A COGNITIVE CONSTRUCTIVIST APPROACH
TO EARLY SYNTAX ACQUISITION

A thesis submitted to the University of Manchester for the degree of
Doctor of Philosophy in the Faculty of Medical and Human Sciences

2010

PAUL IBBOTSON

SCHOOL OF PSYCHOLOGICAL SCIENCES
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. List of contents.</td>
<td>2</td>
</tr>
<tr>
<td>2. List of tables.</td>
<td>5</td>
</tr>
<tr>
<td>3. List of figures.</td>
<td>6</td>
</tr>
<tr>
<td>4. Abstract.</td>
<td>8</td>
</tr>
<tr>
<td>5. Declaration.</td>
<td>9</td>
</tr>
<tr>
<td>6. Copyright statement.</td>
<td>9</td>
</tr>
<tr>
<td>7. Acknowledgements.</td>
<td>10</td>
</tr>
<tr>
<td>8. Rationale for submitting the thesis in an alternative format.</td>
<td>11</td>
</tr>
</tbody>
</table>

### CHAPTER ONE

<table>
<thead>
<tr>
<th>Section</th>
<th>Page Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Preface.</td>
<td>15</td>
</tr>
<tr>
<td>2. Introduction</td>
<td>17</td>
</tr>
<tr>
<td>2.1. Historical and conceptual foundations.</td>
<td>17</td>
</tr>
<tr>
<td>2.2. Background.</td>
<td>20</td>
</tr>
<tr>
<td>3. Formalism</td>
<td>21</td>
</tr>
<tr>
<td>4. Psychological reality.</td>
<td>27</td>
</tr>
<tr>
<td>5. Poverty of the stimulus revisited.</td>
<td>32</td>
</tr>
<tr>
<td>6. Universals, linking and continuity.</td>
<td>35</td>
</tr>
<tr>
<td>7. Functionalism.</td>
<td>45</td>
</tr>
<tr>
<td>7.1. Associationism in sheep’s clothing?</td>
<td>48</td>
</tr>
<tr>
<td>7.2. Adults’ knowledge of constructions.</td>
<td>53</td>
</tr>
<tr>
<td>7.3. Learning constructions: some preliminary issues.</td>
<td>54</td>
</tr>
<tr>
<td>7.4. Language-internal and cross-linguistic generalisations.</td>
<td>57</td>
</tr>
<tr>
<td>8. Future challenges.</td>
<td>65</td>
</tr>
<tr>
<td>9. Beyond dichotomies.</td>
<td>66</td>
</tr>
<tr>
<td>10. Synopsis of the thesis.</td>
<td>71</td>
</tr>
</tbody>
</table>

### CHAPTER TWO

**Paper: Learning argument-structure constructions: a usage-based approach.**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Introduction.</td>
<td>74</td>
</tr>
<tr>
<td>2. The context in which acquisition takes place.</td>
<td>75</td>
</tr>
<tr>
<td>3. Learning language concepts.</td>
<td>76</td>
</tr>
<tr>
<td>4. Analogy and Categorisation.</td>
<td>78</td>
</tr>
<tr>
<td>5. Bootstrapping meaning.</td>
<td>79</td>
</tr>
<tr>
<td>6. Abstracting over instances.</td>
<td>80</td>
</tr>
<tr>
<td>7. Linguistic and non-linguistic cues to argument-structure constructions.</td>
<td>82</td>
</tr>
<tr>
<td>8. Event Construal.</td>
<td>88</td>
</tr>
<tr>
<td>9. Generalisations of function.</td>
<td>96</td>
</tr>
<tr>
<td>10. Schema Productivity.</td>
<td>100</td>
</tr>
<tr>
<td>11. Item-specific knowledge and generalisations.</td>
<td>102</td>
</tr>
<tr>
<td>12. The distribution in the input.</td>
<td>105</td>
</tr>
<tr>
<td>13. Summary and Conclusions.</td>
<td>105</td>
</tr>
</tbody>
</table>
## CHAPTER THREE

Paper: *Prototype constructions in early language acquisition.*

1. Introduction. 110
2. Prototypes in cognition and language. 111
3. Acquiring the transitive construction in English.
   3.1. The developmental pattern. 120
   3.2. What English-speaking children hear. 123
4. Recent cross-linguistic studies.
   4.1. German and Polish. 127
   4.2. Cantonese, German and English. 134
5. Relative clauses. 137
6. Discussion. 139

## CHAPTER FOUR

1. Introduction. 145
2. Generalisations in general. 145
3. Linguistic categories. 148
4. Categorising argument-structure constructions: the example of the transitive construction. 151
5. Experimental Evidence for Prototype Representations. 157
6. Constructing Prototypical Transitive Semantics in Development 164


1. Introduction. 168
   2.1. Participants. 171
   2.2. Materials. 172
   2.3. Procedure. 174
3. Results. 176
4. Discussion. 179

## CHAPTER FIVE

1. Introduction. 184
2. The cue-competition model and development. 189
3. The English transitive construction.

Paper: *The Role of Pronoun Frames in Early Comprehension of Transitive Constructions in English.*

1. Introduction. 200
   2.1. Participants. 205
   2.2. Materials. 206
   2.3. Design. 207
   2.4. Procedure. 209
   2.5. Coding. 211

---

Page Number

109
110
111
114
120
120
123
126
127
134
137
139
145
145
148
151
157
164
167
168
171
172
174
176
179
184
189
191
198
200
205
206
207
209
211
3.0 Results.
   3.1. Actives versus Passives. 212
   3.2. Pronoun Frames: Comparison to Chance. 213
   3.3. Pronoun Frames: Comparison across Conditions.
       3.3.1. Grammatical Sentences. 217
       3.3.2. Ungrammatical Sentences 218
4. Discussion. 218
5. Acknowledgements. 223

CHAPTER SIX 224

1. Introduction. 224
2. The role of skewed distribution and cognitive anchoring. 224
   Paper: A Usage-Based Acquisition Model of Argument Productivity
       in Subject-Verb-Object Constructions.
       1. Introduction. 246
       2. Methodology and Model.
           2.1. The corpus. 249
           2.2. Extraction procedure. 250
           2.3. The model. 251
3. Results and Discussion. 254
4. Conclusions. 265

CHAPTER SEVEN 269

1. Review of the thesis 269
2. Some common themes: constraints on cognitive processing and properties of the input.
    2.1. Schema abstraction and elaboration 274
    2.2. Token frequency 278
    2.3. Type frequency. 279
    2.4. Type/Token Distribution. 280
    2.5. Schema complexity. 282
    2.6. Reliability. 283
    2.7. Analogy and categorisation. 285
3. Strengths and weaknesses of the approach. 287
4. Future Directions. 290

References

Word count 71,120
2. List of Tables

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>1</td>
<td>Ranked frequency distributions of constructions from several English Corpora.</td>
<td>64</td>
</tr>
<tr>
<td>TWO</td>
<td>1a</td>
<td>In the beginning local contingencies emerge first.</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>1b</td>
<td>Only later do opportunities for abstraction occur as a function the distributions and raw frequency.</td>
<td>86</td>
</tr>
<tr>
<td>FOUR</td>
<td>1</td>
<td>Comparisons between prototype, old, new and foil conditions.</td>
<td>179</td>
</tr>
<tr>
<td>FIVE</td>
<td>1</td>
<td>Within-subject variables.</td>
<td>208</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Mean scores (in bold), standard deviations (in curly brackets) and t-test statistics for pronoun frame interpretation by age.</td>
<td>215</td>
</tr>
<tr>
<td>SIX</td>
<td>1</td>
<td>The top ten ranked verbs and verbs from the tail of the distribution.</td>
<td>238</td>
</tr>
</tbody>
</table>

Paper: A Usage-Based Acquisition Model of Argument Productivity in Subject-Verb-Object Constructions. 245
3. List of Figures

<table>
<thead>
<tr>
<th>CHAPTER ONE</th>
<th>Page Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1. The two most historically influential theoretical approaches to language acquisition.</td>
<td>19</td>
</tr>
<tr>
<td>Figure 2. The academic impact of Chomsky's language acquisition theory (notice the log-scale on the y-axis).</td>
<td>26</td>
</tr>
<tr>
<td>Figure 3. A formal analysis of the sentence “It was in the garden that he met the panther”.</td>
<td>29</td>
</tr>
<tr>
<td>Figure 4. The main historical antecedents to cognitive linguistics.</td>
<td>47</td>
</tr>
<tr>
<td>Figure 5. The idiolect of a speaker.</td>
<td>60</td>
</tr>
<tr>
<td>Figure 6. The idiolects of two speakers. The intersection represents shared ways of using language: words, argument-structure constructions, agreement, case and so on.</td>
<td>61</td>
</tr>
<tr>
<td>Figure 7. Frequency distribution of linguistic items.</td>
<td>62</td>
</tr>
<tr>
<td>Figure 8. Two speakers’ linguistic knowledge in comparison with the usage-frequency of that knowledge.</td>
<td>63</td>
</tr>
<tr>
<td>Figure 9. The vastly different timescales over which linguistic structure, systematicity and function emerge.</td>
<td>68</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER TWO</th>
<th>73</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper: Learning argument-structure constructions: a usage-based approach.</td>
<td>74</td>
</tr>
<tr>
<td>Figure 1. Events associated with utterances, mapped onto a plane of usage-events.</td>
<td>82</td>
</tr>
<tr>
<td>Figure 2. The functional alignment of constructions based on their event construal; the relevant domain of analogy for argument-structure constructions.</td>
<td>91</td>
</tr>
<tr>
<td>Figure 3. Multiple domains of competing analogies.</td>
<td>95</td>
</tr>
<tr>
<td>Figure 4. The second tier of abstraction generalises over caused-motion event-utterance co-occurrences.</td>
<td>96</td>
</tr>
<tr>
<td>Figure 5. The generalisation of animacy over caused-motion schema.</td>
<td>99</td>
</tr>
<tr>
<td>Figure 6. A four-plane abstraction of the X cause Y event, the caused-motion event forms part of this abstraction, the symbolic assembly is viewed through the four planes and incorporates the activated schemas it dominates.</td>
<td>104</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER THREE</th>
<th>109</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper: Prototype constructions in early language acquisition.</td>
<td>110</td>
</tr>
<tr>
<td>Figure 1a. Ditransitive distribution.</td>
<td>119</td>
</tr>
<tr>
<td>Figure 1b. Transitive distribution.</td>
<td>119</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER FOUR</th>
<th>145</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1. Transitivity Features (Hopper and Thompson, 1980).</td>
<td>152</td>
</tr>
<tr>
<td>Figure 2. Agent and Patient as maximally distinct categories.</td>
<td>155</td>
</tr>
<tr>
<td>Figure 3. ‘Distortions’ of the Transitive Prototype according to the features defined by the maximally distinct argument hypothesis.</td>
<td>156</td>
</tr>
</tbody>
</table>
Figure 4. Examples of stimuli and results of the Franks and Bransford experiment.


Figure 1. Percentage confidence rating across conditions for Adults.
Figure 2. Percentage confidence rating across conditions for Children.
Figure 3a. Median confidence rating across conditions for Adults.
Figure 3b. Median confidence rating across conditions for Children.

CHAPTER FIVE

Figure 1. A simplified SVO schema that unifies active and passive forms in the transitive event.
Paper: The Role of Pronoun Frames in Early Comprehension of Transitive Constructions in English.

Figure 1. Example of test stimulus.
Figure 2. Mean Scores over Active and Passive Conditions for 3 Age Groups.
Figure 3. Mean Scores of Interpretation for Grammatical Pronoun Frames.
Figure 4. Mean Scores of Interpretation for Ungrammatical Pronoun Frames.

CHAPTER SIX

Figure 1. Schematisation under skewed input conditions.
Figure 2a. The distribution of all verb tokens in SVO constructions in Becky’s mothers’ speech.
Figure 2b. The distribution of all verb tokens in SVO constructions in Becky’s speech.

Paper: A Usage-Based Acquisition Model of Argument Productivity in Subject-Verb-Object Constructions.

Figure 1. Development in the SVO network by age.
Figure 2. The distribution of verb collocations.
Figure 3. The development of subject-object variability in both mother and child.

CHAPTER SEVEN

Figure 1. A taxonomic hierarchy of instances and schemas; a hypothesised architecture of the linguistic symbolic assembly.

Paper: 

Figure 2. Abstraction over time, network elaboration.
Figure 3. Contrasting token frequencies in two schematic networks.
Figure 4. Contrasting type frequencies in two schematic networks.
Figure 5. Contrasting type/token distributions in two schematic networks.
Figure 6. Contrasting schema complexity in two schematic networks.
Figure 7. Contrasting reliability in two schematic networks.
Figure 8. Analogical mapping between items and schemas in two networks.
4. Abstract

This thesis is concerned with a central question in any construction-based, usage-based theory of language acquisition: how children get from more concrete and item-based constructions to more abstract constructions. The overall approach places central importance on meaning and the role of cognition to categorise chunks of linguistic experience into conventional grammatical units. Chapter 1 outlines the historical and conceptual foundations in the field of language acquisition and justifies the usage-based approach taken in this thesis. Chapter 2 then considers what usage-based theories mean when they characterise language acquisition as ‘a developing inventory of constructions that are more-or-less schematic’. By bringing together findings from categorisation and analogy, social cognition and construction grammar the aim is to show how argument-structure constructions are learnable. Chapter 3 investigates the topic of how infants construct grammatical categories by taking a cross-linguistic look at the transitive construction in the context of a prototype theory of categorisation. Chapter 4 applies the theoretical investigations of previous chapter to an empirical experiment, specifically, making developmental comparisons of the prototypical semantics of the transitive construction in English. Chapter 5 considers further the role of different cues in children understanding of argument-structure constructions by examining the role of pronoun frames in early comprehension of transitive constructions in English. Chapter 6 focuses on how infants and mothers actually use language in a corpus study. It begins by looking at the role of skewed distribution and cognitive anchoring in schematising the Subject Verb Object construction in English. It then presents a usage-based acquisition model of argument productivity in subject-verb-object constructions. Chapter 7 concludes the thesis by summarising the experimental and theoretical work; identifying some cognitive features and properties of the input that seem to be important in all the studies; providing a critique of the usage-based approach; and finally, suggesting some key issues that future work in usage-based approaches to the acquisition of grammar needs to address.
5. Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

6. Copyright Statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the
prior written permission of the owner(s) of the relevant Intellectual Property
and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and
commercialisation of this thesis, the Copyright and any Intellectual Property
and/or Reproductions described in it may take place is available in the
University IP Policy (see
http://www.campus.manchester.ac.uk/medialibrary/policies/intellectual-
property.pdf), in any relevant Thesis restriction declarations deposited in the
University Library, The University Library’s regulations (see
http://www.manchester.ac.uk/library/aboutus/regulations) and in The
University’s policy on presentation of Theses.

7. Acknowledgements

This thesis has benefitted enormously from discussions with those at the Max Planck
Institute for Evolutionary Anthropology, Leipzig, and the Max Planck Child Study
Centre at the University of Manchester. Notable thanks go to Dr. Claire Noble, Dr.
Danielle Matthews, Dr. Grzegorz Krajewski, Dr Silke Brandt, Eileen Graf, Thomas
Grünloh, Barbara Stumper, Daniel Schmerse, Sanjo Nitschke, Mickie Glover, Bianca
Junge, Keith Austin and Anne-Kristin Siebenborn. Thank you also to those parents,
children and students who participated in the studies – obviously without their
involvement this thesis could not have been written. I am grateful to those who have
read extracts of this thesis and provided insightful comments on earlier drafts,
including: Prof. Adele Goldberg, Dr. Caroline Rowland, Dr. Inbal Arnon, Dr.
Ludovica Serratrice, Dr. Louise Connell and Dr. Amanda Campbell. The greatest debt
of gratitude is reserved for my supervisors: Dr. Anna Theakston, Prof. Elena Lieven and Prof. Mike Tomasello, who have been very generous with their knowledge while allowing me the freedom to pursue my own lines of research. Thank you most of all to my parents, Christine and Jeffrey Ibbotson, who have provided endless support and unconditional love over many years. This thesis was supported by a bursary from the Max Planck Institute for Evolutionary Anthropology, Leipzig.

8. **Rationale for submitting the thesis in an alternative format and an account of how the thesis format has been constructed.**

This thesis is submitted in the Alternative Format which incorporates sections that are in a format suitable for submission for publication in a peer-reviewed journal. The formatting of papers has been altered to make them consistent throughout the thesis but in all other respects they are as they were published or submitted for publication. This alternative format thesis conforms to the same standards expected for a standard thesis. As in the traditional format, the introduction reviews previous research (Chapter 1) and the summary draws together the various outcomes of the work (Chapter 7). In addition, two theoretical papers and three empirical papers are also presented. As of the date of this thesis submission, one paper has been published (Chapter 3), one paper is in press (Chapter 5), one paper is submitted (Chapter 4) and two chapters are in preparation for submission (Chapter 2 and 6).

My own contribution to these investigations was as follows. I researched the relevant literature for each paper; conceived the research question; designed the experimental methodology including procedure and materials (animations, picture cards etc.); conducted recruiting, piloting and testing of participants.; coded, analysed and conducted statistical analysis of the data (except where indicated); wrote the papers, submitted to the relevant journals and corresponded with the journals.
regarding revisions. My collaborators, Anna Theakston, Elena Lieven and Mike Tomasello all played a central advisory role throughout the process.

Since the studies are presented in published (or publishable) form, they do not include the more extensive theoretical or empirical background that motivated them. They are, therefore, prefaced by sections that are designed to a) set the papers in the context of the wider philosophical and empirical debates in the literature b) improve the continuity between chapters and the coherence of the thesis as a whole.

THE UNIVERSITY OF MANCHESTER

Faculty of Medical and Human Sciences

PERMISSION TO SUBMIT A Ph.D. THESIS IN ALTERNATIVE FORMAT
(SUITABLE FOR PUBLICATION OR DISSEMINATION)

We confirm that the following student in the Faculty of Medical and Human Sciences, has
been granted permission by his/her supervisor to submit a Ph.D. thesis in alternative format
approved under the regulations, including sections which are in a format suitable for
submission for publication or dissemination.

Name:          _______ Paul IBBOTSON_______________________________
Degree programme:  ___________ PhD Psychology________________________
Registration Number:  ______________ 7222617 ______________________
School:  ______Psychological Sciences____________________________

This form should be submitted with the thesis.

Signed: ___________________________  Date: ____2/10/09______
Graduate Education Manager
Faculty of Medical and Human Sciences Graduate Office

TWO COPIES OF THIS FORM SHOULD BE GIVEN TO THE STUDENT, WHO
SHOULD SUBMIT ONE COPY WITH THE THESIS AND RETAIN THE OTHER
COPY. FURTHER COPIES SHOULD BE HELD BY THE FACULTY GRADUATE
OFFICE AND BY THE SUPERVISOR.
for Sophie
1. Preface

Learning to speak the language into which we are born is such a ubiquitous feature of childhood it seems an inevitable – and perhaps on the face of it – a trivial aspect of human development. After all, language develops in every normal infant with apparent effortlessness and in the absence of formal instruction. However, as with other cognitive abilities such as vision, memory and attention, the ease and ordinariness with which we appear to learn our first language is a result of our (over)familiarity with something extraordinary. This leads to a position where “we’re least aware of what our minds do best, accordingly, we’re more aware of simple processes that don’t work well than of complex ones that work flawlessly” (Minsky, 1987: p29).

In the past hundred years language has become immensely important and problematic for philosophers and psychologists. This can be seen in the context of a wider trend towards no longer taking language for granted. By studying language in more and more detail we have been able to raise our awareness of how powerful, complex and subtle it really is. In fact, everything we know about language suggests a system for conceptual thinking and communication that is versatile and intricately structured. We also know that children show spontaneous development towards adult mastery over this system. Clearly there is something to explain here, namely, how we manage to do this.

Language development takes place in the context of many other simultaneous psychological developments so by focusing here on language acquisition and even
more specifically, on the acquisition of grammar, this thesis is only looking at a subset of cognitive development and only part of how the infant comes to understand the world around them. That said, there is a strong case to be made that by understanding the principles of language processing we will be able to shed light on other cognitive domains and vice versa. This is because language-learning seems like a case study \textit{par excellence} for the integration of a number of cognitive domains, including perception, memory, attention and social reasoning. Above all, language acquisition involves learning a skill and part of this skill involves solving an abstract problem: comprehending that a string of sounds is intended to direct attention in a particular way, and that a particular pattern of conceptualisation is conventionally expressed by such-and-such a linguistic construction. It is this problem that this thesis is mainly concerned with. The overall approach taken here places central importance on meaning and the role of cognition to categorise chunks of linguistic experience into conventional grammatical units. Fortunately, the input and output to this system is all around us in the form of child directed speech and child language. With a greater understanding of the input and output, we are then in a more sensible position to ask questions about the nature of language in the developing mind. How, for instance, are linguistic representations built up during the process of language acquisition? How abstractly are linguistic constructions represented and at what age? How do one-year olds proceed from the single word expressions to more complex forms of linguistic competence in which they comprehend and produce multi-unit linguistic utterances? How do children know what they hear \textit{is} language and not just noise? The exercise of modelling language acquisition reveals the scale of the challenge facing the language-learner. When one tries to answer these questions, the non-trivial nature of what the average pre-schooler has already achieved with her language soon becomes clear.
2. Introduction

The purpose of this introduction is to outline the historical and conceptual foundations in the field of language acquisition and justify the usage-based approach taken in this thesis. The introduction concludes with a synopsis of the contents of the following chapters.

2.1 Historical and conceptual foundations

Sometimes it seems de rigueur in the developmental psychology literature to make reference to the ‘contributions’ of nature and nurture. This introduction will, eventually, move away from this binary division and suggest this is not the most insightful framework in which to explore the issues of language acquisition. For the time being, however, it is worth pursuing this distinction as the major theories that have attempted to explain language acquisition are historically rooted on either side of this divide and are themselves a specific instance of a more general polarisation in science, that of rationalist and empiricist approaches.

For any theory of language acquisition, it is necessary to look at precisely what it is the mature speaker has command of. Linguistic theory is therefore an essential part of the study of language acquisition. For example, if one considers language to have certain structural properties that are so abstract, so complex, and the principles that govern them are so subtle and highly specific to their purpose that they could not possibly be learned by an infant, then one needs some way of explaining how infants overcome this. One celebrated and highly influential proposal, which we can call formalism, defines a general theoretical framework that is adult-centred in its approach to language acquisition and is typified by Chomskian generative grammar and the syntax of American structural linguistics. If, on the other hand, one thinks
human linguistic competence is best described in terms of its symbolic dimensions, conceptualisations and communicative function then the process of acquisition will be structured with this end-point in mind (Figure 1). Thus, developmental theories of language acquisition are often trying to answer two related questions simultaneously: what does it mean to know $x$ and how do humans acquire $x$. The two main approaches to studying the acquisition of grammar have reached different conclusions on the learnability of grammar based on their conceptualisation of what grammar is (which in turn influences how it is learnt and so on).
Figure 1. *The two most historically influential theoretical approaches to language acquisition.*

**The general questions**

What does it mean to know a grammar? How is that grammar acquired?

A set of abstract rules and principles

Impossible to learn therefore innate

Conventions employed for structuring and symbolizing thought

Building a repertoire of these conventions through usage events

**Formalism’s general answers**

**Functionalism’s general answers**

The implications of this second view of language acquisition are best described by a family of approaches we can call *functionalism*, a potentially more child-centred approach to development; this includes usage-based accounts of language acquisition and cognitive linguistics.

Although this is a somewhat oversimplified division, with many researchers positioning themselves somewhere between these extremes, the two basic positions (for better or worse) still structure much of the current debate in language acquisition. Drawing the theoretical dividing lines in this way also serves to illustrate how these
approaches arrived at fundamentally different conceptualisations of language and thought.

2.2 Background

With language being so central to what it means to be human it is not surprising that generations have reflected on its place in the natural world, what kind of biological system it is, and how it relates to other systems in our own species and others. It is clear that people have long been interested in the relationship between language and the mind, however, it was events in the middle of the last century that shaped the theoretical landscape against which many of the modern controversies are still set. The dominant doctrine in psychology at the time was behaviourism. In their search for scientific respectability the behaviourists argued psychology ought to be concerned only with the formulation of laws relating observables such as stimuli and responses; not with unobservable mental processes and mechanisms such as attention, intention, memory and motivation. This view made several assumptions about learning that would have important implications for the way the behaviourists thought about language acquisition: (1) organisms have no innate principles that guide their learning; (2) learning is the result of a general purpose process, not of a collection of mechanisms tailored to the demands of different kinds of problems; and (3) learning is a change in relation between responses and the stimuli that control or elicit them (Routledge Encyclopaedia of Philosophy, 2000). It was clearly with these principles in mind that B.F. Skinner published *Verbal Behaviour* in 1957, a work which owed much to his behaviourist contemporaries; Leonard Bloomfield, J.B. Watson, and J.R. Kantor. *Verbal Behaviour* argued that young children learn pieces of language by means of instrumental conditioning (based on principles of association) and that they
generalise to new instances by means of stimulus generalisation (based on the principles of induction). This position is closely allied philosophically to empiricism, stressing that all knowledge, including linguistic knowledge, is ultimately based on experience and the influence of this school-of-thought is noticeable today. For example, many of the behaviourist ideas about learning survive in connectionism; a computational theory of cognition that models human learning capabilities using a network of information-processing units. Methodological behaviourism also implicitly underpins much of modern day experimental cognitive science; the data on which a psychological science must rest are behavioural data, not the data provided by introspection of the contents of an observer’s consciousness. Radical behaviourism however – whereby mental terms like ‘know’ and ‘think’ were branded unscientific – began to decline as psychologists started to talk about psychological phenomena as solutions to complex and abstract problems of information processing, freely using computational metaphors borrowed from the emerging field of computer science.

3. **Formalism**

Noam Chomsky’s 1959 review of Skinner’s *Verbal Behaviour*, it is often said, was the catalyst to the ‘cognitive revolution’; a new doctrine that would come to prevail in late twentieth-century psychology that believed mental representations were valid topics of scientific enquiry. Rejecting the empiricism that was dominating Anglo-American philosophy, psychology and linguistics at the time, Chomsky argued from a nativist-rationalist position. In direct opposition to Skinner and the behaviourists he contended that inborn knowledge facilitated learning that was domain-specific. The fact that Chomsky thinks linguistic competence is distinct from other cognitive
abilities is underscored by his use of the terms “language faculty” and “mental organ”.

For example:

“It would be surprising indeed if we were to find that the principles governing [linguistic] phenomena are operative in other cognitive systems, although there might be a certain loose analogies, perhaps in terms of figure and ground, or properties of memory, as we see when relevant principles are made explicit. Such examples illustrate…that there is good reason to suppose that the functioning of the language faculty is guided by special principles specific to this domain…” (Chomsky, 1980, p. 44).

What’s more, language is a complex ability for which the human brain, and only the human brain, is specialised:

“A human language is a system of remarkable complexity. To come to know a human language would be an extraordinary intellectual achievement for a creature not specifically designed to accomplish this task. A normal child acquires this knowledge on relatively slight exposure and without specific training. He can then quite effortlessly make use of an intricate structure of specific rules and guiding principles to convey his thoughts and feelings to others, arousing in them novel ideas and subtle perceptions and judgements” (Chomsky, 1975, p. 4).

Thus human linguistic capacity rests on dedicated mental structures, many specific details of which are an innate biological endowment of the species. This once-
revolutionary, and still controversial collection of ideas, emerged from observations about language and learning that Chomsky found impossible to reconcile within the traditional tenets of behaviourism. A finite repertoire of conditioned stimulus-responses is just not capable of generating the infinite productivity and systematicity we see in language. No amount of association and induction (the same learning mechanisms that could be studied with rats pressing bars and dogs salivating to tones) could explain how virtually every sentence that a person utters or understands is unique. Chomsky argued that it was this expressive power that lies at the heart of human language. It is our possession of recursive procedures or a ‘mental grammar’ that allows us to generate an unlimited set of sentences out of a finite list of words – and he would later argue that recursion is the only feature particular to language (Hauser, Chomsky & Fitch, 2002). Universal grammar is what Chomsky called ‘the basic design underlying the grammars of all human languages and the circuitry in children’s brains that allows them to learn the grammar of their parent’s language’ (Pinker, 1994, p. 538).

Essentially, the motivation for positing a universal grammar is born out of his second observation. Chomsky considers language to have certain structural properties that are so abstract, so complex, and the principles that govern them so subtle and highly specific to their purpose, that in principle they could not possibly be learned. This is the argument that there is a poverty of stimulus, otherwise known as ‘the logical problem of language acquisition’ (Baker & McCarthy, 1981), ‘Plato’s Problem,’ ‘Chomsky’s Problem,’ ’Gold’s Problem,’ or ‘Baker’s Paradox.’ The acquisition of language is a particular instance of the more general process of the acquisition of knowledge. The philosopher Bertrand Russell rhetorically asked
“How comes it that human beings, whose contacts with the world are brief and personal and limited, are nevertheless able to know as much as they do know? Is the belief in our knowledge partly illusionary? And if not, what must we know otherwise through the senses”

(Russell, 1948: p.5)

In the last line Russell alludes to the possibility that much of our knowledge is not learned. Applied to the linguistic domain this argument-from-the-poverty-of-the-stimulus means that grammar is in principle unlearnable. Essentially, the sentences heard by the child are so full of retracings, errors and incompletions that they provide no clear indication of the possible sentences of the language. Coupled with this is the problem of unavailability of negative evidence. According to this view children have a hard time knowing which forms of their language are acceptable and which are unacceptable, because parents fail to provide consistent evidence regarding the ungrammaticality of unacceptable sentences. Worse still, when such evidence is provided, children appear to ignore it (MacWhinney, 2004). If, as adults, we know more about our language than we acquire or ever could acquire through experience, then it is a straightforward step – but one that has profound implications – to attribute the ‘missing’ linguistic knowledge to an innate, specialized algorithm for learning language. In fact, if we accept the premise of the nativist-formalist theory, then ‘learning’ no longer seems the appropriate way to talk about the development of grammatical knowledge. ‘A certain level of cognitive ability is essential for tasks like playing chess which involve learning, but there are no prerequisites for developing the ability to see or touch. That is, we do not “learn” how to see in colour or in 3-D, even
though the development of three-dimensional vision takes place only around the age of sixteen to eighteen weeks…to the extent that language is a module then, we would expect learning to be irrelevant to its development’ (Smith, 1999, p.26).

The empiricist movement in psychology grew out of the nineteenth-century development of physiological psychology, viewing learning as a system of adjustments in a network without any intervening representations. As the so-called ‘cognitive revolution’ implied, Chomsky helped to re-introduce mental representations as valid topics for scientific enquiry and in some sense a return to a more Platonic view on the nature of thought: there may be innate constraints on the way we view the world therefore we need to examine the extent to which innate ‘knowledge’ may be revised as a result of learning. To be clear, that means learning the specifics of a particular language – the surface regularities – not the learning of the machinery of grammar which is impossible according to the theory.

*Syntactic Structures* (1957) and *Aspects of the Theory of Syntax* (1965) presented Chomsky’s alternative position to the behaviourist conceptualisation of language. In these works, Chomsky (following Harris) outlined his transformational grammar; a very powerful formal tool when applied to look for distributional patterns in languages. This opened up new avenues of inquiry in the study of child development, neurology and genetics. As Figure 2 shows the impact on the study of language was unquestionable (whether these were approving or disapproving citations is another matter), spawning a myriad of linguistic theories that differed from the Chomskian orthodox by varying degrees – or what was orthodox at the time as Chomsky has continued to revise his theory in works such as *Remarks on Nominalization* (1970), *Conditions on Transformations* (1973), *Lectures on*
While Chomsky wasn’t the first to propose that all languages are built on a common grammar he did instigate something of a paradigmatic shift in linguistics, which, in retrospect, was more of a formal revolution rather than a cognitive one. For example, his work on transformational grammar in the early 1960’s motivated hundreds of scientists to take forward his ideas and use the formal tools he had created. The generativist framework continues to inspire research that explores the notion of abstract grammatical rules as the basis for generating sentential structure (I am not
setting generativism as a straw-man that no one believes in anymore). Chomsky’s revolution breathed fresh life into linguistic research that was for a long time restricted to exhaustive taxonomy and naïve methodology. Hockett, a leading figure of American structuralism of that time, defined linguistics as a classificatory science (Hockett, 1942). Champions of the generative enterprise say that one of Chomsky’s achievements has been to move linguistics beyond the historical and comparative studies of the nineteenth-century and provide a theory of why languages are the way they are – namely, because each language is a particular example of a universal faculty of the mind, whose basic properties are encoded in universal grammar. It is the validity of this claim that we now examine.

4. Psychological reality

Despite Chomsky arguing that linguistics is a branch of cognitive psychology, communicative function was still considered irrelevant for characterising grammatical constructions. For American structural linguists such Bloomfield psychological reality was of no interest. How people use language to express meaning and communicate were problems for the psychologist to deal with. Unfortunately, this was not the first, nor was it to be last time linguists and psychologists pursued their research agendas in relative independence of one-another; ‘the attitude favoured by most of [psychology’s] founding fathers was of simple neglect. Other mental processes were given higher priority and language was left to professional linguists’ (Miller, 1990: p. 9). It is peculiar that the ideas of Chomskian grammar, which did so much to precipitate the cognitive revolution, repeatedly resisted attempts to find psychological validation for the types of syntactic processes being described. To take one example, the derivational theory of complexity suggests a simple hypothesis: The more
transformations used to derive complex sentence types the longer they should take to
process (Smith, 1999: p.107). Initial work was encouraging: passives took longer to
process than actives; passive questions took longer still; and negative passive
questions took longest of all (e.g., Miller, 1962; Savin & Perchonock, 1965). However
when a wider variety of structures were tested there turned out to be no simple
correlation between the number of derivations involved and processing time (see
Fodor & Garrett, 1966 for early critical discussion). No matter, Chomskian linguistic
theory continued to evolve more elaborate levels of abstractions until there was a level
of linguistic representation, that is, a syntax that was independent of all other levels of
linguistic description including semantics and all other aspects of cognition too
(Tomasello, 1998).

Critics of this approach say that the ‘autonomy of syntax’ is a theoretical
commitment motivated more by mathematical economy than a desire to explain a set
of empirical data. In fact, however, once one takes on board the mathematical
metaphor, the result is often not economical or elegant by anyone’s standard (Figure
3); instead, it licenses a proliferation of derivations where, typically, syntactic trees
contain more structure than there are words.

The general idea is that there is a ‘basic’ sentence structure and then all other
sorts of sentences can be derived via movement and merge rules which are contained
in the universal grammar. The letters on tree nodes in Figure 3 generally refer to
grammatical constituents but the details aren’t important here. Given that for complex
sentences, a lot of intricate changes have to be made, the formal analysis of a sentence
like “It was in the garden that he met the panther” looks like this:
Figure 3. A formal analysis of the sentence “It was in the garden that he met the panther”.

Contrast this approach with that of the experimental psycholinguists who were thinking about language in cognitive and processing terms and paid more respect to the roles of attention, memory and learning in the process of acquisition. For example “One of the fundamental problems is that the [generative] model derives a tree starting from all the lexical items and working up to the top-most node [of the syntactic tree], which obviously is difficult to reconcile with left-to-right incremental parsing” (Ferreira, 2005: p.6)

Much earlier, Bever (1970) pointed out that in the cases of ambiguity, when more than one sentence structure might be permissible, there appeared to be one preferred interpretation. Moreover, this consistency appeared to hold not only across different examples of the same kind of ambiguity, but across different people too. The ambiguity of the garden path sentence the horse raced past the barn fell shows that the grammaticality of a sentence and its processability are not the same thing. Readers
consistently read *raced* as the main verb instead of the past participle so that it is only by the time they reach *fell* at the end of the sentence that they realise something has been parsed incorrectly. The important point is that this ambiguity is not just a feature of odd garden-path sentences like *the prime number few* or *fat people eat* *accumulates*, most sentences are underspecified to convey the full weight of what is being communicated if one ignores how language is used in a social context or how unspoken premises are inferred. One of the first computer parsers developed at Harvard in the 1960s, provides a famous example. When put to work on the sentence *time flies like an arrow*, it returned no less than five different interpretations:

- Time proceeds as quickly as an arrow proceeds. (the intended reading)
- Measure the speed of flies in the same way that you measure the speed of an arrow.
- Measure the speed of flies in the same way that an arrow measures the speed of flies
- Measure the speed of flies that resemble an arrow.
- Flies of a particular kind, time flies, are fond of an arrow

In the past two decades the use of computational models has seen a dramatic increase as a way to understand the acquisition process. This has paralleled a maturing of computational theories of learning, the renaissance in neural networking (connectionist) approaches and more recently, the linkages with mathematical approaches such as Bayesian inference, information theory, and statistical learning (to such an extent that the mathematics involved are largely incomprehensible to those in other fields). Issues that have been the subject of modelling include the so-called
“burst” phenomena; U-shaped curves in learning; discovery of words, categories, and grammatical structure; alternatives to rule-based accounts; and the poverty of the stimulus hypothesis (Elman, 2005). Computational models have been useful tools to explore complex interactions and emergent phenomena where it is only through running a simulation that one can know how a model will behave. They also offer the opportunity to explore aspects that would be difficult to justify to an ethics committee (e.g., creating a virtual impairment in a model’s network to simulate a brain lesion). At the very least, computational models can be thought of as enforcing a level of detail and descriptive rigour in a theory that a verbal description might not possess. And in some cases, it requires the act of designing a computational model, which distances us from a phenomenon with which we are all-too-familiar, to realise we even have a particular linguistic function (or else we are only made aware of it when it malfunctions, such as in a rare neurological condition).

The lesson from the study of information processing in natural and artificial cognitive systems is that ‘theorists would be ill-advised to ignore the questions of the psychological reality of the syntactic structures they posit, of the algorithmic feasibility of their learning and manipulation, or of possibility of their implementation in a neural substrate’ (Edelman & Waterfall, 2007: p. 3).

In a similar way, and to pre-empt a later theme of emergent phenomena, MacWhinney (2005: p. 196) has made the point:

“Emergentism holds that domain-general processes operate across a variety of domain-dependent landscapes or substrates, including the brain, the body and society. To attempt to construct an emergentist psycholinguistics that ignores the topology of these landscapes would be like attempting to build an account
of honeycomb formation that ignores the honey and the wax. The importance of the interaction between domain-general processes and specific landscapes has become more and more apparent as emergentist accounts increase their emphasis on the embodied and grounded nature of language and cognition”

Chomsky’s linguistic theories are theories of what constitutes knowledge of a language rather than of how knowledge of language is put to use; a distinction between what he called competence and performance. For example, in one of the later revisions of the generative framework, the Minimalist Program (Chomsky, 1995) operates entirely at the abstract competence level. Minimalism refers to a theoretical desire for the derivational distance between thoughts and utterances to contain only those representations that are ‘conceptually necessary’, that is, the number of steps needed to link meaning and sound should be as small as possible. Given the focus on competency, the symbolic computations of idealised speaker-listeners take place “in a completely homogenous speech-community… unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors in applying his knowledge of the language in actual performance” (Chomsky, 1965: p. 3-4). Spontaneous spoken speech shows how removed this idealisation is from the way in which people actually use language. The question is whether the competence/performance division is useful or justified or more likely whether it is an artefact of viewing language through the lens of generative grammar.

5. Poverty of the stimulus revisited

To recap, the poverty of the stimulus argument is concerned with ‘how a learner can correctly generalise from a finite sample of sentences in context to the infinite set of
sentences that define the language from which the sample is drawn’ (Pinker, 2004: p. 1). Pinker frames the problem in this relatively ‘theory-neutral’ language in the hope that whether one is a nativist, empiricist, behaviourist, connectionist, constructivist or emergentist, one can at least agree on the logic of the problem, as originally spelt out by Quine (1969). Part of this problem is knowing what sentences are possible and what are not. Consider the following sentences, which are presented in the ‘minimal pairs’ format, a common analytic tool used in linguistics to draw attention to a pattern. They often take the form of ‘if A is to B as C is to D, how then do we account for D?’

1a) The ninja turtle danced while he ate pizza.

b) While he danced, the ninja turtle ate pizza.

c) His archrival danced while the Ninja Turtle ate pizza.

d) He danced while the Ninja Turtle ate pizza.

(examples from Crain, 1991)

The noun-phrases $i$ and $j$ can pick out the same referent in 1a-c but co-reference in 1d is impossible. The question for general-purpose learning mechanisms is “why don’t learners who have heard 1a-c generalise to 1d?” What constraint on generalisation prevents a learner from co-referencing in 1d? This question is addressed in Chapter 2.

As already noted, one of the main reasons for proposing a universal grammar can be traced back to the logical problem of language acquisition. It can’t be that the input to the learner is too inconsistent and incomplete to determine language acquisition because children do acquire grammar. This is part of a more general problem of induction where many different laws are consistent with a set of
observations. Specific to language, if many grammars are consistent with a set of sentences what is to prevent the learner hypothesising too large a grammar – from which they would never recover given the lack of negative evidence. ‘People attain knowledge of the structure of their language for which no evidence is available in the data to which they are exposed as children’ (Hornstein & Lightfoot, 1981: p. 9).

Nativists argue the pattern of errors (and non-errors) found in children makes traditional “learning” implausible as the basis for acquisition. Is this really the case? If it can be shown that the logical problem of language acquisition is surmountable then this severely weakens the need for a universal grammar. MacWhinney (2004) argues that although the logical problem of language acquisition provided a useful focus for child language research in the 1970s and 1980s, it has not kept pace with advances in theory, experimentation, and observation.

MacWhinney notes that there is little evidence for truly error-free learning in the absence of exposure to relevant positive data, although there are areas of the grammar where error rates are low. Recovery from overgeneralisation is supported by a set of four powerful processes (competition, cue construction, monitoring, and indirect negative evidence) that provide redundant and complementary solutions to the logical problem. Alternative characterisations of the nature of the target grammar can take much of the logical part out of the logical problem. For example, there is no poverty of the stimulus when a structured inventory of constructions is the adult endpoint (e.g. Tomasello, 2003). Of course, what exactly is meant by ‘a structured inventory of constructions’ needs to be spelt out in detail if this statement is going to change people’s minds on the poverty of the stimulus, so this thesis will return to this characterisation of linguistic representations in later theoretical and empirical chapters.
The truth of Chomsky’s impoverished input claim depends on what one means by impoverished. A number of studies suggest that adults’ speech to children contains fewer false starts and hesitations and is therefore more grammatically “correct” than speech between adults (Drach, 1969). On the other hand it is clear that children are receiving a very skewed input by comparison with the range of structures in principle available to adults: for instance more questions and fewer complex sentences (Broen, 1972; Furrow, Nelson, & Benedict, 1979; Newport, Gleitman, & Gleitman, 1977; see Pine 1994; Richards, 1994 for reviews). Cameron-Faulkner et al. (2003) provided further detail about the role of frequency in child directed speech and children’s language learning. Based on their sample – and broadly corroborated by similar analysis of a more diversely collected sample (Wells, 1981) – they estimate that English-speaking children hear, every day, something in the order of 7,000 utterances – including about 2,000 questions, about 1,500 fragments, about 1,000 copulas, and about 400 complex utterances, if nothing else, casting doubt on Chomsky’s ‘relatively slight exposure’ claim. Without further studies of this type, of course it is possible that this may not characterise the input to all children but to pick an example of one grammatical construction; this means an average infant born into an English-speaking environment will have heard 766,500 examples of a Subject-Verb-Object construction by the time they are two (whether they are hearing them as Subject-Verb-Object constructions is another matter).

6. **Universals, linking and continuity**

Children are born into any one of 6000-plus languages so universal grammar must have a way of linking itself to a particular language. The difficulty is that any universal that might be a candidate for pre-existing knowledge looks different in
different languages. Slobin (1997) and Braine (1992) have made the case that there is too much variability across languages for any static and innate look-up table to function in the way it would need to, to solve the problem of linking. Of course many universal properties of language are not specific to language but are simply reflections of universals of human experience. In fact, most universals are expressed as ‘if…then’ dependencies, for example, if a language has both derivational suffixes and inflectional suffixes, then the derivational suffixes are always closer to the word stem than the inflectional ones (Kiparsky, 1982). So to capture the diverse range of linguistic phenomena they need to explain, universals (elements of universal grammar), are often described in quite general, abstract terms. Formally, they have been conceptualised as hierarchically ordered sets of binary parameters, invoking a ‘switch-setting’ metaphor. According to the principles and parameter theory (Chomsky, 1981b), the setting of higher-order switches (e.g., left vs. right branching) cascades through the hierarchy constraining the settings of other parameters that are more deeply embedded (e.g., pro-drop). However, the difficulty of recognising a particular parameter trigger in the linguistic form (head, complement etc.) requires a level of linguistic knowledge that is logically prior to any act of parameter setting, in which case, if the language-learner already has the knowledge there is no need to set the parameter. Principles and parameters theory attempts to solve the linking problem (which can be traced back to the logical problem of language acquisition) at the expense of making extremely strong claims about the nature of human language. For parameter-setting to work three conditions must be met:

“First, there must be a small set of parameters constituting the set of possible human languages. Second, there must be a clear specification of the unmarked
settings of these parameters. Third, there must be a clear specification of the surface structure triggers that would lead the child to move from an unmarked parameter setting to a marked parameter setting for each of the hypothesized parameters. Despite over two decades of work within the framework of principles and parameters, none of these three conditions has yet been met” (MacWhinney, 2004: p. 30).

