AN INVESTIGATION OF THE USE OF CASE MARKED PRONOUNS IN ENGLISH SPEAKING CHILDREN

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

2015

Stacey Mcknight

School of Psychological Sciences
LIST OF CONTENTS

Chapter 1: Introduction to Child Language Research 11
  1.1 The Wonder that is Language 11
    1.1.1 The Generativist approach 12
    1.1.2 The Constructivist approach 13
    1.1.3 Thesis overview 15
  1.2 Language Acquisition 16
    1.2.1 Variation in acquisition 16
    1.2.2 The Competition model 17
    1.2.3 Form-Function mapping 20
    1.2.4 Abstraction 21
  1.3 Children’s Errors 23
    1.3.1 Frequency 24
    1.3.2 Strings in the input 28
    1.3.3 The rate of error 32
  1.4 Pronoun Case Marking Errors 33
    1.4.1 Case marking 33
    1.4.2 Pronoun case marking errors 34
    1.4.3 Agreement/tense omission model 35
    1.4.4 Paradigm building model 37
    1.4.5 Phonological properties 39
    1.4.6 Input base explanations of pronoun case marking errors 40
  1.5 My-For-I Errors 42
    1.5.1 Pronoun modelling 42
    1.5.2 Functional approach 44
  1.6 The Current Thesis 46

Chapter 2: Children’s My-for-I Errors: Analysing the Input 49
  2.1 Abstract 51
  2.2 Introduction 52
    2.2.1 The current study 58
  2.3 Method 59
    2.3.1 Corpora 59
Chapter 4: Children’s Pronoun Use and Protest in a Conflict Over Agency

4.1 Abstract

4.2 Introduction

4.2.1 The current study

4.3 Method

4.3.1 Participants

4.3.2 Materials

4.3.2.1 Puppet

4.3.2.2 Game

4.3.2.3 Stories

4.3.2.4 Parent questionnaire

4.3.2.5 Sticker sheet

4.3.3 Design

4.3.4 Procedure

4.3.4.1 Warm-up and training

4.3.4.2 Priming

4.3.4.3 First condition

4.3.4.4 High conflict condition

4.3.4.5 Low conflict condition

4.3.4.6 Priming

4.3.4.7 Second condition

4.4 Study One

4.4.1 Pronoun and proper name coding

4.5 Results

4.6 Discussion

4.7 Study Two

4.7.1 Protest coding

4.7.1.1 Normative protest

4.7.1.2 Imperative protest

4.7.1.3 Agreement

4.7.2 Action coding

4.7.2.1 Preventative action

4.7.2.2 Active action

4.7.2.3 Corrective action

4.7.2.4 Agree

4.8 Results

4.9 Discussion

4.10 General Discussion

4.11 References

4.12 Appendices
Chapter 5: Discussion and Conclusions

5.1 Summary of Findings

5.1.1 The relation of input to errors

5.1.1.1 Frequency and strings in the input

5.1.1.2 Proportional frequency of different forms in the input

5.1.2 Children’s prior use of forms

5.1.3 The role of function

5.1.4 Errors in protest

5.1.5 Types of protest

5.2 How Our Results Fit with Previous Literature from the Constructivist Perspective

5.2.1 Input

5.2.2 Function

5.2.3 The competition model

5.2.4 The role of function in other pronoun case marking errors

5.2.5 Children’s ability to protest

5.2.6 Multifaceted explanation of children’s errors

5.3 Limitations of the Present Studies and Suggestions for Future Research

5.3.1 Corpus analysis

5.3.2 Experimental studies

5.3.3 Future research

5.4 Conclusion

Appendices

References

Word count: 55,620
LIST OF TABLES

Table 1.1 Pronoun forms for the Nominative, Accusative and Genitive case. 33
Table 2.1. Defined my-error period for all sixteen children 63
Table 2.2. Total number of my, me and I utterances during the Defined error period, including number of each different error type. 65
Table 2.3. All fifteen children’s defined input files. 66
Table 2.4. Number of each first person pronoun uttered by caregivers in the defined input files. 67
Table 2.5. Total number of am I and specifically am I-verb utterances spoken in the defined input files. 69
Table 2.6. Proportion of mother’s my-utterances from the defined Input files that contain nouns that can also be used as verb (father’s utterances indicated in parenthesis). 70
Table 3.1. Defined my-error period for the fifteen children with total number of my-for-I errors, me-for-I errors and correct I-verb utterances during this period. 95
Table 3.2. Number of first person pronouns in the defined input files. 96
Table 3.3. Error period for the children in study two. 108
Table 3.4. Error period for the four own-name children. 114
Table 3.5. Error period for the four me-errors children. 114
Table 3.6. Mixed Effects Models for me-errors and own-name data. 115
Table 4.1. Total pronoun or own name and verb use for the two age groups. 143
Table 4.2. F values, significance level and partial $\eta^2$ for the comparison between the two year old and three year old children for each reference type. 145
Table 4.3. Total self-reference and reference to Cat in both age groups in both conditions. 146
Table 4.4. F values, significance level and partial $\eta^2$ for the Comparison between the high-conflict and low-conflict conditions for each reference type. 147
Table 4.5. Number of each protest types for both the two-year-old and three-year-olds in both the High-Conflict and Low-Conflict conditions. 157
Table 4.6. Number of each response type for both the two-year-olds and three-year-olds in the High-Conflict condition. 158
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Scatterplot of the fifteen children’s my-error frequency in relation to the <em>my</em>, <em>am-I</em> and <em>my-noun/verb</em> input frequency.</td>
</tr>
<tr>
<td>3.1</td>
<td>Example target utterance with context. This particular utterance is one which the prior ten utterances do not give enough information to understand the context, so the analysis was taken back to twenty utterances.</td>
</tr>
<tr>
<td>3.2</td>
<td>An example target interaction between mother and child in which the child is attempting to bring about a change in the agent of an action.</td>
</tr>
<tr>
<td>3.3</td>
<td>The raw number of <em>my-error utterances</em> and <em>I-utterances</em> that functioned to attempt to gain agency or control (out of a possible 60, 53 for Ruth) (predictors 1, 2 or 3).</td>
</tr>
<tr>
<td>4.1</td>
<td>Number of Each Protest Type in total for all children combined, across the Nine Rule Violations in the High-Conflict Condition.</td>
</tr>
<tr>
<td>5.1</td>
<td>A Multifaceted Explanation of Children’s <em>my-for-I errors</em>. A model of the several factors all combining to increase the chance of error.</td>
</tr>
</tbody>
</table>
ABSTRACT

Case marked pronouns in English have received an increasing amount of attention from child language researchers from a Constructivist or Usage-based perspective (Tomasello, 2000; 2003; Kirjavainen et al., 2009) and also from a Generativist perspective (Schutze & Wexler, 2000; Rispoli 1998). However, genitive-for-nominative errors specifically, have received less attention.

A detailed Corpus Analysis investigated English-speaking children’s my-for-I pronoun case marking errors. The results revealed no direct link between frequency of my in the input either independently or as part of a larger construction and children’s my-for-I error rates. However, the modelling of the pronominal system was found to be related to children’s error rates. Children receiving mostly pronouns in the input, in contrast to a high proportion of proper name replacement, were found to have a lower rate of errors. Further, children with a less entrenched use of I, before their error period, were those who made a higher my-for-I error rate. Further analysis revealed an association between the function of agency, control and possession in children’s erroneous uses of my as subject in verb phrases.

An experimental study with two and a half and three and half year old children found that a range of pronoun case marking errors were being made alongside the correct use of I in both age groups. No age differences were found for the rate of correct I-verb utterances. This shows that children do not go through a period of time 100% making a certain error type and then change to 100% correct I, but that actually both forms are accessible and will compete for use. Within the study the claiming of agency and control was in focus, which elicited a high level of protest from the children. Analysis revealed that the children as young as two and a half were capable of protesting normatively, making use of normative language to show their awareness of the game rules and how things “should” be done. Further, developmental differences showed that the younger children may not be as sophisticated as the older children in altering their protest, according to context. This development trend shows that children may be learning more linguistic forms, but their understanding of when it is necessary to use them, may not yet be as advanced, in two year old children.

This current thesis supports a Constructivist or Usage-based perspective of language acquisition. The importance of input, function and competition between forms has been exemplified. If children assign a form-function mapping between agency, control and possession and the form my, they are likely to make a higher level of error in these contexts. This error rate is then increased if the modelling of the pronominal system is low and not offering a high level of correct use of my and I, to compete with this erroneous mapping. This lack of modelling plus a lower entrenched use of I from the outset, will mean this competition will therefore be won out more often and for longer by this incorrect my form.
DECLARATION

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

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 she 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, trademarks 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://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487), 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.
ACKNOWLEDGEMENTS

I would like to thank my supervisors, Anna Theakston and Elena Lieven for their support and guidance throughout my doctorate research and also the thesis writing process. I would like to thank Anna Coates for her help in recruitment and for her excellent role as ‘Cat’. A big thank you to Mickie Glover, for her help with recruitment and for all of the support in the lab. Also a special thanks goes to Knitting Club, Henna Lemetyinen, Rachel Ashworth, Anna Coates and Helen Allwood for their continued support and for making the journey such a fun one. Also thanks to Ryan Blything for thought provoking discussions and to Laura Núñez for her role as second coder and for being so patient in the coding stages. A huge thank you to Nichola Grogan for all of her support throughout the whole four years, for putting up with me through my moments of stress and for always understanding. To Bobby, for sitting beside my computer for days on end with me and keeping me company. Also to my parents and brother for their encouragement and support. Finally I would like to thank all of the families who have taken part and took time to come in to the Max Planck Child Study Centre and for the nurseries who allowed me to conduct my research. Further, to the number of people who have made corpus data available on CHILDES. Without their commitment I could not have conducted the research reported in my thesis.

The research in this thesis was supported by a studentship from the Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany.
Chapter 1: Introduction to child language acquisition research

The purpose of the introduction is to outline the importance of investigating children’s errors when researching the process of typical language acquisition and to justify the constructivist approach taken in this thesis. There will then be an overview of the literature with a focus on the past literature surrounding children’s errors and in particular, children’s pronoun case marking errors.

1.1 THE WONDER THAT IS LANGUAGE

The human ability to speak a language is unique and differs dramatically from the communication systems of other animal species. Most human communication is symbolic and therefore can be arbitrary. We pattern these linguistic symbols together into structures which then have their own meaning. This is what we call grammar. Language develops in every typically developing child without formal instruction and with apparent ease. This language learning process is of upmost importance and relates to children’s ability to read and write and levels of education. The learning of one’s native language could almost appear to be a straightforward and simple process. We start by expressing our first words at around twelve months, but in what seems like no time at all, children have an array of vocabulary and have also learnt many syntactic features of their language. By twenty four months children have acquired on average 200 – 300 words and by thirty six months are understanding and using simple questions, can use negation and are understanding and using first, second and third person pronouns. But just exactly how do we master the adult system of our native language? This is a question that has been investigated in great depth for many years and which has driven a great deal of empirical research. How we acquire such a huge collection of
vocabulary and such complex and varied syntax has been the subject of great theoretical debate between two major theoretical standpoints, namely the Constructivists or Usage-Based (UB) approach and the Generativist approach. There are many variations and sub-theories to both the Constructivist and the Generativist approach, but the main tenets of their frameworks, to which most theorists belonging to the respective paradigms subscribe, will now be highlighted.

1.1.1 The Generativist Approach

The Generativist approach assumes that children are born with some language knowledge and that this knowledge will help them acquire grammar. This innate knowledge is said to consist of formal rules that will operate on abstract linguistic categories, such as VERB and NOUN, as well as phrases such as NOUN PHRASE. This knowledge is held to be part of a Universal Grammar (a general grammar that applies to all the world’s languages), which is believed to be innate and biologically programmed in the human brain (Lenneberg, 1967) and therefore children have a knowledge of this Universal Grammar (UG) from birth. It is debated within the Generativist framework, what exactly the innate knowledge is, from just grammatical information (Radford, 1990; Wexler, 1998) to semantic knowledge also (Pinker, 1984). The linking problem (Pinker, 1984) has noted that UG would have to be extremely general in order to link to all languages of the world. So general in fact that the task would be as difficult as learning the language without this innate knowledge. Therefore, Pinker (1984) suggests that children also have innate linking rules that link the syntax and semantics.

The principles and parameters theory of the Nativist perspective attempts to explain how children acquiring their native language use UG. Universal innate rules
(principles) are claimed to combine with the child setting innate parameters according the input of their native language (Radford, 1990; 2004). The parameters are usually thought to be binary, children have to set these parameters as positive or negative based on the language they hear. For example, whether sentence subjects are compulsory (positive for English children as subject is required). This role of input is important for children to learning their individual native language, however, as the parameters are innate, these just need to be set and therefore this is an easy process, which is seen to be exemplified by the rapid pace of language learning. Legate and Yang’s (2007) Variational Learning Model is a generativist model that attempts to explain the function of input in children’s language development. The child is said to converge on the correct grammar and the correct parameter settings and abandon other possibilities, based on the linguistic environment. The child will try different settings in their speech and the input data will go towards the strengthening of one of the parameter settings, until one of the settings ‘wins out’. The settings that don’t win out will eventually no longer be used. This role of the input differs to that of Constructivist approaches, discussed below, in that the input strengthens or weakens each parameter setting. These parameters are already present and just waiting to be set, to match that of the input (Yang, 2002; 2004).

1.1.2 The Constructivist Approach

In contrast to the generativist approach, the constructivist or usage-based approaches do not assume that children have an innate knowledge of grammar or are born with grammatical categories such as VERB and NOUN, but rather that they must acquire these by generalizing across the speech that they hear. Therefore,
the linguistic environment of the child is crucial in most constructivist approaches, in that the input that children hear is closely related to the language acquisition process. Constructivists would therefore often predict that the most frequently encountered words will be most easily acquired because of this repeated exposure to these particular words. Ambridge, Kidd, Rowland & Theakston (2015) provide evidence for the importance of frequency and show that frequency effects are inescapable when studying children’s first language acquisition. Ambridge, et al. (2015) acknowledge that frequency is not everything, for example by pointing out that when considering simple token frequency, “the” would be much higher-frequency than “cake”, however, children’s desire to obtain cake may influence their effort in storing this word over the frequency of hearing the word. Therefore a frequency driven mechanism is not argued for, but a mechanism that is frequency sensitive. It is also pointed out that there are many different types of frequency effects, as well as the fact that high-frequency forms can both prevent and cause errors, so there is by no means a straightforward relationship between frequency of specific lexical items and their acquisition. Finally Ambridge, Kidd, Rowland & Theakston (2015) show that frequency effects in word learning often interact with other factors such as serial position and utterance length.

The Constructivist perspective does not believe that children have an innate knowledge of grammar. However, the ability to learn language is considered to be innate. This ability relies on the human cognitive system, which is the general cognitive system that is responsible for many types of learning, other than language. The general cognitive processes such as the ability to categorise, memorise, imitate and so on are considered to be innate and these are the mechanisms which give us the ability to learn language (Tomasello, 2003). Once
we have the mechanisms in place to learn a language, we need to have motivation to learn it. A usage-based or functional approach to language acquisition believes that the driving force behind this language learning process is the child’s desire to use language. This usage-based (UB) approach forms a large part of the constructivist perspective and sees language as a communicative function. A child’s wish to communicate with others, for example to request an object, to disagree with an activity or to express a need, will be the main driving force behind learning your native language (Ambridge & Lieven, 2015). These two proposals are aligned and the current thesis takes a constructivist/usage-based position on children’s language learning. For the purposes of clarity, the broader term ‘Constructivist approach’ will be used throughout this thesis as this includes the broader scope of literature that will be covered and it is this broader perspective that the current thesis will sit. However, the term usage-based or functionalist approach may be used when there is a specific focus on the child’s own language use or the function of language being spoken, under focus.

1.1.3 Thesis Overview

Research into child language acquisition needs to not only explain how children master the adult system of language, but how this process begins as well as explaining how, along the way, children’s language often differs from the native adult language and many different types of errors are made. Whether children’s language acquisition is based on innate linguistic structures or built from their linguistic experience and cognitive skills is still debated today. However, this current thesis does not aim to adjudicate between the two major theoretical standpoints, but aims to work within the Constructivist stance on language
acquisition. Therefore the current thesis will take this Constructivist stance on children’s language acquisition and aims to elaborate on explanations within this framework, in investigating children’s pronoun case marking errors. The current chapter will introduce the Constructivist literature surrounding a range of children’s errors and then introduce both Generativist and Constructivist theories of pronoun case marking errors, which is the focus of this thesis. Chapter two contains paper one, which is a detailed corpus study investigating the input in relation to children’s my-for-I pronoun case marking errors. Chapter three contains paper two, which includes two studies. First, an overview of the corpus study contained in paper one (for publication purposes) plus additional analyses surrounding input factors and second, a functional analysis of children’s my-for-I pronoun case marking errors. Chapter four contains paper three which also contains two studies. The first looks experimentally at children’s pronoun use in differing contexts. The second focuses on children’s protest in these different contexts. Chapter five is the final chapter and will discuss the results overall in relation to past research and discuss what this tells us about children’s language learning.

1.2 LANGUAGE ACQUISITION

1.2.1 Variation in acquisition

The process of language acquisition is by no means identical for any two children. From looking at the first 50 words acquired by 18 children (Nelson, 1973), children varied on a continuum, from being highly referential and focussed a lot on object naming or highly expressive, focussing on the social uses of language. Bloom, Lightbown and Hood (1975) noted variation on a grammatical level, at the very first stages of grammatical development. Bloom et al. (1975) saw a
nominal/pronominal dimension when studying four children and found that two referred to themselves by name e.g. “Kathryn sock” and two referred to themselves by first person pronouns, e.g. “I finish”. This contrast was short lived, as by the time the children had reached an MLU of 2.5, this distinction had gone. These distinctions may be short lived, but have been shown to link to other dimensions of variation throughout later stages of language acquisition (Bates & Macwhinney, 1987) Bates and MacWhinney (1987) highlighted connections between referential or expressive style in the early, one word stages and the nominal or morphological contrasts in multiword speech. Horgan (1981) further claimed that the nominal or referential children are “noun lovers” whereas pronominal or expressive children are more “noun leavers” in later stages of language learning. A model of language learning that attempts to address this variation in language learning is the Competition Model (MacWhinney & Bates, 1989; MacWhinney, 2004).

1.2.2 The Competition Model

The Competition Model (MacWhinney & Bates, 1989; MacWhinney, 2004) is a model that attempts to provide the mechanisms for understanding variability in words, segments and constructions across languages and how this then impacts on children’s learning of these different aspects of language. As well as this, the model also attempts to address individual variation within a single language and how this can then impact children’s learning of their own native language. The statistical properties of the input play a major role and because of this the model can be applied to both cross-linguistic variation and also the variation between individual language learners. The competition model comes from a functionalist approach to language acquisition. This approach holds that forms in languages are created,
acquired and used to serve communicative functions. The model focuses on the mapping between form and function in language production, comprehension and acquisition. The model is based on two levels; the functional level, representing all the meanings and intentions of an utterance and the formal level, representing the devices available (forms) in that particular language. The mapping is assumed to be direct, but usually several forms can map on to many functions. This means that analysing which cues people rely on is a difficult task, especially as the different cues then vary in weight across languages, for example word order is more reliable in English than in many other languages. Under this model, when interpreting a sentence, children will rely on what the model calls cue validity. A cue’s validity is the function of its availability and reliability. Availability “represents the extent to which a cue is there when you need it” (MacWhinney and Bates, 1989: 41) and this is defined by how often it is present over all exemplars of that particular task domain. An example in English, would be to identify the AGENT and PATIENT in a NOUN VERB NOUN utterance. The word order cue is 100 per cent available as there will always be a word order when there are two nouns and a verb. Reliability “represent the extent to which a cue leads to the correct interpretation when you count on it” (MacWhinney and Bates, 1989: 41). Reliability is defined by the amount of time the cue correctly indicates the classification such as when the cue correctly indicates the agent or patient when it is present. For example in English transitive sentences, the first noun is almost always the agent/subject of the sentence and therefore this cue is reliable. This is less true in languages with a more variable word order.

The interesting question is how children use these factors to develop a correct and adult like cue use of their language. Many act-out studies have shown
that while adults almost always rely on the most valid cues when acting out a sentence, young children do not. It seems early in development, children can over rely on highly available cues and rely less on more reliable, but less available cues. Dittmar, Abbot-Smith, Lieven and Tomasello (2008) showed, in an act-out study using novel verbs, that German children aged 7 years relied on case marking, whereas children aged 2 years and 4 years, relied on the less reliable word order, when these two cues were placed in conflict. A factor that has been seen to affect children’s performance in these studies is cue cost (MacWhinney and Bates, 1989). This relates to how difficult a cue is to use. For example subject-verb agreement may be a highly reliable and available cue, however, storing the subject, whilst waiting for an agreeing verb form may be difficult for young children and therefore cause them to have a less sophisticated grasp of cue use.

From a usage-based perspective, the use of language is to communicate a certain function. In children’s search for the correct form, to communicate their desired function, errors can be made. In looking at the variation between individual language learners, we can attempt to uncover what it is that causes these differences. In particular in studying errors, this variation can link back to the input that children have received, as well as this, other factors that have caused them to choose one particular form over another, when expressing their communicative function. The Competition Model (MacWhinney & Bates, 1989; MacWhinney, 2004) can attempt to explain what factors cause children to choose this one particular form over another when they are learning their language. How there may be forms competing to be used to express a certain function; the competition between a particular frequently heard form in the input and a form that has been
associated by that child with specific function already (whether this assignment is correct or not).

1.2.3 *Form-Function Mapping*

MacWhinney and Bates’ (1989) Competition Model and the theory that the frequency of a form interacts with its consistency to affect the speed of learning, was exemplified by a detailed corpus analysis (Cameron-Faulkner, Lieven & Theakston, 2007). Cameron-Faulkner et al. studied the development of one child’s negative utterances. Initially all of the child’s negative utterances involved “no+verb”, which reflected the fact that “no” was heard frequently in the mother’s speech and mostly as a single word utterance. The child soon started to produce “not+verb”. This reflected the mother’s second most frequent use of negation, but this was often heard in multiword utterances. It was claimed that the child started to use “no” first in multiword negative utterances, as they would have been using it in a single utterance because of the frequency of use in the mother’s speech. Soon, “can’t+verb” and “don’t+verb” started to appear in the child’s speech, and eventually other “aux+n’t+verb” constructions began to appear. Cameron-Faulkner et al. (2007) claimed that the speed that the child progressed through the different forms was related to the range of negators that the mother was using for any particular function. In other words, if a function was consistently related to one form, the child could map this and learn this earlier, however, if the function can map on to several forms, this may take longer for the child to learn and they will rely on the trustworthy “no+verb” or “not+verb”. This relates back to the notion of competition (MacWhinney & Bates, 1989; MacWhinney, 2004) for the child, between the forms heard in the input in similar utterances such as “can’t” and
“don’t” and the already strongly engrained form-function mappings between the trusty “no” and “not” forms and a similar function.

1.2.4 Abstraction

A central tenet of the usage-based approach is the idea that children move beyond learning that is lexically specific to construct more abstract schemas (Dąbrowska, 2000). This then allows children to produce utterances which will not have been heard in the input, in any form. The constructivist approach argues that children’s earliest linguistic representations are not adult like categories, but rote-learned “holophrases”, for example I+WANT+IT and also low-level, lexically specific slot and frame patterns such as I’M (X)ING IT (Tomasello, 1992; Pine & Lieven, 1997; Pine, Lieven & Rowland, 1998; Ambridge & Lieven, 2015). These ‘constant – variable’ patterns (Braine, 1976) are limited and involve individual words in a single position and these are then combined with a variety of other words. It has been argued that children rely on these low-level templates for a lot longer than previously thought (Dąbrowska, 2000) and that these may still be represented in the adult system. It is a very slow and gradual process that involves children abstracting across these to arrive at adult-like and more abstract, schematic constructions.

Children are not consciously trying to learn language, they are using it, as a communicative tool and in this use they build abstract linguistic categories. According to the constructivist approach, children start out with these constant-variable patterns, or holophrases. These develop first into lexical schemas and then into adult-like abstract constructions. This process can begin as soon as children have two stored exemplars over which to abstract (Ambridge & Lieven, 2015). An
example is to look at English determiners, ‘the’ and ‘a(n)’. Once children have formed the lexically specific slot-and-frame schema ‘a [X]’ and ‘the [Y]’ they can then produce determiner+noun combinations that haven’t been heard in the input. As evidence that children begin with rote learned strings, Pine, Freudenthal, Krajewski and Gobet (2013) found that children had a significantly lower overlap of nouns used with both ‘the’ and ‘a’ than their caregivers. This shows that these determiner and noun categories and knowledge of how to combine them cannot be present from birth, as some generativist theories would suggest (Valian, 1986). The suggestion is that the children go through a stage where they only use the rote learned form and therefore some nouns have only been stored with ‘a’ and others with only ‘the’. Therefore before the productive schema has been formed, they will not produce a combination not heard in the input. Pine and Lieven (1997) have also found that children’s early determiner use shows lexical specificity such as “that’s a X”, “Where’s the Y?” This supports the theory that children start off with lexically specific frames, which are highly frequent in the input and then these categories broaden and as children notice overlap between categories, they will abstract and eventually form adult-like generalisations.

Children may abstract from lexically specific frames to produce a broader more generalised language. However, children could also abstract from their errors, which may cause them to make a greater number of these errors. It could be the case that children may make an error because of a learnt string heard in the input and used out of context from the larger grammatical utterance (Tomasello, 2003), but because a child’s learning will involve them abstracting across stored examples, they will then produce this string with a lexical item that would not have been heard in the input. We are now going to discuss explanations for a range of children’s
error types, from a constructivist perspective. Throughout the next section, it is important to keep in mind, that although often the input may explain the starting point of many of children’s errors, children will abstract away from what they have heard in the input and therefore input-based explanations cannot be the whole picture.

1.3 CHILDREN’S ERRORS

In studying children’s acquisition of language, an important feature that emerges is children’s grammatical errors. In learning multi-word speech children will make a wide variety of errors and these are of great interest to us when trying to establish how they learn the system in the first place. If a variety of children make the same type of error, it is an interesting question to investigate the factors involved in leading them to make this same mistake. True, not all children make the same errors, however, some errors are relatively frequent amongst children speaking the same language. For example, in English-speaking children there is a wide variety of common errors from overextension errors, when children overextend the use of a grammatical rule, such as the past tense regular –ed, and used this on an irregular form to create utterances such as “goed”; to agreement errors, when children use incorrect agreement/tense marked verb/noun forms, such as using a verb marked for 3sg, for example “likes” with a pronoun marked for 1sg, for example “I”, therefore producing the error “I likes”; optional infinitive errors, when children produce utterances that lack tense and agreement markers that are obligatory in adult speech, such as “That go there”; pronoun case marking errors, where children have used an incorrect pronoun in place of the correct form such as “Her like it”. Researchers working within the constructivist approach have focussed
on looking for the source of these errors in the input that children are exposed to. This approach to language acquisition attempts to address the individual variation in error patterns, and is based on the view that children learn their native language from observing patterns and making generalisations from the input (Bannard & Matthews, 2008; Theakston, Lieven, Pine & Rowland, 2001; 2002; Theakston, Maslen, Lieven & Tomasello, 2012; Theakston, Ibbotson, Freudenthal, Lieven & Tomasello, 2015). Differences in children’s input are expected to result in differences in children’s patterns of learning and error (Theakston, Lieven & Tomasello, 2003).

1.3.1 Frequency

Across a variety of construction types, there is increasing evidence that constructions heard less frequently in the input attract a higher rate of errors. Ambridge, Kidd, Rowland & Theakston (2015) provide an overview to evidence this and argue that there is a frequency sensitive mechanism in place. And although there is by no means a linear relationship between frequency in the input and rates of all types of errors in children’s speech, frequency is still a hugely important factor. There is a large number of studies highlighting the link between frequency in the input and children’s rate of errors.

The entrenchment hypothesis (Braine & Brooks, 1995) claims that repeated exposure to a particular verb, for example “giggle” in a particular construction, for example the intransitive, “Mummy giggled” will strengthen the probabilistic evidence that it cannot be used in other constructions, in which it has not been heard, for example the transitive construction. Therefore the higher the frequency of a verb heard in one construction, the less likely a child will think it is acceptable in
another construction that it has not previously experienced in the input. There has been considerable support for this hypothesis, often looking at argument structure overgeneralisation errors. It is argued that many overgeneralisation errors such as past-tense overregularisation (overextending the regular past tense form –ed heard in the input, for example “runned”) will eventually cease as the child acquires the correct form (“ran”). However, argument structure overgeneralisation errors are more problematic, as there is often no direct competitor heard in the input that will block this incorrect use. These errors are when a particular verb, such as the intransitive verb “disappear” is used in an argument structure construction such as the transitive causative construction, which is incorrect and not licensed in the adult grammar, for example “The magician disappeared the rabbit”. Given that it seems children do not receive negative evidence, which is feedback when their utterances are incorrect and ungrammatical (Lieven, 1994), it is unclear how children learn to not produce these errors. Brooks, Tomasello, Lewis & Dodson (1999) found evidence for entrenchment when investigating argument structure overgeneralisation errors. Children aged 3 years, 4-5 years and 8 years were exposed to exclusively transitive and exclusively intransitive verbs. They were then asked questions which may encourage them to overgeneralise usage of these verbs. The children saw four actions performed by puppets and each of these actions had two verbs with similar meaning, one that was frequent and more familiar to young children and a less frequently heard, less familiar verb. It was found that children were more likely to overgeneralise the use of the less familiar verbs. This shows that as children learn the transitivity status of verbs and as their usage of these verbs in particular construction types becomes entrenched, they then become reluctant to use these verbs in other argument structure constructions.
Ambridge, Pine, Rowland & Young (2008) also investigated children’s argument structure overgeneralisation errors. Ambridge et al. (2008) attempted to investigate two proposals for how children restrict these errors, the entrenchment hypothesis and Pinker’s (1989) semantic verb class hypothesis. As well as investigating the effects of the frequency of verbs, on children’s (and adults’) judgements of these errors, the verb semantic class was also investigated as Pinker (1989) theorised that each verb in a speaker’s lexicon is assigned to one of a ‘narrow-range’ set of semantic classes. Pinker claimed that certain classes were therefore semantically consistent with certain construction types. For example, a ‘motion in a particular direction’ class is said to be semantically consistent with the intransitive construction and contains verbs such as ‘ascend’ and ‘rise’. A ‘manner of motion’ class, is said to be semantically consistent with both intransitive and transitive constructions and contains verbs such as ‘swing’ and ‘bounce’. Twenty-seven children aged 5-6 years, twenty-four children aged 9-10 years and forty-two adults used a five point scale to rate grammatical (correct intransitive) and overgeneralised (incorrect transitive causative) uses of high frequency, low frequency and novel intransitive verbs from each of three semantic classes. The three semantic classes used were ‘directed motion’ (fall, tumble), ‘going out of existence (disappear, vanish) and ‘semivoluntary expression of emotion’ (laugh, giggle). It was found that for all age groups, the preference for grammatical (intransitive) over ungrammatical (transitive) verb uses was significantly greater for high frequency verbs compared to both low frequency and novel verbs and for low frequency, more than novel verbs. This shows an entrenchment effect, where for high frequency verbs, considerable evidence regarding correct usage has built up. It was also found that the semantic verb class had an effect on the grammatical
ratings, with the preference for grammatical (intransitive) verbs being greatest for the ‘semivoluntary expression of emotion’ class, which is associated with the “lowest degree of direct external causation (the prototypical meaning of the transitive causative construction)” Ambridge et al. (2008; p. 88). This shows that children do form semantic verb classes and will associate meaning with certain construction types. It also highlights the fact that frequency is not the whole picture. This study shows the importance of frequency as well as semantics, in the retreat from overgeneralisation errors.

Frequency has also been shown to be an important factor in children’s correct wh-question production and subject-auxiliary inversion errors. Rowland, Pine, Lieven and Theakston (2003) tested the claim that wh-question acquisition is influenced by the frequency that children hear wh-words and verbs in the input (see below; Rowland & Pine, 2000). Twelve two to three year old children’s data as well as their caregiver’s data was examined. It was found that, above syntactic and semantic complexity, the frequency of particular wh-words and auxiliaries in the input predicted the order of acquisition of wh-questions. This again offers further support regarding the role of frequency in language acquisition.

