a relatively better performance on ‘that’, “C” pattern denoting a bad performance on ‘that’ with a relatively better performance on ‘the’, and finally “D” pattern denoting a bad performance in both ‘the’ and ‘that’. ‘Good’ performance was considered if participants got 4/6 or greater ‘same’ actions on the same object. ‘Bad’ performance was 3/6 or less items. For ‘B’ and ‘C’ patterns participants had to get a difference of 2 ‘same’ actions between ‘the’ and ‘that’ for the difference to count. Thus difference between 3 and 4 does not count, but a difference between 3 and 5 does. This is was done to figure out whether participants are showing adult-like knowledge (A pattern), showing deficits especially in semantics (but not pragmatics) (B pattern, found in TD children), showing deficits especially in pragmatics (but not in semantics) (C pattern, not found in TD children), or showing deficits across the board (D pattern).

![Figure and Table 6.1.3. Proportion of participants showing patterns of performance](image)

<table>
<thead>
<tr>
<th></th>
<th>WS</th>
<th>ASD</th>
<th>KBIT control</th>
<th>PPVT control</th>
<th>TROG control</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>D</strong></td>
<td>0.50</td>
<td>1.00</td>
<td>0.00</td>
<td>0.17</td>
<td>0.50</td>
</tr>
<tr>
<td><strong>B</strong></td>
<td>0.33</td>
<td>0.00</td>
<td>0.33</td>
<td>0.00</td>
<td>0.17</td>
</tr>
<tr>
<td><strong>A</strong></td>
<td>0.17</td>
<td>0.00</td>
<td>0.67</td>
<td>0.83</td>
<td>0.33</td>
</tr>
</tbody>
</table>

The proportion of participants showing these patterns is illustrated. It becomes immediately clear that WS participants are not showing a random performance – their 3/6 average responses for ‘the’ are not equally distributed across participants. As a group, 50% of participants with WS show no knowledge of articles (pattern “D”), with only 15% (one subject) showing good knowledge (“A” pattern), and the remaining 35% (two subjects) showing the “B”
pattern where they do better on ‘that’ than on ‘the’. Only the TROG control group has similar
distribution of participants showing that many poor responses. The TROG control group (which
is the youngest) is showing 35% “A” pattern. In the KBIT and PPVT control groups 70-80% of
participants are showing good proficiency with ‘the’ and ‘that’. Observe that not a single
participant shows the “C” pattern. The pattern that ASD group shows contrasts with WS and all
controls – not a single participant with ASD, when matched on age, nonverbal reasoning,
vocabulary and grammar to WS participants, shows knowledge of articles.

ANOVA (between group factor: WS vs ASD vs three control groups, dependent variable
= proportion participants showing a particular pattern) showed significant effect of group
(F(4,25)=6.1, p=.001), with Fisher’s LSD post hoc tests showing WS approaching significant
difference from the ASD group (p=.072), and WS differing from KBIT group and PPVT group
(p=.024 and .042 respectively), with R2 for the model .495 (adjusted .415).

This analysis illustrates two of the main points of the data. Participants with WS are more
deficient in their knowledge of articles than expected given their nonverbal reasoning abilities,
and more deficient than expected given their vocabulary levels. However, participants with WS
are performing at the levels expected by their grammar levels – levels shown by typically
developing five year olds. Participants with WS outperform only one comparison group – those
with autism spectrum disorders. Even though WS and ASD participants are at similar ages and
have similar nonverbal reasoning, vocabulary and grammar levels, half of WS show no
knowledge whereas ALL of ASD so no knowledge.

<table>
<thead>
<tr>
<th>AGE</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘A’</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>‘ANOTHER’</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>‘THE’</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>‘THAT’</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>‘SHARE’</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>‘SAME’</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>KBIT RAW</td>
<td>15</td>
<td>20</td>
<td>19</td>
<td>26</td>
<td>30</td>
<td>36</td>
</tr>
<tr>
<td>KBIT SS</td>
<td>90</td>
<td>82</td>
<td>75</td>
<td>77</td>
<td>86</td>
<td>101</td>
</tr>
<tr>
<td>PPVT RAW</td>
<td>59</td>
<td>92</td>
<td>95</td>
<td>129</td>
<td>148</td>
<td>153</td>
</tr>
<tr>
<td>PPVT SS</td>
<td>78</td>
<td>74</td>
<td>73</td>
<td>79</td>
<td>84</td>
<td>86</td>
</tr>
<tr>
<td>PPVT AE</td>
<td>4.58</td>
<td>6.92</td>
<td>7.20</td>
<td>10.08</td>
<td>12.67</td>
<td>13.67</td>
</tr>
<tr>
<td>TROG RAW</td>
<td>4</td>
<td>7</td>
<td>5</td>
<td>12</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>TROG SS</td>
<td>69</td>
<td>55</td>
<td>55</td>
<td>71</td>
<td>76</td>
<td>81</td>
</tr>
<tr>
<td>TROG AE</td>
<td>4</td>
<td>4.92</td>
<td>4.42</td>
<td>6.5</td>
<td>7</td>
<td>7.92</td>
</tr>
</tbody>
</table>

Table 6.1.4. Details of performance and functioning of individual WS participants

Since only six WS participants are presently included, it is worth to have a closer look at
their individual performances. Values for ‘a’, ‘another’, ‘the’, and ‘that’ show actions on the
‘sam’ objects out of 6. Values for ‘same’ and ‘share’ show the same out of 4, and only four WS
participants were tested on these. Three WS participants (2,5,6) showed perfect knowledge of
‘sam’ and ‘share’, while at the same time showing ‘B’ and ‘D’ patterns in determiner
knowledge, and one participant (3) understood ‘share’ but not at all ‘same’ and showed ‘D’ pattern. So far it seems that knowledge of WS grows with age – those younger than 11 years showing poor knowledge, and those 16 and older are at least showing partial knowledge. It is an interesting question whether knowledge of articles matures in WS, and whether the majority of WS attain knowledge of articles by adulthood. Of course, we observe a delay in article knowledge relative to chronological age, but relative to grammatical age equivalents, WS participants are exactly where they should be.

6.1.7. Discussion

In summary, we find that children and adolescents with Williams syndrome show multiple levels of knowledge of definite and indefinite articles, with the youngest WS participants showing no knowledge (D pattern), and the oldest WS participants showing either adult-like (A pattern) or child-like knowledge (B pattern), where ‘that’ is known better than ‘the’.

Relative to their TD controls, WS perform worse than expected given their nonverbal-reasoning and vocabulary levels, but WS perform similarly to their grammar controls who are on average 6 years old. This finding suggests that knowledge of articles in WS tracks more closely with their overall grammatical development which is known to be delayed (Perovic, Wexler, 2006, 2007), and does not track with their vocabulary levels which are known to be a relative strength in WS.

Relative to participants with ASD matched on age, nonverbal, verbal and grammatical levels, WS perform much much better. Where all ASD participants show poor knowledge of articles, only 50% of WS participants do so.

So far, the results are consistent with articles being mediated by the computational semantic system of language and not solely pragmatics. This is because WS and ASD participants show widely different performances, and if their deficits were in pragmatics, and not in semantics, it could be expected they would show similar performance. This is also because some WS show better performance on ‘that’ than on ‘the’ (a child-like pattern of performance), suggesting they are sensitive to the saliency of referent with ‘that’ which involves additional pragmatic features (of familiarity or interpretation relative to nondefault context situation), but their knowledge of ‘the’ is deficient – suggesting a deficit in the semantic property of uniqueness/Maximality. It is notable that 50% of WS show intact performance on article ‘that’ despite their known difficulties in pragmatic communication. Knowledge of the indefinite article ‘a’, and anaphor ‘another’ seems intact, as we are not seeing random performance which can result in at least 30% performance.