Furthermore, there have been virtually no successful proposals for what the specific aspects of the parameters or principles might look like (for discussion of this failure see, e.g. Culicover, 1999; Jackendoff, 2002; Newmeyer 1998; Webelhuth and Ackerman, 1998, Tomasello, 2003). A sample of candidates for universal grammar ‘content’ includes:

- Lexical categories (N, V, A, P, Adv) and functional categories (Det, Aux, Deg, Comp, Pro, Conj) (O’Grady, 1997)
- X-bar syntax and linking rules ‘NP = object’, ‘VP = action’ (Jackendoff, 2002)
- As Jackendoff with the addition of ‘subject’, 'object', movement rules, and grammatical morphology (Pinker, 1994).
- Polysynthesis, ergative case, serial verbs, null subject. (Baker, 2001)
- V to I movement, subject initial, affix hopping, pied piping, topic marking, I to C movement, Q inversion, and oblique topic. (Fodor, 2003)
- wh-movement, island constraints, the subset principle, head movement, c-command, the projection principle, and the empty category principle (Crain & Lillo-Martin, 1999).
• Well-formedness constraints such as stay, telegraph, drop topic, recoverability, and MaxLex (see Haspelmath for a review, 1999).
• Distinctive features, double articulation, prediction and reference, lexical categories, argument hierarchy, adjunction, and quantification (Wunderlich, 2004).
• Recursion (Hauser, Chomsky, and Fitch, 2002)
• Merge (Chomsky, 2004).

More worrying than the relative lack of overlap in this list is the absence of a debate in the literature as to which ones are correct and by what criteria. One might imagine that one criterion by which you should be able to decide membership to universal grammar is that of being universal. However, Jackendoff (2002) encourages us to think of universal grammar as a tool-box from which you can draw. When a language is found that doesn’t contain one of these tools, it is not evidence that that it doesn’t exist. When a language is found with a new set of tools, they are simply added to the list of universals. Indeed, one might expect the existence of significant cross-linguistic variation in such things as basic grammatical categories to be potentially damaging to universal grammar. Moreover many languages show no evidence of having any form of movement rule or configurational pattern of phrase structure organisation yet these are still assumed to be universal (e.g., by Radford, 1997). Of course, generative linguists can always posit more covert movement and unseen elements but there must come a point when they need to explain what kind of evidence could refute the universal grammar hypothesis. In summary, parameters assume a range of formally specified lexical and functional categories whose cross-linguistic status is problematic at best (Croft, 2001). It is also worth mentioning that innately pre-specified binary
parameters seem like an extremely implausible biological mechanism with no analogies in other human skills, either cognitive or social (Bates and Goodman, 1998; Deacon 1997; Elman et al. 1996; Sampson 1997, Tomasello, 2003).

There is another problem with universal grammar, one of developmental change. Because there is a single universal grammar, the basic linguistic representations must be the same throughout all stages of child development, the continuity assumption (Pinker, 1984). Given this, it is not clear how one moves from ‘concrete’, baby utterances such as “more juice” and “doggie gone” to the kind of formal descriptions of adult language being proposed by Chomsky and others (see the abstract constituent categories in the syntactic trees of Figure 3; see also Radford (1990) and Pinker (1989) for attempts to address this).

When generativist theorists talk of universals it is instructive to keep in mind what they mean by universal. If we extrapolate back through time at the rate we have been losing languages since western colonisation, a conservative estimate would be that there once existed 500,000 languages. Of this half-million, only about 5-8000 languages are still with us today. Of these remaining languages about 1000 have been studied in anything like the detail one would need to make claims of universality. So a statement like ‘all languages contain \( x \)’ might only be true of 0.2% of the total languages that exist or have existed. However, the claim of universals is actually weaker than this. As already noted, universals are generally not of the form ‘all languages have an \( x \)’ but conditional tendencies of the type ‘in languages where there exists an \( x \) it is likely there is also a \( y \)’. Superficially this may appear similar to Greenbergian typological dependencies (Greenberg, 1963), however Greenberg’s source of such dependencies – the issue at stake here – are, for example, principles of information structure and economy and therefore radically different from the
traditional UG account. The precarious status of absolute linguistic universals is due to the fact that one counter-example downgrades the universal to a very strong trend. Obviously, whatever theoretical background one is arguing from, we only have the languages that are currently in use to work with in any depth. The point is that claims of linguistic universals rest on an extremely small sample. Imagine how distorted our view would be of what is possible in language (or the contents of UG) if English, for example, was the only language left to study.

In conclusion, Chomskian linguistics has undergone many internal revisions and upheavals since Syntactic Structures and Aspects and it would be impossible to review them all here, however, several enduring characteristics can be identified:

- **Formalism** – seeks to specify with maximum precision the rules and principles which generate all and only the grammatical sentences of a language.
- **Modularity** – it was always assumed within generative grammar that in any act of language use there is an interaction of linguistic knowledge with other cognitive capacities such as memory and perception. But linguistic knowledge itself is regarded as a separate cognitive faculty, informationally encapsulated and structured according to its own specific principles.
- **Sub-modularity** – the mental grammar itself has been attributed a modular structure. That is, X-bar principle, Theta principle and so on, are thought of as smaller components of a grammar module. The complexity of grammar emerges from the interaction of these simpler modules.
- **Abstractness** – a trend towards entities and processes that have no overt manifestation in actual linguistic expressions. E.g., underlying structure which
bears little resemblance to surface structure, structures which contain invisible entities like ‘traces’ and ‘empty categories’, and movement operations.

- **Search for high-level generalisations** – a focus on phenomena which are believed to be subject to a small number of very general principles, such as *wh*-movement (as manifested, for example, in question formation), raising, and anaphora.

- **Restriction to ‘core’ phenomena** – facets of linguistic knowledge which are manifestly idiosyncratic and not subject to general rules have tended to be pushed aside as theoretically uninteresting. Expressions based on individual words or phrases, such as ritualised greetings, idioms, metaphors, and noncanonical phrasal collocations, are relegated to the ‘periphery’. This division between core and periphery leads to the dual process route approach to language acquisition (e.g., Pinker’s words and rules, 1999).

(adapted from Taylor, 2002: p 7)

In light of these general characteristics of Chomskian linguistics, one might ask in what sense such an enterprise might legitimately be described as cognitive. The theory has tended to be driven by its own internal logic, not by any considerations deriving from independently established facts about human cognition. So, to what extent has the formalist approach answered the question of how children develop grammatical competence? The evidence we have considered so far is summarised below:

1. The formalist framework has produced some insights into language structure (the discovery of island constraints using the formal tools of generative grammar

41
being a notable example) but the psychological reality of the wider formalist structures posited has not been demonstrated. This is hardly surprising when one considers the attitude of its chief architect “I think a linguist can do perfectly good work in generative grammar without ever caring about questions of physical realism or what his work has to do with the structure of the mind” (Chomsky, 2004: p. 56).

This is because Chomsky never set out to describe performance, his is a theory of competence. We have already seen that this distinction allows generative grammar to maintain the ‘core-and-everything-else’ idealisation of grammar but this creates its own problem of falsification.

2. Chomskian linguistics and American structural linguistics both ignore the semantic and pragmatic dimensions of grammatical constructions in order to concentrate on competence; essentially they are interested in defining the universal principles which constrain the class of all languages and without which, language, would in principle be unlearnable. The focus by those working in the generativist framework on defining what is the initial state of universal grammar comes at a price, ignoring as it does how people actually use language and what is known about the mind.

3. Skinner’s conceptualisation of learning – the target of Chomsky’s withering review, which was, in turn, the birth of modern formalism – is not the same as the modern conceptualisation of learning and the mind. There is a growing body of evidence that the nature and properties of at least certain patterns in language are learnable on the basis of general categorization strategies, explored below in § 7 (e.g. Bybee and Slobin, 1982; Bybee and Modé, 1983; Goldberg, Casenhiser &
Sethuraman, 2004; Jackendoff, 2002; Lakoff, 1987; Pullum and Scholz, 2002; Taylor 1995, Tomasello, 2003). Furthermore, the cognitive and social learning skills that children bring to the acquisition process are much more powerful than previously believed. There is no poverty of stimulus if we assume that children are cognitive beings seeking to make sense of the linguistic communication addressed to them (Tomasello, 2003; 2008).

4. It has been over four decades since Aspects of a Theory of Syntax, and there is less agreement than ever before about precisely what is and what is not in universal grammar, it is not surprising then that many modern generative grammarians are abandoning the idea of parameters all together. Universals appear to be very elusive – it is very hard to find two linguists who agree on specific contents of universal grammar and maybe this is telling. There is very little evidence from those who carry out detailed analysis of individual languages cross-linguistically for the existence of contentful language universals of the type one would expect with an innate universal grammar (see Evans & Levinson, 2009 summarising decades of cross-linguistic work by typologists and linguists, showing just how few and unprofound the universal characteristics of a language are, once we examine the diversity offered by 6,000-8000 languages).

5. As we have seen, universal grammar seems to replace one problem – how children acquire their language – with two other problems – how can the child link her abstract universal grammar to the grammatical peculiarities of a given language (the linking problem) and how can we understand the changing nature of children’s
language across development if universal grammar is always the same (the problem of continuity)?

6. Chomskians have made a point of stressing that a grammar, as conceived within the formalist tradition, could not in principle be learned by any general learning mechanism. For this reason the basic architecture of the grammar has to be pre-wired into the human mind. The evidence for this claim is none other than linguistic theory itself (Taylor, 2002). In fact, there are no viable algorithms for learning Minimalism or other transformational generative grammars. But recall, principles and parameters, the continuity assumption and the linking problem are only problems if we accept the theory that proposes them. This section has introduced some arguments that begin to undermine the need for universal grammar by showing that its *raison d’être* – the logical problem of language acquisition – is not a problem given the correct conceptualisation of learning and by including function in a theory of language acquisition. In science logical demonstrations are only as good as their premises, which are demonstrably false in the case of at least some poverty of the stimulus arguments (Pullum & Scholz, 2002).

This introduction has devoted enough space to why this thesis *doesn’t* follow in the formalist tradition; we now need to build a case of why we should adopt the major alternative theoretical approach to language acquisition, functionalism.
7. Functionalism

*The greatest challenge to any thinker is stating the problem in a way that will allow a solution.*

- attributed to Bertrand Russell (1872-1970)

I believe Chomsky set in motion an intellectual diversion in the field of language acquisition research because he framed the problem in a way that didn't allow a solution (at least not a very sensible one). He laid language out as an object – decontextualised and ungrounded – where trying to capture linguistic patterns paper-and-pen style really does create a ‘logical’ problem and a poverty of the stimulus. But as a communicative agent embedded in communicative context this isn't the kind of problem the child has to solve. In recent years a new view of language and human linguistic competence has begun to emerge. This view is represented by a group of theories most often called cognitive-functional linguistics but sometimes also called usage-based linguistics to emphasise their central processing tenet that language structure emerges from language use (e.g., Bybee, 1985, 1995, 2002; Croft, 1991, 2001; Dąbrowska, 2005, 2006; Givón, 1995; Goldberg, 1995; Langacker, 1987, 1991, 2003, 2006; Tomasello, 2003). This view of language takes the need to learn each language’s idiosyncrasies seriously and proposes that there is no innate grammar. Instead all of grammar – from the most regular of structures (e.g., the cat sat on the mat) to the most peculiar of idioms (e.g., the more the merrier) – is learned. It is worth briefly reminding ourselves of how this is deeply unlike the classic formalist vision: “in certain basic respects we do not really learn language; rather, grammar grows in the mind” (Chomsky, 1980: p. 134). Functionalism says that this is fundamentally the
wrong way to think about language. Instead, ‘linguistic structures are generated by a large, open set of constructions of varying degrees of abstraction and complexity, which embody both form and meaning and are acquired through socially situated experience in a given language community, by probabilistic learning algorithms that resemble those at work in other cognitive modalities’ (Edelman & Waterfall, 2007: p.1). The idea is that by conceptualising language in this way it turns acquisition from a logical problem to an empirical one (Tomasello, 2003).

Figure 4: The main historical antecedents to cognitive linguistics.
Figure 4 shows some of the main historical antecedents to cognitive linguistics, one modern instantiation of the view we have called functionalism. The rise of a particular approach has often been in response to the prevailing doctrine of the time: Behaviourism is widely considered a reaction to introspective psychology, cognitivism a reaction to behaviourism, grounded cognition a reaction to cognitivism and so on. It is informative to consider the philosophical lineage of cognitive linguistics, as its influences have come from across a continuum between rationalism and empiricism.

7.1 Associationism in sheep’s clothing?

It is a misunderstanding to assume that when usage-based theorists talk of learning they are committing themselves to the same vision of learning as the early behaviourists (e.g. trial and error, rewards for successes). The modern conceptualisation is something much more sophisticated and diverse than that. For example, by attending to transitional probabilities children are able to find ‘words in a sea of sounds’ and treat these newly acquired words as part of their lexical inventory (Saffran, 2001). Eight-month-old infants are sensitive to the statistical tendency that transitional probabilities are generally higher within words than across words (Saffran, Aslin, and Newport, 1996). Children are also able to discover syntactic regularities between categories of words as well as the statistical regularities in sound patterns (Marcus et al. 1999). By twelve-months-old, infants can use their newly discovered word boundaries to discover regularities in the syntactic distributions of a novel word-string grammar (Saffran and Wilson, 2003). And by fifteen months old, infants are able to combine multiple cues in order to learn morphophonological
paradigmatic constraints (Gerken, Wilson, and Lewis, 2005). For example, consider the following sentences.

(1) Here is a pum (≈ count noun).
(2) Here is Pum (≈ proper noun).
(3) I am pumming (≈ intransitive verb).
(4) I pummed the duck (≈ transitive (causative) verb).
(5) I need some pum (≈ mass noun).
(6) This is the pum one (≈ adjective).

(examples from MacWhinney, 2005)

If we assume the learner already knows something of the distributional properties of items of their language, the suggestion is that children are able to bootstrap, or at least hypothesise a meaning from the syntactic context – ‘hypothesise’ to emphasise that this is a probabilistic process (represented by ≈ above), the initial bootstrap offers the best hypothesis given the evidence (input) awaiting modification or rejection in light of subsequent evidence.

Although nonsense verse, the following passage of Jabberwocky highlights the same sensitivity to distributional regularities in language that allows us to have an intuition that slithy is probably a property that modifies toves in some way and that gyre and gimble is something that toves can do in a particular place or time.
'Twas brillig, and the slithy toves
Did gyre and gimble in the wabe;
All mimsy were the borogoves,
And the mome raths outgrabe.

Lewis Carol, 1872

This is illustrative of another point: it highlights that language and thus language acquisition is a division of labour between acquisition of structure and acquisition of content, traditionally described as the difference between the closed class system (prepositions, conjunctions, demonstratives, determiners, pronouns, auxiliary verbs, grammatical patterns) and the open class system (nouns, verbs, adjectives). Knowing that *a* or *the* prototypically precede entities or things in English begins to limit the degrees of freedom as to the kinds of things that a new entity appearing in that slot is likely to be. Encyclopaedic knowledge can also structure a scene by applying what you know about the way the world works in a communicative context, for example there are only so many plausible interpretations if you hear a situation being described with the words *feeds*, *mother* and *baby*.

The question is whether these kinds of statistical heuristics, learning biases and probabilistic strategies are powerful enough to take infants all the way through to adult-like grammatical competence? This question is dealt with in some detail in Chapter 2, but for now, notice the emphasis has shifted. Many researchers are not arguing over whether or not such things are learnable in principle – the evidence suggests that many of them are – but, out of the possible strategies, which ones do children actually use, how and when.
If we are interested in acquisition and performance (as opposed to being only concerned with competence) we need a theory with room for serious developmental change and one that is integrated generally into the rest of cognition. For example, as Pinker (1995; pp.199) points out:

“children selectively pick up information at the ends of words (Slobin, 1973), and at the beginnings and ends of sentences (Newport et al., 1977) presumably because these are parts of the sentence string that are best retained in short term memory. We also need a theory that accounts for the speed and accuracy with which children acquire their language. For example, apart from constructions that are rare, predominantly used in written language, or mentally taxing even to an adult (like The horse that the elephant kicked kissed the pig), the fundamental parts of all languages are acquired before the child turns four (Slobin, 1985,1992). When researchers focus on a single grammatical rule and count how often a child obeys it versus how often he or she flouts it, the results are just as impressive: for just about every rule that has been looked at, three-year olds obey it a majority of the time (Crain, 1991; Marcus et al., 1992; Pinker, 1984, 1989; Stromswold, 1990).”

Goldberg might counter “Surely it is premature to give up hope that humans, with our rich cognitive abilities, complex social skills, predilection to imitate, and 100-billion-neuron brains, can learn language from the available input. Children have ample motivations to learn language, hear thousands of utterances every day, and receive indirect statistical feedback” (Goldberg, 2006, p. 69). But surely, the nativist may respond, language learning looks like a very different kind of learning to anything else, for example the difference between speaking and writing suggests that language
has all the hallmarks of an instinct. Learning to write requires formal instruction, explicit feedback and years of structured tuition, whereas speaking spontaneously emerges and unfolds in the same way that other instincts do. The question is then ‘what would be a fair comparison?’ What other activity that the child faces provides 5000-7000 exemplars per day on which to learn, each one directing and sharing attention, with its own communicative intention? Formalism has to acknowledge that the learning capacities are strong enough to learn any one of the 6000-plus languages in the world, each with its own set of grammatical conventions, therefore universal grammar, if it exists at all, could only provide a starting point (Matthews & Tomasello, in press).

In the generativist framework, knowledge of language is innate with two exceptions: the lexicon and the setting of the binary parameters of universal grammar, which, under these assumptions, are all that distinguishes, say, Mandarin from English or Swahili (Edelman & Waterfall, 2007). Learning a language then is ‘just’ a matter of setting parameters and learning words: their pronunciation, features and meaning. Functionalist approaches do not have a problem, per se, with organisms having adaptive specialisations that contain pre-existing knowledge of something they are likely to encounter in their environment – the problem is language cannot work this way because of the problems already described in the previous section.

Any theory that has descriptive as well as explanatory adequacy needs the right balance between constraints and freedom to account for diversity between languages and development within individuals. It is worth looking in depth at one account of functionalism to see how far it meets these criteria. Goldberg is a key proponent of the usage-based approach to language which shares the same theoretical desiderata for descriptive and explanatory adequacy as other accounts of language
however there is much more emphasis on psychological plausibility. Goldberg argues that language forms an integral part of human cognition and as a result, any insightful analysis of linguistic phenomena needs to be rooted into what is known about human cognitive abilities such as memory, learning and attention. Given this cognitive commitment it is natural that it is an increasingly empirically driven enterprise drawing on corpus resources and psycholinguistic experimentation. Goldberg’s particular variety of usage-based theory – cognitive construction grammar – has been most systematically described in *Constructions: a new theoretical approach to language* (2003) and more recently in *Constructions at Work, the Nature of Generalization in Language* (2006) which focuses more on the relationship between cognition, constructions and language in general.

### 7.2 Adults’ knowledge of constructions

The constructionist approach to language stresses that speakers’ knowledge of language consists of systematic collections of constructions, i.e. form-function pairings that are learned from the input with domain-general processes. Unlike traditional generativist approaches, in the constructionist view of language there is no distinction between the ‘core’ and ‘periphery’ of grammar, neither are there null-elements, abstract movements or derivations. Instead, words, morphemes, idioms and argument structure are all considered the same kind of entity – constructions. The constructionist account acknowledges that we retain an impressive amount of item-specific knowledge, including relative frequencies of usage, and that we also categorise the input we hear into patterns based on form and function.

This raises the question of how learners acquire generalisations such that they can produce an open-ended number of novel utterances based on a finite amount of
input. Goldberg makes the case that insights gained from research in general categorisation can explain what makes learners create a constructional category instead of treating each utterance as an unrelated idiom. As language users we have item-specific knowledge, for example, a specification of what kinds of phrases a verb can appear with. A familiar example is the distinction between a transitive verb like devour, which requires a direct object (*He devoured the steak vs. He devoured*) and an intransitive verb like dine, which does not (*He dined vs. He dined the steak*). Likewise, abstract forms can also convey meaning, independently of the main verb. For instance, if we are only presented with the sentence *he mooped her a flower* without any further context, we might interpret *moop* as meaning something analogous to ‘give’ (and the majority of people do). The reason we have intuitions about unfamiliar verbs is because we can understand the meaning of the verb based on its relationship to the other items. The construction (in this case the ditransitive) must be signalling what *kind* of verb *moop* can be, i.e. one that behaves like *give*. Incidentally, when usage-based theorists use conventional labels such as ‘theme’ ‘verb phrase’, ‘recipient’, they emphasise that this linguistic notation is a convenient shorthand for roles that are determined by the meaning of the construction and by using such terms they are not committing themselves to any theory of phrase structure, X-bar theory etc. Hence they are used tentatively and, ideally only when the construction has a foundation in functional cognitive schemas, for example Fillmore's (1977a, 1977b) ideas on the everyday interactional scenes and frames that structure human language.

7.3 Learning constructions: some preliminary issues
Assuming that constructions are at least as reliable cues to overall sentence meaning as verbs, the question then for those interested in the developmental trajectory of the construction is at what degree of abstraction is knowledge represented and at what age. We know that two-year-old children appear to have an intuition that the typical transitive construction at a very basic level represents asymmetrical activities rather than two participants engaging in the same act simultaneously, and young two-year-olds already distinguish between such things as ‘x is tickling y’ and ‘y is tickling x’ (Hirsh-Pasek & Golinkoff, 1996; see Gertner, Fisher & Eisengart, 2006 establishing this for novel verbs). It seems that here we must be cautious about attributing to young infants highly abstract knowledge, such as the transitive construction with slots for ‘subject’ and ‘object’. This is because a more parsimonious account, such as some schema of the type ‘x does something to y’ could equally accounts for children’s performance on these types of tasks.

Goldberg argues that through a combination of the properties of the input (contra poverty-of-stimulus arguments) and general cognitive processes, we are able to account for how we learn our systematic and complex knowledge of constructions. It is clear that children and adults do not rampantly generalise over a set of elements as if they were logicians cranking through every possible permutation. In order to explain why we make certain generalisations and not others from ‘every logical possibility’ the formalist approach has been to impose very specific top-down solutions (UG). The functionalist approach starts from an embodied cognition perspective and therefore ‘every logical possibility’ is not the problem space to begin with (more on this in Chapter 2). For a start, most humans (apart from theoretical physicists) move and think in three dimensions of space and one of time, we see and hear between a particular spectrum and bandwidth, we have difficulty imagining that
one thing can exist in two places at once, we interpret the actions of animate entities as purposeful and so on. We share these properties with other people and we *act* as if we share these properties with other people. This is important for language as many everyday concepts are indicated by linguistic conventions. Goldberg makes the more specific claim that only certain generalisations are entertained because generalisations are constrained by conservative learning and statistical pre-emption. It is also true that children and adults readily generalise beyond their experience. This is encouraged by shared similarity among items, the predictive value of forming a category and priming. Categorisation in particular will be explored in greater depths in later chapters. These processes all play a role in guiding us to make the right kinds of analogies and if one pays attention to the communicative *function* of the utterance, some improbable generalisations are never entertained to begin with. To give one example:

(1a) She painted the green house.
(2a) She painted the house green.
(3a) She knew the green house.
(4a) *She knew the house green.

On the basis of (1) and (2) one might correctly generalise (3), but overgeneralise to incorrectly infer (4). If we look at each sentence and examine their meanings, we can paraphrase them as:

(1b) She painted a house that happened to be green.
(2b) She caused the house to become green by painting.
(3b) She knew a house that happened to be green.

(4b) ?She caused the house to become green by knowing.

It is clear that the unacceptability of (4) and (4b) becomes obvious with the benefit of adult-state semantic competence and a rich network of ‘world-knowledge’.

Understanding exactly how we acquire and deploy this knowledge in the way that we do is a question currently being addressed by many developmental psychologists. The difficulty of their task will be familiar to those in Artificial Intelligence who have endeavoured to capture these facts in formal semantics. ‘Common sense is not a simple thing. Instead, it is an immense society of hard-earned practical ideas – of multitudes of life-learned rules and exceptions, dispositions and tendencies, balances and checks’ (Minsky, 1987: p 22).

7.4 Language-internal and cross-linguistic generalisations

As Ibbotson (2007; http://iascl.talkbank.org/bulletins/bulletinV27N2.html#report) discusses:

‘If there is no universal grammar why are languages the way they are? Goldberg points out, as we have already noted, that there is very little agreement amongst linguists over what constitute language universals and, in instances of stated universals, that these are tendencies. Instead, she argues that what look like linguistic universals or tendencies are in fact underlying universals of human cognition. Of particular importance in the context of language use and its acquisition are the principles of cooperative communication, processing constraints, and attentional biases.

Many syntacticians are sympathetic to the usage-based, cognitive linguistics approach but some (mainly from the more formal approaches) are
still unclear…if there is no universal grammar, what then, determines the set of well-formed sentences or unacceptable constructions? Ross (1967) developed the notion of syntactic ‘islands’ whose constituents cannot be questioned or extracted. For example, sentential subjects, as in *That she knew Bob bothered John*, do not allow questioning of their internal object *Who did that she knew bother John?* Goldberg’s discourse-based account of island constraints is an attempt to work through an example of when constructions may come into conflict. In the spirit of Goldberg’s earlier work *Constructions* (1995), phenomena that were thought to be purely syntactic (and classic generativist territory) receive a functional analysis. The ‘backgrounded constructions are islands hypothesis’ states that the well-formedness is determined by the compatibility of the constructions that are combined, specifically, constructions that require that a constituent be discourse-prominent cannot combine with constructions that require that the constituent be backgrounded in discourse. Backgrounded constituents are neither the primary topic nor part of the focus domain. Elements involved in unbounded dependencies are positioned in discourse prominent slots, thus it is pragmatically anomalous to treat an element as both backgrounded and discourse prominent. So in the sentence introduced above, *that she knew Bob bothered John*, the proposition *that she knew Bob* is asserted or given information and is unavailable for extraction or questioning. Contrast this with the Principles and Parameters account of the same phenomena that states that the well-formedness of wh-questions is determined by the syntactic trees involved, i.e. the wh-phrase and the gap cannot be separated by more than one Sentence or Noun-Phrase node (a condition known as “subjacency”). Using a
pragmatic/information structure account predicts a range of subtle facts that are difficult to account for with binary parameters, such as the shades of acceptability judgements in island constructions and the fact that previous discourse can reverse unacceptability judgements (Ambridge & Goldberg, 2008). Approaching the problem from this perspective seems to buy more explanatory and descriptive power while satisfying the usage-based desideratum of psychological plausibility – figure-ground being motivated as a domain general psychological process”.

A different, but not necessarily incompatible, explanation for restrictions on questions and other questions with long distance dependencies is proposed by Dąbrowska (2008). The lexical template hypothesis begins with the observation that the questions with long-distance dependencies (LLD) that people actually use tend to be quite stereotypical (e.g. What do you think you’re doing?) and rather different from the textbook examples linguists have traditionally worried about (e.g., what does the lawyer suppose that the policeman believes that Peter stole?). Incidentally, not looking in detail at the sentences people spontaneously use may create its own problem when the intuitions of linguistic experts start to diverge with the vast majority of ‘non-expert’ speakers the longer and harder one studies such examples (see Dąbrowska, 2009 for discussion). To return to LLDs, in an examination of attested corpus uses it was found that the matrix clause usually consists of a WH word, the auxiliary do or did, the pronoun you, and the verb think or say, with no other elements; and they virtually never contain more than one subordinate clause. This has lead usage based theorists such as Dąbrowska (see also Verhagen 2005) to hypothesise that ‘speakers’ knowledge about such constructions is best explained in
terms of relatively specific, lexical templates rather than general rules that apply
“across the board”. Prototypical LDD questions are produced simply by inserting new
material into the appropriate slots in a pre-existing template. Non-prototypical LDD
questions such as *What does she hope she’ll get?* require additional cognitive effort,
since the speaker has to adapt the template – in this case, substitute *she* for *you* and
*hope* for *think*, and modify the auxiliary so that it agrees with the subject”
(Dąbrowska, 2008: pp392). In a sentence acceptability test, “prototypical instances of
this construction, i.e. those which fit one of the templates postulated on the basis of
corpus research (*WH do you think S-GAP?*, *WH did you say S-GAP?*) were judged to
be the most acceptable of all sentences. Departures from the prototype (use of a
different auxiliary or verb in the matrix clause, addition of a complementizer or an
extra complement clause) resulted in lower acceptability ratings” (Dąbrowska, 2008:
pp392).

Dąbrowska has also brought to attention the variation in intuitions between
speakers of the same language in rating the acceptability of LLD questions and for
other grammatical constructions such as the passive (Dąbrowska and Street, 2006; see
Dąbrowska, submitted, for a review of individual differences in native language
attainment). This raises interesting questions about the extent to which speakers
converge on the same adult grammar; an implicit assumption held by very different
theoretical approaches in linguistic, psycholinguistic and developmental work. One
way to think about the issue is like this. The shaded area below in Figure 5 represents
the idiolect of one speaker of a language community; the content of the area is the
sum uses of their grammar, pronunciation, vocabulary and so on.
Figure 5. *The idiolect of a speaker.*

The relationship between two members of that speech community is shown below in Figure 5. There is not a perfect overlap – there is inter-speaker variation in pronunciation, vocabulary and grammatical competence – but the variation is clearly not arbitrary either, that is why two speakers can be thought of as belonging to the same speech community. The shared intersection of the two idiolects (the hashed area) are those conventions and cultural norms that say ‘if you want to express such-and-such a thought, in this community we do it this way’, for example if you want to talk about a transitive event in English, the speaker orders the referents of the transitive scene in this way to communicate who-did-what-to-whom.

Figure 6. *The idiolects of two speakers.* *The intersection represents shared ways of using language:* words, argument-structure constructions, agreement, case and so on.
What role does the frequency of constructions, lexical material and usage events play in all of this? All the content of a particular idiolect are shared in some sense (individual items are used to communicate with at least one other member) but some linguistic item are more shared than others (active versus passive voice for example). Figure 6 shows a distributional pattern plotted in three dimensions and approximates the overall frequency of constructions or words in a given language; the more frequent the use the higher the area is on the graph.

Figure 7. *Frequency distribution of linguistic items.*

If we overlay this distribution over the idiolects of particular speakers (Figure 7) then we see the interaction between frequency and inter-speaker variability. The variation between members of the same speech community is greatest at the tails of the distributions and more homogeneous towards the centre, that is, repeated use
entrenches the convention and infrequent uses result in more variable representations across the community.

Figure 8. Two speakers’ linguistic knowledge in comparison with the usage-frequency of that knowledge.

Table 1 below ranks the frequencies of various grammatical patterns from three different corpora (adapted from Roland, Dick & Elman, 2007). The usage-based prediction is that the between-speaker variation in intuitions (as shown by grammaticality judgements for example) will be correlated with this ranked frequency, such that there will be greater heterogeneity the lower down the rank list. Examining distributions of this type in combination with tests of grammatical competence could help determine the extent to which learners converge in adulthood on the same grammar.
Table 1. Ranked frequency distributions of constructions from several English Corpora.

<table>
<thead>
<tr>
<th>Construction Type</th>
<th>British National Corpus</th>
<th>British National Corpus Spoken</th>
<th>Brown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple Transitive</td>
<td>30%</td>
<td>31%</td>
<td>32%</td>
</tr>
<tr>
<td>Prepositional Phrase</td>
<td>17%</td>
<td>14%</td>
<td>18%</td>
</tr>
<tr>
<td>Simple Intransitive</td>
<td>11%</td>
<td>13%</td>
<td>15%</td>
</tr>
<tr>
<td>Passive</td>
<td>9%</td>
<td>9%</td>
<td>11%</td>
</tr>
<tr>
<td>Transitive + Prepositional Phrase</td>
<td>7%</td>
<td>6%</td>
<td>8%</td>
</tr>
<tr>
<td>To Infinitive Verb Phrase</td>
<td>6%</td>
<td>6%</td>
<td>5%</td>
</tr>
<tr>
<td>WH Clause</td>
<td>4%</td>
<td>5%</td>
<td>3%</td>
</tr>
<tr>
<td>Sentential Complement (no Complementizer)</td>
<td>4%</td>
<td>5%</td>
<td>2%</td>
</tr>
<tr>
<td>Sentential Complement with Complementizer</td>
<td>3%</td>
<td>3%</td>
<td>1%</td>
</tr>
<tr>
<td>Perception Complement</td>
<td>3%</td>
<td>3%</td>
<td>1%</td>
</tr>
<tr>
<td>Transitive + WH clause</td>
<td>2%</td>
<td>2%</td>
<td>1%</td>
</tr>
<tr>
<td>Gerundive Verb Phrase</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>Ditransitive</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>Transitive + To Infinitive Verb Phrase</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>Transitive + Sentential Complement (no Complementizer)</td>
<td>1%</td>
<td>1%</td>
<td>0%</td>
</tr>
<tr>
<td>Prepositional Phrase + To Infinitive Verb Phrase</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Transitive + Sentential Complement with complementizer</td>
<td>0%</td>
<td>Gerundive Verb Phrase</td>
<td>0%</td>
</tr>
<tr>
<td>------------------------------------------------------</td>
<td>----</td>
<td>-----------------------</td>
<td>----</td>
</tr>
<tr>
<td>Transitive + Sentential Complement (no complementizer)</td>
<td>0%</td>
<td>Transitive + Sentential Complement with complementizer</td>
<td>0%</td>
</tr>
<tr>
<td>TOTAL 100%</td>
<td>TOTAL 100%</td>
<td>TOTAL 100%</td>
<td>TOTAL 100%</td>
</tr>
</tbody>
</table>

8. Future challenges

As usage-based theorists, Goldberg and Dąbrowska use parsimony to point out that what can be achieved with existing cognitive and social learning skills doesn't need to appeal to ‘additional’ mental machinery such as a Language Acquisition Device. Attributing language acquisition to domain-general abilities, such as pattern finding and intention reading, also usefully sidesteps universal grammar-specific problems; for example linking universal grammar to cross-linguistic diversity and the problem of developmental change given the constancy of universal grammar. Goldberg acknowledges the challenges for the constructionist account of language. Two issues which are in particular need of further research are a) the need to specify in greater detail the learning mechanisms that underpin generalisations and b) defining the nature of construction interaction in relation to preferred argument structure and discourse-pragmatics.

Matthews and Tomasello (in press) note that those who argue grammar is not learnable must articulate what they propose children know innately and how this combines with the children’s experience of particular languages in a way that yields the developments we observe. Those who argue that children could plausibly learn
Thus far, this chapter has laid out why formalism has not provided a satisfactory answer to the question of language acquisition and has given a brief introduction to the most well-thought-out alternative of the functional approach, exemplified by the cognitive construction grammar of Goldberg. Ultimately, one would hope that the future of the debate rests on the empirical evidence. This raises the question what kind of empirical evidence could decide between the two positions? Under the nativist account, innate knowledge of abstract linguistic categories, emerges gradually throughout development as it is revealed by decreasing restrictions on performance (leaving aside the issue that no-one really agrees on what would be in universal grammar to begin with). The problem for child language acquisition research is that the developmental trajectory looks very similar under a usage-based view where both competence and performance develop together (in as much as the division between competence and performance exists in the usage-based approach). We are left in a position where we resort to the meta-theoretical stance of parsimony. We might not be able to directly disprove the alternative hypothesis, rather, we can force it onto ground where it needs so many internal assumptions and patches that it a) dilutes the power of the original idea to such an extent that it no longer seems reasonable and b) the alternative explanation is less ad hoc and independently supported from what we know about the mind from research in other domains.

9. **Beyond dichotomies**

At the start of this introduction I said that framing the question in terms of innate versus learnt is unhelpful (as we have seen above) and potentially three things could
have gone wrong here; no-one has thought of the right experiment yet that will
distinguish between the two major approaches, the division as it stands is ill-posed, or
both. In line with other developmental theorists, I argue a better question than ‘is a
particular ability innate or is it learned?’ is ‘over what time period does the
structure/function emerge?’ In this way we might stand a better chance of achieving
‘both-and’ thinking rather than ‘either-or’ dualism, the later of which seems to be the
modus operandi in cognitive science (structure vs. function, nature vs. nurture, brain
vs. behaviour, perception vs. cognition, mind vs. body, competence vs.
performance…).

If we pursue the emergentist approach, we can ask what are some of the most
influential sources of complexity in studying language development? To start with, if
we want to explain language’s structure, systematicity and function we must
acknowledge that organisation emerges over a timescale that involves several
different orders of magnitude: The hundreds of thousands of years of the evolution of
cooperative communication in our species, the thousands and hundreds of years of the
evolution of language (e.g. the grammaticalisation from Old English SOV to Modern
English SVO), the tens of years that an individual builds a vocabulary, the years of
morpho-syntactic development, the minute-to-minute and second-to-second decisions
of discourse pragmatics and information structure (Figure 9).
Figure 9. The vastly different timescales over which linguistic structure, systematicity and function emerge.

- Evolution of Cooperative Communication in Humans e.g. requesting, sharing, informing
- Evolution of different languages e.g. grammaticalisation
- Lexical learning
- Morpho-syntactic learning
- Information structure/Discourse pragmatics (e.g. topic, focus)
- Unified at the moment of comprehension and production
If we take grammatical gender and some richly inflected case systems as an example, some of these features are difficult to explain as a straightforward mapping between form and function. One needs to consider different magnitudes of time-scale in the context of the frequency of use of these forms in the language as well as studying how these phonological forms can become detached from their semantic pole and ‘fossilised’ in a language over time.

Examples of dynamic self-organising systems that show emergent, adaptive complexity are not difficult to find in nature, these include: language evolution, the hexagonal structure of honey comb, weather and climate systems, ant colonies, growth of cities, brain development, the immune system, length of queues at supermarket, the rock weathering thermostat that controls the CO\textsubscript{2} in the atmosphere, embryology, traffic jams, EBay and the behaviour of slime mould (\textit{Physarum polycephalum}) which, impressively, can be trained to find the shortest route through a maze (Nakagaki, 2000). What unites these different phenomena is a network of neighbour interaction, pattern recognition, feedback, and indirect control (Johnson, 2001). There is a movement from low-level rules to higher-level sophistication that is adaptive for the system as a whole without anyone or thing (e.g. universal grammar) being in charge. Dawkins, discussing the flocking behaviour of starlings (and making an analogy with embryology) puts it like this:

“The key point is that there is no choreographer and no leader. Order, organisation, structure – these all \textit{emerge} as by-products of rules acting \textit{locally} and many times over, not globally”. (Dawkins 2009: p. 220)
All of these structures can be explained, not by genes for supermarket queue length or grammar for that matter, but by more basic human goals and desires, and physical properties of the world – e.g. people will self-sort themselves by going to the shortest queue. Emergentist accounts often emphasize the extent to which a complex set of behaviours and forms can arise from a few simple mechanisms. Some examples, discussed in MacWhinney (2005; pp.10-11), of general cognitive mechanisms that exhibit this behaviour include:

1. “Learning through error propagation, as in the back-propagation algorithm (Rumelhart and McClelland 1986).


3. Statistical learning that allows both children and adults to learn patterns of sounds, tones, or visual forms (Jusczyk 1997; Saffran et al. 1996), as constituted in sequential groups (Gupta and MacWhinney 1997; Houghton 1990).

4. Analogy, which allows problem-solvers and learners to structure domains and relations in parallel alignment (Gentner & Markman, 1997).

5. Consolidation of short-term memories into long-term memories (McClelland et al. 1995).”

One can almost hear the generativist incredulity that such an account will ever reach descriptive adequacy when it comes to explaining the subtleties of grammatical patterns. The challenge therefore for usage-based accounts is to develop the descriptive tools and explain the learning mechanisms with enough precision and
detail as to be a) falsifiable or at least more falsifiable than UG – one practical implication of this is that there should be, in theory, some outcomes of experiments that the theory cannot explain; b) capture the phenomena, i.e. it has descriptive adequacy, and c) genuinely predictive of new findings, otherwise the explanations will always have a post hoc feel.

10. Synopsis of the thesis

For the time being, if we are going to conceptualise language acquisition as a developing inventory of constructions that are more-or-less schematic – as usage-based accounts do – then as a starting point we need to show how, in principle, grammatical constructions are learnable and what these representations might look like. Chapter 2 examines this question by looking at one construction in detail, the caused-motion construction. Chapter 3 further investigates the topic of how infants construct grammatical constructions by taking a cross-linguistic look at the transitive construction in the context of a prototype theory of categorisation. Chapters 2 and 3 are theoretical reviews that build on the usage-based approach to language acquisition and generate predictions for the subsequent empirical studies. Chapter 4 applies the theoretical investigations of chapter 3 in an empirical experiment, specifically, making developmental comparisons of the prototypical semantics of the transitive construction in English. Chapter 5 considers further the role of different cues in children’s understanding of argument structure by presenting an experimental study examining the role of pronoun frames in early comprehension of transitive constructions in English. Chapter 6 focuses on how infants and mothers actually use language in a corpus study. It begins by looking at the role of skewed distribution and cognitive anchoring in schematising the Subject Verb Object construction in English.
It then presents a usage-based acquisition model of argument productivity in subject-verb-object constructions. Finally, the thesis concludes drawing together the various outcomes of the work into an attempt at a coherent synthesis and indicating directions for future work.
CHAPTER TWO

Publication 1

Learning argument-structure constructions: a usage-based approach

Page: 73
Learning argument-structure constructions: a usage-based approach

Abstract

By focusing on one construction – the caused-motion construction – the aim of this paper is to show how, in principle, form-function parings are learnable. The suggestion is that the co-occurrence of events and utterances represents a grounded way-in for infants to learn the function of argument-structure constructions. By taking a closer look at the processes of schematisation, each shift on the continuum from input to abstraction is justified with respect to the cognitive abilities of the infant and the properties of the input at that phase when the shift happens. Especially important in this explanation will be the fact that humans possess a unique set of social-cognitive and social motivational-skills that allows language to happen. Furthermore, early linguistic categories are formed around the underlying functional core of concepts and on the basis of their communicative discourse function. This, combined with powerful pattern-detection skills, enables distributional regularities in the input to be paired with what the speakers intend to communicate.

Keywords

Language acquisition, schematisation, event-structure, argument-constructions
1. Introduction

*Putting together novel expressions is something that speakers do, not grammars. It is a problem-solving activity that demands a constructive effort and occurs when linguistic convention is put to use in specific circumstances.*

(Langacker, 1987: pp. 65)

What do usage-based theories mean when they characterise language acquisition as a developing inventory of constructions that are more-or-less schematic? One of the main challenges to any usage-based approach is to specify in greater detail the learning mechanisms that underpin generalisations. Therefore the aim of this paper is to make the usage-based approach to learning more explicit and by focusing on one construction I want to show how, in principle, form-function pairings are learnable. In order to do so requires bringing together findings from categorisation and analogy, social cognition and construction grammar in a synthesis that is best described as developmental cognitive linguistics. Specifically, I hope to explain how categorisation and analogy can be unified, explicate the role of analogy in the construal of events, and integrate social reasoning as part of grammar learning process.

As the most historically controversial issue is whether infants can ever get to an adult grammar inductively, each shift on the continuum from input to abstraction needs to be carefully justified with respect to the cognitive abilities of the infant and the properties of the input at that phase when the shift happens. I will use the example of the caused-motion construction, the meaning of which can be paraphrased as X causes Y to move to Z, or more formally where a subject works as an agent of the caused-motion event, the direct object as a moved or a moving thing, and the oblique object within the preposition phrase as a goal, for example “Frank sneezed the tissue
off the table” (Goldberg, 1995). As a geneticist would justify using Drosophila at the expense of studying other species, so I hope to show how learning argument-structure constructions could work in a simplified, isolable and frequent instance. I begin by sketching out the broader context in which this acquisition happens.

2. The context in which acquisition takes place

A popular question, historically-speaking, among psycholinguists has been to ask “what is the relationship between language and cognition?” The approach taken here begins with a slightly different assumption. As Langacker’s quote above suggests, it might be more productive to recognise language as a kind of cognition with a particular job to do – the kind of cognition that is put to use when a speaker intends to communicate a targeted conception. We can think about grammar as a shared repository of different ways of packaging these concepts. The language learner’s task is to become an expert in those grammatical conventions which they are surrounded by, so that they can share and make use of this repository in communication. Any given language then is one solution to the problem of communicating concepts with other minds and the possible kinds of things human language can be is best described with reference to these communicative needs.

