Paradigms and the Acquisition of Agreement
Morphology in German

by

William Karl Thompson

B.A., The University of Chicago, 1991

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS
in
THE FACULTY OF GRADUATE STUDIES
(Department of Linguistics)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
July 1995
©William Karl Thompson, 1995
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Linguistics

The University of British Columbia
Vancouver, Canada

Date August 4, 1995
ABSTRACT

In this thesis, I present arguments for a model of language acquisition with three characteristics. These are (1) Continuity in the abstract principles of Universal Grammar; (2) Lexical Learning, or the setting of syntactic parameters based upon the acquisition of morphology; and (3) Morpholexical Learning, which is the abstraction of morphological patterns and generalizations from a lexical database. Continuity accounts for what is invariant in language development. Lexical Learning accounts for what is language-particular, and which therefore must be learned. Morpholexical Learning accounts for the sequence of developmental stages observed in child language data. The main goal of this thesis is to demonstrate that Morpholexical Learning, in conjunction with paradigmatic structure in the lexicon, provides a model for the acquisition of inflectional morphology. I demonstrate this proposal with data on the acquisition of subject-verb agreement morphology in German.

In Chapter One, I present an introduction to the concerns and main proposals of this thesis. In Chapter Two, I motivate the existence of paradigmatic structure with three diachronic case studies. In Chapter Three, I return to the acquisitional debates introduced in Chapter One. I argue that the Continuity Hypothesis represents a preferable alternative to the Maturational Hypothesis. Next, I show that Lexical Learning of clausal representations is superior to the Lexical Projection Hypothesis and the Full Competence Hypothesis. I argue that Morpholexical Learning provides an answer to the Developmental Problem which Continuity and Lexical Learning create.

In Chapter Four, I provide an extended case study of the acquisition of subject-verb agreement in German. The construction of word-specific paradigms during the stages under examination accounts for the pattern of agreement errors which German children produce. In Chapter Five, I continue the analysis of paradigm mixture begun in
Chapter Two. The patterns of paradigm mixture attested Latin, German, and Icelandic are in essence identical to one another, which suggests universal principles of inflectional organization. In Chapter Six, I conclude the thesis with a sketch of how children develop from the "word-specific paradigm" stage to the "general paradigm stage".
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>iv</td>
</tr>
<tr>
<td>List of Tables</td>
<td>vi</td>
</tr>
<tr>
<td>List of Figures</td>
<td>viii</td>
</tr>
<tr>
<td>Chapter One: Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Chapter Two: Paradigms and the Mapping Problem</td>
<td>7</td>
</tr>
<tr>
<td>2.1. Mapping Between Meaning and Form</td>
<td>8</td>
</tr>
<tr>
<td>2.2. Paradigmatic Structure and Diachronic Change</td>
<td>13</td>
</tr>
<tr>
<td>2.2.1. Paradigm Pressure</td>
<td>15</td>
</tr>
<tr>
<td>2.2.2. Patterns of Paradigm Mixture</td>
<td>19</td>
</tr>
<tr>
<td>2.2.3. German Modal Inflection</td>
<td>27</td>
</tr>
<tr>
<td>Chapter Three: Continuity and Lexical Learning</td>
<td>31</td>
</tr>
<tr>
<td>3.1. The Status of UG Principles in Child Grammar</td>
<td>33</td>
</tr>
<tr>
<td>3.2. The Status of Clausal Structures in Child Grammar</td>
<td>36</td>
</tr>
<tr>
<td>3.2.1. Lexical Projection Hypothesis</td>
<td>37</td>
</tr>
<tr>
<td>3.2.2. Full Competence Hypothesis</td>
<td>39</td>
</tr>
<tr>
<td>3.3. Morpholexical Learning</td>
<td>43</td>
</tr>
<tr>
<td>Chapter Four: Acquisition of German Verbal Inflection</td>
<td>46</td>
</tr>
<tr>
<td>4.1. Morphosyntax of German Verbs</td>
<td>49</td>
</tr>
<tr>
<td>4.2. Subjects and Methodology</td>
<td>52</td>
</tr>
<tr>
<td>4.3. The Developmental Paradox</td>
<td>54</td>
</tr>
<tr>
<td>4.3.1. Root Infinitives in German Child Speech</td>
<td>57</td>
</tr>
<tr>
<td>4.3.2. Segmentation of Affixes</td>
<td>67</td>
</tr>
<tr>
<td>4.3.3. Lexical Storage in Paradigms</td>
<td>71</td>
</tr>
<tr>
<td>4.3.4. Discourse Properties of Parental Input</td>
<td>79</td>
</tr>
<tr>
<td>4.4. Conclusions: The Word-Specific Paradigm Stage</td>
<td>84</td>
</tr>
</tbody>
</table>
Chapter Five: Inflection Class Mixture and Acquisition