The observed performances reveal varying degrees of impaired, delayed, and normal semantic and pragmatic knowledge in our sample of children with WS. No participants showed any atypical patterns of article comprehension, in fact patterns found in typical development were exaggerated in some children with WS. In further investigations our goal is to increase the present sample size. It would also be amazing to investigate whether knowledge of articles in WS tracks with any atypical deletions, in a similar way that visuospatial and social behaviors do.

6.1.8. Acknowledgements

Our heartfelt thanks go to all the participants and their families, Alexandra Perovic, and the wonderful undergraduates of the Wexler ab/Normal Language Lab who made this WS study possible: Margaret Echelbarger, Aimi Watanabe, Ariel Glover, Ashli Davis, Jessica Wooton, Kenny Aronson.
6.1.9. References


6.2. Initial characterization of GTF2I family of gene homologues in mouse neocortical development

6.2.1. Abstract

People with Williams syndrome have a profound disorganization of grey and white matter, with region specific thickening and thickening of the cortex which correlates with their cognitive strengths and weaknesses. This indicates that some genes in Williams syndrome region contribute to abnormalities in brain development and hence can contribute to abnormalities in laminar patterning, which can be directly investigated in mouse models using the method of in utero electroporation. We investigated effects of overexpression of mouse homologues of Williams syndrome region genes GTF2I and GTF2IRD1, which are known to contribute to the visuospatial, social and perhaps even linguistic deficits in the disorder, in mouse embryonic neocortical development using in utero electroporation. We find that individually Gtf2i and Gtf2ird1 do not significantly affect laminar patterning, but when overexpressed together, they act synergistically to cause a severe depletion of cells in the cortical plate phenotype. Primary neurite length at the end of embryogenesis was found increased relative to Gfp control in both Gtf2i and Gtf2ird1, but was decreased relative to Gfp in double overexpression. The abnormalities in orientations of primary neurites were increased in Gtf2i and Gtf2ird1, and in the double overexpression. These results indicate that GTF2I and GTI2IRD1 act synergistically to contribute to cortical abnormalities in Williams syndrome and Williams syndrome duplication phenotypes, which may underlie their cognitive and behavioral deficits.

6.2.2. Background on the Williams syndrome region genotypes and neurological phenotypes

Williams syndrome results from a microdeletion on chromosome 7q11.23 (referred to herein as WS region) resulting in haploinsufficiency in about 25 genes (Ewart et al 1993). The incidence is 1 in 7,500 births (on the basis of FISH test). Physical characteristics include unusual facies -‘elfin faces’; hypercalcemia, heart and kidney abnormalities, hypersensitivity to sound, auditory and visual problems, growth retardation. Neuropsychologically, people with WS are hypersociable and easily approach strangers (Doyle, Bellugi, Korenberg, Graham 2004), yet they have Theory of Mind deficits and pragmatic deficits similar to those found in autism (e.g. Laws Bishop 2004). WS also have severe to mild cognitive impairment (mean IQ 55, range 40-90); deficits in visuospatial perception, spatial memory and processing, but preserved face recognition, all unlike people with autism. Language abilities in WS were believed to be delayed but intact, or at least preserved relative to their cognitive deficits, however recent studies show that WS have problems with complex aspects of grammar involving dislocation of sentential elements, e.g. passive sentences with psychological verbs and raising sentences with ‘seem’ that involve similar syntax (Perovic & Wexler, in press, 2007). From a syntactic point of view, even adult individuals with WS linguistically resemble typically developing 5-6 year old children. Other skills in adults with WS also remain at 6-8 year age equivalent (Howlin, Davies, Udwin, 1998).

Several human (and mouse) imaging studies suggest profound disorganization of grey and white matter in WS as contributing to the WS phenotype. In addition to global reductions in
white matter (up to 18%) and in grey matter (6%) (e.g. Thompson et al 2005), there is region specific thickening and thickening of the cortex and corresponding abnormalities in the gyral patterns. There are increased numbers of thinner gyri (Thompson et al 2005, Gaser et al 2006) and reduced (Kippenhan et al 2005) and abnormal (van Essen et al 2006) sulcal depths in both hemispheres, with reductions in grey matter correlating with reductions in sulcal depth. Imaging studies have shown alterations in brains of WS that are consistent with their relative strengths and weaknesses. Reiss et al (2004) find overall WS brains are smaller than controls, but “after adjusting for overall brain volume, find that participants with WS showed reduced thalamic and occipital lobe gray matter volumes and reduced gray matter density in subcortical and cortical regions comprising the human visual-spatial system compared with controls [also found by Boddaert et al 2006 in children with WS; Campbell et al 2009; by Chiang et al 2007 correlating positively with performance IQ]. The WS group also showed disproportionate increases in volume and gray matter density in several areas known to participate in emotion and face processing, including the amygdala, orbital and medial prefrontal cortices, anterior cingulate, insular cortex, and superior temporal gyrus” (also found by Campbell et al 2009). One postmortem study (Galaburda et al 2002) found increased cell density in left layer four, and increased numbers of small neurons in layers four, five and six in the visual cortex, supporting the neuroimaging findings. Another postmortem study found abnormalities in the auditory cortex: WS had more larger neurons bilaterally, and showed a lack of asymmetry in cell packing density found in controls (higher on the left than on the right) (Holinger et al 2005). Notably, planum temporale (surrounding Sylvian fissure and involved in language) showed reversed asymmetry in WS, with larger planum temporale on right hemisphere, which itself was due to abnormal horizontal Sylvian fissure which did not ascend into the parietal lobe – a pattern usually found in the left hemisphere of typical people (Eckert et al 2006). The same area additionally shows significant grey matter thickening (by 5-10%) (Thompson et al 2005). Of the studies focusing on white matter abnormalities in WS, Marenco et al (2007) used “diffusion tensor imaging to demonstrate alteration in white matter fiber directionality, deviation in posterior fiber tract course, and reduced lateralization of fiber coherence in WS” and suggest that “abnormalities are consistent with an alteration of the late stages of neuronal migration”. In many cases, WS had not only thicker white matter tracts, but also multiple ‘streams’ of such projections.

The WS neural networks are disorganized at a fundamental level as shown by a study of auditory perception in WS via fMRI (Levitin et al 2003) where auditory regions (temporal lobe) had reduced activation (relative typical controls) and greater amygdalar activation. Additionally, the activation was more widely distributed across cortical and subcortical structures during music processing. Thus the fact that people with WS have relatively preserved music abilities and hypersensitivity to sound does not mean that brain areas underlying these abilities are spared — in fact the reverse is true. A similar picture emerges in other domains. During a global processing task, WS showed reduced parietal and visual cortical activation, normal activation in ventral occipitotemporal cortex, but elevated activation of the thalamic nuclei (Mobbs et al 2007a). In a response inhibition task, while behaviorally similar to controls, WS showed reduced activation of critical cortical and subcortical structures involved in behavioral inhibition (such striatum, dorsolateral prefrontal and dorsal anterior cingulated cortices) (Mobbs et al 2007b). Hoeft et al (2007) used diffusion tensor imaging to show that fractional anisotropy was higher in the right superior longitudinal fasciculus – associating between deficits in visuospatial construction and higher FA in WS. Similarly, Meyer-Lindenberg et al (2004) using fMRI showed isolated
hypoactivation in WS in the parietal portion of the dorsal stream which correlated with gray matter volume reductions.