To be clearer on what is meant by a ‘developing inventory of constructions’ involves taking the development part of this proposition seriously. How a child learns $x$ is inseparable from what $x$ is, and a comprehensive theory of language structure, function and organisation (i.e. what $x$ is) needs room for not only for development within the individual but diachronic development as well. The usage-based model is particularly well-suited to addressing this as it draws on a unified set of concepts to account for both acquisition and change. Domain-general cognitive principles and
general communicative norms (§3 below) define the kinds of things that languages can be. For example, the unique trade-off that any given language arrives at between what conceptual structure is grammaticalised and what communicative content is left to social cognition to cash out, is best explained in terms of the developmental history of that language. Typological variation can be thought of as the result of cultural drift within this space of possibilities: on the one hand, high fidelity transmission of cultural conventions between generations is powered by imitation and learning – saving us from literally re-inventing the wheel every generation; on the other hand a small amount of linguistic innovations (mutations) can enter a language depending on a number of criteria, two of which are that a novel form must be useable and learnable if it is to survive beyond a generation and spread throughout the speech community. Under this account typological similarities emerge as the result of a kind of convergent cultural evolution, where different language communities (mutually unintelligible systems) have arrived at similar communicative solutions – e.g., pronouns, word-order, morphology, anaphora – to similar communicative problems – e.g., sharing, informing and requesting. Setting language acquisition in a diachronic context buys some room for manoeuvre, particularly for cognitive linguistics, as it unburdens theory from having to explain all of language’s systematic and organisational features on a purely symbolic account, for example some linguistic features can be fossilised structures.

---

1 This can be thought of as a branching tree evolving through a multi-dimensional space. The tree represents language evolving over time and the space is defined by the possible things language can be for humans, beyond which communication is either too costly or impossible. The closer any language is to the edge of a particular dimension – the more extreme it has evolved in one particular direction – the more one can expect a greater reliance on other (linguistic) strategies to compensate in the opposing direction such that the net communicative package remains stable. For example, the freer the word-order is, the greater reliance on morphology; as the role of syntax is reduced, the greater the reliance on contextual support. It is because of this, typological statistical tendencies emerge of the type ‘if a language has x it is also likely to have y’.
3. **Learning language concepts**

It needs to be stated from the outset that, crucially, where language differs from categorising and schematising the world on the basis of perceptual stimuli alone is that associations between utterances and events are made within the arena of cooperative norms and reasoning, established between a communicator and recipient. Other authors have pointed out that possessing a powerful set of social-cognitive and social-motivational tools is logically prior – both in the development of the species and the development of the individual – to recipients treating utterances as language and not just a string of sounds (e.g. Tomasello, 2008). This characterisation of the ‘top-down’ psychology is crucial if there is going to be a convincing account of how grammar is constructed ‘bottom-up’ from usage events.

There are important implications of this view of cognitive organisation for learning in general and learning abstract grammatical structures in particular. For a start, if an arena of communicative cooperation has been established, the default psychological status is that a communicator’s utterances will have referential intention; a communicator wants the recipient to attend to something: an action, an object, some aspect of the scene, the speaker’s attitude towards a scene or a proposition or whatever. Secondly, the communicator is also likely to have a social motive for doing this: they want the recipient to do something, to feel something, to know something or to share something. Finally, in ongoing discourse these communicative acts are modified by what the speaker-recipient know together, the common ground and joint attention that they have established over the course of their communicative history.

To have this in mind when looking at the schematization of grammatical categories is important as many of the claims about the processes involved here –
statistical learning, pattern recognition and implicit categorisation for example – are likely to be true of other species. But once our view of language, and more generally of communication, is one of a social act, it is possible to say why domain-general cognitive processes and even species-general processes might be necessary but not sufficient for language to evolve.

4. Analogy and Categorisation

Examples of organising experience into clusters of things that look the same or do the same thing can be found everywhere: from the relatively abstract generalisations of grammar to more prosaic examples; when a six-year-old realises that tulips must need water, because people do (Inagaki & Hatano, 1987). The two things being compared here or ‘analogised’ fall under a category of things that behave in a particular way (things that need water), a category/concept which may or may not be lexicalised in the language. At first glance one might want to draw a distinction between processes of analogy and categorisation, and indeed they are often talked about in different terms in the cognitive literature. There seems to be a difference between categorisation – typically depicted as ‘vertical’ relationships in a taxonomic hierarchy where X and Y are instances of Z and Z is schematic of X and Y, and analogy – typically ‘horizontal’ relationships that do not involve a literal re-labelling of what X is an instance of, rather, the operation is more like X is like Y in way Z. However, I would like to suggest there is no principled cognitive distinction between these processes assuming that (linguistic) knowledge is organised in something like a structured inventory of schema-instance relationships. For example, in analogy, if X is

---

2 Inagaki and Hatano also reported that 40% of kindergarteners believe that a tulip can “feel happy” and 72% believe that a tulip can “feel pretty”.

79
like Y in a way Z, then Z can be considered schematic of X and Y and if Z is schematic of X and Y then these have been treated as instances of the category Z. Although there maybe pragmatic differences between analogy and categorisation, such as their explicitness in the discourse, in terms of the output of the cognitive operations (X is like Y in way Z) I treat them here as equivalent. One relevant observation here is that people often find comparisons much more similar in one direction than the other (Tversky, 1977). For example we prefer “a scanner is like a copy machine” to “a copy machine is like a scanner”. As Tversky points out, this directionality seems odd if one sees similarity as a symmetrical relationship; if $a$ is similar to $b$ then $b$ should be equally similar to $a$. Viewing categorisation, analogy and similarity as different manifestations of the same underlying cognitive process (pattern finding) provides a natural explanation: scanners are but one kind of copy machines; the inferences in this case are derived in a particular direction from the schema to the instance not the other way around. So for the purposes of describing grammatical category learning I use the terms categorisation and analogy somewhat interchangeably here as I assume they are both subject to the same general cognitive processes of schematisation that can operate at all levels of the representational continuum. In terms of construction grammar this means at all levels of the symbolic assembly from form to function.

5. **Bootstrapping meaning**

Something that has fur like a dog, smells like a dog and barks like a dog, probably is a dog. But in order to work out the function of a grammatical category, a system based entirely on the perceptual similarity of the utterances themselves will obviously not work. *The cat ate the monkey*, and *a goat hit a woman* are analogous in grammatical
function – they share an interpretation of who-did-what-to-whom on the basis of subject, verb and object – yet there are no perceptual surface features that are shared between these sentences that could be reliably mapped onto function. Indeed the same sentence *the cat ate the monkey* can be said by the same person with completely different speech stream characteristics to communicate the same proposition.

So at some point the learner needs a way into function and a reasonable starting point for a usage-based theory of verb argument acquisition is to look at those moments where a perceptual event might become associated with an utterance. At the bottom of Figure 1 there are simplified scenes that co-occur with utterances, linked by a solid line in the diagram. At the risk of taking the construction metaphor too far, the events that take place in a communicative context are the ‘ground floor’ of the symbolic assembly, and are especially important in learning early on in development. The events in question all share something in common, shown in the diagram as a dashed line; the participants behave in a similar way, that is, their actions are intended to accomplish something similar. Thus right from the start intention-reading skills provide important cognitive foundations on which later abstractions are formed, including those that will go on to become grammatical schemas. Key to the argument is the proposal that the association between form and function takes place within the norms of cooperative communication and in the diagram these co-occurrences are linked to blobs on a two-dimensional plane with each blob representing a perceptually grounded usage event. I am going to focus on what happens to the part of this plane that contains the contingencies for a particular kind of event, a caused motion event.
6. Abstracting over instances

Suppose a man is drowning in the sea, what an object is made of – and thus whether it floats or not – will be of greater relevance to his goals at that time than what it tastes like, and if he then makes it to a desert island, what something tastes like – and thus whether it is edible or not – will be of greater importance to him than whether it floats. Children face a similar problem when constructing grammatical schemas\(^3\), namely, which correspondences are important and for which categories?

---

\(^3\) Whether one talks of Rumelhart’s (1977) concept of “schemata”, Fillmore’s (1968) “case frame”, Minsky’s (1981) “frames,” Schank and Abelson’s (1977) “scripts”, Lakoff’s (1987) “idealised cognitive models” or “cognitive templates” is not crucial to the argument here in the sense that what they share in common – pattern finding models used to structure and make sense of our experience – is more interesting than what differentiates them.
Firstly, as preferential looking studies have shown, the perceptual system of infants can extract category-specific clusters of exemplars from a familiarisation phase so that when subsequently tested infants have learned to differentiate members of familiar and novel categories (Quinn, 1993; 2000). One important point to note here is that many studies have emphasised that the type variation and the token variation and the ratio between them have important implications for the character of the resulting taxonomy of that category (Casenhiser & Goldberg, 2005; Goldberg, Casenhiser, & Sethuraman, 2004; Rovee-Collier & Boller, 1995). There seems to be a ‘Goldilocks’ zone where there is neither too much nor too little variation but ‘just the right amount’ of variation in the learning set (maybe a skewed type/token distribution multiplied by overall frequency of learning opportunities) that facilitates category acquisition and thus generalisation. For example, skewed distributions have been shown to have a beneficial role in learning linguistic and non-linguistic categories (Avrahami et al., 1997; Casenhiser & Goldberg, 2005; Elio and Anderson, 1984; Genter, Loewenstein, and Hung, 2002). Cognitive anchoring is one mechanism that has been suggested to play a role in this effect. According to the theory a high frequency exemplar which emerges early in learning provides a salient standard of comparison against which subsequent instances are judged. This statistical skew in the input facilitates relatively rapid category construction compared to flatter distributions because a minority of high frequency types provide a relevant anchor that serves to organise memory and reasoning about other types. Similar skewed distributions have been reported in Child Directed Speech for English (Naigles & Hoff-Ginsberg, 1995; Theakston, Lieven, Pine & Rowland, 2004) and for Hebrew (Ninio, 1999a; 1999b).

Secondly, studies have focused on categories and induction in young children and have shown that infants are able to flexibly adjust the attentional weights based
on the psychophysical properties of the context (e.g., Deak & Bauer, 1996; Gelman & Markman, 1986; Mak & Vera, 1999). At one level the idea of attentional weights is just another facet of habituation or salience, for example, when an individual has been habituated to a baseline noise level which is consistently high in a room, a moment of silence becomes salient (it is foregrounded); conversely, when silence is the baseline (which is habituated and backgrounded), attention will be directed towards a sudden noise. A slightly more sophisticated example shows this process can keep track of multiple contingencies; when there is a red triangle and a red circle, shape is a more salient property, and when there is a blue triangle and a red triangle, colour is a more salient property. The key point for the argument here is that attentional bias towards a salient property is particularly predisposed to being reinforced throughout development if it is diagnostic in partitioning a set of entities according to how they behave or what they are for— their function. Put very generally, human concepts are defined by their function in a more encompassing conceptual system adapted to patterns of events and socio-cultural activities in the human world. The implications for development are that children’s understanding of what one can do with an object and what an object can do forms the functional core of the concept (Nelson, 1974; 1996). This view has been supported by the theoretical work of Mandler (e.g., 2000) who argues that what something is, is defined most fundamentally by what it does, with perceptual features serving a somewhat ancillary or identifying role only.

Thirdly, and importantly for putting together novel expressions, infants can detect regularities and thus abstract over those, but this is dependent on the character of what they are trying to learn, in short, the more abstract the pattern or correlation that one needs to detect, the more evidence one needs. For example, suppose an infant, over time, is exposed a variety of objects that are a variety of colours: most of
the balls are red, most of the cars are yellow, most of the tables are white and most of the birds are green. In the beginning (Table 1a), local contingencies emerge as a result of the distributions of sampling, for example most of balls encountered in the life of that child up until this point are red. As the corpus of evidence grows (Table 1b) and as a function of the within- and between-category variance, the opportunity to abstract regularities arises, and this abstract regularity, in turn, makes shape and colour a more important predictor of category membership. One could imagine further levels of abstraction in Table 1a/b, for example, if the more schematic category ‘wavelength’ were in some sense predictive of the way shapes behave or vice versa. Before this second tier could be extracted above colour we might need to add into the matrix, microwaves, x-rays and gamma rays. Colour would then be an instance of wavelength (the visible spectrum) as would the other wavelengths (non-visible spectrum) with each of these categories being schematic of those instances they dominate in the hierarchy – red, green, microwaves etc. The hypothesis is that the ‘depth’ of the abstraction, represented in construction grammar as the taxonomic hierarchy, is proportional to the size of the matrix, which in the case of learning grammar means the size of the corpus of utterances. The usage-based prediction for learning is that the more abstract the schema (whether that be ‘wavelength’ or ‘caused motion’) the more instances one needs to construct it. This refers to the acquisition of a fully productive schema (as opposed to those at the idiomatic end of the spectrum), where a speaker converges with other speakers on the instances that schema can license and the conventions that define its relationship to other schemas.
Table 1a. In the beginning local contingencies emerge first. Table 1b. Only later do opportunities for abstraction occur as a function of the distributions and raw frequency (example adapted from Sloutsky, 2003)

The implications for language are that as the infant’s representations become richer they can detect more abstract correlations between functional dimensions and use these correlations in the development of categories. This internal feedback mechanism is a major source of complexity for any theory trying to explain the development of language-learning as the system begins internal re-descriptions and self-organisation on the basis of what a speaker knows (which itself is a shifting target). Emergent phenomena such as this require nonlinear interaction among components of a system undergoing qualitative change, leading some cognitive scientists to seek insights from complexity science to explain language development (van Geert, 1991; Thelen & Smith, 1996; Ninio, 2006).
In the usage-based view of language acquisition, children first acquire a number of low-scope verb schemas (e.g., X hits Y, X kisses Y, X pushes Y, X pulls Y, etc.), then by forming analogies between the roles that participants are playing in these events, these item-specific constructions eventually coalesce into a verb-general abstract construction (Tomasello 1992, 2003). Thus the process of verb-specific, ‘local’ knowledge being re-described and re-categorised at a later stage by more abstract ‘global’ schemas is equivalent – in representational redescriptional terms (Karmiloff-Smith, 1979; 1992) – to knowledge being initially in the system, but not yet available to the system, a shift that sees linguistic knowledge moving along a implicit/explicit continuum.

There is some point when the input detaches itself from the event and concept-to-concept mapping (analogies of conceptual structure) bootstraps new levels of abstraction and analogy. This flexibility of analogy-making is extremely impressive for many reasons, two of which are (i) it takes place at multiple levels of representation – from surface generalisations to functional generalisations and everything in between, and as a result of this (ii) the things being compared need not share any functional characteristics (a blue dog and a blue pencil are both blue) nor any perceptual characteristics (a blue taxi and a red boat are both means of transport). Thus analogies made on the basis of how things behave can be made over vastly different scales and domains (Genter, Ratterman, & Forbus, 1993; Holyoak & Koh, 1987); the orbiting of the planets around the sun is like the orbiting of electrons around a nucleus. Gentner and Markman (1997: p 48) observe “The contrast between analogy and literal similarity is in fact a continuum not a dichotomy. Yet it is an important continuum psychologically, because overall similarity comparisons are far easier to notice and map than purely analogical comparisons”. Perhaps this is because,
at the start of learning, perceptual similarity is all the infant has to work with until things start to do something or until people start to use something, only then can infants begin categorising artefacts (including language) on the basis of functional as well surface properties. Developmentally, if we assume that the more ‘abstracted-over’, conceptual knowledge is organised more slowly than perceptual information for this reason, then the view would account for the characteristic-to-defining shift (Keil & Batterman, 1984) in children’s schemas from local features (e.g., a postman = that nice man with the beard who puts letters through my front door) to deeper relational commonalities (a postman = any person who delivers mail (including that nice man with the beard)).

7. Linguistic and non-linguistic cues to argument-structure constructions

The number of referents realised in speech is certainly one cue to assigning the semantic roles of argument structure (among many other competing cues) but this can’t be sufficient for many reasons, for example (i) There is no one-to-one mapping of the number of arguments to the semantic roles; for example in English the active-passive voice, agent-patient assignment is reversed yet there are obviously still only two role slots to be filled (e.g., she kicked him vs. he was kicked by her) or conjoined intransitives and regular transitives (e.g., Mary and John slept vs. Mary tickled John) (ii) in English, and many other languages, speakers can also drop the subject or object if these elements are given information in the discourse and (iii) in null-argument languages it would be very misleading to assume that, for example, one argument equals intransitive predicate where a sentence like “jane blinked” could be a transitive

4 Of course the child also needs to have reached some threshold of social-cognitive competence in order to schematise the behaviour of others (intention-reading).
with an omitted object or “blicked it” could also be a transitive with an omitted subject.

However, there is evidence that, like adults, children are able to exploit the referential context of the scene (what I have called the ground floor of the construction in learning), particularly the apparent roles of the scene’s participants and moreover they can use this information to constrain the assignment of semantic roles to the entities mentioned in the utterance. (e.g., Chapman & Miller, 1975; Clark, 1973; Huttenlocher, 1974; Knoeferle, Crocker, Scheepers, 2005; Shatz, 1978).

8. Event Construal

How events are construed in mental space (the conceptualisation) presumably bears some relation to how they are construed in perceptual space (the simplified scenes at the bottom of Figure 1). The suggestion is that the embodied form of the event representation or construal is redescribed forms of the perceptual scenes (the starting point for utterance-event associations) tagged with the psychological characteristics of the actors involved (Figure 2). Argument-structures can then start to be differentiated into different families of use on the basis of these properties. At no point along this continuum of descriptions and re-descriptions are there the meaningless rules posited by more formal accounts interacting with meaningful words, rather, at some point children begin make categorisation decisions based on the relationship between participants in an event – for the caused-motion construction this means they realise this particular kind of utterance is an example of a convention that is put to use where the intention is to communicate ‘X causes Y to move to Z’, this is schematisation based on ‘things are what they do’.
Figure 2 shows how the functional alignment of arguments would work for some utterances that fall under a general caused-motion schema, elaborated from Lakoff and Johnson’s source-path-goal schema (1999: 32). There is good evidence that the alignment of relational structure and mapping between representations is a fundamental psychological process underlying analogy and similarity (Genter & Markman, 1993; 1994; 1995; Goldstone 1994; Goldstone & Medin 1994; Goldstone, Medin & Genter, 1991). One manifestation of this domain-general process allows inter-construction mappings and is shown by the vertical dotted lines shown in the diagram. They are analogies/categorisations in the sense that X is like Y in way Z: utterance a) is like utterance b) in that they construe the agent, thing and path in similar ways; utterance b) is like c) in that they construe path and goal in a similar way and utterance c) is like d) in that they construe agent, thing and path in a similar way. They are placed in this order as a) is more like b) than it is like c) or d) with respect to the relationship between participants; b) is more like c) than it is like d) and so on.

Figure 2. *The functional alignment of constructions based on their event construal; the relevant domain of analogy for argument-structure constructions.*
a) John sent the package to Mary

b) John sent the package to London

c) Mary carried the post to John

d) Mary drove the car to Manchester
The sequential aspects of this event representation are analogous to the *syntagmatic* dimensions of Saussure (1916) — the way participants in an event relate to one another spatially, temporally and causally. The *paradigmatic* dimension is analogous to the space where concepts are formed on the basis of substitutability in events. These have been called slot-filler categories and emerge in development when infants learn about what kinds of things can participate in the roles in such events (see Nelson, 1996 for a review). Thus the formation of linguistic categories — both syntagmatic categories such as agent and patient and paradigmatic categories such as noun and verb — can be seen in the same basic terms as the formation of non-linguistic (or semantic) categories. This account is especially attractive in explaining how young children form superordinate categories, such as food and furniture whose members share little in common perceptually. Food consists simply of those items that play a certain role in children’s breakfast, lunch and dinner scripts. So, schematisation of argument-structure constructions is driven at the conceptual level by shared elements of the event representation, for example the manner and path of objects, the distinguishability and number of the participants, and somewhat more abstract relationships like the attitudes of participants to what is being communicated (e.g. *I think X, I know X*). These meanings have been formalised into the theoretical constructs (e.g. profile, base, domain, trajectory and landmark) of the conceptualist approach to semantics (e.g., Fillmore, 1985; Langacker, 1990). The power of this approach recognises meaning is more than sets of relations between linguistic items. In an important sense, knowing what a word means involves knowing how to combine that word with other words, but the meaning in language is more than a vast calculus of language-internal relations. There are many studies emphasising that children show a sophisticated appreciation of the statistical landscape of the language
they are acquiring (Marcus et al. 1999; Saffran, 2001; Saffran, Aslin, and Newport, 1996). This, no doubt, is useful information for the child to know but theories of development should be careful not to equate this knowledge with learning what words mean; if someone were to download all the statistical knowledge that a speaker has about their language into a spreadsheet, no one, I assume, is suggesting that a matrix of collocation probabilities alone would be able to tell you anything about what the words, or grammatical constructions for that matter, mean. The limitation with the language-internal approach is made clear when we ask how language learners could ever bootstrap the conceptual content of linguistic expressions. “Observation of the semantic relations between dead and alive, between tall and short, between buy and sell actually tell us very little about the conceptual content of these words…language-internal relations must be regarded as symptomatic of meaning, not as meaning itself” (Taylor, 2002) Thus, linguistic intuitions such as noun-hood verb-hood and complement-hood must emerge from the embodied event representations (conceptualisations). Embodied in this context simply emphasises that cognition is a situated activity: abstract symbols acquire real-world meaning by ultimately being grounded in terms of the agent’s experience and physical characteristics. Developmental psychologists and linguistics have asked ‘given an infinite number of generalisations a child could make, why do they make the ones they do?’ The embodied view of cognition combined with the social-cooperative model of communication shows that there simply isn’t an infinite number of generalisations available to the learner to begin with.

The language-learner is constantly changing roles between recipient and communicator and this is important as for the child, learning their language is the cumulative acquisition of de-constructing and re-constructing utterances on the basis
of which forms are associated with which parts of the event representation, and which functions are associated with which forms. Although the mapping between these different representations is, for want of a better word, messy, it is not without regularities either, allowing the infant to triangulate reliabilities between referential context, linguistic form and semantic content. For example, this means the infant might construct an association between the morphological -ing and the ongoing element in the event representation – typically an association that at the start of development is grounded in a communicative joint-attention context and where the perceptual event is ongoing. This idea – that infants’ early speech is grounded in the here-and-now and that associations are time-locked – is compatible with Johnson’s (1999) notion of ‘constructional grounding’, the hypothesis that more concrete source constructions, whose “interpretations are more easily demonstrated by and inferred from non-verbal cues” (Johnson 1999:1), are acquired earlier by children than their metaphorical counterparts.

To return to the example of the caused-motion construction, the idea that argument-structure analogy needs to take place at this level of abstraction doesn’t deny that infants are also developing rich lexical representations; clearly, one utterance is potentially analogous to other utterances across multiple domains (Figure 3). The main point here is that when it comes to schematising the relationships between participants, the functional alignment based on participant roles is the only domain of analogy which will lead to argument-structure grammaticalisation.

Figure 3 shows that for one utterance, the man sneezed the foam off the latte, there are formal-surface analogies to the lexical representations of sneezing and functional analogies that operate on the alignment of event-roles – key to the schematization of argument-structure construction. When the participants are
identified and aligned with previous utterances (cf. Figure 2), *she pushed the glass into the sink* out-competes *she sneezed* when the goal of the system is to find an analogous functional template (c.f. relational mapping, e.g. Genter & Markman, 1993; 1994; 1995).

Figure 3. *Multiple domains of competing analogies.*

So in the beginning, the infant’s attention is directed by a number of processes to the role that participants are playing in an event. By accumulating a history of these usage events – which are temporal associations of multimodal information made within the area of cooperative communication and reasoning – the infant has begun to learn the particular functional distributions of utterance-event co-occurrences in her language.
9. Generalisations of function

Assuming that infants have started to accumulate some of the distributional characteristics of events and utterances, I hypothesise they are able to make the generalisation over the events that X causes Y to move to Z.

Figure 4. The second tier of abstraction generalises over caused-motion event-utterance co-occurrences.

By adding an extra plane of abstraction I am saying that a number of usage-events have been identified as having properties in common that are in some way predictive.
of how they behave – it is worth abstracting over them (in the figure, the lines connecting one plane of abstraction to another). For the story I am telling here for a caused-motion construction, a reasonable inductive generalisation that follows from ‘X cause Y to move Z’ might involve animacy (Figure 5). It is a generalisation in the sense that the category of animate things affecting inanimate things is greater than the class of caused-motion events. There is no one-to-one relationship here, not all caused-motion expressions are going to have animate subjects lexically realised (e.g., *the ball hit the cup off the table*) and clearly not all animate subjects are involved in a caused motion event, however, as actions are prototypically interpreted as an attempt to accomplish a goal, and goals are correlated with animate things, it seems an animacy association that would have some abstraction value. Moreover, infants’ knowledge of animacy could in theory constrain assignment of semantic object roles in the caused-motion construction to plausible candidates. This is a probabilistic non-deterministic influence on role assignment; compare *John throws the ball to Mary* versus *Mary throws John to the ball*. So what do I mean by ‘some abstraction value’?

Firstly, it is resource efficient to incorporate a potentially infinite number of individuals into a smaller number of classes: this is something that is presumably important to a mind/brain with a finite amount of energy at its disposal and a finite amount of time, for practical communicative reasons, in which to comprehend and produce an utterance in the discourse. Secondly, this ability supports an organised schema-instance taxonomy and by knowing some properties you can infer some other properties for free because members of the same class are assumed to share some unobserved properties.

Is this abstraction a reasonable assumption given what is known about the cognitive capacities of young infants? Children have strong expectations about the
capacities of animate and inanimate entities from an early age, for example toddlers
know that animate entities can move on their own while inanimate things cannot
(Golinkoff, 1975; Massey & Gelman, 1988), moreover, they can also recruit this
information for the purpose of role assignment in argument-structure constructions
(Becker, 2007; Gelman & Koenig, 2001). It is not surprising then that a number of
theorists have proposed that the linguistic system has a tendency to align particular
semantic roles with different levels of animacy (e.g. animate → agent, inanimate →
patient (Aissen; 1999, Dowty, 1991). Presumably, the fact that languages with such
different lineages as Hindi, Finnish, Russian, Samoan, Dyirbal, Apachean and Papuan
have evolved to grammaticalise this feature, reflects the fact that carving the world
into animate and inanimate categories is powerfully predictive of how things behave
and thus buys some information for free (see Comrie, 1981; Kibrik, 1985; Mallinson
& Blake, 1981; Song, 2001 for various discussions on the communicative function of
animacy). This is particularly so in the realm being discussed here of the caused-
motion event where acting and being acted upon are salient features of the scene.
Figure 5. *The generalisation of animacy over caused-motion schema.*

The most abstract super-schema ‘X cause Y’ considered here (Figure 6) will of course be schematic of many other constructions in this network such as the ditransitive (X cause Y to receive Z – *John emailed Susan the attachment*) and the resultative (X causes Y to become Z – *she drank herself to death*). Using these abstraction planes makes it clearer to see how families of constructions or patterns of use are related by
the overlap of connections in the network, e.g., ditransitives, resulatives and caused motions are all ‘dominated’ by ‘X cause Y’. One of the advantages of construction grammar is that it recognises languages use grammaticalised prefabricated packages as time-saving devices that structure the contents of what is communicated in a schematised form. This insight – that constructions themselves carry meaning – can neatly explain why some utterances have the meaning they do. For example, the verbs in “Sam helped him into the car” “She let the water out of the bathtub” and “Frank sneezed the tissue off the table” cannot be said to mean “X cause Y to move Z” – they are not causative independent of the construction. The way the overall scene is construed is, in part, structured by the schematized meaning of the construction (Goldberg, 1995).

10. Schema Productivity

So the key question is what happens when the child hears “the dax meeks the gazzer to the pumbo”? How do the various levels of abstraction in the form-function pairing of the caused-motion construction handle this novel utterance? (Figure 6). Of course, the point of developing a step-by-step abstraction has been to try to show that at a schematic level of representation, by this stage in development, this utterance is not novel. Analogous meaning-form-event assemblies will have been encountered before (represented in the diagram by a links to other utterance forms and events). Is this justified? The lesson from the matrix example (Table 1a/b) is that it is, if the infant has had sufficient opportunities to extract such regularities.

Whether the infant has had opportunity to schematize the form will be a function of the frequency of the pattern in the input and how abstract the pattern is. Focussing on frequency, an extrapolation from one estimate of English child directed
speech (but broadly corroborated by Wells, 1981) projects that infants hear something in the order of 7,000 utterances per day of which 15% have the ‘canonical’ English SVO pattern (Cameron-Faulkner, Lieven & Tomasello, 2003). This works out as 383,250 exemplars of SVO utterances per year. This, of course, may not characterise the input heard by all children and the truth is no one really knows what the critical mass is in order to schematize a pattern given that infants are simultaneously learning everything else, but given this input frequency, surely, theories of learning have to consider the possibility that it might not be beyond the parallel processing power of 50-100 billion neurons with 100 trillion synaptic connections to detect such patterns in this distribution, especially when those patterns have socio-cognitive pay-offs in being able to detect them – understanding and being understood. In this example, frequency is calculated at a constructional level, but consider the type variability within the construction; without knowing the referents of dax, meek, and gazzer the recipient can still gain some information from the number of arguments in the discourse and the closed-class items which are highly likely to have been encountered before given their frequency. The morphological markers like progressive-ing, prepositions and determiners provide a skeleton structure from which a scene could in theory be reconstructed. Of course what would be missing would be the role that the participants were playing in such a scene, which is precisely what the caused-motion construction has grammaticalised. So in a sense, for the infant that hears this novel sentence the burden of processing is really on working out what dax, meeks, gazzer

---

5 Thus, the question is often asked ‘how is that children learn language so quickly?’ (given that it is so fiendishly abstract and complicated) but one could as reasonably ask, ‘why does it take them so long?’ (given the experimental evidence that fast learning of novel construction is possible? (Casenhiser & Goldberg, 2005)). The problem with this line of reasoning is that it rarely spells out what pace of development is being compared to. It takes children as long as it does to accomplish x and this is only surprising if it happens quicker/slower than a theory would predict for a certain input condition. The somewhat more interesting question is what factors predict an individual’s developmental trajectory and developmental patterns in general.
and *pumbo* refer to. With reference to nothing else at all, the learner could only conclude that the referents are involved in some sort of caused-motion event. That said, it would be quite pragmatically odd to mention these things without respecting whether the recipient knows what these refer to, hence speakers spend time, at least at the start of discourse, establishing “what we know together” (not explicitly of course). Once communicators and recipients have anchored the referents, the constructional frame, in combination with the verb, structures the relationship between them – if all is well, then in such a way the communicator wants the scene to be construed in the mind of the recipient\(^6\).

11. **Item-specific knowledge and generalisations**

Notice that as the schematicity has increased from level to level the semantic content has decreased (Figure 6). This obviously represents a problem for a model of schematization if it cashes-out meaning as the output of the most abstract node: at the most general level it is schematic of everything and predictive of nothing. However, the symbolic assembly is a composite of all the levels it dominates, a form-function pairing (shown in Figure 6 as a red line). This is consistent with the evidence for graded representations of linguistic knowledge, in fact, one instantiation of this view – radial prototype conceptual structure – is produced as a by-product of ‘seeing’ through the cumulative layers of abstraction, shown in Figure 6 as the cluster of elements through which the red line runs. Thus this is not a situation where we need to choose between supposedly dichotomous views of representations: we retain item-specific knowledge (e.g. exemplars) and we can abstract over them (e.g. prototypes)

---

\(^6\) On this account cognitively incongruent utterances arise when there is a mismatch between the meaning potential (Allwood, 2003) or purport (Cruse, 2000) of the verb and the schematic meaning of the construction in which it appears, e.g. *He sleeps Mary, John received.*
Importantly, after a period of development, the formal pole of this symbolic assembly may stop short of phonological content, that is, after sufficient evidence on which to generalise the construction can detach from the input. Thus the caused motion schema comes to represent a grammaticalised event, free of phonological content in the sense that whether or not the noun-phrase can enter the argument slots is not determined by its phonological properties.

Figure 6. A *four-plane abstraction of the X cause Y event, the caused-motion event forms part of this abstraction, the symbolic assembly is viewed through the four planes and incorporates the activated schemas it dominates.*
“She throws it over there”
“He move ball to the store”
“He push it to Tom”

Norms of Cooperation and Cooperative Reasoning
12. The distribution in the input

The other lesson from the matrix example is that the levels of abstraction are strongly determined by the training set and of course, what is functionally relevant to the individual’s goals, e.g., communicative motives. In the example above I have considered the case where the caused-motion construction emerges in development, but the stages of this abstraction (Figures 1, 4, 5 & 6) are obviously dependent on the structure and distribution of the input. I have tried to model these abstractions in stages that assume there are many other abstractions happening simultaneously (other blobs on the abstraction planes are not connected just because it would make it difficult to see what was going on in one section). One can imagine an instance where caused-motion events are dominated directly by an X cause Y schema if there were no variation in the learning set, for example we lived in a world where animacy was not predicative of any linguistic or event behaviour. Indeed, as a by product of sampling from the input one would expect some categories to start out this way (i.e. they are ‘under-extended’ or ‘over-extended’ with respect to an adult end-state), before being reorganised under the weight of new input evidence, a process which is similar to the conceptual reorganisation of semantic knowledge in Parallel Distributed Processing models (e.g., Rogers and McClelland, 2008).

13. Summary and Conclusions

Of course this is an oversimplified account of the caused-motion construction; there are many finer-semantic gradations that adult speakers have access to that are not represented on the various planes of abstractions in the diagrams. I consider this to be a quantitative difference though, a matter of more development within the individual; the point was to show that at each stage the abstraction is cognitively justified with
respect to the abilities of the infant and the input they are exposed to. Especially important in this explanation has been the fact that humans possess a unique set of social-cognitive and social motivational-skills that allow language to happen.

Secondly, early language acquisition can only be explained with explicit reference to function – both in the sense that underlying concepts have a functional core and in the sense that linguistic categories are formed on the basis of communicative discourse functions (Nelson, 1996; Tomasello, 2003). This, combined with powerful pattern-detection skills, enables distributional regularities in the input to be paired with what the speakers intend to communicate. This account has also emphasised understanding the acquisition of function in the wider socio-cultural context and physical grounding in which language is situated. This approach is answering Wittgenstein’s call to look at language in action: “don’t ask for its meaning, look for its use”. Not doing so is like ripping a plant up to see how the roots work: in one way it makes it easier to see the detail of the ‘object’ of study and in another way it makes it harder to understand how it does what it does. I have tried to describe a psychologically plausible escape from the circularity of concept learning that Plato first described: in order to learn a category one needs to know what the relevant features are; in order to know what the relevant features are assumes some knowledge of what the category is going to become. The main ideas presented here – for example, what something does forms the functional core of a concept – have all existed in the literature in various forms and (re)appeared at various times. The hope is that by bringing them together here I have been able to make what ‘abstracting over instances’ means for a usage-based theorist a little clearer and more concrete. After all, if usage-based theories characterise acquisition as a developing inventory of form-function pairings, it is incumbent on those approaches to outline what they think develops and what form
linguistic representations take. The development of representations proposed here is consistent with the mainstream view of grammar under a cognitive linguistics framework, whereby abstract entities are permissible, but only to the extent that these are schematic for actually occurring structures, and which can be abstracted from actually occurring instances (Taylor, 2002).

An obvious objection to an experiential account of concept learning is that the caused-motion construction might be an example of ‘looking for the car keys under the street light’, that is to say, out of all the constructions this is the most obviously grounded in an event. There is an analogy here with the embodied cognition which has faced similar objections, accused of focussing on concrete terms with tangible physical referents or instantiations. What about the concepts of _justice, art, truth, crime, government_ and so on? This is actually an argument in favour of the approach taken here. It could be argued that the meaning of abstract concepts like _life, liberty, and the pursuit of happiness_ can only be fully appreciated when considered in a wider body of knowledge about how people, institutions and societies work. Likewise, I have emphasised that abstract grammatical patterns are conventionalised patterns of shared experience; patterns of experiences that are organised on the basis of how the participants behave, whether their actions are _intended_ to accomplish something similar. Thus right from the start, wider knowledge of how people work is brought to bear on the schematisation process.

The co-occurrence of events and utterances offers a grounded way-in for infants to learn argument-structure constructions. Once the process of pairing form with function has begun, the redescription of representational levels takes over such that there is a complex interaction between events schematising language and language schematising events, ratcheting higher and higher levels of abstraction. By
this process, over time, the caused-motion event is grammaticalised into the caused-motion construction, so that when a novel construction shares formal properties with previous usage events – that started life as utterance-event pairings – the construction by itself is enough to simulate the event in the mind of the recipient, and the targeted conception is communicated.

The next chapter further investigates the topic of how infants construct grammatical constructions by looking at the transitive construction cross-linguistically in the context of a prototype theory of categorisation. The paper was published in Language and Cognition, 2009: issue 1 volume 1, authored by Paul Ibbotson and Michael Tomasello.
CHAPTER THREE

Publication 2

Prototype constructions in early language acquisition

Page: 109
Prototype constructions in early language acquisition

Abstract

In this paper we bring together several lines of cross-linguistic research to demonstrate the role of prototypicality in young children’s acquisition of the transitive construction. Much research has shown that young children are slow to form abstract constructions because they fail to see the more general applicability of syntactic markers such as word order and case marking. Here we attempt to explain this fact by investigating the nature of the language children do and do not hear, specifically, the reliability and availability of the linguistic cues they are exposed to. We suggest that constructions redundantly marked with multiple cues could have a special status as a nucleus around which the prototype forms—which makes it difficult for them to isolate the functional significance of each cue. The implications of this view for language acquisition are discussed within a usage-based framework.

Keywords

language acquisition, input, construction, prototype, grammar
1. Introduction

One of the great theoretical advances in modern linguistics is the recognition that grammatical constructions are conventionalized pairing of complex forms with complex semantic/pragmatic functions. Phrasal patterns and the rules of syntactic combination not only have meaning but also have the capacity to change the meanings of the words they govern (Croft 2001; Fillmore 1989; Fillmore et al. 1988; Goldberg 1995, 2006; Kay and Fillmore 1999; Lakoff 1977; Langacker 1987; Michaelis and Lambrecht 1996). This is the key insight that enables us to abandon the unproductive mathematical metaphor for grammar (as, for example, in traditional phrase-structure-based theories of grammar) in which words have meanings but grammatical “rules” are totally formal and without meaning or function (Tomasello 1998, 2005). In this more functional view, a person’s grammatical competence is best characterized as a taxonomy of constructions that are organised from fully specified instances to more abstract representations, along with skills for using those constructions to communicate with other persons in particular usage events (Langacker 2000).

The problem for language acquisition is that children do not experience constructions but only utterances; they must (re-)construct for themselves the constructions of their language from the individual utterances they experience. Of particular importance as targets of acquisition are the relatively abstract utterance-level constructions that enable children, at some point, to generate an almost unlimited array of particular utterances following the same general form: for example, transitive, intransitive, ditransitive, passive, cleft, yes-no question, wh-question, identificational, attributional, caused motion, and intransitive motion constructions. But in the usage-based view of language acquisition, children do not begin with
abstract constructions such as these, but rather with constructions that are concrete and item-based with only limited and local abstractions (Tomasello 1992, 2003). Thus, for example, children first acquire a number of different verb-island constructions (e.g., \textit{X hits Y}, \textit{X kisses Y}, \textit{X pushes Y}, \textit{X pulls Y}, etc.) and use these in generating utterances before these coalesce into a verb-general transitive construction.

The central question in any construction-based, usage-based theory of language acquisition is thus how children get from more concrete and item-based constructions to more abstract constructions. Tomasello (2003) has emphasized the general cognitive processes of (i) analogy, for forming abstract syntagmatic schemas across phrases; and (ii) distributional analysis, for forming paradigms of items that might go into particular slots in the schemas. Goldberg (2006) has proposed that this process is facilitated by the fact that in adult language certain verbs are prototypical for certain argument-structure constructions; for example, the verb give is closely associated with the ditransitive construction. Thus, young children hear exemplars of the ditransitive construction most often with the verb give and so, in an important sense, acquiring the meaning of the verb give and acquiring the ditransitive construction are part of the same process. Although skewed input of this type may facilitate learning (Goldberg 2006: 89), it isn’t necessary for children to acquire abstract constructions. Perhaps a case in point is the transitive construction, which in English is so general in function—something like, prototypically, Talmy’s (1988) force-dynamic schema—that no one verb may be considered prototypical (and there is no evidence, to our knowledge, that one transitive verb in English is most frequently experienced by all children). Of course that does not imply that children comprehend the semantic roles of all transitive verbs in the same way. Pyykkönen and colleagues (submitted) showed fifteen 3-year-olds a screen displaying two characters and a
location while they listened to a story of the following form: 1. The X [verb]ed the Y near the [location]. 2. Do you know what happened next? 3. He did something very silly. 4. He [verb]ed. The children had stronger expectations about referents with high-transitivity verbs such that they looked for both subject and the object significantly longer for prototypically transitive verbs such as kissed, kicked and hit, than for those with low-transitivity verbs such as bumped, found and loved. The fact that the effect of verb semantics was detected quickly in the children’s eye-movements (between 520 to 2600ms) suggests that they, like adults, can rapidly generate expectations about the upcoming discourse (Koornneef and Van Berkum 2006) and that these expectations are guided to some extent by a cline of transitivity. So, in the case of the transitive construction— and to some degree in the case of all argument-structure constructions— the prototype is formed not by a verb but by a gestalt of grammatical markers such as word order, marking on the noun and/or verb, as well as semantic cues.

In the current paper, we develop a prototype-based view of the acquisition of argument-structure constructions, with particular reference to the transitive construction. The transitive construction is important as it is present in almost all languages in one form or another (Hopper and Thompson 1980), and ontogenetically-speaking, it is the earliest in which comprehension rests crucially on being able to successfully identify which participants are playing which roles in the event (who is doing what to whom).

We focus here mostly on experimental research, as experiments are necessary to determine the level of abstraction at which children are working. Most of the early research was done in English, but recently there have appeared some cross-linguistic experimental studies that are especially helpful in clarifying the processes involved.
In Section 2 we begin by clarifying what we mean by prototype constructions and how this applies in the particular case of transitivity. Prototype constructions Section 3 focuses on acquiring the transitive in English, and explaining how the language children hear around them influences the nature of their generalizations and abstraction. In Section 4 we systematically compare German, Cantonese & Polish, highlighting the way in which the different morphosyntactic features of these languages affect the development of the prototype form of the transitive construction. Section 5 briefly discusses the effect of frequency and prototypicality in relative clauses, and how the prototype attracts errors. In section 6 we summarise what we see as the role of prototypical constructions in early language acquisition.

2. Prototypes in cognition and language

The introduction of the notion of prototype into the categorization literature by Rosch and colleagues was a revolution (e.g. Mervis and Rosch 1981; Rosch 1983). The basic idea was that a concept, for example, bird, was not defined by a set of necessary and sufficient features—with all members that met the criteria being equals—but rather that the concept had a graded structure, with fuzzy boundaries, in which some members played a privileged role. Thus the prototypical bird is one that shares the most features with other birds and is maximally distinct from non-birds. Similarly, the prototypical transitive clause, with two conceptually prominent participants, is taken to be maximally distinct from a prototypical intransitive.

An important finding in prototype theory is that the prototype—either an actual exemplar or a composite entity—comprises a maximal number of features common to the category, often ‘averaged’ across exemplars. Franks and Bransford (1971) performed an experiment in which they constructed stimuli by combining
geometric forms such as circles, stars, and triangles into structured groups of various kinds. Some of these were then shown to participants—who were then later asked if they recognized these and other shapes they had not seen previously. Importantly, one of the exemplars shown at test contained all of the geometric forms together, an exemplar that had actually never been shown previously (but could be considered the prototype if all of the experienced exemplars were averaged). The participants not only thought that they had seen this prototype, but they were actually more confident that they had seen it than the other previously seen exemplars (or distracter items which they had not seen). The idea of averaging across exemplars contains within it a notion of frequency. The conceptual space that prototypes occupy includes the most representative member but may also be distorted, skewed or weighted towards its most frequent members. Thus, one might say the prototypical outfit for a businessman is a suit—merely on the basis of its high frequency. Importantly, a prototype assigns membership to a category by means of a judgement of similarity to a central exemplar so that an essential property of a prototype category is that it is gradable.