5.1. Latin 3rd Declension
5.2. The "Mixed" Classes of German Verbs
5.3. Icelandic Monosyllabic Feminine Declension

Chapter Six: Conclusion: From Word-Specific to General
LIST OF TABLES

Table 1: Latin 'mixed' 3rd Declension 20
Table 2: V to INFL Movement in the Germanic Languages 40
Table 3: Ages and MLU of Dorothy, Nicole, Katrin, and Andreas 52
Table 4: Proportion of Correct Use of Agreement Suffixes (Dorothy, Nicole, Katrin, and Andreas) 54
Table 5: Proportion of Correct Use of Agreement Suffixes (Daniel and Mathias) 55
Table 6: Use in Obligatory Contexts of Agreement Markers (Dorothy, Nicole, Katrin, and Andreas) 56
Table 7: Use in Obligatory Contexts of Agreement Markers (Daniel and Mathias) 56
Table 8: Distribution of Andreas' Finite and Nonfinite Verb Forms 58
Table 9: Clausal Position of Verbs Marked with [en] (Dorothy, Nicole, Katrin, and Andreas) 58
Table 10: Use of Nonfinite and Finite Forms in Modal Contexts 64
Table 11: Occurrence of Verbs in Finite and/or Nonfinite Variants (Dorothy, Nicole, Katrin, and Andreas) 69
Table 12: Instances of Multiple Subject Agreement Marking in Verbs (Dorothy, Nicole, Katrin, and Andreas) 70
Table 13: Inflections which Appear with 3rd Sing. Subjects (Daniel and Mathias) 78
Table 14: Verbs Used at Least Three Times with 3rd Sing. [t] 82
Table 15a: Classical Latin Masculine and Feminine Noun Declensions 89
Table 15b: Classical Latin Neuter Noun Declensions 89
Table 16a: Classical Latin Masculine and Feminine Noun Declensions: Pattern of Syncretism 90
Table 16b: Classical Latin Neuter Noun Declensions: Pattern of Syncretism 90
Table 17: Possible Ways for the Latin 3rd Declension to have Mixed
Table 18: Hypothetical Example of Paradigm Mixture
Table 19: Affixal Classes of German Verbs
Table 20: Possible Ways for the German 'Strong' and 'Weak' Conjugations to have Mixed
Table 21a: Original Set of Paradigms
Table 21b: Set of Paradigms after Loss of [s]
Table 21c: Set of Paradigms after Loss of [p]
Table 22: Icelandic Monosyllabic Feminine Nouns
Table 23: Possible Ways for Icelandic Monosyllabic Feminine Declensions to have Mixed
LIST OF FIGURES

Figure 1: Continuum of Morphological Fusion 10
Figure 2: Paradigm Structure Conditions 22
Figure 3: Latin 3rd Declension 24
Figure 4: Latin 3rd Declension (Revised) 92
Figure 5: Latin 3rd Declension: Maximum Possible Mixture 95
Figure 6: Hypothetical Paradigm Mixture 96
Figure 7: German Conjugational Classes 99
Figure 8: German Conjugational Classes: Maximum Possible Mixture 99
Figure 9: Hypothetical Paradigm Mixture: Diachronic Change 101
Figure 10: Icelandic Monosyllabic Feminine Declensions 103
1. Introduction