Despite these brain abnormalities in WS, of the genes involved in WS none have been investigated for their direct effects on embryonic brain development.

Discoveries of atypical deletions of the WS locus contributed to understanding of some genes via genotype-phenotype correlations. Notably, atypical deletions involving (or missing) the GTF2I family of genes (GTF2IRD1 & GTF2I) suggested that those are the genes strongly contributing to the hyposocial, the visuospatial and even maybe the linguistic deficits in WS. These atypical deletions spare just the GTF2I family of genes, while deleting all other WS genes. Although the details vary across several studies involving from all over the world subjects, all these patients have improved visuospatial abilities, milder facial abnormalities, milder to no mental retardation, and close to normal social skills (without overfriendliness or anxiety) relative to ‘typical’ WS. Language was addressed in only four of the studies, and in two of them normal language skills were found, while other two showed mild delay in language acquisition (both vocabulary levels and syntactic comprehension), similar to typical WS. (van Hagen et al 2007, Howald et al 2006, Morris et al 2003, Hirota et al 2003, Doyle et al 2004, Tassabehji et al 2005, Gagliardi et al 2003, Blyth et al 2008, Ferrero et al 2009). Dai et al (2009) find a unique individual with atypical WS deletion that includes GTF2IRD1, but not GTF2I, who has physical and visuospatial WS features but does not have language delay and has no overt abnormal social behavior phenotype. These data are consistent with the possibility that GTF2IRD1 is associated with WS facial dismorphology and visuospatial deficits, and that GTF2I is associated with the hypersociability and linguistic deficits.

Another study of an atypical deletion (Edelmann et al 2007), this one involving only the GTF2I genes and those more distal (but excluding majority of WS genes), found autism and Williams syndrome profile (including impaired visuospatial ability, an overly friendly personality, excessive non-social anxiety and language delay) in the patient, suggesting that GTF2I genes can, essentially by themselves, cause most features of WS. That is not the first study of people having both autism features and Williams syndrome (Reiss, Feinstein et al 1985; Herguner, Mukaddes 2006). Notably, Schubert & Laccone (2006) find a patient with a deletion spanning entire WS region and extending by a few genes on each side, who had severe WS phenotype, with motor and equilibrium disturbances as well as lack of speech/language abilities and autistic behavior. These studies suggest that genes bordering the WS region likely contribute to ASD.

The recently discovered duplication of the WS region shows a different phenotype than WS deletion. There are no visuospatial deficits, no phobias, no heart abnormalities, but there is a severe speech and language delay, affecting both production and comprehension of language, and gait and motor problems – both uncharacteristic of WS deletion. Unfortunately, detailed linguistic description of WS duplication is unavailable, but language difficulties experienced by these individuals are orders of magnitude worse than language impairments found on WS deletion (Mervis, p.c.). Social (autistic) deficits and some mental retardation appear in some, but not all cases of WS duplication. (Sommerville et al 2005, Depienne et al 2007, Torniero et al 2007, 2008, Berg et al 2007). GTF2I was among the genes whose expression is increased in the duplication (Sommerville et al 2005).

Brain imaging of one case of WS duplication (in Torniero et al 2008) showed increased cortical thickness and simplified gyral pattern in parietal lobes, increased cortical thickness in insular cortex, and some ectopic gray matter formations in subcortical white matter extending
from ventricles to beneath left parietal lobe, which the authors suggest represent columns of radial migration abnormality. Another case (Torniero et al 2007) showed cortical thickening and simplified gyri in left sylvian and temporal cortex (as in typical WS, consistent with the severe language phenotype); ectopic formations were not observed.

In summary, we conclude that language development (although likely not in same quality) is something that is consistently affected in both WS duplications and WS deletions (as do Osborne, Mervis 2007), suggesting that, perhaps, GTF2I is one of the dosage-sensitive gene contributing to the cortical abnormalities underlying linguistic deficits. Torniero et al (2007) reach a similar conclusion but implicate other genes: “Two such genes, linked to microtubule-associated proteins, are present in the duplicated region: LIMK1, encoding a protein kinase which regulates actin filament dynamics, and CYLN2, encoding a protein presumably linking microtubules to dendritic lamellar body.”

6.2.3. GTF2I family of genes: GTF2I and GTF2IRD1

In mouse embryonic development, both GTF2I and GTF2IRD1 are widely expressed. In adult mice (and humans), GTF2I is present exclusively in neurons showing especially high levels in the hippocampus, Purkinje cells, and dentate gyrus, but GTF2IRD1 is expressed throughout the body in adults, in brain in olfactory bulb, purkinje cells and neurons in piriform cortex. (Danoff et al 2004, Bayarsaihan et al 2003, Enkhmandakh et al 2004).

Animal models of Gtf2i and Gtf2ird1 were found to have particular phenotypes, replicating some features of Williams syndrome.

Animal models of Gtf2ird1 knockout in mice, in one case showed growth retardation and craniofacial abnormalities similar to those in WS (Tassabehji et al 2005). In another case (Young et al 2008), which did not replicate the craniofacial dismorphology, Gtf2ird1-targeted mice showed reduced fear, anxiety and aggression (like in WS), but with spared spatial memory, slightly impaired auditory-cued fear memory, and normal synaptic activity in the hippocampus. Additionally there was an altered serotonin metabolism, with alteration of postsynaptic serotonin turnover without increase in overall serotonin production – serotonin metabolite 5-HIAA (5-hydroxyindole acetic acid) was found to be elevated in amygdala, frontal cortex and parietal cortex. Van Hagen et al (2007) find that Gtf2ird1 heterozygote mice have significantly enlarged ventricles, mildly enlarged corpus callosum, freeze more than WT in fear contexts, and spend almost twice as much time on accelerating rotorod than WT. Staining brains with “hematoxylin/eosin did not reveal any gross abnormalities in the distributions of neurons, neuropil, glia or white matter. In addition, the cyto-architectonic layers in both the hippocampus and cerebellum appeared unaffected.”

Animal model of Gtf2i is so far only grossly characterized: microarray data from Gtf2ird1 and Gtf2i homozygotes, who showed embryonic lethality, brain hemorrhage, and vasculogeneic, craniofacial, and neural tube defects, suggested impairment of genes involved in the TGFbetaRII/Alk1/Smad5 signal transduction pathway, although the effect on these genes by Gtf2i knockout was less pronounced; heterozygotes of each genotype selectively showed microcephaly as well as craniofacial and skeletal abnormalities, and retarded growth – all features found in WS (Enkhmandakh et al 2008).

What are the molecular mechanisms that can underlie these phenotypes? Many findings, mostly from the cancer literature, suggest involvement of GTF2I in regulation of cell cycle. GTF2I family of genes, including GTF2I and GTF2IRD1, have been under scrutiny not only for their involvement in Williams syndrome, but also in cancer, in cell cycle regulation and
in transcriptional regulation of growth-promoting genes and signal transduction. Upon modulation by extracellular signals, GTF2I protein (TFII-I) regulates activation of the c-fos promoter (Kim, Cochran, 2000), which is part of the early-immediate response gene cascade involved in, among other things, learning of bird song (e.g. Kimpo et al 1997). TFII-I is a unique transcription factor because it provides a link between the basal transcriptional machinery and specific transcription factors (Roy, 2001) (quote from Stasyk et al 2005). Involvement of TFII-I in cell cycle makes it a great candidate for investigation in neuronal migration. Desgranges et al (2005) show that usually TFII-I activates transcription of cyclin D1 by binding to its promoter. Upon cell cycle arrest (from genotoxic stress and p53 activation), TFII-I is ubiquitinated and degraded, however overexpression TFII-I increases cyclin D1 levels, and causes accelerated entry to and exit from S phase, and overcomes p53-mediated cell cycle arrest.