Räsanen, Ambridge and Pine (2013) have found that English speaking children may make errors such as “he play” which have a missing 3sg marker –s, because of defaulting to a high-frequency and/or the most phonologically simple form. These type of errors are common in early children’s speech and are often referred to as Optional Infinitive errors because the verb is in a non-finite form (Wexler, 1998). Räsanen et al. (2013) investigated these specific errors with twenty two children aged three to four years. It was found that there was a negative correlation between the proportions of bare verse –s marked forms in the input
corpus and the rate of –s production. This shows that children may often default to the form that they have heard most from their input and again highlights the importance of frequency in children’s language learning and their errors.

1.3.2 Strings in the input

As well as simple frequency of constructions and lexical items heard in the input, some error patterns are thought to derive from strings heard in the input. In this context when we refer to a construction, we are referring to an entire grammatical sentence. However, when we refer to a string in this context, we are referring to a smaller part of this construction. Therefore we will use ‘string’ when referring to a group of words that were used together within a larger, grammatical construction, which may not be grammatical in themselves when used independently. An example of a construction in this context is “Will you let me have the ball please?” Whereas a string could be “me have” or “ball please”. These strings can then be used by children independently from the larger construction in which they were heard, causing a grammatical error. There are a range of studies that have found a link between strings in the input and children’s early errors. Theakston, Lieven and Tomasello (2003) investigated whether children’s use of non-finite verb forms such as “it go there” used in finite contexts “it goes there” can be explained because of the presence of questions such as “Where does it go?” in the input. This would be explained by the non-finite verb form immediately following the grammatical subject in this string in the input. Two and a half to three year old children were exposed to known and novel verbs produced in either declaratives, questions or both and were then questioned to elicit use of these verbs. It was found that for the novel verbs, the children’s pattern of verb use was very
close to that of the verb use modelled, even when this resulted in an ungrammatical utterance. This shows that children will learn from strings in the input in larger constructions. This can therefore lead to errors in children’s early speech. Further, Theakston and Lieven (2008) conducted an experimental and naturalistic study to investigate whether there is a relationship between the presence of questions in the input modelling S-V word order (e.g. Is he playing?) and children’s auxiliary verb omissions in declaratives, for example in utterances such as “he playing” where auxiliary BE is omitted. The experimental study involved exposing two year olds to novel verbs modelled in questions only or declaratives only. The naturalistic study involved looking closely at one child’s corpus data and looking at the influence of declaratives and questions in the input on the child’s use of auxiliary BE. Both studies revealed that the presence of questions in the input lowered the levels of auxiliary BE in the child’s declaratives when compared to the presence of declaratives in the input. As well as this, the child’s own prior usage or non-use was an important factor and also influenced subsequent use.

Strings in the associated input have been shown to be a relevant factor in the simulation of the development patterning of children’s optional infinitive (OI) errors in English, Dutch, German and Spanish. (Freudenthal, Pine, Aguado-Orea & Gobet, 2007). The model of syntax acquisition in children (MOSAIC) learns from child-directed speech and was able to successfully simulate the data from all four languages. MOSAIC incrementally learnt input utterances from the ends of utterances, storing more words, each time the utterance is presented. The model gives output at various points in training and this is compared to the real speech produced by the child who had received that particular input. MOSAIC simulated the difference between Spanish, in which OI errors are virtually absent and the
languages (For example English, Dutch and German) in which subject is obligatory and do display high levels of OI error. The successful simulation of the proportion of finite and non-finite utterances for each particular language shows that the model’s utterance final bias and the distributional statistics of the input can explain the variation in the occurrence of OI. MOSAIC produces OI errors, as it learns preferentially from the utterance-final phrases, leading to the use of non-finite verbs where children are learning a language that place these in sentence-final position when complex verb phrases (tense-marked auxiliary + infinitive) are used. This shows further evidence that children can learn from strings heard in the input and produce errors when they produce these strings out of the context of the larger construction.

Strings in the input have also been shown to be a relevant factor in children’s errors in wh-questions. Rowland and Pine (2000) investigated one child’s early wh-question data and subject-auxiliary inversion errors in a corpus analysis of one child between the age 2;3 years and 4;10 years. Rowland and Pine tested two, movement rule-based theories that claimed children had access to a subject-auxiliary inversion rule and used this in order to produce correctly inverted questions. The subject-auxiliary inversion rule is part of a generativist rule-based theory that involves children applying a set of rules when transforming a declarative sentence into a wh-question and is specifically the subject and the auxiliary switching positions so that the auxiliary moves to the head of the phrase. The two movement theories (DeVilliers’s, 1991; Valian, Lasser & Mandelbaum, 1992) tested by Rowland and Pine, claimed that children are failing to apply this rule when they are producing uninversion errors such as “What he can ride in?” It was suggested that, in fact, an alternative theory that does not assume the children’s
knowledge of grammatical categories or rules, could explain this data better. It was investigated if actually the wh-questions that are found to occur with inversion are specific wh-word-auxiliary combinations combined with a variety of lexical items. Further, whether the frequency of wh-word-auxiliary combinations in the child’s input related to the questions that were inverted or uninverted. This was found to be the case and the data therefore suggested a distributional learning mechanism as an explanation to the corpus data, better than a rule governed one. The results from Rowland and Pine (2000) suggested that children learn specific wh-word-auxiliary strings from the input. The strings with higher frequency are learnt first and it is with these learnt strings that the child was producing a correctly inverted wh-question.

Ambridge, Rowland, Theakston & Tomasello (2006) investigated children’s acquisition of non-subject wh-questions using four wh-words (what, who, how and why) and three auxiliaries (BE, DO and CAN). It was found that children may be initially acquiring the wh-word+lexical auxiliary combinations (e.g. What can X? Why are X?) from strings heard in the input. Similarly Ambridge and Rowland (2009) found a significant correlation between the frequency of strings in the input and correct production of 11 of 12 lexical frames when investigating children’s production of positive and negative, what, why, yes, no questions. It was also found that some of the negative question types had particularly high error rates, higher than one might expect based on input frequency alone. This further highlights that input frequency is by no means the only cause of these error patterns.
1.3.3 The Rate of Error

The variation is vast between children as well as for each particular child, for the type of error, number of each particular error produced, number of types of errors, the onset of errors, the ceasing of errors and the rate of recovery from these errors. Children do not produce an error and then produce this error consistently. But actually children are often producing errors alongside the correct grammatical utterance. Children will therefore have a competition between the erroneous form and the correct form and during a certain stage in language development, these erroneous forms may win out as the form of choice in a selection of utterances. This links us back to the notion of the Competition Model (Macwhinney & Bates, 1989) with two competing forms available to express a certain function. The frequency of certain characteristics of the input can often be related to the rate of errors (as exemplified above). This can explain how certain modelling in the input can lead children to make form-function mappings that are not actually present in the adult grammar. These errors will eventually cease from children’s language. Again, this is varied, this doesn’t happen instantly, but gradually different errors will disappear and no longer appear in children’s speech. This will happen as the child’s weighting of correct forms outweighs the weighting of the erroneous forms. The competition between the choices will be won out by the correct form. This will happen as the evidence in the input increases the modelling of the correct form and therefore this form-function mapping will intensify. As this happens, the weighting of the incorrect form will gradually decrease, as the lack of modelling of this will leave this form-function mapping unsupported and eventually redundant.
1.4 PRONOUN CASE MARKING ERRORS

Another frequent and interesting type of error is pronoun case marking errors and this will be the focus of the current thesis. These are commonly made by English speaking children and seen at around two to four years.

1.4.1 Case Marking

Case refers to the grammatical function of a word within a sentence. Within English there are just three cases. In their old English form they are known as the Nominative, Accusative and the Genitive case. However in modern English these three cases are more commonly known as Subjective, Objective and Possessive.

<table>
<thead>
<tr>
<th>Nominative Case (Subjective)</th>
<th>Accusative Case (Objective)</th>
<th>Genitive Case (Possessive)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Me</td>
<td>My</td>
</tr>
<tr>
<td>You</td>
<td>You</td>
<td>Your</td>
</tr>
<tr>
<td>He</td>
<td>Him</td>
<td>His</td>
</tr>
<tr>
<td>She</td>
<td>Her</td>
<td>Her</td>
</tr>
<tr>
<td>We</td>
<td>Us</td>
<td>Our</td>
</tr>
<tr>
<td>They</td>
<td>Them</td>
<td>Their</td>
</tr>
</tbody>
</table>

Unlike nouns, pronouns have distinctive forms for each of the three cases:

Subjective/Nominative forms refer to the subject within a sentence,

Objective/Accusative forms refer to the object in a sentence, and

Genitive/Possessive forms are used to express ownership (see Table 1.1 for examples).
1.4.2 Pronoun Case Marking Errors

A pronoun case marking error is when a child uses an incorrect pronominal case in place of the correct pronominal case. For example, in English, children use me or my in nominative position instead of I, resulting in a me-for-I (accusative-for-nominative) pronoun case marking error such as “Me do it” or a my-for-I (genitive-for-nominative) pronoun case marking error such as “My have it”. Another common pronoun case marking error is when children use her instead of she, resulting in a her-for-she (accusative/genitive-for-nominative) pronoun case marking error such as “Her like it”. These first person accusative-for-nominative (me-for-I), genitive-for-nominative (my-for-I) and third person feminine accusative/genitive-for-nominative (her-for-she) errors are the most common types of pronoun case marking errors (Budwig, 1989; Rispoli, 1998a; 1998b; Vainikka, 1994). Other error types such as using him instead of he, resulting in a him-for-he (accusative-for-nominative) pronoun case marking error (e.g. “Him can’t see”) are seen much less often and even less frequent are nominative-for-accusative (I-for-me) or nominative-for-genitive (I-for-my) pronoun case marking errors. Case selection is a complex task, however, this clear observed bias towards specific error types suggests that it is not just the case that children are confused in this complex task (as Maratsos, 1979 suggested), but that there is a more specific explanation.

The rates of children’s pronoun case marking errors can vary dramatically from child to child, as well as the type of error that they predominantly make and there have been many differing explanations into this variation.
1.4.3 Agreement/Tense Omission Model

One very influential account of English pronoun case-marking errors is the Agreement/Tense Omission Model (ATOM) (Schutze & Wexler, 1996; Wexler, 1998). This generativist theory predicts that children are able to mark both tense and agreement by the time they begin to produce multiword speech as the parameters of basic inflection and clause structure are set. Schütze and Wexler (1996) argue that children do have case assignment mechanisms at this stage, but errors occur when children fail to check for both tense and agreement in their utterances. As agreement is assumed to licence nominative case, if the child fails to check for agreement, the result may be that accusative and genitive pronouns such as me, my, her and him are incorrectly used in contexts requiring nominative pronouns such as I, she and he. This model also crucially predicts that, if agreement and tense are present in the underlying representation of the sentence, pronoun case marking errors should not occur. This is because children have checked both tense and agreement and therefore should assign the correct case. Thus, errors such as “me am going”, “him goes” or “her’s gone” should not be found (Pine, Rowland, Lieven & Theakston, 2005). These utterances are overtly marked for agreement through the use of person marked (auxiliary) verb forms which, according to the ATOM model, should licence the use of the nominative pronoun. Therefore, according to Schutze (1997), the rate of agreeing verb forms occurring with non-nominative subjects will be so low, that these errors are anomalies that can be disregarded as “noise” in the data (p.236).

The ATOM (Schütze and Wexler, 1996) has been supported by evidence showing that nominative pronouns are used frequently with agreeing verbs whereas non-nominative pronouns are very rarely used with agreeing verbs (Wexler,
Schütze & Rice, 1998). Wexler et al. also found a significant association between
the use of correct subject case and levels of overt agreement marking. However,
these findings do not explicitly test their main prediction that children will use
agreeing verbs with non-nominative subjects at a rate so low that these errors can
be disregarded as noise. Pine et al. (2005) pointed out that some children may have
high rates of non-nominative subjects used with agreeing verbs; however this effect
may not be seen, as their rate of nominative subjects used with agreeing verbs is
even higher. Pine et al. (2005) set about testing this main prediction of the ATOM
and therefore whether the rate of agreeing verb forms occurring with non-
nominative subjects are so low, that these errors can be disregarded as noise. Pine et
al. did this by setting an upper limit on the rate of non-nominative subjects
occurring with agreeing verb forms that counted as ‘noise’ (10%). They then tested
to see if the rate that these errors actually do occur, is above or below this. Pine et
al. noted that the low rate of non-nominative subjects with agreeing verbs may be a
consequence of the overall low rate of non-nominative subjects appearing with any
verbs. To explore this, they identified children who produced a reasonably large
number of non-nominative subjects. It was found that for the feminine case-marked
subjects, the rate of agreeing verbs with non-nominative feminine subjects was over
30%. This high rate cannot be simply disregarded as noise and this raises serious
doubts about the ATOM and the theory that children produce non-nominative
subjects because of lack of agreement.

Another argument is that when children make pronoun case errors, there is
always a default pronoun case form that they use (Schütze & Wexler, 1996). In
English, the accusative case (me, her, him) is argued to be the default form as me,
can be seen in “syntactically unrestricted environments in the adult grammar”
(Rispoli, 1998a, p. 386) resulting in high rates of me-for-I errors. Me is considered to be less syntactically restricted because it can be used in several positions within a sentence grammatically, for example “Let me get it” and “Did you see it get me?”

Although this approach can potentially explain why accusative-for-nominative pronoun errors are so frequently observed, as children can potentially just default to the form that they have seen used in most sentence positions, it does not explain the variability across children, or why children frequently produce genitive-for-nominative pronoun errors, in apparent contradiction of the claim for a single default form. Schütze (1997) suggests that genitive subjects are possible in sentences that are lacking both tense and agreement. However, there are counterexamples in which tense is clearly present, for example, “My taked it off” (Jeffrey, Budwig, 1995, p.73) and such examples contradict this theory. Data such as this are again simply disregarded as ‘noise’, however as Rispoli (1998b) points out, “the source of that noise is never fully addressed” (p. 535).

1.4.4 Paradigm Building Model

Another, alternative theory to explain children’s case marking errors is Rispoli’s (1994; 1998a; 1998b; 1999) Paradigm Building Model. The paradigm building model is a model which suggests that children’s linguistic systems are made up of cells defined by person, gender, number and case. Each pronominal form is said to fill a particular cell(s), for example, the form ‘He’ represents the third person, singular, masculine form in nominative case and the form ‘My’ represents the first person, singular form in genitive case. Whereas ‘Her’ represents the third person, singular, feminine form in both accusative and genitive case. Rispoli (1998b) argues that errors are more likely to occur when children are
learning this paradigm but then attempt more cells than they can handle. If a child attempts to refer to a particular referent when the form required is one they are still in the process of learning, Rispoli claims this can result in errors or the overuse of other forms in place of this unknown form. One advantage of this approach is its potential to explain the different error rates observed with different pronoun forms.

Rispoli (1998b) found that the rate of her-for-she errors was considerably higher (49%) than other accusative-for-nominative pronoun errors (11% him-for-he and 12% they-for-them), and concluded that this is due to the ‘double cell effect’: for each pronominal form there are three cells, one that corresponds to the nominative form, one to the accusative and the third to the genitive. The third person feminine pronoun her occurs as both the accusative and genitive form and therefore fills two of the three available cells for third person singular feminine reference (one that corresponds to the nominative form ‘she’, one to the accusative ‘her’ and the third to the genitive ‘her’), facilitating its retrieval from memory. Other pronouns have three different forms, such as he, him and his or I, me and my, making the likelihood of incorrect selection of any particular error form lower.

Rispoli (1998a) also attempted to account for children’s different error patterns and for the fact that as well as accusative-for-nominative (me-for-I) first person pronoun case errors, genitive-for-nominative (my-for-I) errors are also quite commonly found, for example, “my open that, Mommy” (Megan: Budwig, 1995). He (1998a) investigated this differential pattern of use by studying naturalistic data of twelve children, observed monthly, from the age of 1;0 to 3;0 years. The data suggested that “me-for-I and my-for-I replacements were antagonistic” (p. 390) with children having a preference for one or other error type. Rispoli found that children who produced high levels of correct me, when my-for-I errors first emerged, were
more likely to make *me*-for-*I* errors than *my*-for-*I* errors, whereas children who produced few correct *me* utterances at the onset of *my*-errors, were likely to favour *my* in their *I*-replacements, supporting a previous case study by Vainikka (1994). This suggested that the preference to use *my* as a replacement of *I* is due to the low levels of correct *me* production during this period. This finding is particularly interesting because it suggests that children’s pronoun use prior to the emergence of any errors may have an impact on the type of subsequent errors and this will be explored further in the following chapters. However, this can predict which type of errors children may make, but does not explain why children are replacing *I* with any incorrect form to begin with, so other factors also need to be explored.

1.4.5 Phonological Properties

Rispoli (1994; 1998b) offers another possible explanation for the great variation in pronoun case marking errors and the frequently observed first person singular case marking errors, by considering the phonological properties of the different pronominal forms. Two out of the three possible choices for a first person pronoun, *me* and *my*, share an initial phoneme and *I* is phonetically the ‘odd one out’. From this, Rispoli claims that when children are unsure of the correct pronominal form, they will choose one of the phonetically similar pronoun choices in substitution of the odd one out. Rispoli notes that children also make some nominative-for-accusative pronoun case errors with the third person masculine form, *he*-for-*him* as well as the plural form *they*-for-*them*. He attributes these errors to the pronouns’ ‘morphophonetic structure’ as the three pronoun possibilities for the third person masculine form, *he, him* and *his* all have a common phonetic core, as do the plural forms, *they, them and their*. As there is no ‘odd one out’, there can
possibly be confusion in either direction when making overextensions. However, if phonological similarity means there can be overextensions in either direction, then a question arises as to why *his* is not also overextended. The lack of *his* overextensions in comparison to *he* overextensions could be an error in perception and actually *his* is used in pronoun case marking errors but assumed to be *he’s*. Or it could indicate that this phonological explanation is limited and other factors play a part in explaining pronoun case marking errors. So again this does not offer a complete picture and other possible causes need to be considered.

1.4.6 Input based explanations of pronoun case marking errors

Many researchers from a constructivist approach (Kirjavainen, Theakston & Lieven, 2009; Tomasello, 2000; 2003) have looked at the relations between the input received by children and the pronoun case marking errors that they produce. One input-based explanation for pronoun case marking errors is that particular error patterns derive from strings heard in the input, but these are used in children’s speech independently from the larger construction in which they typically occur. This has also been seen in other error types and discussed above (Theakston et al., 2003; Ambridge et al., 2006; Freudenthal et al., 2007; Theakston & Lieven, 2008; Kirjavainen & Theakston, 2011). An example of this in pronoun case marking errors comes from Tomasello (2000; 2003). He suggests that children’s accusative-for-nominative pronoun case marking errors might be explained by children’s exposure to input containing *me* or *her* followed by a verb, within larger complex sentences. In some constructions, the English objective case *me* or *her* can appear as the object of the main clause and subject of the ‘subordinate’ clause, for example “Let her do it” or “Help me find it” (Tomasello, 2003, p. 135). These kinds of
utterances may lead children to produce pronoun errors such as “her do it” or “me find it”. Children are not simply imitating a simple sentence from the input, but are ‘omitting’ the beginning of complex sentences. This is perhaps due to children learning from the ends of utterances due to a more general bias to remember things heard most recently, and only over development are the longer, more complex utterances learnt and errors cease (Freudenthal, Pine, Aguado-Orea & Gobet, 2007).

This was investigated in a naturalistic study by Kirjavainen et al., 2009 and is discussed in detail, see Chapter two, page 54. An experimental study also supports this, see Mcknight & Theakston, 2009, for further details, Chapter two, page 56.

The modelling of the me-verb string in the input (Tomasello, 2003) as well as the modelling of correct I-verb strings, may lead to children having a competition between forms when wishing to express a first person pronoun-verb utterance. During a certain point, a form-function mapping may be formed with me and a certain verb form being associated with expressing a certain function. There will then be competition when the child wishes to express this function between the erroneous form-function mapping made and the form-function mapping modelled correctly in the input (I). The form chosen during this stage of competition will depend on certain frequencies and aspects of the input, the modelling of forms and the evidence received to strengthen each form-function mapping. The competition between the error me and the correct form I will eventually cease when the correct form is strengthened and outweighs the error in the child’s system. This just again shows that it is not just a direct link between the frequency of a form within a certain construction in the input and the rate of errors, but an interplay between this,
the strength of competing forms and the availability of certain forms when children are searching for a form to express a function.

1.5 **MY-FOR-I ERRORS**

The focus of the current thesis is a specific type of pronoun case marking error, namely genitive-for-nominative, *my-for-I* pronoun case marking errors. Although a number of early child errors, including accusative-for-nominative case errors in English, can be explained in terms of input-based learning, there are no analogous input-based explanations for other observed pronoun case errors such as genitive for nominative *my-for-I* errors (“*My do it*”), as children do not hear complex sentences in the input containing *my* in pre-verbal position. These errors therefore pose a challenge for constructivist approaches to language learning as, currently, they have no clear explanation within this framework. While children cannot get *my-for-I errors* directly from strings in the input as they can for *me-for-I errors*, other aspects of person reference in the input could be contributing and we now turn briefly to these.

1.5.1 **Pronoun Modelling**

As well as pronoun case marking errors, about twenty percent of children commonly make pronoun reversal errors, such as using *I* and *me* instead of *you* and vice versa (Dale & Crain-Thoreson, 1993). Learning first and second person pronouns is a difficult task for children as the referent of *I* in the input, is not the same person who is *I* in their own linguistic productions. The most common type of error is the child’s incorrect substitution of *you*, for *me* or *I*, when referring to themselves, for example “Pick you up” to mean “pick me up” (Tomasello, 2003),
which is reported to appear four times more than the reverse *I/me-for-you* error (Dale & Crain-Thoreson, 1993). The frequency of *you* in the input is considered an important explanation of this pattern of reversal errors, as *you* is much more frequent in the speech to children than is *me* or *I* (Lieven, Pine & Baldwin, 1997), and is used by mothers to refer to their children during parent-child interaction. Frequent input of the word *you* coupled with reference to the child may cause the child to associate the second person pronoun with reference to themselves (Tomasello, 2003). Evidence shows that children with siblings find first and second person pronouns less confusing, and second born children are more advanced than firstborns in their pronoun production (Oshima-Takane, 1988; 1999), which has been argued to reflect their observation of adult-child interactions modelling the contrastive use of pronoun forms such as *I/my* and *you*. Thus, the relative frequencies with which children are exposed to particular form-function mappings appears to be relevant in determining the acquisition of pronoun forms, at least for first and second person pronouns.

Smiley, Chang and Allhoff (2011) argued that the specific pattern of pronoun and proper noun use in the input determines the speed with which children work out correct pronominal reference. Previous studies have shown that children vary in their patterns of self-reference, with some children favouring pronouns and others using proper names (Bloom et al., 1994). Parental input is shown to also sometimes include *own name* (“mummy”) and proper names for siblings and child, in subject position, rather than conventional pronouns (Conti-Ramsden, 1989). Smiley et al. (2011) is discussed in further detail in Chapter three, page 90.
1.5.2 Functional Approach

Another approach to child language development is the Functional Approach. This approach explores the many functions that language serves and suggests that children have a set of functions and they search for the linguistic form to map on to these functions (Budwig, 1995). Researchers within the field of child language development vary in the degree to which they believe language develops in relation to its communicative function. However, Budwig (1995) points out, “it is possible to believe that children begin with functionally oriented language systems, and yet over development these systems become more abstract” (p. 9). Research from Slobin (1981; 1985) shows that children’s linguistic forms do not merely copy what they hear from the adult system, but are creative in their early speech (see also Cameron-Faulkner, Lieven & Theakston, 2007, on early negation). So, children can allocate their own forms to serve a particular communicative function and these forms may not serve this same function within the adult grammar. Each pronominal form used by the child therefore may have a different function and this may lead to errors (Karmiloff-Smith, 1979). Deutsch and Budwig (1983) noted the contrastive use of specific linguistic forms based on the pragmatic force of the utterance (see also Carter, 1975). Children’s early contrastive and differing use of possessive referents, my+noun or own name+noun is claimed to be dependent on the communicative force of the utterance, therefore whether the utterance was indicative or ‘violitational’ (Deutsch and Budwig, 1983).

Budwig (1989; 1995) specifically investigated children’s pronoun case marking errors by studying the pragmatic distinctions between the pronominal forms that children use in their utterances. She claimed that children are extremely systematic in the different self-reference forms they use, with each form marking a
different perspective on agency: Proper names such as “Ruth do it” are used when children are involved in ‘intransitive physical actions’, the use of *my* as subject is used to mark the child’s participation in actions high in agency and control and conventional self-reference *I* is used in terms that are low in agency (Budwig, 1995).

In an empirical study, Budwig (1995) looked at six children’s use of varying self-reference terms over a period of four months. The three younger children, aged between 1;8 and 2;6 months; Jeffrey, Megan and Grice, were classified as ‘ego-anchored children’. These children referred to themselves as the main participant over 75% of the time, with very few references to others as the main participant. In addition these children had multiple non-conventional self-reference terms, therefore as well as using *I* to refer to self in subject position in utterances with verbs, these children commonly made pronoun errors, using *me, my* and also *own name*, in subject position. Budwig ran a detailed semantic analysis of the three younger children’s utterances and predicted that the differing self-reference forms used by children will serve contrastive functions. She concentrated on the degree of semantic agency expressed by the utterances as well as analysing the level of control by the speaker. Quite specifically she predicted that children will use *my* when expressing control and *I* when asserting inner beliefs of the self. To code for level of agency, Budwig used a subset of Hopper and Thompson’s (1980) transitivity parameters that were most relevant. This was to identify if utterances containing self-reference occurred with a verb ranked high or low in agency. Budwig concluded that, in fact, children were systematically using differing self-referent forms to communicate differing semantic and pragmatic meanings. Therefore children’s pronominal reference was determined by pragmatic
considerations and children make language distinctions that are not present in adult speech. Jeffrey very clearly exemplifies this distinction and supports Budwig’s predictions with his ‘erroneous’ use of my in subject position. These utterances were found to be high in semantic agency and also functioned as a control act. The use of I in subject position was found in utterances that were ranked low in semantic agency and functioned as assertions. The detailed case studies on Megan and Grice also exemplify the differing self-reference forms serving contrastive functions. The use of proper names as subject, such as “Megan do it” were used when the children were involved in ‘intransitive physical actions’. It is clear that children do sometimes use different self-referent forms in representation of different semantic and pragmatic functions as Budwig shows. These results are of considerable potential interest however a more systematic examination of this proposal is needed because, in her detailed case studies, some of the children were using different forms for the same function and in addition, coding reliabilities were not obtained and there were no tests of significance provided. However, there is clearly a phenomenon in which at least some children do use different forms to express different functions. Therefore the pragmatic and semantic functions should definitely be considered when analysing children’s pronominal case errors.

1.6 THE CURRENT THESIS

The current thesis attempts to offer a constructivist explanation for children’s genitive for nominative my-for-I errors. We have highlighted the importance of frequency and strings from the input when considering children’s patterns of a range of errors. We now aim to investigate various factors in relation to genitive for nominative my-for-I errors. First in chapter two we consider the
basic frequency of the pronoun *my* in the input. Also in Chapter two we then consider possible sources for *my*-for-*I* errors in the input in the form of phonological strings. Children may think that they are hearing *my* in a pre-verbal position in the input as certain constructions can appear this way. For example, some nouns are also verbs and therefore a child hearing a *my*-noun sequence such as “*my drink*”, may think that they have heard a *my*-verb utterance. Or children hearing questions such as “*What am I doing?*” or “*Where am I going?*” could also think they have heard a *my*-verb utterance as the sequence *am I* is phonologically very similar to *my* (e.g. *Where my going?*), especially when produced quickly and fluently in questions, with stress falling after the onset of ‘*am*’. In chapter three we will look at another aspect of the input, namely the proportion of first person pronouns compared with own name replacements such as “*mummy do it*” and how this related to error patterns. We then turn to the child’s own use of pronouns and investigate the possibility that the frequency here will influence their error patterns. From this we will also investigate the claim that children are productive and can often assign certain functions to specific forms and therefore investigate the link between function and errors. Chapter four then includes an experimental study which will continue to examine children’s pronoun usage, but in differing contexts, to extend this focus on the function of particular forms. As well as this, the different levels of children’s protest in these differing contexts, will also be investigated. The current thesis aims to elaborate on explanations from a constructivist perspective on children’s errors and apply these to *genitive-for-nominative* pronoun case marking errors specifically. It is hypothesised that aspects of the input and the frequency of certain forms in the language that is heard will influence children’s error patterns. As well as this, the importance and saliency of certain forms in certain contexts is
also hypothesised to have an effect on children’s rate of errors. We now turn to
paper one of the thesis, chapter two. This second chapter aims to specifically
investigate children’s rate of *my*-*for*-I pronoun case marking errors in direct relation
to the frequency of forms heard in the input.
Chapter 2: Children’s my-for-I errors: Analysing the Input

The following three chapters include three journal style papers. Chapter two contains the first paper in this alternative format thesis.
Children’s my-for-I errors: Analysing the Input

Stacey Mcknight¹,², Elena Lieven²,³ & Anna Theakston²,³

¹Max Planck Child Study Centre, University of Manchester, UK
²School of Psychological Sciences, University of Manchester, UK
³ESRC International Centre for Language and Communicative Development (LuCiD)
2.1 ABSTRACT

English speaking children commonly produce pronoun case marking errors around the age of two-to-four-years, for example, uttering “me open it” instead of “I open it”. The most common type of pronoun error is when accusative/objective pronouns (me and her) are spoken in place of nominative/subjective pronouns (I and she). These errors can be (at least partly) explained by children hearing complex sentences in the input such as “Let me open it”. We are investigating another common case marking error; genitive-for-nominative errors, when children incorrectly use my (possessive pronoun) in subject position, e.g. “my make it”. We are investigating whether my can be explained in the same way as me and her-errors are. My-verb strings are not grammatical in the same way as me-verb strings are, within complex sentences. However, children could potentially think they are hearing these my-verb strings in the input in sentences such as “That is my drink” (noun/verb), “What am I like?” (phonological similarity). We therefore investigate whether the raw frequency of my, the frequency of my-noun/verb utterances, or the frequency of am I-verb utterances, relate to the rates of children’s my-for-I errors. Longitudinal naturalistic data from fifteen English-speaking two-to-four-year-olds was searched for first-person-singular-form (1psg) genitive-for-nominative case errors and all other 1psg preverbal pronominal contexts, including own name+verb utterances. As well as this, caregiver data for fifteen children was searched for all 1psg preverbal pronominal contexts, also including own name+verb utterances such as “mummy do it”. It was found that none of the frequency based input analyses investigated related to the rate of children’s verb utterances that were my-verb utterances. This shows that children’s my-for-I verb errors are not explained by the frequency of hearing my, or these specific strings containing my, in the input, in the same way of me-for-I errors are.
2.2 INTRODUCTION

Pronoun case marking errors have been noted in literature surrounding children’s early language acquisition from as early as Leopold (1939) (see also Bloom, 1970; Brown, 1973; Menyuk, 1969) and are frequently found in many languages. Specifically, English children around the age of 2 to 4 years commonly make pronoun case marking errors (Bloom, 1970; Budwig, 1989; Loeb & Leonard, 1991; Rispoli, 1994; 1998a; 1998b; Tomasello, 2003; Vainikka, 1994). Maratsos (1979) suggested that these pronoun errors simply should be expected, due to the complexity of the rules involved in case selection. This view is countered by the fact that some errors occur often and some extremely little, suggesting a more specific cause. The most common type of pronoun case errors, in English speaking children, are those in which children replace the subject/nominative pronoun with the accusative and the genitive form. This causes children to produce utterances such as, *me-for-I errors*; “Me open it” (Grice: Budwig, 1995, p. 124), *my-for-I errors*; “My make red table” (Nina: Vainikka, 1994, p. 268), *her-for-she errors*; “Her have some tea” (Hannah: Aldridge, 1989, p.178) and *him-for-he errors*; “Him can’t see” (Nina: Vainikka, 1994, p. 295).

In specifically looking at English children’s pronoun case marking errors, it is clear that the rates of errors across children vary dramatically and the types of errors that they produce also can differ greatly. A high proportion of English children’s pronoun errors are made with first person accusative-for-nominative (*me-for-I*), genitive-for-nominative (*my-for-I*) and third person feminine accusative-for-nominative (*her-for-she*) errors (Budwig, 1989; Rispoli, 1998a; 1998b; Vainikka, 1994). Thus, it is quite often found that children overextend *her-for-she* considerably more often than the overextension of *him-for-he*, which is actually
quite uncommon (Rispoli, 1994; Moore, 1995). There are many other less common types of pronoun case errors, such as replacing accusative pronouns (e.g. *him*) with the nominative form (e.g. *he*; “I ride he” (Sarah: Vainikka, 1994, p.286)). Rispoli (1994) found that these nominative overextensions are significantly more likely with the third person masculine pronoun as exemplified above, than with the third person feminine pronoun. These nominative overextensions were not found for first person pronouns at all, therefore suggesting *I-for-me* and *I-for-my errors* are extremely rare, if not non-existent (Rispoli, 1994). It therefore seems unlikely that it is simply children being confused over which pronoun to use in different contexts that can explain these pronoun case errors, as there is a clear bias for certain types of error (Rispoli, 1998a; 1998b).