Traditional approaches to linguistic typology took the morphological properties of a language as its most important characteristic. Indeed, for 19th century linguists, a typology of language was tantamount to a typology of morphological systems. Sapir's well-known classificatory scheme falls within this rubric.1 Recent debate in linguistics has returned to this concern with morphology, and its role in accounting for linguistic diversity. Two main factors are responsible for this renewed interest: (1) the intimate connection between inflectional morphology and Functional Projections (cf. Pollock 1989), which in some current frameworks serve as engines for syntactic movement; and (2) the hypothesis that overt differences in morphology play a crucial role in the setting of syntactic parameters, which may themselves be restricted to properties of Functional Categories (Borer 1984; Chomsky 1989). In this thesis I will take these issues as the backdrop against which I investigate the mechanisms behind the acquisition of inflectional morphology. In particular, I claim that the paradigmatic dimension of language is as highly structured as the syntactic dimension, and that universal principles of language operating both in the syntax and in the lexicon combine to make the acquisition of language possible.

This investigation is guided by three main hypotheses on the nature of language acquisition. The first of these is the Continuity Hypothesis of Pinker (1984), which asserts that universal principles of grammar do not develop over time, but remain constant from the very beginning. Due to its restrictiveness, this assertion increases the explanatory power of an acquisitional theory. The second is the Lexical Learning Hypothesis of Borer (1984), in which the acquisition of morphological knowledge interacts with syntactic principles in such a way as to account for linguistic variation. Lexical Learning has the obvious appeal of placing the burden of language acquisition precisely where overt evidence is easily obtainable.

1 Sapir (1921) classifies languages according to three areas of potential variation (cf. Anderson 1990): (1) morphological concepts expressed; (2) formal processes used in expressing these concepts; and (3) internal complexity of words. The last of these areas in turn ranges over three separate linguistic types: analytic, synthetic, and polysynthetic.
A theory of language acquisition which asserts Continuity and Lexical Learning must also be able to account for the Developmental Problem, or why language acquisition is not instantaneous (Clahsen 1992). This motivates the third hypothesis, which is that the mapping between meaning and form (that is, the creation of a lexicon and morphological systems) constitutes the major acquisitional problem for the language-learner (Slobin 1985). I term this aspect of language acquisition Morpholexical Learning.² It is Morpholexical Learning which requires the existence of a powerful, innate learning strategy, and (I claim) is largely responsible for the sequence of empirically observable developmental stages which the child moves through en route to the adult grammar.

Each of these three hypotheses plays a crucial role in this conception of language acquisition. I use the Continuity Hypothesis in order to account for what is structurally invariant in human language, especially those aspects which could not plausibly be learned on the basis of easily observable properties of the input language (cf. Chomsky 1986). Lexical Learning accounts for those aspects of grammar which are open to variation, and which must therefore be learned. Morpholexical Learning is required to account for the developmental stages exhibited by child language during the course of language acquisition. These stages would otherwise be mysterious, since the principles underlying grammar are, by hypothesis, always present.³

A crucial part of this approach is that “learning” in a real sense takes place when the child acquires the morphology of his or her language: morphological acquisition is not adequately modeled by a parameter-setting procedure, nor by a simple listing of individual content-form pairs. Morpholexical Learning involves the acquisition of lexical items, followed by the extraction of patterns and rules of varying generality, which relate these items to each other and to the contexts in which they are used (Bybee 1988; Spencer 1988,

² The term ‘Morpholexical Learning’ emphasizes that the creation of the lexicon and morphological systems is in and of itself a significant learning problem for the child. This is in contrast to the Lexical Learning Hypothesis, which claims that acquiring pieces of morphology accounts for syntactic variation.
³ In the maturational approach to language acquisition advocated by (among others) Borer & Wexler (1987), Morpholexical Learning is not required, since the principles of grammar are claimed to develop over time, thereby accounting for the Developmental Problem. See below for arguments against this position.
1990). In this thesis, I investigate this hypothesis by looking at the acquisition of agreement morphology in German. My central claim is that language-learners initially construct word-specific paradigms (Pinker 1984), subsequently extracting general, implicational rule schemas, based upon information stored in these paradigms. Word-specific paradigms crucially provide mechanisms whereby inflectional regularities may be discovered. Implicational rule schemas provide the child with a means for inflecting novel lexical items. Evidence for a word-specific paradigm stage is taken from German child language, while evidence for the existence of general implicational rules is taken from data on paradigm mixture in German, Icelandic, and Latin.