The GTF2I family of genes shows mutually antagonizing effect – usually both are found in the nucleus, but their mutual overexpression results in all GTF2IRD1 locating to the nucleus and in all TFII-I locating outside of the nucleus, preventing the latter’s transcriptional activation of c-fos promoter (Tussie-Luna et al 2001). The GTF2I family of genes also modulates regulation of the transforming growth factor beta (TGF-beta)/activin target gene goosecoid (Gsc) - suggesting they are involved in TGF-beta signaling (Ku et al 2005). Downregulating TFII-I abolished TGF-beta mediated induction of Goosecoid, as does overexpressing GTF2IRD1. Ku et al suggest GTF2IRD1 protein (also known as BEN) plays a repressor or a baseline regulator role (it is constitutively recruited to distal promoters of Goosecoid), whereas TFII-I in complex with SMAD2 plays an upregulating role – upon effect of TGF-beta signaling, TFII-I displaces BEN and activates transcription. (SMAD2 has the ability to interact with gene promoters via DNA-binding proteins; SMAD3 (which in Ku et al did not bind with TFII-I) can bind to DNA directly (reviewed by Stasyk et al 2005). Stasyk et al (2005) however find that when TGF-beta1-dependent phosphorylation of TFII-I is cancelled, SMAD3 cooperates with TFII-I to increase expression of cyclin D2 and D3.

![Figure 6.2.1](image)

**Figure 6.2.1. Reproduced from Ku et al (2005)**

“Proposed model for TGF-beta signal-dependent transcriptional regulation of Gsc at a steady state. DE is occupied by BEN, which suppresses the transcription of Gsc possibly by recruitment of HDAC3. Upon TGF-beta/activin stimulation, Smad2 translocates to the nucleus, interacts with TFII-I, displaces BEN, and activates Gsc transcription.”

140
Upstream regulators and Downstream targets of the GTF2I family of genes

It is notable that through their downstream and upstream targets, the GTF2I family of genes connects on a molecular level with other neurodevelopmental disorders.

Upstream regulators of GTF2I family of genes include TGF-beta signaling, which plays a role in proliferation and differentiation and inflammation, and which has been implicated in autism. Serum levels in patients with autism had significantly decreased levels of TGF-beta1 (Okada, Hashimoto et al 2007), but brain levels of TGF-beta1 were elevated (Cohly, Panja 2005). MADH9, a member of the SMAD family of proteins that mediate the TGF-beta signaling pathway which is expressed in the brain, was implicated in a patient with autism and auditory processing deficits (Smith, Woodroffe, Smith, Holguin et al 2002). The genetic link between autism and Williams syndrome is not limited to the TGF-beta signaling pathway. Other genes within Williams syndrome region overlap with gene pathways involved in autism. For example, FZD9, a gene deleted in WS, may act as receptor for WNT signaling proteins involved in tissue polarity and development (in Tassabehji 2003). WNT2, involved in development of central nervous system, was implicated in two families with parent-child autism and severe language abnormalities (Wassink et al 2001).

The downstream targets of the GTF2I family of genes have been investigated via molecular and bioinformatics analyses, which showed their involvement in immunity response, cell cycle, catalytic activity, signaling pathways and transcriptional regulation (Chimge et al 2007a, Chimge et al 2007b, Chimge et al 2008). Notably, Chimge et al (2008) identify, among others, Tgfb2 as a potential downstream target of Gtf2ird1. Most recently, TFII-I, in complex with PARP1 and SFPQ proteins1, was found to regulate the DYX1C1 gene (Tapia-Páez et al 2008), which is a candidate gene for dyslexia (e.g. Taipale et al 2003), a neurodevelopmental disorder where children are unable to learn to read despite having good linguistic skills; the problem seems to be due to abnormal phonological processing implicating auditory brain areas. DYX1C1 has been implicated in abnormal laminar patterning (neuronal migration) in rat neocortex (Wang et al 2006) and in auditory processing and spatial learning in rats (e.g. Threlkeld et al 2007). Additionally, downregulation of DYX1C1 causes heterotopias (pockets of unmigrated neurons along lateral ventricles – interfering with white matter) in the cortex and the hippocampus (Threlkeld et al 2007, Rosen et al 2007). These neocortical heterotopias (or ectopias) have been found in brains of people with dyslexia (Humphreys, Kaufmann, Galaburda, 1990), Williams syndrome (Ferland et al 2006), duplication of Williams syndrome region (Torniero et al 2008), autism (Bauman & Kemper, 2005:130-131), language impairment (Preis et al 1998) and fragile X syndrome (Parrini et al 2006). Notably, neocortical ectopias/heterotopias are associated with attenuated neurophysiological responses to rapidly changing auditory stimuli in mice - a key phenotype in dyslexia (Frenkel et al 2000). This molecular connection between WS and dyslexia is quite intriguing. People with WS are known to be susceptible to dyslexia, and have auditory processing difficulties (e.g. Temple 2003, Levitin et al 2003). That DYX1C1 is a downstream target of TFII-I suggests a common molecular mechanism across disorders. Another interesting downstream target of Gtf2ird1 (as noted by Tassabehji et al 2005) is Goosecoid.

1 PARP1 is intracellular protease with a role in development of neuronal cell death, apoptosis, necrosis; Activated in traumatic brain injury, Alzheimer’s, Parkinson’s; involved in learning and long-term memory in rats (dyslexics have working memory problems!). SFPQ is involved in neuronal differentiation and maturation, lies in a region on chromosome 1p34-p36 linked to dyslexia and SLI.
which in humans is on 22q11 - DiGeorge/Velocardiofacial syndrome region which is associated with autism and visuospatial construction deficit.

6.2.4. Aim

No studies address the effect of the *GTF2I* genes on neocortical development, and the goal here is to characterize the consequences of modulation of these genes on laminar patterning, growth and cell morphogenesis. Laminar patterning can be modulated by altering cell cycle (in which TFII-I is clearly involved), e.g. cell output at the ventricular zone, or by altering neuronal migration directly (which is not known but possible for TFII-I, e.g. in a manner similar to Reelin (another autism candidate gene) (Persico et al 2001). Another goal is to further evaluate the significance of the interaction of TFII-I with TGF-beta pathway. Given findings that TFII-I is phosphorylated upon TGF-beta activation to activate transcription, we investigated the effects of inhibiting its activation by TGF-beta with SMAD7. SMAD7 is a potent inhibitor of TGF-beta activity that acts by binding to TGF-beta receptors and preventing phosphorylation of SMAD3 and SMAD2 (which themselves interact with TFII-I) and is a mediator of its proinflammatory signalling (Yan, Liu, Chen 2009; Fraser, 2007).

6.2.5. Methods

We investigated the effects of overexpression of Gtf2i, Gtf2ird1, as well as co-overexpression of Gtf2i and Gtf2ird1. We also investigated the effects of co-overexpression these genes with Smad7.

Using the *in utero* electroporation technique which allows specific targeting of DNA misexpression and maintaining of the normal course of development, and which targets specifically migrating neuroblasts of pyramidal cells (Tabata, Nakajima, 2001; Saito, Nakatsuji, 2001), these target genes (along with green fluorescent protein gene to clearly visualize the resulting changes) were misexpressed in the developing mouse neocortex at E14.5.