Certain types of pronoun case marking errors, such as first person (*me-for-I*), and third person feminine accusative-for-nominative (*her-for-she*) errors have been explained from a constructivist approach (Kirjavainen et al., 2009; Rowland & Pine, 2000; Theakston et al., 2001; 2002; 2003). This approach looks at the relations between the input received by children and the errors that they produce. This is based on the view that children learn their native language from observing patterns and making generalisations from the input (Bannard & Matthews, 2008; Theakston, Lieven, Pine & Rowland, 2001; 2002; Theakston, Ibbotson, Freudenthal, Lieven & Tomasello, 2015). Therefore the errors children commonly make could be linked to their linguistic experiences (Theakston et al., 2003). More specifically, commonly heard words or constructions are more likely to be learnt early by children and appear in their speech (Rowland & Pine, 2000; Theakston et al., 2001; 2002). This may lead to less frequently heard constructions attracting a higher error rate compared to more frequently heard combinations (see Ambridge,
et al., 2005, for an overview; also Ambridge, Pine, Rowland & Young, 2008; Brooks, Tomasello, Lewis & Dodson, 1999 for argument structure overgeneralisation errors; Rowland & Pine, 2000 for wh-questions; Rowland & Theakston, 2009; Theakston & Rowland, 2009; Theakston, Lieven, Pine & Rowland, 2005 for auxiliary verb constructions). An additional input-based explanation of errors is that particular error patterns derive from strings heard in the input, but these are used in children’s speech independently from the larger construction. An example of this comes from Tomasello (2000; 2003). He suggests that children’s accusative-for-nominative pronoun case marking errors might be explained by children’s exposure to input containing me or her followed by a verb, within larger complex sentences. In some constructions, the English objective case me or her can appear as the object of the main clause and subject of the ‘subordinate’ clause, for example “Let her do it” or “Help me find it” (Tomasello, 2003, p. 135). These kinds of utterances may lead children to produce pronoun errors such as “her do it” or “me find it”. Children are not simply imitating a simple sentence from the input, but are ‘omitting’ the beginning of complex sentences, perhaps because they are learning from the ends of utterances.

Kirjavainen et al. (2009) investigated whether first person accusative-for-nominative me-for-I errors in children’s speech, could be explained by the input that children hear from their caregivers. In a corpus analysis of longitudinal data from seventeen two-to-four year old, monolingual English-speaking children, the researchers found that the proportion of pre-verbal pronouns that were me in the input, such as “Let me do it” and “Did you see me doing it?” correlated with the proportion of me-verb errors that the children produced during a pre-defined error period. However, there was no relation between the proportional use of me (as a
function of total accusative and nominative pronouns) in the input and children’s *me-for-I* errors. This shows that it is not the overall use of *me* in the input that causes children to use *me* erroneously in preverbal position, but rather hearing *me* in pre-verbal constructions within complex sentences in the input. Further, the complexity of parental input (in terms of MLU) also did not correlate with children’s error rates. Therefore it cannot be claimed that it is simply the complexity of the input causing children difficulty in acquiring sentence subject grammar.

In an additional analysis, Kirjavainen et al. (2009) found that the verbs children use in their *me-verb* errors were more likely to appear in complex utterances with *me* in the input, than verbs children used correctly with the nominative pronoun (*I-verb*). This suggests that children are directly learning from these complex utterances containing *me* in a pre-verbal position, that *me* can precede a verb. This input approach can explain the higher rates of *me-for-I* errors, compared to error rates for pronoun forms that do not appear pre-verbally in the input (Kirjavainen et al., 2009). However, after directly learning this *me-verb* construction from the input, some children then start to generalise them to produce other *me-verb* constructions that will not have appeared in the input. For example, eleven of the fourteen children in the Kirjavainen et al. study were observed to produce *me-errors* with verb forms that would not be produced in complex *me-verb* sentences in the input (for example “me want” or “me got” which are not likely to appear in the input in a *me+verb* sequence), thus demonstrating that the *me-error* pattern had become productive for these children. It is important to note that a central tenet of the usage-based approach is the idea that children move beyond learning that is lexically specific to construct more abstract schemas such as
me+verb (Dąbrowska, 2000) allowing them to produce utterances which will not have been heard in the input. Although the input may provide a starting point for errors of this type, a more sophisticated account is required to explain the developmental trajectory of these errors.

A further study (Mcknight & Theakston, 2010) experimentally investigated this relationship found by Kirjavainen et al. (2009) between the me-verb pairings heard in the input within complex sentences and children’s me pronoun case marking errors. The relationship between the input and children’s errors for the third person singular pronoun her, was also investigated, as her-for-she pronoun errors are very commonly made in children’s early speech (Rispoli, 1998b). Forty children between 3;0 and 4;0 years were tested. Children received input from a talking horse (Dobbin) describing corresponding pictures, in complex sentence structures. Twenty children heard sentences each containing a pronoun+verb construction, such as “watch me/her surprise the bear”. The other twenty children received input containing a verb+pronoun construction such as “watch the princess tease me/her”. This between-groups design meant that half of the children heard the pronouns her and me pre-verbally and half heard them post-verbally. The input for each of the four verbs was followed by three test sentences for each verb, in which children were asked to tell the experimenter what Dobbin said. These twelve test sentences, regardless of input group, were the same and always contained a pronoun case marking error, for example “And then me spray the tiger”. The verb in the test sentence always corresponded to the verb previously heard in the input sentence. All of the children’s responses to the test sentences were coded as to whether the child repeated the pronoun case marking error, omitted it, or corrected it to the correct pronoun form. It was found, that as a proportion of total correct and
repeat responses, children repeated the pronoun error significantly more and
corrected proportionally less when they had received pronoun+verb input. These
results provide support for the input account of me-for-I errors (Kirjavainen et al.,
2009; Tomasello, 2000; 2003). In addition Mcknight and Theakston (2010) found
an effect of pronoun, with children repeating the errors significantly more and
correcting significantly less for the pronoun her compared to me, in support of
previous work showing that children commonly produce her-for-she errors
(Rispoli, 1998b).

The input that children hear has been shown to have an effect on children’s
nominative-for-accusative case marking errors., However, there are no analogous
input-based explanations for other observed pronoun case errors such as genitive
for nominative my-for-I errors (“my do it”), as children will not hear complex
sentences in the input containing my in a pre-verbal position in the way that me-for-
I errors are. Several possible explanations for these my-for-I errors have been put
forward. Budwig (1995) looked at the function of children’s utterances (see also
Deutsch and Budwig, 1983; Karmiloff-Smith, 1979) and claimed that there are
pragmatic distinctions between the pronominal forms that children use. Budwig
specifically claimed that my is used pre-verbally to express high levels agency and
control. Rispoli (1994; 1998a) offered a phonological explanation of me and my
errors, by pointing out that children, when faced with the choice of me, my or I and
are unsure of the correct form, will choose one of the phonetically similar choices
(me or my) in place of the odd one out (I). Further, Rispoli (1998b) claimed that
children had a preference for either my-errors or me-errors and that this relates to
their correct pronoun production at the onset of their errors. However, the role of
the child’s linguistic experience and the specific input containing my and the rate of
these *my-for-I errors* has not yet been thoroughly investigated and this is the focus of the current study.

### 2.2.1 The Current Study

The focus of the current study is to investigate English children’s genitive-for-nominative first person pronoun errors. These *my-for-I errors* such as ”my go there” and “my want that” are quite commonly produced by children at around 2-4 years. These *my-for-I errors* cannot be explained by children hearing the exact pronoun and verb construction in the input within complex sentences, which has been shown to at least partly explain *me-for-I* and *her-for-she errors* (Kirjavainen et al., 2009; Tomasello, 2000). However, it is possible that these errors can be explained by the frequency of *my* in the input, or by specific strings that children hear from the input, in the same way as *me-for-I errors*. Children could potentially think they are hearing these *my-verb* strings in the input in sentences such as “Have you seen my drawing”. Words such as “drawing” and “drink” can be used as both nouns and as verbs. Children may then learn this string directly from the input and produce an error such as “My drawing a fishy” or “my drink my juice”. Further, children may also think that they are hearing these specific *my-verb* strings in the input in utterances that contain *am I-verb*, such as “What am I doing?” These utterances sound very similar to “what my doing?” and therefore children could then produce utterances such as “my doing it” or “my going the shop” after hearing “Where am I going?” in the input. In the current study, we therefore investigate whether; the raw frequency of *my*; the frequency of *my-noun/verb* utterances; or the frequency of *am I-verb* utterances, relate to the rates of children’s *my-for-I errors*. Longitudinal naturalistic data from fifteen English-speaking two-to-four-year-olds was searched for first-person-singular-form (1psg) genitive-for-nominative case
errors and all other 1psg preverbal pronominal contexts, including *own name+verb* utterances. As well as this, caregiver data for fifteen children was searched for all 1psg preverbal pronominal contexts, also including *own name+verb* utterances such as “mummy do it”. A corpus analysis of the longitudinal data of fifteen children available on CHILDES (MacWhinney, 2000) was carried out to investigate if any these genitive-for-nominative case marking errors can be explained by the frequency of certain strings in the input that the children hear, motivated directly by the relation found between the strings in the input and children’s *me-for-I* pronoun case marking errors.

### 2.3 METHOD

#### 2.3.1 Corpora

The twelve children from the Manchester Corpus (Theakston, Lieven, Pine & Rowland, 2001) were included in the corpus analysis to investigate children’s *my-for-I* pronoun case marking errors. A search was conducted using CLAN on the morphological tier for occurrences of *my* followed directly by a verb or auxiliary verb. Only two of the children produced a high number of these errors (above forty errors). Therefore, in order to include more children in the analysis who make a high number of these errors, the CHILDES database was searched. All English speaking datasets that included children above age 1;6 and below age 3;6 years were analysed. Those datasets that had a morphological tier were searched in the same way as the Manchester Corpus; by searching for *my* followed by a verb or auxiliary verb. Any children who made a high number of these genitive-for-nominative errors were included in the analysis (see below). Those datasets that did not include a morphological tier were searched using the *cooccur* command in
CLAN. This extracted all words that co-occurred with *my*. This data was then searched for *my* followed by a verb, which included any utterances that contained intervening words such as “my just do that” or “my can have it”. Again, those children who made a high number of these *my*-verb errors were included in the analysis. Four further children were found to have more than forty errors and were included in the corpus analysis.

As a result, a corpus analysis was conducted on the longitudinal data from sixteen monolingual English-speaking children taken from the CHILDES database (MacWhinney, 2000). These children included:
- The twelve children from the Manchester corpus (Anne, Aran, Becky, Carl, Dominic, Gail, Joel, John, Liz, Nic, Ruth and Warren). These children ranged from 1;8.22 to 2;0.25 at the beginning of the study and from 2;8.15 to 3;0.10 at the end. The children were recorded interacting with their mothers in their own homes for two separate hours in every three week period throughout (Theakston, Lieven, Pine & Rowland, 2001).
- Thomas, who was recorded in Manchester, UK by researchers at the Max Planck Child Study Centre, University of Manchester. Thomas was recorded interacting primarily with his mother in his home between the ages 2;00.12 to 3;02.12 for five separate hours per week, then between the ages 3;02.12 to 4;11.20, for five hours in a single week in every month (Lieven, Salomo, Tomasello, 2009).
- Betty, who was recorded in her own home for twenty-four spontaneous ninety-second intervals throughout the whole day. There are a total of eight recordings, taken every three months, from age 1;6.03 to 3;03.02. A further recording was made at age 4;11.2 years (Wells, 1981).
Nina, recorded interacting primarily with her mother for fifty-two, hour long recording sessions between age 1;11.16 and 3;3.21 (Suppes, 1974).

Laura, whose data consists of a diary study as well as audio recordings. The diary data available is from age 1;03.21 to 3;05.00. This consists of a total 700 files, which are all separate speech events from any given day. The recordings were made from age 1;05.09 to 7;00.14. This consists of a total of 200 recordings, sometimes several from one day. The recordings are made at frequent intervals but then become less so after age 5 (Braunwald, 1997). The diary and audio data were combined, unless otherwise stated. Any differences between audio and diary data are highlighted.

2.3.2 Data Extraction and Coding

In order to uncover every my-error that each child makes, every utterance containing my throughout every file for all sixteen children, was extracted using the KWAL function in CLAN. These my-utterances were transferred into an excel workbook and coded separately for each child. Utterances were coded for the lexical item that follows my in the child’s utterance and its grammatical category, for example noun or verb. If there were more than two instances of my within an utterance, this utterance was coded twice, once per instance of my. Any instances of my-errors and their type (e.g. “my want it” is a my-for-I error, whereas “let my in” is a my-for-me error) were also coded.

In order to calculate my-error rates for the children, it was also necessary to extract utterances containing the other first person pronoun forms I and me. Because different amounts of speech were available for the different children, my-error rates were calculated as a proportion of all first person pronoun+verb,
auxiliary, modal, neg-verb or adv-verb utterances. That is to say, how many of their verb utterances are *my-errors* out of all *my-errors, me-errors* and *correct-I* uses in nominative contexts. To identify when children have established the correct use of *I* and also to compare the number of *my-for-I verb* errors that children are making with the number of correct *I-verb* utterances, every utterance containing *I* was therefore extracted using the KWAL function in CLAN, along with all contracted forms such as *I’m, I’ll, I’ve* and *I’d*1. These *I-utterances* were coded for the lexical item following *I*, just as with the *my-utterances*, and for whether the use of *I* was correct (*I* used with verb, auxiliary, modal, neg-verb or adv-verb). In addition, every utterance containing *me* was extracted and coded for the lexical item following *me* or, as *me* is often found in sentence final position, for no following item. As *me* can be followed by a verb in a correct utterance such as a complex sentence; “Let me do that”, these instances were differentiated from those in which a *me-verb error* occurred. Further, to make sure of a complete analysis of all first person pronouns, any child utterances of *mine* were extracted and coded. However, in a search of every child’s data, only one instance2 of *mine+verb* was found in which the child was referring to themselves and using *mine* in place of *I*. This isolated instance of a *mine-for-I error* across all files shows the rare nature of this type of error.

2.3.3 Individual error periods

To define the period of time in which each child made *my-for-I errors*, an error period was defined for each child. First, we established that each child was

---

1 Every reference made to child or parental *I-verb* utterances, always included the contracted forms of *I’m, I’ll, I’ve* and *I’d*.
2 This instance of a *mine-for-I errors* “mine have it” was in Domin file 21a, line 732.
aware of the correct use of *I* and of *my*, during this ‘error period’. This was to eliminate the possibility that children simply do not know the correct form and this is the cause of these prounon errors.

Table 2.1. Defined *my*-error period for all sixteen children

<table>
<thead>
<tr>
<th>Name</th>
<th>First 3 correct use of <em>I</em></th>
<th>First 3 correct use of <em>my</em></th>
<th>First Error AFTER correct <em>my</em> and <em>I</em></th>
<th>Last error</th>
<th>Error period Age (years; months; days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anne</td>
<td>1a</td>
<td>3b</td>
<td>13a</td>
<td>-</td>
<td>29b</td>
</tr>
<tr>
<td>Aran</td>
<td>1b</td>
<td>5a</td>
<td>13b</td>
<td>-</td>
<td>32a</td>
</tr>
<tr>
<td>Beck</td>
<td>2a</td>
<td>7a</td>
<td>5a</td>
<td>8a</td>
<td>33b</td>
</tr>
<tr>
<td>Carl</td>
<td>1a</td>
<td>1b</td>
<td>11b</td>
<td>-</td>
<td>32b</td>
</tr>
<tr>
<td>Dom</td>
<td>6a</td>
<td>5a</td>
<td>2a</td>
<td>16b</td>
<td>26a</td>
</tr>
<tr>
<td>Gail</td>
<td>3a</td>
<td>8a</td>
<td>12a</td>
<td>-</td>
<td>31a</td>
</tr>
<tr>
<td>Joel</td>
<td>4b</td>
<td>3b</td>
<td>13b</td>
<td>-</td>
<td>29b</td>
</tr>
<tr>
<td>John</td>
<td>1a</td>
<td>4b</td>
<td>1b</td>
<td>XXX</td>
<td>2b</td>
</tr>
<tr>
<td>Liz</td>
<td>5b</td>
<td>1a</td>
<td>8b</td>
<td>-</td>
<td>22b</td>
</tr>
<tr>
<td>Nic</td>
<td>3a</td>
<td>7b</td>
<td>23a</td>
<td>-</td>
<td>33a</td>
</tr>
<tr>
<td>Ruth</td>
<td>2a</td>
<td>13a</td>
<td>24a</td>
<td>-</td>
<td>34b</td>
</tr>
<tr>
<td>Warr</td>
<td>1a</td>
<td>1a</td>
<td>1a</td>
<td>6a</td>
<td>33a</td>
</tr>
<tr>
<td>Nina</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>35</td>
</tr>
<tr>
<td>Thomas</td>
<td>2-02-29</td>
<td>2-01-30</td>
<td>2-07-23</td>
<td>-</td>
<td>3-04-02</td>
</tr>
<tr>
<td>Betty</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>-</td>
<td>7</td>
</tr>
<tr>
<td>Laura</td>
<td>1-5-13</td>
<td>1-5-11</td>
<td>1-6-07</td>
<td>-</td>
<td>2-5-30</td>
</tr>
<tr>
<td>(audio)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laura</td>
<td>1-7-10</td>
<td>1-04-08</td>
<td>1-9-00</td>
<td>-</td>
<td>2-4-16</td>
</tr>
<tr>
<td>(diary)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laura</td>
<td>1-5-13</td>
<td>1-04-08</td>
<td>1-6-07</td>
<td>-</td>
<td>2-5-30</td>
</tr>
<tr>
<td>(both)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a* John only made two errors in total, in files 1b and 2b. John therefore made no *my*-errors after he had produced three correct uses of *my*. John had no error period and was therefore excluded from the analysis.

The file by which each child had produced three correct uses of *I* and *my*, was identified (see Table 2.1). Next, the file in which each child made their first *my*-for-*I* error was identified. If this first error was AFTER the three correct uses of

---

3 The three correct uses had to be three unique utterances with three different lexical items. Utterances of the single pronoun alone such as “*my*” were not counted, nor were idioms, songs or interjections such as “oh my god” or “this little finger on my right”. Also uses of *I* in contracted forms such as *I’m, I’ve, I’ll* or *I’d* will be counted as a correct use of *I*, as long as they are used with another lexical item. (So the single contracted form “*I’m*” will not be counted as a correct use, as it suggests an incomplete utterance. However, “*I’ll go*” and “*I’d go*” would be considered two unique utterances). Any utterances coded with a [?] will not be included in the three correct uses.
my and I, then the file in which the error occurred was defined as the start of the error period. If this first error was in files BEFORE the three correct uses of my and I then the first error in the files AFTER the three correct uses was identified and the file in which this later error occurred was defined as the start of the error period. The error period then ran up until each child’s last my-for-I error plus two additional files. See Table 2.1 for each child’s defined error period.

From this defined error period, each child’s total number of my, me and I utterances was calculated along with the total number of errors made, and the number of these errors which were my/me-for-I verb errors (i.e. incorrect my/me-verb/auxiliary/modal/neg-verb/adverb-verb utterances, for example “my not do it” or “me just go there”). Other my/me-for-I error types included verb omission errors, for example me/my-for-I adjective sequences (“my happy”) or me/my-for-I adverb sequences (“me just sad”). In addition, there were a variety of other my and me errors in which my/me were not substituted for I. The total number of I-utterances that are verb utterances was calculated, which included all contracted I’m, I’ll, I’ve and I’d utterances. All other I-utterances such as I used in tag questions or pre-adjectively were coded as other I-utterances. From the total my-for-I verb, me-for-I verb and correct I-verb utterances, the proportion that are my-errors or me-errors from all first person pronominal verb utterances can be established.

2.4 RESULTS

4 The error period was defined on my-for-I errors so errors that were not my-verb errors (for example “my happy”), were still included.
5 For four of the children in the Manchester Corpus (Becky, Nicola, Ruth and Warren), it was not possible to add two files after the child’s last error, as they had made errors either to the end of their recordings or almost to the end.
6 From this point on, the children’s first person verb utterances will indicate just me, my and I as mine does not occur except for one instance, which has no impact on any results.
Fifteen children had a defined *my-error period* and were therefore included in the corpus analysis. Table 2.2 shows the total number of *my, me* and *I* utterances during this period. In addition, the total number of *my-errors* and *me-errors* is given with indication to the different type of errors.

**Table 2.2. Total number of my, me and I utterances during the defined error period, including number of each different error type.**

<table>
<thead>
<tr>
<th>Name</th>
<th>Total <em>my</em> utterances</th>
<th>Total <em>my</em>-errors</th>
<th>Total <em>me</em>-errors</th>
<th>Total <em>I</em>-utterances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anne</td>
<td>308</td>
<td>5</td>
<td>1</td>
<td>155</td>
</tr>
<tr>
<td>Aran</td>
<td>372</td>
<td>2</td>
<td>0</td>
<td>237</td>
</tr>
<tr>
<td>Beck</td>
<td>243</td>
<td>8</td>
<td>2</td>
<td>215</td>
</tr>
<tr>
<td>Carl</td>
<td>182</td>
<td>22</td>
<td>5</td>
<td>70</td>
</tr>
<tr>
<td>Dom</td>
<td>248</td>
<td>6</td>
<td>0</td>
<td>78</td>
</tr>
<tr>
<td>Gail</td>
<td>357</td>
<td>15</td>
<td>6</td>
<td>170</td>
</tr>
<tr>
<td>Joel</td>
<td>320</td>
<td>4</td>
<td>2</td>
<td>144</td>
</tr>
<tr>
<td>Liz</td>
<td>409</td>
<td>9</td>
<td>0</td>
<td>83</td>
</tr>
<tr>
<td>Nic</td>
<td>182</td>
<td>4</td>
<td>0</td>
<td>106</td>
</tr>
<tr>
<td>Ruth</td>
<td>312</td>
<td>53</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Warr</td>
<td>465</td>
<td>98</td>
<td>1</td>
<td>69</td>
</tr>
<tr>
<td>Nina</td>
<td>915</td>
<td>155</td>
<td>5</td>
<td>379</td>
</tr>
<tr>
<td>Thomas</td>
<td>2409</td>
<td>118</td>
<td>39</td>
<td>546</td>
</tr>
<tr>
<td>Betty</td>
<td>133</td>
<td>60</td>
<td>1</td>
<td>49</td>
</tr>
<tr>
<td>Laura</td>
<td>1719</td>
<td>316</td>
<td>39</td>
<td>931</td>
</tr>
</tbody>
</table>

**2.4.1 Input Analysis**

**2.4.1.1 Input files**

In order to investigate an input approach to children’s pronoun errors, the input files that will be used in such analysis were defined, see Table 2.3.

The amount of data files for each child’s error period varies from child to child, dependent on the density of recordings and the duration of the error period. Therefore, the decision was taken to select the number of input files for each child to vary in proportion with this. The files selected covered at least one month’s worth of recordings prior to the error period to sample the input the child received.
immediately prior to the error period and covered no more than two months in total, to capture a selection of files that was proportional to the number of error files. For some of the children, their input files overlapped into the error period, however this was never more than a two week overlap.

Table 2.3. All fifteen children's defined input files.

<table>
<thead>
<tr>
<th>Name</th>
<th>Error period (no. of files)</th>
<th>Input files (no. of files)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anne</td>
<td>13 – 31 (19)</td>
<td>8-12 (5)</td>
</tr>
<tr>
<td>Aran</td>
<td>13 – 34 (22)</td>
<td>8-12 (5)</td>
</tr>
<tr>
<td>Beck</td>
<td>8 – 34 (27)</td>
<td>3-7 (5)</td>
</tr>
<tr>
<td>Carl</td>
<td>11 – 34 (24)</td>
<td>6-10 (5)</td>
</tr>
<tr>
<td>Dom</td>
<td>16 – 28 (13)</td>
<td>11-15 (5)</td>
</tr>
<tr>
<td>Gail</td>
<td>12 – 33 (22)</td>
<td>7-11 (5)</td>
</tr>
<tr>
<td>Joel</td>
<td>13 – 31 (19)</td>
<td>8-12 (5)</td>
</tr>
<tr>
<td>Liz</td>
<td>8 – 24 (17)</td>
<td>3-7 (5)</td>
</tr>
<tr>
<td>Nic</td>
<td>23 – 34 (12)</td>
<td>18-22 (5)</td>
</tr>
<tr>
<td>Ruth</td>
<td>24 – 34 (11)</td>
<td>19-23 (5)</td>
</tr>
<tr>
<td>Warr</td>
<td>6 – 34 (29)</td>
<td>1-6a (5)</td>
</tr>
<tr>
<td>Nina</td>
<td>4 – 37 (30)</td>
<td>1-6 (6)</td>
</tr>
<tr>
<td>Thomas</td>
<td>2-7-23 – 3-4-04 (135)</td>
<td>2-6-12 – 2-7-22 (28)</td>
</tr>
<tr>
<td>Betty</td>
<td>5 – 9 (5)</td>
<td>4-5 (2)</td>
</tr>
<tr>
<td>Laura</td>
<td>1-6-7 - 2-6-4 (100+diary data)</td>
<td>1-5-09 – 1-6-21 (23)</td>
</tr>
</tbody>
</table>

a Warren had half of one file (file 3b) missing due to him falling asleep during recording. Therefore his input file was extended to also include file 6a.

b Laura’s diary data contained no input, therefore the input files are defined from the audio data. The input period is of the month before the error period. We can therefore assume the input in the audio files had as much influence over child utterances in the diary data as it did on the audio data.

2.4.2 Frequency of the use of my in the input

The caregiver’s utterances throughout the defined input files for each child were analysed. As the amount of input available for each child differed, the proportion of my as a function of all word tokens and of all first person pronominal reference in the input files was calculated (see Table 2.4). We first investigated whether the overall amount that children hear the pronoun my in the input influences their overall my-for-I error rate. The raw frequency of all first person
pronouns in the input files was extracted. As the amount of input available for each child differed, both the proportion of my as a function of all word tokens, and as a function of all first person pronouns (me, my, mine and I) in the input, were calculated. The data violated the assumptions for parametric tests and therefore rank-order correlations were conducted throughout.

Table 2.4. Number of each first person pronoun uttered by caregivers in the defined input files.

<table>
<thead>
<tr>
<th>Name</th>
<th>Total my</th>
<th>Percent my/all words</th>
<th>Total me</th>
<th>Total mine</th>
<th>Total I (inc. I’m, I’ll, I’ve, and I’d)</th>
<th>Total 1psg (my, me, mine &amp; I)</th>
<th>Percent my 1psg (of total me, my, mine and I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anne</td>
<td>66</td>
<td>0.2882%</td>
<td>104</td>
<td>1</td>
<td>336</td>
<td>507</td>
<td>0.13017751</td>
</tr>
<tr>
<td>Aran</td>
<td>14</td>
<td>0.0479%</td>
<td>70</td>
<td>0</td>
<td>231</td>
<td>315</td>
<td>0.04444444</td>
</tr>
<tr>
<td>Becky</td>
<td>17</td>
<td>0.0908%</td>
<td>80</td>
<td>2</td>
<td>306</td>
<td>405</td>
<td>0.04197531</td>
</tr>
<tr>
<td>Carl</td>
<td>2 (0)</td>
<td>0.0135%</td>
<td>10 (3)</td>
<td>4 (0)</td>
<td>149 (0)</td>
<td>168</td>
<td>0.01190476</td>
</tr>
<tr>
<td>Dom</td>
<td>21</td>
<td>0.1107%</td>
<td>101</td>
<td>2</td>
<td>612</td>
<td>736</td>
<td>0.02853261</td>
</tr>
<tr>
<td>Gail</td>
<td>15</td>
<td>0.0931%</td>
<td>26</td>
<td>0</td>
<td>122</td>
<td>163</td>
<td>0.09202454</td>
</tr>
<tr>
<td>Joel</td>
<td>13</td>
<td>0.0872%</td>
<td>69</td>
<td>0</td>
<td>177</td>
<td>259</td>
<td>0.05019305</td>
</tr>
<tr>
<td>Liz</td>
<td>11</td>
<td>0.0809%</td>
<td>47</td>
<td>1</td>
<td>300</td>
<td>359</td>
<td>0.03064067</td>
</tr>
<tr>
<td>Nic</td>
<td>22</td>
<td>0.1141%</td>
<td>95</td>
<td>4</td>
<td>353</td>
<td>474</td>
<td>0.0464135</td>
</tr>
<tr>
<td>Ruth</td>
<td>30</td>
<td>0.1472%</td>
<td>133</td>
<td>0</td>
<td>347</td>
<td>510</td>
<td>0.05882353</td>
</tr>
<tr>
<td>Warr</td>
<td>11</td>
<td>0.0635%</td>
<td>26</td>
<td>1</td>
<td>232</td>
<td>270</td>
<td>0.04074074</td>
</tr>
<tr>
<td>Nina</td>
<td>7</td>
<td>0.0461%</td>
<td>51</td>
<td>0</td>
<td>105</td>
<td>163</td>
<td>0.04294479</td>
</tr>
<tr>
<td>Thom</td>
<td>98 (0)</td>
<td>0.0843%</td>
<td>280 (0)</td>
<td>7 (0)</td>
<td>1863 (2)</td>
<td>2250</td>
<td>0.04355556</td>
</tr>
<tr>
<td>Betty</td>
<td>3 (0)</td>
<td>0.2191%</td>
<td>5 (1)</td>
<td>0 (0)</td>
<td>16 (10)</td>
<td>35</td>
<td>0.08571429</td>
</tr>
<tr>
<td>Laura</td>
<td>25 (1)</td>
<td>0.2129%</td>
<td>30 (6)</td>
<td>0 (0)</td>
<td>196 (31)</td>
<td>289</td>
<td>0.0899654</td>
</tr>
</tbody>
</table>

A Spearman’s rho (two-tailed) test was run to determine if there was a relation between the proportional use of my in the input as a function of either all word tokens (range: 0.000 – 0.003, M = 0.001, SD = 0.001), or all first person pronominal reference (range: 0.012 - 0.130, M = 0.056, SD = 0.031), and the proportion of pronom+verb utterances that contain my-for-I errors in the children’s speech (range: 0.001 – 0.244, M = 0.045 , SD = 0.066). However, both relations were non-significant, rs = 0.196, N = 15, p = 0.483 and rs = 0.196, N = 15, p = .483
respectively. Thus there was no indication that the frequency with which caregivers used *my* overall or as a proportion of all first person pronoun use was related to children’s rates of *my-for-I* errors.

### 2.4.3 “What Am I doing?”

It has been shown that children can make *me-verb* errors from hearing *me-verb* utterances within complex sentences in the input (Kirjavainen et al., 2009; Tomasello, 2000; 2003). Children are very unlikely to hear *my-verb* in any constructions in the input because the sequence is ungrammatical. However, children may often hear the construction *am I* within sentences in the input, in a pre-verbal position. This construction can appear in front of many verbs, such as “what am I doing?”, “where am I going?” or “how am I getting there?” It may sound to the child just like “what my doing?” and “where my going?” These questions containing *am I* followed by verb, could have the same effect as *me-verb* utterances do, such that the child may think they are hearing *my-verb* constructions in the input. Therefore, whether these *am I-verb* constructions in the input are having an effect on children’s *my-verb* errors, was examined. The total number of *am I* utterances was extracted from the caregivers’ input in the defined input files. These instances were then coded as to whether they are followed by: a verb (“Am I going to the car?”); another lexical item (“Am I the dog?”); nothing; (“Who am I?”); nothing in a tag (“I’m having some of Anne’s, am I?”) The number of *am-I verb* utterances was calculated. Further, the total *pronoun-verb* utterances in the input files was calculated to include *am I verb* utterances and the proportion of this total that is *am I* will be given (see Table 2.5). This is in order to look at how often
children are hearing this *am I* pre-verbally in comparison with other pronoun-verb utterances.