Much of this thesis is dedicated to motivating the notion of paradigm, and its role in solving the Mapping Problem with respect to inflectional morphology. A paradigm is defined here as an n-dimensional, hierarchically organized matrix (Pinker 1984). Following Pinker (1984) and Wunderlich & Fabri (1994), I claim that the acquisition of inflectional morphology is guided by innate procedures for paradigm construction. Though these procedures are universal, the end product is not, since paradigms are open to variation in both the number and type of dimensions (morphological categories) and values (morphological features) which are instantiated. This variation is driven by overt morphological contrasts among the inflected forms of a lexeme, and is strongly analogous to the sort of syntactic variation hypothesized in the Lexical Learning approach.

The rest of the thesis is organized as follows. In Chapter 2, I begin with a discussion of the Mapping Problem and the role of paradigms in helping to solve it (Section 2.1). Since paradigms have not received much attention within the generative tradition, I will present a series of diachronic case studies in Section 2.2 which support their existence. The first of these (Section 2.2.1) deals with the process of analogical leveling, which serves to regularize morphophonemic alternations within a paradigm. The second (Section 2.2.2) deals with the phenomenon of paradigm mixture in Latin. The Latin 'mixed' 3rd declension consists of five sub-classes, resulting from the historical falling together of the
‘i-declension’ and the ‘consonantal-declension’ (cf. Janson 1971). Taking Wurzel (1987) as a starting point, I show that the inflectional Case morphemes are organized into hierarchically structured rule schemas, which are able to account precisely for the composition of these declensions. This analysis is developed further in Chapter 5. The final diachronic case study (Section 2.2.3) presents the development of German modal inflection from the class of Proto-Germanic \textit{preterit-presents} (Wurzel 1987). This development indicates that language-learners are sensitive to statistical patterns in the input, and that they actively seek out correlations between inflectional patterns and the extramorphological properties of lexical items.

Chapter 3 presents an analysis of two separate, but related debates in acquisitional theory. The first debate (Section 3.1) concerns the status of principles of Universal Grammar in child language. The Continuity Hypothesis, which makes the claim that these principles remain invariant, is in direct opposition to the \textit{Maturational Hypothesis} of Borer & Wexler (1987). The Maturational Hypothesis attempts to account for developmental stages in child language by adding principles to the linguistic competence of a child in a step-wise fashion. I conclude that Continuity represents a conceptually preferable alternative, while the “Developmental Problem” which this position creates can be accounted for with a learning component, which is required for the acquisition of the lexicon and morphological systems of knowledge.

The second of these debates (Section 3.2) concerns the nature of clausal representations in child language. The \textit{Lexical Projection Hypothesis} asserts that there are stages of early child grammar in which functional categories are absent (Radford 1990; Guilfoyle & Noonan 1992), while the \textit{Full Competence Hypothesis}\footnote{The term ‘Full Competence’ is somewhat misleading, since it does not refer to the linguistic competence (as standardly defined in generative linguistic theory) of the child per se, but refers instead to the type of clausal structures produced by the child. It would be more accurate to call this position the “Invariant Structure Hypothesis”.} makes the opposite assertion that clausal representations remain invariant throughout language acquisition (Poeppel & Wexler 1993; Guasti 1994). I present arguments that both of these positions
are incorrect, and that a combination of Continuity and Lexical Learning (cf. Clahsen, et al. 1994, among others) is preferable. Lexical Learning implies that the number and type of Functional Categories represented in syntactic structures is open to cross-linguistic variation, driven by overt differences in morphological expression. This position is supported by Rohrbacher’s (1994) findings that V-to-INFL movement is present in exactly those languages with an articulated agreement paradigm. Section 3.3 concludes the chapter with an overview of the Morpholexical Learning approach to language acquisition.

Chapter 4 contains an extended case study of the acquisition of subject-verb agreement morphology in German. This case study is meant to serve three goals. First, I present data and arguments for Continuity and against Maturation. Second, I show that the acquisition of morphological knowledge is able to account for the developmental patterns in the use of agreement inflections in German child language. Third, I argue that paradigms play a role in how inflectional morphemes are discovered and used. The sections of this chapter are as follows.