In linguistics, Lakoff (1987) applied the notion of prototype to both lexical semantics and grammatical constructions. In applying the notion to linguistic constructions, we must attend not only to function—for example, the ditransitive construction prototypically involves transfer of possession—but also to linguistic form—the ditransitive construction prototypically has the form of NP1 + VERBditrans. + NP2 + NP3. In Goldberg’s (1995) version of construction grammar there is a focus on the fact that a given form often has a prototypical meaning as well as conventional extensions of that meaning. This is not just confined to metaphorical extensions of prototypical constructions to ‘similar’ conceptual situations (e.g., the
use of the ditransitive construction for acts of information transfer and for benefactives), but also to negation, enablement, and future transfer.

In language acquisition, new exemplars will be assimilated to the prototype by analogy. For instance, when the learner is trying to comprehend the two sentences the car is towing the boat and the truck is towing the car, they do not begin by aligning elements on the basis of the literal similarity between the two cars, but match the car and the truck because they are doing the same job from the perspective of the functional interrelations involved. There is much evidence that people, including young children, focus on certain kinds of relations in making analogies, the most important being spatial and causal relations (Gentner and Markman 1995, 1997; Gentner and Medina 1998). Thus, crucial for making analogies across linguistic constructions is the meaning of the relational words involved, especially the verbs, and the spatial, temporal, and causal relations they encode (Tomasello 2003). It may be helpful to think of prototypes as a kind of prediction-generator. What this means is that once an abstraction is created, it allows generalisations about novel items: the more similar a new instance is to the prototype the more likely it will be to behave in a similar way (this includes linguistic behaviour). This means that the dimensions over which we compute similarity is massively constrained if children are focusing their resources on predicting the function of the forms they actually hear.

With particular regard to the transitive construction, Hopper & Thompson (1980, 1984) laid out the prototypical semantics underlying transitive constructions across many languages. They list what they call the ‘‘component parts of the Transitivity notion’’, representing a scale according to which clauses can be ranked as more or less transitive. More recently, Naess (2007) proposes that a basic criterion for transitivity is Prototype constructions formulated in the maximally distinct argument
hypothesis: A prototypical transitive clause is one where the two participants are maximally semantically distinct in terms of their roles in the event described by the clause. The two-participants of a transitive clause are most commonly labelled ‘agent’ and ‘patient’. Therefore, another way of formulating the maximally distinct argument hypothesis would be to say that the agent and patient categories should be defined in maximal opposition to each other. Thus this event is prototypically realised as an agent **intentionally instigating** an action that directly results in the patient being **affected**. The empirical predictions of Naess’ model are fairly straightforward: any deviation from the prototype will lead to the use of a structure distinct from the fully transitive clause in some language. Conversely, if the semantic features of the prototype are to be taken as a sufficient definition of prototypical transitivity, then any differences in formal transitive marking must in principle be explainable in terms of deviations from this feature configuration. There should be better-or-worse examples of transitivity where the sentence *John broke the plate* semantically overlaps with all the prototypical features described by Næss, while *John accidentally broke the plate*, *John entered the room*, *John didn’t break the plate* (but wanted to) *John didn’t break the plate* (and didn’t want to), all depart from the prototype along the dimensions of intentionality, instigation and affectedness of the agent or patient. So, prototypical transitivity is defined primarily as a semantic/pragmatic concept, but one with obvious structural implications, of which, the clearest are the traditional grammatical relations subject and object (which may themselves be generalisations from the core notions of agent and patient).

While the prototypical semantics underlying transitive constructions may be universal and express something fundamental to human experience, the morphosyntactic resources that are available to express the transitive scene vary from
language to language. Furthermore, within languages the balance between the cues shifts depending on the sentence’s context, the language’s history, or both.

Nevertheless, in many languages (including all of those we will deal with here) there are three major syntactic devices for helping to indicate who-did-what-to-whom in the transitive construction: word order, marking on one or both of the nouns (case), and marking on the verb (agreement). Of course, at some level these cues need to interface with world-knowledge, such as the animacy of the participants (Haspelmath et al. 2005) and semantic plausibility, both of which will contribute to making probabilistic formulations of the event being described. Focusing on the averaging function characteristic of prototypes, we may thus specify the prototype of the form of the transitive construction as one employing all or most of these syntactic devices and cues—at least those that are reasonably frequent. And so in English, something like He’s pushing it or He’s eating the berries can be considered prototypical exemplars of the transitive construction in that they are marked by several of the key syntactic devices and cues. If exemplars of this type are also frequent then they might have special importance in children’s early acquisition of the transitive construction. In the early stages of development, when the type/token ratio is low, the prototype will be closer to the most frequent item. As the type/token ratio increases, with more instances of that category, the average will begin to stabilize, and as the set approaches adulthood levels of exemplars, the prototype of that category will become increasing entrenched and insensitive to new members. If the type/token ratio remains low (as in the ditransitive), the prototype will remain skewed towards the mode (see Figure 1.a. below). Another way of putting this is to say the prototype of a particular functional set is weighted towards its most frequent members, so that you wouldn’t conclude read or pass is as good an example of a ditransitive verb as give was. The
frequency profile for transitives is much flatter (Figure 1.b.), although, as we have seen, they will differ in the extent to which they encode characteristic transitive semantics (Pyykkönen et al. in press).

Figure 1a. *Ditransitive distribution.*

![Ditransitive distribution](image)

ditransitive {give, give, give, give, read, pass}

Figure 1b. *Transitive distribution.*

![Transitive distribution](image)

transitive {get, have, want, take, find, put, do, eat, play, see, like, say...}

A potential problem this presents for acquisition, however, is that if a construction is often marked redundantly with multiple cues, it may be difficult for children to isolate exactly what job each of these markers is doing—and so to generalize these markers productively. From the perspective of the memory and cognition literature this is analogous to the problem of compound cues for function and category learning. From the perspective of an adult conceptualization of construction grammar, redundancy is a manifestation of multiple constructions (e.g., transitive and ditransitive) coming together (through unification) to license a clausal token. The main point concerning us here is that, when a prototype construction is redundantly marked, children may not be able to isolate the syntactic markers involved—what Tomasello (2003) calls the ‘blame assignment’ problem—which will limit their productivity with them.
3. **Acquiring the transitive construction in English**

English children’s relatively slow acquisition of the transitive construction as an abstract, verb-general construction is well established across many studies. Here we review that literature very briefly, and then attempt to explain this developmental pattern in terms of the language children hear around them and what this does and does not afford in terms of generalizations and abstraction.

3.1. **The developmental pattern**

Chan et al 2010 (pp. 30) provided a review of the evidence:

“\(\text{In their spontaneous speech young English-speaking children use canonical SVO word order in transitive constructions fairly consistently from quite early in development (Brown 1973). Similarly, in comprehension tasks, children as young as two years of age appropriately act out ‘the doggie bites the cat’ (reversible transitives) that depend exclusively on a knowledge of canonical word order (e.g., DeVilliers & DeVilliers 1973). We cannot conclude from these facts, however, that very young children have full productive control of English SVO word order. If we do not know what children have and have not heard, it might be that they are simply reproducing the ordering of the particular words they have heard adults using, or they might be marking SVO relations syntactically, but only locally for some highly familiar verbs (Tomasello 1992). One source of evidence for productivity is overgeneralization errors such as “She falled me down” or “Don’t giggle me”—in which intransitive verbs are used productively in the SVO transitive frame (Bowerman 1982; Pinker 1989). The basic fact here is that children seldom make such errors prior to about 3 years of age. Another source of}\)
evidence comes from experiments with novel verbs, enabling researchers to control what the child has and has not heard. In the case of English SVO transitive word order, the experiments often involve elicited production. In these experiments, children are exposed to a novel verb in a syntactic construction such as an intransitive or passive and then tested to see if they can later use that verb productively in the canonical SVO transitive construction—with cues to syntactic roles other than word order (e.g., animacy of the S and O participants, use of case marked pronouns) carefully controlled.

Experiments of this type have clearly demonstrated that by 3.5 or 4 years of age most English-speaking children can readily assimilate novel verbs to an abstract SVO schema (e.g., Maratsos et al. 1987; Pinker et al. 1987). That is, the vast majority of children from 3.5 years of age can produce a canonical transitive SVO utterance with a completely novel verb (e.g., the cat’s meeking the dog) that they have never heard before in a transitive SVO construction. However, this is not true of younger children. In a series of studies by Tomasello and colleagues, 2- to 3-year-old children readily use a novel verb in the construction in which they have heard it modelled, but they do not use it in constructions, including the SVO transitive construction, in which they have not heard it modelled (see Tomasello 2000, 2003, for reviews). Very similar findings have come from other elicited production methods with very different response demands, such as the weird word order method of Akhtar (1999) (see also Abbot-Smith et al. 2001, and Matthews et al. 2005), and the priming methodology of Savage and colleagues (2003) (see also Shimpi et al. 2007). In all cases, children below about 3 years of age are generally very poor at
producing newly learned verbs in constructions in which they have not heard them used previously.

Despite these very consistent findings using several different production methodologies, some researchers argue that production experiments are too demanding from a performance point of view (although this criticism has not been levelled against the priming methodology). The classic test of comprehension is the act-out task. For example, Akhtar and Tomasello (1997) created a modified act-out task for use with novel verbs. Children first learned to act out a novel action on a novel apparatus with two toy characters, and then the adult handed them two new characters and requested Show me: Cookie Monster is meeking Big Bird (while placing the apparatus in front of them). Since every child knew the names of the novel characters and on every trial attempted to make one of them act on the other in the appropriate way, the only question was which character should play which role. The under-3-year-old children were, as a group, at chance in this task, whereas older children were quite good. The other major technique used to assess children’s comprehension of English word order is preferential looking. In this technique, a child is shown two displays and hears a single utterance (through a centrally located loudspeaker) that describes only one of the displays felicitously. The question is which display the child will prefer to look at. Although there are findings of children’s comprehension of English word order using familiar verbs (e.g., Hirsh-Pasek & Golinkoff 1996), and findings concerning verb semantics using novel verbs (e.g., the difference between causative and inchoative meanings; Naigles 1990), there is only one study testing SVO comprehension with novel verbs. Gertner, Fisher and
Eisengart (2006) found that children as young as 21 and 25 months old preferred watching the causative event which matched with the reversible SVO sentence they heard."

But this study had a training phase with familiar verbs in SVO sentences that used the exact same characters (and nouns to describe them) used in the test phase (in canonical SVO order), which could have taught the children essential linguistic and/or non-linguistic strategies. Clearly these discrimination effects are highly fragile and while they may point to an initial sensitivity to possible mappings between aspects of the sentence and events much more work needs to be done to investigate precisely what these mappings might be and whether and how they are related to children’s much later ability to use a full representation of the abstract transitive construction.

Overall, then, we see a slow and gradual development of skills with English word order in the transitive construction, in both production and comprehension. We would now like to ask the question why this development is so slow, with children taking several years for a highly frequent construction.

3.2. What English-speaking children hear

English children’s relatively slow acquisition of word order as an indicator of a verb-general transitive construction is due, we believe, to two main factors: the diversity of configurations in which S, V, and O are realised in the input, and the redundancy of the syntactic cues involved.

First, in terms of natural input, Cameron-Faulkner and colleagues (2003) looked at the child directed speech (CDS) of twelve English-speaking mothers and analysed their utterance-level constructions, amounting to nearly 17,000 utterances of CDS. Based on their sample— and broadly corroborated by similar analysis of a more
diversely collected sample (Wells 1981)—they estimate that English-speaking children hear, every day, something in the order of 7,000 utterances. Over 30 percent of the utterances have no subject (e.g., imperatives, fragments), another 40 percent have the subject after either the auxiliary or main verb (e.g., questions), and of the remaining 30 percent only about half have the prototypical patient following the verb. What this means, at least for the age range sampled (1;9–2;6), is that an English-speaking child would have difficulty figuring out that SVO is supposed to be the canonical English form. The fact that children hear other combinations of S, V, and O much more often than they hear SVO utterances may help to explain why English-speaking children take so long to construct an abstract transitive construction based on word order.

Evidence to support this interpretation comes from a training study. Childers and Tomasello (2003) gave 2 1/2-year-old children several hundred transitive utterances of various types over three training sessions and then, in a fourth session, assessed their ability to use a novel verb productively in the transitive construction. For some children, the transitive utterances were modelled with nouns only (The boy’s chopping the tree); for other children the transitive utterances were modelled both with pronouns (He’s chopping it) and with nouns (The boy’s chopping the tree); and some children were in a control condition with no relevant training. Only 20 percent of the children with no training produced transitive utterances with novel verbs during testing (in line with previous studies). The children who received training were much better, with roughly two-thirds generalizing with the novel verb at test—with children trained with both pronouns and nouns being especially skilful.

These results demonstrate two key points relevant to current purposes. First, when the range of SVO configurations is reduced in the input and children are given
lots of exemplars in close temporal proximity, young children are capable of creating any verb-general transitive construction at younger ages than normal. Second, it is likely that many factors influence this process, at least two of which are: (i) type variation in the NP slots (as in the condition with nouns and pronouns) and (ii) case marking on pronouns (e.g., he is the nominative form).

The second reason that English-children may be relatively slow to acquire the word-order cue for indicating agent/patient relationships is that, as in most languages, the simple transitive construction is redundantly marked with word-order, case marking agreement, and animacy, as for instance in He’s eating it. If all of these cues co-vary with high frequency then the learner may treat the cues as a kind of ‘linguistic gestalt’ with no internal structure, only breaking down into its component parts when placed in a sufficient number of different linguistic contexts where the cues are placed in conflict. This is a similar idea to a compound cue for function and category learning, a well-established phenomenon in the memory and cognition literature.

The idea is that in the novel verb experiments of Tomasello and colleagues, in almost all cases the only syntactic cue available to children was word order, with no case marking, agreement, or animacy cues available— probably a fairly rare event in young children’s early linguist experience (since 80 percent of the transitive sentences they hear have pronoun subjects; Cameron-Faulkner et al. 2003). To define itself against those other cues, word order would have to vary independently of case, animacy, and agreement with sufficient regularity—which in the real world it does not do very often. This suggests the possibility that young children could acquire a verb-general transitive construction a bit earlier if it were presented to them in prototypical form with all cues available—and that their slow acquisition is in reality a slow acquisition of the significance of the word or a cue working alone. In other
words, redundancy masks that the whole has many parts, and consequently learning to use these parts productively in acquisition is delayed.

In summary, English children may be slow to acquire verb-general comprehension and production of transitive word order for two basic reasons. First, the diversity of constructions across the input makes it difficult to isolate SVO agent-patient as the most productive sentence interpretation strategy. Second, because word-order appears redundantly in the input with animacy and case-specific pronouns (and sometimes agreement), it takes time for word order to define itself against these cues, which often functionally co-vary as an indication of agent/patient relationships.

One way to test these hypotheses is to systematically compare the prototype form of the transitive construction—using multiple cues redundantly— with other less informative forms, and to do this in languages whose morphology is more important to sentence interpretation than it is in English.

4. Recent cross-linguistic studies

One way to investigate the interaction of different syntactic cues in construction learning in different languages is through the theoretical framework of the competition model (Bates and MacWhinney 1987). Since the 1980s, the model has motivated cross-linguistic research in over fifteen languages with both adults and children. The model seeks to account for the differential ‘weights’ among various combinations of competing and converging cues including word-order, stress, animacy, agreement, topicalisation, prepositions, and case. The model has a functionalist approach to sentence processing; ‘the forms of natural languages are created, governed, constrained, acquired and used in the service of communicative
functions’ (MacWhinney et al. 1984: 128). It argues that the strength of a particular cue is a product of how frequent it is present when it is needed (cue availability) and how consistently it is mapped onto a particular form whenever it is present (cue reliability). Cue strength is also affected by the processing limitations imposed by the perceptual and working memory systems (cue cost). At its simplest, the classical competition model conceptualises language development as a process whereby cue strength (mediated by cue cost) comes to vary as a function of that cue’s availability and reliability in the input.

One prediction from this is that children should find it especially easy to comprehend prototypical transitive sentences with both word-order and case marking (and perhaps other cues) working in coalition: the coalitions-as-prototypes model (Bates and MacWhinney 1987). But since case marking is not really a consistent feature of English (only present in the pronominal system), to do this we need other languages. Of special note in the following experiment is that, as in the English experiments reported above, novel verbs are used. In the classic competition model experiments, familiar verbs are used and so it is impossible to tell how general children’s knowledge is of the syntactic markers being tested.

4.1. German and Polish

An excellent language for testing children’s use of both case marking and word order is German. Dittmar et al (2008: pp1154) describe the relevant facts for German:

“In active transitive sentences in German the agent of the action is subject and is marked with nominative case marking, and the patient is direct object and is marked with accusative case marking. For both of these, the case marking is either a special form of pronoun or a noun with a special form of determiner.
For example, if a dog is agent the form is *der Hund* (the-nominative dog) or *er* (he), whereas if a dog is patient the form is *den Hund* (the-accusative dog) or *ihn* (him). Additional complexity comes from the fact that nominative and accusative marking take different forms when applied to nouns of different genders, and in some cases they are not distinct. For example, unlike the example of dog above (which is masculine), if a cat is the agent the form is *die Katze* (the-cat-nominative), but if a cat is the patient the form is exactly the same *die Katze* (the-cat-accusative). This means that in some instances case marking is not an available cue in the sense that it does not identify case role unambiguously. Finally, although in German transitive sentences agents typically come before the verb and patients after the verb to highlight the patient pragmatically the reverse order may be used, as in English, with the case roles marked by case marking and unaffected by the reverse order. Thus, “*Den Hund beisst der Mann*” has the first noun, *Hund*, marked as accusative and the second noun, *Mann*, marked as nominative and so, despite word order, it is the man who is biting the dog.”

Dittmar and colleagues (2008) used novel verbs to test German 2-, 5-, and 7-year olds’ comprehension of case and word order in transitive sentences. For the reasons noted above, in German case marking is 100 percent reliable (although not always available, e.g., with feminine NPs) whereas word-order is not reliable because of object first transitives (though, in some sense, always available). Below are examples of the three experimental conditions and their associated sentences in which the children had to identify the agent/patient to successfully act-out or point to the appropriate participant.
Prototype

Der Hund wieft den Löwen

(The masculine-nominative dog is weefing the masculine-accusative lion)

Word-order-only condition

Das Schwein tammt das Zebra

(The neuter pig is tamming the neuter zebra)

Conflict condition

Den Hasen bafft der Frosch

(The masculine-accusative bunny is baffing the masculine-nominative frog)

The results showed an astoundingly clear pattern. The 2-year-olds comprehended transitive sentences only in their prototypical form with redundant marking of agent and patient. In other words, they were the most sensitive group to cue omission, failing to comprehend transitive sentences for which the diagnostic case marking was absent or those in which the word order was non-canonical (object first). The findings suggest that in a language like German children do not begin by attending to cues individually, but rather they learn to comprehend the prototype and are impaired whenever there is deviation from it. Interestingly, the prototypical form in German is also the most frequent (calculated from a corpus of child-directed speech).

The 5-year-olds comprehended the transitive sentences mostly in terms of word-order. Their performance in the word order condition was as high as with the full prototype, and they chose at random in response to the object-first sentences in which word-order and case conflicted, with a slight tendency to go, incorrectly, with word order. Only the 7-year-olds performed like adults in going with case marking in
all conditions, including in object-first sentences in which case marking and canonical word order conflicted.

This pattern of results presents a challenge to the standard cue-competition analysis; word-order is less reliable and valid than case marking in German. One possibility is that German children do not use case marking in a completely general way. This is based on the fact, noted above, that German has three noun classes so that, for example, nominative case marking has three different forms in the singular and another in the plural. If children at a particular age have not yet discovered that all these forms mark the same case, then how the cue reliability is calculated (how reliable a cue is, when it is present, in indicating the correct interpretation) needs to reflect this. In the Dittmar and colleagues study the children were tested on the particular case markers \textit{der} and \textit{den} used as determiners (masculine nominative and accusative). But in children’s natural input these particular items are available in only 21 percent of all transitive sentences. Therefore, children’s comprehension of these may not benefit from their experience with case marking using pronouns or the case markers as expressed in other genders, in which case the cue availability of “case marking” in German is not particularly high.

Recalibrating the cue availability in this more item-based way results in the availability of case marking (as instantiated by the masculine form only) being much lower (21 percent) than that of word order (87 percent)— and this means that overall cue strength is lower for this item-based case marking as well. In line with Sokolov (1988) this suggests that “young German children rely on different input parameters at different stages of development; specifically they rely more on cue availability” early in development and cue reliability later on in development (Dittmar et al, 2008: pp1163). We will return to the idea of children learning to connect the different case
equivalent forms when we consider data from Polish children in third and fourth year of life.

It is worth noting that adults have difficulty in processing non-canonical word orders, at least measured by reaction times (Ferreira 2003; Kaiser and Trueswell 2004). Thus, when German adults are confronted with object-first sentences which are ambiguously marked on the first noun-phrase, they initially interpret these as subject-first sentences until they hear the second noun phrase. (Weber et al. 2006). Perhaps it is not surprising then, that it was only the 7-year-olds in the Dittmar and colleagues study that succeeded in the conflict situation, weighting the case marking cue over the word cue as adults do. Following the reasoning from above this would mean that by seven years of age, children should know the grammatical equivalence of all (or at least most) of the different gendered case markers serving the same grammatical function (and should ignore ambiguities based on other information). For the 7 year-olds, the cue reliability of case marking resembles that of adults so they finally rely on case marking over word order.

In summary, older 2-year-olds understood only sentences with both cues supporting each other, but not sentences with either cue on its own. Five year-olds were able to use word order by itself but not case. Only 7-year-olds behaved like adults comprehending both cues on their own and relying on case when they conflicted. So it seems that it is only when children are somewhat older—when they have had sufficient exposure to the grammatical cues in various combinations—that they are able to isolate and weigh them appropriately in terms of their reliabilities for signalling specific functions.

The German 2-year-olds showed that neither case nor word order presented on its own is sufficient, to guarantee comprehension. More evidence of how grouping of
linguistic cues develop in learning morphological paradigms comes from a recent study by Dąbrowska and Tomasello (2008) in Polish. As Dabrowska and Tomasello describe, (pp. 534)

“Polish is a morphologically rich language with a very elaborate system of case inflections. It has seven cases each signalled by different suffixes that also mark the number on the noun. The single most important determinant of the choice of ending is gender, which can be fairly reliably predicted from the phonological form of the nominative; nearly all feminine nouns end in -a or -i; the vast majority of masculines end in a consonant, and neuters typically end in -o, -e, or ę. Other factors, such as the phonological make-up of the stem and semantics (especially animacy) come into play when there is more than one ending for a particular gender. The instrumental case endings are -em [em] for the masculine and neuter singular and -q [o] for the feminine singular and for the masculine nouns which end in -a (which decline like feminines in all cases, not just the instrumental).”

Like other Polish cases, the instrumental is a polysemous category with a number of different functions, the most important of which are instrument (body parts), material/substance, means of transport, companion, subject predicate, ground object, and manipulated object (as in English ‘play with X’). In the experiment children were exposed to two novel verbs of manipulation which govern the instrumental case. One of the verbs was modelled in a conjunction with three masculine patient nouns, and the other in conjunction with three feminine nouns. Sentences were then elicited with the novel verbs and new patient nouns of the same gender (the matching gender condition) and a different gender (the non-matching gender condition). The rationale
for doing so was as follows. If children rely on concrete generalizations on specific case markers, they should be able to apply the correct ending to nouns of the same gender they were trained with, but not in the non-matching gender condition. On the other hand, if they have access to a more abstract instrumental construction which subsumes these concrete endings, they should be able to use the novel verb with nouns of both genders. Children could also ‘correct’ to canonical (i.e. accusative) case marking; this would indicate that they have acquired a verb-general nominative-accusative construction. The main question was thus “how much children know about the instrumental as a syntactic category, specifically, whether they know that -em and -a, are both exponents of the same case” (Dąbrowska and Tomasello, 2008: pp534). Polish “2½-year-olds were able to supply the correct instrumental marking on the object of a novel verb governing the instrumental case even when the noun belonged to a different class than the nouns in the training set” just under half the time (Dąbrowska and Tomasello, 2008: pp534). The 3½-year-olds were able to do so approximately 85 percent of the time, showing that they have much stronger access to an abstract instrumental category. Thus, Polish children are able to use the instrumental to mark important case relationships with novel verbs considerably earlier than English-speaking children learn to use word order productively for the same purpose, in spite of the fact that the instrumental case is relatively infrequent. The reason for this, Dąbrowska and Tomasello suggest, is that case markers are local cues in the sense that one can determine the role the noun plays in the event described in the sentence from the case marker alone, without having to hold the whole sentence in working memory.

Children’s performance on these kinds of tasks, though far from perfect, was clearly systematic and reveals that the children have formed some kind of linguistic
generalisation about the verbs they learned during the experiment. Polish children in third and fourth year of life learned to connect different forms of instrumental case, meaning that the child’s perception of the input changes in the sense that the three different forms of the instrumental all count as evidence for the same thing—whereas before this connection they did not. This may give an answer as to why the German children in the Dittmar and colleagues study are slow to realise case is a much more valid cue in German than word order. Polish 3-year-olds have learnt to group (and so count) the instrumental endings (-em and -q) to their polysemous functions as a single cue in a way the German 3-year-olds have not yet done so with der and den. Recall that if the connection has not been made across genders and nominative and accusative in German the reliability is actually lower than word-order. This shows that some of the most critical stages in linguistic developmental are those where cues are grouped and regrouped to maximise the predicative power of a syntactic category to infer a function or functions. At the beginning the grouping of cues is conservative, and many items stay grammatical islands, prone to extensive revision and maybe even abandonment. Thus, children may sometimes have constructions that are somewhat incompatible with one another, so that they will vacillate between competing ways of saying the same thing (e.g. Akhtar 1999). This process continues until the weight of examples is such that the prototypes are understood in terms of the role that each of the different cues is playing in the whole construction.

4.2. Cantonese, German and English

Chan and colleagues (2009) “examined young children’s general understanding of word order and animacy contrasts as cues to the agent-patient relations in the transitive construction” (Chan et al., pp. 290). They compared children acquiring
different languages at different points in development using essentially the same act-out paradigm with the same sentence, object, and novel action stimuli. They tested children acquiring monolingually Cantonese, German, and English at three age levels (2;6; 3;6 & 4;6).

Cantonese, German and English are similar in having the same basic SVO word order (in pragmatically neutral situations) of simple active sentences but vary in the extent to which these cues are available or reliable and hence informative as to the meaning of their utterance. Based on calculations of cue strength, one would expect word order to be the dominant cue to agenthood in English, a prediction that has been empirically verified many times within the cue-competition model (for example Bates et al. 1984, 1987; MacWhinney et al. 1984). The situation is somewhat different for German so that apart from the canonical SVO it also allows OVS, SOV OSV, VOS and VSO, for both grammatical reasons (e.g. the verb-final rule for subordinate clauses) and for pragmatic purposes. Because the subject/agent or the object/patient is more loosely tied to position we would expect that word-order as a cue to mark agent/patient relations is not as reliable as in English. Cantonese is similar to German in this regard as it permits OVS, OSV, SOV and VOS word orders. In addition it also allows a high degree of argument noun ellipsis in natural discourse, reducing that cue’s availability. Consequently, one would expect that word order is not highly reliable in Cantonese.

Animacy occupies a slightly different status as a cue to agent/patient relationships than the cues we have been considering so far because, by definition, it is an inherent semantic property of an entity rather than a grammatical convention, such as case or word order. “Prototypically, agents tend to be animate and patients tend to be inanimate”, so where “there is an animacy contrast between two entities in
a transitive sentence it is prototypical that the animate one is the agent and the
inanimate one is the patient”, and this should be “highly reliable across languages—
though certainly not always as in the sentences the ball hit John” (Chan et al., 2009:
pp 274). In this sense the animacy cue is just a by-product of the semantics of
intentionality, instigation and affectedness proposed by Næss (2007) as prototypical
of transitives. In the scenario where both nouns in a sentence are animate, the animacy
cue is not contrastively available (either noun could act as agent). The key here is that
Cantonese is characterised by a lack of morphological cues to agent-patient
relationships (no subject-verb agreement or case) which are more readily available in
German and English—and word order is not always available due to ellipsis—and so
we might expect Cantonese children to rely more on animacy.

In this study, again, novel verbs were used to address the question of whether
children could use their understanding of word order and animacy as cues to the
agent/patient relations to interpret novel sentences. Of particular interest was whether
children would find especially easy sentences in which both cues were used
redundantly in conjunction (Bates & MacWhinney 1987) and they would find
especially difficult sentences in which the two cues conflicted.

Cantonese-, German-, and English-speaking children aged 2;6, 3;6 and 4;6
acted out transitive sentences containing novel verbs in three conditions: (1) agent and
patient were marked redundantly with both word order and animacy; (2) agent and
patient were marked only with word order, and (3) agent and patient were marked in
conflicting ways with word order and animacy.

When word-order was the only cue, English children showed the earliest
comprehension at 2;6, then German children, and then Cantonese children at 3;6.
When the cues conflicted, none of the 2;6 children in any language comprehended in
adult-like ways, whereas all of the children at 3;6 and 4;6 preferred word order over animacy. These results showed that across languages, children aged 2;6 comprehend transitive sentences when they have support from the coalition of word-order and animacy, even with novel verbs. Not only are the convergent cues helpful, the animate-verb-inanimate sentences are also frequent in the input, and they encode the prototypical causative scene which should be highly familiar to young children in their experience regardless of their target language (see the Manipulative Activity Scene in Slobin 1985, 1997).

Reflecting on Cantonese children in particular, we may propose that they take a particularly long time to acquire word order marking of ‘agent/subject’ vs. ‘patient/object’ because (i) often there are no arguments realised in the input, and (ii) word order often appears redundantly with animacy in the prototype. These developmental results correspond well with the different properties of the languages children experience, suggesting that children’s learning of syntactic marking of agent-patient relations is strongly influenced by the nature of the language they hear.

5. Relative clauses

The simple transitive construction is relatively frequent in all of the languages studied here. Relative clauses in some cases use transitive syntax, and here we can see another effect of frequency and prototypicality, namely, that the prototype attracts errors in its direction.

Diessel and Tomasello (2005) gave English and German 4-year-olds relative clauses to imitate. It turns out that even just repeating syntactically difficult relative clauses such as genitive relatives (e.g., “This is the woman whose cat caught a mouse yesterday.”) is extremely difficult for children this young. The same is true, though
less so, for oblique relatives (e.g., “This is the dog that the cat ran away from this morning”) and even object relatives (e.g., “This is the girl who the boy teased at school this morning”). Subject relatives (e.g., “This is the man who saw Peter on the bus this morning”) are easiest for children in both languages.

There are various reasons why subject relatives should be easiest for children, but it is not just frequency, as object relatives are just as frequent in the language children hear. Diessel and Tomasello stress that in subject relatives the basic syntax of the relative clause matches the transitive syntax (assuming transitivity in all cases) of the simple transitive construction used in main clauses. Strong evidence for this comes from the fact that in both languages when children made mistakes in repeating the more difficult types of relative clauses, they almost always reverted to subject relatives employing transitive or intransitive syntax (approximately 80–90 percent in both languages). This is despite the fact that in the two languages the transformations the children had to effect to make this mistake were completely different. In English the children had to transform the word order they heard to get to the easier subject relative: for example, if they heard “Here is the man that the woman kissed”, they produced “Here is the man who kissed the woman”. In contrast, in making this same mistake the German children left word order the same but had to change the case of the relativizer: for example, if they heard “Hier ist der Mann den die Frau küsste” (object relative: den = accusative case), they produced “Hier ist der Mann der die Frau küsste” (subject relative: der = nominative case). Similar phenomena have been reported in questions with long-distance dependencies (Dąbrowska in press) and complement taking verbs (Kidd et al. 2006). The main point for current purposes is simply this. The simple transitive construction serves as a kind of an already established prototype or template that has its influence as children are acquiring more
complex syntactic constructions such as relative and other subordinate clauses. Prototypes attract errors in their direction.

6. Discussion

We have attempted here to explore the usefulness of the notion of prototype constructions for theories of language acquisition. There is a story to tell both for the language children hear from the adults around them, as well as for children’s own construction of linguistic representations.

We have characterized prototypicality in terms of two dimensions: frequency and maximal marking. But actually these two dimensions do not always go together in the language children hear from adults. In our proposal, adults have a prototype representation of, for example, the transitive construction in which all forms of syntactic marking (word order, case, agreement, and perhaps animacy as a semantic cue) are present. But this fully marked version may not be the one they use most frequently in speaking to young children or anyone else. This fact may have been a bit obscured in our account, given the construction and languages we chose to investigate here—as the transitive construction in both English and German is fairly frequent in something close to maximally marked form in the language children hear.

But in Cantonese and Japanese, for example, young children quite often hear transitive sentences with missing arguments, in which case none of the marking devices or semantic cues could even potentially be present. But then they hear other transitive sentences with other overt manifestations of marking, and—as the adults did in the Franks and Branford study with nonsense shapes—they glue these all together into one prototype construction. They presumably do this based on the similar meaning in the different cases. And this is where Goldberg’s (2006) proposal of the
special role of certain verbs might be especially useful. If children hear a variety of
different transitive sentences using very different patterns of syntactic marking—but
all with the exact same verb and closely related meaning—this might be the perfect
situation for them to perceive the different sentences as all exemplars of the transitive
construction.

And so our proposal is actually that frequency and maximal marking play very
different roles in acquisition. Children will acquire first the instantiation of the
construction they hear most frequently. But then they will bring together different
instantiations of the construction on the basis of similar function, such that a prototype
is formed including all of the marking options.

And so we might actually propose two different routes children might use to
get to their own prototype representation of a linguistic construction (as well as
various possible mixtures of these strategies). On the one hand, they might hear quite
frequently maximally marked exemplars, for example, of the English or German
transitive construction. In this case, as we have stressed above, while the redundant
marking may in some sense help initial acquisition, it hinders children’s ability to
identify the syntactic work being done by each of the cues separately. Children will
need to do some kind of “blame assignment” by observing a particular marking
device in sentences that use only it—or in different constructions. Cue availability and
reliability will play a crucial role in this process.

On the other hand, if children were to hear most frequently early in
development sentences with, for example, only one type of marking, then they would
have to create a prototype representation by somehow bringing together with this
singly marked instantiation, other singly or multiply marked instantiations of this
same construction. But in this case, too, our proposal is that they end up at some point
with a prototype representation involving maximal marking. One may also, of course, imagine various combinations of these two routes. For example, Cantonese children probably use some combination as they hear very frequently both transitive sentences with no arguments and transitive sentences with redundancy between animacy and word order. Despite the different specific processes involved, in all of these various developmental routes, cue availability and reliability will always play crucial roles.

Given that there are different developmental routes possible to a prototypical representation of a construction, one important avenue for future research will be to explore other construction types and in other languages. One interesting domain in which this has already been done to some extent is tense-aspect marking within the verb phrase. For example, Shirai and Andersen (1995), and Andersen and Shirai (1996), argue and provide evidence for a prototype account—based on many different languages from many different language families—in which children start using past inflections predominantly with achievement verbs, and progressive inflections with activity verbs. This is presumably because this distributional bias is found in the speech they hear around them. But in the end the children will learn to more readily use all tenses and aspects with all kinds of verbs.

Many questions still remain, such as what level of abstraction is necessary to characterise young children’s grammatical competence. A parsimonious way forward would be to suggest that the same principles govern prototypicality across linguistic and non-linguistic domains, i.e., the level of abstraction represents a trade-off between the cost of identifying a category versus its informativeness. For example, there is evidence that it is at the basic level at which entities are most likely to be named, and that these are the first words children learn for objects. As originally described by Barrett (1986) and demonstrated by Meints and colleagues (1999), prototypicality
plays an important role in early word learning as children connect their first words
(e.g., bird) to prototypical items (e.g., a sparrow) before they connect them to atypical
items (e.g., an ostrich). This is a closely related (but different) idea to that of schema-
instances discussed in the cognitive grammar literature (e.g., see discussions of the
‘categorisation triangle’ in Taylor 2002). Schemas abstract what is common to all its
instances and the recursive application of the schema-instance relation can be
represented in a taxonomic hierarchy. Interestingly, there is evidence that multi-level
taxonomies are constructed gradually in the course of acquisition from the basic-level
upwards (Anglin 1986: 91). One way to think about basic levels is to say that it is the
highest level in a taxonomy at which one is able to form a mental image of a concept.
For instance, there may be a prototypical bird but there is no prototypical furniture,
only instances of that schema. This fits with thinking about prototypes as prediction-
generators; we would not expect a prototype ‘furniture’ because a piece of furniture
does not have characteristic parts, nor is there a characteristic function of furniture in
general. Since it has no predictive value it is not worth the processing cost of forming
an abstract category (this may be analogous to constructions towards the more
idiomatic end of the spectrum). There are, however, characteristic parts to a chair
(seat, legs and back) and function (to sit on), therefore it would be worth abstracting
over these instances and we can easily call to mind what this prototype would look
like. So to reiterate an earlier point, instances of a schema are related by similarity
from the perspective of the functional interrelations involved. In the current case,
instances that are closer to the prototypical transitive schema should be more
productive, that is, able to withstand more substitutions of features yet still be
recognisable as a member of that category. The further towards the periphery of a
prototype the instance is, the more it should require specialized pragmatic contexts to
license such as departure. Examples of such usage-events would be the prototype ‘licensing’ metaphorical extensions or a change in figure-ground organisation, as in the passive. Clearly prototypes will take time to assemble in development, that is, the learner must have experienced a certain threshold of exemplars from which to form an abstract category, otherwise the set of things that the category refers to is too large and uninformative. By informative we mean, able to make productive generalisations on the basis of an abstraction that are compatible with the conventions of the language. In other words, units above the basic level are generally so schematic that they are applicable to a very wide range of entities and tell you very little about the entity in question. There has been very little systematic research on the proposal that a certain number of exemplars is needed—a critical mass—before totally abstract analogies can be made (Marchman and Bates 1994). We need to take a closer look at the nature of this critical mass, for example, the reliability and availability of the linguistic cues children experience as input to the language acquisition process.

Our aim in this paper has been to add a developmental account to existing prototype theories of language processing by taking a more detailed look at one syntactic construction in four languages. Overall, we hope to have shown (i) that an abstraction is formed initially on the basis of frequent overlapping cues, so that the initial abstraction constitutes what will eventually be the prototype of a more complex, more abstract category; (ii) the importance of high-frequency forms in providing ‘anchor-points’ from which more abstract generalisation will gradually emerge; (iii) isolating individual cues is difficult to begin with as they often occur redundantly in the input; (iv) the construction redundantly marked with multiple cues could have a special status as a nucleus around which the prototype forms; and (v) the nature of the input, as characterised by reliability and availability, is a strong predictor
of cross-linguistic differences in language acquisition. Our theoretical claim here is that there are some basic principles of frequency, reliability, and conceptual development that will be critical and play similar roles in all cases.

The next chapter applies the theoretical review of this chapter to an empirical experiment, specifically, making developmental comparisons of the prototypical semantics of the transitive construction in English.
CHAPTER FOUR

1. Introduction

This chapter begins by setting the empirical work in more background context than is possible to include in the submitted paper that appears at the end of this chapter. Hopefully this will make the motivations for the questions addressed in the paper even clearer and also help the reader to see the continuity with the preceding chapters.

2. Generalisations in general

One way in which we can investigate the nature of generalisations in language is to start by considering what motivates generalisations within a system, that is, their function. Framing the question in this way allows us to talk of the ‘why’ of generalisation phenomena not only of the ‘where’ and ‘what’ of purely descriptive accounts. The motivation for generalisations in the language system can be laid out as a kind of syllogism:

1. The goal of the language learning system is to understand and be understood.
2. The sentences that the system learns to comprehend and produce is an open-ended set.
3. In order to satisfy (1) and given (2) the system must generalise patterns to some extent in order to comprehend and produce new utterances (as it is not sufficient to memorise all sentences heard).
So what is the nature of generalisations in language? Bates and MacWhinney regard language as a “new machine built out of old parts” (1988: p147). While there is no general agreement in the emergentist literature on just what those parts might be, the point is that they are not modular to language, for example, those parts typically include features of physiology and perception, processing and working memory, pragmatics and social interaction, properties of the input and of the learning mechanisms. So, following in the emergentist tradition, instead of asking ‘what is the nature of generalisation in language?’ one could reasonably begin by asking ‘what is the nature of generalisation?’

Cognitive systems do not generalise randomly or completely, they are driven by some functional pressure, typically the need to predict or infer certain properties on the basis of perceived characteristics (Anderson 1991; Holland et al., 1989; Kersten and Billman 1997; Leake and Ram 1995; Murphy 2002; Ross and Makin 1999; Wisniewski 1995). The main message here is that as a general rule people pay attention to the features that their prior knowledge says are the important ones. When faced with a new object, situation, or problem, people often access knowledge to categorise it as being of a certain type and then use this knowledge of this type to determine how to deal with it. For example, suppose someone sees something that barks, has four legs, a furry coat and moves around in a particular way. Presumably that person will not stand in wonder at the sight of this strange new beast, unable to make any guesses about how it might behave. They reason, hopefully, that things that bark, have four legs, move around, have a furry coat and so on typically behave in similar ways. It is the behaving-in-a-similar-way that makes generalising possible and category formation worthwhile. If four-leggedness, barking and so on have been the experience of dogs for that person then they might categorise this animal as a dog
(until they have evidence to the contrary). Given the importance of categories then, a critical research question is to understand how category knowledge is represented.

Traditionally, theoretical approaches to the problem have either taken a prototype view (see Rosch, 1973, 1975; also Hampton, 1993, 1995) or exemplar view (Brooks, 1978, 1987; Hintzman, 1986; Medin & Schaffer, 1978). Continuing with the dog example highlights the difference between the two views. The prototype view would claim that knowledge about dogs in general is used classify this new animal – in particular, knowledge consists of a kind of average of dog features, built up from having seen many dogs in the past. The exemplar view would claim that this dog evoked memories of one or more specific earlier dogs and that you used the similarity of the new animal to your memory of dog exemplars to decide that this new animal was a dog.

The most important difference between the models rests on the question of abstraction, specifically, at what level of abstraction is it necessary to represent category knowledge. Prototype representations can be thought of as containing exemplar representations (i.e., they have all the exemplar information plus abstractions), so in this sense the exemplar model is more parsimonious and therefore to be preferred in circumstances where it describes the data equally as well as the prototype model. However, if there are regularities in a category that occur across instances, in many cases it would seem plausible, and indeed very useful from a categorisation point of view, to encode these regularities at some level for aiding the classification and use of new instances. Therefore, whether we need the additional level of abstract representation that the prototype model offers will most likely depend upon the category structure and the nature of the categorisation task. Overall the evidence favours a view in which both exemplars and more abstract representations
are used in categorisation (e.g. Elio and Anderson, 1981; Estes, 1986; Homa, 1984; Malt, 1989; Medin, Dewey & Murphy, 1983; see also review by Ross & Makin, 1999). Very much in line with cognitive and construction grammar, this recognises that people generalise over category instances and that they also retain a lot of item specific information (see the symbolic assembly developed in Chapter 2, specifically Figure 6). In line with the arguments presented in Chapter 2, exemplars and abstractions may represent points along a continuum and that the same set of processes (analogy, categorisation, schematisation = pattern finding) could use either type of representation in classifying new instances.

3. **Linguistic categories**

A defining feature that sets cognitive linguistics apart from other approaches to studying language is that it attributes symbolic status to grammatical constructions (Croft 2001; Fillmore 1989; Fillmore et al., 1988; Goldberg 1995, 2006; Kay and Fillmore 1999; Lakoff 1977; Langacker 1987). The problem for language acquisition is that children must (re-)construct for themselves the symbolic assemblies of their language from the individual utterances they experience. Of particular importance as targets of acquisition are the relatively abstract utterance-level constructions that enable children, at some point, to generate an almost unlimited array of particular utterances following the same general form: for example, the transitive construction.

Cognitive linguistics also has a more general commitment. Aspects of language structure, function and systematicity must be consistent with other facets of cognitive processes. An interesting empirical question follows from this commitment: are the relatively abstract argument-structure categories, like transitive, subject to the same categorisation processes as other non-linguistic abstract categories? There is
reason to suspect they might be as sorting tokens of experience into separate
categories and forming analogies between them is such a basic and pervasive
cognitive activity. Given this, categorisation processes are obvious psychological
candidates where we would expect similar processes to operate in linguistic and non-
linguistic domains.

The theoretical stance one takes on the nature of categorisation and
generalisation obviously affects one’s view of what exactly is learnt. A long line of
logic-orientated rationalists (Descartes, Leibnitz, Russell, Frege, Fodor and Chomsky)
have sympathised with the Platonic tradition of categories as discrete and well
bounded idealisations, viewed through a lens of (more-or-less interesting)
psychological performance. By contrast, for Wittgenstein (1953) there is no level at
which speakers calculate the necessary-and-sufficient conditions in order to
understand what concepts mean. For nearly all cases the meaning of a word is its use
in language. In a complete metaphysical reversal from his earlier position (1921), he
argued that rather than the structure of reality determining the structure of language,
instead the structure of the language determines what we think of as reality. Thus,
what a word or a sentence means is no longer what it pictures (depicts) but rather the
sum total of its possible uses.