The genes were individually cloned into pEF6 plasmid. Plasmid DNA was purified with a mixed protocol (designed by Damon Page) combining QIAGEN EndoFree plasmid kit and QIAGEN HiSpeed Plasmid kit (to attain maximal possible yields while making sure to remove all endotoxins), dissolved in 10% TE buffer in at concentration of 2 µg/µl. For injection, total plasmid concentration was 1µg/µl with 1% Fast Green solution for visualization of injections. Time-pregnant mice were anesthetized with Isoflurane, the abdomen was opened and the uterine horns were exposed. Approximately 1-1.5 µl of plasmid soluion was injected via picospritzer into the lateral ventricle with glass micropipette. The embryo in the uterus was placed between the tweezers-type electrodes, which has disc electrodes of 5-7 mm in diameter at the tip. Five electronic pulses (25-55V, 50ms) were delivered with an electroporater. The wall and the skin of the abdominal cavity were sutured, and the embryos were allowed to develop normally.

The embryos were extracted at E18.5 (birth), immersion fixed in 4% paraformaldehyde overnight, subsequently extracted, postfixed for 4-6 hours, cryoprotected in 20% sucrose overnight, embedded in OCT compound, and sectioned into 50 µm slices in cryostat, and mounted on slides, and after removing the OCT with phosphate-buffered saline (PBS), were coverslipped for viewing under Zeiss confocal microscope (Laser Scanning Microscope LSM 5 Pascal). Images were taken with a 10x objective, at 1024-1024 resolution, with no digital zoom. The resulting images were normalized in Photoshop and analyzed in ImageJ to measure the distribution of transfected GFP signal across the migratory paths from the ventricular zone to the top levels of the cortex. The GFP signal was divided into 10 bins for initial analysis. Bins 1-3
(corresponding to the signal from the ventricular zone (VZ)), bins 4-6 (corresponding to the signal from the intermediate zone (IZ)), and bins 7-10 (corresponding to the signal from the cortical plate (CP)) were later summed for final statistical analysis (in SPSS).

A total of 81 pregnant Black Swiss mice underwent IUEP procedure, with surgeries on 26 mice (about 223 embryos) resulting in no surviving embryos (due to issues with electrodes, inconsistencies in plasmid preparation, and other experimental issues), and with surgeries on 55 mice (521 embryos) resulting in 246 surviving embryos (47% survival rate), of which 131 ended up being transfected (53% transfection rate). Analyzable transfections (clear lateral migrations showing cortical plate through ventricular zone) formed a minor subset of the total, with many transfections labeling thalamic cells, midline cells, or weakly labeling lateral cells.

Laminar migration data presented herein are based on the following (images = adjacent sections; technical replicates = nonoverlapping regions of interest defined by a box of 150 pixels in width across cortex by between 600-800 pixels and covering the area from the ventricular zone to the pial surface): GFP (2 brains, 4-5 images per brain, 1-2 technical replicates per image), Gtf2ird1 (2 brains, 1-5 images, 2-4 technical replicates per image), Gtf2i (3 brains, 3-4 images each, 1-2 technical replicates per image), Gtf2i+Gtf2ird1 (1 brain, 4 images, 1-3 technical replicates per image), Smad7 (2 brains from Damon Page, 7 images total, 2-4 technical replicates per image), Smad7+Gtf2ird1 (1 brain, 10 images, 2-5 technical replicates per image), Smad7+Gtf2i (1 brain, 9 images, 1-5 technical replicates per image).

Neurite morphology at the end of embryogenesis data presented herein are based on the following lateral transfections: GFP (3 brains, 5 images, 30 cells), Gtf2i (6 brains, 6 images, 32 cells), Gtf2ird1 (3 brains, 4 images, 30 cells), Gtf2i+Gtf2ird1 (2 brains, 4 images, 30 cells), Smad7 (2 brains, 3 images, 28 cells), Smad7+Gtf2ird1 (1 brain, 2 images, 29 cells), Smad7+Gtf2i (2 brains, 2 images, 30 cells). Analysis was not done in blind fashion, but every care was taken to evaluate all cells with clearly discernible neurites in each image. Gfp, Gtf2i, Gtf2ird1 cell analysis was done on different brains from the laminar migration data analysis. The angle of the apical neurite, as compared with a vertical line to the pial surface (measure of neurite orientation), and the length of primary neurite, as measured from the junction of the primary neurite to the cell body, were approximated using line measure in ImageJ. For analysis, angles outside of 15 degrees from the vertical (-15 to 15 degrees) were considered abnormal (as in Meikle et al 2008), yielding percent abnormal neurite orientations. Raw length (in pixels), given the uniform sizes of images, yielded the percent change relative to Gfp controls in length of primary neurite. These primary neurites are likely primary dendrites, however MAP2 antibody for dendrites and TAU-1 for axons needs to be done.
6.2.6. Results

*Gtf2i* and *Gtf2ird1* act synergistically to modulate laminar patterning of pyramidal neurons

Figure 6.2.2. Laminar patterning in *Gtf2i*, *Gtf2ird1* and *Gtf2i+Gtf2ird1* overexpression

Distribution of GFP-only cells was similar to that reported elsewhere (e.g. Hand et al 2005).

MANOVA (plasmid type as independent factor (*Gfp*, *Gtf2ird1*, *Gtf2i*, *Gtf2i+Gtf2ird1*), three bins (1-3 (VZ), 4-6 (IZ), 7-10 (CP)) as dependent variables) showed significant effect of plasmid for bins 1-3 (F(3,54)=4.8, p=.005) and bins 7-10 (F(3,54)=5.8, p=.002). Fisher’s LSD posthoc comparisons showed significant differences in bins 1-3 (ventricular zone) between the double plasmid (*Gtf2i+Gtf2ird1*) and *Gfp* (p=.001), and *Gtf2ird1* (p=.003), and *Gtf2i* (p=.011). In bins 4-6 (intermediate zone) no differences between plasmids were observed. In bins 7-10 (cortical plate), double plasmid *Gtf2i+Gtf2ird1* significantly differed from *Gfp* (p<.001), and from *Gtf2i* (p=.001), and from *Gtf2ird1* (p=.003).

Comparative analysis of modulation of laminar patterning by overexpression of *Gtf2i* or *Gtf2ird1* by themselves vs GFP controls shows no statistically significant differences on laminar patterning. There are subtle effects in overexpression of *Gtf2ird1* with slightly less cells making it up to the cortical plate and more cells remaining in the intermediate zone, especially relative to *Gtf2i* overexpression. There is a subtle effect in overexpression *Gtf2i* with slightly more cells remaining in ventricular zone. Thus both *Gtf2ird1* and *Gtf2i* show minor laminar patterning effects, but those are different in quality.
Simultaneous overexpression of both Gtf2ird1 and Gtf2i however has a drastic effect on laminar patterning, causing many cells to remain behind in the ventricular plate, and shrinking the cortical plate.

**Smad7, a Potential Upstream Regulator, worsens laminar patterning effects of Gtf2ird1 and Gtf2i**

MANOVA (plasmid type as independent factor (Gtf2i+Gtf2ird1, Smad7, Smad7+Gtf2i, Smad7+Gtf2ird1), three bins (1-3, 4-6, 7-10) as dependent variables) showed significant effect of plasmid for bins 1-3 (F(3,93)=3.5, p=.018) and bins 7-10 (F(3,93)=2.98, p=.035); the differences in bins 4-6 approached significant trends (F(3,93)=2.5, p=.066). Fisher’s LSD posthoc comparisons showed significant differences in bins 1-3 (ventricular zone) between Smad7+Gtf2ird1 and Gtf2i+Gtf2ird1 (p=.01), and Smad7 alone (p=.016). In bins 4-6 (intermediate zone) there were no significant differences between plasmids. In bins 7-10 (cortical plate), Smad7+Gtf2ird1 differed from Gtf2i+Gtf2ird1 (p=.031), and from Smad7+Gtf2i (p=.013).