The relevant facts of German verbal morphosyntax are presented in Section 4.1, while details on the child subjects and methodology used in this study are presented in Section 4.2. Next (Section 4.3.1) I present an analysis of the child language phenomenon of Root Infinitives, which are matrix clauses containing a [-finite] main verb and no auxiliary. Arguing against the Maturational explanation offered in Wexler (1994), I present evidence from Ingram & Thompson (1995) that Root Infinitives result from performance constraints on sentence production, leading to the pragmatic omission of modals and auxiliaries (cf. P. Bloom 1990; Weissenborn 1992). In Sections 4.3.2 and 4.3.3, I present evidence that inflected verbs are initially stored as wholes in word-specific paradigm matrices, where they are then subject to morphological analysis and segmentation (Pinker 1984). Section 4.3.4 shows how discourse properties of parental input accounts for the

---

5 To be more precise, he shows that Germanic and Romance languages which have V-to-INFL raising are exactly those which distinctly mark the Person features [1st] and [2nd] in at least one number of one tense of the regular verbs (Rohrbacher 1994:117).
observation in Clahsen (1986), that use of 3rd Sing. inflection on the verb correlates with intransitivity. This analysis of German child language accounts for more of the acquisitional data than other current approaches, including that of Clahsen (Clahsen 1986, 1990; Clahsen & Penke 1992) and that of Wexler (Wexler 1994; Poeppel & Wexler 1993). Section 4.4 concludes this chapter with a review of the results.

Chapter 5 continues the analysis of paradigm mixture begun in Section 2.2.2. The goal of this chapter is to show what happens after the word-specific paradigm stage; in particular, I present an account of the Latin ‘mixed’ 3rd declension which relies upon the hierarchical organization of inflectional markers into implicational rule schemas, which I claim are abstracted from word-specific paradigm matrices during the course of acquisition (Section 5.1). These schemas, inspired by the Paradigm Structure Conditions of Wurzel (1987), bear strong resemblance to the Morpholexical Rules argued for in Spencer (1988, 1990). In order to account for the ways in which mixed paradigms are structured, I propose a restriction on inflectional class assignment, such that lexical items may contain at most one morphological specification per word. This analysis makes predictions on the direction of diachronic change in mixed inflectional classes, which the Latin data bear out. I proceed to demonstrate that similarly structured rule schemas are able to account for data from mixed paradigms in both German (Section 5.2) and Icelandic (Section 5.3), suggesting the existence of universal principles of morphological organization. Chapter 6 concludes the thesis with a discussion on how the acquisitional data of Chapter 4 relate to the data on paradigm mixture in Chapter 5.
2. Paradigms and the Mapping Problem

One traditional view of the lexicon is that it is the repository of the idiosyncratic and irregular portions of language (cf. Bloomfield 1933). If this were all there were to it, the lexicon would constitute an uninteresting domain of study. However, in the past few years, it has become clear that the lexicon constitutes a highly structured and intricate domain of knowledge. For example, the on-line production and comprehension of speech demand extremely quick and efficient lexical access mechanisms. Experiments have shown that so-called “fast shadowers” can reproduce auditory speech with as little as a 250-275 millisecond delay. This accomplishment is all the more impressive given the fact that these subjects maintain full comprehension of the input (Marslen-Wilson 1975).\(^6\) This minuscule time-frame suffices for these subjects to process and segment the auditory signal, extract its semantic content, and issue the proper articulatory commands to the vocal apparatus. The quickness and ease of lexical processing demonstrated by this act makes implausible the idea that the lexicon is an unstructured list of irregularities.

Moreover, the acquisition of the lexicon constitutes a remarkable achievement by anyone’s standards: by the time they reach adulthood, children will have acquired a vocabulary of perhaps 50,000 words, although some estimates go much higher (cf. Nagy & Anderson 1984). Performing the relevant calculations, it can be estimated that, at the onset of the “naming explosion” (sometime during the second year), children learn words at a rate of about one every waking hour. This involves several non-trivial tasks: (1) segmenting the speech stream into a set of linguistically relevant phonological units; (2) picking out the proper referential and propositional content from the non-linguistic context; and (3) mapping between these two.