Prototype theory is a psycholinguistic model of conceptual organisation that
navigates between these positions with features that are somewhat reconcilable with
both sides of the debate (Collins and Quillian 1972; Posner and Keel 1968; Rosch
1973, 1975). For example, multiple criterial features and graded membership appeals
to Wittgenstein’s dynamic aspect of natural types, providing for variation and change,
whereas a strong association between features (having one feature implies having
many of the others) and strong clustering around the mean (the majority of members
are ‘closer to the prototype’) is consistent with the Platonic idea of discreteness, stability and cognitive coherence.

Prototype theory has been deployed in a wide range of linguistic contexts, used in various different ways and to various extents to account for lexical semantics and grammatical constructions (Lakoff 1987); tense-aspect marking (Shirai & Andersen 1995; Andersen & Shirai 1996); relative clauses (Diessel and Tomasello 2005); questions with long-distance dependencies (Dąbrowska, Rowland & Theakston 2009); subject auxiliary inversion (Goldberg 2006; see also Lakoff and Brugman 1987; Lambrecht 1994); the transitive construction (Chan et al. in press; Dąbrowska & Tomasello, in press; Dittmar et al., 2008); and the lexical reorganisation that leads to semantic over-generalisation and recovery from overgeneralisation, as modelled by an unsupervised neural network (Schyns, 1991).

Whether one believes category knowledge is best described by exemplar or prototype or a hybrid of these models, the variation in the input and the wider cultural context in which concepts are learnt will obviously alter the resulting category content and the thus the generalisations that are made on the basis of the way that content is organised. Lakoff (1987) famously explored the role of culture and categorisation in Women, fire, and dangerous things: what categories reveal about the mind. He discusses the example of Dyirbal, an aboriginal language from Australia which apparently uses just four classes for all things. Whenever a Dyirbal speaker uses a noun it must be preceded by one of just four words bayi, balan, balam or bala. Bayi includes men, kangaroos, opossums, bats, most snakes, some birds, some insects, the moon, storms, rainbows, boomerangs, and some spears. Balan included women, bandicoots (an Australian rat species), dogs, platypuses, echidna (spiny anteaters), some snakes, some fishes, most birds, fireflies, scorpions, crickets, the hairy mary
grub, anything connected with fire or water, suns and stars, shields, some spears, and some trees. Balam included all edible fruit and the plants that bear them, tubers, ferns, honey, cigarettes, wine, and cake. Finally, bala included parts of the body, meat, bees, wind, yamsticks, some spears, most trees, grass, mud, stones, noises, and language.

Lakoff emphasised that linguistic categories derive from the experience the Dyirbal speakers have with the external world. The categories mark what is important to the Dyirbal in the world and help them to conceptualise their experience and to survive. From this analysis of the Dyirbal categories flows a more general hypothesis about the function of categorisation in language. For Lakoff and other cognitive-functional theorists, categories are derived from the interactions and experience of human beings with the world – their situated use – and these interactions are constrained by biological and perceptual biases along with the social-cognitive skills needed to communicate them in particular usage events (Langacker, 1987; Tomasello, 2008). This interpretation, that experience structures our conceptual structure so that language reflects our conceptualisation of the world, constitutes the leitmotif of cognitive linguistics.

4. Categorising argument-structure constructions: the example of the transitive construction

Hopper and Thompson (1980, 1984) laid out the prototypical semantics underlying transitive constructions across many languages. They list what they call the “component parts of the Transitivity notion”, representing a scale according to which clauses can be ranked as more or less transitive; each property on the list can have either a ‘high’ or a ‘low’ value, and the more ‘high’ values a clause shows, the higher it is in Transitivity. In Figure 1, A refers to the most agent-like participant of a two-
participant clause, O refers to the second, non-agentive participant of a two-participant clause.

Figure 1. *Transitivity Features* (Hopper and Thompson, 1980).

<table>
<thead>
<tr>
<th>Component</th>
<th>HIGH</th>
<th>LOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. PARTICIPANTS</td>
<td>2 or more participants, A and O</td>
<td>1 Participant</td>
</tr>
<tr>
<td>B. KINESIS</td>
<td>Action</td>
<td>Non-action</td>
</tr>
<tr>
<td>C. ASPECT</td>
<td>Telic</td>
<td>Atelic</td>
</tr>
<tr>
<td>D. PUNCTUALITY</td>
<td>Punctual</td>
<td>Non-punctual</td>
</tr>
<tr>
<td>E. VOLITIONALLITY</td>
<td>Volitional</td>
<td>Non-volitional</td>
</tr>
<tr>
<td>F. AFFIRMATION</td>
<td>Affirmative</td>
<td>Negative</td>
</tr>
<tr>
<td>G. MODE</td>
<td>Realis</td>
<td>Irrealis</td>
</tr>
<tr>
<td>H. AGENCY</td>
<td>A high in potency</td>
<td>A low in Potency</td>
</tr>
<tr>
<td>I. AFFECTEDNESS OF O</td>
<td>O totally affected</td>
<td>O not affected</td>
</tr>
<tr>
<td>J. INDIVIDUATION OF O</td>
<td>O highly individuated</td>
<td>O non-individuated</td>
</tr>
</tbody>
</table>

Developing this idea further Næss (2007) condenses these ten component parts into the key notion of syntactic prominence, in the sense of the encoding of noun phrases as independent syntactic arguments. This reflects pragmatic prominence and the desire to draw particular attention to certain aspects of an event. Naess proposes that the distinctness of participants as a basic criterion for transitivity is not only a distinctness of physical entities, but a distinctness of the semantic roles involved in a transitive event. The semantic roles that are played by the participants of a transitive event can be formulated in the maximally distinct argument hypothesis (Naess 2007: p.7):
A prototypical transitive clause is one where the two participants are maximally semantically distinct in terms of their roles in the event described by the clause. This event is prototypically realised as an agent intentionally instigating an action that directly results in the patient being affected. It is worth briefly exploring what is meant by these terms as they are critical to the theory and the experimental work to follow. The relevance of intentionality or volition to the concept of agency has been subject to some debate, with terms such as volitionality, control and causation recurring in discussions about typical agent properties (e.g., Delancy, 1984; Givon, 1985; Langacker 1991). Here the focus is on defining intentionality as a relational property between agent and patient rather than an inherent property of a participant itself. The difference can be illustrated with an example sentence, John murdered Peter. On the assumption that Peter refers to a human being, he is in possession of volitionality as an inherent property of human beings, but presumably he does not exercise this volitionality where his role in this particular event is concerned; more importantly, his being volitional is not a prerequisite for his being able to fill this role. By contrast, where John is concerned, his involvement in the event requires the active exercise of volitionality, as murdering someone is an act which necessarily involves a volitional agent; in other words, John’s role in the event is partly defined by the property of volitionality. So what is at issue here is the exercise of a capacity to interact with an event in a certain way. This definition is intended to include experiencer arguments of verbs of cognition or perception; it is a prerequisite for a participant’s functioning as an experiencer that its cognitive capacities are actively engaged in the event (e.g. Langacker 1991: 238-9, where agents and experiencers are grouped together as ‘active participants’). The exercise of volition in carrying out an
event may be seen as the (proto)typical way in which participants involve their
cognitive capacity in interacting with an event, even though this is not the only
possible way.

The property of instigating or causing an event is central to our whole
understanding of what an agent is; a simplistic description of a transitive event might
refer to an act where one participant is ‘doing something to’ another, where we
understand the participant ‘doing something’, that is, instigating the event, as the
agent. Thus the prototypical transitive has its agent causing the event, rather than
some other intermediate, external source.

As far as the notion of patient is concerned, the main property typically taken
to define this category is that of affectedness; a patient is generally defined as a
participant which in some way undergoes a change of state as a result of the event.
Affectedness of the patient is a central factor in the encoding of two-participant
clauses in a number of languages. For instance, Finnish encodes highly affected
objects in fully transitive clauses with accusative case, but less-affected objects take
oblique (partive) case-marking, whether their reduced affectedness stems from only
part of the object being affected (1.1a-b), or the nature of the verbal action being such
that it impinges less strongly on the object (1.1c-d):

1.1 Finnish (Uralic, Finno-Ugric; Kittilä 2000: 113-114):
a. Hän jo-i maito-n
S/he drink-PAST.3SG milk-ACC
‘S/he drank (all) the milk’
b. Hän jo-i maito-n
S/he drink-PAST.3SG milk-PART
‘S/he drank (some) the milk’
c. Hän tappo-i miehe-n
S/he kill-PAST.3SG man-ACC
‘S/he killed the man’
d. Hän lö-i mies-tä
S/he hit-PAST.3SG man-PART
‘S/he hit the man’

The two-participants of a transitive clause are most commonly labelled ‘agent and ‘patient’. Therefore, another way of formulating the maximally distinct argument hypothesis would be to say that the agent and patient categories should be defined in maximal opposition to each other. Intentionality, instigation and affectedness can be construed as binary properties which can be assigned the values of ‘+’ or ‘-’. In reality the picture is, of course, rather more complex; affectedness in particular is clearly a gradable concept, and we can conceive of entities as more or less affected. Feature assignment should therefore be understood as referring to high or low values on what is in principle a semantic scale, that is ‘+’ essentially means ‘showing this property to a high extent’, and a ‘-’ ‘showing this property to little or no extent’. The semantic opposition of agent and patient can be represented schematically as follows:

Figure 2. *Agent and Patient as maximally distinct categories.*

<table>
<thead>
<tr>
<th></th>
<th>Agent</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intentionality</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Instigation</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Affectedness</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>
The empirical predictions of Naess’ model are fairly straightforward: any deviation from the feature configuration presented in Figure 2 will lead to the use of a structure distinct from the fully transitive clause in some language. Conversely, if Figure 2 is to be taken as a sufficient definition of prototypical transitivity, then any differences in formal transitive marking must in principle be explainable in terms of deviations from this feature configuration. So, there should be better-or-worse examples of transitivity where the sentence *John broke the plate* semantically overlaps with all the prototypical features described by Næss, while *John rolled down the hill, John accidentally broke the plate, John entered the room, John didn’t break the plate (but wanted to)* all depart from the prototype along the dimensions of intentionality, instigation and affectedness of the agent or patient. If we distort the transitive prototype of Figure 2 we see various permutations of the core semantic elements shown in Figure 3.

**Figure 3. ‘Distortions’ of the Transitive Prototype according to the features defined by the maximally distinct argument hypothesis.**

<table>
<thead>
<tr>
<th></th>
<th>John broke the glass (on purpose)</th>
<th>John didn’t break the glass (and didn’t intend to)</th>
<th>John broke the glass (accidentally)</th>
<th>The ball broke the window</th>
<th>John loves Mary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intentionality</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Instigation</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Affectedness</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Naess (2007) describes other prototype deviations within his framework which were not included in the current study for the following reasons. Volitional Undergoers,
describes the concept of a patientive participant that voluntarily submits to some action. The only cited examples are from Icelandic and Sanskrit and apparently these languages mark for this semantic category, but when translated the examples feel more than a little contrived e.g. *the elephant lets itself be mounted*. Another category I have avoided is the Frustrative, that is a participant which is volitionally involved in some state of affairs but does not actually instigate anything. Cited examples are from Hindi, Meithei (Sino-Tibetan), Malayam (Dravidian) and Tsez (Nakh-Daghestanian) which mark for this category. Again the examples do not translate easily into the kind of transitive I will be using throughout the rest of the experiment because the meaning can only be conveyed in English with additional lexical content rather with than case or agreement (e.g. *I did not (have not been able to) read this book*). The meaning of this construction in these instances is not just that of not doing something, but of not being able to do something – the participant wants to act but for some reason is not able to carry out the action.

In summary, prototypical transitivity is defined primarily as a semantic/pragmatic concept, but one with obvious structural implications, of which, the clearest are the traditional grammatical relations subject and object (which may themselves be generalisations from the core notions of agent and patient).

5. Experimental Evidence for Prototype Representations

Bransford and Franks’ (1971) classic study in prototype formation was primarily concerned not with memory for individual words or sentences but rather with memory for wholistic, semantic ideas. Specifically, they were interested in whether participants could acquire a complete idea from exposure to partial ideas, and whether they would think that they had heard that complete idea before when in fact they had
not. They presented adults with sentences such as *the sweet jelly was on the table* and *the ants in the kitchen ate the jelly*. There then followed an unexpected recall test. Participants were asked whether they recognised sentences presented previously as well as novel sentences that were agglomerations of these ideas, such as *the ants in the kitchen ate the sweet jelly which was on the table*. These complex ideas resulted from integrating information expressed by many different sentences experienced successively and often non-consecutively over time, that is, the process involved the abstraction of linguistic ideas. The focus here is on memory for sets of sentences expressing common semantic content. The ‘prototype idea’ is defined as the stimulus pattern which on average is the most similar to all of the acquisition stimulus patterns. Similarity is defined by the number of transformations required to generate a stimulus pattern from one to which it is being compared.

The results showed that adults spontaneously stored together in memory related semantic propositions that had been presented individually and in no logical order. They were most confident that they had heard the complex prototypical idea before, when in fact they had not. This means they acquired something more general or abstract than simply a list of those sentences experienced during acquisition. They integrated the information communicated by sets of individual sentences to construct wholistic semantic ideas.

Franks and Bransford (1971) also conducted an analogous experiment with visual input. They constructed stimuli by combining geometric forms such as circles, stars, and triangles into structured groups of various kinds (Figure 4). As before, some of these were then shown to participants - who were then later asked if they recognized these and other shapes they had not seen previously. Importantly, one of the exemplars shown at test contained all of the geometric forms together, an
exemplar that had actually never been shown previously (but could be considered the prototype if all of the experienced exemplars were averaged). The participants not only thought that they had seen this prototype, but they were actually more confident that they had seen it than the other previously seen exemplars (or distracter items which they had not seen). The idea of averaging across exemplars contains within it a notion of frequency. The conceptual space that prototypes occupy includes the most representative member but may also be distorted, skewed or weighted towards its most frequent members. Importantly, a prototype assigns membership to a category by means of a judgement of similarity to a central exemplar so that an essential property of a prototype category is that it is gradable.

Figure 4. Examples of stimuli and results of the Franks and Bransford experiment. Note how confidence that participants saw an item declined systematically as the distortion from the prototype information increased. Note also that the highest confidence ratings were given to prototype stimuli (A) even though none of these were presented during learning.
Lasky (1974) tested the ability of six year-olds, eight year-olds, and adults to abstract visual patterns in what was essentially a replication of the Franks and Bransford (1971) visual task with children. Lasky used geometric forms in a prototype-plus-distortions design. In the acquisition phase participants were presented with various...
shapes and ask to reproduce them. In the recognition phase they were asked whether they had seen them before, rating their confidence on a scale from -5 to +5 or a happy-sad face scale for the children. There were six acquisition patterns and fifteen recognition patterns – three identical to acquisition and twelve novel items (including the prototype). They replicated the findings of Franks and Bransford with eight year-olds and adults but failed to find prototype effects in the six year-olds: unlike the other two groups, the youngest children’ ratings were a function of the number of transformations the stimulus patterns were removed from the prototype. Lasky suggests the reason for this is that the 6 year-olds abstracted, and hence were only productive, with part of the prototype.

Boswell and Green (1982) investigated the abstraction and recognition of prototypes with four and six year-olds as well as adults. The authors used line-drawing polygons in a prototype-plus-distortions design. They were particularly interested in exploring the role of specific experienced exemplars versus a generalised, abstracted prototype. Five figures of prototype distortions were chosen as a learning set. The participants had to say whether each shape belonged to one of two ‘extraterrestrial space families’, which both had nonsense names. They were given feedback as to whether they had correctly guessed the right name. Then they were randomly assigned to one of two groups. Group REM received instructions stressing the accuracy in identifying only those figures that had been previously presented. Group CAT was instructed to identify those figures to which they had been exposed, and also to indicate any new “family members” which they had not met. Participants were instructed that figures which seemed inappropriate to either family should be classified as ‘strangers’. They were then assessed using all 10 acquisition figures, the
10 new categorical distortions, the 2 prototypes and the 4 foils. The different configurations of shapes were defined as follows:

**Prototype**: most representative of the experienced distortions

**New**: a distortion that was not in the acquisition phase

**Old**: a distortion that was experienced in the acquisition phase

**Foil**: not derivable from the distortions

They found that adults failed to discriminate Olds from Prototypes, even with accuracy instructions, suggesting that attention to the commonalities among figures during prototype construction may function to effectively limit processing of specific item identification. In contrast to the robustness of the prototype categorization effect for adults, children in the REM condition differentiated Olds from Prototypes, News, and Foils, which indicates the availability of specific exemplar information not evidenced by adults. The authors conclude children’s knowledge appears to encompass both prototypic and exemplar-specific knowledge, which may be selectively accessed in relation to the demands of the concept assessment task (REM or CAT). In contrast adults were unable to differentiate the prototype as non-acquisition members. The children’s capacity for greater discrimination than adults, evidenced by their rejection of NEWS and PROTOTYPES in the REM condition, is subject to their control. When instructed to generalise (CAT), children categorise the same stimuli with great accuracy, moreover, they seem to have greater access to exemplar specific knowledge, and are more flexible at processing this information.

Most of these early studies used the prototype-plus-distortion method with graphical stimuli because of certain perceived methodological problems with its
application to the study of linguistic information processing. Lasky (1974: p. 4) notes that “[the] utilization of natural language categories may yield results that are significantly influenced by the idiosyncrasies in the individual’s past experience. The problem is particularly acute with regard to the limitations of children’s experience…the child with a pet St. Bernard may have a radically different concept of dog than the child with a pet Chinese Crested Hairless”. Using visual stimuli affords control of children’s experience with these categories. This perceived limitation made the application of the prototype-plus-distortion method to questions of language acquisition seem problematic.

Casenhiser and Goldberg (2005) investigated how children map a novel meaning onto a new phrasal form, and how this process could be facilitated by constructions that are centred around prototypical verbs, such as give in the ditransitive construction. 51 five- to seven-year-old children watched a short set of video clips depicting objects appearing in various ways. Each scene was described using a novel verb embedded in a novel construction. The novel construction NP1 NP2 NovelVerb + O was indicative of the meaning of APPERANCE. The different verbs indicated manner of appearance. For example, children heard the spot the king moopo-es, they then saw a video-clip where a spot appears on the king’s nose, and then they heard the spot the king moopo-ed. There were 5 novel verbs and 16 examples of the construction so that in the balanced condition the verb types were arranged as: 4-4-4-2-2 and in the skewed frequency condition they were arranged as: 8-2-2-2-2. There was also as control condition with no sound.

Children who watched the videos and heard the accompanying description were able to match new descriptions that used the novel construction with new scenes of appearance. There was a facilitative effect for the disproportionately high
frequency of occurrence of a single verb in a particular construction (such as has been
found to exist in naturalistic input to children (Goldberg, 2006: 3). The authors
concluded that the input is structured in such a way as to make generalization of
constructions straightforward and associations are made between the meaning of the
dominant verb in the construction and the construction itself.

6. Constructing Prototypical Transitive Semantics in Development

The following experimental study integrates the theoretical and experimental work
discussed above in several ways.

1. To the extent that Naess (2007) has correctly captured the semantic prototype
   of the transitive clause with the feature-decompositional account it should be possible
to empirically validate this and track its development in language acquisition.

2. To test this we will use the well established prototype-plus-distortion method.
   Specifically we will use the procedure of Acquisition Phase followed by Recognition
   Phase using Old, New, Prototype and Foil stimuli.

The novel aspect to this study is that no-one has used this methodology applied to
questions of language acquisition. The question of whether children are able to use
knowledge of the abstract transitive semantics in order to perform prototypicality
judgements feeds into a wider and more fundamental question for any usage based
theory: how children get from more concrete and item-based constructions to more
abstract constructions. There is evidence to suggest that early in language
development abstractions are formed initially on the basis of frequent overlapping
cues, so that the initial abstraction constitutes what will eventually be the prototype of a more complex, more abstract category (see review by Ibbotson and Tomasello, 2009, Chapter 3). This account acknowledges a) the importance of high-frequency forms in providing ‘anchor-points’ from which more abstract generalisation will gradually emerge and b) that the construction redundantly marked with multiple cues could have a special status as a nucleus around which the prototype forms.

There also seems good reason to use the transitive clause as a test case construction. Ontogenetically the transitive is important as it is one of the most basic constructions in which comprehension rests crucially on being able to successfully identify which participants are playing which roles in the event (who is doing what to whom). Secondly, it codifies a fundamental pattern of human experience; “In many languages (and perhaps covertly in all languages) the transitivity relationship lies at the explanatory core of most grammatical processes” (Hopper and Thompson 1982). So if this is true and acquiring the transitive is a key building block in constructing one’s language, the way in which children are building lexical, syntactic, semantic and phonological prototypes needs to be much more deeply understood. The kinds of processes that this study is investigating are also of relevance to the role of similarity in the development of categorization. For instance, how do mechanisms of inductive generalization change in the course of development? What are the origins and the developmental course of abstract relational concepts?

The following paper is currently submitted and is authored by Paul Ibbotson, Anna Theakston, Elena Lieven, and Mike Tomasello.
Publication 3

Prototypical Semantics of the Transitive Construction:
Developmental Comparisons.

Page: 166
Abstract

There is much controversy about whether linguistic categories and schemas are acquired in the same basic manner as categories and schemas in other domains of human cognition. In this study we investigated the English transitive construction for evidence of the kind of prototype effects characteristic of non-linguistic categories, in both adults and young children. Adapting the prototype-plus-distortion methodology of Bransford and Franks (1971), we found that whereas adults were lured towards false positive recognition of sentences with prototype transitive semantics, young children showed no such false positive recognition. Thus the children were, counterintuitively, better than the adults in this task, which suggests an interesting new fact about how linguistic prototypes are formed ontogenetically. Children recall individual sentences better than adults because the adults are assimilating the sentences to their already formed prototype construction, effectively limiting the processing of specific items, whereas young children do not have such a strong pre-existing construction and so focus more on the individual sentences. This suggests that acquiring syntactic constructions is a process of categorisation similar to that occurring in other cognitive domains, and that the initial developmental stages of the process are more dominated by exemplar-specific information.
1. Introduction

Humans are flexible and pervasive categorisers, with many everyday concepts being indicated by linguistic conventions. In virtually all linguistic theories, lexical and morphological conventions, indicated by words and grammatical morphemes, are formed by domain general and well-known processes of categorisation such as distributional analysis and analogy. In some linguistic theories, however, syntax is something different, comprising “rules” that are not a result of categorisation processes (e.g., Pinker, 1999). In cognitive linguistics approaches, in contrast, syntactic constructions such as the transitive or ditransitive are considered complex categories or schemas. As such, syntactic constructions should display many of the characteristics of other kinds of categories, including lexical categories, for example, prototype effects. Thus, Goldberg (1995) argues for a network of ditransitive constructions, with instances like “He gave her the book” being prototypical based on linguistic evidence like frequency and markedness.

A basic construction in almost all the world’s languages is the transitive construction, as in “He kicked the ball”. Building on Hopper and Thompson’s (1980, 1984) investigation of transitivity across the world’s languages, Naess (2007) proposes that the prototypical transitive sentence is characterized by the ‘maximally distinct argument hypothesis’: a prototypical transitive clause is one where the two participants are maximally semantically distinct in terms of their roles in the event described by the clause. Thus this ‘motion event’ (Talmy 1985: 85) is prototypically realised as an agent **intentionally instigating** an action that directly results in the patient being **affected**. In line with the gradable nature of concepts advocated by prototype theory, there should be ‘better-or-worse’ examples of transitivity. For
example, the sentence *John broke the plate* semantically overlaps with all the prototypical features described by Næss, while *John dropped the plate accidentally*, *John received the plate as a present*, and *John saw the plate*, are all distortions from the prototype along the dimensions of intentionality, instigation, and affectedness of the patient.

Although there have been many analyses in cognitive linguistics of the prototypical features of various linguistic constructions (e.g., Taylor, 1995), there are almost no experimental studies. Closest are the classic studies of Bransford and Franks (1971) who looked at semantic propositions. They argued in general that the prototype comprises a maximal number of features common to the category, often "averaged" across exemplars. They were interested to see whether participants would falsely recall a full proposition when they had only seen some parts in various admixtures. They presented adults with sentences such as *the sweet jelly was on the table* and *the ants in the kitchen ate the jelly*. There then followed an unexpected memory test. Participants were asked whether they recognised sentences presented previously as well as novel sentences that were agglomerations, such as *the ants in the kitchen ate the sweet jelly which was on the table*. These complex propositions resulted from integrating information expressed by different sentences experienced successively and often non-consecutively over time. The ‘prototype idea’ was defined as the proposition which on average was the most similar to all of the propositions, and similarity was defined by the number of transformations required to generate a proposition from one to which it was being compared. Results showed that adults spontaneously stored together in memory related semantic propositions that had been presented individually and in no logical order. In the memory test they were most confident that they had heard the complex prototypical idea previously, when in fact
they had not. This means that they acquired something more general or abstract than simply a list of those propositions experienced during acquisition. They integrated the information communicated by sets of individual propositions to construct prototypical propositions.

In the current study, we used the prototype-plus-distortion methodology of Bransford and Franks to investigate experimentally whether syntactic constructions – specifically, the transitive construction – are subject to prototype effects. Moreover, because in some theories of language acquisition development proceeds from more item-based and exemplar-based constructions to more general and abstract constructions (e.g., Abbott-Smith & Tomasello, 2002; Tomasello, 2003) we tested both adults and children to see whether prototype effects might differ as a function of stage of development.

Very little is known about prototype effects across development, and what is known concerns non-linguistic materials. Lasky (1974) tested the ability of six year-olds, eight year-olds, and adults to abstract geometric patterns in what was essentially a visual analogue of the Bransford and Franks (1971) experiment. They replicated the prototype effect with eight year-olds and adults but failed to find prototype effects in the six year-olds. Lasky suggested – based on a follow-up test - that the reason the six year-olds didn’t show strong false positive recall on the prototype was that they abstracted only one part of the prototype. Boswell and Green (1982) explored the role of specific experienced exemplars versus a generalised, abstracted prototype with four and six year-olds as well as adults. The configurations of line drawings they used to investigate this were: **Prototype**: most representative of the experienced exemplars; **New**: a distortion that was not in the acquisition phase; **Old**: a distortion that was experienced in the acquisition phase; **Foil**: not derivable from the distortions. Adults
failed to discriminate olds from prototypes, regardless of increased exposure to the acquisition set or level of task complexity, suggesting that attention to the commonalities among figures during prototype construction may function to effectively limit processing of specific item identification. Children, in contrast, differentiated olds from prototypes, news, and foils, which indicates the availability of specific exemplar information not evidenced by adults.

The current study thus investigated two questions. First, we wanted to know if the transitive construction would show prototype effects with the prototype-plus-distortion methodology. If it did, this would provide experimental evidence that syntactic constructions are a result of domain general categorisation processes. Second, we wanted to know if prototype effects with the transitive construction were similar at different stages of development. If, as in the studies of non-linguistic categorisation, adults showed prototype effects whereas children did not (i.e. the children have better memories for specific sentences), this would provide suggestive evidence that children’s syntactic constructions retain more exemplar- and item-based information than do those of adults.

2. Methods

2.1 Participants

Adults: Thirty-nine monolingual English-speaking undergraduates from the University of Manchester took part in the study (29 women, 10 men: age range 18-43 years old, Mean = 18.8).

Children: Thirty monolingual English-speaking children (20 boys; 10 girls: age range 3.9-4.9, Mean = 4.2) took part in this study. A further four children were tested but
were excluded from the analysis due to not using the scale during the recognition phase (1), saying that they remembered all of the test sentences (2) or none of them (1). Testing took place in a nursery in Manchester, UK.

2.2 Materials

According to the maximally distinct argument hypothesis, anything that reduces the distinguishability of participants reduces its transitivity. We can think of ‘transitivity reducing transformations’ as distortions away from the prototypical semantics of the transitive. In the current study there are 3 semantic categories that are distortions of the transitive prototype which are discussed below.

**Distortion 1. Instrument:** an instrument is an entity being manipulated by another entity in order to achieve an effect on a third entity. For example, *The hammer broke the window*. This is a distortion from the prototype in the sense that *the hammer* has neither intentionality nor instigated the action. Thus the only overlap with the prototype is that it results in an affected patient.

**Distortion 2. Force or Involuntary Agent:** This category covers self-driven natural forces with no capacity for volitional action, that is, they bring about events by virtue of their own inherent power. It also includes human or animate actors which are capable of volitional action but whose involvement in the particular action in question is non-volitional. Example: *The wind closed the door* or *John (accidentally) broke the plate*.

**Distortion 3. Neutral:** Neutral is an object which is not directly involved with the event either in terms of participating in its instigation or in registering its effect. This is different from the above as ‘force’ or ‘non-volitional agents’ are not in full control of the event but nevertheless contribute to it being brought about. In ‘neutral’
the object may be presented as relatively unaffected compared to other, highly affected objects. Example: *Peter climbed the mountain*. Alternatively, the clause may be negated so that no actual effect is produced (*John didn’t hit Mary*). The idea is not whether intentionality, instigation and affectedness are present or absent (+ or -) but whether there is a contrast in these semantics in the relationship between agent and patient. The way the methodology is hypothesised to work is by ‘priming’ the semantics of the distortions around prototype in the acquisition phase and then presenting the prototype as a ‘target’ in the recognition phase.

*Test Sentences for Adults:* There were two phases to the experiment. In the **Acquisition Phase** there were 6 examples of each of the 3 semantic categories that represent distortions from the prototype (instrument, force/involuntary action and neutral). There were also 4 foils. In the **Recognition Phase** there were 2 examples of each of the distortion categories that appeared in the Acquisition Phase which are labelled as ‘old’ and 2 examples of each of the distortion categories that were not in the Acquisition Phase labelled as ‘new’. In this sentence type the lexical items were changed but the underlying semantics is the same with respect to the semantic category they belong to (instrument, force/involuntary agent, or neutral). In the Recognition Phase there were also 4 examples of the prototype which importantly were not in the Acquisition Phase. Finally, there were also 4 new foils in the Recognition Phase. These were included to provide an additional benchmark to judge false positive recall.

Clearly, the more lexical items a sentence in the Recall phase shares with items in the Acquisition phase, the more confident participants would be that they have heard that sentence before regardless of semantics. Therefore the degree of lexical overlap between *new* and *prototype* (the critical comparison) was
counterbalanced by (i) the total number of items in that grammatical slot (S, V or O) in the acquisition phase and (ii) all items in the acquisition phase.

Test sentences for Children: The design of prototype-plus-distortion plus recognition test was exactly the same for the children as it was for the adults, the only difference was that the number of items in the Acquisition phase was reduced from 6-items per semantic category to 4 items in order to make the test length more appropriate for the age group.

2.3 Procedure

Adults:

Acquisition Phase. Adults were instructed that they should read a list of sentences aloud into a microphone. In order to prevent adults anticipating a memory test participants were told that “we are interested in recording samples of people reading these sentences aloud for a language processing study”. This was irrelevant to the aims of the study but was designed to encourage them to pay attention to each sentence as well as stop them mentally rehearsing the test items.

Recognition Phase. After participants had read all sentences they were then given a 5-point Likert-style scale: | definitely not heard before | probably not heard before | unsure | probably heard before | definitely heard before |. Participants were told by the experimenter “now I’m going to read some more sentences and I want you to tell me whether you’ve heard them before from the list I have just read out. I want you to tell me how confident you are that you’ve heard the sentence before. I’d like you place a mark in the box that most closely matches your level of confidence that you have heard the sentence before”. The experimenter then read the sentences
aloud. After the rating sheets were collected in participants were debriefed as to the true purpose of the task.

*Children:*

*Acquisition Phase.* In the first half of the experiment the child played a sentence repetition game in which s/he was asked to copy sentences that the experimenter produced. Every time s/he copied a sentence s/he was given a sticker to maintain interest.

*Recognition Phase.* There then followed a recognition game where the experimenter read out more sentences, some of which were from the previous list and some of which were sentences the child had not heard before. The child’s task was then to judge whether or not they thought that they had heard the sentence before. They did this by placing a counter on a 5-point scale that represents how ‘happy’ they were that they had heard the sentence before (Ambridge, Pine, Rowland & Young, 2008). This was intended to be comparable to the adult 5-point Likert-style scale.

If children thought they had heard the sentence from the previous list, they placed the counter somewhere on the right-hand side depending on how confident they were, if they thought it was new they placed it somewhere on the left-hand side and if they didn’t know they placed it in the middle. The scale was modelled by the experimenter in a brief ‘memory game’ played between the Acquisition and Recognition phase. This involved the experimenter, with the child, looking at pictures of monsters printed on a card (intended to be analogous to the Acquisition phase). A second, different set of monsters was then used and the experimenter had to remember which ones they had seen from the previous card (analogous to the Recognition phase). Some monsters on the second card appeared on the first card and some didn’t; others had various degrees of resemblance to monsters on the first card. All points of
the scale were used during this demonstration. The experimenter then said “Now you are going to play the game with words instead of pictures”. The Recognition phase then began.

3. Results

The outcome measure was the confidence rating obtained in the recognition phase of the experiment. The matrix bar charts below show the percentage response for each point on the rating scale across prototype, old, new and foil conditions.

Figure 1. Percentage confidence rating across conditions for Adults.
Figures 1 and 2 show that both adults and children were able to discriminate between new sentences (sentences that were not in the Acquisition phase) and old sentences (sentences that were in the Acquisition phase). The interesting and critical differences lie in the way adults and children treated the prototype condition, which was only presented in the Recognition phase and so methodologically identical to new sentences. In summary adults’ responses on prototype items are skewed towards recognising these sentences as old whereas children’s responses indicate that these are new.
Median confidence rating across conditions for Adults (Figure 3a, left) and median confidence rating across conditions for Children (Figure 3b, right).

Figure 3a/b summarises the distributional frequencies of results by taking the median response for each of the rating points across conditions. Figure 3a clearly shows that adults are treating prototype sentences more similarly to old sentences than they are to new sentences. Conversely Figure 3b shows that children are on average able to discriminate between old and prototype items.

A Kolmogorov-Smirnov test of normality showed prototype, old, new and foil conditions to be all significantly non-normally distributed (p<0.001). Wilcoxon Sign Ranked Tests were therefore used to assess significant differences in the way participants rated the sentences across conditions.
Table 1. Comparisons between prototype, old, new and foil conditions.

<table>
<thead>
<tr>
<th></th>
<th>Prototype-Old</th>
<th>Prototype-New</th>
<th>Prototype-Foil</th>
<th>Old-New</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adults Median Pair</td>
<td>4-5</td>
<td>4-1</td>
<td>4-2</td>
<td>5-1</td>
</tr>
<tr>
<td>Adults Z</td>
<td>-2.45</td>
<td>-9.73</td>
<td>-7.94</td>
<td>-13.90</td>
</tr>
<tr>
<td>Adults Significance (2-tailed)</td>
<td>p=0.014**</td>
<td>p&lt;0.001**</td>
<td>p&lt;0.001**</td>
<td>p&lt;0.001**</td>
</tr>
<tr>
<td>Adults Effect Size</td>
<td>-0.20</td>
<td>-0.78</td>
<td>-0.64</td>
<td>-0.81</td>
</tr>
<tr>
<td>Children Median Pair</td>
<td>1-5</td>
<td>1-1</td>
<td>1-1</td>
<td>5-1</td>
</tr>
<tr>
<td>Children Z</td>
<td>-8.45</td>
<td>-0.005</td>
<td>-2.43</td>
<td>-8.20</td>
</tr>
<tr>
<td>Children Significance (2-tailed)</td>
<td>p&lt;0.001**</td>
<td>p=0.996</td>
<td>p=0.015**</td>
<td>p&lt;0.001**</td>
</tr>
<tr>
<td>Children Effect Size</td>
<td>-0.77</td>
<td>0.00</td>
<td>-0.22</td>
<td>-0.75</td>
</tr>
</tbody>
</table>

Table 1 shows that, for the adults, all pairs are significantly different, including, crucially, prototype-new, indicating a significantly greater “recognition” of prototype sentences than new sentences. For children, they too were able to discriminate between old and new items, so they were paying attention to the stimuli and knew what they had and hadn’t heard. Unlike adults, there was no significant difference between prototype and new item mean recognition and thus no effect of the prototype eliciting false positive recall in the children⁷.

4. Discussion

We found that adults were lured into false positive recognition of sentences with prototypical transitive semantics significantly more than other sentences they had not

⁷ In order to make the test length appropriate for the four year-olds there were four items per distortion category in the acquisition phase compared with six for the adults. Obviously recognition performance will be affected by the overall number of items in the acquisition phase. For that reason, we also tested twenty-eight (different) adults with the same acquisition stimuli as the children. In this condition there was also a significant difference between the recognition rates of new and prototype items (Z = -6.56, p<0.001) albeit the overall effect was weaker, such that the median recognition rates for new items was “definitely not heard before” versus “unsure” for prototype items. Obviously, as the overall number of items to remember is reduced this improves performance in the recognition phase. Even in this condition the prototype items are being treated differently from other new items by adults in a way that was not the case with the four year-olds with the same stimuli.
heard before (new sentences). This effect was not seen in the four year-olds who were able to discriminate between prototype and new items (recall that neither new nor prototype items were in the Acquisition phase).

The fact that we found a prototype effect in the adults is consistent with the idea that abstract syntactic constructions such as the transitive are subject to the same categorisation processes (namely prototype effects) as other non-linguistic abstract categories. This effect persists when overall frequency and possible lexical memory effects are accounted for. That this effect was seen in the prototype sentences – and not other new sentences – gives the maximally distinct argument hypothesis some experimental support, in addition to the typological evidence with which previous accounts have mainly been concerned (Næss, 2007).

Prototype conceptual structure is only one way to achieve prototype effects – see Lakoff (1987) for a discussion on how these effects could arise out of a more indirect interaction between a number of cognitive models, for example salience, radial structures, and metonymy. Leaving aside which of these interpretations is correct, from a cognitive processing perspective, the higher false positive recognition rates in the prototype condition are consistent with the idea that the semantics of the prototypical transitive (Naess, 2007) is relatively more entrenched than those of the distortion categories. The implications are that this leads to sentences with prototypical semantics being easier to retrieve in the Recognition phase when related but different (distortion) sentences have been activated in the Acquisition phase.

It seems that the adult performance with linguistic stimuli reflects similar processing effects to those described by Boswell and Green (1982) in a non-linguistic task, namely, that attention to the commonalities among figures during prototype construction may function to effectively limit processing of specific item
identification. To be clear, the median recognition rate in the prototype was not as high as old items, that is, the adults were still more confident they had heard old items than prototype items. In this sense, the prototype effect is not as strong as in Franks and Bransford’s original study where prototype items were actually recognised more confidently than all other stimuli, including old items. That said, adults in this study are clearly treating prototype sentences differently from other new sentences and more like sentences they have heard in the Acquisition phase.

In line with previous developmental studies that have used this methodology (Boswell and Green 1982; Lasky 1974), the children were better able to discriminate between prototype and new items, or put another way, there was nothing in the prototype and new conditions that caused children to treat them differently in the Recognition phase. Note that previous studies using this methodology with children and adults have primarily used visual stimuli and to our knowledge this is the first study using the prototype-plus-distortion methodology to experimentally explore the semantic nature of a syntactic construction.

Exactly why the children do not show the prototype effect is clearly an issue that needs much further investigation. One possibility is that the accessibility of knowledge along an exemplarPrototype continuum is more susceptible to the demands of the task in children than it is in adults (Boswell & Green, 1982; Ford, 2003). Using a similar prototype-plus-distortion methodology, Boswell & Green (1982) found that when children were asked to remember which stimuli they had and hadn’t heard – an analogous condition to the current study – they successfully avoided false recognition of prototype items whereas adults did not – an analogous result to the current study. It is worth noting that the same pattern of results was obtained using very different stimuli; the participants had to say whether each shape belonged to one
of two ‘extraterrestrial space families’, underlining the domain generality of these effects. When the children were asked to remember they seemed to have greater access to exemplar specific knowledge and were more flexible at processing this information. However, when instructed to generalise the same stimuli in a categorisation task, children’s performance provided evidence of prototype formation. Thus children’s knowledge appeared to encompass both prototypic and exemplar-specific knowledge, which may be selectively accessed in relation to the demands of the concept assessment task. There is evidence that children of a similar age tested in this study (and younger) have prototypical knowledge of this construction (e.g., Pyykkönen, Matthews & Järvikivi, 2009) and more generally show prototype effects in different linguistic constructions and across different languages (Ibbotson & Tomasello, 2009). Therefore because this study involved an explicit memory task, the suggestion is that this encouraged children to focus on and access more exemplar-based information than abstract information. Taken together, the data from this study and others that have investigated abstract knowledge of the transitive construction (e.g., Pyykkönen, Matthews & Järvikivi, 2009; Dittmar et al 2008) suggests children have greater access to exemplar specific knowledge as well as some generalized knowledge of transitives. In this study they were able to recruit this knowledge of what they have and haven’t heard to avoid the false positive recognition (prototype effects) and were encouraged to do so by the recognition nature of the task. Although other studies using the prototype-plus-distortion methodology have found an increase in false recognition with age (Johnson & Scholnick, 1979; Prawatt & Cancelli, 1976) exactly why the adults weren’t able to recruit this information remains unclear but as suggested by Boswell and Green, the attention to the ‘gist’ or commonalities among
figures during prototype construction may function to effectively limit processing of specific item identification.

The fact that children showed better verbatim memory than the adults when distinguishing between prototype and new items is an important fact for theories of language acquisition. Indeed, for any usage-based model of language it is a theoretical prerequisite that learners retain significant detailed memories for language over which generalizations and abstractions may be formed (see Goldberg 2006: 5). Under this framework the grammatical system performs a structuring function, providing schematic meaning. Abstract entities are permissible, but only to the extent that these are schematic for actually occurring structures and can be abstracted from actually occurring instances (Taylor, 2002). Thus, the focus is on language acquisition being driven by linguistic experience and generalisations over usage events. If young children didn’t posses a powerful and subtle ability to encode and store these events it would be hard to see how they could ever build up a threshold of examples that could potentially be grouped under the same category. Nor is it obvious how they could conduct the powerful statistical learning that we know takes place from a very early age (Gerken, Wilson, and Lewis, 2005; Marcus et al., 1999; Saffran, Aslin, and Newport 1996).

In contrast with approaches to language acquisition that emphasise the performance limitations of children, in this study we have shown children to out-perform adults in item-specific memory, and that such good performance is an essential prerequisite for any usage-based theory of language acquisition. We have also, and importantly, found prototype effects for adults at the level of a grammatical category, namely, the transitive construction. We suggest the results are most naturally accommodated in a theoretical framework that views language function,
structure and organisation as deeply integrated with processes that are common to the 
rest of cognition (Goldberg, 2006; Langacker, 1987; Tomasello, 2000).

The following chapter considers further the role of different cues in children’s 
understanding of argument structure by examining the role of pronoun frames in early 
comprehension of transitive constructions in English.
CHAPTER FIVE

1. Introduction

As with previous empirical chapters, the published paper is preceded by a review of the literature relevant to the themes of the study.

A key milestone in any child’s language development is learning to identify and use the linguistic conventions of their language to communicate fundamental conceptual information, such as of who-did-what-to-whom. Within languages the balance between the conventions shifts depending on the sentence’s context, the language’s history, or both. In many languages there are three major syntactic devices for helping to indicate who-did-what-to-whom in the transitive construction: word order (e.g., SVO, OVS, VSO) marking on one or both of the nouns (case), and marking on the verb (agreement). Of course, at some level these cues need to interface with world-knowledge, such as the animacy of the participants (Haspelmath et al. 2005) and semantic plausibility, both of which will contribute to making probabilistic formulations of the event being described.

One way to investigate the interaction of different syntactic cues in different languages is through the theoretical framework of the competition model (Bates & MacWhinney, 1987). Since the 1980s, the model has motivated cross-linguistic research in over fifteen languages with both adults and children. The model provides a framework within which one can quantify the differential ‘weights’ among various combinations of competing and converging cues including word-order, stress, animacy, agreement, topicalisation, prepositions, and case. The model has a functionalist approach to sentence processing and language in general; ‘the forms of
natural languages are created, governed, constrained, acquired and used in the service of communicative functions’. The theory operationalises the strength of a particular cue as the product of how frequently it is present where it could be used to determine a form-meaning mapping (cue availability) and how consistently a function is mapped onto a particular form whenever it is present (cue reliability). Cue strength is also affected by the processing limitations imposed by the perceptual and working memory systems (cue cost). At its simplest, the classical competition model conceptualises language development as a process whereby cue strength (mediated by cue cost) comes to vary as a function of that cue’s availability and reliability in the input.