Table 2.5. *Total number of am I and specifically am I-verb utterances spoken in the defined input files.*

<table>
<thead>
<tr>
<th>Name</th>
<th>Total 1Am I</th>
<th>Total Am I - verb</th>
<th>Proportion am I-verb / all am I</th>
<th>Total number of utterances</th>
<th>Proportion am I/all utterances</th>
<th>Total pronoun + verb utterances, Inc am I verb utterances</th>
<th>Prop am I verb / all 1psg/am i+verb input</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anne</td>
<td>2</td>
<td>1</td>
<td>0.5</td>
<td>5931</td>
<td>0.003372</td>
<td>508</td>
<td>0.002</td>
</tr>
<tr>
<td>Aran</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>5558</td>
<td>0.0005398</td>
<td>315</td>
<td>0.000</td>
</tr>
<tr>
<td>Becky</td>
<td>10</td>
<td>6</td>
<td>0.6</td>
<td>4990</td>
<td>0.0020040</td>
<td>411</td>
<td>0.015</td>
</tr>
<tr>
<td>Carl</td>
<td>2</td>
<td>1</td>
<td>0.5</td>
<td>3906 (19)</td>
<td>0.0005096</td>
<td>169</td>
<td>0.006</td>
</tr>
<tr>
<td>Dom</td>
<td>3</td>
<td>1</td>
<td>0.333</td>
<td>5409</td>
<td>0.0005546</td>
<td>737</td>
<td>0.001</td>
</tr>
<tr>
<td>Gail</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4808</td>
<td>0.0002080</td>
<td>164</td>
<td>0.006</td>
</tr>
<tr>
<td>Joel</td>
<td>2</td>
<td>1</td>
<td>0.5</td>
<td>4138</td>
<td>0.0004833</td>
<td>260</td>
<td>0.004</td>
</tr>
<tr>
<td>Liz</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3407</td>
<td>0.0005870</td>
<td>361</td>
<td>0.006</td>
</tr>
<tr>
<td>Nic</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>4924</td>
<td>0.0004062</td>
<td>476</td>
<td>0.004</td>
</tr>
<tr>
<td>Ruth</td>
<td>6</td>
<td>2</td>
<td>0.333</td>
<td>5235</td>
<td>0.0011461</td>
<td>512</td>
<td>0.004</td>
</tr>
<tr>
<td>Warr</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4432</td>
<td>0.0002256</td>
<td>270</td>
<td>0.000</td>
</tr>
<tr>
<td>Nina</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>3101</td>
<td>0.0019349</td>
<td>169</td>
<td>0.036</td>
</tr>
<tr>
<td>Thom</td>
<td>8</td>
<td>7</td>
<td>0.875</td>
<td>21745 (12)</td>
<td>0.0003677</td>
<td>2257</td>
<td>0.003</td>
</tr>
<tr>
<td>Betty</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>224 (196)</td>
<td>0.0000000</td>
<td>35</td>
<td>0.000</td>
</tr>
<tr>
<td>Laura</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3263 (606)</td>
<td>0.0000000</td>
<td>289</td>
<td>0.000</td>
</tr>
</tbody>
</table>

A Spearman’s rho correlation to determine if there was a relation between the proportional use of *am I-verb* in the input (range: 0.000 – 0.036, $M = 0.006$, $SD = 0.009$) and the proportion of *pronoun+verb* child utterances that contain *my-for-I* errors (range: 0.001 – 0.244, $M = 0.045$, $SD = 0.066$) was non-significant $r_s = -0.133$, $N = 15$, $p = .636$. This indicates that children’s *my-for-I* error rates were not related to the rate of hearing *am I-verb* constructions in the input.

### 2.4.4 My-noun/verb

Children will not hear any *my-verb* pairings in the input. However many nouns are also used as verbs, such as *drink* and *building*. Therefore children may
hear these lexical items used in their noun form after *my*, in the input, for example “do you like *my* drawing?” or “that is *my* drink”. The child may then use these same lexical items after *my*, but use them in verb form, for example “*my* drawing a picture” or “*my* drink all my milk”, resulting in *my*-verb errors.

Table 2.6. Proportion of mother’s *my*-utterances from the defined input files that contain nouns that can also be used as verb (father’s utterances indicated in parenthesis).

<table>
<thead>
<tr>
<th>Name</th>
<th>Total <em>my</em> in defined input files</th>
<th>No. of different lexical items following <em>my</em></th>
<th>No. of lexical items that can also be used as verbs</th>
<th>Total use of noun/verb lexical items</th>
<th>Proportion of <em>my</em>-noun items that also can be verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anne</td>
<td>66</td>
<td>24</td>
<td>3</td>
<td>12</td>
<td>0.1818</td>
</tr>
<tr>
<td>Aran</td>
<td>14</td>
<td>11</td>
<td>1</td>
<td>2</td>
<td>0.1429</td>
</tr>
<tr>
<td>Becky</td>
<td>17</td>
<td>12</td>
<td>1</td>
<td>1</td>
<td>0.0588</td>
</tr>
<tr>
<td>Carl</td>
<td>2 (0)</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0.0000</td>
</tr>
<tr>
<td>Dom</td>
<td>21</td>
<td>12</td>
<td>2</td>
<td>3</td>
<td>0.1429</td>
</tr>
<tr>
<td>Gail</td>
<td>15</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0.0000</td>
</tr>
<tr>
<td>Joel</td>
<td>13</td>
<td>9</td>
<td>2</td>
<td>3</td>
<td>0.2308</td>
</tr>
<tr>
<td>Liz</td>
<td>11</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>0.0909</td>
</tr>
<tr>
<td>Nic</td>
<td>22</td>
<td>17</td>
<td>1</td>
<td>1</td>
<td>0.0455</td>
</tr>
<tr>
<td>Ruth</td>
<td>30</td>
<td>10</td>
<td>2</td>
<td>6</td>
<td>0.2000</td>
</tr>
<tr>
<td>Warr</td>
<td>11</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>0.0909</td>
</tr>
<tr>
<td>Nina</td>
<td>7</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0.1429</td>
</tr>
<tr>
<td>Thom</td>
<td>98 (0)</td>
<td>62</td>
<td>6</td>
<td>9</td>
<td>0.0918</td>
</tr>
<tr>
<td>Betty</td>
<td>3 (0)</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0.6667</td>
</tr>
<tr>
<td>Laura</td>
<td>25 (1)</td>
<td>19</td>
<td>3</td>
<td>3</td>
<td>0.1154</td>
</tr>
</tbody>
</table>

All lexical items following *my* in the parental input were coded as *nouns only* or as *nouns/verbs* if they can be grammatically used as both. For an item to be considered both a *noun* and *verb* in child directed speech, at least one instance of the noun being used as a verb had to be found in the data files⁷. The numbers of those lexical items that can also be used as verbs was calculated. This is a measure

---

⁷ A search was run on all of the input data across all files. If the noun was found being used as a verb, it was considered a *noun/verb*. Lexical items that could be considered both a noun and verb grammatically in adult speech, but that were not found being used as a verb in child directed speech was not included.
of token frequency. The proportion of all my utterances spoken in the input that are my-noun/verb utterances was calculated, see Table 2.6.

A Spearman’s rho correlation was run between the proportion of my-noun/verb utterances as a function of all my utterances in the input (range: 0.000 – 0.667, $M = 0.147$, $SD = 0.159$) and the proportion of children’s pronoun-verb utterances that are my-for-I errors (range: 0.001 – 0.244, $M = 0.045$, $SD = 0.066$). There was a non-significant correlation $r_s = 0.054$, $N = 15$, $p = .849$, showing that higher levels of my-verb/noun utterances in the input are not related to higher rates of my-for-I errors.

![Figure 2.1](image)

Figure 2.1. Scatterplot of the fifteen children’s my-error frequency in relation to the my, am-I and my-noun/verb input frequency.

2.5 DISCUSSION

This paper explored the question of why some children make genitive-for-nominative first person pronoun errors. These my-for-I errors such as “my want that” occur around 2-4 years. A corpus analysis on the longitudinal data of fifteen children’s my utterances was calculated using every token of my-noun/verb over every token of my in the input files.

---

8 The proportion of all my utterances that were my noun/verb was calculated using every token of my-noun/verb over every token of my in the input files.
English speaking children available on CHILDES (MacWhinney, 2000) analysed the frequency of *my* and the frequency of *my* within various constructions in the input, to investigate if these *my-for-I* pronoun case marking errors can be explained by children hearing specific strings in the input, just as *me-for-I* errors have been partly explained by (Kirjavainen et al., 2009). The raw frequency of *my* in the input, as a function of all word tokens during the pre-error period did not relate to the subsequent rate of children’s errors. Nor was the proportional frequency of *my* in the input, related to the children’s rate of *my-for-I* errors. This is in line with past research (Kirjavainen et al., 2009) on *me-for-I* errors, which suggests it is not simply the rate of hearing the specific pronominal form in the input that influences children to make pronoun errors with that pronoun. Thus, we can conclude that children making *my-for-I* errors were not simply hearing *my* more (or less) in general, or relative to other self-referring forms in the input and consequently overextending its use in their utterances.

Previous research has shown that children who hear a higher proportion of *me-verb* utterances in the input as part of a complex sentence such as “Let me do it” have higher levels of *me-for-I* errors. In the same way, children hearing *am I-verb* in the input may think they have heard *my-verb* utterances leading to *my-verb* errors. However, no relation between this kind of input and children’s *my*-errors was found, and in fact *am I-verb* sequences were, in general, very infrequent in the input relative to other kinds of pre-verbal first person pronominal reference. Similarly, *my-noun* utterances in the input may appear to children as *my-verb* utterances, because some nouns can also be used as verbs. However, there was no relation between the rate of use of mixed noun/verb forms in the input and children’s *my-error* rates. It is therefore not the case that some children are hearing
more my-noun utterances in the input with words that they are also familiar with as
verbs and from these establishing my-verb constructions in their speech. Together,
these two analyses seem to rule out the possibility that children’s my-for-I errors
result from the direct learning of strings (either in a misheard “what am I doing?” or
in a misunderstood noun phrase such as “that is my drink”) from the input.

The current study shows that an explanation of my-for-I errors that is very
similar to that of me-for-I errors does not appear to exist and the frequency of my in
specific strings in the input does not relate to children’s rate of my-for-I errors.
Future work should attempt to investigate other theories of these genitive-for-
nominautive pronoun case marking errors. It has been suggested that input that
contains a mix of both person names (“Mummy/Mommy”, sibling and child’s
names) and pronouns, will assist children in their acquisition of first person
pronouns as it assists children in mapping pronouns on to the speaker (Smiley,
Chang & Allhoff, 2011). This could be explored by looking at the contrastive use of
pronoun and proper name in the input compared with children’s rates of my-for-I
errors. Further, Deutsch and Budwig (1983) and Budwig (1989; 1995) have claimed
that children do not always imitate what they hear from the input and can be quite
creative. This can cause them to use non-conventional self-reference terms in
subject position. It was claimed that children’s choice of pronominal reference can
be determined by pragmatic considerations and the use of my as subject was used to
mark their participation in actions high in agency and control. Future work
investigating children’s my-for-I errors should attempt to explore this functional
role of my in children’s errors. Further, Rispoli (1998b) claimed that children, who
had a higher level of correct me at the onset of their my-for-I error period, resulted
in making more me-for-I errors than my-for-I errors. Therefore it would be of
interest to investigate the child’s own prior usage of these first person pronouns in more detail in relation to their error rates.

2.5.1 Conclusion

The present study has looked at specific strings in the input and its relationship with children’s *my-for-I* pronoun case marking errors. Basic raw frequencies of *my* in the input were not related to children’s errors, as expected. Further exploration of the frequency of *my* almost pre-verbally, either misheard in an *am I-verb* utterance or in a *my-noun/verb* utterance were also not related to children’s error rates. These *my-for-I* errors cannot be explained in the same way that *me-for-I* errors can. Future work should attempt to investigate other possible theories; the contrasting input of pronoun and proper name and if this relates the children’s error levels; analysing children’s own prior pronoun usage; or investigating if children’s errors have a specific function.
2.6 REFERENCES


Chapter 3: Children’s my-for-I errors: A Corpus Analysis

Chapter three contains the second paper of the thesis. Chapter two included paper one which investigated whether children’s my-for-I pronoun case marking errors can be explained in a similar way as children’s me-for-I pronoun case marking errors. Children’s me-for-I pronoun case marking errors have been shown to be partly explained by the frequency of specific me-verb strings in the input. The first paper therefore explored the relationship between the frequency of my in the input and the frequency of specific strings containing my in the input and children’s level of my-for-I pronoun case marking errors. These factors did not relate to the rate of children’s errors and therefore children’s my-for-I pronoun case marking errors cannot be explained in the same way as me-for-I pronoun case marking errors. Paper two will take into consideration a range of other measures, derived from the literature and investigate the relationship between these factors and the rate of children’s my-for-I pronoun case marking errors. Given their lack of significance, the measures examined in paper one are not suitable for journal submission. Therefore, paper one will be summarised and included within paper two, to give a complete picture within a single paper, for publication purposes. Because of this, the first part of paper two will involve some degree of repetition of paper one, but will then go on to examine several other factors, in relation to children’s my-for-I pronoun case marking errors.
Children’s my-for-I errors: A Corpus Analysis

Stacey Mcknight\textsuperscript{1,2}, Elena Lieven\textsuperscript{2,3} & Anna Theakston\textsuperscript{2,3}

\textsuperscript{1}Max Planck Child Study Centre, University of Manchester, UK
\textsuperscript{2}School of Psychological Sciences, University of Manchester, UK
\textsuperscript{3}ESRC International Centre for Language and Communicative Development (LuCiD)
3.1 ABSTRACT

English-speaking children commonly produce case marking errors. Although constructivist approaches to language learning explain some error types with reference to the presence of similar strings in the input (e.g. “Me do it” from “Let me do it”), there is currently no satisfactory explanation for why children make other error types such as the use of my (possessive) in subject position (e.g. “My make it”). In Study 1, a corpus analysis on longitudinal data from fifteen, English-speaking, two-to-four-year-olds was conducted to investigate if other distributional aspects of children’s speech and their input can explain these errors. Children’s my-error rates were not related to the input frequency of: my; am-I+verb strings; or my+noun/verb strings. However, children receiving less pronoun modelling and more proper-name input made more my-for-I-errors. Further, the higher relative frequency of my and lower relative frequency of I in children’s speech before the onset of errors predicted children’s my-for-I-error rates. In Study 2, we investigated whether children use my pre-verbally to express high levels of agency and control (Budwig, 1989). For four children, the prior discourse context was analysed to establish the pragmatic force of my and I utterances and coded for three aspects of agency and control. All were found to predict the presence of my-for-I-errors demonstrating that my-errors mark the specific function of agency and control. These errors are therefore a result of a variety of factors; lack of pronoun modelling, a low entrenchment of the correct I form and the overextension of the possessive meaning of my.

KEYWORDS: pronoun case marking errors, acquisition, language input, constructivist approaches
### 3.2 INTRODUCTION

Children are known to produce a variety of errors in their early speech. One interesting type of common error is pronoun case marking errors, when a child uses one pronominal case, in place of another. For example, children produce utterances such as “Me open it” (me-for-I error, Grice: Budwig, 1995, p. 124), “My make red table” (my-for-I error, Nina: Vainikka, 1994, p. 268), “Her have some tea” (her-for-she error, Hannah: Aldridge, 1989, p.178) and “Him can’t see” (him-for-he error, Nina: Vainikka, 1994, p. 295). It has been suggested that these errors reflect the complexity of case selection (Maratsos, 1979). However, in looking specifically at English children’s pronoun case marking errors, it is clear that the rates of error across children vary dramatically and the types of errors that they produce also differ greatly. A high proportion of English children’s pronoun errors are made with first person accusative-for-nominative (me-for-I), genitive-for-nominative (my-for-I) and third person feminine accusative-for-nominative (her-for-she) forms (Budwig, 1989; Rispoli, 1998a; 1998b; Vainikka, 1994). In contrast, him-for-he, I-for-me and I-for-my errors are extremely rare (Rispoli, 1994; Moore, 1995). It therefore seems unlikely that these errors can be explained by children simply being confused over which pronoun to use in different contexts. The clear observed bias towards certain types of error suggests a more specific explanation is required (Rispoli, 1998a; 1998b).

There have been many attempts to explain English children’s pronoun case marking errors from a generativist perspective. Radford (1990) claimed that English children make pronominal case errors in their early speech, because their case assignment system, part of their Universal Grammar, has not yet matured: me-for-I errors are explained as instances of noun phrases lacking case with me being the...
preferred form of the first person singular in the lexicon. However, his theory provided no clear explanation for genitive subject *my-for-I* errors (although see Radford, 1998 for a possible explanation). Another, influential account of English pronoun case-marking errors is the Agreement/Tense Omission Model (ATOM) (Schutze & Wexler, 1996; Wexler, 1998). This generativist theory argues that children do have case assignment mechanisms from the beginning stages of language acquisition, but errors occur when children fail to check for both tense and agreement in their utterances. Schutze & Wexler (1996) argued that when children make pronoun case errors, they use a default pronoun form which, in English, is the accusative case (*me, her, him*). Although this approach can potentially explain why accusative-for-nominative pronoun errors are so frequently observed, it does not explain the variability across children, or why children frequently produce genitive-for-nominative pronoun errors, in apparent contradiction of the claim for a single default form. Schutze (1997) further argued that genitive subjects are possible in sentences lacking in both tense and agreement, however examples such as “*My taked it off*” (Jeffrey, Budwig, 1995, p.73) in which tense is clearly present, challenge this theory.

Another theory of case marking errors is Rispoli’s (1994; 1998a; 1998b; 1999) Paradigm Building Model, which suggests that children’s linguistic systems are made up of cells defined by person, gender, number and case, with each pronominal form filling a particular cell(s). For example, the form ‘*He*’ represents the third person, singular, masculine form in nominative case. Rispoli (1998b) argued that errors are more likely to occur when children are learning this paradigm. During this time they may attempt to refer to particular referents where the form required is one they are still in the process of fully learning, resulting in
errors or the overuse of other forms. One advantage of this approach is its potential to explain the different error rates observed with different pronoun forms. The high observed rate of *her-for-she* errors is explained by the ‘double cell effect’, as the third person feminine pronoun *her*, fills two of the three available cells for third person singular feminine reference (one that corresponds to the nominative form ‘*she*’, one to the accusative ‘*her*’ and the third to the genitive ‘*her*’), facilitating its retrieval from memory. Similarly, Rispoli (1994; 1998b) offered a phonological explanation for the frequently observed first person singular case marking errors, pointing out that two out of the three possible choices for a first person pronoun, *me* and *my*, share an initial phoneme and *I* is phonetically the ‘odd one out’. He claimed that when children are unsure of the correct pronominal form, they will choose one of the two phonetically similar pronoun choices in substitution of the odd one out. Finally, Rispoli (1998a:390) attempted to explain why different children demonstrate different error patterns. He argued that “*me-for-I* and *my-for-I* replacements are antagonistic” with children having a preference for one or other error type. He found that children who produced high levels of correct *me*, when *my-for-I* errors first emerged, were more likely to make *me-for-I* errors than *my-for-I* errors, whereas children who produced few correct *me* utterances at the onset of *my-errors*, were likely to favour *my* in their *I-replacements*, supporting a previous case study by Vainikka (1994). This finding is particularly interesting because it suggests that children’s pronoun use prior to the emergence of any errors may have an impact on the type of subsequent errors.

This claim is broadly consistent with usage-based approaches to language learning, outlined in more detail below, in that children’s existing linguistic knowledge is thought to impact on the learning trajectory for later acquired forms
(e.g. see Ambridge, Kidd, Rowland & Theakston, 2015, on frequency effects).

However, Rispoli’s (1998a) and Vainikka’s (1994) studies focussed on children’s preference for a particular error type, but not on their individual error rates. From a usage-based perspective, it could be that children’s pronoun use prior to the onset of errors influences their error rate too, because well-established forms can either promote or protect against error (Ambridge et al., 2015). For example, one possibility is that children who produce a lot of correct my at the onset of errors but relatively less use of I, and for whom my is consequently more readily available, produce a higher proportion of my-for-I errors relative to correct I-verb sequences. In contrast, children who do not favour my may produce more correct utterances. In the current study we investigate the relation between children’s early pronoun use and their subsequent my-for-I error rates to cast light on the mechanisms underlying errors.

An alternative approach has been to look for the source of case errors in the input that children are exposed to. The constructivist approach to language acquisition attempts to address the individual variation in error patterns, and is based on the view that children learn their native language from observing patterns and making generalisations from the input (Bannard & Matthews, 2008; Theakston, Lieven, Pine & Rowland, 2001; 2002; Theakston, Maslen, Lieven & Tomasello, 2012; Theakston, Ibbotson, Freudenthal, Lieven & Tomasello, 2015). Differences in children’s input are expected to result in differences in children’s patterns of learning and error (Theakston, Lieven & Tomasello, 2003). For example, there is growing evidence, across a variety of language domains, that less frequently heard constructions are more likely to attract errors (see Ambridge, et al., 2015, for an overview; also Ambridge, Pine, Rowland & Young, 2008; Brooks, Tomasello,
Lewis & Dodson, 1999 for argument structure overgeneralisation errors; Rowland & Pine, 2000 for wh-questions; Rowland & Theakston, 2009; Theakston & Rowland, 2009; Theakston, Lieven, Pine & Rowland, 2005 for auxiliary verb constructions). Furthermore, some error patterns are thought to derive from strings heard in the input, which are then used in children’s speech independently from the larger construction in which they typically occur. For example, Tomasello (2000; 2003) suggested that children’s accusative-for-nominative pronoun case marking errors might be explained by their exposure to input containing me or her in pre-verbal position, within certain kinds of complex sentences, such as “Let her do it” or “Help me find it”, leading children to produce pronoun errors such as “Her do it” or “Me find it”. One possibility is that this reflects children learning from the ends of utterances due to a more general bias to remember things heard most recently, and only over development are the longer, more complex utterances learnt and errors cease (Freudenthal, Pine, Aguado-Orea & Gobet, 2007). In support of this claim, Kirjavainen, Theakston & Lieven (2009), in a corpus analysis of data from seventeen two-to-four year old, monolingual English-speaking children, found that the proportion of first person pre-verbal pronouns that were me in the input, such as “Let me do it” and “Did you see me doing it?” correlated with the proportion of me-verb errors that the children produced. Furthermore, the specific verbs used in these strings in the input were more likely to be produced erroneously in me-verb sequences in the children’s speech. However, there was no relation between the proportional use of me (as a function of total accusative and nominative pronouns) in the input and children’s me-for-I errors. This demonstrates that the children were not using me erroneously simply because they heard me used frequently in the input, but rather were producing errors based on specific strings they heard. A
number of other studies have found similar relations between strings in the input and children’s early errors for third person verb marking errors (Croker, Pine & Gobet, 2003; Theakston et al., 2003), infinitival-to omissions (Kirjavainen et al., 2009; Kirjavainen & Theakston, 2011), auxiliary verb omissions (Theakston & Lieven, 2005; 2008), questions (Rowland & Pine, 2000; Ambridge, Rowland, Theakston & Tomasello, 2006; Ambridge & Rowland, 2009), and auxiliary use in SLI children (Leonard & Deevy, 2011). However, there is also evidence that errors can be caused by children defaulting to the form heard most frequently in the input, (Räsänen, Ambridge and Pine, 2013) without the presence of a specific string.

Although a number of early child errors, including accusative-for-nominative case errors in English, can be explained in terms of input-based learning, there are no analogous input-based explanations for other observed pronoun case errors such as genitive for nominative my-for-I errors (“My do it”), as children do not hear complex sentences in the input containing my in pre-verbal position. These errors therefore pose a challenge for constructivist approaches to language learning as, currently, they have no clear explanation within this framework. In this study we present a systematic analysis of a number of possible input-based explanations for genitive-for-nominative errors within a constructivist framework in an attempt to cast light on the processes involved.

First we consider possible sources for my-for-I errors in the input in the form of phonological strings. Children may think that they are hearing my in a pre-verbal position in the input as certain constructions can appear this way. For example, some nouns are also verbs and therefore a child hearing a my-noun sequence such as “my drink”, may think that they have heard a my-verb utterance. Or children hearing questions such as “What am I doing?” or “Where am I going?”
could also think they have heard a *my*-verb utterance as the sequence *am I* is phonologically very similar to *my* (e.g. *Where’re my shoes?*), especially when produced quickly and fluently in questions, with stress falling after the onset of ‘*am*’.

Second, we will consider frequency-based distributional factors. As well as pronoun case marking errors, about twenty percent of children commonly make pronoun reversal errors, such as using *I* and *me* instead of *you* and vice versa (Dale & Crain-Thoreson, 1993). Evidence shows that children with siblings find first and second person pronouns less confusing, and second born children are more advanced than firstborns in their pronoun production (Oshima-Takane, 1988; 1999), which has been argued to reflect their observation of adult-child interactions modelling the contrastive use of pronoun forms such as *I/my* and *you*. Thus, the relative frequencies with which children are exposed to particular form-function mappings appears to be relevant in determining the acquisition of pronoun forms, at least for first and second person pronouns. Smiley, Chang and Allhoff (2011) argued that the specific pattern of pronoun and proper noun use in the input determines the speed with which children work out correct pronominal reference.

They proposed that input that contains a mix of both person names (‘*Mummy/Mommy*’, sibling and child’s names) and pronouns, will assist children in their acquisition of first and second person pronouns as it allows “children to more easily map pronouns to the conversational roles of speaker and listener” (Smiley et al., 2011; p.82). Their longitudinal analysis of naturalistic corpus data from young children provided some support for this claim. Children who received ‘high-contrast’ input, defined as less than 90% pronoun use alongside use of proper nouns (of first and second person subject reference), acquired first and second person
pronouns *I* and *you* earlier than those whose input was more heavily dominated by pronominal forms. Further, those children not receiving this beneficial contrast become innovative in their utterances and used their own name as well as the first person pronouns *my* and *me* erroneously in place of the correct nominative pronoun *I*, before they acquired the correct subject pronoun.

However, there are limitations with this analysis. The results showed that only children (*N*=2) who received high contrast input for both first and second person subject reference made fewer errors; there was no difference in correct first person reference between those who received high vs. low contrast first person input, alongside low contrast second person input. Smiley et al. concluded that high contrast of both *I* and *You* was necessary to master the system, but this conclusion may be premature given that they only considered utterances in action contexts, potentially excluding a lot of important input data. Furthermore, the relation between input contrast and individual error types was not investigated in any detail. It is therefore unclear whether the effects of input contrast operate across the board to promote correct use of *I*, or whether the relation between input contrast and particular error types might vary. In this study we investigate whether children’s *my-*for-*I* errors reflect a low or high contrast pattern in first person pronoun and proper name input.

Finally, a different approach, based more loosely on the input, is to consider that each form used by the child has a different function and this may lead to errors (Karmiloff-Smith, 1979). Children do not always imitate exactly what they hear from the adult system, but can be creative in their early speech (e.g. see Cameron-Faulkner, Lieven & Theakston, 2007, on early negation). Deutsch and Budwig (1983) noted children’s contrastive use of specific linguistic forms based on the
pragmatic force of the utterance (see also Carter, 1975). Budwig (1989; 1995) specifically investigated children’s pronoun case marking errors by studying the pragmatic distinctions between the pronominal forms that three children aged between 1;8 and 2;6 used in their utterances. These children had multiple non-conventional self-reference terms, and commonly made pronoun errors, using me, my and also own name, in subject position. A detailed semantic analysis, concentrating on the degree of agency expressed in the children’s utterances as well as the children’s level of control, suggested that the children’s pronominal reference was determined by pragmatic considerations: proper names such as “Ruth do it” were used when the children were involved in ‘intransitive physical actions’; the use of my as subject was used to mark their participation in actions high in agency and control; and conventional self-reference I was used in terms that were low in agency. However, although these results suggest that children’s early pronoun use may be pragmatically governed, the children in Budwig’s study often also used different forms for the same function. Furthermore, coding reliabilities were not obtained and there were no tests of significance provided. Thus, a more systematic examination of this proposal is needed to determine whether pragmatic and semantic functions play a role in children’s pronominal case errors.

3.2.1 The Current Study

The present study is an investigation into English children’s genitive-for-nominative first person pronoun errors. In Study one, a corpus analysis was carried out to investigate if any constructivist theories of pronominal case marking errors can explain my-errors. Various aspects of the input were examined to see if the language that children are exposed to shortly before the onset of these errors has
any effect on their prevalence in children’s speech. Further, the children’s pronoun use before the onset of errors was analysed to uncover any relationships between the two variables. In Study two, Budwig’s (1989) claim that children use my preverbally to express high agency and control was investigated.

3.3 STUDY 1

The aim of Study one was to determine whether the rate of children’s my-errors can be explained by various aspects of the language that children are exposed to, or produce, shortly before the onset of these errors.

3.3.1 METHOD

3.3.1.1 Corpora

A corpus analysis was conducted on the longitudinal data of fifteen children available on CHILDES (MacWhinney, 2000). These children included the twelve children from the Manchester Corpus (Theakston, Lieven, Pine & Rowland, 2001). One child was excluded\(^9\) and the remaining eleven children were included in the study, although only two produced a high number of my-for-I errors. In order to include more children who make my errors, the CHILDES database was searched and four further children were included in the analysis: Thomas (Lieven, Salomo, Tomasello, 2009); Betty (Wells, 1981); Nina (Suppes, 1974); Laura (Braunwald, 1997).

\(^9\)John made no my-errors after he had produced three correct uses of my and was excluded from the analysis.
3.3.1.2 Determining the error period

Every child utterance containing *my* was extracted and coded for the lexical item that followed *my* and its grammatical category. Because different amounts of speech were available for the different children, *my-error* rates were calculated as a proportion of all first person pronoun+verb, auxiliary, modal, negative-verb or adverb-verb utterances. That is to say, how many of the children’s verb utterances are *my*-errors out of all *my-errors*, *me-errors* and *correct-I* uses in nominative contexts?

An error period was defined for each child, by first establishing that s/he had correctly used *I* and *my*. This was to eliminate the possibility that pronoun errors simply reflect the child’s lack of knowledge of the correct form. The file by which each child had produced three correct uses\(^{11}\) of *I* and *my* was identified and the first *my-for-I error* after the three correct uses was taken as the start of the error period. The error period then ran up until each child’s last *my-for-I error* plus two additional files when possible. From this defined error period, each child’s total number of *my-for-I verb* errors, *me-for-I verb* errors (i.e. incorrect *my/me-verb/auxiliary/modal/neg-verb/adverb-verb* utterances, for example “*My not do it*” or “*Me just go there*”) and total number of *I-verb* utterances was calculated, as well as their total frequency of use of the form *my* (see Table 3.1). The proportion of *my-\(^{10}\)

---

\(^{10}\)Every reference to child or parental *I-verb* utterances, also including the contracted forms of *I’m, I’ll, I’ve* and *I’d.*

\(^{11}\)The three correct uses had to be three unique utterances with three different lexical items. Utterances of the single pronoun alone were not counted, nor were idioms, songs or interjections. Uses of *I* in contracted forms such as *I m, I’ve, I’ll or I’d* were counted as a correct use of *I*, as long as they were used with another lexical item and not just as “*I’m*” on its own. “*I’ll go*” and “*I’d go*” were considered two unique utterances. Any utterances coded with a [?] were not included as a correct use.
errors and me-errors from all first person pronominal verb utterances\textsuperscript{12} was then established.

\begin{table}
\centering
\begin{tabular}{llllll}
\hline
Name & Error period & Total & my-for-I & me-for-I & I-verb & Percent \\
& Age (years; months.days) & & my-for-I verb & errors & errors & my-for-I / me-for-I + I-verb \\
\hline
Anne & 2;2.10 to 2;8.24 & 308 & 5 & 28 & 866 & 0.6\% \\
Aran & 2;3.15 to 2;10.28 & 372 & 2 & 2 & 2050 & 0.1\% \\
Becky & 2;2.22 to 2;11.15 & 243 & 8 & 7 & 2123 & 0.4\% \\
Carl & 1;11.29 to 2;8.15 & 182 & 22 & 2 & 2103 & 1.0\% \\
Dom & 2;4.04 to 2;8.00 & 248 & 6 & 8 & 889 & 0.7\% \\
Gail & 2;3.17 to 2;11.05 & 357 & 15 & 3 & 930 & 1.6\% \\
Joel & 2;3.04 to 2;9.13 & 320 & 4 & 7 & 1163 & 0.3\% \\
Liz & 2;1.25 to 2;7.03 & 409 & 9 & 13 & 1103 & 0.8\% \\
Nicole & 2;8.20 to 3;0.10 & 182 & 4 & 21 & 228 & 1.6\% \\
Ruth & 2;7.17 to 2;11.21 & 312 & 53 & 342 & 364 & 7.0\% \\
Warren & 2;0.03 to 2;9.20 & 465 & 98 & 10 & 1169 & 7.7\% \\
Nina & 2;0.03 to 2;10.21 & 915 & 155 & 14 & 1672 & 8.4\% \\
Thomas & 2;7.23 to 3;4.04 & 2409 & 118 & 13 & 6511 & 1.8\% \\
Betty & 2;3.02 to 3;3.02 & 133 & 60 & 15 & 171 & 24.4\% \\
Laura & 1;6.07 to 2;6.04 & 1719 & 316 & 36 & 2475 & 11.2\% \\
\hline
\end{tabular}
\caption{Defined my-error period for the fifteen children with total number of my-for-I errors, me-for-I errors and correct I-verb utterances during this period.}
\end{table}

\textbf{3.3.1.3 Input Data}

In order to investigate children’s pronoun errors, the input files to be used in the analysis were defined (see Table 3.2).