In the sections which follow, I describe these tasks in greater detail. I also show how a learning model for inflection which incorporates paradigms (as in Pinker 1984) enables the language-learner to solve (1), (2), and (3) in an efficient way. Paradigms are

\(^6\) In comparison, the average (English) word has a duration of around 375 milliseconds.
defined here as hierarchically organized n-dimensional matrices, with each dimension corresponding to a morphological category such as Tense, Person, or Number, and each value of the dimension corresponding to a feature such as [Pret(erior)], [2nd], or [Plural]]. The discovery that there are two or more forms of a single lexeme leads to the creation of a paradigmatic dimension corresponding to a hypothesized morphological category. This in turn leads to the expectation that all of the feature values of this dimension, corresponding to individual cells, will be filled by other forms. Because storage of lexical items in paradigm cells provides the child with information on exactly how form and meaning co-vary, he or she can use this information in segmenting lexical items into separate morphemes.

2.1. Mapping between Meaning and Form

In creating a lexicon, children must first of all be able to segment the acoustic signal into the units relevant to the language they are learning. Children are not systematically presented with words in isolation: input consists by and large of whole phrases and sentences (cf. Peters 1983, 1985; Cutler 1994). Children must have some means of using prosodic (or other) cues in order to discover the phonological “words” of their language. Secondly, the child must be able to pick out from the context the proper propositional or referential content which is to be paired with this acoustic signal. The semantic content of the utterance must then be mapped onto the proper lexical units (cf. Clark 1993; Pinker 1984, 1989). In other words, children must, under conditions of extreme indeterminacy, be able to decompose a proposition and a string of sounds into a series of Saussurean signs, which we may define as “a bidirectional mapping between a phonological form and some representation of a concept” (Hurford 1989:187). That children do so extremely rapidly, and surprisingly well, constitutes one of the major mysteries of language acquisition.
Once we take into consideration the acquisition of the lexicon and the morphological systems of language, a strong, innate learning component becomes a logical necessity. This is so even if we continue to maintain that the acquisition of syntax remains a relatively automatic and simple affair for the child. That is, as Slobin (1985:1162) puts it: “The child’s fundamental task is to construct mappings between meaning and form.” For learning individual words, this task consists (partially) of the brute-force pairing of a phonological representation with a lexical semantic representation.

Yet matters are more complex than this, since these lexical representations must be integrated into the densely interconnected network which constitutes the lexicon. Additionally, language-specific morphological rules and patterns must be constructed on the basis of these stored lexical items (Bybee 1988). For instance, the semantic notion of causation is represented in various, language-specific ways.

(1) a. Mtsikana a -na -u-gw-ets -a mtsuku
   girl SP-PAST-OP-fall-CAUS-ASP waterpot
   ‘The girl made the waterpot fall.’ (Baker 1988:11)

   b. The farmer kills the duckling.

(2) a. Ron shafax[P1] et ha-mits.
   R. spilt Obj. Marker the juice
   ‘Ron spilt the juice.’

   b. ha-mits nishpax[P2]
   the juice spilt
   ‘The juice spilt.’ (Berman 1994:192)

In (1a), we see an example of a morphological causative construction from the Bantu language Chichewa (from Baker 1988). Here, causation is expressed via a bound morpheme which appears in composition with the main verb. As can be seen by the gloss, the English equivalent requires an analytical construction with a separate lexical item (make) to express the same concept. In (1b) the notion of causation is inextricably tied to the meaning of the verb kill itself, receiving no separate phonological representation. These
examples show that causation can be expressed in at least three ways: morphologically, lexically, and with an independent lexical item.

In (2), I give an example of a diathesis alternation from Hebrew (from Berman 1994). We see here that the verb obligatorily appears in a particular "pattern" (denoted by [P1] or [P2]) depending on whether it expresses causation (2a) or not (2b). This contrasts with the English glosses for these sentences, where the change in the argument structure of spill does not affect its phonological shape. This is a systematic difference between the morphological systems of these languages. Also important is the contrast between (1a) and (2a): Chichewa uses a bound affix, whereas Hebrew employs non-concatenative morphology. These examples show that how languages perform the mapping between meaning and form can produce systems which on the surface appear quite different from each other.