To empirically test the relative strengths of different cues in any particular language, participants in cue-competition-style experiments typically hear or read sentences in conjunction with some sort of comprehension or production task. For example, participants hear a transitive sentence consisting of two nouns and a verb, the task is then to identify the scene out of a number of options (usually two) that most closely matches their intuition of what the sentence means. Since the 1990s a number of measures of online processing, such as reaction time, have also been developed. As well as using consistency of interpretation as a measure of comprehension as in the offline measures, the idea is that cue interactions in collaboration and in competition will facilitate or inhibit the time course of sentence processing (e.g., Berger, Wulfeck, Bates & Fink, 1996).

Here is a typical set of test sentences from the cue-competition paradigm:

(1) *The horse* is kicking the cow
(2) *The dog the cat* is chasing
(3) Is kissing *the boy the girl*
(4) The ball is pushing the elephant
(5) The tiger are chasing the bears
(6) The rocks is hitting the pig
(7) Is hitting the rabbit the pencils
(8) The boy are pushing the blocks.

adapted from Bates, Devescovi & D’Amico (1999)

The underlined arguments in the sentences above are favoured as the agent of the action by English-speakers; Italian-speakers prefer the italicised arguments (and choose at chance when both are italicised). Contrasting the intuitions of English- and Italian-speakers demonstrates two points (i) it shows how different languages use different linguistic conventions (discussed below) to resolve the referential ambiguity and indicate which participant is which, and (ii) perhaps more specifically and relevant to the following study, it illustrates the widely replicated finding that word order is the dominant cue to agenthood in English (see review by Bates et al., 1982; Bates, Devescovi & D’Amico, 1999; also Devescovi et al., 2003).

Most theories of language processing would predict that English and Italian speakers should interpret the who-did-what-to-whom of sentence (1) as agent-verb-patient. However, in addition to this the competition model also makes quantitative predictions: for example, on semantically and morphological reversible items, English-speakers choose the first noun on average 10% more than Italians. To explain this, one needs to know something about the relative strength of cues in interpreting Subject Verb Object constructions in English and Italian, something that would be more difficult to explain in theories that use ‘all-or-nothing’ criteria for grammatical items and the relationships they enter into. In sentences (2) and (3) English-speakers behave as if they are using the partial overlap with semi-grammatical and well-formed
phrase structure types that do occur in English. Specifically, in (2) the OSV interpretation could be modelled on relative clause constructions like ‘the boy that the girl kicked’ and on left-dislocated structures that are permitted in informal speech within some dialects ‘now this one, I like’. In (3) the VOS interpretation could be an analogy made with imperative constructions such as ‘hit the ball, John’ and in right-dislocated ‘afterthought’ structures occasionally observed in informal speech ‘plays a mean game of cards, that guy’. One could also argue that SV and VO are separate elements of a surface coalition that overwhelming appears as SVO. This way one strategy for interpretation is always available. Italian speakers perform at chance in their interpretation of (2) and (3), presumably because OSV, SOV, VOS and VSO are all permitted in informal Italian discourse under some pragmatic and/or morphological conditions (Benica, 1993; Simone, 1993). In (4) English word order out-competes animacy as the dominant agent assignment cue. In (5) word-order out-competes agreement on the second noun and in (6) it out-competes both animacy and agreement. English listeners even trust their non-canonical OSV and VOS strategies more than they trust semantic or morphological information, evidenced in the fact that “the pencils” is chosen more often as the actor in sentences like (7), where VOS must compete against animacy and agreement. In contrast, Italians show dramatically different patterns of sentence interpretation with equivalent stimuli. In (8) English-speakers follow the SVO bias but Italian-speakers choose the inanimate noun as actor, showing that for Italian-speakers, overall subject-verb agreement is the most important cue in this context. This study of English and Italian morphosyntactic cues is typical of the way various linguistic parameters have traditionally been manipulated in the cue-competition model to gain a deeper understanding of the relative strengths of cues to sentence interpretation in different languages.
So the cue-competition model appears to be an empirically productive and theoretically plausible framework in which to investigate the relative strengths of linguistic cues, particularly as a tool for quantifying cross-linguistic comparisons. However, the sentences we have considered thus far have been rather simple (e.g., *the horse is kicking the cow*) by morphosyntactic standards. How does the model perform with more complex constructions that are closer to the language a learner actually hears? Bates, Devescovi & D’Amico (1999) tested word-order and agreement in English and Italian adults. They were interested in the strategies used to interpret the matrix clause when a relative clause modifies either the first or second noun. They aimed to challenge the claim that the competition model is only applicable for simple sentences. They replicated earlier findings that English-speakers rely exclusively on word order and Italians rely overwhelmingly on agreement. For English and Italian speakers these conventions only begin to break-down under the worst processing conditions, such as when nouns pile-up before the main verb in “the waitress the cowboy that the baker hears killed”, the syntactic structure of which can be paraphrased as (Noun Noun (Noun Verb) Verb). The general idea is that English-speakers find heavily centre-embedded sentences more difficult to interpret than Italian as a by-product of English’s heavy reliance on word order. Italian speakers are more likely to recover the intended referents of the actions in these sentences as they are potentially able to use the gender cue on the past participle, a product of agreement in general being much more important in Italian than English for assigning agentivity.

Many of the studies above have noted the influence of animacy as a nonlinguistic cue to thematic role assignment (see also Comrie, 2005; Dryer, 2005; Siewierska, 2005) and it is worth noting that it seems to occupy a slightly different
status as a cue than case and word-order because by definition it is an inherent semantic property of an entity rather than a grammatical convention. Prototypically, agents tend to be animate and patients tend to be inanimate, so where there is an animacy contrast between two entities in a transitive sentence it is prototypical that the animate one is the agent and the inanimate one is the patient, and this should be highly reliable across languages – though certainly not always as in the sentence the ball hit John. In the scenario where both nouns in a sentence are animate, the animacy cue is not contrastively available (either noun could act as agent).

2. The cue-competition model and development

While it is clearly important for the cue-competition model that it is able to describe adult grammatical reasoning as a result of competing and converging cues, we are primarily concerned here with understanding the implications of this model for development, specifically, how the infant might be using multiple cues in constructing their language.

Slobin & Bever, (1982) tested English, Italian, Serbo-Croatian & Turkish 2-4 year olds, varying cues of word-order, stress, and agreement. Each child tested on their ability to understand reversible transitive sentences containing familiar words, ‘e.g., the cat scratched the dog’ and their task was to demonstrate the action using toy animals. The most general result of this investigation is that children seemed equally prepared to learn both inflectional and word order languages. The results indicated that children are attuned to canonical sentence form, and that, early on, they develop schemas embodying the most typical features of such forms in their language. Slobin & Bever point out that the acquisition of Turkish is not at all impaired (relative to other languages in the study) by the fact that word order is not a cue for semantic
relations in that language. Serbo-Croatian children, however, show a strong
dependence on word-order comprehension strategies, even in inflected sentence forms
where such strategies would be unnecessary. Apparently the fact that the word-order
strategy is required for the comprehension of some forms that are not recognisably
inflected stimulates the child to overgeneralise the use of the strategy to all forms.
While it might seem uncontroversial (perhaps even trivial) for those working in the
usage-based approach to point out that children seem equally prepared to learn both
inflectional and word-order languages, this should be seen in the context of arguing
against one influential nativist theory that predicts the primacy of word order in
acquisition (Pinker, 1981, 1982). In another study that focused on infants’ ability to
comprehend sentence meaning, Bates, MacWhinney, Caselli, Devescovi, Natale &
Venza, (1984) looked at the cues of word-order, animacy and stress in 2-5 year-old
English- and Italian-speaking children. At every age tested English-speaking children
relied on word-order and Italian-speaking children relied on animacy.

In summary, advocates of the cue competition model argue that sensitivity to
the ‘information value’ of cues in a particular language (as estimated by cue
reliability, validity and cost) is the single most important factor that accounts for the
sentence comprehension strategies and syntactic competence of speakers of that
language. This is wholly consistent with usage-based approaches which take the role
of the input seriously in the process of language acquisition; not ‘just’ as a trigger for
underlying innate representations that are themselves meaningless (in the technical
sense that that they are so abstract they perform no communicative function) but as
the raw material from which an infant constructs their language.

These studies also add to a massive amount of evidence from
psycholinguistics that shows there is simply no principled way to draw a line around
the ‘core’ grammatical competence (as traditional generative approaches to syntax have tried to do) and what was once thought to be relatively unimportant ‘periphery’ performance phenomena such as features of memory and attention, the frequency of linguistic items and constructions in the input, or the reliability with which a particular form is paired with a particular function and vice versa. The empirical evidence supports a fuzzy boundary between pragmatics, semantics, syntax and morphology and this is more naturally accommodated in a theory of language which sees different psychological levels of representations as a continuum (instantiated in the construction/cognitive approach to grammar as symbolic assemblies).

3. The English transitive construction

As has already been discussed, cue-competition style experiments have repeatedly found that word-order is the most dominant cue in English. However, one problem with these studies is that they used familiar verbs, therefore they cannot rule out the possibility that children responded on the basis of only verb specific knowledge, for example knowing that in ‘canonical’ SVO sentences the ‘tickler’ comes before ‘tickles’ and the ‘tickle-ee’ comes after the word ‘tickle’; as we know, in contrast, adult grammatical knowledge productively generalises across item-specific knowledge. To control for this Gertner, Fisher, & Eisengart (2006) used novel verbs and tested English-speaking children’s comprehension of word-order when it was the only cue available (so did not conflict with any other cues). They found that even young two-year olds already distinguish between such things as ‘x is tickling y’ and ‘y is tickling x’. Hirsh-Pasek & Golinkoff (1996) found a similar effect using familiar verbs. At two-years-of-age, children appear to have an intuition that the typical
transitive construction at a very basic level represents asymmetrical activities rather than two participants engaging in the same act simultaneously (Naigles, 1990).

In the empirical work that follows shortly we wanted to go beyond these findings and look at two properties of English transitives that have not been considered in this detail together before – active/passive voice and case-marking.

1. Like in many languages English-speakers are able to passivise utterances. In the English transitive this means speakers can alter the focus of the utterance from the agent acting on the patient in the active voice (e.g., John hit Mary) to the action as experienced by the patient in the passive voice (e.g., Mary was hit by John). Note that the scene, in terms of who-did-what-to-whom, has not changed, thus at some conceptual level where the scene is construed, the different surface forms probably share an underlying event schema (though with a different landmark and trajector, Langacker, 1987). For this reason the following study looked at active and passive sentences so that we could pull apart the weights of the different cues, particularly the agent-verbing-patient mapping that exists in English (Figure 1).

Figure 1 shows part of a simplified transitive construction with active and passive voice alternations. Red highlighting indicates that the information is in discourse focus. The dotted line hypothesises that the different surface forms and semantic constituents are unified by an abstract transitive schema. For a usage-based approach it is plausible that early on in development, active and passive voice constructions could be relatively independent form-function pairings, only unified at an abstract level at a later phase. What this ‘later phase’ depends on will be the type of things discussed in Chapter 2. To recap, some of the important factors in the schematisation process are: i) the more abstract the pattern or correlation that one
needs to detect, the more evidence one needs, ii) the facilitative role of variation in the learning set (a skewed type/token distribution multiplied by overall frequency of learning opportunities) and iii) recognising that the alignment of relational structure and mapping between representations is a fundamental psychological process underlying analogy and similarity. Thus categorisation decisions are based on the relationship between participants in an event – schematisation based on ‘things are what they do’.

Figure 1. A simplified transitive schema that unifies active and passive forms.

It is interesting to briefly look at passivisation behaviour in English transitive utterances. Consider the following sentences, which are each tested for containing noun-phrases with participant status. For example, bona fide participants can be topicalised (a, b), one can enquire as to the role they played in the process (c) and they can be modified (d).

(1) a. The farmer shot the rabbit.
b. The rabbit was shot by the farmer.
c. ...as for the rabbit, the farmer shot it.
d. What happened to the rabbit? He was shot.

(2)  a. The heat suffocated him.
b. He was suffocated by the heat.
c. ...as for him, the heat suffocated him.
d. What happened to him? He suffocated.

(3)  a. The tent sleeps six people.
b. Six people are slept by the tent.
c. ...as for six people, the tent sleeps them.
d. What happened to the six people? The tent slept them.

(4)  a. My guitar broke a string
b. A string was broken by my guitar.
c. ...as for my string, the guitar broke it.
d. What happened to the string? My guitar broke it.

(5)  a. My guitar broke a window
b. a window was broken by my guitar
c. ...as for the window, my guitar broke it.
d. What happened to the window? My guitar broke it.

(6)  a. My car burst a tyre
b. A tyre was burst by my car.

c. ....as for the tyre, my car burst it.

d. What happened to the tyre? My car burst it.

(7)  
a. My car burst a balloon.

b. A balloon was burst by my car.

c. ....as for the balloon, my car burst it.

d. What happened to the balloon? My car burst it.

All these sentences are on a continuum of acceptability-unacceptability to English-speakers and it is beyond the scope of what we are investigating here to explain in any detail all the factors that might be at work in passivisation phenomena (other authors have looked at this phenomena in book-length depth; see Langacker, 1982). One purely needs to note that the alternations show subtle patterns which are sensitive to (among other things) intentionality, focus and animacy. The following study tests only a subset of the competence an adult-speaker has regarding active-passive alternations, all of which the child has to learn.

2. In English, nominative-accusative case marking on pronouns carries additional information for semantic role assignment. Yoshimura & MacWhinney (2010) have recently shown even when the case-marking cue is available, word order remains the strongest cue in English. However, for noncanonical word orders, the case-marking cue had a strong effect on sentence interpretation. Word order only dominates case in canonical NVN order (*the boy kissed the girl*). In that order, there are two word order cues present: preverbal positioning and postverbal positioning. In the noncanonical
VNN (pushes the boy the marble) and NNV (the boy the girl kissed) orders, only one of these word order cues is present at a time. In those orders, case is able to compete with the single word order cue and often dominates it.

There is also evidence that 2 ½-year-olds’ comprehension of transitives is boosted when pronominal case-marked nouns are used as opposed to full noun phrases (Childers and Tomasello, 2003). By using the highly reliable role information available from pronouns, children can move more quickly to learning about the thematic roles and positions of NPs that cooccur with new verbs. However, case marking works in opposing ways to indicate agent-patient relations in active and passive transitives. For this reason the following study looked at case marking in active and passive sentences so that we could investigate the relative weights of case marking and word order in determining the meaning of sentences. The role of case (as instantiated in ‘pronoun frames’) in the acquisition of English is discussed in detail in the following paper.

One final point that should be noted is that in order to expose the children to all permutations of these variables this meant that some of the utterances they heard were ungrammatical. Putting cues into conflict in this way allows us to disentangle the relative weights of cues.

The following study set out to explore the development of transitive comprehension by assessing what cues young children are sensitive to in the input and how children use these cues to comprehend agent-patient relationships. The paper has been accepted for publication in Language, Learning and Development (2010) and is authored by Paul Ibbotson, Anna Theakston, Elena Lieven, Mike Tomasello.
Publication 4

The Role of Pronoun Frames in Early Comprehension of Transitive Constructions in English.

Page: 198
The Role of Pronoun Frames in Early Comprehension of Transitive Constructions in English

Abstract

Case marking in English is available only on some pronouns and only in some cases. It is unknown whether young children acquiring English nevertheless make use of this highly restricted marking as a cue to sentence interpretation. The current study therefore examined how 2 and 3 year-old English children use case-marked pronoun frames and constructional cues (actives versus passives) to understand agent-patient relations in transitive sentences containing novel verbs. In a pointing comprehension test, 2 year-olds used pronoun frames containing 2 case marked pronouns to help them interpret grammatical sentences, both actives and passives, but they were unable to assign agent-patient relationships in any consistent way with ungrammatical pronoun frames. Three year-olds also used pronoun frames to interpret grammatical active and passive sentences (with either 1 or 2 case-marked pronouns), but varied in their interpretation of ungrammatical sentences according to pronoun frame. These results suggest that the role of case-marked pronouns has been underestimated in English language acquisition, and that even very young English children use multiple cues to comprehend transitive sentences.
1. Introduction.

The transitive construction, in one form or another, is present in almost all the world's languages (Hopper & Thompson, 1980) as it codifies a fundamental pattern of human experience in which, prototypically, an agent intentionally instigates an action that affects a patient (Næss, 2007). The different roles of the two participants are typically marked with either word order, case-marking, or some combination of these. Ontogenetically the transitive construction (usually in its active form) is important as it is the most basic construction in which comprehension rests crucially on being able to successfully identify via this marking which participants are playing which roles in the event (who is doing what to whom).

English is typically considered a word order language because case marking operates over a small set of lexical items, marking only pronouns (not full nouns) for their subject and object position with nominative and accusative case respectively. If case is so restricted perhaps English-speaking children ignore it and rely more strongly on other cues such as word order. However, analysis of English child directed speech shows that 83% of active transitive sentences have a pronoun in the subject position and 36% have a pronoun in the object position (Cameron-Faulkner,Lieven, & Tomasello, 2003; see also Valian, 1991). Given this, we may have underestimated the extent to which English-speaking children are learning to integrate case-marking information with other cues to assigning agent-patient relations. The suggestion is that pronouns, such as I, me, he and it, may have a special role to play because they are highly frequent and, in many instances, have different forms for subjects (he/she preverbally) and objects (him/her postverbally). Even where a pronoun is not case-marked it can occur in one role much more often than the other (e.g. it as the object of a transitive and as the subject of an intransitive).
In support of this view, Jones, Gobet and Pine (2000) used a computational model to perform item-based distributional analyses on a corpus of child directed speech. Of special significance is that the model formed a number of “pronoun islands” because some high-frequency pronouns, such as *I*, *he* and *it*, occur with regularity in certain utterance positions. Other work also suggests the importance of pronouns in anchoring children’s early transitive constructions (Childers & Tomasello, 2001 for actives; Dodson & Tomasello, 1998; Lieven, Pine, & Baldwin, 1997; Savage, Lieven, Theakston & Tomasello, 2003 for actives & passives) but, to our knowledge, there has been no systematic research into the ‘value added’ by different case-marked pronoun frames in signifying agent-patient roles in English.

In contrast, the role of word-order in determining agent-patient relationships has been much more widely researched. Several developmental studies that looked at children between the ages of 2-5 have established word-order as the dominant sentence interpretation strategy for English-speaking children such that they interpret the preverbal noun in active transitives as agent even when the cues of agreement, stress and animacy point to the post-verbal noun as the agent (e.g., Bates, MacWhinney, Caselli, Devescovi, Natale & Venza, 1984; MacWhinney, Bates, & Kliegl, 1984; Slobin & Bever, 1982). In the cue-competition model, Bates & MacWhinney (1987) attempted to quantify the interaction of different syntactic cues for constructions in different languages. Thus the differential ‘weights’ of various combinations of competing and converging cues including word-order, stress, animacy, agreement, topicalisation, prepositions, and case are measured. The strength of a particular cue is a product of how frequently it is present when it is needed (cue availability) and how consistently it is mapped onto a particular form whenever it is present (cue reliability).
This idea can be illustrated by contrasting the active and passive forms of the transitive construction in English. Word order is almost always available as a sentence interpretation strategy in English but it is not reliable across active and passive constructions because the form-function mapping changes between passive (patient-verbed-by-agent) and active (agent-verbing-patient). Cue strength is a measure of how informative a cue is in a language. As well as being a product of validity and reliability it is affected by the processing limitations imposed by the perceptual and working memory systems (cue cost). At its simplest, the classical competition model conceptualises language development as a process whereby cue strength (mediated by cue cost) comes to vary as a function of that cue’s availability and reliability in the input.

One problem with these early studies on cues to the active transitive construction is that familiar verbs were used and therefore the possibility that children responded on the basis of verb specific knowledge (for instance, knowing that the ‘tickler’ comes before ‘tickles’ and the ‘tickle-ee’ comes after) cannot be ruled out. In contrast, adult grammatical knowledge productively generalises across item-specific knowledge. Dittmar, Lieven, Abbot-Smith and Tomasello (2008) avoided this limitation by using novel verbs to test German 2-, 5-, and 7-year olds' comprehension of case and word order in active transitive sentences. The results showed a very clear pattern. The two year-olds comprehended active transitives only with redundant marking of agent and patient. Redundancy in this context means that multiple cues indicate the same agent-patient relationship (i.e. nominative-accusative case-marking in SVO word order). The two year-olds in the Dittmar et al study were the most sensitive group to distortions from this, failing to comprehend transitive sentences for
which diagnostic case marking was absent\(^8\) or those in which the word order was non-canonical but grammatical (object first). Dittmar et al invoke the concept of a ‘prototype’ to explain these results. In the cue-competition model a prototype represents the centre of a common functional group of cues, that is, prototypical transitive sentences with both word order and case marking (and perhaps other cues) working in coalition: the coalitions-as-prototypes model (Bates & MacWhinney, 1987). The five year-olds comprehended the transitive sentences mostly in terms of word-order. Their performance in the word order condition was as high as with the full prototype and they chose at random in response to the object-first sentences in which word-order and case conflicted, with a slight tendency to go, incorrectly, with word order. Only the seven year-olds performed like adults and went with case marking in all conditions, including in object-first sentences in which case marking and canonical word order conflicted (Dittmar et al., 2008). Interestingly, the prototypical form in German is also the most frequent (calculated from a corpus of child-directed speech). This does not have to be the case. For instance in the model of Franks and Bransford (1971), the prototype is an instance that has all the cues but it does not have to be the most frequent exemplar.

In an extension of this work, Chan, Lieven and Tomasello (2009) compared young Cantonese-, German-, and English-speaking children’s understanding of word order and animacy contrasts as cues to agent-patient relations in the active transitive construction. Children aged 2;6, 3;6 and 4;6 acted out active transitives containing novel verbs in three conditions: (1) SVO word order with animate agents and inanimate patients (‘redundant’ cues); (2) SVO word order with both agent and patient inanimate (word order cue only) and (3) SVO word order with an inanimate\(^8\) For example, Das Schwein tammt das Zebra (The neuter pig is tamming the neuter zebra) which is grammatical but lacks any cues to agent patient assignment based on case marking.
agent and an animate patient (conflicting cues). Of particular interest was whether children would find especially easy sentences in which both cues were used together redundantly and especially difficult sentences in which the two cues conflicted (Bates & MacWhinney 1987). The results showed that when word-order was the only cue, English children showed the earliest comprehension at 2;6, then German children, and then Cantonese children at 3;6. When the cues conflicted, none of the 2;6 children in any language comprehended in adult-like ways, whereas children at 3;6 and 4;6 were significantly more likely to follow word order rather than animacy. These results indicate that across languages, children aged 2;6 comprehend active transitive sentences when they have support from the coalition of word-order and animacy, even with novel verbs. However it is not only the convergence of the two cues but also their frequency and salience which may be helpful. Sentences of the form animate-verb-inanimate are frequent in the input, and they encode the prototypical causative scene which should be highly familiar to young children regardless of their target language (see the Manipulative Activity Scene in Slobin 1985, 1997).

In the current study we were interested in the value added by case in marking the semantic roles of transitives over and above that of word order and other constructional cues in active and passive argument structures. To investigate this we systematically varied the constructional cues (NP is verbing NP vs. NP is getting verbed by NP) and case-marking cues on the pronouns (he/she vs. her/him vs. it). These can be configured as different ‘pronoun-frames’, for example: [NP\textsubscript{nom} – Verb – NP\textsubscript{acc}] vs. [NP\textsubscript{neu} – Verb – NP\textsubscript{acc}]. In the present study, to test for the ‘value added’ by different case-marked pronoun frames, the child viewed two pictures simultaneously, and the child’s task was to correctly identify who is doing what to whom by pointing to the cartoon that corresponded to the test sentence. In
order to expose the children to all permutations of these variables some of the
utterances they heard were ungrammatical, as is typical of cue-competition-style
experiments because putting cues into conflict allows us to disentangle their relative
weights (e.g., MacWhinney & Pléh, 1988; Smith & Bates, 1987). Sceptics might
wonder about the use of ungrammatical sentences, arguing that children employ extra-
linguistic strategies to decode such stimuli. But in fact there is cross-linguistic
evidence that the processing of both grammatical and ungrammatical sentences in
competition-style experiments proceeds by reference to the same sets of cues and
processing patterns (MacWhinney, Pléh, & Bates, 1985; Smith and Mimica, 1984).
We were interested in seeing how two and three year-olds’ varied in their use of
pronoun frames versus constructional cues (active and passive constructions) to
successfully comprehend agent-patient relationships in English.

2. Method

2.1 Participants

Forty-two monolingual English-speaking children took part in the study, twenty-one
from each of the two age bands: 2;8-3;0 (M = 2;10) which included 8 boys and 13
girls and 3;4-3;8 (M = 3;7) which included 10 boys and 11 girls. A further eighteen
children were tested but were excluded from the analysis due to bilingualism (1); the
child only pointing to the left on every trial (3); pointing only to the right on every
trial (5); pointing to both sides on every trial (3); or refusing to point at all (6). The
excluded children were evenly distributed across the two age bands. Children were
recruited from a database of parents who volunteered to participate in language
acquisition studies and testing took place in the Max Planck Child Study Centre,
Manchester, UK.
2.2 Materials

Two novel verbs *tamming/tammed* and *keefing/keefed* were used to describe two novel transitive actions that were performed by animated cartoon characters. Two novel verbs were chosen to vary the actions that children watched and so keep their interest over the duration of the experiment. Both verbs referred to prototypical causative transitive actions that were reversible, involving direct contact between a volitional agent and an affected patient (Hopper & Thompson, 1980).

*Tamming/tammed* referred to an action where an agent flew through the air towards the patient and then moved up and down on top of a patient *causing* the patient to rotate one-and-a-half turns. This action resulted in the affected patient coming to rest upside-down while the agent returned to its original starting position. *Keefing/Keefed* referred to an action where an agent moved horizontally until it made contact with the patient, *causing* it to move in the same direction and rotate ninety degrees onto its side as if it had been pushed. Agents and patients could be either human or non-human depending on the condition. Human agent/patients were represented by two male and two female cartoon characters. Non-human agent/patients were represented by a cartoon pencil, cup, and a chair. The present progressive form of the verb *(tamming/keefing)* appeared only in the active sentences while the past participle *(tammed/keefed)* appeared only in the passives (see Table 1).

Although the discourse function of *get* and *be* passives is equivalent – they both serve to background the agent – there is some suggestion that *get*-passives (the ones used in this experiment) tend to dominate in early comprehension and production (Turner & Rommetviet 1967a, 1967b; Menyuk 1969; Harris & Flora 1982; Crain & Fodor 1993; Slobin 1994). There is even some evidence that young children were more likely to use *get*-passives to describe prototypically transitive scenes which
involved a clear change of state (of the kind modelled in this experiment) whereas be-passives tended to be used for non-prototypical transitive scenes\(^9\) (Marchman, Bates, Burkhardt & Good, 1991). Given that the youngest age group was 2;10 and the task was relatively demanding we chose get-passives to boost performance on passives across the board.\(^2\) This should not significantly influence the pattern of results as in this study the relative performance of active versus passive (regardless of pronoun-frame or grammaticality) is not as important as how pronoun frames affect understanding within active and passive performance and with various degrees of grammaticality.

### 2.3 Design

This was a mixed design with three within-subject factors: Transitive construction type (active, passive); Grammaticality (grammatical, ungrammatical); Pronoun Frame (Nominative/Accusative, Nominative/Neuter, Neuter/Accusative); and one between-subject factor: Age (2;8-3;0 & 3;4-3;8). Given the evidence that all the independent variables may interact – affecting the outcome measure in complex and unpredictable ways – a factorial approach was chosen. When the within-subject variables are crossed there are twelve conditions, listed in Table 1.

---

\(^9\) There is conflicting evidence as to the domination of get- vs. be-passives in British English children’s speech (Meints 2003; Messenger 2009). It was therefore deemed more reliable to go with just one type.
<table>
<thead>
<tr>
<th>Transitive Construction Type</th>
<th>Grammaticality</th>
<th>Pronoun Frame (agent/patient)</th>
<th>Test Sentence Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active</td>
<td>Grammatical</td>
<td>Nominative/Accusative</td>
<td>She is tamming him</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nominative/Neuter</td>
<td>She is tamming it</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Neuter/Accusative</td>
<td>It is tamming her</td>
</tr>
<tr>
<td>Ungrammatical</td>
<td></td>
<td>Accusative/Nominative</td>
<td>Her is tamming he</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Accusative/Neuter</td>
<td>Her is tamming it</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Neuter/Nominative</td>
<td>It is tamming she</td>
</tr>
<tr>
<td>Passive</td>
<td>Grammatical</td>
<td>Accusative/Nominative</td>
<td>She is getting tammed by him</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Accusative/Neuter</td>
<td>It is getting tammed by her</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Neuter/Nominative</td>
<td>She is getting tammed by it</td>
</tr>
<tr>
<td>Ungrammatical</td>
<td></td>
<td>Nominative/Accusative</td>
<td>Him is getting tammed by she</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nominative/Neuter</td>
<td>It is getting tammed by she</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Neuter/Accusative</td>
<td>Her is getting tammed by it</td>
</tr>
</tbody>
</table>

*Table 1*. Within-subject variables.

There was a total of twenty-four test sentences such that each child received four examples of each of the 4 different 1 case-marked pronoun conditions and two examples each of the 4 different 2 case-marked conditions \((4\times4) + (4\times2) = 24\). All children heard the same test sentences which were counterbalanced in the following ways: the pairing of human/non-human agent/patients with all four forms of the two novel verbs; the left/right appearance on the computer monitor of the grammatical sentence; the left/right appearance in the ungrammatical sentences against first noun/second noun (a plausible strategy children might resort to in such cases); and finally, the pronoun gender with the verb forms.

Each sentence type had one example of each of the two novel verbs. For example, *X is tamming Y* and *X is keefing Y*. To balance the gender across conditions, the 1 case-marked pronoun condition had two sentences of the same female gender (e.g. *She is X-ing it*), and two of male gender (e.g. *He is X-ing it*). For the 2 case-
marked pronoun conditions one sentence was female-first male-second (e.g. *She is X-ing him*) and one was male first and female second (e.g. *He is X-ing her*). No sentences ever had matching genders as it would be impossible to know which was the correct interpretation given the experimental design with two sets of reversible transitive scenes. The order that sentences appeared in the test was then randomised.

### 2.4 Procedure

During the session, the child sat in front of a computer monitor with the experimenter sitting next to them.

*Warm-up and pre-test:* To begin with, the child saw cartoon pictures of boy/girl couples. The experimenter then said “here’s a picture, can you point to where she/he is” and on the second set “here’s another picture, can you point to her/him” The purpose of this was to screen out those individuals who were not competent in using *he/she* and *him/her* pronouns to refer to stereotypical boy/girl characters that were similar to the ones used in the test-phase. Each child was given 2 opportunities to correctly identify the referent of each pronoun and those that scored less than 75% correct on this task were excluded from the analysis (in fact, all of these children would have been excluded anyway on other criteria such as only pointing to one side). After this, the child watched a series of animated scenes that introduced the actions and novel verbs. The experimenter guided the attention of the child by saying “Look, tamming/keefing!” and “Look what’s happening now, tammed/keefed!” thus the two verbs occurred in both forms with both human and non-human agents and patients to avoid associating any particular thematic role with subject or object position. The warm-up and pre-test contained no human or non-human characters that would appear in the test phase. Preceding each of the test slides, there were two slides that
introduced the participants involved in the actions one couple at a time but with no accompanying audio (the left/right appearance of this was also counterbalanced). This reduced the processing burden of integrating new visual and audio information on the pointing trial. The sequence duration of pre-test slide presentation was automatically controlled by PowerPoint at three seconds per character couple. The length of time that the test slide remained on screen was under the control of the experimenter to allow for individual response times in pointing.

Test phase: In the test phase the child viewed two pictures simultaneously, and the child’s task was to correctly identify who is doing what to whom by pointing to the cartoon that corresponded to the test sentence. In this phase the sentences were pre-recorded to ensure phonological consistency across participants. The audio began when the action started and was timed to finish when the action stopped. The duration of the actions was four seconds for both sides. Figure 1 shows a schematic representation of a scene the child watched while hearing, for example, “look what's happening now...can you point to, she is tamming him?” The left-hand picture shows the direct of action from agent to patient and the right-hand picture shows the outcome.

Figure 1. Example of test stimulus.
The experimenter verbally praised the child for engaging with the task and pointing, regardless of the correctness of the response.

2.5 Coding

We were interested in the extent to which the children’s answers showed any systematicity and the whether this would be in the same direction as those of an adult under the same circumstances. In a pilot, six adults were tested on the same stimuli and they performed at ceiling, rating all active sentences on the basis of preverb=agent/postverb=patient and all passive sentences as preverb=patient/postverb=agent, regardless of grammaticality or pronoun frame. Therefore, the ‘right’ answer to ungrammatical stimuli was taken to be that systematically produced by adult speakers of the language.

For every trial, the side (left/right) that the child pointed to was recorded by the experimenter. In a sentence such as It is tamming her, if the child pointed to the cartoon couple that were acting out [It tamming her] then he scored 1 and if he pointed to [she tamming it] he scored 0. A score of ‘1’ represents an ‘active’ interpretation of the sentence so that the thematic role of agent is mapped onto the subject position and patient onto the object. A score of ‘0’ represents a ‘passive’ interpretation where this mapping has been reversed, so that the agent now corresponds with object position and patient with subject position. This means that in an ungrammatical sentence such as Him is tamming she, if the child pointed to [he tamming her] they scored 1 and if they pointed to [she tamming him] they scored 0. For passives such as It is getting keefed by her, if the child pointed to the cartoon couple that were acting out [she keefing it] they scored 0 and if they pointed to [it
keefing her] they scored 1. In an ungrammatical sentence such as *him is getting keefed by she*, if the child pointed to [she keefing him] they scored 0 and if they pointed to [he keefing her] they scored 1. In other words, assigning the agent-patient relationship was recorded with reference to the active/passive frame with 0.5-1 scores indicating an active interpretation, and 0.5-0 scores representing a passive interpretation.

### 3.0 Results

#### 3.1 Actives versus Passives

Figure 2 shows the overall developmental picture for active and for passive sentence comprehension, collapsing across grammatical and ungrammatical sentences, and across pronoun frames. The dependent measure is a mean pointing score calculated so that 0 represents passive interpretation by all children and 1 represents complete active interpretation by all children (for all sentences in a particular condition). In other words, this score reflects the proportion of times children interpreted any given sentence as agent-first, patient-first or no clear preference. So, if equal numbers of children interpreted the sentences of one condition as active and as passive the score would be 0.5, and this, therefore, represents the chance baseline. As expected, overall children are at a developmental advantage when interpreting the actives over the passives and this difference is most pronounced in the younger children. Adults interpreted all sentences, regardless of grammaticality or case-marking, with respect to the construction cue (active/passive).
3.2 Pronoun Frames: Comparison to Chance

The analysis of pronoun frame comprehension involved collapsing across conditions with one case-marked pronoun (e.g. *it* taming *her* and *she* taming *it*). This was done because from the perspective of case-marking, these pronoun frames both carry the same information value when indicating the grammatical role of the pronoun with respect to the verb, that is, only one of the pronouns is case-marked in either sentence. This means sentences with ‘*it*’ appearing in either subject or object position are treated as one pronoun frame (in comparison to the two case-marked pronoun frame), reducing the conditions from 12 (as described in Table 1) to 8\(^\text{10}\). To determine

\^[10] We have examined the sentences separately from the one pronoun frame condition, however we found no clear or consistent advantage for sentences with ‘*she/he/her/him*’ as agent and ‘*it*’ as the patient over sentences with ‘*it*’ as the agent and ‘*she/he/her/him*’ as the patient. We analyse the data in the way that we do for two reasons. Firstly, the motivation for grouping the conditions in the way that we did was on the basis of the relative information content carried by the different pronoun frames (1 case-marked versus 2 case-marked) – the main theoretical motivation of the experiment. Secondly, we argue in the Discussion that animacy effects are neutralised in this experiment, which is perhaps why we don’t see any additional performance boost when we do break-apart the 1 case-marked condition and analyse the results separately.
whether children were using the pronoun frames in their sentence interpretation, one- sample t-tests (against chance) were conducted on the 8 different sentence types at both ages.

The results can be seen in Table 2 where the values in the boxes refer to the mean proportion of pointing to the first noun as agent across children (in bold), standard deviations (in curly brackets), t-values and p-levels. The shaded boxes indicate conditions in which the pointing was significantly biased away from chance in one direction or the other (chance = .50).

The results show that, on grammatical sentences (top two rows), both groups of children were significantly different from chance on the active sentences, regardless of whether they contain one or two case-marked pronouns. On grammatical passives, both age groups are significantly different from chance for correct comprehension of sentences with two grammatical case-marked pronouns, but only the older children are significantly different from chance for the sentences with one case-marked pronoun (it in the other argument role). This means that for the younger children, more case marking (i.e., on two versus one argument) was a help for passive interpretation.

For the ungrammatical sentences, the younger children were not significantly different from chance for either pronoun frame in either the active or passive. The older children, in contrast, were different from chance for both actives and passives with two incorrectly marked pronouns but not with one case-marked pronoun (again with it in the other argument role). This means that the younger children are confused by the conflicting (ungrammatical) marking, and so perform randomly, whereas the older children are able to assign agent/patient relations with regard to the construction
(active/passive) using word-order cues despite having two case-marked pronouns that are in ungrammatical positions.

*Table 2.* Mean scores (in bold), standard deviations (in curly brackets) and t-test statistics for pronoun frame interpretation by age.

<table>
<thead>
<tr>
<th>Pronoun frame</th>
<th>Example sentences</th>
<th>2;8-3;0</th>
<th></th>
<th>3;4-3;8</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Acts</td>
<td>Passives</td>
<td>Acts</td>
<td>Passives</td>
</tr>
<tr>
<td><strong>Grammatical</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nom. V-ing Acc.</td>
<td><em>She is taming him</em></td>
<td><strong>0.895</strong></td>
<td>(0.294)</td>
<td><strong>0.333</strong></td>
<td>(0.329)</td>
</tr>
<tr>
<td></td>
<td><em>She is getting tammed by him</em></td>
<td><em>t</em>[20]=4.81</td>
<td>p&lt;0.001</td>
<td><em>t</em>[20]=2.32</td>
<td>p=0.031</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>0.857</strong></td>
<td>(0.280)</td>
<td><strong>0.261</strong></td>
<td>(0.339)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>t</em>[20]=5.83</td>
<td>p&lt;0.001</td>
<td><em>t</em>[20]=3.21</td>
<td>p=0.004</td>
</tr>
<tr>
<td>Θ V-ing Acc.</td>
<td><em>It is taming her</em></td>
<td><strong>0.631</strong></td>
<td>(0.257)</td>
<td><strong>0.488</strong></td>
<td>(0.320)</td>
</tr>
<tr>
<td>or</td>
<td><em>He is tamming it</em></td>
<td><em>t</em>[20]=2.32</td>
<td>p=0.030</td>
<td><em>t</em>[20]=-0.23</td>
<td>p=0.815</td>
</tr>
<tr>
<td>Nom. V-ing Θ</td>
<td><em>It is getting tammed by her</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>She is getting tammed by it</em></td>
<td><strong>0.726</strong></td>
<td>(0.235)</td>
<td><strong>0.262</strong></td>
<td>(0.185)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>t</em>[20]=4.39</td>
<td>p&lt;0.001</td>
<td><em>t</em>[20]=5.89</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td><strong>Ungrammatical</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acc. V-ing Nom.</td>
<td><em>Him is tamming she</em></td>
<td><strong>0.595</strong></td>
<td>(0.339)</td>
<td><strong>0.428</strong></td>
<td>(0.396)</td>
</tr>
<tr>
<td></td>
<td><em>Him is getting tammed by she</em></td>
<td><em>t</em>[20]=1.28</td>
<td>p=0.214</td>
<td><em>t</em>[20]=-0.83</td>
<td>p=0.419</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>0.833</strong></td>
<td>(0.288)</td>
<td><strong>0.190</strong></td>
<td>(0.334)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>t</em>[20]=5.29</td>
<td>p&lt;0.001</td>
<td><em>t</em>[20]=-4.24</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>Θ V-ing Nom.</td>
<td><em>It is taming she</em></td>
<td><strong>0.595</strong></td>
<td>(0.243)</td>
<td><strong>0.428</strong></td>
<td>(0.195)</td>
</tr>
<tr>
<td>or</td>
<td><em>Her is tamming it</em></td>
<td><em>t</em>[20]=1.79</td>
<td>p=0.088</td>
<td><em>t</em>[20]=-1.67</td>
<td>p=0.110</td>
</tr>
<tr>
<td>Acc. V-ing Θ</td>
<td><em>It is getting tammed by she</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Her is getting tammed by it</em></td>
<td><strong>0.571</strong></td>
<td>(0.251)</td>
<td><strong>0.416</strong></td>
<td>(0.266)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>t</em>[20]=1.30</td>
<td>p=0.208</td>
<td><em>t</em>[20]=-1.43</td>
<td>p=0.167</td>
</tr>
</tbody>
</table>

Figures 3 and 4 illustrate the mean interpretation scores for each pronoun frame condition (with 95% confidence intervals) for grammatical and ungrammatical sentences. Here, it is perhaps clearer to see the value added by different pronoun...
frame configurations and which conditions pull away from chance at what age. Asterisked conditions are significantly ($p<0.05$) different from chance.

Figure 3. *Mean Scores of Interpretation for Grammatical Pronoun Frames.*

Figure 4. *Mean Scores of Interpretation for Ungrammatical Pronoun Frames.*
Comparisons to chance indicated in which conditions participants showed a systematic preference for a certain interpretation (active or passive). In order to establish which conditions were significantly different from each other (i.e. relative performance) we conducted comparisons across conditions.

3.3 Pronoun Frames: Comparison across Conditions

To look more closely at how the children’s performance varied across conditions, we conducted two separate 2 x 2 (pronoun frame x transitive construction type) mixed ANOVAs for each age group - one with grammatical sentences and one with ungrammatical sentences. Bonferroni adjustments were applied for pairwise comparisons.

3.3.1 Grammatical Sentences

For the younger age group, there was a main effect of transitive construction type such that children comprehended actives better than passives, $F(1,20) = 14.69, p = .001$, but this interacted with pronoun frame, $F(1,20) = 12.7, p = .002$. Bonferroni-adjusted t-tests revealed that children comprehended active sentences with two case-marked pronouns better than active sentences with one case-marked pronoun, $t(20) = 2.97, p = .008, (M=80.95 > 63.09)$ and they also comprehended passive sentences with two case-marked pronouns better than passive sentences with one case-marked pronoun, $t(20) = -2.35, p = .029, (M=33.3 < 48.8)$. In other words, the younger children benefited more from the pronoun frame with two case-marked pronouns than the pronoun frame with one case-marked pronoun in both the active and passive sentences. The same ANOVA (transitive construction type x pronoun frame) was conducted for the older age group. There was a main effect of transitive construction
type, $F(1, 20) = 82.16, p < .001$, such that children comprehended actives better than passives. There was no effects of pronoun frame and no interaction. This analysis shows that unlike the younger children, the older children were not affected by pronoun frame in their interpretation of grammatical active or passive sentences.

### 3.3.2 Ungrammatical Sentences

In a 2 x 2 (transitive construction type x pronoun frame) ANOVA on the ungrammatical sentences for the younger age group, there was a main effect of transitive construction type such that children performed better on actives than passives (although they comprehended neither at a rate different from chance), $F(1, 20) = 5.38, p = .03$, but no main effect of pronoun frame and no interactions. The same ANOVA (transitive construction type x pronoun frame) was conducted for the older age group. Again they comprehended actives better than passives, $F(1, 20) = 29.43, p < .001$, and there was no main effect of pronoun frame but there was an interaction between transitive construction type and pronoun frame, $F(1, 20) = 29.13, p < .001$. Analysing this interaction, Bonferroni-adjusted t-tests revealed that with active sentences the older children performed significantly better with two case-marked pronouns than with one case-marked pronoun, $t(20) = 3.00, p = .007$ ($M = 83.3 > 57.14$). Similarly with passive sentences, they performed better with two case-marked pronouns than with one case-marked pronoun $t(20) = -2.57, p = .018$ ($M = 19.04 < 41.66$).

### 4. Discussion

In the current study we investigated how young children use pronoun frames versus constructional cues (word order and some morphological marking) to successfully
identify agents and patients in both active and passive transitive sentences. The main finding was that children at two different ages used pronoun frames in their sentence interpretations. Children just under three years of age used pronoun frames to interpret grammatical active and passive sentences, as evidenced by the fact that they were better when two case-marked pronouns were present rather than just one (for both actives and passives). In the active, two case-marked pronoun condition they had an 80% successful comprehension rate which shows that despite the relatively high processing burden of the task the youngest children were still able to successfully assign agent/patient relationships when case and word-order cues coalesced. The older children, just under the age of four, interpreted both types of grammatical sentences easily with either one or two case-marked pronouns and only used the pronoun frames cues to interpret ungrammatical sentences (for both actives and passives) - and they did this in a somewhat surprising way, discussed below.