\textsuperscript{12}The children’s first person verb utterances include just me, my and I as mine does not occur in the data except for one instance, which has no impact on any of the results.
The number of data files for each child’s error period varied from child to child, dependent on the density of recordings and the duration of the error period. Therefore, the decision was taken to select the number of input files for each child to; (1) vary in proportion with the number of files in the error period, (2) cover at least one month’s worth of recordings prior to the error period, (3) cover no more than two months of input data, and (4) overlap into the error period by no more than two weeks. For all subsequent analyses, the children’s data were taken from the error period, and the input data was taken from the associated input period, as defined above. Input data was predominantly from the mother, but included any caregiver interaction and in some recordings the father was present.
3.3.2 RESULTS

3.3.2.1 Frequency of my in the input

We first investigated whether the overall amount that children hear the pronoun my in the input influences their overall my-for-I error rate. The raw frequency of all first person pronouns in the input files was extracted. As the amount of input available for each child differed, both the proportion of my as a function of all word tokens, and as a function of all first person pronouns (me, my, mine and I) in the input, were calculated. The data violated the assumptions for parametric tests and therefore rank-order correlations were conducted throughout. A Spearman’s rho (two-tailed) test was run to determine if there was a relation between the proportional use of my in the input as a function of either all word tokens (range: 0.000 – 0.003, M = 0.001, SD = 0.001), or all first person pronominal reference (range: 0.012 - 0.130, M = 0.056, SD = 0.031), and the proportion of pronoun+verb utterances that contain my-for-I errors in the children’s speech (range: 0.001 – 0.244, M = 0.045 , SD = 0.066). However, both relations were non-significant, $r_s = 0.196$, $N = 15$, $p = 0.483$ and $r_s = 0.196$, $N = 15$, $p = .483$ respectively. Thus there was no indication that the frequency with which caregivers used my overall or as a proportion of all first person pronoun use was related to children’s rates of my-for-I errors.

3.3.2.2 Strings in the input: “What Am I Doing?”

Children may hear the construction am I within sentences in the input, in a pre-verbal position (e.g. What am I doing?), which could sound to the child just like my (e.g. What my doing?), thus licensing the use of my pre-verbally. Therefore, we examined whether am I-verb constructions in the input have an effect on children’s
my-verb error rates. The proportion of am-I verb utterances from the total number of am I-verb and all other pronoun+verb input utterances was calculated. A Spearman’s rho correlation to determine if there was a relation between the proportional use of am I-verb in the input (range: 0.000 – 0.036, M = 0.006, SD = 0.009) and the proportion of pronoun+verb child utterances that contain my-for-I errors (range: 0.001 – 0.244, M = 0.045, SD = 0.066) was non-significant $r_s = -0.133$, $N = 15, p = .636$. This indicates that children’s my-for-I error rates were not related to the rate of hearing am I-verb constructions in the input.

### 3.3.2.3 Strings in the input: My-noun/verb

Many nouns such as drink and drawing are also used as verbs. Therefore children may hear these lexical items used in their nominal form after my, in the input, for example “Do you like my drawing?” or “That is my drink”. One possibility is that the child then uses these same lexical items after my, but in verb form, for example “My drawing a picture” or “My drink all my milk”, resulting in my-verb errors. All lexical items following my in the parental input were coded as nouns only or as nouns/verbs if they can be grammatically used as both. For an item to be considered both a noun and verb in child directed speech, at least one instance of the noun being used as a verb had to be found in the entire corpus of input data, across all files, so that there was at least one example of the lexical item being used as a verb in child directed speech. The proportion of all my utterances spoken in the input that contained a post-my element classified as a mixed noun/verb was calculated separately for each child’s input. A Spearman’s rho correlation was run between the proportion of my-noun/verb utterances as a function of all my utterances in the input (range: 0.000 – 0.667, M = 0.147, SD = 0.159) and the
proportion of children’s pronoun-verb utterances that are my-for-I errors (range: 0.001 – 0.244, $M = 0.045, SD = 0.066$). There was a non-significant correlation $r_s = 0.054, N = 15, p = .849$, showing that higher levels of my-verb/noun utterances in the input are not related to higher rates of my-for-I errors.

### 3.3.2.4 Pronominal Modelling

Smiley et al. (2011) argued that children who do not receive enough informative input regarding the referent associated with first and second person pronouns become innovative when they are unsure of the correct pronoun to use, causing them to make my-errors and me-errors, or even to rely on using their own-name pre-verbally. To investigate whether the distribution of pronominal and proper name forms in the input might influence children’s rates of my-errors, the children’s data and their input was searched for own name-verb utterances and self-referring terms such as Mummy/Mommy$^{13}$/Daddy$^{14}$ respectively. This included Mummy (as object, e.g. Give that to Mummy instead of Give that to me); Mummy (as subject, e.g. Mummy do it, instead of I do it); Mummy’s (as possessive in a noun phrase, e.g. It’s Mummy’s ball, instead of It’s my ball); Mummy’s (as possessive e.g. That’s Mummy’s instead of That’s mine); Mummy’s (as subject, e.g. Mummy’s doing it instead of I’m doing it).

First, Spearman’s rho correlations were run to determine whether the proportion of pronoun vs. proper name input, both overall and in subject position only, was related to the proportion of correct utterances ($r_s = -.054, N = 15, p$

---

$^{13}$The appropriate mummy/mommy term was searched for depending on whether the child was from the ENG-UK or ENG-USA database.

$^{14}$All utterances such as “Mummy’ll do it” were included in the appropriate category, just in the same way as I’ll, I’m, I’d and I’m. Any input that was Mummy/Daddy but not a self-referring term, or that was ambiguous, was not included in the count, for example, “a mummy crocodile, is it?”
= .850 and \( r_s = -0.157, N = 15, p = .576 \), respectively) and my-for-I errors in children’s first person verb utterances \( (r_s = 0.357, N = 15, p = .191 \) and \( r_s = -0.046, N = 15, p = .869 \) respectively). All correlations were non-significant, showing that there is no direct linear relationship between the contrast of proper name and pronoun use, either overall or specifically in subject position, and children’s correct subject-verb utterances or my-verb errors.

Following Smiley et al. (2011), the children were then categorised as receiving high (over 90% pronoun use) or low (less than 90% pronouns) contrastive input, based on all first person pronominal and nominal input. We conducted two Mann-Whitney tests to compare (a) the proportion of correct I-verb utterances, and (b) the proportion of my-verb errors, as a function of all subject-verb utterances (including my, me, I and own name) between children receiving high \( (N = 9) \) and low \( (N = 6) \) contrast input. There was no significant difference in the proportion of utterances that were correct I-verb utterances between the two groups, \( Z < 1, ns \), but there was a significant difference between the two groups in the rates of my-verb errors, \( Z = -2.36, p = .018 \). Children receiving less than 90% pronoun use and therefore a higher contrast, produced a higher rate of errors \( (M = 7.02\%) \) compared to those receiving low contrast input \( (M = 0.72\%) \). A further analysis was run with children’s input categorised as high or low contrast specifically on the subject position referents in their input and also on their possessive input, as children receiving high contrast of my modelled in possessive position may not receive enough correct informative modelling of the pronoun my, which may lead them to use this erroneously. But both these analyses revealed no difference in my-error rates between those categorised as receiving high and low contrast input, \( Z = -1.27, p = .203 \) and \( Z = -1.16, p = .248 \) respectively.
3.3.2.5 *Onset of Errors*

The pronoun use in children’s speech before the onset of errors was analysed to establish if there was a relationship between the entrenched use of one pronoun and the level of *my-errors*. The total number of *I, me* and *my* in all utterance positions was extracted from all available data files, (averaging at approximately three months of data) before the error period for each individual child.

A theoretical challenge is to establish exactly which forms should be available to compete for production under any given circumstances in the child’s developing linguistic system. One possibility is that all first person reference forms compete, and the relative frequency of each should determine which form is produced in any first person context, in a probabilistic manner. Clearly, this is not representative of the adult end-state, in which particular forms are associated with specific grammatical roles and functions. Another is that the availability of forms is determined both by their relative frequencies, and their form-function mappings, such that forms that are tightly associated with other functions are less activated, and the relative frequencies of remaining forms should determine production. Currently we have insufficient evidence to distinguish between these possibilities in the child’s developing linguistic system. Thus, we conducted analyses to examine the impact of both relative frequency (of all first person forms), and relative frequency (of pairs of available forms) to better understand when *my-errors* are produced.

First, a Spearman’s rho correlation was run between the children’s proportional use of *I* of all first person pronouns before the error period and their
proportion of *my-for*-I errors during the error period. This was significant, \( r_s = -0.67, N = 15, p < .01 \), showing that children who use *I* relatively infrequently in comparison to other first person pronouns are more prone to *my-for*-I errors.

However, a further Spearman’s correlation to determine whether the proportional use of *my* of all first person pronouns before the error period was related to *my-for*-I error rates was non-significant, \( r_s = 0.31, N = 15, p = .254 \), showing that children who make more *my-for*-I errors are not simply those who use *my* relatively more often than other first person pronouns overall.

Second, we tested Rispoli’s (1998) claim that children who produce high levels of correct *me* are less likely to be ‘*my-error* children’ because they will instead favour use of *me* in their errors. However, we found no relation between either children’s raw frequency of *me*, \( r_s = 0.004, N = 15, p = .990 \), or its use relative to other first person pronouns, prior to the error period, \( r_s = 0.25, N = 15, p = .369 \), and their subsequent *my-error* rates. This, coupled with the previous analysis, suggests that although Rispoli found *my-errors* and *me-errors* to be antagonistic, the frequency of use of these pronouns relative to all other first person pronominal reference prior to the onset of *my-errors*, does not determine the error rates observed.

Third, we investigated whether there might be direct competition between *I* and other specific first person pronouns that related to the children’s error rates. Spearman’s rho correlations between the proportional use of *I* as a function of all uses of *my* and *I*, and children’s *my-error* rates revealed that children who produced a high rate of *my* relative to *I* had higher *my-error* rates \( (r_s = 0.52, N = 15, p < .05) \). In contrast, there was no relation between the relative use of *I* vs. *me*, and
subsequent my-error rates ($r_s = 0.25, N = 15, p > .05$). Thus, prior knowledge of my and me seem to have independent effects on the subsequent my-error rate.

Finally, one possibility is that those children who acquire the pronoun my before I may be the children who begin to later produce a higher number of my-for-I replacements. Unfortunately, many of the children were producing both pronouns by the earliest available data file and therefore no detailed analysis could be run.

Note, however that when we look at only those children for whom my and I are first observed at different developmental points, the three children with the highest error rates, Nina, Thomas and Laura, all produced my before I whereas the two children with the lowest error rates, Becky and Joel, produced I before my.

3.3.3 DISCUSSION

Study one investigated several constructivist approaches to explaining children’s my-for-I pronoun case marking errors with the aim of identifying if any input driven theories can explain these errors, and/or if the individual children’s pronoun use relates to the observed error rates.

The raw frequency of my in the input, as a function of all word tokens during the pre-error period did not relate to the subsequent rate of children’s errors. Nor was the proportional frequency of my in the input, either as a function of all first person pronouns, or of all first person referents (including own name) related to the children’s rate of my-for-I errors. This is in line with past research (Kirjavainen et al., 2009) on me-for-I errors, which suggests it is not simply the rate of hearing the specific pronominal form in the input that influences children to make pronoun errors with that pronoun. Thus, we can conclude that children making my-for-I errors were not simply hearing my more in general, or relative to
other self-referring forms in the input and consequently overextending its use in their utterances.

Previous research has shown that children who hear a higher proportion of *me-verb* utterances in the input as part of a complex sentence such as “*Let me do it*” have higher levels of *me-for-I* errors. In the same way, children hearing *am I-verb* in the input may think they have heard *my-verb* utterances leading to *my-verb* errors. However, no relation between this kind of input and children’s *my*-errors was found, and in fact *am I-verb* sequences were, in general, very infrequent in the input relative to other kinds of pre-verbal first person pronominal reference. Similarly, *my-noun* utterances in the input may appear to children as *my-verb* utterances, because some nouns can also be used as verbs. However, there was no relation between the rate of use of mixed noun/verb forms in the input and children’s *my-error* rates. Together, these two analyses seem to rule out the possibility that children’s *my-for-I* errors result from the direct learning of strings from the input.

Further analysis looked at the rates of pronoun input in parental speech, in relation to *proper name* input. Smiley et al. (2011) suggested that receiving a high contrast of pronoun and proper name referents benefits children in learning *I* and *You*, consequently leading to fewer errors. They argued that children who have fewer opportunities to observe pronouns and proper names used contrastively take longer in mastering pronominal reference and may become innovative, leading to pronoun case marking errors. However, our analysis showed that children receiving high contrast (below 90% pronouns) for all first person pronominal input had higher levels of *my-for-I* errors. There are several differences between the studies which may explain these results. For example, the input samples varied, with
Smiley et al. focussing exclusively on action contexts whereas all contexts were included in the current study. In addition, Smiley et al. focussed on children’s correct use of I and observed a variety of error types in their children’s data, whereas we focus on one specific error-type. Finally, Smiley et al. (2011) found that only children who received high contrast in both first and second person reference (I/own-name and you/child’s name) showed more accurate use of the first person pronoun I, whereas those who only received high contrast in first person reference, the focus of the current study, showed higher error rates. The present analysis shows that children receiving a higher level of proper name input in first person self-reference are the children who had a higher rate of my-for-I errors, suggesting that higher contrast per se is not necessarily beneficial. Furthermore, our results hold when considering all first person reference together, but not if we look only at subject reference (I), or possessive reference (my) in isolation. One possibility is that these children were not receiving sufficient pronoun modelling across the range of first person forms to establish the unique distributional patterning of each, and therefore became innovative in their utterances.

Finally, we examined the levels of pronoun use by each individual child before the onset of the error period. We found that children who have a lesser entrenched use of I as a function of all first person pronoun use before the onset of errors, make a higher proportion of my-for-I replacements during the error period. Neither the relative use of my nor me as a function of all first person pronoun reference predicted the children’s my-error rates, suggesting that children are not simply producing a single strongly entrenched form. Note, that Rispoli’s (1998a) findings in relation to early pronoun use; that an early use of me was related to high me-error rates, concerned a preference for a particular error type, rather than to the
absolute rate of errors observed. However, when we examined direct competition between pairs of pronominal forms, we found that children who had higher levels of *my* relative to *I* in the pre-error period were those who subsequently made more *my-for-I* errors. In contrast, there was no relation between children’s use of *me* relative to *I* during the pre-error period and their subsequent *my-for-I* error rate. Thus, prior knowledge of different pronoun forms seems to have an independent influence on subsequent error rates.

In summary, Study 1 showed that children who had a high proportion of *proper name* input relative to pronominal input, made a higher proportion of *my-for-I* errors. Further, children who produced *I* relatively less frequently than other first person pronouns, who favoured *my* over *I* in the pre-error period, and who apparently learned *my* before *I*, were the children who subsequently made a higher proportion of *my-for-I* errors. Thus, an early preference for *my* over *I*, and a lack of modelling of *I* and *my* in the input may lead children to be unsure over which pronoun to use, causing them to become innovative in their language. In this innovation, children use erroneous forms and it has been claimed that these forms may be assigned individual functions (Budwig, 1989). In Study 2 we consider the functional explanations of *my-for-I* pronoun case marking errors.

### 3.4 STUDY 2

The aim of Study two was to investigate Budwig’s (1989) claim that children use *my* pre-verbally to express high agency and control. Four children who made a high number of *my-for-I* errors were included in the analysis. Every target utterance was analysed in context to code for the function of *my-verb* errors in comparison with correct *I-verb* utterances.
3.4.1 METHOD

3.4.1.1 Corpora

To investigate the function of children’s my-for-I errors it was necessary to have a direct comparison with their correct I-utterances, that is, I-verb utterances available from the same recording files as the errors. This was to establish whether children specifically use my in place of I to express a certain function. The children identified as making the highest number of my-errors in Study one were Ruth, Warren, Nina, Thomas, Betty, and Laura (see Table 3.1). However, for Study 2, of these 6 children; Betty could not be analysed as her errors occurred in two single recordings and there was not enough I data for comparison; and Laura’s data comprised of audio data with input, and diary data without input, making it impossible to determine the context, and consequently the function, for many of her errors (although sufficient input data was available to establish caregiver rates of pronoun use for Study 1). Therefore Ruth, Warren, Nina and Thomas were the four children included in the detailed functional analysis.

3.4.1.2 Error Period

The period in which children made my-for-I errors was defined as in Study 1 with the additional criteria that there had to be a matching I-utterance for comparison with each my-error. However, Nina produced some my-errors at the start of her error period for which there were insufficient matching I-utterances for comparison. Therefore, these errors were excluded from the analysis. To explore the function of the errors further and across the error period, we identified an early and a late stage in each child’s error period, comprising the first 30 errors with recording-matched I-utterances, and the last 30 errors within the defined error
period for each child, or the remaining utterances if the child produced fewer than 60 errors in total (Ruth), see Table 3.3. This was to determine whether children start off using these errors to serve a certain function, but then generalise this use.

Table 3.3. Error period for the children in Study 2.

<table>
<thead>
<tr>
<th>Name</th>
<th>Early Stage of Errors (Age years;months.days)</th>
<th>Late Stage of Errors (Age years;months.days)</th>
<th>Total my-verb errors included in analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ruth</td>
<td>24-30 (2;7.17 to 2;9.23)</td>
<td>31-34 (2;10.9 to 2;11.21)</td>
<td>53</td>
</tr>
<tr>
<td>Warren</td>
<td>6-16 (2;0.3 to 2;3.8)</td>
<td>27-33 (2;7.8 to 2;9.16)</td>
<td>60</td>
</tr>
<tr>
<td>Nina</td>
<td>5-10(^{15}) (2;0.10 to 2;1.15)</td>
<td>17-35 (2;3.14 to 2;10.6)</td>
<td>60</td>
</tr>
<tr>
<td>Thomas</td>
<td>2;07.23 to 2;09.10</td>
<td>2;10.21 to 3;04.2</td>
<td>60</td>
</tr>
</tbody>
</table>

3.4.1.3 Coding

For each child, for every *my-for-I verb* error during the early and late stages, a correct *I-verb* utterance was chosen at random from the same data file. Every *my-for-I* error and matching *I-utterance* for both stages then became the target utterances. Every target utterance was analysed in context to determine its function, by extracting the ten utterances before and after every target.

There were a small number of instances in which these utterances were not sufficient to establish the context and in such situations, twenty prior utterances were examined, (see Figure 3.1). In this example, because Warren’s mother was

\(^{15}\)Nina had a total of 30 *my-for-I* errors in files 5 and 6 alone. However, only 3 of these errors were used and her first 30 errors were from files 5 to 10, as these had matching *I-utterances*. 
counting for the ten prior utterances, it was necessary to look further back to
establish the context.

**Utterance**

8  warren10b: line 860.
   1 *MOT:  oh six pounds ?
   2  *CHI:  yeah .
   3 *MOT:  do you need some money back ?
   4 *MOT:  do you need some change ?
   5  *CHI:  Warren change .
   6 *MOT:  oh well thank+you very much .
   7  *CHI:  give the money .
   8 *MOT:  I'll go and put the six pounds in that nice money box over here .
   9 *MOT:  one .
  10 *MOT:  chink .
  11 *MOT:  two .
  12 *MOT:  chink .
  13 *MOT:  three .
  14 *MOT:  chink .
  15 *MOT:  four .
  16 *MOT:  chink .
  17 *MOT:  five .
  18 *MOT:  chink .
  19 *MOT:  six .
  20 *MOT:  chink .
  21  *CHI:  my [*] do it now .
  22  *MOT:  listen to all that money .
  23  *MOT:  are you going to put some money in ?
  24  *CHI:  &a:h .
  25  *MOT:  thank+you .
  26  *MOT:  all the money's away safe .
  27   *CHI:  buy shopping again .
  28  *CHI:  buy shopping again .
  29  *MOT:  go shopping again .
  30  *MOT:  all right .
  31  *MOT:  <what're you going to have> [/] what're you going to buy today ?

Figure 3.1. Example target utterance with context. This particular utterance is one
which the prior ten utterances do not give enough information to understand the
context, so the analysis was taken back to twenty utterances.
In order to identify the pragmatic force of each target utterance, a specific coding scheme was developed. This coding scheme focussed on the level of agency and control in each utterance and was based on previous work by Budwig (1995). It was claimed by Budwig (1995) that children were using *my* in place of *I* in utterances which were marking their participation in actions that were high in agency and control. This was measured by analysing the broader discursive context and actions were coded as either control or non-control acts and also whether the action was conscious or just a ‘happening’ (Budwig, 1995). For the current study, three central questions were identified to establish whether the child was attempting to express agency in his/her target utterance or attempting to take or express control. Every target utterance was analysed using its discursive context in order to answer *yes, no or unsure* to the three coding questions. In addition, the main verb in every target utterance was identified and recorded. These three questions were:

1. Does the target child utterance, bring about a change, or attempt to bring about a change in the agent of an action or suggested action, to become the sole agent of the action?

2. Does the target child utterance, bring about a change or attempt to bring about a change in the action being carried out, to dismiss the other action or to dismiss the suggested action?

3. Does the target child utterance, bring about, or attempt to bring about a change in the owner or possessor of an object?

An example in which Warren protests at the suggestion of his mother making a tower is shown in Figure 3.2. He suggests repeatedly that he will make a tower and his *my-for-I* error expresses this.
A second coder, blind to the aims of the study, coded 100% of one child’s data and 20% of the other three children’s data to check inter-rater reliability. All target utterances were presented in a random order, mixed across different children’s utterances, with no emphasis on the pronoun in the utterance and further own name and me utterances were mixed in to avoid any obvious indications that the aim was to analyse my-utterances vs. I-utterances. A reliability check was run using a separate Cohen’s Kappa for each of the three questions (as they were not mutually exclusive). The two coders matched in over 95% of their ratings for all three coding questions, with Kappa = 0.807, 0.903 and 1.000 respectively.

Figure 3.2. An example target interaction between mother and child in which the child is attempting to bring about a change in the agent of an action.
3.4.2 RESULTS

3.4.2.1 My-errors analysis

First we compared the number of my-for-I errors that showed an aspect of agency and control with the number of matched I-utterances that showed agency and control, see Figure 3.3.

![Figure 3.3. The raw number of my-error utterances and I-utterances that functioned to attempt to gain agency or control (out of a possible 60, 53 for Ruth) (predictors 1,2 or 3).](image)

A mixed effects model was fitted to the data to investigate the effects of the three predictors (which corresponded to the three coding questions) on the likelihood of children producing a my-for-I error. A Logistic Regression model was created and run in R to include random effects of child and of verb (only random intercepts were modelled, slopes were not included in the model). Verb corresponded to the identity of the main verb in each target sentence and was included to control for the possibility that children may use my more than I with specific verbs. The analysis was run with fixed effects of stage (early vs. late) and predictors 1, 2 and 3 (corresponding to the three coding questions). Interactions
were included between stage and predictors 1, 2 and 3 to determine whether there was a change in the function of *my-errors* over development. The three predictors were run separately in order to highlight the different types of agency and control that children may be attempting to acquire.

Results revealed a significant main effect of predictor 1, $\beta = 2.90$ ($SE = 0.81$), $p < .001$, predictor 2, $\beta = 2.10$ ($SE = 0.61$), $p < .001$ and predictor 3, $\beta = 1.86$ ($SE = 0.62$), $p < .01$, with *my-errors* matching the coding categories significantly more than the *I*-utterances. There was no significant interaction between developmental stage and any of the three predictors showing that the difference in agency and control between *my-errors* and the *I*-utterances, identified by the three coding categories, did not change significantly between the early and later stages of children’s my-errors.

### 3.4.2.2 Me-errors and own-name-verb analysis

The results show that *my-for-I*-errors have a function of agency and control in comparison with *I*-utterances. In order to conclude that this function is specific to *my-errors* and not applicable to all unconventional self-referential forms the same analysis was conducted on four children who predominantly made *me-for-I* errors and four children who made *own name-for-I* replacements.

The four children selected to be included in the *own name* replacement vs. *I*-utterances functional analysis were Anne, Liz, Nic and Warren. These four children were all from the Manchester Corpus (Theakston, et al., 2001) and included in Study 1. Just as with the *my* analysis, individual error periods were defined, based on the child’s use of *own name-for-I* replacements and on there being sufficient
matched *I-utterances* for comparison. An early and late stage of errors was again identified, see Table 3.4.

**Table 3.4. Error period for the four own-name children.**

<table>
<thead>
<tr>
<th>Name</th>
<th>Early Stage of Errors (Age years; months.days)</th>
<th>Late Stage of Errors (Age years; months.days)</th>
<th>Total own name utterances included in analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anne</td>
<td>1;10.07 to 2;0.15</td>
<td>2;4.14 to 2;9.08</td>
<td>60</td>
</tr>
<tr>
<td>Liz</td>
<td>2;8.28 to 2;1.25</td>
<td>2;8.28 to 2;7.17</td>
<td>60</td>
</tr>
<tr>
<td>Nic</td>
<td>2;1.15 to 2;4.21</td>
<td>3;0.10</td>
<td>60</td>
</tr>
<tr>
<td>Warren</td>
<td>1;10.06 to 1;11.04</td>
<td>2;3.25 to 2;8.24</td>
<td>60</td>
</tr>
</tbody>
</table>

The selection of four children who made a substantial number of *me-utterances* was not as straightforward. Only two of the children included in Study 1, Anne and Ruth, had a high enough number of *me-errors* to be selected so CHILDES was searched for two more children. Peter (Bloom, Lightbown & Hood, 1975) and Lara (Rowland & Fletcher, 2006) were chosen as the two additional children with the highest number of *me-errors*. Again, individual error periods were defined and an early/late stage of errors identified, see Table 3.5.

**Table 3.5. Error period for the four me-errors children.**

<table>
<thead>
<tr>
<th>Name</th>
<th>Early Stage of Errors (Age years;months.days)</th>
<th>Late Stage of Errors (Age years;months.days)</th>
<th>Total me-errors included in analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anne</td>
<td>2;1.20 to 2;4.12</td>
<td>2;4.12 to 2;6.29</td>
<td>28</td>
</tr>
<tr>
<td>Lara</td>
<td>2;00.08 to 2;10.14</td>
<td>2;10.21 to 3;02.11</td>
<td>30</td>
</tr>
<tr>
<td>Peter</td>
<td>2;1.00 to 2;1.18</td>
<td>2;1.18 to 2;8.12</td>
<td>38</td>
</tr>
<tr>
<td>Ruth</td>
<td>2;1.10 to 2;5.06</td>
<td>2;11.14 to 2;11.21</td>
<td>60</td>
</tr>
</tbody>
</table>
For both analyses, matched *I-utterances* were extracted for every *own-name* or *me-utterance*. Every target utterance was given with the ten proceeding and ten following utterances in order to establish context. Twenty percent of the utterances were coded by a second independent coder, using the same coding scheme as identified above. Inter-rater reliability was over 95% agreement, *Kappa* = 0.749, 0.702 and 0.707 for predictors 1, 2 and 3 respectively.

Mixed effects models were fitted to the data to investigate the effects of the three predictors on the likelihood of children producing a *me-for-I* error or *own name-for-I* replacement. The models included random effects of child and verb and fixed effects of stage (early vs. late) and predictors 1, 2 and 3 (corresponding to the three coding questions). Interactions were included between stage and predictors 1, 2 and 3. Table 3.6 shows that there was a significant main effect of predictor 2 in both analyses, with the *I-utterances* matching the coding categories significantly more than the *me-errors* and *own name-replacements*, but no effects of predictors 1 or 3. There were no interactions, showing that this preference to use *I* instead of *me* and *own name* when attempting to change the action being carried out, does not significantly differ between the early and late stage of errors.

| Table 3.6. Mixed Effects Models for me-errors and own name data. |
|--------------------------------|----------------|-----|-----|
| Predictor  | Beta (β) | SE  | p    |
| **Me-errors** |          |     |      |
| 1           | 0.42     | 0.52| 0.416|
| 2           | -2.11    | 1.08| 0.496|
| 3           | 0.37     | 0.93| 0.688|
| **Own name* |          |     |      |
| 1           | 0.25     | 0.46| 0.58 |
| 2           | -0.78    | 0.36| 0.03 |
* There is no result for question 3 for the own name analysis due to no utterances being judged to have that pragmatic function.

### 3.4.3 DISCUSSION

Study two investigated the claim that children use *my* pre-verbally when they are attempting to claim agency or acquire control (Budwig, 1995). Four children’s *my*-errors were extracted and studied in context to establish if children’s genitive-for-nominative pronoun case marking errors have a specific function of agency and control. The data showed that *my*-errors were associated with the child becoming the agent of an action against a previous suggestion; changing the action, or attempting to continue an action against a previous suggestion to stop; or changing the possessor or owner of an object, in order to control the object. These results provide direct support for Budwig’s (1995) claim that children use *my* as subject to mark their involvement in actions high in agency and control, in comparison with *I* as subject.

We found no effect of the children’s stage of development (early vs. late) on the pragmatics of their *my*-errors. This demonstrates that children are continuing to use *my* in attempts to show agency or gain control throughout their error period. If children have attached a specific function to a certain erroneous construction type, then it makes sense that they will use it for this function throughout the error period, before adopting the correct grammatical form. This distinction is not present in the adult grammar. Therefore, as children learn the adult system better and learn that they can communicate the same meaning by using a different (correct) form, its use will disappear from their speech. This would explain the relatively short-lived nature of these errors.
One important consideration is the fact that only about a third of *my*-for-*I* errors have a function of agency and control, thus it is clear that children were using *my*-for-*I* replacements to encode a variety of functions. On the other hand, the children’s *me* and *own name* replacements of *I* did not have a function of agency and control, demonstrating that there appears to be some systematicity in the association between form and function, in line with Budwig’s (1995) argument that different non-conventional self-referencing forms are associated with different pragmatic functions. It may be that children’s *own-name* and *me-errors* are also associated with a specific function, but one that is different to that of *my*. However, it is clear that all of these pronominal case marking errors require multifaceted explanations. In the case of *me-verb* constructions, these are in part learned as direct strings from the input (Kirjavainen et al., 2009), and thus the form-function mapping for these errors may be much less concrete. This is also true of *own-name utterances* as children hear parents referring to themselves pre-verbally as “Mummy”, “Daddy” etc. and to the child with their name. Furthermore, caregivers vary in their tendency to use *own-name* first and second person reference (Smiley et al., 2011), thus any possible relation between these forms and a specific function is unclear. But in both cases, it is possible to find children who appear to generalise beyond the patterns present in the input, producing errors that cannot derive from input strings in a direct sense (e.g. *Me can’t do it* – Kirjavainen et al., 2009), or extending frequency of use beyond the relative use in the input (e.g. Dominic’s reliance on *own-name* self-reference in comparison to relatively high levels of pronoun use in his input, Smiley et al., 2011). This again suggests that simple explanations for these errors are unlikely to capture the full range of variation seen in children’s speech.
In the present paper, we demonstrate that children making a higher number of *my*-*for-*I errors are those who receive less modelling of the first person pronominal system and higher levels of proper name input. In addition, the relative familiarity of particular pronoun forms prior to the onset of errors influences the rate of pronominal case marking errors observed. These factors, together, lead children to make *my*-for-*I* errors, sometimes without the assignment of a specific function. However, it is possible that *my* has a tighter association with function due to its transparent meaning in possessive contexts than some other pronominal forms. Children may recruit this pronoun in instances where they wish to claim possession or agency over an action, rendering its use in place of the more opaque *I* appropriate until they have developed a more adult-like mapping between form, function, and grammatical role.