Various languages choose one or more of these options in order to express causation: none of these morphological regularities are pre-determined. They must each be learned by the child when acquiring his or her language. How the mapping between meaning and form is done has important consequences for the organization of the language as a whole. Bybee (1985) suggests that data such as those in (1) and (2) support a continuum of morphological fusion, in which the monomorphemic (lexical) expression of a conflation of semantic elements represents the greatest degree of fusion, while the separate, syntactically independent expression of a semantic element represents the least degree of fusion. The hierarchy she proposes is given in Figure 1:

Figure 1.

LEXICAL -- DERIVATIONAL -- INFLECTIONAL-- FREE GRAMMATICAL -- SYNTACTIC

<-- greater degree of fusion =>

lesser degree of fusion
One of the problems facing the language-learner is to discover where on this continuum the morphological categories of the target language are being expressed.

Successfully solving this problem entails that the child has already discovered which of a pre-determined set of morphological categories (cf. Talmy 1978; Slobin 1982) are relevant to the target language. By this, I mean those categories which have overt, distinctive realizations. This idea is made precise in Nicholl (1992:55). Nicholl there defines *Closed-Category Relevance* as a two-place predicate $R$, which takes a word as its first argument, and a feature dimension (e.g. Person, Number, or Tense) as its second:

\[ R\text{ (word, dimension) is true iff word can encode some but not all of the values in dimension.} \]

So, for example, the category Person is (Closed-Category) Relevant to a particular verb (that is, $R\text{ (verb, Person) is true}$) if and only if the verb has at least two separate phonological realizations (through the addition of affixes, through ablaut, etc.) when shifting along the Person axis of a paradigm, keeping all other dimensions and features constant. A hypothetical example of a Person/Number paradigm in which Person is Relevant and Number is not is given in (4), while a paradigm in which Number is Relevant and Person is not is given in (5).

(4) 

\[
\begin{array}{c|c|c}
\text{Person} & \text{[Sing.]} & \text{[Plural]} \\
\hline
\text{[1st]} & \text{umpa} & \text{umpa} \\
\text{[2nd]} & \text{umpi} & \text{umpi} \\
\text{[3rd]} & \text{ump} & \text{ump} \\
\end{array}
\]

(5) 

\[
\begin{array}{c|c|c}
\text{Person} & \text{[Sing.]} & \text{[Plural]} \\
\hline
\text{[1st]} & \text{lumpa} & \text{lumpi} \\
\text{[2nd]} & \text{lumpa} & \text{lumpi} \\
\text{[3rd]} & \text{lumpa} & \text{lumpi} \\
\end{array}
\]
This notion of Relevance is crucial to the learning account proposed here, since overt differences in morphological expression are what drive linguistic variation. This means that the primary task of the language-learner is to discover how meaning and form co-vary.

It should be clear by now that this task involves two parts: discovery of those categories which are relevant to the target language, as well as discovery of those formal processes which are used to express the category in question. These two parts of the mapping procedure are in principle distinct from each other. It is possible to construct a learning model in which multiple, distinct procedures operate on the linguistic input. The first of these would use phonological and prosodic cues to segment words and morphemes from each stored utterance. The second of these would construct a complete set of morphosyntactic categories which are compatible with these utterances. The third procedure would be to match up the segmented phonological units with morphosyntactic categories with some type of mapping algorithm.

An alternative approach would be to posit a single mechanism corresponding to all three of those above. The paradigm learning model of Pinker (1984) is exactly such an approach. In his model, paradigms are n-dimensional matrices, where each dimension is a morphological category, and each value of this dimension is a morphological feature. This is demonstrated by examples (4) and (5), which are composed of two dimensions each (Person X Number). The values are [1st] and [2nd] for Person, and [Sing.] and [Plural] for Number. As shown above, Person but not Number is Relevant to the paradigm in (4), and Number but not Person is Relevant to the paradigm in (5). We expect then, that the language-learner will “cancel out” the non-Relevant dimensions (or more likely, the language-learner will not posit such a dimension in the first place).