In terms of specific pronoun frames, the younger children were better than chance on all grammatical actives. In contrast, they were better than chance with passives only if they had two correctly case-marked pronouns in subject and object position (e.g., *She is getting tammed by him*). The older group also showed the above chance comprehension of all grammatical sentence types with two correctly case-marked pronouns (e.g, *She is tamming him, She is getting tammed by him*). While the younger group were at chance in the passive condition on sentences that had one case-marked pronoun occurring in either the grammatical subject or object position ([NP\text{neuter} - V\text{ing} - NP\text{accusative}] or [NP\text{nominative} - V\text{ing} - NP\text{neuter}]), the older children succeeded in this condition.

The pronoun frame that had two case-marked pronouns occurring in the ungrammatical subject/object position ([NP\text{accusative} - V\text{ing} - NP\text{nominative}]) resulted in at
chance performance for younger children on both active and passive. The older children were able to overcome the ungrammaticality of the pronoun-marking and assign agent-patient relationships on the basis of the constructional cue.

Overall, for the ungrammatical sentences, the younger children are puzzled and assign agent/patient relationships with no clear or consistent strategy. Although performance with actives is slightly better than with passives, their performance does not differ from chance regardless of construction or pronoun frame. The older children, however, use the pronoun frames. Perhaps surprisingly, for these older children having two (incorrectly) case-marked pronouns aids comprehension over having one (incorrectly) case-marked pronoun. So, only in the older children is there an effect of pronoun frame, which is opposite to what we saw with grammatical sentences where the effect was in the younger children only.

These data suggest, at the very least, that young children learning a language not known for robust case marking nevertheless make use of case marking in sentence comprehension. In the present study we have shown that there is a complex interaction between case-marked pronoun frames and constructional cues. We have also shown that the older children show the same general pattern in comprehension as the younger children and this extends over the ungrammatical case-marked pronoun frames (2 case-marked pronoun frame > 1 case-marked pronoun frame; active > passive). Furthermore, following the notion of degrees of grammaticality (Chomsky, 1961) the fact that him is taming she is not as bad as her is getting tammed by it reflects that in English the dominant active word order is a stronger cue than case-marking, perhaps explaining the younger children’s slightly better performance with ungrammatical actives over ungrammatical passives. The idea that cues are processed
due to their relative strength is exactly the analysis that the competition model supports.

There are two alternative explanations for our results that we must consider. First, it is possible that children did not treat all scenes as equally interesting. For example, the same action performed by a boy on a chair might be more (or less) interesting than one performed by a chair on a boy. Once attention has been focused in this way children might point to the chair-verbing-boy \([\text{NP}_{\text{neuter}} - \text{Ving} - \text{NP}_{\text{accusative}}]\) picture more often than the boy-verbing-chair \([\text{NP}_{\text{nominative}} - \text{Ving} - \text{NP}_{\text{neuter}}]\) leading to differences either above or below chance that do not reflect linguistic processing. This would mean that performance could be lowered in the one case-marked pronoun condition over the two case-marked pronoun condition, simply because although the children understand the sentences in the same way, they have a tendency to look to the \([\text{NP}_{\text{neuter}} - \text{Ving} - \text{NP}_{\text{accusative}}]\) picture leading to errors when they hear an \([\text{NP}_{\text{nominative}} - \text{Ving} - \text{NP}_{\text{neuter}}]\) sentence. In order to test for this possibility the pointing data of the children was re-coded in the critical conditions (all sentences except \([\text{NP}_{\text{nominative}} - \text{Ving} - \text{NP}_{\text{accusative}}]\)) such that responses would reflect whether children had a bias towards a particular pronoun frame. The proportion of times children point to \([\text{NP}_{\text{neuter}} - \text{Ving} - \text{NP}_{\text{accusative}}]\) versus the proportion they point to \([\text{NP}_{\text{nominative}} - \text{Ving} - \text{NP}_{\text{neuter}}]\) (irrespective of the pronoun frame they hear) was 51.78% for the younger children and the 48.21% for the older children. Children were thus not biased towards pointing on the basis of the ‘interestingness’ of the scene.

Second, we know that animacy also affects passive interpretation for both children and adults (Angiolillo & Goldin-Meadow, 1982; Budwig, 2001; Ferreira, 1994). However, in this experiment all test items were animated (literally) to perform the novel verb actions. Therefore it is not clear in what sense we could be said to be
truly manipulating animacy. Because animacy does not play its normal role here, the effects we see between conditions can be more confidently attributed to the morphosyntactic properties of the different sentence types.

The results reported here add an important data point to what we know about the development of the passive construction. Even the younger children in this study showed evidence of sensitivity to word order and passive morphology. Tomasello and Brooks (1998) and Brooks and Tomasello (1999) observed that children as young as two years to two and a half years old could use novel verbs in transitive, intransitive and passive constructions but they had a tendency to use each experimental verb in exactly the same construction in which they had heard the experimenter use it. The fact that children in this study are able to comprehend the novel verbs when they appear in either active or passive constructions is most likely due to two features of the experimental procedure. Firstly, both novel verbs were introduced in the warm-up phase in both forms (tamming/keefing and tammed/keefed) with both human and non-human agents and patients to avoid associating any particular thematic role with subject or object position. Secondly, each novel verb was modelled by the experimenter in the spoken sentence accompanying the cartoons in passive and active constructions throughout the test phase. This clearly helped to license the use of that verb in either passive or active constructions (and in all types of pronoun-frames).

Infants must work out how case marking is realised in the language they are learning on the basis of the language they hear, helping them to map the form and function of argument structure constructions. However, one might assume from the fact that English is relatively limited in its use of case-marking (as opposed to, say, Turkish or Finnish) that children relegate the importance of this information, attending to more powerful predictors of argument structure in English such as word
order. Taken together, these results show that although English is globally impoverished with case, children are able to capitalise on highly frequent, local and reliable cues such as pronouns in order to understand argument structure. In this sense the role of case-marked pronouns has been previously underestimated in English language acquisition. Even very young English children use multiple cues when learning language-particular morphosyntactic forms, including pronoun-frames, to interpret transitive constructions. At different stages in development children are making use of cues in different ways which suggests that at the beginning of development the grouping of cues is prone to extensive revision. This suggests that the cluster of cues associated with transitive sentences take time to coalesce into fully productive constructions. This process continues until the weight of examples is such that transitive schemas are understood in terms of the role that each of the different cues is playing in the whole construction.

5. Acknowledgements

This study was made possible by the parents and children who came in to the Max Planck Child Study Centre in Manchester. Many thanks to Dr. Grzegorz Krajewski who assisted in transformations of the data and statistical analyses.

The following chapter focuses on how children and mothers actually use language in a corpus study. It begins by looking at the role of skewed distribution and cognitive anchoring in schematising the Subject Verb Object construction in English. It then presents a usage-based acquisition model of argument productivity in subject-verb-object constructions.
CHAPTER SIX

1. Introduction.

Cognitive grammar is usage based; it assumes the input to language acquisition consists of encounters with actual linguistic expressions, specified in their phonological, semantic and symbolic aspects. Therefore a major goal of usage based research is to explain the course of acquisition in terms of the semantic, phonological, syntactic and general statistical properties of child directed speech (CDS) – which is, in the western cultures most studied, the predominant source of linguistic experience in the early years of language acquisition. This chapter begins by briefly discussing an example of where bottom-up properties of the input interact with top-down psychological processes – skewed distributions and cognitive anchoring – and concludes with a corpus study that assesses a usage based model of grammatical productivity.

2. The role of skewed distribution and cognitive anchoring.

Skewed distributions have been shown to have a beneficial role in learning linguistic and non-linguistic categories (Avarahami et al., 1997; Casenhiser & Goldberg, 2005; Elio and Anderson, 1984; Genter, Loewenstein, and Hung, 2007). Cognitive anchoring is one mechanism that has been suggested to play a role in this effect (e.g., Goldberg, 2006: 88; Goldberg, Casenhiser, Sethuramen, 2004). According to the theory a high frequency exemplar which emerges early in learning provides a salient standard of comparison against which subsequent instances are judged. This statistical skew in the input facilitates relatively rapid category construction compared to flatter
distributions because a minority of high frequency types provide a relevant anchor that serves to organise memory and reasoning about other types.

Similar skewed distributions have been reported in CDS for English (Goldberg, Casenhiser, Sethuramen, 2004; Naigles & Hoff-Ginsberg, 1995; Theakston, Lieven, Pine & Rowland, 2004 and for Hebrew (Ninio, 1999a; 1999b). For example, Goldberg, Casenhiser, Sethuramen, (2004) conducted an extensive corpus study of children’s and mothers’ speech and showed that tokens of one particular verb are found to account for the lion’s share of instances of each argument frame considered. In the graphs below the verbs that appear in the constructions are on the $x$-axis, and the percentage of the constructions containing that verb is displayed on the $y$-axis. The results for intransitive motion and caused motion constructions are as follows:

<table>
<thead>
<tr>
<th>Form</th>
<th>Meaning</th>
<th>Construction label</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Subj V Obl$_{\text{path/loc}}$</td>
<td>X moves Y$_{\text{path/loc}}$</td>
<td>Intransitive motion</td>
</tr>
</tbody>
</table>

Example: I went to the store.
2. Subj V Obj Ob1\textsubscript{path/loc}  

<table>
<thead>
<tr>
<th>Form</th>
<th>Meaning</th>
<th>Construction label</th>
</tr>
</thead>
<tbody>
<tr>
<td>X causes Y to move Z\textsubscript{path/loc}</td>
<td>Caused motion</td>
<td></td>
</tr>
</tbody>
</table>

Example: He put the book on the shelf.

In a follow-up experiment to the corpus study, eighty-one undergraduate participants were randomly and equally divided into three conditions: the no-training condition, the balanced condition and the high-token frequency condition. Subjects in both conditions other than the no training condition saw a short film that consisted of 16 film clips in which puppets acted out scenes. The descriptions involved novel verbs in a novel construction, “SOVo”, e.g., the king the ball moopo-ed. The construction was formally novel in that it had SOV word order as opposed to standard English SVO order. Additionally, the construction was morphologically marked in that all nonsense verbs had the nonsense suffix -o attached. The construction could be said to designate a scene of appearance and since there is no construction in English devoted to this meaning, the semantics of the construction was novel as well.
In the balanced condition, subjects heard five different novel verbs, each with a relatively low token frequency (4-4-4-2-2). In the high-token frequency condition, subjects again heard the same five novel verbs, but this time one had especially high token frequency (8-2-2-2-2). To be consistent with the finding that the most frequent verb had a very general meaning, the high token-frequency exemplar was designed to encode the meaning of “appearance” in a very general way, without designating a particular manner. The test was designed to determine whether subjects had learned anything about the construction’s semantics by watching the scenes and listening to the novel constructional descriptions of each scene, examined by a forced choice comprehension task. As expected, subjects in the control no-training condition did no better than chance at choosing the correct scene. The balanced condition showed a statistically significant improvement over the control condition, indicating that they had learned something about the construction’s semantics from the training film. As predicted by the skewed input hypothesis, the high-token-frequency condition showed a statistically significant improvement in accuracy over the balanced condition. The results show that the high token frequency of a single exemplar facilitates the learning of constructional meaning. Interestingly, Casenhiser and Goldberg (2005) showed a similar pattern of results for 51 five- to seven-year-old children. The hypothesis is that it is the high frequency of particular verbs in particular constructions that allows children to note a correlation between the meaning of a particular verb in a constructional pattern and the pattern itself, giving rise to an association between meaning and form. In this way, grammatical constructions may arise developmentally as generalisations over lexical items in particular patterns. Implications of this work are potentially far reaching insofar as tokens of constructions are typically centred around one or a few specific words, or around a semantic prototype, even when they...
potentially occur with a much broader range of words or meanings (Diessel, 2002; Goldberg, 1996; Hunston and Francis, 1999; Scheibman, 2002; Thompson and Hopper, 2001). What might this mean for the form that linguistic representations take during acquisition? Whether a number of high frequency items provide the same single anchor for a construction (a kind of coalition of items into a prototype) or whether they might provide a number of different anchors for a given construction (clusters of exemplars) will be largely dependent on the stage of development we are talking about – one cannot form a prototype or schema over an instance of one; the prediction value of the schema would be the same as the instance itself so, in this parsimonious explanation, the cognitive effort of forming a schema outweighs its prediction value. In the usage-based approach, representations are built from the bottom-up in a piecemeal fashion, interacting with top-down social-cognitive reasoning abilities, so the prediction is that exemplar specific representations grounded in experience play a greater role early on in development (cf. results of the prototype study, Chapter 3) before more abstract, schematic knowledge is abstracted over these instances. In Chapter 2 and 4 I have tried to make the argument that, for adults at least, prototypes and exemplar representation are not models we need to choose between. Overall the evidence favours a view in which adult-like categorical competence uses both exemplars and more abstract representations (e.g. Elio and Anderson, 1981; Estes, 1986; Homa, 1984; Malt, 1989; Medin, Dewey & Murphy, 1983; see also review by Ross & Makin, 1999). For example, Figure 6 in Chapter 2 shows how radial prototype conceptual structure is produced as a by-product of ‘seeing’ through the cumulative layers of the symbolic assembly, with exemplars grounded at the bottom of this assembly: we retain item-specific knowledge (e.g.
exemplars) and, eventually, we can abstract over them (e.g., prototypes and abstract schemas) to generate and understand novel instances.

To my knowledge, no author that has investigated the role of skewed type/token distributions in learning has stated that category learning is impossible in with flatter distributions, the emphasis has been to investigate why might skewed distributions have the facilitatory effect they do and simply to note that many grammatical constructions in CDS corpora follow this distribution (e.g., Goldberg, Casenhiser, Sethuramen, 2004). With that caveat in mind, how might skewed distributions actually help categorisation in the process of language acquisition? Figure 1 proposes a cognitive model of schematisation for the caused-motion construction but importantly the principles of the model hold for any construction with a skewed type-token distribution of this type. The y-axis shows the token frequency of particular verb types; the x-axis show types ranked by token frequency. Of course this raises the critical question: types of what? If the infant is treating tokens as types of verbs that appear in the caused-motion construction then the child already has a level of knowledge about this construction that renders the role of skewed input in categorisation somewhat redundant. The default usage-based approach, so to speak, is to assume linguistic conventions are learnt and linguistic categories are constructed unless there is good reason to suspect otherwise. Therefore, in answer to the problem of ‘types of what’, we would not want to suggest infants are ‘linking’ the input to any innate [X causes Y to move Z] representation let alone [Subj V Obj Obl] but rather infants construct these schemas as social agents interacting with other social agents in particular usage events. The question still remains whether the x-axis represents all uses of ‘put’, for example, for the caused-motion construction.
Firstly, the expanded notion of reliability advocated in Chapter 2 is helpful here. The argument is that the traditional notion of reliability (e.g., Bates & MacWhinney, 1984) – how consistently a function is mapped onto a particular form whenever it is present – is too narrow to get the job done and is not the problem infants have to solve anyway. However, if we expand the notion of reliability to cover usage events, then infants can triangulate their reliabilities between referential context, linguistic form and semantic content. The hypothesis is that infants don’t need an innate representation of the type [X causes Y to move Z_{path/loc}] or [Subj V Obj Obl_{path/loc}] but rather, because linguistic forms become associated with particular referential contexts (in the beginning grounded in the perceptually available here-and-now you-and-me) this begins to limit the scope of instances that could appear on the x-axis. In other words, one way of bootstrapping out of this problem without innate abstract linguistic representation is to say in the beginning verbs are grouped together on the basis that they appear concurrently with a caused-motion event perceptually available in joint attention. The events in question all share something in common, the participants behave in a similar way [X caused Y to move to Z], that is, their actions are all intended to accomplish something similar. Thus right from the start intention-reading skills provide important cognitive foundations on which later abstractions are formed, including those that will go on to become grammatical schemas. Of course this strongly suggests infants – if not innately then from a very early age – are already equipped with the shared intentionality and social reasoning skills that this inference is built on, but that is another story (Tomasello, 2008; 2009).

Secondly, I am unaware of any corpus studies that have specifically looked at this, but a factor which would obviously help the infant would be if the ‘lion’s share’ verbs of one construction were not also the ‘lion’s share’ verbs of other constructions.
The $x$-axis could then represent all instances of $x$ in the corpus from the beginning of acquisition without danger that the ‘lion’s share’ verbs of one construction are polysemous with the meaning of two or more, other constructions. Of course, even if this were true the situation isn’t irreparable, because if the constructions have different meanings then presumably they refer to different events, states, attitudes or whatever. Thus under the usage-event definition of reliability, the referential context would help disambiguate between the constructions with overlapping most-frequent forms.
Figure 1. Schematisation under skewed input conditions.

CDS input distribution

Infant at $t_1$

Infant at $t_2$

Infant at $t_3$

Parent Population

Emergent distribution under random sampling
The graph at the top-left represents the verb types and tokens in CDS. To simulate what the child hears, we propose a model that randomly samples from the CDS input distribution, equivalent statistically-speaking to sampling from a parent population. We will then place those samples on the child’s own graphs (t_1-t_3) as the child cumulatively builds their own distributions over time. The skewed input of types in CDS makes it extremely probable that at t_1 the child’s first emerging cognitive anchors will be the most frequent types in the CDS. To illustrate, imagine we place all maternal caused-motion utterances in a bag and withdraw verb tokens one at a time. The probability of any verb type being withdrawn is a function of that verb’s token frequency in the bag. Say there are ten caused-motion maternal utterances, the probability of extracting the verb ‘put’ might be 5/10, ‘get’ 2.5/10, ‘take’ 1.5/10, ‘do/pick’ 1/10. This is a kind of Bayesian reasoning of the expected probabilities: given a caused motion event, what are the probabilities that a particular verb will be used in this context. In the model the probabilities are the same for each sample (equivalent to replacing the tokens after each draw) as it makes no sense to say that because a mother has used a verb she can’t use it again. The dashed lines in the diagrams represent the semantic contribution of each verb type to the emerging schema. The skewed distribution means the most frequent members of the set have a higher probability of being represented in the early sampling and thus forming the core of the proto-schema. The hypothesis is that the frequency of use reinforces the representation of linguistic expressions in memory, which in turn influences their expectation and interpretation in language use. At t_2 the infant is using this schema to categorise new instances (the shaded bars under the curly brackets) on the basis of the emerging caused-motion schema at that point in time. Those instances that have been categorised at t_2 now contribute to the schema and the categorisation of new instances.
at $t_3$ but due to the skew of the distribution the semantic weight that these contribute
to the overall schema is less than those early exemplars. Thus, the strength of the
verb’s role in defining the character of schema is proportional to its frequency within
that construction. At $t_3$ instances in the long tail of the distribution are now
categorised with respect to cognitive anchor verbs. The graphs at the bottom of Figure
1 summarise this sampling process.

To be explicit about the logic of this: the probability of extracting an item is
proportional to its frequency in the set; the more frequent an item is the more likely it
is to be sampled on any occasion; the more likely it is to be sampled the more likely it
will emerge first and semantically contribute to the proto-schema. For the same
reasons, high frequency verbs (if recognised as instances of the same verb) will also
be periodically reinforced in the sampling procedure. We have modelled the infant
distribution in a cumulative way as if all instances are remembered equally. If we
factor in a more realistic memory retrieval and decay process (e.g., Brown, Vousden
& McCormack, 2009) this actually skews the distribution further towards the most
frequent items, again, for the same reasons of probability. The low frequency items
cannot afford to remain unreinforced for very long before their memory trace is so
weak they are no longer available for retrieval. Hypothetically, if forgetting occurred
on a token-by-token basis, high frequency items would be the last ones left in memory
for the same reasons they were the first to arrive. The interspersing of more highly
frequent verbs in between less frequent verbs will help to keep the semantic core of
schema periodically activated and reinforced in memory, which in turns makes it a
more salient cognitive anchor for subsequent categorisations.

While this model suggests why skewed distributions facilitate schematisation
for the caused motion construction where the meaning can be paraphrased as ‘X

causes Y to move to Z’, there is some doubt as to whether cognitive anchoring can account for the acquisition of the English transitive Subject-Verb-Object construction in the same way (e.g. Sethuraman & Goodman, 2004). One reason is that this grammatical pattern instantiates a semantically diverse set of items including verbs that are prototypically transitive (break, cut, hit) in the Hopper and Thompson sense (1980) and those that express psychological experiences (want, see, know). Thus compared with the caused-motion construction for example, there is no one verb like put that accounts for the lion’s share of the tokens which is also schematic for the function of the construction as a whole. For example, consider the range of thematic roles (a-f) that can fall under the general syntactic pattern Subject–Verb-Object.

a) the farmer killed the duck [agent, patient]
b) the lawyer received the summons [recipient, theme]
c) the rock broke the window [instrument, patient]
d) many tourists saw the accident [experiencer, stimulus]
e) the mugger robbed sam of £50 [agent, source]
f) the clown amused the child [agent, experiencer]

Before concluding that skewed distributions are not important in learning SVO constructions it is worthwhile looking at the corpus data. Figures 2a and 2b show the verb type-token distribution for one mother infant pair taken from the Manchester corpus (Theakston, Lieven, Pine & Rowland, 2001) of the CHILDES database (MacWhinney & Snow, 1990). The y-axis shows the token frequency of particular verb types in SVO constructions (the extraction procedure for the SVO utterances in described in detail in the paper to follow), the x-axis show the verb types ranked by
token frequency. For both the CDS and the infant data we can see that there is an excellent fit to a power-law distribution: $R^2 = 0.93$; $R^2 = 0.96$. This means that there are a few verbs that appear with high frequency and most verbs appear with low frequency. By comparison, for a perfectly flat distribution, where for each type there are equal numbers of tokens, $R^2 = 0$.

Figure 2a. The distribution of all verb tokens in SVO constructions in Becky’s mothers’ speech.
In theory, both CDS and infant speech could fit a power-law distribution without there necessarily being any relationship between the particular items used in that distribution. To explore the extent to which the infant’s production preserves the rank order of items in CDS, Table 1 compares the top ten ranked verbs in the SVO construction between CDS and infant speech and verbs from the tail end of the distribution. It is also important not only look at the distribution but the semantic character of the verbs that appear most frequently in the distribution. Semantic generality of the verb is based on Ninio’s (1999a, b) classification system as implemented in Theakston, Pine, Lieven and Rowand (2004).
Table 1. The top ten ranked verbs and verbs from the tail of the distribution.

<table>
<thead>
<tr>
<th>Rank order of verbs</th>
<th>Becky’s mother</th>
<th>Semantic Generality</th>
<th>% of total verb use</th>
<th>Becky</th>
<th>Semantic Generality</th>
<th>% of total verb use</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>want</td>
<td>light</td>
<td>11.65</td>
<td>got</td>
<td>light</td>
<td>12.53</td>
</tr>
<tr>
<td>2</td>
<td>like</td>
<td>part light</td>
<td>9.71</td>
<td>want</td>
<td>light</td>
<td>9.49</td>
</tr>
<tr>
<td>3</td>
<td>do</td>
<td>light</td>
<td>8.74</td>
<td>do</td>
<td>light</td>
<td>9.29</td>
</tr>
<tr>
<td>4</td>
<td>see</td>
<td>light</td>
<td>7.44</td>
<td>like</td>
<td>part light</td>
<td>6.67</td>
</tr>
<tr>
<td>5</td>
<td>got</td>
<td>light</td>
<td>6.15</td>
<td>find</td>
<td>light</td>
<td>6.06</td>
</tr>
<tr>
<td>6</td>
<td>done</td>
<td>light</td>
<td>4.53</td>
<td>see</td>
<td>light</td>
<td>4.44</td>
</tr>
<tr>
<td>7</td>
<td>eat</td>
<td>light</td>
<td>3.88</td>
<td>need</td>
<td>part light</td>
<td>3.84</td>
</tr>
<tr>
<td>8</td>
<td>find</td>
<td>light</td>
<td>3.88</td>
<td>done</td>
<td>light</td>
<td>3.03</td>
</tr>
<tr>
<td>9</td>
<td>break</td>
<td>non-light</td>
<td>3.56</td>
<td>have</td>
<td>light</td>
<td>2.63</td>
</tr>
<tr>
<td>10</td>
<td>lost</td>
<td>non-light</td>
<td>2.59</td>
<td>get</td>
<td>light</td>
<td>2.22</td>
</tr>
<tr>
<td>Sum % of verbs ranked ≤ 10</td>
<td></td>
<td></td>
<td>62.14</td>
<td></td>
<td></td>
<td>60.20</td>
</tr>
<tr>
<td>94</td>
<td>ripping</td>
<td>non-light</td>
<td>0.32</td>
<td>stick</td>
<td>non-light</td>
<td>0.20</td>
</tr>
<tr>
<td>95</td>
<td>says</td>
<td>non-light</td>
<td>0.32</td>
<td>lick</td>
<td>non-light</td>
<td>0.20</td>
</tr>
<tr>
<td>96</td>
<td>scratch</td>
<td>non-light</td>
<td>0.32</td>
<td>talk</td>
<td>non-light</td>
<td>0.20</td>
</tr>
<tr>
<td>97</td>
<td>shake</td>
<td>non-light</td>
<td>0.32</td>
<td>throwing</td>
<td>part light</td>
<td>0.20</td>
</tr>
<tr>
<td>98</td>
<td>shuffling</td>
<td>non-light</td>
<td>0.32</td>
<td>told</td>
<td>non-light</td>
<td>0.20</td>
</tr>
<tr>
<td>99</td>
<td>stick</td>
<td>non-light</td>
<td>0.32</td>
<td>wash</td>
<td>non-light</td>
<td>0.20</td>
</tr>
<tr>
<td>100</td>
<td>talk</td>
<td>non-light</td>
<td>0.32</td>
<td>watch</td>
<td>non-light</td>
<td>0.20</td>
</tr>
<tr>
<td>101</td>
<td>told</td>
<td>non-light</td>
<td>0.32</td>
<td>mend</td>
<td>non-light</td>
<td>0.20</td>
</tr>
<tr>
<td>102</td>
<td>knock</td>
<td>non-light</td>
<td>0.32</td>
<td>wipe</td>
<td>non-light</td>
<td>0.20</td>
</tr>
<tr>
<td>103</td>
<td>wipe</td>
<td>non-light</td>
<td>0.32</td>
<td>pull</td>
<td>non-light</td>
<td>0.20</td>
</tr>
<tr>
<td>Sum % of verbs ranked &gt;10</td>
<td></td>
<td></td>
<td>37.86</td>
<td></td>
<td></td>
<td>39.80</td>
</tr>
</tbody>
</table>

In summary, there are significantly skewed distributions for verb types for both mother and child in SVO constructions (indicated in % columns above and Figures 1a/b). Furthermore, as Theakston, et al., (2004) found, the generic verbs appear most frequently in the mother’s and the child’s speech. This raises the obvious question of why this distribution is seen in the adult and why are these verbs the most frequently used by the adult speaker? There seems to be two obvious answers (i) the linguistic
distribution reflects the pattern of experience: after some time of development most situations are variations on a theme, new experiences are rarely entirely new but a mixture of repetition and invention and thus the events that the utterances describe (including those of the SVO grammatical pattern) might themselves represent a power-law distribution (ii) the verbs that symbolise these events or experiences (go, got, want) apply to wider range of contexts than verbs with a tighter scope (mend, shuffle, wipe) and thus can appear with a wider range of subjects and objects. It is important to state that there is nothing intrinsically special about the semantics of these early verbs in terms of their learnability. All other things being equal, we would expect verbs of much tighter semantic scope (e.g., whittle, decant, claw) to be acquired first if these were ranked highest in the frequency distribution of CDS. But all other things aren’t equal, and verbs with wider semantic scope tend to be more frequent in English (Theakston, Lieven, Pine & Rowland, 2004). So given that the distribution of items is the way that it is for the English SVO construction, this still leaves open the possibility that the semantic character of these early anchoring verbs means they serve as hypernyms for verbs categorised subsequently under the SVO schema.

It is interesting to compare some experimental work on skewed input learning in children (Casenhiser and Goldberg, 2005) with the corpus data presented here. To recap, Casenhiser and Goldberg reported that there was a facilitative effect for the disproportionately high frequency of occurrence of a single verb in a particular construction (such as has been found to exist in naturalistic input to children). The authors concluded that the input is structured in such a way as to make generalisation of constructions straightforward and associations are made between the meaning of the dominant verb in the construction and the construction itself. If we fit the
experimental data of Casenhiser and Goldberg (2005) to the same power-law distribution we have used to estimate skew in this study, we get $R^2$ values of 0.7 and 0.6 for the skewed and balanced conditions respectively. Compared to the skewed experimental condition, the input children actually hear for the English SVO data is more skewed ($R^2 = 0.9$). It is also worth noting that in this experiment the increased skew (from 0.6 to 0.7) resulted in an approximately 33% increase in performance in a novel construction learning task (compared to chance). Thus relatively small increases in skew distribution can lead to improved categorisation performance.

Skewed distributions are not limited to language networks, the same pattern, for example, emerges in Internet growth, academic citations and protein-interaction; a set of findings which have led to the power-law’s status as a global statistical feature of such network (first formalised by Mandelbrot, 1966; Zipf, 1935, 1965). Given the disparate domains in which power-laws have been observed one might be sceptical as to what this distribution can tell us about how children learn language, which is, as far as we know the acquisition of a unique cognitive-social skill. The cognitive anchoring hypothesis gives us a psychological reason to be interested in such distributions from a usage-based approach to language acquisition. Skewed input has a facilitatory role in category formation (and thus generalisation) because the organisational features of memory and inference benefit from a high frequency exemplar which emerges early in learning (Avarahami et al., 1997; Elio and Anderson, 1984; Genter, Loewenstein, and Hung, 2007). These instances provide a salient standard of comparison against which subsequent instances are judged. A challenge to this hypothesis might be categories that represent a diverse set of items where no one item emerges early in category development as cognitive anchor. One such linguistic category which potentially fits this profile is the English SVO construction. This preliminary
investigation has shown tentative evidence that the infant learning the SVO construction can still capitalise on the benefits of cognitive anchoring because (i) the frequency of items in CDS of this construction is significantly skewed. This construction may not need to have one verb that accounts for the lion’s share because the fact that the input is significantly skewed is enough to create some verbs that are used as cognitive anchors (ii) the probabilities associated with sampling from this distribution (equivalent to the infant hearing CDS) makes it extremely likely that an early SVO schema is built around these items and (iii) the semantic character of these early anchoring verbs means they could serve as hypernyms for verbs categorised subsequently under the SVO schema. The results are consistent with results reported for Hebrew where ‘path-breaking’ verbs happen to be general purpose verbs in SVO and VO constructions (Ninio, 1999) and more generally with the idea that early verbs act as templates for the acquisition of later verbs on the basis of their semantics (Clark, 1996; Kay, 1996; Goldberg, 1995).

There are two immediate issues arising from these claims. Firstly, Ninio (1999) has argued that from the very first uses these verbs provide a route in to other verbs. The account presented here is different. Here we want to argue the emerging type/token distribution that infants use to build their constructional schema replicates the type/token distribution of the input because of sampling probabilities (exposure to input). Thus it is not the first use of a verb that is important but how the verb’s ranked frequency among other verbs emerges over time. Verbs which are continually used are likely to emerge first as higher ranked verbs and lead to stronger representations and form-meaning mappings. Because the model samples from CDS there is no reason why these representations couldn’t be built up on the basis of frequency-based comprehension prior to the child learning to say anything, consistent with the usage-
based definition of reliability discussed above as referential context-form-meaning mapping.

Secondly, we need to look at a wide range of type/token distributions in grammatical categories across languages to establish how widespread this process is. For example, Penelope Brown (1998a) has examined the development of argument representation in the Mayan language Tzeltal, in both its lexical and verbal cross-referencing forms, analysing the semantic and pragmatic factors influencing the form argument representation takes. The first 500 multi-morpheme combinations of 3 children (aged between 1;8 and 2;4) were examined. On the basis of this Brown argues that there is no evidence of semantically light ‘pathbreaking’ verbs (Ninio 1996) leading the way into word combinations, at least not in the transitive. She does note that most intransitive roots in Tzeltel are general in the sense that they do not restrict the referents of the arguments that can occur with them. However, transitive roots in children’s vocabularies are predominantly highly specific, with, for example, different verbs for eating depending on what is being eaten, and different carrying verbs depending on what is being carried (and how). This is not very surprising given that this specificity is characteristic of most basic-level transitive and positional verbs in Tzeltel. To be clear, the cross-linguistic prediction of the hypothesis advocated here is that cognitive anchoring can play a facilitatory role in learning argument structure constructions where a language has (a) an argument-structure construction with a significantly skewed distribution of verbs appearing in that construction and (b) the ‘lion’s share’ of the verbs can be considered to have meanings which are prototypical of the constructional as a whole. Brown (1998a) does not provide the specific (distributional) data needed to truly examine this claim but it seems unlikely that Tzeltel CDS, for the transitive at least, satisfies a) or b) of these premises. This does
not mean cognitive anchoring plays no role here, it simply means, as Brown (1998b: 1) acknowledges elsewhere, that how a child acquires a grammatical category is centrally influenced by the structural properties of the input, and the semantic structure of the language plays a fundamental role in addition to distributional facts. For example, cognitive anchoring could still play a role in building proto-transitive schemas in Tzeltel around highly frequent eating verbs, if, for example, there is a skewed distribution of eating verbs and the lion’s share of eating verbs are prototypical of an ‘eating transitive schema’ in general. This is still consistent with the cognitive anchoring hypothesis where a minority of high frequency types provide a relevant anchor that serves to organise memory and reasoning about other types, but it is operating at a level of abstraction beneath that of transitive argument structure, as we might expect from the type of transitives Tzeltel children hear.

The model of schematisation under skewed input conditions (Figure 1) is somewhat predicated on the psychological plausibility of cumulatively storing instances from a distribution. So before we advocate the role of anchor verbs in constructional schematisation any further we need establish the validity of this process. To do this the following paper uses verb collocation behaviour in Subject Verb Object constructions as a test case. We chose this domain because a crucial part of the grammatical schematisation process involves learning what arguments a verb typically appears with, that is the disposition of a particular item to combine with other items, and this is fundamentally associated with the semantic/pragmatic contexts in which that verb is used. We pilot the procedure with one adult-infant pair to test how well this approach works with the aim of extending this to other mother-infant pairs.
Publication 5

A Usage-Based Acquisition Model of Argument Productivity

in Subject-Verb-Object Constructions.

Page: 244
A Usage-Based Acquisition Model of Argument Productivity

in Subject-Verb-Object Constructions

Abstract

The flexibility with which children use their first verbs is taken to be a key indicator of
how their grammatical competence is developing. There is some debate as to whether
young children begin acquiring their grammar as ‘conservative learners’ or whether
they are more ‘avid generalisers’. The alternative views have implications for how
abstractly linguistic knowledge is represented and at what age of development. Using
corpus data we test the usage-based hypothesis that the variability in English-
speaking children’s verb-specific patterns of use is strongly predicted by the
variability they hear. Specifically, we build a computational model that uses child
directed speech as input to create verb-specific developmental trajectories for
subject- and object-slot variability in SVO constructions. With minimal model
architecture (cumulative sampling) we account for the verb-specific developmental
trajectory of productivity in subject and object slots in SVO constructions. The rate at
which productivity increases in these slots is what would be expected if children were
cumulatively storing (viz. remembering) verb-specific argument productivity attested
in the input. The implications are that (i) the variability in the argument slots grows
as a function of the instances of variability children are exposed to and (ii) the verb-
specific nature of this learning suggests that children’s early SVO representations are
lexically based.
1. **Introduction.**

The flexibility with which infants’ use their first verbs is taken to be a key indicator of how their grammatical competence is developing. There is some debate as to whether young children begin acquiring their grammar as ‘conservative learners’ (e.g., Tomasello, 1992, 2000, 2006) or whether they are more ‘avid generalisers’ (e.g., Naigles, Hoff & Vear, 2009). The alternative views have implications for how abstractly linguistic knowledge is represented and at what age of development. If infants are rapidly and readily extending uses beyond their experience then this suggests a level of representation that is sufficiently abstract to generate this use (e.g. “Verb Phrase”, “Subject”, “Direct Object” and so on). A strong version of this theory predicts that all linguistic structures that fall under a certain formal description, such as subject-verb-object (SVO), should emerge at the same time in development and be applied productively across all lexical and grammatical items in the child’s repertoire more-or-less immediately (Radford, 1990). If, however, infants’ generalisations are more local and lexically-based, it suggests that linguistic constructions build up during ontogeny in a piecemeal, more concrete fashion, beginning with fewer and weaker abstractions. This predicts that the acquisition of particular linguistic structures is heavily dependent on the specific language to which a particular child is exposed, and with patterns of early grammatical productivity in the child showing a close correlation with the patterns of use they hear.

There is evidence that there are close links between the way adults use particular words, morphemes, and phrases in Child Directed Speech (CDS) and the way children learn them. For example, the way children use particular verbs (as either transitives, intransitives, or both) is strongly related to the way their mothers used those particular verbs (Theakston et al., 2001). Whether children mark particular
verbs for tense and agreement is related to the way they hear those same verbs used (marked in finite clauses or unmarked in non-finite clauses) by their mothers (Pine et al. 2005). Children's acquisition of some particular grammatical morphemes in English (e.g., past tense –ed, plural –s, progressive –ing) is facilitated when mothers use these morphemes as immediate recasts of the child's utterances that are missing them (Farrar, 1990, 1992). The acquisition order of wh-questions is predicted from the frequency with which particular wh-words and verbs occur in children's input (Rowland & Pine, 2000; Rowland, Pine, Lieven & Theakston, 2003). Children's proportional use of me-for-I errors (e.g., 'me do it') correlates with their caregivers' proportional use of me in 1psg preverbal contexts (e.g., 'let me do it'). Furthermore, the verbs that children produce in me-error utterances appear in complex sentences containing me in the input more often than verbs that do not appear in me-for-I errors in the children's speech (Kirjavainen, Theakston & Lieven, 2009). The pattern of negator emergence (no→not→'nt) follows the frequency of negators in the input; that is negators used frequently in the input are the first to emerge in the child's speech (Cameron-Faulkner, Lieven & Theakston, 2007). Overall then, in support of the 'conservative learner' and the more usage-based approach to language acquisition, the evidence suggests that the environment provides the raw materials out of which young children construct their linguistic inventories – not just the trigger for underlying representations proposed by generative linguists (e.g. Chomsky, 1975, 1981,1995; Crain & Lillo-Martin, 1999).

However, there are still claims that the language-learning child is best characterised as an avid grammatical generaliser right from the start of development. For example, Naigles, Hoff & Vear (2009) report the results of a diary study where eight mothers kept records of their children’s first 10 uses of 34 target verbs. The
children were between 16 and 20 months when the study began (depending on when the children started to produce verbs) and were followed for between 3 and 12 months. These diary records were coded to provide measures of the pragmatic, semantic, and grammatical flexibility of children’s first verb uses. To quantify productivity, they assumed that the production of three to five verbs in a given frame (e.g. transitive, intransitive) displays some amount of productivity. On the basis that all 8 children displayed semantic and grammatical flexibility (as defined by the metric above) before 24 months of age, they argue that this shows avid generalisation – more consistent with early abstraction accounts – rather than production which initially stays close to the input – more consistent with usage-based accounts. As well as the verbs-per-frame measure the authors also recorded the extent to which verbs were used with different lexical items in the subject or object positions. They found that children demonstrated flexibility in filling all these grammatical slots with multiple lexical items. In particular, the children showed lexical subject flexibility, in which a verb was used with at least two different subjects, with approximately one third of their verbs; just over one third of the children used verbs with multiple subjects. Lexical object flexibility was demonstrated by verbs for just under half of the children and by children for just under half of their verbs.

However, crucially, notions of ‘conservativeness’ and ‘productivity’ need to be defined with respect to the input. If we do not know what children have and have not heard, it might be that they are simply reproducing the ordering of the particular words they have heard adults using, or they might be marking SVO relations syntactically, but only for some highly familiar verbs (Tomasello, 1992). Naigles et al. acknowledge that in the absence of knowing more about the language children hear, the conclusions they can draw about grammatical productivity and the
The present study is designed to address this limitation by providing a characterisation of the input to explore the issue of early productivity in a grammatical construction. We investigate the possibility that the variability in children’s verb-specific patterns of use is strongly predicted by the variability they hear. Specifically, we build a model that uses CDS as its input to create verb-specific developmental trajectories for subject- and object-slot variability in SVO constructions. We then test the performance of the model against child corpus data and by doing so we create a baseline against which to judge the productivity claims of competing accounts of children’s early grammatical representations.

2. Methodology and Model

2.1. The corpus

We present the data from one monolingual English-speaking child, Becky, and her mother, taken from the Manchester corpus (Theakston, Lieven, Pine & Rowland, 2001) available from the CHILDES database (MacWhinney & Snow, 1990). At the beginning of recording Becky was 2;0.07 and at the end of recording she was 2;6.19, and for this period was visited at home twice every three weeks. The duration of recording for each visit was one hour, divided into two parts with a 10-minute break in between. During the visits, she was engaged in normal play with her mother.
2.2. Extraction procedure

The spontaneous SVO utterances for Becky (2;0.07-2;6.19) and her mother (2;0-3;0) were automatically extracted from the corpus using CLAN. In order to maximise the sample size of utterances (and thus giving the model the best chance of generating a developmental trajectory) we used the entire corpus of maternal SVO utterances (2;0-3;0) then matched this token frequency, 523 SVOs, with Becky’s first 523 SVO utterances, a period which spanned the first six months of development – typically a period when children are constructing their first multiword utterances.

The first sweep returned utterances that matched the SVO template (utterances had S, V and O in the right sequence on the grammatical tier) but were different from ‘canonical’ SVO in English in the sense that they both receive different formal analysis in the linguistics literature and differ on the basis of their communicative function. For that reason, the following construction types were manually excluded from the data: Questions – *did you watch the circus?*; Subject-Predicate – *that’s like a pineapple*; Infinitival complements – *I want to stroke him*; Prepositional Objects – *I put the chair on the table*; Double-Object – *I’ll make you a gate*; and utterances with unknown or untranscribable elements. Reliabilities between two coders on the exclusion criteria resulted in a Cohen’s Kappa of 0.81 conducted on a 10% sample of 500 utterances. The utterances that remained after the exclusion criteria had been applied constitute the data for the present study and are utterances that lexically realise the ‘basic’ English SVO construction (e.g. *I want it, you do it, Becky wants cake*). For the purposes of this study we were not concerned with the distribution of other lexical material such as auxiliaries, negation or adjectives in the SVO construction, so as far as the subject-object frequency is concerned, an SVO sentence
such as *Becky really wants a big cake* is treated as *Becky wants cake*. After this procedure was applied we were left with 523 SVO utterances for the mother (her entire simple SVO use over the corpus) and 523 utterances for Becky, a period which covers over six months of recording from 2;0 -2;6.

2.3. The model

The aim of the model is to simulate the developmental trajectory of individual verbs’ collocation behaviour, as defined by the type frequency of the subjects and objects they combine with. To summarise, the way it does this is by randomly sampling from the SVO CDS – a proxy for the child’s SVO input – and cumulatively storing these patterns of use – a proxy for the child’s memory for item-specific collocations. This gives a cumulative sequence of numbers that represents the ‘substitutability’ of the verb argument slots as sampled from CDS. For example, if the model samples on three successive occasions from CDS, ‘*I eat it*, ‘*I eat it*, ‘*you eat it*’, the cumulative type frequency for the subject slot is 1, 1, 2 (*I, I, you*) and the object slot is 1,1,1, (*it, it, it*) for the verb ‘*eat*’. The more substitutions that appear in these slots – as measured by type frequency – the more productive a given verb is taken to be. The model’s output, therefore, consists of two parameters per verb: subject and object type frequency. To build up a picture of how these parameters change over time the model repeatedly samples from the pool of SVO utterances spoken by the mother in the corpus. Obviously, the child does not have access to the global statistics of the corpus (from 2;0-3;0) at the start of the corpus recording (2;0), so the model keeps track of individual verbs’ subject-object collocations by cumulatively adding the type frequency counts from the previous sample (1 type at T1 and 1 type at T2 = cumulatively 1 type at T2 if types are same and 2 types at T2 if they’re different).
For the purpose of evaluating the performance of the model, Becky’s utterances were divided into ten stages with equal numbers of SVO utterances. This resulted in 10 stages each with 52 SVOs (with 3 utterances omitted randomly). In exactly the same way as we did for the model, the cumulative subject-object collocations were calculated per verb for Becky. We then track this development over a six month period beginning at two years-old. The difference is that for the child, the collocation statistics are calculated for the utterances in the sequence Becky actually uttered them (beginning at 2;0 and ending at 2;6) whereas for the model the SVOs were extracted randomly from any time in the maternal corpus (2;0-3;0). The number of SVOs per stage was then matched between child and model, so that the model randomly sampled the same amount of SVO utterances per stage as the child uttered (52). To be clear, one stage equals a set of 52 SVOs, ten stages equal six months of development even though this could mean that data from the same recording session (and the same age) ends up in two different stages. This gives us a way of directly comparing productivity between model and child; we ask ‘for \( x \) amount of SVO utterances how well does the model predict the development of subject-object variability for individual verbs’? If the model shows a good fit with the child’s developmental trajectory for individual verbs, then this shows Becky’s verb-specific argument productivity can be modelled as a statistical sampling process that correlates highly with the contexts in which she has heard those verbs used – which is what we would expect under the usage-based approach. Where the model departs from the actual developmental trajectory of the child, this suggests that Becky is combining verbs with subject and object types in a way that is different (either ‘over productive’ or ‘underproductive’) with respect to SVO input she has been exposed to.
To be clear, we are not comparing particular items found in the subject and object roles, we are primarily interested in the extent to which a) the development of argument productivity can be modelled by sampling productivity patterns in the input and b) that productivity is verb-specific. For example, a) to what the extent can an asymmetry in subject and object productivity in an infant’s use be modelled as a statistical sampling process from the input and b) to what extent do each of the verbs have their own subject/object asymmetric developmental pathway? Thus productivity is defined at the level of semantic/pragmatic valency of that verb – who is doing what to whom is less important than the variety of participants attested in the input. So infants could be productive at the lexical level – they are using referents unattested in the input – yet they could be as productive as would be expected from their exposure to the input at the level of verb valency.