3.5 GENERAL DISCUSSION

This paper explores the question of why some English-speaking children make genitive-for-nominative first person pronoun errors. These *my*-for-*I* errors such as “*My want that*” occur around 2-4 years and have been explained by generativist theories as being a result of children having an immature case assignment mechanism (Radford, 1990) or failing to check for tense and agreement in their utterances (Schutze & Wexler, 1996). These theories have proven problematic because they do not account for the huge individual variation between children in which errors they make and at what rate these errors occur. In this paper we explored constructivist approaches to explaining these errors in an attempt to cast further light on the nature of children’s linguistic representations, and the processes involved in these errors.
In Study one we show that individual pronoun frequency as well as the frequency of certain input strings does not explain *my-verb* errors. It is the children who receive a high contrast of pronoun and proper name use in the input who make a higher number of errors. One possibility is that these children have not had the pronoun *my* modelled sufficiently and therefore are more prone to overextend its use. Further, the children’s pronoun use before the onset of errors is related to subsequent error rates. It is the children who produce relatively less *I* in this period and relatively more *my*, that go on to produce a higher number of *my-for-I* replacements. In Study two, in line with early suggestions from Budwig (1989) and Karmiloff-Smith (1979), we found that some children’s errors seem to have functional motivations. *My-for-I errors* serve a specific communicative function: to show possession and control of an action and to be the agent of this action. This could be a result of the child overextending the function of *my* as a possessive pronoun into contexts of claiming control and agency over an action.

These many factors all play a role in causing children to make these overextensions. However, at present it is difficult to determine exactly how the various factors interact, and how these interactions might change over the course of development. Computational modelling approaches are likely to be needed to fully explore these issues, because in pronoun case errors, as with other error patterns, individual variation in their frequency of occurrence limits the number of children with sufficient corpus data for analysis, and these errors are notoriously difficult to elicit in experimental situations. Therein lies the challenge to language acquisition researchers; the need to provide models of development that can predict and account for significant individual variation in a principled and detailed manner.
In summary, the present paper adds to our understanding of the complex and multifaceted explanations for children’s *my-errors* by investigating a number of possible explanations for these errors in data from the same children across development. In order to build a comprehensive account of children’s linguistic representations at any given point in development, input patterns, children’s prior usage, and functional considerations all need to be taken into account. This is broadly consistent with constructivist approaches to acquisition, but the need for better specified models of acquisition to account for these data presents a clear challenge to researchers, whatever their theoretical orientation.
3.6 REFERENCES


Processing interclausal relationships in the production and comprehension of text (121-137). Hillsdale, NJ: Lawrence Erlbaum Associate.


Chapter 4: Children’s *My-for-I* Errors and Protest in a Conflict Over Agency

This chapter will present an experimental study examining the functionalist explanation of *my-for-I* errors as well as an analysis of children’s protest types.

The focus of chapter two and chapter three was investigating why children make *my-for-I* pronoun case marking errors. These in depth corpus analyses have revealed that the function of an utterance partly explains children making these errors. The following chapter adopts an experimental method in attempting to elicit pronoun case marking errors. Because of the importance of function, the experiment involves creating a context involving conflict, in an attempt to elicit specifically *my-for-I* pronoun case marking errors. However, it seems that eliciting these errors in a one shot experimental context is rather difficult. Because of this we also chose to look at the types of protest that children were producing in this context.

There is a wealth of prior literature surrounding children’s protest and their ability to understand normativity and protest against rule violations. The next chapter attempts to build on this prior literature by looking at two and three year olds’ ability to protest to a puppet breaking a rule in a game play situation. Here we focus not only on if the two year olds can produce normative language, but also the specific nature of the language used and on pronoun use in such contexts.
Children’s *My-for-I Errors* and Protest in a Conflict Over Agency.

Stacey Mcknight\(^1,2\), Elena Lieven\(^2,3\) & Anna Theakston\(^2,3\)

\(^1\)Max Planck Child Study Centre, University of Manchester, UK
\(^2\)School of Psychological Sciences, University of Manchester, UK
\(^3\)ESRC International Centre for Language and Communicative Development (LuCiD), University of Manchester, UK
4.1 ABSTRACT

Research has shown that children’s *my*-*for-I* pronoun case marking errors can be functionally motivated; *my* can be used in place of *I* in contexts surrounding the child’s claims of agency and control (Mcknight, Theakston & Lieven, Chapter three). In study one we experimentally investigate children’s pronoun use and specifically *my*-for-*I* errors in a game that creates situations of high and low conflict over agency. Twenty six two-and-a-half to three-year old children and twenty six three to three-and-a-half year old children participated and their verbal responses were coded for pronoun use and pronoun errors. We found that children used self- and other-reference more in the high conflict situation compared to the low conflict situation. No age differences were found and no differences in pronoun case marking errors or correct *I*-verb use were found across ages or conditions. Very few *my-for-I* errors were found in the experimental study and this finding will be discussed in relation to possible contextual explanations. In study two, we investigated the children’s verbal responses further to focus on their protests. Past research has found that children as young as three years of age will normatively protest to rule violations (Rakoczy, Warneken & Tomasello, 2008) therefore showing knowledge of norms and rules. Children’s protest was coded as either verbal (normative, imperative or as an agreement) or as an action protest (preventative, active, corrective or agreeing). It was found that children as young as two-and-a-half years were protesting normatively. Further, the two and three year old’s differed in their ability to adjust their level of imperative response in reaction to context as well as two year olds relying more on verbal only responses compared to the three year olds. The present work highlights two year old’s ability to understand norms and their ability to vocalise their normative awareness, while
showing slightly less sophistication than the three years old children in adapting their protest types.
4.2 INTRODUCTION

English speaking children will often make pronoun case marking errors in their early speech. It has been claimed that each form used by the child has a different function and this may lead to errors (Karmiloff-Smith, 1979). Children do not always imitate exactly what they hear from the adult system, but can be creative in their early speech (e.g. see Cameron-Faulkner, Lieven & Theakston, 2007, on early negation). Budwig (1989; 1995) specifically investigated children’s pronoun case marking errors and the claim that children use contrasting forms based on the pragmatic force of the utterance (Deutsch and Budwig, 1983). She found preliminary evidence that the use of my, as subject appeared to be related to the child’s participation in actions and contexts that were high in agency and control. A recent corpus study investigating children’s my-for-I pronoun case marking errors (McKnight, Theakston & Lieven, Chapter three) has further shown that children are inclined to use my in place of I in situations involving the child attempting to gain agency and control over an action or attempting to claim control over what they want to happen. A higher proportion of my-for-I errors were made in contexts when the child is attempting to be the agent of an action or control the course of action, in comparison to correct I-verb utterances. The following example shows a child making a my-for-I error in a context in which they wish to claim agency; Mother: “Can mummy make a tower please?” Child: “No. My make a tower. My make a tower. My, my, my, My make a tower”.

In the current paper, we explore the functional basis for children’s self-reference further, in an experimental settings. We also extend this work to consider how the use of self- and other-referring forms is influenced by conflict and rule violations. In Study one, we focus on the number of my-for-I as well as other
pronoun case marking errors and correct pronoun use made in a situation of high vs. low-conflict of agency, and how this changes with age. In Study two we investigate the types of protest that the children produce in response to rule violations, how this differs between the two age groups, and also how this changes (does protest increase or decrease) across trials.

Children are aware of social norms from an early age and research has shown that children have the ability to not only protest to any rule break but to also then attempt to enforce these rules and social norms to third parties from as young as two and three years (Rakoczy, 2008). Work has since shown that children can distinguish between conventions and morals and that young children can grasp the norms associated with both ‘regulative’ (rules that regulate existing activities, see Kalish, 1998, Kalish & Shiverick, 2004, and Harris, Nunez & Brett, 2001) and ‘constitutive’ rules (rules that bring into existence the activity or game that they are playing). These rules bring about status functions on objects and actions, such as “in this game the block now is a sponge, but in this game the block is a car”.

Rakoczy, Warneken & Tomasello (2008) investigated two and three year olds’ awareness of the normative structure in conventional games. In this set up, each child interacts with an experimenter in performing two actions, one labelled as a novel verb such as ‘daxing’ and another that is marked as a mistake. The child then gets to ‘dax’. Next, a puppet comes along and performs the incorrect action that had previously been marked as a mistake and claims to be ‘daxing’. The control condition involves the experimenter performing two actions, but marked neutrally, the child then has a turn and the puppet then comes to play but just says “I am going to show you something” and therefore this does not constitute a mistake. The children’s responses (protest, critique or teaching) were scored for the two
conditions. There was a main effect of condition and of age, however planned comparisons revealed that only the three year olds performed more normative responses in the experimental condition compared to the control condition. Normative protest was defined as the child intervening in a normative way and therefore using normative vocabulary that refers to how things should be. For example, “No! It does not go like this”. Imperative protest does not involve this normative vocabulary and is defined as positive, e.g. “Take the stick!” or negative, e.g. “No! Not in this hole!” protest or direction. The two year olds protested very infrequently in a normative manner and this didn’t differ across the conditions.

Study two investigated children’s normative interventions with a more conventional rule game that tested children’s understanding of the context specificity of these rules. The method was similar to study one except the experimenter and child played with some known objects (such as bricks) in phase one in the usual manner, such as building. Phase two involved the experimenter showing the child a rule governed game called ‘daxing’ (a slightly simpler version of the game was used for the two year olds) and explaining the rules. The puppet then appeared and performed the usual action done with the objects, for example building and either claimed to be ‘daxing’ (experimental condition) or to be building (control condition). Results revealed that the three year olds performed significantly more protests (normative and imperative) in the experimental rather than the control condition, showing that they understood the normative space that constituted a mistake. The two year olds did not protest normatively, but did protest more imperatively in the experimental condition compared to the control condition. It was concluded that young children clearly understand the normative structure of simple conventional acts and are able to install a normative framework of what is
an appropriate act and what is a mistake. They can then criticise and/or protest against these mistakes. They claimed that the two year olds may hold similar normative awareness as the three year olds but in a yet to be developed form, therefore the two year olds may be aware that this is against the normative rules, but are not as advanced as the three year olds in vocalising this. Rakoczy (2008) investigated children’s normative responses to a third party making mistakes in joint fictional activities. Rather than introducing explicit rules, children were simply involved in a joint pretence game with a play partner and a third party entered and either played appropriate pretence acts (control) or inappropriate pretence acts (experimental). Rakoczy (2008) found that children as young as three (and to a lesser extent the two year olds) “not only act in accordance with the fictional stipulations of joint pretence games, but grasped their normative structure” (p.1198). Three year old children protested normatively when a puppet confused the pretend status functions within a game, for example, protesting when the puppet pretended to eat a toy knife, but not when the puppet pretended to eat a toy carrot. The two year olds did also protest but on a less explicit or normative level and less so overall.

These studies both had two experimental conditions and two control conditions in which the focus was the child’s reaction to one central ‘mistake’ by the puppet. This one-shot reaction to the norm-violation is possibly not enough to extract the maximum level of protest from these children. Perhaps initially these young children assume the puppet knows more than them or they are too shy to protest. Rakoczy (2008) pointed out that spontaneous intervention is quite demanding and the absence of this does not mean that children do not understand the mistake. Also the younger children may not have been as engaged in general
with the game and therefore the age difference could be due to the three year olds’ increased willingness to play. In order to further investigate this, one possibility would be to investigate if children of such a young age will protest more, whether normatively or imperatively, when the puppet repeatedly breaks the rules of a game. If they are given more than just two one-shot rule breaks, there may be a higher number of reactions and their abilities and normative understanding may be exemplified.

Wyman, Rakoczy and Tomasello (2009) further examined children’s understanding of normative rules and their context-specificity with two studies. Testing three year old children they compared children’s interventions to a puppet when he claims to be playing a pretend game with an object (for example, playing the ‘toothbrush game’ in which a pen was used as a pretend toothbrush), but then plays with the object using its real function, to when he says he doesn’t like the pretend game (and says he would prefer to draw) and then uses the pen as it is ‘meant’ to be used. The children’s responses were coded as Explicit (normative instruction, such as “No! You must brush your teeth), Implicit (such as saying “tooth brush”) and descriptive (describing the puppet’s action or something relating to the pretend game, for example “this is toothpaste”). The results showed that the three year olds understood the context-specificity of the normative rules and protested normatively more in the condition in which the puppet declared that he was joining in with the pretence game but then acted out the functional purpose of the item, compared to when he acted functionally but hadn’t declared that he was part of the pretend game. In their second study Wyman, Rakoczy and Tomasello (2009) investigated if the same children could apply different norms to two different pretend games (rather than compared to reality). One game was played at
location A and one at location B, but the target object had a different ‘pretend’
identity in each game. This time the puppet did not declare any action, but just
joined in the game at one location performing the same pretend action irrespective
of the game. The action was thus either consistent or inconsistent with the context.
Results showed that these three year old children understood that the object can
have two different pretend statuses and intervened in a normative way more when
the pretend action was inconsistent with the game context. This shows that three
year olds can apply normative rules context-specifically. It was noted in this study
that the overall intervention rate was below 50% and again the measure was quite
demanding. Rakoczy, Brosche, Warneken, & Tomasello (2009) also found that
three year olds were able to intervene in a rule break to protest normatively, but two
year olds did not. Further, Rossano, Rakoczy and Tomasello (2011) show that
children will protest to an actor taking possession of an object that belongs to the
child or a third party and/or attempting to throw this object away. It was found that
three year olds protest in both conditions, protesting most regarding their own
items, but also more when the puppet took or threatened to throw away a third
party’s item of clothing than when the puppet did this with their own clothing.
However it was found that two year olds, did not stand up for the property rights of
a third party and this was a consistent finding, even when nonverbal reactions are
considered. The two year old children only protested when it was their object that
was being taken (Rossano, Rakoczy & Tomasello, 2011). In their protests, the three
year olds were seen to use normative language, such as “You can’t do that. It’s
hers”. This further supports that children as young as three are aware of normative
rules and in this instance the normative dimension of property rights and they
understand when this rule is violated.
The current study examines whether two year old children have the ability to not only know when a rule is being broken, but also to use normative language and express this. Therefore it is of interest to see if these younger children will protest normatively in a situation in which there is more than two one-shot rule breaks, that may take the child by surprise and they are not prepared. By repeating the same rule break, it will give children an increased opportunity to express this. Further, we investigate pronoun use in the same context and in particular my-verb errors. To attempt to illicit an increased number of my uses, a priming technique will be used. As well as studies showing that adults and children are more likely to produce a syntactic structure if they had heard it or themselves produced it shortly preceding the target utterance (Bloom, Rispoli, Gartner & Hafitx, 1989; Savage, Lieven, Theakston & Tomasello, 2003; Theakston & Lieven, 2008) it has also been shown that adults and children can be primed for specific lexical items (Ferreira, 2003; Savage et al., 2003; 2006). Therefore we will attempt to prime the pronoun my before the experimental trials begin.

4.1.1 Current Study

In study one, we experimentally investigate children’s my:for-I pronoun case marking errors, with the aim of eliciting these errors in an experimental setting whilst manipulating the level of agency and control the children will have in a game setting. We further plan to look at overall levels of self-reference in conditions of high agency and control and low agency and control. We designed a game in which the child plays a board game with ‘Cat’ (a soft toy puppet) and there are clear rules regarding agency, in which the child has to place all of their pieces on the game board and Cat must place all of his. There are then two conditions, one in which Cat
will violate the rules and claim the child’s blocks and attempt to take the child’s turn three times out of six (high conflict condition) on three games and therefore making a mistake in accordance with the game rules. This will create a context in which the child may wish to (re)claim agency and control. In the low conflict condition, Cat will suggest that he could take the child’s turn, but without actively taking control of the child’s piece and therefore without actively making a mistake. The two conditions will be compared for pronoun use and specifically any *my*-for-*I* errors made or other pronoun case marking errors produced. In setting up this methodology, children were also using a high level of protest and therefore study two will then explore this. The level of protest used by the two and three year old children in the same game will be analysed, how this differs across the two conditions, between the two age groups and how this changes across trials.

### 4.3 METHOD

#### 4.3.1 Participants

Fifty two monolingual English 2;6 to 3;6-year old children were included in the study. An additional seven children were tested but were excluded due to them not completing the test sessions. Twenty six children were recruited through the Max Planck Child Study Centre database and tested in the Max Planck Child Study Centre. Twenty six were recruited from local nurseries and pre-schools in the Greater Manchester Area and tested in a quiet area of the nursery.

#### 4.3.2 Materials

4.3.2.1 *Puppet* - All children interacted with a soft cat puppet, called “Cat”, controlled by a second experimenter, who was also able to talk for him.
4.3.2.2 Game - Children played a colour matching game, which involved three animal pictures (Snake, Snail and Caterpillar) with coloured stripes, which alternate between either red and yellow (the child’s colours) and blue and green (Cat’s colours). There were also six blocks for the child to play with (three red and three yellow) and six blocks for Cat to use (three blue and three green). Each stripe on the game boards had a Velcro piece and each block had Velcro on several sides, so that the pieces could easily be attached to the game board (see Appendix A). There were also two additional snake pictures used in training, to teach the child the rules of the games.

4.3.2.3 Stories - For priming of the pronoun my there was an audio recording of four stories, all containing six sentences, with my modelled in each sentence (see Appendix B). Each story had corresponding pictures for each sentence, which resulted in a picture book of 24 pictures (see Appendix C).

4.3.2.4 Parent Questionnaire - Parents were asked to fill in a short questionnaire which asked them if they had noticed their children making any type of pronoun case marking error (see Appendix D).

4.3.2.5 Sticker Sheet - A red and yellow striped snake was designed, with six spaces. This was used as a motivating tool to encourage the child to complete the sticker sheet by playing 6 games with Cat (see Appendix E).

4.3.3 Design

There were two conditions; high-conflict and low-conflict, which related to how the puppet acted in the game. In the high-conflict condition the cat claimed agency over the blocks in 3 of the child’s 6 turns, by saying “Cat do it” and also physically taking the child’s block. In the low-conflict condition, the cat vocalised
agency, by saying “Cat do it”, but did not claim or go near the child’s block. These two conditions were counterbalanced, with half of the children receiving the high-conflict condition first (order 1) and half of the children receiving the low-conflict condition first (order 2). The three games (Snake, Caterpillar and Snail) within each condition were also counterbalanced for order of presentation, as were the four priming stories.

4.3.4 Procedure

4.3.4.1 Warm up and training - First the child completed a jigsaw with Cat, to help them feel comfortable in talking to the experimenter and to Cat. They were then shown the coloured blocks and asked to name the colours\(^\text{16}\). They were given the red and yellow blocks and explicitly told that those are their blocks to play with in the game. Cat had the green and blue blocks and this pattern stayed the same throughout the experiment, so that it was very clear to the child which blocks they should use. They were then introduced to the game and shown the first practice snake and told they were to go along the snake matching the blocks from tail to head without missing any stripes out. For each of the child’s turns, Cat will tell them to “put the yellow one there” or “stick the red piece on next” etc. The child was encouraged to do the same and tell Cat what to do for his turns. The experimenter helped the child play the game to teach them the rules. On one of the child’s turns, instead of saying “Put the red one there”, Cat said “oooh, red next, Cat do it?” If the child didn’t object, the experimenter explained that the child has to do all of their pieces and Cat isn’t to do any, as those are the rules of the game.

\(^\text{16}\)The majority of the children were comfortable with the four colours, however for the few children who were not sure, the focus of the game wasn’t to name the colour each time, but to show which colour to do next and say “do that one next”, which a lot of the children did anyway. Naming the colour wasn’t an essential part of the game.
Once the first practice game had finished, the child received a sticker for “doing all of their own red and yellow pieces”. A second practice game was then played to reiterate the rules and ensure the child understood the game and that they are to do their own pieces.

4.3.4.2 Priming - Next, the child was asked to look at the pictures and listen to the stories and the four priming stories were shown/played in a randomly counterbalanced order.

4.3.4.3 First Condition - The child then played three more games with Cat, either the three high-conflict or three low-conflict games (described below). The child was asked if they wished to get the rest of the stickers on their sticker sheet (they have two so far out of six from the two warm up games). To do so they should play some more games with Cat. The experimenter got the first game, (the snail, caterpillar or snake in a counterbalanced order) and the child was reminded what to do. The experimenter explained that they have some writing to do now, so will be busy doing that, but as the child knows how to play the game, they are to play with Cat and shout to tell the experimenter once they have finished. The experimenter then turned around and did not watch the game as it was played, to encourage the child to verbally inform her of any rule violations.

4.3.4.4 High Conflict Condition - Cat played the game but on three of the child’s six turns Cat took the child’s piece and said “Cat do it? Who, Cat put it there” and stuck down the child’s piece for them. At the end of the game the child called the Experimenter who then turned around. The child was then asked if they did all of their own pieces. If they said yes, Cat informed the experimenter that he did some of the child’s colours. The child was reminded that to get a sticker they have to do all of their own pieces (to encourage the child’s attempts to regain agency/protest
against Cat’s rule violations) and the experimenter then got the next game ready to play. The experimenter then turned around again. This was repeated for the third game.

4.3.4.5 Low Conflict Condition - The Cat asked on three of the child’s six turns if he should do it, “Cat do it?” but he did not attempt to take the block or actually take control of the action. Instead he said, “Who, Ok, you put the red one on” for example. The child got a sticker at the end of each low-conflict game because they had put all of their blocks on themselves.

4.3.4.6 Priming - After the first high/low-conflict condition, the child was then shown two of the four priming stories that they had previously seen. This was always the first two that were shown out of the four previously (which varied depending on the story order that the child received).

4.3.4.7 Second Condition - The second of the two conditions was now played, again with the experimenter not watching. After this second condition the game finished. The child was praised and allowed to take their sticker sheet home with them.

4.4 STUDY 1

Study one investigated children’s my-for-I pronoun case marking errors in particular and also the self and other-reference forms (pronoun and proper names) used by the children, across conditions and between the two age groups. Therefore all self-reference and reference to Cat was coded.

4.4.1 Pronoun and proper name use Coding

All of the data for each child was transcribed. The child’s speech in each game was transcribed from the first instance of “Cat do it?” through to the end of
each game. In addition, all video data was analysed and any actions the child made during the game recorded alongside the utterances produced. For example, if the child covered their blocks or handed the block to Cat, this was recorded with separate action codes, along with the verbal data.

Every instance of a self-referring pronoun or proper name use was coded. This included any use of *I, my, me, own name* and *mine*. Further, any use of these forms with a verb phrase was coded. This allowed any correct *I-verb* utterances to be coded as well as any *my-for-I errors* and other pronoun case marking errors or own name replacements of *I*. In addition, any reference to Cat was coded. This included, *he, his, him, Cat, you and your*. Again any use of these forms with a verb phrase was also coded. From the coding we derived the total frequency of use of the different forms per game and as a function of conflict condition, for each child.

### 4.5 RESULTS

The main aim of the current study was to experimentally investigate children’s *my-for-I* pronoun case marking errors in relation to the function of the utterance. The current methodology had two conditions creating differing contexts within the same colour matching game. One condition in which a puppet was claiming a high level of agency and control and one condition in which the puppet was only claiming a low level of agency and control.

<table>
<thead>
<tr>
<th>Total pronoun or own name and verb use for the two age groups.</th>
</tr>
</thead>
<tbody>
<tr>
<td>I+Verb</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>3 year olds</td>
</tr>
<tr>
<td>2 year olds</td>
</tr>
</tbody>
</table>
In this experimental paradigm, the total number of pronoun case marking errors produced by the two and three year old children, was very low, see table 4.1. There were only two instance of a my-for-I pronoun case marking error in this experimental setting (Note: these were both made in the high conflict of agency condition). However, with such a low number of the errors under focus being produced, no analysis could be conducted on this specific data. This low number of errors will be discussed further and possible explanations for such a low number of errors will be highlighted in the discussion section.

As well as this main focus of the current study on children’s my-for-I errors, we also aimed to investigate if there was a difference in the number of correct pronoun verb utterances and pronoun case marking errors overall, between the two age groups regardless of condition. A one-way MANOVA with main effect of Age (younger group and older group) was conducted on this data. Each individual self-reference+verb was compared between the two age groups; I, My, Me, Name and Mine (therefore including pronoun case marking errors). Further, the total number of I-replacements (My, Me and Name) and specifically My and Me replacements are compared between the two age groups. All results were non-significant, see table 4.2. This shows that there is no difference in individual pronoun production between the two age groups. More specifically, there is no difference in total error rate or correct I-Verb production between the two age groups.

To establish whether either of the two conflict conditions resulted in the children’s production of more self-referring forms, a three way mixed effects ANOVA was conducted on the total number of self-referring forms used between the two conditions.
Table 4.2. *F* values, significance level and partial $\eta^2$ for the comparison between the two year old and three year old children for each reference type.

<table>
<thead>
<tr>
<th>Reference</th>
<th>$F(1,50)$</th>
<th>Sig.</th>
<th>partial $\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>1.047</td>
<td>.311</td>
<td>.021</td>
</tr>
<tr>
<td>I + Verb</td>
<td>.194</td>
<td>.661</td>
<td>.004</td>
</tr>
<tr>
<td>All I</td>
<td>.358</td>
<td>.553</td>
<td>.007</td>
</tr>
<tr>
<td>My</td>
<td>.020</td>
<td>.888</td>
<td>.000</td>
</tr>
<tr>
<td>My + Verb</td>
<td>0.000</td>
<td>1.000</td>
<td>0.000</td>
</tr>
<tr>
<td>All My</td>
<td>.020</td>
<td>.889</td>
<td>.000</td>
</tr>
<tr>
<td>Me</td>
<td>.470</td>
<td>.496</td>
<td>.009</td>
</tr>
<tr>
<td>Me + Verb</td>
<td>1.536</td>
<td>.221</td>
<td>.030</td>
</tr>
<tr>
<td>All Me</td>
<td>.108</td>
<td>.744</td>
<td>.002</td>
</tr>
<tr>
<td>Name</td>
<td>.795</td>
<td>.377</td>
<td>.016</td>
</tr>
<tr>
<td>Name + Verb</td>
<td>.637</td>
<td>.429</td>
<td>.013</td>
</tr>
<tr>
<td>All Name</td>
<td>1.125</td>
<td>.294</td>
<td>.022</td>
</tr>
<tr>
<td>Mine</td>
<td>.500</td>
<td>.483</td>
<td>.010</td>
</tr>
<tr>
<td>All My, Me + Name Verb</td>
<td>1.887</td>
<td>.176</td>
<td>.036</td>
</tr>
<tr>
<td>All My-for-I and Me-For-I Errors</td>
<td>1.535</td>
<td>.221</td>
<td>.030</td>
</tr>
<tr>
<td>He</td>
<td>1.000</td>
<td>.322</td>
<td>.020</td>
</tr>
<tr>
<td>He + Verb</td>
<td>1.435</td>
<td>.237</td>
<td>.028</td>
</tr>
<tr>
<td>All He</td>
<td>1.194</td>
<td>.280</td>
<td>.023</td>
</tr>
<tr>
<td>Cat</td>
<td>.321</td>
<td>.574</td>
<td>.006</td>
</tr>
<tr>
<td>Cat + Verb</td>
<td>.258</td>
<td>.614</td>
<td>.005</td>
</tr>
<tr>
<td>All Cat</td>
<td>.019</td>
<td>.891</td>
<td>.000</td>
</tr>
<tr>
<td>You</td>
<td>1.860</td>
<td>.179</td>
<td>.036</td>
</tr>
<tr>
<td>You + Verb</td>
<td>.148</td>
<td>.702</td>
<td>.003</td>
</tr>
<tr>
<td>All You</td>
<td>.032</td>
<td>.858</td>
<td>.001</td>
</tr>
<tr>
<td>Your</td>
<td>1.860</td>
<td>.179</td>
<td>.036</td>
</tr>
</tbody>
</table>

There were two between-group factors, both with two levels; Order (High-Conflict first, Low-Conflict first) and Age (younger group, older group) and one within-subjects factor, also with two levels; Condition (High-Conflict, Low-Conflict). This analysis revealed a significant main effect of condition $F(1,48) = 4.11, p = .048$, partial $\eta^2 = 0.079$, with children using more self-referring forms in the high-conflict ($M = 4.81, SE = 0.711$) compared to the low-conflict condition ($M$...
= 3.52, \( SE = 0.517 \)). The main effects of Age and Order, were not significant, \( F(1,48) = 0.14, p = .707 \), partial \( \eta^2 = 0.003 \) and \( F(1,48) = 1.37, p = .248 \), partial \( \eta^2 = 0.028 \) respectively. There were no significant interactions.

Table 4.3. Total self-reference and reference to Cat in both age groups in both conditions.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Age</th>
<th>Total self-reference</th>
<th>Total Cat-reference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>My</td>
<td>Me</td>
</tr>
<tr>
<td>High conflict</td>
<td>3 year olds</td>
<td>52</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>2 year olds</td>
<td>43</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Both</td>
<td>95</td>
<td>19</td>
</tr>
<tr>
<td>Low Conflict</td>
<td>3 year olds</td>
<td>44</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>2 year olds</td>
<td>34</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Both</td>
<td>78</td>
<td>17</td>
</tr>
</tbody>
</table>

A three way mixed effects ANOVA was conducted on the number of references to Cat between the two conditions. There were two between-group factors, both with two levels; Order (High-Conflict first, Low-Conflict first) and Age (younger group, older group) and one within-subject factor, also with two levels; Condition (High-Conflict, Low-Conflict). This analysis revealed a significant main effect of condition \( F(1,48) = 10.72, p = .002 \), partial \( \eta^2 = 0.183 \), with children using more other (Cat) referring forms in the high-conflict (\( M = 1.25, SE = 0.324 \)) compared to the low-conflict condition (\( M = 0.42, SE = 0.132 \)). The main effects of Age and Order, were not significant, \( F(1,48) = 0.25, p = .621 \), partial \( \eta^2 = 0.005 \) and \( F(1,48) = 0.59, p = .445 \), partial \( \eta^2 = 0.012 \) respectively. There were no significant interactions.

As there were no effects of age and order on children’s overall use of self and other-referring forms, we conducted a series of MANOVA analyses with a
main effect of condition (with two levels; High-Conflict, Low-Conflict) to investigate if the children’s use of specific self and other-referring forms differed between the two conditions\(^{17}\). First each individual self-reference was compared across the two conditions; *I, My, Me, Name* and *Mine*. Secondly self-reference use with a verb (and therefore including pronoun case marking errors) was compared across the two conditions. Third, self-reference use without verbs (individual use) was compared across the two conditions. All results were non-significant, see table 4.4.

Table 4.4. *F* values, significance level and partial $\eta^2$ for the comparison between the high-conflict and low-conflict conditions for each reference type.

<table>
<thead>
<tr>
<th>Reference</th>
<th>$F(1,48)$</th>
<th>Sig.</th>
<th>partial $\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0.047</td>
<td>.83</td>
<td>0.001</td>
</tr>
<tr>
<td>I + Verb</td>
<td>0.657</td>
<td>.421</td>
<td>0.013</td>
</tr>
<tr>
<td>All I</td>
<td>0.577</td>
<td>.451</td>
<td>0.011</td>
</tr>
<tr>
<td>My</td>
<td>0.000</td>
<td>1.000</td>
<td>0.000</td>
</tr>
<tr>
<td>My + Verb</td>
<td>2.040</td>
<td>.159</td>
<td>0.038</td>
</tr>
<tr>
<td>All My</td>
<td>0.039</td>
<td>.844</td>
<td>0.001</td>
</tr>
<tr>
<td>Me</td>
<td>0.835</td>
<td>.365</td>
<td>0.016</td>
</tr>
<tr>
<td>Me + Verb</td>
<td>1.000</td>
<td>.322</td>
<td>0.019</td>
</tr>
<tr>
<td>All Me</td>
<td>1.378</td>
<td>.246</td>
<td>0.026</td>
</tr>
<tr>
<td>Name</td>
<td>0.000</td>
<td>1.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Name + Verb</td>
<td>3.419</td>
<td>.070</td>
<td>0.063</td>
</tr>
<tr>
<td>All Name</td>
<td>1.043</td>
<td>.312</td>
<td>0.020</td>
</tr>
<tr>
<td>Mine</td>
<td>2.641</td>
<td>.110</td>
<td>0.049</td>
</tr>
<tr>
<td>He</td>
<td>1.000</td>
<td>.322</td>
<td>0.019</td>
</tr>
<tr>
<td>He + Verb</td>
<td>1.423</td>
<td>.238</td>
<td>0.027</td>
</tr>
<tr>
<td>All He</td>
<td>1.662</td>
<td>.203</td>
<td>0.032</td>
</tr>
<tr>
<td>Cat</td>
<td>1.619</td>
<td>.209</td>
<td>0.031</td>
</tr>
<tr>
<td>Cat + Verb</td>
<td>2.878</td>
<td>.096</td>
<td>0.053</td>
</tr>
<tr>
<td>All Cat</td>
<td>5.144</td>
<td>.028</td>
<td>0.092</td>
</tr>
<tr>
<td>You</td>
<td>1.829</td>
<td>.182</td>
<td>0.035</td>
</tr>
<tr>
<td>You + Verb</td>
<td>6.385</td>
<td>.015</td>
<td>0.111</td>
</tr>
<tr>
<td>All You</td>
<td>7.933</td>
<td>.007</td>
<td>0.135</td>
</tr>
<tr>
<td>Your</td>
<td>1.000</td>
<td>.322</td>
<td>0.019</td>
</tr>
</tbody>
</table>

\(^{17}\)There were no instances of *Cat* being referred to as *His* or *Him* and therefore these pronouns were removed from the analysis.
Next each individual reference to Cat was compared across the two conditions; *He, Cat, You* and *Your* (again, first for all uses, secondly for use with a verb and third any individual use). As table 4.4 shows, there was a significant main effect of condition with children referring to Cat as *Cat* (all uses) in the high-conflict condition (*M* = 0.308, *SE* = 0.097) more than in the low-conflict condition (*M* = 0.135, *SE* = 0.055). Further, children refer to Cat as *You* (all uses) and specifically as a *You+verb* utterance, significantly more in the high conflict condition (*M* = 0.654, *SE* = 0.223 and *M* = 0.596, *SE* = 0.206 respectively) than in the low conflict condition (*M* = 0.250, *SE* = 0.106 and *M* = 0.250, *SE* = 0.106 respectively).