Smad7 by itself causes a severe depletion of cells in the cortical plate phenotype, more severe than caused by synergistic overexpression of Gtf2i+Gtf2ird1. There is a trend that does not reach significance where Smad7 overexpression with either Gtf2ird1 or Gtf2i causes even
more cells to remain at the bottom of the ventricular zone. Thus either Smad7 acts in dominant fashion, or the effects of these genes run in parallel.

Figure 6.2.4. Summary of Laminar Patterning in overexpression experiments

<table>
<thead>
<tr>
<th></th>
<th>Gfp</th>
<th>Gtf2ird1</th>
<th>Gtf2i</th>
<th>Gtf2i+Gtf2ird1</th>
<th>Smad7+Gtf2ird1</th>
<th>Smad7+Gtf2i</th>
<th>Smad7</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>VZ (1-3)</strong></td>
<td>36.66 (2.24)</td>
<td>38.79 (2.73)</td>
<td>41.37 (4)</td>
<td>52.22 (3.8)</td>
<td>61.56 (2.98)</td>
<td>57.11 (3.29)</td>
<td>55.68 (3.14)</td>
</tr>
<tr>
<td><strong>IZ (4-6)</strong></td>
<td>26.4 (1.77)</td>
<td>27.86 (1.74)</td>
<td>24.62 (2.77)</td>
<td>26.47 (1.99)</td>
<td>22.67 (1.52)</td>
<td>22.81 (1.64)</td>
<td>25.98 (1.56)</td>
</tr>
<tr>
<td><strong>CP (7-10)</strong></td>
<td>37.04 (3.71)</td>
<td>33.34 (3.17)</td>
<td>33.84 (3.51)</td>
<td>21.22 (3.65)</td>
<td>15.79 (1.57)</td>
<td>19.83 (2.06)</td>
<td>18.39 (1.14)</td>
</tr>
</tbody>
</table>
Neurite morphology at the end of embryogenesis is abnormal with overexpression of *Gtf2i*, *Gtf2ird1*, and *Gtf2i+Gtf2ird1*

![Image of neurite morphology](image)

**Figure 6.2.5. Neurite Morphology at E18.5**

MANOVA showed significant effect of plasmid on both length (F(6,202)=8.6, p<.001) and abnormal angle of orientation(F(6,202)=3.9, p=.001).

Fisher’s LSD post hoc comparisons showed significant differences in length between *Smad7+Gtf2i* and *Gtf2i* (p<.001), and *Gtf2ird1* (p=.008), and *Smad7+Gtf2ird1* (p=.004), and *Gtf2i+Gtf2ird1* (p=.04); *Gfp* and *Gtf2i* (p<.001), *Gtf2ird1* (p=.007), and *Smad7+Gtf2ird1* (p=.003); *Gtf2i* and *Smad7* (p=.03); *Gtf2i+Gtf2ird1* and *Gtf2i* (p<.001), and *Gtf2ird1* (p<.001), and *Smad7+Gtf2ird1* (p<.001), and *Smad7* (p=.03), and *Gfp* (p=.046).

Differences in abnormal angles of orientation showed significance between: *Smad7+Gtf2i* and *Gtf2i* (p=.008), and the difference with *Smad7* approached significance (p=.06); *Gtf2i+Gtf2ird1* was different from *Smad7+Gtf2i* (p=.015), *Smad7+Gtf2ird1* (p=.009), and *Gfp* (p=.01), and almost different from *Gtf2ird1* (p=.058); *Gtf2i* differed from *Gfp* (p=.001), *Gtf2ird1* (p=.033) and *Smad7+Gtf2ird1* (p=.005); *Smad7* differed from *Gfp* (p=.008) and from *Smad7+Gtf2ird1* (p=.04).

Neurite length at the end of embryogenesis was found increased relative to *Gfp* control in both *Gtf2i* and *Gtf2ird1*. There is a trend that does no reach significance in the decrease of neurite length in double overexpression relative to *Gfp*, but a significant difference in neurite length...
between Gtf2i+Gtf2ird1 and Gtf2i and Gtf2ird1. The orientation of primary neurites was not significantly different from Gfp with Gtf2ird1, but was significantly different with Gtf2i and with double Gtf2i+Gtf2ird1 transfection.

Overexpression of Smad7 and Smad7+Gtf2ird1 resulted in longer neurites relative to Gfp control and to double Gtf2i+Gtf2ird1 transfection, but similar length as in Gtf2i and Gtf2ird1. Length of neurites in overexpression of Smad7+Gtf2i was no different from Gfp control. The orientations of primary neurites tell a different story: overexpression of Smad7 had similar abnormal neurite angles to Gtf2i+Gtf2ird1 and to Gtf2i overexpression, however overexpression of Smad7+Gtf2ird1 and Smad7+Gtf2i seemed to bring this abnormality towards levels observed in Gfp controls and in Gtf2ird1 overexpression.

6.2.7. Discussion

Williams syndrome region, given opposing phenotypes in deletion and duplication syndromes, has dosage-sensitive genes that are contributing to the cortical abnormalities underlying behavioral deficits. Based on the data on atypical WS deletions available in the literature, it is possible to suggest that GTF2I contributes to hypersociability and linguistic deficits, and that GTF2IRD1 contributes to the facial dismorphology and visuospatial deficits. However prior to this study, the direct effect of these genes on brain development was not investigated (or published).

We find that overexpression of each of these genes individually in mouse neocortex does not significantly affect laminar patterning, but does increase the length of neurites and increases the abnormal neurite orientations at the end of embryogenesis. Cooverexpression of both genes causes drastic depletion of cells in the cortical plate, shorter neurite lengths, and deviant primary neurite orientations. These results are consistent with previous findings – mouse Gtf2ird1 heterozygotes show normal brain histology (Van Hagen et al (2007), but WS duplication (which involves overexpression of both GTF2I and GTF2IRD1) in humans results in thickened cortex and pockets of ectopic neurons around the ventricles (Torniero et al 2008). We find that when both are simultaneously overexpressed, Gtf2ird1 and Gtf2i act synergistically to cause abnormal laminar patterning.

We further investigated effects of modulating TGF-beta signaling, which is known to interact with GTF2I and GTF2IRD1. Overexpression of Smad7, a potent inhibitor of TGF-beta signaling, caused severe depletion of cells in the cortical plate, similar to co-overexpression of Gtf2ird1 and Gtf2i, and similar abnormal orientation of neurites, but slightly longer than normal neurite lengths. Co-overexpression of Smad7 with either Gtf2ird1 or Gtf2i did not rescue migration effects, and in fact seemed to slightly worsen them, however such co-overexpression seemed to bring the abnormalities in neurite orientation towards levels observed in Gfp controls; Smad7+Gtf2ird1 overexpression however increased neurite length more than in Smad7 alone, but Smad7+Gtf2i overexpression brought neurite length in line with Gfp controls.

In further experiments, we plan on increasing the number of animals in each of the present analyses, and investigating synergistic effects of DYX1C1 downstream candidate gene and TGF-beta regulator candidate gene.