There is evidence that children are sensitive to the semantic homogeneity of the words observed to fill a slot in a low-scope schema (Bannard & Matthews, 2010). Bannard & Matthews predicted that children should be more likely to detect a productive slot when it has tended to be filled with semantically similar items in the past. For example ‘a piece of X’ is generally filled with a semantically more diverse range of words than ‘back in the X’, which tends only to be filled with words denoting containers. 2 and 3 year–olds found it easier to produce unfamiliar variants of these constructions if that the slot was highly semantically homogeneous in the input. So the point here is that particular verbs occur with particular semantic groups of objects, and that it is the size of this potential group rather than the particular items in CDS that determines relative productivity.
3. **Results and Discussion.**

This section is divided into two parts. Firstly, there is a general quantitative characterisation of the distributions of subject and object types in SVO constructions for Becky and her mother. Secondly, we focus on how well the model can account for the development of these distributions for specific verbs.

We started by implementing Becky’s collocations for subject-verb and verb-object in a dynamic network (Figure 1). By doing this we can collect the frequency and connectivity statistics for the SVO network at different points in development. It is also a useful visualisation tool that allows us to see entrenchment of particular collocations in a very literal sense as repeated patterns of meaningful use (as indicated by heavier connections between items in the network). It condenses six months worth of Becky's recorded SVOs (523 utterances) into nine slides. It is a cumulative representation of the utterances and they appear in the network in the order Becky uttered them. It is a directed graph in the sense that the sequence in which the words are uttered matters, indicated by the arrows in the graph (S → V → O). Nodes represent individual words (red = subject, green = verb & blue = object). The size of the nodes grows over time as a function of item token frequency. Lines between nodes represent the use of a particular sequence of SVO. The weights of lines grow as function of the frequency with which any one node (word) is followed by another node (i.e. collocation).

The nodes on the left are subject types, verbs in the middle and objects are on the right. S, V and O are aligned in this way to make it more interpretable but the axes are not meaningful here as this is a network graph; the important features are the number of nodes for a particular grammatical slot (type frequency) size of nodes...
(token frequency) the connectivity (collocation or ‘degree’ distribution); and the arrows (the direction).

Figure 1. Development in the SVO network by age.

It is clear to see that there is differential growth in the SVO network: there are more object types than verb types, and more verb types than there are subject types; the token frequency of items in the S, V and O slots is also non-evenly distributed; and most importantly for characterising the collocation distribution, a handful of the nodes
have very large number of links (the hubs, like ‘I’, ‘want’ and ‘it’) whereas most others are connected to one or two other nodes. To examine this distribution in greater detail we plotted the collocation statistics associated with each verb for Becky’s Mother and for Becky (Figure 2). The y-axis plots the frequency of subjects (blue dots) and objects (pink dots) that collocate with a verb; the x-axis show verbs ranked by subject/object token frequency. When we plot the subject-verb collocations (‘in-degree’ in network terminology) and verb-object collocations (out degree) there is close similarity between the child’s and the mother’s use. To describe the shape of this curve we used a measure of skew which estimates the variance in the distribution explained by a power-law function\(^{11}\) (Mandelbrot, 1966; Zipf, 1935, 1965). Both show an excellent fit to a power-law distribution, Becky’s Mother: Verb-Object $R^2 = 0.93$; Subject-Verb $R^2 = 0.89$; Becky: Verb-Object $R^2 = 0.94$; Subject-Verb $R^2 = 0.91$.

\(^{11}\) Formally this is given as the probability distribution function: $p(k) \propto k^{-\gamma}$ where the probability that a randomly chosen vertex (a verb in this case) has $n^{th}$ rank $p(k)$ is proportional to that vertex and the exponent $\gamma$, a scaling parameter usually $2 \leq \gamma \leq 3$. When plotted on a log-log scale power law distributions follow a straight line.
This confirms the pattern seen developing in Figure 1, namely i) there are relatively few verbs that combine with a lot of subjects and objects, and many verbs that combine with relatively few subjects and objects, and ii) on average, the verbs are relatively freer to combine with objects than they are with subjects in the SVO construction (as shown by the different gradients of the power-law fit lines). The fact that both the mother and Becky have very high R values for overall collocations and show a very similar dissociation between subject and object collocations (Figure 2) indicate that there are close links between the mother’s and Becky’s overall productivity distributions in subject and object slots. However, this is a static picture of the collocation distribution and gives us no indication of how these patterns are emerging in ontogeny, therefore we now discuss how well the model predicts the development of subject-object variability for individual verbs.
We used the model to predict the subject-object type frequency for the 10 most frequent verbs uttered by Becky. Because of the skewed distribution of use (Figure 2), the majority of the verbs have relatively low overall token frequency in the corpus. Therefore we chose the most frequent verbs to give the model the greatest opportunity to fit a developmental trajectory to the verbs. Obviously we do not have access to all the SVOs Becky has heard so the absolute values of the model are not important – recall that the sampling procedure by the model is a proxy for Becky’s SVO CDS. Rather, given the limited maternal input we do have, we are interested in how well the model reflects *the shape of the developmental trajectory*. Precisely how to interpret the trajectory is discussed in detail later but broadly speaking if there is a good fit between model and child then the model has successfully replicated a) the growth in slot type variation for that verb and b) the way in which the ratio between subject and object productivity changes over time. In terms of making claims about the way Becky is learning the SVO construction, this does assume the relative frequencies of subject-object collocation between verbs in the corpus are representative of what Becky has heard outside of the corpus. There is some evidence that this assumption is justified. For example, Theakston et al., (2001, 2002), using the same Manchester corpus used in this study, report strong correlations between different mothers in the rank order frequency distribution of individual verbs in their samples. This obviously did not examine the collocation distribution we are interested in here but does suggest that individual samples are in some ways representative of wider use.

Figure 3 compares the output of the model with Becky’s production data. For each verb there is a plot which charts the development of productivity of the argument slots; cumulative subject types are on the y-axis and cumulative object types are on
the $x$-axis. The points on the lines represents the 10 stages\textsuperscript{12}, so the first point represents the cumulative subject and object types for one verb after 52 spoken SVOs for Becky and 52 sampled SVOs for the model. The next point shows the cumulative subject and object types for the same verb after 104 spoken SVOs for Becky (the next 52 SVOs Becky utters in the corpus plus the previous 52) and 104 sampled SVOs for the model (the next 52 SVOs randomly sampled from the SVO CDS plus the previous 52). For each verb we also show its rank frequency and the bivariate correlation between the model and child for both subject and objects (* indicates significance at $p<0.05$, **indicates significance at $p<0.01$).

\textsuperscript{12}Note that there aren’t always 10 points visible as some stages do not add anything to the cumulative type frequency and in such cases the points lay on top of one another.
Figure 3. The development of subject-object variability in both mother and child.

Cumulative subject types are on the vertical axis and cumulative object types are on the horizontal axis.

<table>
<thead>
<tr>
<th>Rank token frequency of Verb</th>
<th>Verb token</th>
<th>Verb</th>
<th>Subject-Verb Model-Child Correlation</th>
<th>Object-Verb Model-Child Correlation</th>
<th>Model (pink)</th>
<th>Becky (blue)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1= 54</td>
<td>DO</td>
<td></td>
<td>R=.68*</td>
<td>R=.91**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2= 33</td>
<td>GOT</td>
<td></td>
<td>R=.73*</td>
<td>R=.73*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2= 33</td>
<td>WANT</td>
<td></td>
<td>R=.81*</td>
<td>R=.90**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>27</td>
<td>DONE</td>
<td>R=1***</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>----</td>
<td>-----</td>
<td>------</td>
<td>--------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>26</td>
<td>LIKE</td>
<td>E^{13}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>23</td>
<td>FIND</td>
<td>R=.64*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td>22</td>
<td>SEE</td>
<td>R=.62</td>
</tr>
</tbody>
</table>

---

^{13} Because there is no variation in Becky’s object type it is impossible to conduct a correlation with the model.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th>HAVE</th>
<th>R=.86**</th>
<th>R=.94**</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>8</th>
<th>13</th>
<th></th>
<th>FOUND</th>
<th>R=.27</th>
<th>R=.54</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>8</th>
<th>13</th>
<th></th>
<th>GET</th>
<th>R=.38</th>
<th>R=.44</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>9</th>
<th>12</th>
<th></th>
<th>EAT</th>
<th>R=.94**</th>
<th>R=.95**</th>
</tr>
</thead>
</table>
The first thing to notice is that the shape of the trajectories are different for all of Becky’s verbs, that is, the subject and object slot of her emerging SVO constructions are developing productivity at different rates for different verbs. This is true for verbs with identical token frequency (‘got’ and ‘want’ both have a token count of 33) and those with very similar token frequencies (‘find’ has a token frequency of 23 and ‘see’ has a frequency of 22). Secondly, the model also produces unique developmental trajectories for each verb, so for both Becky and the model, there is differential growth in the productivity of subject and object slots, and this growth is verb-specific for a syntactic structure that falls under the same formal description (SVO).

If the gradient of the model’s slope is shallower than that of Becky’s (as it is in the case of ‘got’ and ‘get’) then this suggests that the productivity of Becky’s subject slot is becoming productive at a slower rate than we would expect given the input and the object slot is becoming productive at a faster rate than we would expect given the input. If the gradient of the model’s slope is steeper than that of Becky’s (as it appears to be marginally in ‘eat’) then this suggests that the productivity of Becky’s object slot is becoming productive at a slower rate than we would expect given the input and her subject slot is becoming productive at a faster rate than we would expect given the input. Where the developmental trajectories of model and child do not differ in their shape, the ratio of subject-object productivity is predicted by the model.
The verbs ‘done’ and ‘like’ provide interesting examples. Both have very similar token frequencies in Becky’s overall SVO production whereas the shapes of the developmental trajectories are very different. For ‘done’, in the early stages of development the subject and object slots are developing in productivity at equal rates, as indicated by the straight line for both Becky and the model. In the later stages of development the line is vertical, which means the subject slot in the ‘done’ SVO schema is continuing to use new lexical material whereas the object slot is recycling items already encountered – a change in the developmental trajectory that that model seems to capture. For ‘like’, Becky’s developmental trajectory is vertical and the trajectory begins at a high object type. What this indicates is that Becky has a productive range of object types from the first recording in the corpus but she does not add to these throughout the corpus sample. Becky fixates on a set of object types when using this particular verb, whereas the subject slot becomes productive gradually. In this case the model underestimates the early productivity of the object slot for ‘like’. One way in which this difference between the model and the infant could occur would be if Becky talked about a smaller range of things she likes more often than her mother. From a syntactic perspective this suggests at the start of the corpus Becky already has the partially productive lexical template, such as ‘I like X’ which is gradually elaborates into ‘X like X’ over a period of months. Presumably another source of the variation between model and child is due to the fact that we haven’t controlled for the number of utterances of a given type (e.g., matching the number of SVO utterances with the verb ‘like’ between model and child), thus there is no guarantee the model has the same number of utterances with each verb as in Becky’s data, either at a particular stage of development or overall. If we were to control for this however, we would be weakening the analogy between the model’s
sampling process and the infant hearing CDS – the child obviously does not control for type variation in this way. The fact that model does well overall at predicting developmental trajectories suggests that this source of variation is limited, however, one way in which we could further improve the model’s performance would be to sample a larger corpus of SVO utterances. In this way the relative impact on the developmental profile of a particular corpus file where, say, the child fixates on a particular SVO verb sequence, on would be reduced.

In summary, the fact that out of the 12 verbs, 8 had significant correlations between model and child on subject types and 9 had significant correlations on object types suggests that on the whole, the gradient of the developmental trajectories did not differ significantly between model and child. In other words, the model, which samples the statistics of verb argument productivity from SVO CDS explains the majority of the developmental variation in the way the child produces verb argument productivity in SVO constructions.


With minimal model design (cumulative sampling) we have been able to predict the verb-specific developmental trajectory of productivity in subject and object slots in Becky’s SVO constructions. The rate at which productivity increases in these slots is what we would expect if the child were cumulatively storing (viz. remembering) verb-specific collocations attested in the input. In line with the results in Naigles, Hoff & Vear (2009), the data presented here also shows that young infants do show variability in the subject and object slot – the SVO construction is gradually becoming more productive – but what we have shown is that this variability grows as a function of the instances of variability they are exposed to. The verb-specific nature of this learning
suggests that Becky’s early linguistic representations are lexically based schemas that reflect the collocation statistics between items in the SVO construction. For example, early in development the object slot associated with ‘like’ has high variability (or high entropy in information theory) in relation to the subject slot, with the subject slot only later becoming more productive. A usage-based approach characterises this shift by suggesting Becky’s linguistic representation begins as ‘I like X’ and is gradually elaborated into ‘X like X’. The SVO schema emerges when a number of such schemas (X like X, X wants X, X sees X…) are recognised as instances of a more general syntactic pattern (SVO) in which the linguistic items play similar communicate roles, what Tomasello has called functionally based distributional analysis (Tomasello, 2003: p 169).

The success of the model in predicting patterns of use is consistent with the idea that sampling from CDS is a good proxy for the child’s input and cumulatively storing these patterns of use is a good a proxy for the child’s memory for item-specific collocations. There are potentially wider implications in demonstrating this kind of statistical sensitivity to distributions of types and tokens that fall under a grammatical category. For example, skewed distributions have been shown to have a beneficial role in learning linguistic and non-linguistic categories (Avarahami et al., 1997; Casenhis & Goldberg, 2005; Elio and Anderson, 1984; Genter, Loewenstein, and Hung, 2007). Cognitive anchoring is one mechanism that has been suggested to play a role in this effect (e.g., Goldberg, 2006: 88; Goldberg, Casenhis, Sethuramen, 2004). Similar skewed distributions have been reported in CDS for English (Goldberg, Casenhis, Sethuramen, 2004; Naigles & Hoff-Ginsberg, 1995; Theakston, Lieven, Pine & Rowland, 2004 and for Hebrew (Ninio, 1999a; 1999b). The role that the most frequent verbs in these distributions may play as either ‘path-
breaking’ (Ninio, 2006) and/or as ‘anchors’ (Goldberg, 2004) is somewhat predicated on the psychological plausibility of cumulatively storing instances of these verbs and being sensitive to their relationship to other verbs in a distribution. Here we have provided evidence that usage patterns can be effectively modelled as such a process.

The approach presented here improves on previous attempts to discuss the development of productivity in grammatical constructions in two main ways. Firstly, in this study ‘conservativeness’ and ‘productivity’ are defined with respect to the input. We have used a baseline measure (the model) to address the question of whether the child begins learning as an avid generaliser or conservative learner. When we do this, the question is not actually whether they are avid or conservative but whether their generalisations are justified with respect to variation and quantity of the uses they have been exposed to at that point in development. Secondly, studying the developmental trajectory of grammatical slot variability recognises productivity as a continuum, rather than imposing some arbitrary threshold at which a child is either considered productive or not. The approach taken here is less interested in whether or not a particular point in acquisition has been reached but rather in determining what factors can explain the shape of development. The notion of a productivity continuum maybe useful in another sense. Some authors have argued for a binary semantic distinction between light verbs and non-light verbs and that so-called path-breaking verbs provide an important role for infants constructing their early grammatical schemas (e.g. Ninio, 1999). However it is difficult to draw a theoretically or empirically justified distinction between which verbs are light and non-light. By definition light verbs are wider in semantic scope and thus are more likely to combine with a wider range of subject and object types. Ranking verbs on a continuum of type variation, in the way we have done in this study, may offer a more principled way to
define the pragmatic and semantic scope of verbs than trying to decide a priori whether a verb is light or non-light.

One obvious weakness is that we have only tested the model’s performance on one infant-mother pair. It is very unlikely that the skewed distribution of collocation use would be different in different samples of CDS given the ubiquity of power-law distributions in spoken corpora (Zipf, 1949, 1965; Mandelbrot, 1983). The point here was to test whether the principle works, but it has yet to be shown that the model can account for variation between mother-infant pairs, something that we are currently testing.

In summary, our results are consistent with the usage based approach to language acquisition in that we have shown that Becky is learning how individual verbs combine with subjects and objects from the input. It is consistent with the idea that early in development infants’ generalisations are more local and lexically-based, it suggests that the acquisition of particular linguistic structures is heavily dependent on the specific language to which a particular child is exposed, and with patterns of early grammatical productivity in the child showing a close correlation with the patterns of use they hear.

The final chapter of this thesis summarises the experimental and theoretical work; identify some cognitive features and properties of the input that seem to be important in all the studies; critiques the strengths and weaknesses of the usage-based approach adopted here; and finally, concludes by suggesting some key issues future work in usage-based approaches to the acquisition of grammar need to address.
CHAPTER SEVEN

This chapter has four sections: firstly, there is a summary of the experimental and theoretical work presented in this thesis; secondly, I identify some cognitive features and properties of the input that seem to be important in all the studies; thirdly, there is a critique of the strengths and weaknesses of the usage-based approach adopted here; and finally, I conclude by briefly suggesting some key issues that future work in usage-based approaches to the acquisition of grammar needs to address.

1. Review of the thesis

This thesis aimed to address a central question in any construction-based, usage-based theory of language acquisition: how children get from more concrete and item-based constructions to more abstract constructions. The overall approach places central importance on meaning and the role of cognition to categorise chunks of linguistic experience into conventional grammatical units. Chapter 1 outlined the historical and conceptual foundations in the field of language acquisition and justified the usage-based approach taken in this thesis. Chapter 2 then considered what usage-based theories mean when they characterise language acquisition as a developing inventory of constructions that are more-or-less schematic. One of the main challenges to any usage-based approach is to specify in greater detail the learning mechanisms that underpin generalisations. Therefore the aim of this paper was to make the usage-based approach to learning more explicit and by focusing on one construction I wanted to show how, in principle, form-function pairings are learnable. In order to do so required bringing together findings from categorisation and analogy, social cognition and construction grammar in a synthesis that is best described as developmental
cognitive linguistics. Specifically, I hoped to explain how categorisation and analogy can be unified, explicate the role of analogy in the construal of events, and integrate social reasoning as part of grammar learning process. Especially important in this explanation was the fact that humans possess a unique set of social-cognitive and social motivational-skills that allows language to happen. Furthermore, early linguistic categories are formed around the underlying functional core of concepts and on the basis of their communicative discourse function. This, combined with powerful pattern-detection skills, enables distributional regularities in the input to be paired with what the speakers intend to communicate.

Chapter 3 investigated the topic of how infants construct grammatical categories by taking a cross-linguistic look at the transitive construction in the context of a prototype theory of categorisation. Much research has shown that young children are slow to form abstract constructions because they fail to see the more general applicability of syntactic markers such as word order and case marking. In this published paper we attempted to explain this fact by investigating the nature of the language children do and do not hear, specifically, the reliability and availability of the linguistic cues they are exposed to. Our aim in this paper was to add a developmental account to existing prototype theories of language processing by taking a more detailed look at one syntactic construction in four languages (English, German, Polish and Cantonese). Overall, we hoped to have shown (i) that an abstraction is formed initially on the basis of frequent overlapping cues, so that the initial abstraction constitutes what will eventually be the prototype of a more complex, more abstract category; (ii) the importance of high-frequency forms in providing ‘anchor-points’ from which more abstract generalisation will gradually emerge; (iii) isolating individual cues is difficult to begin with as they often occur
redundantly in the input; (iv) the construction redundantly marked with multiple cues could have a special status as a nucleus around which the prototype forms; and (v) the nature of the input, as characterised by reliability and availability, is a strong predictor of cross-linguistic differences in language acquisition. Our theoretical claim here was that there are some basic principles of frequency, reliability, and conceptual development that will be critical and play similar roles in all cases.

Chapter 4 applied the theoretical investigations of the previous chapter to an empirical experiment, specifically, making developmental comparisons of the prototypical semantics of the transitive construction in English. There is much controversy about whether linguistic categories and schemas are acquired in the same basic manner as categories and schemas in other domains of human cognition. In this submitted paper we investigated the English transitive construction for evidence of the kind of prototype effects characteristic of non-linguistic categories, in both adults and young children. Adapting the prototype-plus-distortion methodology of Bransford and Franks (1971), we found that whereas adults were lured towards false positive recognition of sentences with prototype transitive semantics, young children showed no such false positive recognition. Thus this paper presented novel experimental evidence that children are, counter-intuitively, better than the adults at remembering which sentences they have and haven’t heard, suggesting a significant new fact about how linguistic prototypes are formed ontogenetically. This advances the usage-based theory of language acquisition by suggesting that acquiring syntactic constructions is a process of categorisation similar to that occurring in other cognitive domains, and that the initial developmental stages of the process are more dominated by exemplar-specific information while, for adults, the prototype is more important.
Chapter 5 considered further the role of different cues in children’s understanding of argument-structure constructions by examining the role of case-marked pronoun frames in early comprehension of transitive constructions in English (paper in press). Infants must work out how case marking is realised in the language they are learning on the basis of the language they hear, helping them to map the form and function of argument structure constructions. However, one might assume from the fact that English is relatively limited in its use of case-marking (as opposed to, say, Turkish or Finnish) that children relegate the importance of this information, attending to more powerful predictors of argument structure in English such as word order. Taken together, the empirical results of Chapter 5 show that although English is globally impoverished with case, children are able to capitalise on highly frequent, local and reliable cues such as pronouns in order to understand argument structure. In this sense the role of case-marked pronouns has been previously underestimated in English language acquisition. Even very young English children use multiple cues when learning language-particular morphosyntactic forms, including pronoun-frames, to interpret transitive constructions. At different stages in development children are making use of cues in different ways which suggests that at the beginning of development the grouping of cues is prone to extensive revision. This suggests that the cluster of cues associated with transitive sentences take time to coalesce into fully productive constructions. This process continues until the weight of examples is such that transitive schemas are understood in terms of the role that each of the different cues is playing in the whole construction.

Chapter 6 focused on how infants and mothers actually use language in a corpus study. The flexibility with which children use their first verbs is taken to be a key indicator of how their grammatical competence is developing. There is some
debate as to whether young children begin acquiring their grammar as ‘conservative learners’ or whether they are more ‘avid generalisers’. The alternative views have implications for how abstractly linguistic knowledge is represented and at what age of development. Using corpus data we tested the usage-based hypothesis that the variability in English-speaking children’s verb-specific patterns of use is strongly predicted by the variability they hear. Specifically, we built a computational model that uses child directed speech as its input to create verb-specific developmental trajectories for subject- and object-slot variability in SVO constructions. With minimal model architecture (cumulative sampling) we accounted for the verb-specific developmental trajectory of productivity in subject and object slots in children’s SVO constructions. The rate at which productivity increases in these slots is what would be expected if children were cumulatively storing (viz. remembering) verb-specific argument productivity attested in the input. Children are learning the disposition of specific verbs to combine with particular subjects and objects from the semantic and pragmatic contexts in which that verb is used. The implications are that (i) the variability in the argument slots grows as a function of the instances of variability they are exposed to and (ii) the verb-specific nature of this learning suggests that children’s early SVO representations are lexically based.

2. Some common themes: constraints on cognitive processing and properties of the input.

This section attempts to identify some common themes that recur throughout the thesis and are thus important in answering the overarching question: how children get from more concrete and item-based constructions to more abstract constructions. Specifically, I consider the role of token frequency, type frequency, type/token
distribution, schema complexity, analogy/categorisation, and reliability. Before I do that I briefly discuss schema elaboration and the abstraction process itself.

Figure 1. A taxonomic hierarchy of instances and schemas; a hypothesised architecture of the linguistic symbolic assembly.

2.1 Schema abstraction and elaboration.

Figure 1 shows a schema-instance hierarchy that I will use to summarise the main themes of this thesis. The bottom-up (bottom red box) input are actual usage events grounded in context (represented as i’s in Figure 1). These usage events interact with top-down (top red box) uniquely human cognitive-social psychology as the infant tries to make sense of the communicative function of utterances. As social agents interacting with other social agents they form schemas that capture the functional and formal similarities between instances as schemas (represented as s’s). The hierarchy implies schemas can also be instances of more general schemas (represented as s/i’s). For example the word [tri:] is an instance of a more schematic unit [NOUN], which
turn is an instance of [WORD]. These schemas are linguistic units of established and/or recurring patterns of processing activity, in short, linguistic expertise. It should be clear by now how central the notion of ‘schemas abstracting over instances’ is to any usage-based theory of cognitive grammar as it has been discussed in one sense or another in all preceding chapters. The deep question is, for the infant what ‘counts’ as an instance of a schema? The answer to this will obviously depend on the infant’s stage of development, that is, the knowledge available for abstraction that is in the system at that time. For example, in the usage-based approach, the first time an infant hears an utterance of the form *john kicks mary*, this cannot possibly be an instance of a transitive SVO schema as there needs to be at least two transitive utterances (instances of *x*) for a schema to be schematic of anything. In this sense, all utterances start life towards the more idiomatic end on the constructional continuum, the difference is that sentences of the form *john kicks mary* are destined to become subsumed (for an English-learner) under a more general transitive SVO grammatical schema, whereas *the more the merrier* is destined to remain comparatively unproductive as it has comparatively fewer analogies with other families of constructions on the basis of form or function.

Figure 2 is of course a massive oversimplification of what is actually going on in schema elaboration but it serves to illustrate two points about the dynamic nature of language acquisition; schemas inherit the network structure of those before them (in time); and instances are grouped and regrouped into new schematic structure as new information becomes available to the system.
In Figure 2, highlighted in blue, we see that a particular linguistic unit – it could be a word, morphological unit, phonological segment – at the start of development has been categorised in a certain way. If we follow that element through different stages of development, following the arrows in the diagram, we see that as schemas are elaborated and change in response to other schemas, that particular unit is dominated by different schemas at different times, it is subject to being categorised and recategorised throughout development. In Chapter 2 I aimed to show how this categorisation can only proceed as fast as the evidence is accumulated and that the ‘depth’ of the abstraction is proportional to the instances accumulated. In Chapter 3, by looking at SVO constructions in English, German, Polish and Cantonese we saw how some of the most critical stages in linguistic developmental are those where cues are grouped and regrouped to maximise the predicative power of a syntactic category.
to infer a function or functions. This dynamic categorisation predicts that infants might go through phases where they produce competing ways of saying the same thing, because, for a period of development a linguistic item could be part of a polysemous network – sanctioned by two (or more schemas) – and then later, after more linguistic evidence, it is recategorised more in line with the linguistic convention of the community. For example an infant might go through a phase of saying both *he giggled* and *he giggled me* where, under this account, *giggle* appears as an instance of both the transitive and intransitive schema, before later in development being recategorised as a intransitive-only verb (or decategorised as a transitive).

Chapter 5, which looked at pronoun-frames, showed how infants at different ages were using cues in different ways, suggesting cues are being weighted and reweighted as to their relative importance in predicting how these items work in the particular language they are learning. All this emphasises that instances are part of a dynamic hierarchy, classified and reclassified over development as the network becomes increasingly interconnected.
2.2. Token frequency

Throughout the thesis, the frequency with which infants encounter particular linguistic units of their language has proven to be very important in predicting how and when they learn those units. For example, the frequency of SVO constructions relative to other constructions across different languages (Chapter 3), the relative frequency of the active/passive voice within a language (Chapter 5), and the frequency of particular verbs in SVO constructions (Chapter 6). The evidence from the empirical chapters and the theoretical reviews converges on the usage-based view that that frequency of use reinforces the representation of linguistic expressions in memory, which in turn influences their expectation and interpretation in language use. By contrast with a generative-grammar view that maintains a meaningful theoretical disjunction between what constitutes knowledge of a language (competence) and how knowledge of language is put to use (performance).
2.3. Type frequency.

Figure 4. *Contrasting type frequencies in two schematic networks.*

While raw frequency is obviously important, variation is equally important in dictating the taxonomic character of a schema-instance hierarchy. For example, the transitive construction, discussed throughout this thesis, has a verb type frequency of approximately 2,119, that is, in English around two thousand verbs can felicitously appear in this construction compared with approximately 69 for the ditransitive construction. This kind of difference must have an effect on both the structure of the category and how children learn it. For example, how easily new verbs are assimilated to the schema has implications for the learnability of that construction above and beyond the overall frequency of those items that can appear in that schema.

In Chapter 6 we showed that the rate at which productivity increases in the Subject and Object slots of a Subject-Verb-Object construction is what we would expect if the child were cumulatively storing (*viz.* remembering) verb-specific collocations attested in the input. We showed that this type variability grows as a function of the instances of variability they are exposed to. The verb-specific nature of this learning suggested that Becky’s early linguistic representations are lexically based schemas that reflect the collocation statistics between items in the SVO.
construction. For example, early in development the object slot associated with ‘like’ has high variability in relation to the subject slot, with the subject slot only later becoming more productive. A usage-based approach characterises this shift by suggesting Becky’s linguistic representation begins as ‘I like X’ and is gradually elaborated into ‘X like X’. The SVO schema emerges when a number of such schemas (X like X, X wants X, X sees X…) are recognised as instances of a more general syntactic pattern (SVO) in which the linguistic items play similar communicate roles, what Tomasello has called functionally based distributional analysis (Tomasello, 2003: p 169).

2.4. Type/Token Distribution.

Figure 5. Contrasting type/token distributions in two schematic networks.
Chapter 6 most clearly attempted to combine the ideas of type and token frequency and examine the shape of the distribution of instances that are categorised by a schema. Skewed distributions have been shown to have a beneficial role in learning linguistic and non-linguistic categories (Casenhiser & Goldberg, 2005; Genter, Loewenstein, and Hung; Avarahami et al., 1997; Elio and Anderson, 1984). Similar skewed distributions have been reported in Child Directed Speech for English (Naigles & Hoff-Ginsberg, 1995; Theakston, Lieven, Pine & Rowland, 2004) and for Hebrew (Ninio, 1999a; 1999b). For example, young children hear exemplars of the ditransitive construction most often with the verb ‘give’ (left in Figure 5) and so, in an important sense, acquiring the meaning of the verb give and acquiring the ditransitive construction are part of the same process. If one verb accounts for the ‘lion’s share’ of the tokens, the hypothesis is simply that this will have a result on the structure of that schematic hierarchy. Chapter 6 examines how the skewed distribution might actually help categorisation in the process of language acquisition by proposing a model of schematisation under skewed input conditions.
2.5. Schema complexity.

Figure 6. Contrasting schema complexity in two schematic networks.

Schema complexity refers to the problem of pattern detection introduced in Chapter 2. The idea is that the patterns infants can detect is dependent on the character of what they are trying to learn, in short, the more abstract the pattern or correlation that one needs to detect, the more evidence one needs. As a function of the within- and between-category variance (the number of sub-schemas in the diagram above), and as the corpus of evidence grows, the opportunity to abstract regularities arises. The hypothesis is that the ‘depth’ of the abstraction, represented in construction grammar as the taxonomic hierarchy, is proportional to the size of the corpus of utterances. The usage-based prediction for learning is that the more abstract the schema the more instances one needs to construct it. This is an idealisation based on other things being equal, which of course they are not – the analogical mapping from other schemas (§2.7), for example, is one factor which further complicates the picture when trying to predict learnability.
For example, in Chapter 5 we discussed the significance of high-frequency pronouns, such as *I*, *he* and *it*, which occur with regularity in certain utterance positions as ‘pronoun islands’ in child directed speech. Other work also suggests the importance of pronouns in anchoring children’s early transitive constructions (Childers & Tomasello, 2001 for actives; Dodson & Tomasello, 1998; Lieven, Pine, & Baldwin, 1997). In terms of schema complexity and pattern detection, this means that early schemas anchored around lexical items, e.g., *she/he X-ing him/her* emerge before we see a fully productive case-marked schema e.g. Subject\textsubscript{nominative} – Verb\textsubscript{-ing} – Object\textsubscript{accusative}. In short, because nominative/accusative are more abstract notions than lexical items one needs more evidence (linguistic instances) to abstract that category. This fits with the usage-based approach which emphasises that infants’ early generalisations are more local and lexically-based, and suggests that linguistic constructions build up during ontogeny in a piecemeal, more concrete fashion, beginning with fewer and weaker abstractions.

2.6. Reliability.

Figure 7. *Contrasting reliability in two schematic networks.*
Reliability refers to how consistently a given form is associated with a given function and vice versa, and was introduced in the context of the cue-competition model in Chapter 5 (Bates & MacWhinney, 1985). All other things being equal the less reliable a form is the more difficult it is to learn, because the more abstract the pattern is to detect. Pronoun-frames (Chapter 5), statistically-speaking, provide pockets of predictability between form and function – reliable islands that infants can use to decrease the degrees of freedom when trying to interpret what a sentence means. The suggestion is that pronouns, such as I, me, he and it, may have a special role to play because they are highly frequent and, in many instances, have different forms for subjects (he/she preverbally) and objects (him/her postverbally). Even where a pronoun is not case-marked it can occur in one role much more often than the other (e.g. it as the object of a transitive and as the subject of an intransitive).

While still being a valuable concept I think we must expand the cue-competition notion of reliability, from ‘(linguistic) form-function’ to include ‘referential context, (linguistic) form-function’. The suggestion is that calculating a reliability figure for a particular form-function pairing paper-and-pen style from a corpus is missing something from the infant’s perspective. Specifically, as social agents grounded in a communicative context, calculating reliability this way is not the same problem children have to solve. Although the mapping between these different representations is, for want of a better word, messy, it is not without regularities either, allowing the infant to triangulate reliabilities between referential context, linguistic form and semantic content (see Chapter 2 for a discussion of these ideas).
2.7. Analogy and categorisation.

Figure 8. *Analogical mapping between items and schemas in two networks.*

This a central issue of this thesis, in a usage-based approach to language acquisition, what processes guide generalisations? As discussed in Chapter 2, there is good evidence that the alignment of relational structure (schematically shown between two networks above) and mapping between representations is a fundamental psychological process underlying analogy and similarity (Genter & Markman, 1993; 1994; 1995; Goldstone 1994; Goldstone & Medin 1994; Goldstone, Medin & Genter, 1991). One manifestation of this domain-general process allows inter-construction mappings. In chapter 2 I tried to show for one construction – the caused motion construction – the relationship between how events are construed in mental space (the conceptualisation) and how they are construed in perceptual space (the perception). The suggestion was that the embodied form of the event representation or construal is redescribed forms of the perceptual scenes (the starting point for utterance-event associations) tagged with the psychological characteristics of the actors involved. Thus abstract grammatical patterns are conventionalised patterns of shared experience; patterns of experiences that are organised on the basis of how the participants behave, whether their actions are *intended* to accomplish something
similar. Thus right from the start, wider knowledge of how people work is brought to bear on the schematisation process. By accumulating a history of these usage events – which are temporal associations of multimodal information made within the area of cooperative communication and reasoning – the infant has begun to learn the particular functional distributions of utterance-event co-occurrences in her language.

Chapter 2 suggested that categorisation, analogy and similarity are actually different manifestations of the same underlying cognitive process: pattern finding. The hypothesis was that there is no principled cognitive distinction between these processes assuming that (linguistic) knowledge is organised in something like a structured inventory of schema-instance relationships. In short, in analogy, if X is like Y in a way Z, then Z can be considered schematic of X and Y and if Z is schematic of X and Y then these have been treated as instances of the category Z. So if categorisation and analogy are the cognitive mechanisms driving the generalisations, then what these generalisations look like in development will be determined by the characteristics of the input (token frequency, type frequency, type/token frequency, reliability, schema complexity). In other words pattern detection is driving generalisations, properties of the input determine how difficult that pattern is to detect.

While this thesis has confirmed that many of these cognitive properties and properties of the input are important parameters for learning there is clearly much to discover about how these interact in a dynamic system. The hope is that conceptualising the problem in this way provides a framework for altering one parameter at time – perhaps as implemented as a learning algorithm in a computer model – and observing what the implications are for learnability for a particular form-function pairing.
3. **Strengths and weaknesses of the approach.**

Based on the experience of writing this thesis, what follows are some brief observations about what a usage-based approach to language acquisition does well and what it does less well.

*What it does well.* Firstly, because the emphasis in a usage-based approach is on the input as the raw material out of which children build the language (not just the trigger for innate linguistic representations) it naturally accommodates cross-linguistic differences in developmental trajectories of learning. For example, the studies considered in the cue-competition model (introduction to Chapter 5) and the cross-linguistic evidence from the role of prototype constructions in early language acquisition (Chapter 3) show that there are close links between the way adults use particular words, morphemes, and phrases in child directed speech and the way children learn them. The acquisition of particular linguistic structures is heavily dependent on the specific language to which a particular child is exposed, and with patterns of early grammatical productivity in the child showing a close correlation with the patterns of use they hear (the model results of Chapter 6 also add further evidence). The second strength follows from the ‘cognitive commitment’ – the general thrust of the cognitive linguistics enterprise is to render accounts of syntax, morphology, phonology and word meaning consonant with aspects of cognition which are well documented, or at least highly plausible, and which may manifest in non-linguistic activities (Taylor, 2002). The usage-based approach has been generally successful in seeing the links between cognitive and linguistic phenomena; a trend I have tried to continue and develop, for example, describing and explaining argument-structure construction categories in terms domain-general pattern finding abilities (Chapter 4). The belief is that such an approach will enable the linguist to go beyond a
mere description and formalisation of the linguistic facts, and to arrive at a more
insightful explanation of the facts.

What it does less well. Firstly, and perhaps unsurprisingly, some of the
strengths are also related to the weaknesses; being able to describe a large range of
cross-linguistic findings means that stipulating in any detail, \textit{a priori}, which
processing constraints and input properties are most important, and when, is very
difficult (see difficulties in explaining some of the Dittmar. et al. results, Chapter 3).
In such situations explanations begin to have post hoc feel and are in danger of
becoming unfalsifiable. Popper (1934) states that confirmations of a particular theory
should count only if they are the result of risky predictions, a ‘good’ scientific theory
in a Popperian sense is a prohibition: it forbids certain things to happen. The more a
theory forbids, the better it is. For example, what pattern of results in the pronoun-
frame experiment (Chapter 5) wouldn’t be explainable on a usage-based account? In
truth, this situation was probably not made any better by the number of crossed
conditions, making second- and third-order interactions between factors especially
difficult to predict. However, in defense of this approach one could argue in the spirit
of ‘blind empiricism’ that because the results were so difficult to predict is exactly
why we were so interested in running the experiment in the first place. Also, if we
only looked at one aspect in isolation, say case-marking only in active sentences
rather than case-marking in both active and passive, this would be less like the input
children actually hear (a trade-off in all studies between control and ecological
validity). The hope is, I think, that once enough of these types of experiments have
been conducted, and the results synthesised, usage-based theory will evolve into
something with tighter ‘riskier’ predictions in the Popperian sense and with
experiments where there is more theoretically at stake. This would also go a long way
to engaging those from different theoretical approaches to language acquisition who are skeptical of the usage-based approach for this reason.

Secondly, there is a lot of work on infants’ development of social cognition and pragmatic reasoning, but it has yet to be worked through in any detail how this knowledge interacts with and/or constrains syntactic representations and generalisations. In Chapter 2 I suggested that the embodied form of the event representation or construal is a redescribed form of the perceptual scenes (the starting point for utterance-event associations) and is tagged with the psychological characteristics of the actors involved. This is an attempt to integrate social reasoning (knowledge of how people think and behave) into the syntactic representation (the event construal) in order to constrain and generate generalisations. To revisit an example from Chapter 1:

(1a) She painted the green house.
(2a) She painted the house green.
(3a) She knew the green house.
(4a) *She knew the house green.

On the basis of (1) and (2) one might correctly generalise (3), but overgeneralise to incorrectly infer (4). If we look at each sentence and examine their meanings, we can paraphrase them as:

(1b) She painted a house that happened to be green.
(2b) She caused the house to become green by painting.
(3b) She knew a house that happened to be green.

(4b) She caused the house to become green by knowing.

The formal generalisation 4a is not entertained because the encyclopaedic knowledge of how people behave and what they are capable of, rules out the proposition 4b. Thus the syntactic and pragmatic representations interface in complex ways that are not fully understood. It seems important then that usage-based approaches start integrating the wealth of experimental data that has been produced in recent years on social cognition with the syntactic learning literature, for example how a developing understanding of shared intentionality, built on a foundation of recursive mind-reading, interacts with and helps to constrain and generate the learning of statistical distributions in the language.

4. Future Directions.

1. As just mentioned, one of the main challenges to any usage-based approach is to specify in greater detail the learning mechanisms that underpin generalisations. Part of this involves integrating the body of research on social cognition into the statistical learning literature. If usage-based theories are to evolve a level of descriptive and predictive adequacy, they need to address the issue of falsifiability by making more precise conjectures about what findings would and would not be expected under the theory. One kind of approach that seems particularly successful in generating this kind of rigour is combining corpus and experimental methods to answer a focused question about learning (e.g., Bannard and Matthews, 2010; Dąbrowska, 2008; Goldberg, Casenhisser, Sethuramen, 2004). In this approach one can use the distributions in the input as the starting point to generate predictions, and then through
experimental manipulation, get closer to establishing some kind of cause-and-effect. Related to this is the need to build up more of a developmental dimension. Those working in non-usage based theory can become frustrated when usage-based theorists argue lexical specificity supports the theory, but then are also not worried by evidence of abstraction, since the child’s knowledge is predicted to become abstract at some point in development anyway. Clear developmental predictions about how the process of abstraction should develop, including which systems should become abstract first, are needed. This would also help counter claims that it’s a discontinuous theory (lexical specificity first, abstraction later) – which it isn’t: each system is predicted to develop from lexical specificity to abstraction gradually, with some systems developing faster than others.

2. Children’s errors are the one area in which usage-based theory does make clear predictions: errors in children’s data are predicted to occur when children start to go beyond tried and tested lexically specific formulae, and should occur on low frequency items. However, there is evidence that this isn’t the whole story, for example Ambridge & Rowland (2009) find that children make more errors in negative questions than we might expect given their input frequency. Plus there may be areas where we might expect more errors to occur but which children seem to master pretty quickly. This is an area where usage-based theories can distinguish themselves from other theories, but so far children’s errors have probably been under-researched and have a lot more potential than has been realised.

3. There are still surprisingly few studies that have attempted to characterise the input children hear in dense enough detail and at the level of linguistic representation that is plausible in a usage-based theory for infants (Cameron-Faulkner et al., 2003, being a notable exception). Studies of this type are essential if we are to achieve a
better grip on the frequencies and distributions of items that will give usage-based theories more chance of predicting in detail the specific developmental trajectories of linguistic items.

4. One final speculation on the future of usage-based approaches to language acquisition. Taking the ‘developmental cognitive linguistics’ enterprise to its logical conclusion, one of the big questions is whether there are any purely linguistic representations or can all linguistic categories be traced back or decomposed into the functional or communicative roles they play. This is not, as some critics might say, ‘reducing’ language to cognition any more than biology is being reduced to chemistry or chemistry is being reduced to physics. What it does mean is that it is a cognitive-science approach that is also concerned with the social foundations of human communication; a combination which is perversely rare in language research. This framework for studying the acquisition of syntax has come a long way from the generative-grammar enterprise, but is perhaps no less radical: it suggests linguistic categories are artefacts that emerge from a handful of cognitive skills (e.g., perception, attention, pattern-finding, memory) interacting in complex ways with a species-specific set of social skills (e.g., shared intentionality, cooperative reasoning, cultural intelligence).
References


Carroll, L. (1872). Through the looking glass, and what Alice found there. Penguin.


305