### 4.6 DISCUSSION

Study one investigated children’s *my-for-I* pronoun case marking errors, as well as all self and other-reference in a game in which there was conflict over agency and control. The main aim of the study was to elicit *my-for-I* pronoun case marking errors. However, only two of these errors (made by two different children) were elicited in this experimental setting. There were not enough *my-for-I* errors produced to be able to find a difference across the conditions; however it may be worth noting that both of the two *my-for-I* errors were made in the high conflict condition.

This failure to elicit these errors experimentally may highlight the low level of these errors in general, but also may in itself highlight the functional nature of these errors. Of the fifty two children included in the study, only two produced a *my-for-I* pronoun case marking error and only one produced by each child. This finding shows that a one-shot experimental setting is by no means an optimum
setting to elicit such errors. This may be due to the unnatural setting in which children may not feel confident enough to attempt to control the situation. Children may feel comfortable claiming agency and controlling a game in a familiar environment in which their interlocutor is their caregiver, a sibling or a close friend, but with an unfamiliar puppet, controlled by an adult experimenter and whilst being observed by other experimenters, this may be out of bounds. This could highlight the fact that these errors are assigned to this control function and therefore this would explain the lack of them in a setting which doesn’t offer an environment for the child to express this particular function. This could explain why children produced more *me* and *own-name* utterances in the current study, as these are not found to be associated with this function of high agency and control. Further it could reveal that these errors in fact may only be made in contexts in which the child is speaking with a familiar person, one to whom they feel confident enough to “boss them about” almost and take control away from them and be the main agent. In an experimental setting, the other participants are new, unfamiliar and the child may assume the control lies in the hands of the experimenters and even the unfamiliar puppet and therefore they may not feel comfortable challenging this. Further, the function of claiming agency and control, may not be found when children are on their best behaviour, as this is usually in disagreement with a previously suggested agent or action and therefore, again, the experimental setting may not capture this behaviour. This may be because children coming in to a study centre to take part in an experiment or are at nursery but playing in a different quiet space with new people, are more than likely on their best behaviour. It may be the case that the children are explicitly told to behave whilst taking part prior to the experiment.
Function has not been claimed to be the sole explanation of children’s *my*-for-*I* errors. Function, as an explanation has been claimed to interplay with children’s language input; a strong modelling of the pronoun system and children’s own pronoun usage and an increased entrenchment of the correct form *I* (both shown in chapter three) will decrease *my*-for-*I* error rates. But the current study highlights the importance of setting in eliciting these errors and that in fact corpus studies, with a more naturalistic environment are far more likely to see these error types than an unnatural experimental situation. This natural setting in which children do not feel observed, or on their best behaviour, are playing games with familiar interlocutors, with familiar toys, of their choice, seems a much more optimum situation in order to elicit this attempt at claiming agency and control by the child. However, corpus studies may also be limited. They offer a limited setting, with limited interlocutors and children are recorded quite briefly and perhaps only once a fortnight. To really optimise the possibility of eliciting these errors, a natural setting that can be influenced and also that can provide a denser data set, needs to be established. This will be discussed further in chapter 5, pages 189 and 190.

The lack of *my*-errors found, also shows a lack of a priming effect. The four stories heard by each of the children, once before the experimental conditions and then a repeat of two of the stories between conditions, contained 6 utterances of *my* in each. This exposure to twenty-four instances of *my* before condition one and then twelve more instances before condition two, did not cause children to make more *my* utterances than other pronominal or self-referring forms, with children producing 173 *I*-utterances, 36 *my*-utterances, 157 *me*-utterances and 35 *own* name-utterances. It is perhaps a possibility that the modelling of this correct *my* actually could have entrenched the correct use of the form and prevent these errors.
Relating back to Chapter three, we found that the modelling of the pronoun system was related to a lower rate of errors and therefore, this modelling of the correct use of *my* could promote this correct use and therefore other forms win out when the children were searching for a first person pronoun to use in such contexts.

As well as *my*-for-*I* errors, children’s other pronoun case marking errors were examined. The rate of these was still quite low, with twenty four *me*-for-*I* errors (across nine children, with seven children making one or two errors and two children making six and seven errors), and twenty one *name*-for-*I* replacements (across four children, but with one child making sixteen out of the twenty one errors). This just emphasises the relatively rare and short lived nature of these errors in that it was difficult to capture a high number of these errors in a one-off experimental setting. Based on the fact that previous studies (Mcknight, Theakston & Lieven, Chapter three) show that children make a high number of *my*-for-*I* errors in conflicts over agency and control, but that this was not the case for other error types, it would be expected that there would be no difference in *me*-for-*I* and *own name*-for-*I* replacements between the high conflict and low conflict conditions, which is what was found in the current study. From the parental reports, the error rates amongst children appeared to be quite similar in prevalence. A total of nine parents said they noticed their child making *my*-for-*I* errors “a few times a day”, compared to eight parents for *me*-for-*I* and eleven parents for *own name* replacements. Only one parent reported hearing their child produce *my*-for-*I* errors “More than ten times a day” and the same for *me*-for-*I* errors. Also only one child was reported to make *my*-for-*I* errors “All of the time”. However, none of these children were the two that produced a *my*-for-*I* error in the experimental conditions.
We found that children were using more self-referring forms in total, including all uses of *I, my, me* and *own name* in the high-conflict condition compared to the low conflict condition, although we did not find specific differences in the use of individual self-referring forms. This shows that children are referring to themselves more when attempting to prevent the rule violation compared to when Cat just suggests a rule violation but does not actually carry it out. The same was found when children’s overall references to Cat were compared across the two conditions, including all uses of *he, Cat, you* and *your*, and we also found differences in the use of the specific forms *Cat* and *you+verb*. This again shows that children are referring to Cat more in this high conflict context compared to the low conflict. It should be noted that the overall number of utterances was considerably higher in the high-conflict condition compared with the low-conflict condition. Therefore children’s overall pronoun use was increased because they were using more language associated with themselves and Cat as they protested against the rule violation. The specific forms that were higher in the high conflict condition (*Cat* and *you+verb*) suggest that the children’s strategy in this high conflict context was to potentially explain to Cat that he should not do it, with utterances such as “noooo, you can’t do it”, “you can’t do my blocks” and “Cat can’t do it”. For both analyses there were no age or order effects. The lack of order effects shows that there was no difference between the groups dependent on the order they had received the two conditions. Therefore it wasn’t the case that the children who had seen Cat play three games without taking their blocks first and then saw Cat go on to take their blocks a lot (or vice versa), produced more self-reference or reference to Cat.
The lack of age effects shows that there was no difference in the total numbers of self-referring forms between the two age groups. This could be expected as children of two and a half are capable of producing both first and second person pronouns. When taken as a total number of uses of any self-reference or reference to Cat, all uses of any pronoun or own name uses are counted and therefore this includes any form of reference the children were capable of producing (including errors and single one word utterances, such as “me”). Therefore it is not too surprising to not find an age difference in an overall count. However, further detailed analysis investigated the number of each pronoun type, either including any production of the pronoun, or looking at specific verb phrases for all first person and second person reference despite condition. There were no age effects found between the numbers of any individual pronoun, or total number of correct *I*-verb utterances or numbers of errors. This was the case for each individual error or for total *me*, *my* and *own name errors*. The same was found with the references to Cat, with no age differences found between any individual references to Cat, or any verb phrases in reference to Cat. This lack of age difference between the numbers of errors could be explained by the lack of errors produced overall. It seems that children were not simply attempting to claim agency and therefore producing a high number of pronoun verb utterances (and therefore producing a higher rate of errors), but in fact they were protesting and focussing on attempting to prevent this rule violation instead.

The data show that instead of simply claiming agency and producing a high number of self-referring forms, children were protesting in a variety of ways, rather than simply emphasising agency, they were asserting the rules of the game, explaining why Cat should not be doing what he was doing, in utterances such as
“no more pinching this time Cat”, “noooo stop it, “don’t take it”, “it’s wrong” and “they’re mine, don’t do it, don’t put them ones on!” Children were also physically preventing the rule break, without vocalisation for example pulling Cat’s leg in order to stop him from placing the child’s block down on the board. The children’s responses to the rule violations were analysed in more detail, beyond pronoun use, in study two.

4.7 STUDY TWO

Study two investigated the different types of protest produced by the children in response to the high-conflict and low-conflict conditions. The children’s normative, imperative and agreement responses to the rule violations were compared across the two conditions. Further, children’s attempts to prevent the rule break by verbally protesting, actively protesting or both, was analysed and compared across the two age groups.

4.7.1 Protest Coding

Any instances of protest were highlighted. Protest utterances were operationalised as ANY disagreement and/or agreement with the prior action or utterance. Each protest is considered a protest in reaction to a previous turn/rule violation by Cat until the next turn by Cat has been taken. Then any protest after this is taken as in reaction to that next turn, unless explicitly stated otherwise. Any verbal protest by the child will then be coded as normative or imperative protest. These categories are defined based on Rakoczy, Warneken & Tomasello (2008), with rule referents being adapted and defined as;
4.7.1.1 Normative protest - Any utterance involving protest, critique and also teaching, but that contained normative vocabulary and therefore made reference to the rule. In this case the rules are that the child has ‘do’ their own blocks, which are red and yellow and Cat is to do his own blue and green. The puppet and the child go in turns and do not do the next block until the other participant has placed their block. Normative protest;

- Any protest/critique/teaching that contained reference to who should be the agent of an action, for example “No, I should do that one”, “you don’t do it”, “my got to do that one” or who referred to the turn for example “No it is my turn now” or which made reference to the colour or block rule such as “red is my colour” or “no that is my block”, explicitly referring to the rules.

- Any protest/critique/teaching that contains normative adjectives or adverbs such as “That isn’t right” or “that is wrong”.

4.7.1.2 Imperative protest - commanding or disagreeing (but without any explicit reference to the rule), for example “no”, “no, me” or “no don’t put it on”. Just saying “me”, or their own name is also considered imperative protest when it is in disagreement to Cat suggesting he will do it, or is actively doing it. A head shake without verbal protest is coded as a gestural imperative protest (If the child shakes their head and also says “no”, this will be coded as a verbal imperative and not as gesture so that when these categories are combined for a total count, this will only count as one protest).

4.7.1.3 Agree - This is when the child actively agrees for the Cat to ‘do it’, or gives the Cat helpful direction such as “Cat put it there” or “it goes there”. Again, if the
child gestures in agreement such as a nod, then this will be coded as gestural agreement if they do not vocally agree.

4.7.2 Action Coding

There were three coding categories for protesting actions, which show that the child is attempting to prohibit the rule violation but without vocalisation. These three categories are:

4.7.2.1 Preventative action - Attempting to stop the violation before Cat has the block. For example, the child attempting to stop Cat from taking the block in the first place by covering all of their blocks. This involves anticipating the violation and physically attempting to prevent it, even if this fails or the attempt is very short lived.

4.7.2.2 Active action - Attempting to stop the violation after Cat has the block, or before they think Cat may try and get the block, but before he puts it down on the board. For example, the child trying to take the block out of Cat’s paws whilst Cat is picking it up and placing it down. This involves any attempt even if this fails. Also the child quickly taking another block of the correct colour and putting it down on the space before Cat has the chance to, is considered an active action. Further, the child really rushing to get the block even if Cat is not trying to get it but it seems they think he might, is coded as active action.

4.7.2.3 Corrective action - Attempting to correct the violation after it has happened. For example, the child taking the brick off when Cat has placed it down on the board and doing it again themselves or swapping it for another (same colour) block, so as to correct the rule violation.
4.7.2.4 Agree - This is when the child actively agrees and gives the block to Cat or actively assists/agrees that they should take the turn. This includes pointing and showing Cat where the piece goes.

Each target utterance was coded and then what action (if any) that they took and whether they protested and to what level summarised for each of the three prompts per game. This was for the 9 prompts in the high-conflict condition and the 9 prompts in the low-conflict condition. It could be that within a single prompt the child will go from attempting to protest to agreeing and this was also coded.

20% of data was second coded by a second coder and a reliability check run using a Cohen’s Kappa for the protest data; coders matched in 98.38% of ratings, Kappa = 0.973 and for the action data; coders matched in 97.5% of the ratings, Kappa = 0.962.

4.8 RESULTS

The protest data was compared across conditions to establish if children protest differently in the high-conflict condition, when Cat has the block, to the low-conflict condition, when Cat just suggests that he should ‘do it’, but doesn’t actually take the block. The coding conditions of Normative, Imperative (which collapsed any gestural protest such as headshake into the total count) and Agree (which also included any nods into the count) are compared as a total per child for the high-conflict versus low-conflict conditions.

<table>
<thead>
<tr>
<th>Age</th>
<th>Normative</th>
<th>Imperative</th>
<th>Agree</th>
<th>Normative</th>
<th>Imperative</th>
<th>Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>43</td>
<td>302</td>
<td>88</td>
<td>22</td>
<td>168</td>
<td>22</td>
</tr>
<tr>
<td>2</td>
<td>51</td>
<td>229</td>
<td>150</td>
<td>21</td>
<td>188</td>
<td>27</td>
</tr>
</tbody>
</table>

Table 4.5. Number of each protest types for both the two-year-olds and three-year-olds in both the High-Conflict and Low-Conflict conditions.
A three-way MANOVA was conducted with two between-group factors, both with two levels; Order (High-Conflict first, Low-Conflict first) and Age (younger group, older group) and one within-subjects factor, also with two levels; Condition (High-Conflict, Low-Conflict). There were three dependent variables; Number of Normative protests, Imperative protests and Agreements in either condition. Results revealed a significant main effect of Condition for all three protest types. Normative protest was significantly higher \( F(1,48) = 5.406 \ p = .024 \) in the high conflict of agent condition \( (M = 1.808, \ SE = 0.442) \) compared to the low conflict condition \( (M = 0.827, \ SE = 0.248) \). Imperative protest was significantly higher \( F(1,48) = 17.743 \ p < .001 \) in the high conflict of agent condition \( (M = 10.212, \ SE = 0.930) \) compared to the low conflict condition \( (M = 6.846, \ SE = 0.652) \). Agreement was significantly higher \( F(1,48)= 37.509 \ p < .001 \) in the high conflict of agent condition \( (M = 4.577, \ SE = 0.636) \) compared to the low conflict condition \( (M = 0.942, \ SE = 0.203) \). There was a significant interaction between Condition and Age, for the Imperative Protest \( F(1,48) = 5.011, \ p = .030 \) with the two year olds protesting less often, imperatively \( (M=8.808, \ SE = 1.315) \) than the older age group/three year olds \( (M = 11.615, \ SE = 1.315) \) for the high-conflict condition but protesting more often imperatively \( (M = 7.231, \ SE = 0.922) \) than the three year olds \( (M = 6.462, \ SE = 0.922) \) for the low-conflict condition.

Table 4.6. Number of each response type for both the two-year-olds and three-year-olds in the High-Conflict condition.

<table>
<thead>
<tr>
<th>Age</th>
<th>Verbal ONLY response</th>
<th>Action ONLY response</th>
<th>Both</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>30</td>
<td>19</td>
<td>174</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>61</td>
<td>18</td>
<td>147</td>
<td>8</td>
</tr>
</tbody>
</table>

Children protested significantly more in the high-conflict condition and next the data will be analysed to investigate if there is a difference between the age
groups for whether children use an action protest or agreement compared to a verbal protest or agreement. The action data for the low-conflict condition is not compared to the action data in the high-conflict condition as in the high-conflict condition the puppet actively takes the block and so the children’s need to actively protest in the low-conflict condition isn’t the same. Therefore we coded each child’s response in the high-conflict condition as either verbal response only, non-verbal action response only, both verbal and action response or no response and compared this between the two age groups, see Table 4.6. We ran a one-way MANOVA with a between factor of Age (younger group, older group). Results revealed a significant main effect of Age for the verbal only response type. Verbal only response was significantly higher ($F (1,51) = 4.882, p = .032, \text{partial } \eta^2= 0.089$) in the younger children ($M = 2.346, SE = 0.382$) compared to the older children ($M = 1.154, SE = 0.382$). There was no significant difference between the younger and older children for the number of action only responses ($F (1,51) = 0.010, p = .921, \text{partial } \eta^2= 0.000$), both action and verbal response ($F (1,51) = 2.236, p = .141, \text{partial } \eta^2= 0.043$) or no response ($F (1,51) = 0.290, p = .593, \text{partial } \eta^2= 0.006$).

Finally it was investigated if the different strategies changed across the 9 rule violations in the high-conflict condition. See figure 4.1 for the pattern of different protest types across the nine different rule violations. Figure 4.1 reveals a spike for each of the first trials of each of the three games for the Imperative protest and the Action protest as well as for the Normative protest for the first two games. The Agreement, whether verbal agreement or action, was lower on the first trial of each game and increased across the game.
4.9 DISCUSSION

Study two investigated the two and three year olds’ protest and investigated the different levels of protest used. It was investigated if the children were enforcing the rules normatively, if there were age differences between the protest types and further, if the repeated rule violations caused children to ‘up’ their protest type.

Results showed that children were protesting more in the high conflict condition compared to the low conflict condition for all three protest types; Normative, Imperative and Agreements. This shows that children were able to establish the need for protest when Cat was actively taking their block, compared to when Cat just suggested this but did not do it. However, for the agreement category, this shows that children were agreeing to ‘Cat doing it’ more in the condition when Cat had already taken the block. This may have been a strategy in which the children just admitted defeat and let Cat take the turn, or because of a desire to
avoid conflict, or even that the children were unsure and just agreed with Cat because of this. The results did not reveal order effects and therefore there was no difference in the protest types dependent on the order that the children received the conditions, showing that children’s protest types did not change dependent on whether the high-conflict or low-conflict condition came first.

The results revealed that there was no main effect of age and therefore as a whole there was no difference between the age groups in overall level of protest. This shows that both age groups were equally able in protesting normatively, imperatively and also agreed, to the same extent. However, there were developmental effects found, with a significant interaction found between condition and age for the level of imperative protest. The three year olds increased their imperative protest such as “no, I do it” or simply “no”, when Cat actively took the block and used this considerably less when Cat didn’t take the block. Whereas the two year olds used this protest type across both conditions at a similar rate, with only a very small difference between the conditions. This age difference could potentially show a lower level of understanding and awareness of the game and in identifying the level of protest needed. It may be that the three year olds could clearly identify the need for protest more explicitly in the high-conflict context, whereas the two year olds still protest almost as much at the mere suggestion of Cat taking the block. This shows that although the two year olds are capable of using these linguistic forms, they are less sophisticated in understanding and altering their use in the appropriate context, than the three year olds.

The results also revealed a developmental effect with response types, when categorised as either just vocalising, both verbal and action, just action or no response. There were no age differences for the number of action only responses,
showing that the number of children who chose to just actively try and stop the rule breaks, did not differ between the two and three year olds. Also, the number of children who used both strategies to try and prevent Cat from breaking the rule did not differ, which shows that the younger age group were just as capable as the older group in using a range of strategies. However, the younger age group did rely on verbal only responses more often than the older age group. This could again show that the older children had more an awareness and altered their strategies, incorporating more of a variety, trying action protests, verbal protests and both, whereas more of the younger children may have stuck more with a verbal only strategy for longer, even if this wasn’t actually working to prevent Cat.

Children of both ages were normatively protesting more in the high conflict condition and this shows that they understand the rules and are attempting to protest against these in a way that enforces the norm and what should happen. The lack of age difference shows that the younger age group was also able to produce these normative protests as much as the older age group. It has been previously found that three years olds will normatively and imperatively intervene when a puppet makes a mistake, however it has been found that two year old children do not often protest normatively very often (Rakoczy, Warneken & Tomasello, 2008). In Rakoczy, Warneken & Tomasello (2008), the two year olds protest was said to be less distinctively normative, but this was interpreted as potentially showing an “embryonic form” of normative awareness, but with a lesser ability of explicitly vocalising that. However, these studies have rarely given the children more than one opportunity to protest in two separate experimental trials. This methodology only gave the child two opportunities and as the authors (Rakoczy, Warneken & Tomasello, 2008) pointed out, the children may not protest due to shyness and/or a
general reluctance to interfere with others and therefore their full normative awareness may not be vocalised. The current study gave the children three games and within each game there were three violations of the rule. This meant that the children had repeated opportunity to protest and further, more opportunity to maximise their protest type. However, it should be noted, as we see in Figure 4.1, children’s normative protest did not increase as trial number increased. So, although the repeated exposure will help maximise the opportunity for normative protest, it was not the case that this increased as the chances to protest increased. However, the children did have practice trials and were familiarised with the rules and therefore this sets up a situation in which children will grow confident in their knowledge of the rules and therefore perhaps increase the chance of them protesting normatively, from the start. The younger age group, therefore were more than capable of normatively protesting and producing utterances such as, “you should not do my red one”. And showing their awareness of the rules of the game and how it ‘should’ be played.

In reference to figure 4.1, we can see that children protest more Imperatively, Normatively and with Action protest in the first rule violation of the three games in the high-conflict condition and then this decreases over the other two rule violations. This then appears to increase again for the next game. This shows that children are protesting normatively and imperatively from the very first rule violation and it is not that this gradually increases over the nine trials. It is therefore possible that the two practice trials have familiarised them with how the game ‘should’ be played. As well as a reduction in protest over the three trials within each game, figure 4.1 also reveals a slight increase in agreement over the three trials within a game. This may reveal an underlying pattern that children will
not repeatedly explain the rules, but either try alternative strategies, or almost “give up” and agree to Cat’s way of playing. However, the importance of doing all of their own pieces will then be refreshed in between games and therefore children will have a new motivation to protest to any rule violations.

4.10 GENERAL DISCUSSION

In study one we investigated children’s *my*-for-*I* pronoun case marking errors as well as all self and other referring forms used by the two year old and three year old children, across conditions of high-conflict and low-conflict. In study two we further examined the data to analyse children’s protest types across the two conditions and between the age groups.

Children aged two and three years were found to produce a very low number (2) of *my*-for-*I* pronoun case marking errors. This highlights both the very rare nature of these errors as well as the importance of context and setting in children producing such errors. For the children who do assign a function of claiming agency and control to the form *my*, an experimental setting in which they interact with a puppet, may not be the most ideal context in which to elicit such errors. Children need to feel confident in order to attempt to control a situation and claim agency. A setting with observing adults (experimenters) and in which they play a game with a puppet controlled by an experimenter, may not be this optimum setting to create this feeling of confidence. As discussed, a more naturalistic setting increases the likelihood of capturing these errors, in which the child is in a familiar environment, with familiar interlocutors and therefore feels comfortable in attempting to claim agency and control. The resulting lack of errors therefore may actually reinforce the claim that children are using *my* in place of *I* in contexts in
which they wish the claim agency and control and it is in fact just the case that this is a very difficult context to establish in an unnatural, experimental setting.

As well as looking at children’s *my-for-I* errors, children were found to self-refer, refer to Cat more and verbally protest in a situation of high-conflict over agency compared to a situation with lower conflict over agency. Two year old children were just as competent as the three year olds in their correct *I-verb utterances*, there was no difference in error rates and the two year olds were just as capable as the three year olds of protesting normatively in reaction to the puppet’s rule break. This presence of normative protest from the younger age groups shows that in a context in which the rules are explicitly clear, children may protest normatively at a younger age than previously seen. The children in previous studies may have not protested out of shyness, general reluctance or perhaps because of being surprised and a little confused when the puppet first breaks the rule, however when given a game situation in which they have practice, they perhaps have built enough confidence to attempt to enforce this, albeit they still do not attempt to simply claim agency in the same way as they would when creating a *my-error*, there is still some attempt at explaining the rules in a fun, less demanding manner. Therefore the two year olds have the language ability to protest normatively and show their awareness of the norms and how the game should be played. Perhaps the focus of this particular game on agency and who should ‘do’ each of the pieces is also particularly engaging for the child in a way that it is in their interest to claim agency and therefore protest and correct the puppet’s mistakes. These results show that two and three year olds can not only understand rules and how others should act in accordance with certain rules, but enforce these rules and normatively criticise, protest and teach a puppet’s mistake in the game context requiring these
rules. It is not the case that the two age groups acted in identical ways, however, developmental effects were seen, with the three year olds seemingly reacting to the context more so than the two year olds with a bigger increase in imperative protest across conditions, whereas the two year olds use a similar level of this protest type across conditions. This shows a lower level of understanding to some degree and although the younger children show their ability to protest, this shows that their understanding of when this is necessary, isn’t quite as developed as the three year old children. Further, the younger children also used a higher number of verbal only protest types than the older group, again showing perhaps a more inflexible reaction. It could be that the younger children stuck with the same strategy of just saying “no” or “yes” or explaining the rules and if this didn’t work they kept with it, whereas the older children may have seen this wasn’t working in the high-conflict condition and tried differing strategies, again showing a deeper level of understanding and ability to react to the context accordingly.
4.11 REFERENCES


4.12 APPENDICES

4.12.1 Appendix A

The three colour matching games played in the two experimental conditions with corresponding coloured blocks.

4.12.1.1 Game one – Snake

4.12.1.2 Game two – Caterpillar

4.12.1.3 Game three – Snail
### 4.12.2 Appendix B

The four stories used for the priming of *my*.

<table>
<thead>
<tr>
<th>Story</th>
<th>Child</th>
<th>Character</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>No. of words</th>
<th>No. of stressed MY</th>
<th>No. of sentence initial My</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Story 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Line 1</td>
<td>Child 1</td>
<td>Boy</td>
<td>Lost</td>
<td>my</td>
<td>teddy</td>
<td></td>
<td></td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Line 2</td>
<td>Child 1</td>
<td>Boy</td>
<td>Can't</td>
<td>find</td>
<td>my</td>
<td>teddy</td>
<td></td>
<td></td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Line 3</td>
<td>Child 1</td>
<td>Boy</td>
<td>Oh,</td>
<td>there is</td>
<td>my</td>
<td>teddy!</td>
<td></td>
<td></td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Line 4</td>
<td>Child 2</td>
<td>Girl</td>
<td>No,</td>
<td>this is</td>
<td><strong>MY</strong></td>
<td>teddy!</td>
<td></td>
<td></td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Line 5</td>
<td>Child 1</td>
<td>Boy</td>
<td><strong>MY</strong></td>
<td>teddy!</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Line 6</td>
<td>Child 1</td>
<td>Boy</td>
<td><strong>Oh,</strong></td>
<td><strong>MY</strong></td>
<td>teddy is</td>
<td>here</td>
<td></td>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>24</strong></td>
<td><strong>3</strong></td>
<td><strong>1</strong></td>
</tr>
<tr>
<td><strong>Story 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Line 1</td>
<td>Child 1</td>
<td>Girl</td>
<td>Have you</td>
<td>seen</td>
<td>my</td>
<td>football?</td>
<td></td>
<td></td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Line 2</td>
<td>Child 1</td>
<td>Girl</td>
<td>There is</td>
<td>my</td>
<td>ball!</td>
<td></td>
<td></td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Line 3</td>
<td>Child 2</td>
<td>Boy</td>
<td>No,</td>
<td><strong>MY</strong></td>
<td>football!</td>
<td></td>
<td></td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Line 4</td>
<td>Child 1</td>
<td>Girl</td>
<td><strong>MY</strong></td>
<td>ball!</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Line 5</td>
<td>Child 1</td>
<td>Girl</td>
<td>Where is</td>
<td>my</td>
<td>ball</td>
<td>then?</td>
<td></td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Line 6</td>
<td>Child 1</td>
<td>Girl</td>
<td><strong>Oh,</strong></td>
<td><strong>MY</strong></td>
<td>ball is</td>
<td>there!</td>
<td></td>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>24</strong></td>
<td><strong>3</strong></td>
<td><strong>1</strong></td>
</tr>
<tr>
<td><strong>Story 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Line 1</td>
<td>Child 1</td>
<td>Boy</td>
<td>Where have</td>
<td>my</td>
<td>sweets</td>
<td>gone?</td>
<td></td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Line 2</td>
<td>Child 1</td>
<td>Boy</td>
<td>Are my</td>
<td>sweets</td>
<td>over</td>
<td>here?</td>
<td></td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Line 3</td>
<td>Child 1</td>
<td>Boy</td>
<td><strong>My</strong></td>
<td>sweets!</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Line 4</td>
<td>Child 1</td>
<td>Boy</td>
<td>You have</td>
<td>got</td>
<td><strong>MY</strong></td>
<td>sweets!</td>
<td></td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Line 5</td>
<td>Child 2</td>
<td>Girl</td>
<td>These are</td>
<td><strong>MY</strong></td>
<td>sweets</td>
<td></td>
<td></td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Line 6</td>
<td>Child 1</td>
<td>Boy</td>
<td>No,</td>
<td><strong>MY</strong></td>
<td>sweets</td>
<td></td>
<td></td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>24</strong></td>
<td><strong>3</strong></td>
<td><strong>1</strong></td>
</tr>
<tr>
<td><strong>Story 4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Line 1</td>
<td>Child 1</td>
<td>Girl</td>
<td>Looking for</td>
<td>my</td>
<td>book</td>
<td></td>
<td></td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Line 2</td>
<td>Child 1</td>
<td>Girl</td>
<td>Have you</td>
<td>seen</td>
<td>my</td>
<td>book?</td>
<td></td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Line 3</td>
<td>Child 1</td>
<td>Girl</td>
<td>That is</td>
<td>my</td>
<td>book</td>
<td>there!</td>
<td></td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Line 4</td>
<td>Child 2</td>
<td>Boy</td>
<td>No,</td>
<td>this is</td>
<td><strong>MY</strong></td>
<td>book!</td>
<td></td>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Line 5</td>
<td>Child 1</td>
<td>Girl</td>
<td><strong>MY</strong></td>
<td>book!</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Line 6</td>
<td>Child 1</td>
<td>Girl</td>
<td><strong>Oh,</strong></td>
<td><strong>MY</strong></td>
<td>book!</td>
<td></td>
<td></td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>24</strong></td>
<td><strong>3</strong></td>
<td><strong>1</strong></td>
</tr>
</tbody>
</table>
4.12.3 Appendix C

The corresponding pictures for story one.
4.12.4 Appendix D

Parent Questionnaire.

SCHOOL OF PSYCHOLOGICAL SCIENCES

Title of Project: Children’s Pronoun Case Marking Errors

Parent Questionnaire

Our study is investigating children’s use of different pronouns. Children produce utterances such as “I do it”, but quite often they also use unconventional forms such as “my do it”, “me do it” or their own name such as “Sarah do it”.

We are interested in these types of utterances in which a child uses pronouns and/or their own name, before a verb, for example saying things like ‘Me like cake’, ‘My don’t want it’, ‘Sarah needs it’ (but not when children just say things like “my ball” or “it hit me”).

Please could you fill in this very short questionnaire alongside the parent consent form, just to indicate if you have noticed your child using any of these different self-referring forms and how often you think they do this.

Please circle the self-referring forms you have heard your child produce before a verb in the last month, for example: “Sarah want it”, “My hold the dolly”, “Me kick it” or “I want a blue one”.