While the present results do not tell us the specific role of Gtf2ird1 and Gtf2i in brain development, and they are based on data from a small number of animals, these results do indicate that Gtf2ird1 and Gtf2i certainly each contribute to neurite morphology (length and orientation) at the end of embryogenesis, and in synergism they contribute to abnormal laminar patterning. As such, these results suggest that abnormalities underlying brain development in
Williams syndrome are similar to some of those abnormalities underlying brain development in dyslexia and in autism. In dyslexia and WS, it is laminar patterning that seems to be affected (e.g. Wang et al 2006; the present findings), and it is likely that abnormal laminar patterning contributes to abnormal connectivity and macrocephaly in autism (e.g. via PTEN (e.g. Bill & Geschwind 2009). In autism and WS, it is dendrite morphology, which is key for efficient synaptic function, that seems to be affected (e.g. Ramocki & Zoghbi 2008; Bourgeron 2009; the present findings). These abnormalities can well underlie abnormal brain development which can contribute to deficient language abilities. Successful acquisition and processing of language likely requires extensive integration and fine-grained processing of information from across multiple brain areas, and this process can well be disrupted by abnormalities in functional connectivity due to abnormal neuronal morphology and placement.

6.2.8. Acknowledgements
Thank you Orsi Kuti for teaching me about sectioning and plasmid preparations, and thank you Damon Page for teaching me everything else.

6.2.9. References


Torniero, C., Dalla Bernardina, B., Novara, F., Cerini, R., Bonaglia, C., Pramparo, T., Ciccone, R.,
Guerrini, R., Zuffardi, O. 2008. Dysmorphic features, simplified gyral pattern and 7q11.23 duplication
transcription by nuclear exclusion. Proceedings of the National Academy of Sciences of the USA, 98
(14):7789-94.
Van Essen, D.C., Dierker, D., Snyder, A.Z., Raichle, M.E., Reiss, A.L., Korenberg, J. 2006. Symmetry of
cortical folding abnormalities in Williams syndrome revealed by surface-based analyses. The Journal
of Neuroscience, 26 (20):5470-83.
van Hagen, J.M., van der Geest, J.N., van der Giessen, R.S., Lagers-van Haselen, G.C., Eussen,
Koekkoek, S.K.E., Frens, M.A., van Camp, N., van der Linden, A., Jansweijer, M.C.E., Thorgeirsson,
S.S., De Zeeuw, C.I. 2007. Contribution of CYLN2 and GTF2IRD1 to neurological and cognitive
Wang, Y., Paramasivam, M., Thomas, A., Bai, J., Kaminen-Ahola, N., Kere, J., Voskuil, J., Rosen, G.D.,
Galaburda, A.M., LoTurco, J.J. 2006. DYX1C1 functions in neuronal migration in developing
(Shanghai), 41 (4): 263-72.
Young, E.J., Lipina, T., Tam, E., Mandel, A., Clapcote, S.J., Bechard, A.R., Chambers, J., Mount, H.T.J.,
Fletcher, P.J., Roder, J.C., Osborne, L.R. 2008. Reduced fear and aggression and altered serotonin
7. Global Discussion & Conclusion

This work set out to understand the reasons for children’s overuse of the definite determiner ‘the’. At stake are not only linguistic theories, but also the understanding of how children’s semantic and pragmatic language develops. One of the theoretical camps argues the problem is in the correct use of the ‘the’, i.e. children have adult semantic knowledge, however children do not yet have correct pragmatic interpretations, which primarily have to do with distinguishing between speaker/listener knowledge. The other camp argues it is a problem in children’s knowledge of the semantic principles that guide the correct use of ‘the’, specifically that of Maximality/Uniqueness (Wexler 2003), i.e. it is just another stage that maturation of linguistic principles goes through.

The main investigation focused on comprehension of definite determiner ‘the’ in parallel with indefinite determiner ‘a’, anaphors ‘another’ and ‘same’, and demonstrative ‘that’ within a constrained act-out task where typically developing (TD) participants had to rely on their knowledge of determiners to correctly interpret anaphoric reference (referring to a salient or a nonsalient linguistic antecedent (established by e.g. an indefinite description)). In follow-up studies, we investigated whether there is indeed individual correlation between knowledge of theory of mind (as would be expected by pragmatic theories) and knowledge of the definite determiner, and we also investigated whether children show similar deficits as they do with ‘the’ in free relative clauses which require knowledge of semantic Maximality.

We find that children are aware of the salient entity in the context set as early as three years (as manifested by knowledge of ‘another’ and ‘same’), and children are aware of other’s minds by 5 years, however children begin to show adult-like interpretation of ‘the’ and free relative clauses only by the age of 6.5 years, with majority of children showing full knowledge by 8 years. Such pattern is at odds with pragmatic deficit theories, but is in line with specifically semantic deficit theory.

Turning now to language development in individuals with neurodevelopmental disorders, we find no weird and no atypical patterns of acquisition. Language development in autism spectrum disorders and Williams syndrome is known to be variable, with significant delays in syntactic and pragmatic abilities. However these disorders present an ultimate test of the competing theories: in both ASD and WS, especially in ASD, the recurrent deficit is in pragmatics – in awareness of other minds, in formulating coherent discourses, although the severities of such deficits seem to correlate with overall verbal and cognitive functioning. Whether these disorders are impaired in semantic knowledge, prior to work done in this thesis, was unknown beyond studies of omission of functional morphemes (such as ‘the’) in ASD.

We find that participants with autism spectrum disorders and those with Williams syndrome actually look like TD 3-6 year olds in their performance on determiners (Figure 7.1). Over 50% of participants show no knowledge of definite determiners, with less than 30% showing good knowledge, and with remaining children comprehending the saliency of the unique referent better when referred to with ‘that’ than with ‘the’ – a pattern found in TD children, peaking at age 5. If the deficits seen in TD 3-6 year olds are mirrored by the deficits observed in neurodevelopmental disorders, and if we are arguing that deficits in typical acquisition are semantic in nature, it follows that deficits seen in neurodevelopmental disorders are also semantic in nature.
It must be noted that averaging performance across autism spectrum disorders muddles the linguistic endophenotypes within ASD – children with autism showed no knowledge of determiners, while children with Asperger syndrome showed adult-like or partial (child-like) knowledge of determiners – understanding ‘that’ better than ‘the’. Children with PDD-NOS diagnosis formed a mixed group in knowledge of determiners, but a coherent group in that they all had milder autistic features and milder cognitive deficits than those with autism, but were worse off on such measures than those with Asperger syndrome. Actually, children with PDD-NOS as a group look like children with Williams syndrome on measures of determiners, even though they have better (near-normal) nonverbal reasoning and vocabulary abilities, whereas children with WS have below-normal abilities. Another way to interpret the results is to observe that the effect of Williams syndrome disorder on one’s knowledge of definite determiners is much less severe than the effect of autism, which erases all traces of knowledge. Milder form of autism (PDD-NOS) and WS show similar distributions of patterns. Asperger syndrome stands in its own right – outperforming all other developmental disorders.

Figure 7.1. Patterns of Performance on ‘the’ and ‘that’ in all Participants: Typically Developing Children (TD) of 3-6 years of age, 6.5-9 years of age, and teenagers Autism Spectrum Disorders (ASD), and Williams Syndrome (WS)

“A” pattern denotes good performance on both ‘the’ and ‘that’.
“B” pattern denotes a bad performance on ‘the’ with a relatively better performance on ‘that’.
“C” pattern denotes a bad performance on ‘that’ with a relatively better performance on ‘the’.
“D” pattern denotes a bad performance on both ‘the’ and ‘that’.
Another point worth noting, is that in ASD and WS disorders, it is the typically developing, youngest, grammar (TROG) controls who come closest to matching the performance of participants with developmental problems. In other words, deficits in determiners manifested by participants with ASD and WS are worse than expected given their nonverbal reasoning and their verbal ability (vocabulary knowledge) levels. In case of Williams syndrome, there is a very close match between WS and their TROG controls in terms of their semantic/pragmatic knowledge; in case of autism spectrum disorders, TROG controls outperform ASD participants.

In addition to human behavioral investigations, this work set out to investigate the effects of genes involved in Williams syndrome and in Williams syndrome duplication on brain development using mouse models, in an attempt to understand how genetic abnormality contributes to abnormal brain development that underlies abnormal language in Williams syndrome and in the duplication disorder. We find that *Gtf2i* and *Gtf2ird1*, two transcription factors that are causal to many of WS phenotypes, act synergistically to cause abnormalities in laminar patterning and in neurite morphology at the end of embryogenesis, both of which are key for establishing appropriate cortical layering which in turn is key for efficient information processing.

A few questions remain to be answered in the study of semantic and pragmatic language development.

There is a sizeable group of younger typically developing children who show good knowledge of ‘the’ before age 7 - are they precocious in their linguistic development, are they receiving any kind of super special input, what is driving their knowledge ahead of other children of the same age? Similarly, there are some children who show deficits in ‘the’ when they are 7 or 8 years old – are they disadvantaged in any way in their language development? To answer this question, we are investigating children’s nonverbal reasoning, vocabulary levels, overall grammatical comprehension levels, as well as children’s syntactic knowledge of actional and psychological passives, and children’s syntactic and pragmatic interpretation of personal and reflexive pronouns (Modyanova, Perovic, Hirsch, Wexler, in preparation). The answer to this question can also be provided by a behavioral genetic twin study of comprehension of determiners – do monozygotic twins who share 100% of their genes show more similar trends in development of determiners than do dizygotic twins who only share 50% of their genes? To what extent do shared environmental factors contribute to development of determiners?

Are children really that bad at interpreting Maximality in free relative clauses, and is there true individual correlation between knowledge of Maximality in free relative clauses and knowledge of Maximality in the definite determiner? To answer this question, we are investigating comprehension of definite descriptions in parallel with free relatives (items that involve maximality in their interpretation) and in parallel with items involving scalar implicatures, within the context of a picture-matching task that skirts around the issues with presupposition failures.

Are there subgroups of knowledge within WS, and if yes, what determines those subgroups best? In case of children with ASD, it is their overall grammatical development that predicts their knowledge of determiners, and similar delineation in participants with WS was presently impossible due to small number of subjects.
What about comprehension of determiners in other disorders? Children with Specific Language Impairment (SLI) present a perfect contrast to children with autism – both populations have severe language difficulties, but children with SLI have no theory of mind problems! (e.g. Colle, Baron-Cohen & Hill 2006). Children with Down syndrome have atrocious language skills, but have relatively better social awareness capacities than do children with Williams syndrome and autism. Crucially, what about children with the duplication of the Williams syndrome locus?

Are there brain regions that are uniquely involved in interpretation of determiners, and how does maturation of these brain areas progress? For example, adult’s Broca’s area lights up upon encountering syntactic violations of inflection and auxiliary omission of the type that children produce when they are 2-3 years old, but adult’s parietal area lights up when encountering syntactic violations in subject-verb agreement of the kind that children never make (Kovelman et al 2009). Notably, children with SLI show reversed perisylvian activity asymmetry when auditorily processing ‘the’ in stories, depressing left hemisphere temporal activity and enhancing processing in right temporal area – this a direct link between deviant neurophysiology and processing of definite determiner is established (Shafer et al 2001). What kind of brain responses do omissions or mistakes in determiners, or failures of maximality presupposition, elicit in normal and impaired adults and in children?

Are there any polymorphisms in any genes that can be uniquely associated with faster or slower acquisition of determiners (or for that matter, any aspect of the maturing language faculty)? We already know that development of syntactic constructions, such as passives and verbal inflections, has substantial heritability but little shared environment effects, with higher correlations between abilities of monozygotic twins than dizygotic twins (Ganger et al 2005, Ganger 1998). SLI Consortium (2002, 2004; Falcaro et al 2008) identified two quantitative trait loci in families with language impairment: SLI1 locus on 16q (10.5 Mb region: 16q23.1-16q24.2) associating with phonological short-term memory impairments (via non-word repetition test (NWR)), and SLI2 locus on 19q (23.5 Mb region: 19q12-19q13.42) associating with syntactic impairments (via test of verbal inflections by Wexler and Rice). Bishop et al (2005) found via DeFries-Fulker analysis (predicting one twin’s abilities based on the other twin’s abilities) that verbal inflection does not associate with NWR, but verbal inflection does associate with the sentence structure subtest of a standardized grammar test, suggesting 19q may be responsible not only for verbal inflections, but for all sentential syntax. Alarcon et al (2002, 2008) linked delay in producing a first word in males with autism to CNTNAP2 (7q35) – a gene in the neurexin family involved in synaptic function. Vernes et al (2008) show that CNTNAP2 is a downstream target of FOXP2, a transcription factor involved in a human familial speech and language disorder, and that a specific haplotype of CNTNAP2 is associated with the phonological working memory deficit endophenotype in specific language impairment. Thus, a shared genetic mechanism can be the cause of a language impairment/delay endophenotype across neurodevelopmental disorders. Dyslexia studies found several candidate regions (although none overlapping with SLI), and four candidate genes – three involved in neuronal migration (DYX1C1, DCDC2, KIAA0319) and one involved in activity dependent synaptic plasticity and connectivity (ROBO1) (Galaburda et al 2006). Notably, DYX1C1 is a downstream target of GTF2I (a WS gene) (Tapia-Paez et al 2008) – suggesting a genetic link between dyslexia and Williams syndrome. Observe, that the more precise the phenotypic biomarker – e.g. verbal inflections, or ability to read – the easier (relatively) it is to track down the genetic correlates of the behavioral biomarkers.
How do such ‘linguistic’ genes actually contribute to language development on a cellular and molecular basis? *FOXP2* is known to modulate plasticity – the flexibility of connections between neurons which is key for learning (e.g. Haesler et al 2007), and *CNTNAP2* and *ROBO1* are both involved in synaptic function – also key for plasticity. Key genes for dyslexia, and as this dissertation has found, mouse homologues of key genes in Williams syndrome – *GTF2I* and *GTF2IRD1* – affect laminar patterning – the maturing of neurons and the formation of appropriate cortical layering which is key for efficient information processing. In a unique study, Dediu and Ladd (2007) find an association between newly evolved alleles in genes contributing to human brain growth and languages that use tone mark semantic distinctions, also implicating brain growth as a key modulator of linguistic abilities. Notably, none of these genes are ‘new’ in evolutionary terms, i.e. they have key roles in development of other species of the animal kingdom, e.g. *FOX* genes have transcriptional functions conserved from yeast to humans. However, modulations in expressions or in alleles of these genes are present and are the likely contributor to human language abilities. For example, “345 genes were identified as differentially expressed between superior temporal gyrus and the remaining cerebral cortex”, of which *CNTNAP2* showed an unusual pattern of expression not found in other species (Abrahams et al 2007).

Instead of providing a final word on the maturation of definite determiner ‘the’ in children, this dissertation opens many possible avenues for future investigations.
References