I                        My                               Me                                   Own Name

Which two of the self-referring forms does your child produce most often in sentences such as “Sarah hold it”, “my want to hold it”, “me not do it” or “I want a blue one” (if your child only uses one form in this type of sentence then only select this one self-referring form).

I                        My                               Me                                   Own Name

If your child uses any unconventional self-referring forms, such as “me”, “my” or their own name (as in the above examples, not including conventional forms such as “my toy”, “my crayons” or “Sarah’s toy”), then how often would you estimate that they do this?

“my”-                        A few times a week     A few times a day       More than ten times a day        Almost always

“me”-                        A few times a week     A few times a day       More than ten times a day        Almost always

“own name”-                  A few times a week     A few times a day       More than ten times a day        Almost always

Have you noticed if there is any specific situation in which your child is likely to use something other than “I” to refer to themselves?

........................................................................................................................................
........................................................................................................................................

Do you have any additional comments on your child’s self-referring forms?

........................................................................................................................................
........................................................................................................................................

Do you agree for the above information to be used as part of the current study and for future related studies?

Name of child: …………………………………………………………… Child’s D.O.B …………………

Name of parent: …………………………… Signed: …………………………Date: ……/……/……

This project has been approved by the
School of Psychological Sciences Research Ethics Committee
4.12.5 Appendix E

The sticker sheet used for each child. Children had a target of 6 stickers and earned a sticker every time they did all of their pieces in each game.
Chapter 5: Discussion and Conclusions

The present thesis investigated children's pronoun case marking errors during language acquisition. As well as this, children’s protest in specific contexts and the self-referring forms used during protest, was investigated. As discussed in Chapter one, there are two main theoretical frameworks which attempt to explain children’s language acquisition, namely the Constructivist perspective and Generativist perspective. This thesis took the stance of a Constructivist perspective in investigating children’s pronoun case marking errors. As pointed out previously, very little research has specifically investigated children’s genitive-for-nominative first person pronoun case marking errors and it was these my-for-I errors that were the focus of the first two papers of this thesis. Because of the emerging importance of function in these errors, the third paper took a broader focus on children’s pronoun use but with a further investigation into protest and the pronoun use and protest types. I will start by summarising our findings. Next I will evaluate these findings in respect of the broader literature and how they fit with the Constructivist perspective. I will then discuss limitations of the present studies and offer suggestions for future research. I will then conclude.

5.1 SUMMARY OF FINDINGS

5.1.1 The relation of input to errors

5.1.1.1 Frequency and Strings in the Input

Chapter two focussed on the role of frequency and strings in the input in children’s genitive-for-nominative first person pronoun case marking errors. A detailed corpus analysis first found that the raw frequency of my in the input, in relation to all word tokens in the input, and the proportional frequency of my in the
input, (either as a function of all first person pronouns, or of all first person referents) during a defined pre-error period, did not relate to the rate of children’s errors. Further, neither the frequency of *am-I-verb* strings in the input nor *my-noun* strings that with lexical items that can also be verbs related to the rate of children’s errors. Therefore it was concluded that children were not learning *my-for-I errors* from direct learning of strings in the input or from the basic frequency of the first person genitive form.

5.1.1.2 Proportional Frequency of Different Forms in the Input

Chapter three investigated additional aspects of the input and its relation to children’s *my-for-I* error rates. The rate of pronominal input in caregiver speech in relation to proper name input was analysed. It was found that children receiving less than 90% pronoun input and therefore a higher rate of proper name replacements, made a higher number of *my-for-I* pronoun case marking errors. This suggests that a higher contrast of pronouns and proper name is not necessarily beneficial in mastering the pronominal system and that this lower level of modelling of pronouns can relate to a higher level of error. When we looked at specifically the proportion of pronouns versus proper name input for subject or possessive reference in isolation, the relation to the error rate was not found. It was not specifically a higher modelling of just *my* or *I* in relation to “mummy” replacements, but it was the level of modelling of the whole first person pronoun system, that was linked to the rate of errors. The higher levels of replacement with own name forms such as “mummy do it” and “mummy’s ball” appeared to cause children to be potentially unsure about the correct pronominal form and become innovative in their utterances.
5.1.2 Children’s prior use of forms

Chapter three also investigated the levels of pronoun use by each child before the start of the error period. It was found that children replaced I more (with my) when I was less entrenched before the onset of errors. This was calculated by the proportional use of I over all first person pronoun use. The less entrenched I was found to be, the higher the rate of children’s genitive-for-nominative replacements. The relative use of my (and me) was not related to children’s my-error rates. This suggests that it is not children simply overextending one strongly entrenched form, but that when children are unsure of the correct form, they can become innovative and are more likely to make errors.

5.1.3 The role of Function

Chapter three also investigated another important feature of the usage-based approach which is the role of function in children’s language learning and use. Study two of chapter three found that my-errors were associated with high agency and control and specifically with the child wanting to become the sole agent, to change the action being suggested or to change the possessor or owner of an object. This function was found to run throughout the error period, from the earliest my-errors studied, to the last. This lack of differences between the early and late errors shows that this function stayed present throughout the error period. Further, the same analysis on children’s me-errors and own name-replacements did not reveal this same function. Hence, this function has not been assigned to any replacements of I, such as children using me or own name pre-verbally, but specifically when assigning my as the subject, in place of I.
5.1.4 Errors in Protest

Chapter four experimentally investigated children’s pronoun use in a game context in which there was conflict over agency and control. It was found that children were using, on the whole, more self-referring forms and references to the other player, ‘Cat’, in the high-conflict of agent condition compared to the low-conflict condition. However, the rate of pronoun case marking errors was found to be very low with my-errors so low that a comparison across condition could not be made. Me-errors and own name replacements did not differ across condition, which is to be expected and supports the findings in chapter three which find that it is the function of my-errors that is associated with agency and control, rather than me or own name replacements of I. There were no age effects found, showing that the two year old children were just as developed as the three year olds in producing correct I-verb utterances. This shows that children have a firm grasp of the correct use of I as subject, by two and a half years, but the errors are still just as prevalent across the two age groups, with no age difference found between the level of errors.

5.1.5 Types of Protest

Study two in chapter four investigated two and three year olds protest types and investigated the different levels of protest used. It was found children were protesting more both normatively and imperatively in the high conflict condition and there were no age differences in this overall level of protest. This shows that children of two years are aware of the rules of the game and have the linguistic ability to express this normatively. There were some developmental effects found, with the two year olds imperative protests differing a lot less between the high conflict condition and low conflict condition, than the three year olds. This shows
the three year olds more sophisticated grasp on the context and their ability to adjust their protests accordingly. Further, the two year olds had a higher number of verbal only protests in the high-conflict condition only, compared to the three year olds. This could again show that the three year olds had a more flexible system with differing strategies, whereas the two year olds may continue with the verbal only protest for longer, even though this will not have worked.

5.2 HOW OUR RESULTS FIT WITH PREVIOUS LITERATURE FROM THE CONSTRUCTIVIST PERSPECTIVE

5.2.1 Input

The Constructivist approach views the linguistic environment of a child as a crucial factor in explaining any child’s language development. As well as the input, competing form-function mappings and meaning all interplay to explain children’s language learning and errors. This is in stark contrast to the Generativist approach which views case marking errors as a syntactic issue, explained in terms of the interaction between the Universal Grammar and the parameters of basic inflection and clause structure being set (Schütze & Wexler, 1996). The Constructivist perspective emphasises the role of input and the frequencies of forms and strings in this input in explaining children’s errors (Ambridge, et al., 2015, Tomasello, 2003). However, the Constructivist position is not that input is the only driving force. Ambridge et al. (2015) provided evidence for the importance of frequency but pointed out that we can’t expect a linear relationship between simple token frequency in the input and children’s language learning, as other factors such as utterance length in the input and serial position, will all interfere with simple frequency and cause certain forms to be learnt faster than others. Further, function
and saliency of certain lexical items will cause certain forms to be learnt faster, as well as this competition between forms (MacWhinney & Bates, 1989; Cameron-Faulkner et al., 2007) and each form’s level of entrenchment are all relevant factors in relation to children’s language learning and level of error from the Constructivist perspective. The results from chapter two and three, summarised above, show that simple frequency of forms in the input does not explain children’s my-for-I errors, which also adds to previous literature investigating children’s me-for-I pronoun case marking errors (Kirjavainen et al., 2009) in which it was found that children’s me-error rates did not relate to the raw frequency of me in the input.

Previous literature from a constructivist perspective has shown that children can often produce errors from hearing strings in the input and then reproducing part of the string, out of context from the larger utterance thus producing ungrammatical errors (Rowland & Pine, 2000; Theakston et al., 2003; Freudenthal et al., 2007; Theakston & Lieven, 2008; Kirjavainen et al., 2009). The findings in chapter two found that this was not the case for children’s my-for-I pronoun case marking errors. With other error types, the relevant grammatical strings are heard in the input such as “Let me do it” and the child may produce “me do it” (Tomasello, 2003) or children producing auxiliary verb omissions such as “he playing” because of questions in the input such as “Is he playing?” (Theakston & Lieven, 2008). But, in chapter two, we showed that it is not the case that my-verb is actually modelled grammatically in the input. My-noun utterances are modelled with words that can also be verbs, but the function and semantics of these utterances will be quite different between verb use and noun use in the input. Therefore, the modelling of a string in the input such as “Let me get it” and the child producing the utterance “me get it” has the same basic meaning but the latter is grammatically incorrect.
However, the modelling of “That is my drink” and the child then producing “my drink the juice”, would involve the same lexical *my drink* string, but has a different core meaning in the utterance. This may explain why children do not learn from these strings in the input and there is no relationship between the proportional frequency of these and the rate of errors. The *am I-verb* strings in the input sound phonologically like a *my-verb* utterance and if this was being heard as a “What *my doing?”* or “Where *my going?”* utterance, this would have very closely related semantics and it would be as a straight forward modelling of this *my-verb* in the input. However, as pointed out, these utterances were found to be quite infrequent in the input and therefore this very low frequency could explain this lack of relation to the rate of *my-verb* errors.

The frequency of the pronominal forms in the input relative to the replacement of these by proper names was found to relate to children’s error rates. Although these findings differed from Smiley et al. (2011) as discussed in chapter three, this fits with a broader constructivist perspective regarding children learning from the input that they receive. Braine and Brooks (1995) claim that exposure to a form in one construction will strengthen the probabilistic evidence that it cannot be used in other constructions that they have not previously heard in the input. Therefore, the higher the frequency and increased modelling of the different first person pronominal forms in grammatical constructions, for example hearing caregivers utter “I will get it” and “my handbag”, the more likely children will learn this and therefore will be less likely to think that it is acceptable to say “my will get it”, which will not have been heard in the input. The present findings in chapter three are therefore consistent with this theory of entrenchment. The children who received less of this pronominal modelling and heard more “mummy will get it”
and “mummy’s handbag”, were receiving less evidence for the correct use of the
different first person pronouns, because this is modelled less and therefore children
are more likely to overextend these forms and produce errors such as my-for-I
errors. The finding that it was the overall level of modelling of pronominal forms,
rather than just the level of subject or possessive forms individually that related to
error rates is also consistent with the broader Constructivist perspective. Children
are less likely to replace I with my if I itself was highly frequently modelled but my
alone wasn’t. This is because I would be highly entrenched and therefore children
will be less likely to replace it. And the same, vice versa, children would be less
likely to overextend my if it individually had been frequently modelled, but I
hadn’t, because children would be more aware of the correct construction in which
to use it.

This entrenchment is further found with the link between children’s own
prior usage of pronouns and their individual error rates. The children who had a
lower relative use of I compared to me and my produced a higher rate of errors. This
is consistent with the constructivist theories, in that children who had a lower
entrenched use of I were more likely to replace it erroneously with another form as
it has not yet built up a strong enough representation. Thus our results have found
links between the frequency of forms in the input, which add to our understanding
of just how frequency and competition between forms interact in development.
However, just as Pinker (1989) suggested and further work by Ambridge et al.
(2008) supported, semantics also plays a role in the error rates of children and we
will now turn to discuss the role of function in relation to these my-for-I errors.
5.2.2 Function

Researchers working from a usage-based approach argue that children learn to use language as a communicative tool and the function of the utterance is the driving force behind using language. Chapter three investigated the role of function in children’s *my*–*for-I* pronoun case marking errors. It was found that children may be overextending the possessive function of *my* to contexts outside of this noun possession. This was causing children to use *my* as a possessive over actions in contexts involving high agency and control, causing *my-verb* errors. This supports previous work by Budwig (1989; 1995) and also provides evidence for the important role of function in language learning. It is claimed (Slobin 1981; 1985) that children do not always copy the input, but become creative and may assign specific functions to certain forms that are not present in the adult grammar (Karmiloff-Smith, 1979). This is exemplified by the evidence that children making a high number of these *my-for-I* errors had assigned this function of claiming agency and control to the first person genitive form. This then causes errors to be made with the incorrect form used in place of the correct form because children have not yet assigned adult-like form function mappings. This fits in with previous research (Cameron-Faulkner et al., 2007) which concluded that if children assign a function to several forms or if the input does not provide evidence for one-to-one mapping of form to function, then children can take longer to learn these mappings. This functional distinction made between different pronominal forms by children will cause form-function mappings that are not present in the adult system. It should be noted that children will not use this form (*my*) to express high agency and *I* for everything else, because of the competition in the system when children are
wishing to express a certain function. This brings us to discuss the competition model (Bates & MacWhinney, 1989) in relation to these errors.

5.2.3 The Competition Model

Bates and MacWhinney’s (1989) Competition Model focuses on the mapping between forms and function in children’s language learning. As many forms can be mapped onto several functions, the process of learning these form-function mappings can be a difficult one. Chapter three found that children who were making a high number of *my-for-I errors* had made a form function mapping that was not present in the adult grammar, with *my* seemingly having a function of expressing high agency, control and possession. It has been claimed (Karmiloff-Smith, 1979; Budwig, 1995) that once children have identified a function, they like to distinguish this from other functions by using a different word specifically for this function, and in this case, *my* to express agency and control. This distinction can possibly be more salient at this point in development as possession and attempting to claim agency may be particularly important at two years of age. These findings can be directly related to the Competition Model. This assignment of two forms, will cause competition between this form, *my* and the form that is used for this function in the adult grammar, *I*. If children have associated the form *my* with the function of high agency and control this may well win out over *I*, which may be associated with actions with agency and control but on a lower level, which is what was found in Chapter three.

It is important to note that children will not make these errors absolutely, because of the competition between the frequency of hearing the form in the correct way in the input and the child’s own system and operation, aiming for a one-to-one
mapping, which will push for the error. The results summarised above in chapter four found that children’s level of correct pronoun use did not differ between the two and three year olds. As well as this, the levels of pronoun case marking error, although low overall, did not differ between the two age groups. This shows that both age groups are producing correct *I-verb* utterances in the same developmental stage as they are producing pronoun case marking errors. This provides further evidence for the competition of more than one form being used at any point in development to express a certain function (MacWhinney & Bates, 1989) at any one time.

The input that children receive is an important factor in the competition model and the finding that children who receive less pronominal modelling actually make a higher proportion of errors also fits with this model. These children will have had less evidence of mapping each of the first pronominal forms (*I, me, my*) onto their correct grammatical function. They instead received more own name input which is the same across several functions, for example “Mummy do it”, “Mummy’s ball” and “get it for Mummy”. Thus ‘Mummy’ is being used across all three grammatical roles. This doesn’t allow different forms to map the different functions, but just has one form for multiple functions. This lack of modelling can lead to children making form-function relations not present and these will continue until children have received more positive evidence of the correct adult-like form function mappings to win out over the innovative mapping, in this case *my-for-I* replacements.
5.2.4 The Role of Function in other Pronoun Case Marking Errors

This functionalist based investigation into children’s *my-for-I* errors in chapter three showed that children were assigning the function of high agency and control on to the form *my*. However, these findings also showed that this same function was not being assigned to other forms used in place of *I*. Results showed that children were not using *me* or their *own-name* pre-verbally to express high agency or control. However, this does not mean that children have not assigned some other function onto this unconventional use of these forms. These findings highlighting the importance of function in children’s *my-for-I* errors may not be specific to *my-for-I* errors in particular. Children make form-function mappings when they are acquiring language and therefore this erroneous or unconventional form function mapping with *my* has led to a higher rate of errors. This could be the case for children making other pronoun case marking errors, such as “her do it” and “me have it”. It could be that as well as input based explanations to these errors (Tomasello, 2003; Kirjavainen et al., 2009) which show that children make these errors because of hearing strings in the input, children are also assigning a specific non-conventional function to these forms (causing a non-conventional use). It is not claimed that the input explains 100% of these other (*me/her*) pronoun case marking errors and therefore functional based explanations should also be explored. It was found that aspects of the input increased children’s vulnerability to making these *my-errors* and this could be the case for other pronouns in the whole pronoun system. It is possible that previously seen input based explanations (Tomasello, 2003; Kirjavainen et al., 2009) are a part of the explanation, but that function may play an important role for other pronoun case marking errors, as it does for *my-errors*. However, children’s use of *my* may be particularly vulnerable to errors as it
is such a salient form to children at the age of two to four years. Children’s claim of
possession over items is an important function and so is their independence and
desire to control tasks, be the main agent or decide what is happening. From this,
children may assign a form that they have learnt highlights possession and control
over an object, on to control over an action. It could be the case that incorrect form-
function mappings are therefore more likely with my than with other pronouns such
as her, me or him. If this is the case, it may be that my having such a strong
meaning, makes the function of this pronoun more important and therefore more
open to overextension. If this is the case, it is possible that function may not play as
big a part in explaining pronoun errors for other parts of the system, but this is
something that could still be investigated further.

5.2.5 Children’s ability to protest

The results from chapter four also showed that two and three year olds have
the ability to normatively protest. In a context which has rule breaks and norm
violations, children as young as two were able to not only disagree, but search for
the linguistic forms to explain why this rule break should not be happening. The
usage-based approach stresses the importance of function in language use and the
results of chapter four exemplify that children will find the forms to express their
desired function. The children in the study, not only used imperative protest, but
searched for normative language to communicate that the puppet taking their pieces
was ‘not right’. The majority of the children in the study protested with the rule
break and the particularly engaging game motivated children to want to express
what they thought the issue was. The developmental effects observed, showed that
although two year olds were competent in protesting using normative language,
when protesting imperatively, the younger children were less sophisticated than the older group at adjusting their level of protest accordingly. The older age group reduced their level of imperative protest significantly between conditions, whereas the younger age group did not. Just how this ability to protest is acquired and whether this development fits within a Constructivist framework is something that could be investigated further, see future research below.

5.2.6 Multifaceted explanation of children’s errors

The results from Chapters two and three have shown that children’s my-for-I errors do not have one single explanation. This relates to the Constructivist perspective that does not claim a direct link between frequency of forms, or strings in the input and error rates (Ambridge et al., 2015), but that there is an interplay between sentence position and frequency as well as between the function and saliency of certain forms. Further, only a third of children’s my-verb errors were found to have the function of agency and control and this clearly highlights that this assignment of function is not the whole picture.

The current findings add to the literature showing that children’s my-for-I pronoun case marking errors can be explained by a combination of factors. The interaction between children’s linguistic environment, the pronoun modelling that they receive in comparison with just a proper name form in all positions can lead to children being unsure of the correct form. The level of entrenchment of pronominal forms at a certain stage of development then also leads to competing forms playing out. This then can allow for children to assign a form-function mapping which will lead to the use of an incorrect form in place of the correct one. However, this form-function mapping will be competing with the input which models the correct form-
function mapping. See figure 5.1 for a diagram of the different factors all influencing the likelihood of error being increased. This assignment of a function to the form, will also increase the likelihood of errors in certain contexts. This interplay between the input and function will cease as the modelling of the different forms is increased and the child’s representation of these forms becomes more adult-like and leading these innovative form-function mappings to cease.

Figure 5.1. A Multifaceted Explanation of Children’s *my-for-I* errors. A model of the several factors all combining to increase the chance of error.
5.3 LIMITATIONS OF THE PRESENT STUDIES AND SUGGESTIONS FOR FUTURE RESEARCH

5.3.1 Corpus Analysis

A methodological issue found in chapters two and three in the corpus studies is the difficult task of capturing these *my*-for-*I* errors: both capturing children who are making the errors at the right time, as well as sampling densely enough to capture a high number of errors. This can result in only a small number of children who actually make this kind of error being available. This is further reduced by number of these children that have sufficient data available in order to capture enough of these errors as well as having input data available in order to establish the context in which these errors are spoken. It may be the case that the sampling contexts in the corpus data are very narrow and therefore may not be the optimum context to capture this particular error type. For example, playing quietly with their mother and no other children present may not be the optimum situation that will prompt the child to attempt to claim agency and control. It has been shown that these *my*-errors are made more when the child wishes to express control and therefore children may make a considerably higher number of these errors in other conversational contexts. A challenge would be to find data in which children at the age at which they are likely to make these errors are interacting in a wider set of contexts and potentially with more than just their caregiver. However, corpus data with more than one child present can be difficult to use due to issues identifying speaker. Further, the density of recordings in the majority of the corpus data available is on average an hour every fortnight, which is a problem as this causes a lot of errors during the error period to be missed and the number of individual
errors captured to be low. This is an issue with corpus analysis as the error period in several children may be missed or the number of errors captured be dramatically reduced.

5.3.2 Experimental issues

A methodological issue found within chapter four is the issue with an experiment in that it is a one shot setting and therefore capturing a high pronoun case marking error rate proved to be very difficult. As discussed, the priming used in the experimental methodology did not appear to increase the levels of my produced and there were only two my-for-I errors made during the experimental trials. The children in the experimental setting appeared to be more focussed on protesting against the rule break than asserting their own control over the situation. This could perhaps be a result of an unnatural setting in which children are less confident and the unfamiliar environment may make them feel less in control. The children certainly understood that the puppet was doing wrong and protested against this, but not by using the erroneous my+verb structure.

Further, we had a parental questionnaire asking caregivers to state how often they have heard their children make any pronoun case marking errors or own name replacements. This seemed to be quite a difficult task for parents to do in retrospect. They may not have particularly paid attention to these aspects of their child’s grammar or if they had, they may be unsure as to how often these errors are made. This made the parental reports less reliable than they could have been.
5.3.3 Future Research

This thesis took a Constructivist stance on children's language learning, however, in order to thoroughly investigate a generativist explanation to these errors, the suggestion by Schütze (1997) should be thoroughly addressed. Schütze claimed that children may make genitive-for-nominative replacements in sentences that lack both tense and agreement. Examples that counter this argument such as “My taked it off” are disregarded as noise. In a similar approach to Pine et al., (2005) the rate of verbs marked for tense OR agreement occurring with genitive subjects should be analysed and if this was found to be higher than a level that could be considered ‘noise’ (Pine et al. (2005) used 10%) then this claim by Schütze (1997) could be disputed.

In an attempt to capture a higher rate of pronoun case marking errors in corpus data, a random sampling technique cannot really be relied upon. One possible solution is to recruit caregivers who can use diaries to record their child’s self-reference. This will allow children who start to make a high number of errors to be identified relatively soon after the error period begins. From this, these children could then be recorded intensively and for more than an hour or so every fortnight. This would then also allow for targeted recording sessions over a range of different activities. Therefore this could capture speech between caregiver and child, not just in a quiet and quite artificial hour of play time, but during a range of activities, for example during morning routine, meal times, singing time, tidy up time, as well as play times. This perhaps could increase the likelihood of capturing a higher rate of errors and instead of recording every two weeks for over a year, the recordings could be intense for the start of the error period and then intense again as the errors are noted to be reducing. So the recording wouldn’t have to be too much of a
burden, by being intense for a long period of time, but may capture a rich amount of data in the more intense, targeted short term.

Future research could attempt to address the data capturing issue found in the experimental studies, by running a series of observational and experimental studies side by side. One approach could be to record children interacting in contexts involving high and low agency and control in a series of settings and with a wider range of participants. For example, children’s language and behaviour could be analysed interacting in a game play setting with other children, firstly in a naturalistic game and then in a game where they are told rules to play the game. The same children’s behaviour could then also be analysed interacting with a puppet and with a caregiver in a range of game play situations on several different occasions. This could potentially maximise the possibility of case marking errors being captured, as these will be more natural settings, but also as children’s behaviour will be analysed for more than just one context. This set up would allow for a controlled and specified context, as we had in our experiment, but with the more naturalistic setting of a corpus study, as the children will interact with others, have less formal instruction and also will participate several times over several months. This would be an attempt to capture children at a stage in which they are making these pronoun case marking errors.

Further, an additional suggestion for future research could be to recruit parents and caregivers to use the diary technique and monitor their children’s pronoun use before they are invited in for an experiment. This will solve the issue with the parental reports being unreliable as caregivers found the task of remembering their child’s errors, a difficult one. This would mean that the data would be more reliable as parents would be able to keep note of any errors as
children made them. But, as mentioned above, this could help to highlight children at a stage of making a certain error type. This technique could be used and children assigned to an experiment but only invited in to participate during a period of time that they are making errors. This would increase the possibility of capturing children who are prone to making these errors, if the methodology was a one shot experiment.

Future work could attempt to investigate how the different forms of protest develop even further. As discussed, the normative protest seen in both the two year olds and three year olds, did not increase across the nine rule violations. We therefore cannot conclude that the two year olds’ ability to protest normatively was captured because of an increased number of opportunities. However, there were two practice trials in which children were familiarised with the game, to make sure they knew how to play. Future work could look at the levels of protest across a high number of rule violations, without any practice trials. This will enable us to investigate if there is an increase, from the very first rule break, to the second or third. If this was the case, we can conclude that this repeated rule break, increases the likelihood of children protesting normatively, as a one shot setting may pass them by, because of a general shyness, reluctance or lack of confidence.

Future work could also attempt to investigate how this protest develops in children. It would be interesting to analyse the development of the linguistic forms needed to be able to protest normatively. Are children who are producing normative utterances such as “No, you should not do that colour” or “you are not supposed to do that one” hearing more normative explanations in the input that they hear? Or are they acquiring particular lexical items such as “should” earlier and does this relate to the input that they hear? It would be interesting to compare between
children who protest imperatively, just disagreeing but without normative language and those who make use of normative vocabulary and perhaps attempting to analyse at one point this shift occurs and is it because of an acquisition of certain lexical forms.

5.4 CONCLUSION

This thesis sits well within a constructivist perspective. The focus on genitive-for-nominative first person pronoun case marking errors has revealed a multifaceted explanation for children’s errors. The linguistic environment has been shown to influence children’s error rates by the levels of modelling available to allow children to produce adult-like form-function mappings. As well as this, the importance of function in children’s language learning is highlighted. Children’s desire to communicate is a driving force behind learning language and searching for the correct forms to express the desired function is paramount. Children at an age when becoming the agent of many actions is an important issue as independence and their ability to do more things by themselves increases, are likely to search for forms to express this function. At this same time the pronoun system is not yet fully mastered and so a lack of modelling of the correct form-function mappings, as well as a less entrenched I, will all interact and cause children to replace the correct nominative form with my. This form then may be associated with claiming possession and become overextended and associated with high agency and control.
APPENDICES

Chapter 3: Parent Information Sheet

School of Psychological Sciences

Parent Information Sheet

Title of project: Children’s Pronoun Case Marking Errors

Introduction
Many young children make pronoun errors around 2-4 years, in which they replace *I* with other pronouns such as *Me, She or Her* and hence produce utterances such as “me get it” and “her find it”. It is suggested that children produce these errors because they hear utterances such as “*Let me get it*”, “*Watch me do it*” and so on. Children also make errors in which they replace *I* with *My* and produce utterances such as “*my do it*”. We know rather less about why children produce these utterances in which they use *my* in place of *I*. Further, children often use their own name in their utterances and produce sentences such as “Warren do it” which is also of interest.

What will my child be asked to do if s/he takes part?
If your child takes part s/he will be invited to play several games with the investigator and a second experimenter who will be controlling a puppet. First your child will sit down with the experimenter by a table and play some simple games to get them comfortable with the experimenter. They will then be introduced to the puppet and told they will play some games and the puppet will be taking part too. They will play a practice game such as doing a jigsaw and asked to tell the puppet what to do as they complete the game. Next your child will be asked to simply look at several pictures and listen to a corresponding audio tape which will be simply narrating several stories. Your child will then be asked to play several small games with the puppet. The games will be simple tasks where they are to put coloured blocks in the correct place. The games will involve your child interacting with the puppet by telling the puppet what to do as they take turns to complete the game. Your child will receive stickers for completion of each game. Your child will be video recorded during the experiment to allow for their language and the context during the study to be analysed after the completion of the
experiment. We will be looking at which pronominal forms are used during the games with the experimenter and also with the puppet. The study will last approximately 30 minutes.

**Will my child’s data be confidential?**
Yes, only the researchers will have access to the data. If your child takes part s/he will be video-recorded for data coding purposes. The recordings will only be identified by your child’s date of birth. Consent forms and data will be held under responsibility of School of Psychological Sciences for 5 years following completion of the study for publishing purposes and will be stored in a lockable file and then destroyed by a shredder.

**Does my child have to take part?**
You do not have to allow your child to take part in the study. If you decide to allow your child to take part and then change your mind (whether before the child starts the study, half way through it or even after it has finished) you can withdraw without giving your reasons, and, if you wish, your child’s data will be destroyed. Your child will be asked whether s/he is happy to play a game with the researcher. If your child does not want to participate, s/he will not be included, even though you have given your consent for your child to participate.

**Benefits and risks**
There are no direct benefits to your child participating in this study. However, both children and parents generally indicate their enjoyment for taking part in such studies, and are generally enthusiastic about participating in research on child development. Similarly, there are no direct risks to taking part in the study either. If your child demonstrates any form of reluctance s/he will not be required to take part. Children who show any signs of distress will also not be required to take part.

**Where can I obtain further information if I need it?**
If you require any further details, please contact the principal investigator, Stacey Mcknight, by email: stacey.mcknight@postgrad.manchester.ac.uk

If there are any issues regarding this research that you would prefer not to discuss with members of the research team, please contact the Research Practice and Governance Co-ordinator by either writing to (in English) "The Research Practice and Governance Co-ordinator, Research Office, Christie Building, The University of Manchester, Oxford Road, Manchester M13 9PL, UK, by emailing: Research-Governance@manchester.ac.uk, or by telephoning +44 161 275 7583 or 275 8093."
Chapter 3: Caregiver Consent Form

SCHOOL OF PSYCHOLOGICAL SCIENCES

Consent form

Title of Project: Children’s Pronoun Case Marking Errors

*The parent should complete the following part of this sheet him/herself*

Please initial:

<table>
<thead>
<tr>
<th>Question</th>
<th>YES / NO</th>
<th>Initials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Have you read the Parent Information Sheet?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Have you received enough information about the study?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do you understand that your child does not need to take part in the study if s/he does not want to and if s/he does, you are free to withdraw him/her?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>* at any time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>* without having to give a reason for withdrawing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>* and without detriment to you?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do you understand that if your child takes part s/he will be video-recorded for data coding purposes?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do you agree for your child to take part in this study?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do you agree to use short video clips of your child’s participation for research presentations and teaching purposes?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do you agree for your child’s data to be used in future related studies?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Name of child: .................................. Child’s D.O.B .................................

Name of parent: ..................... Signed: ................................. Date: ....../....../......

Name of researcher: ................. Signed: ................................. Date: ....../....../......

This project has been approved by the School of Psychological Sciences Research Ethics Committee
Chapter 3: Headteacher/Nursery Manager Consent Form

Children’s Pronoun Case Marking Errors

HEAD TEACHER / NURSERY MANAGER CONSENT FORM

If you are happy for this study to be run in your school/nursery please complete and sign the consent form below

1. I confirm that I have read the attached information sheet on the above study.
   I confirm that I have had the opportunity to consider the information and ask further questions. I confirm that these have been answered satisfactorily.

2. I understand that children’s participation in the study is on the basis of opt-in parental consent.

3. I understand that the above study will be carried out on school/nursery premises during the normal school/nursery day.

4. I understand that I am free to withdraw the researchers’ access to my school/nursery at any time and without giving a reason.

5. I understand that all data will be confidential. The children will not be identifiable in any publications. Only the researchers will have access to the raw data.

I agree to allow the researchers access to my school/nursery to conduct the above study.

Name of Headteacher / Nursery Manager: __________________________
Signed: __________________________ Date: ______/_______/__________

Name of Researcher: _____________________________________________
Signed: __________________________ Date: ______/_______/__________

This project has been approved by the University of Manchester Research Ethics Committee
REFERENCES


Mahwah, NJ: Erlbaum.

*Journal of child language, 34* (02), 251-282.