The storage of inflected word forms in paradigms enables the language-learner to discover which morphological categories are active in his or her language. Additionally, it allows the language-learner to discover the precise manner in which morphological
categories are formally marked. In (4) the child will extract suffixes marking 1st Person, 2nd Person, and 3rd Person ([a],[i], and [ø] respectively), and in (5) the child will extract suffixes marking Singular and Plural ([a] and [i] respectively). As Pinker (1984:172) notes, segmentation of affixes from stems “...is a non-trivial step that cannot be accomplished by an examination of individual inflected words.” Storage of word-forms in paradigms allows for segmentation in a computationally efficient manner. Furthermore, such segmentation is contingent upon actual meaning contrasts among inflected word forms. This means that “spurious” morpheme segmentations can be avoided.

In this section, I have provided conceptual motivation for utilizing paradigmatic structure as part of the solution to the mapping problem. Yet it remains to be seen whether or not we can provide any empirical support for the psychological and linguistic reality of paradigms. The is done in the next section, which consists of a series of diachronic case studies. These show that some cases of diachronic change can be most easily explained with reference to paradigmatic structure.

2.2. Paradigmatic Structure and Diachronic Change

Inflectional morphology represents a particularly complicated instance of the mapping problem, since deviations from a one-to-one pairing of meaning and form are unusually common in this domain (Carstairs 1987). Both inflectional homonymy (syncretism) and inflectional synonymy (the existence of inflectional classes) are present in large numbers of languages. One of the tasks of morphological theory is to explain why this is so. An important question to ask in this respect is whether or not these deviations are constrained in any principled manner. Various attempts have been made to answer this question by invoking the existence of paradigmatic structure, e.g. Carstairs (1987) and Williams (1981, 1994).

If paradigms actually do play a central role in accounting for various morphological phenomena, then we are forced to assign them status as a genuine part of our language
faculty. In that case, the acquisition of paradigmatic structures constitutes one of the tasks of the language learner (cf. Carstairs-McCarthy 1994; Bybee 1985, 1991; Pinker 1984; Williams 1994). Williams (1994:21) makes a strong claim in this regard:

“A paradigm is a multidimensional array of linguistic forms—for example, a verb conjugation, or a Latin noun declension. A paradigm is not just a convenient way to display linguistic information; rather, it is a basic form of linguistic knowledge. It interests here because it is highly language-particular, and at the same time, quite abstract in structure. Paradigmatic structure is also pervasive.”

Paradigms are an important part of *Morpholexical Learning*, which is the hypothesis that children begin by learning individual lexical items, and proceed to extract rules and generalizations from these stored items (Bybee 1988; Spencer 1988, 1990). Paradigms are important precisely because they provide the necessary structure with which the language-learner may discover the morphological patterns of his or her language. Linguistic theory must provide constraints on the sorts of generalizations made by children in the course of acquisition. In syntax, this is accomplished through the existence of abstract principles of Universal Grammar. In the lexicon, I propose that this is accomplished with innate procedures for paradigm construction. Without such constraints, children would be faced with an insurmountably vast number of possible generalizations when attempting to learn their target language, most of them linguistically irrelevant.

The use of paradigms to organize and represent the inflectional variants of lexical items has an ancient history (cf. Matthews 1991). In the past few decades, however, the notion has not received much support from generative grammarians. Paradigms have typically been either ignored altogether, or treated as epiphenomena, with no real status in linguistic theory (cf. Lieber 1981). On the other hand, a small group of linguists have resisted this general viewpoint, suggesting that paradigms do indeed have a role to play (Bybee 1985, 1988, 1991; Carstairs 1987, 1991, 1994; Halle 1973; Matthews 1991; Williams 1981, 1994; Wurzel 1987, 1991; see also the papers collected in Plank (ed.)
These researchers have brought to attention phenomena both in language acquisition and diachronic change which are most easily explained if we make reference to paradigmatic structures. In addition, linguists such as Andrew Carstairs and Wolfgang Wurzel point to synchronic characteristics of inflectional morphology which support the notion 'paradigm' (Carstairs 1987; Wurzel 1987).

In what follows, I give three diachronic case studies. Section 2.2.1 presents an instance of paradigm-leveling in the history of German and English. In Old English and Old High German, the verbal paradigms for the cognates of choose exhibited morphophonemic irregularity. By the time we reach Modern German and Modern English, the irregular forms of these paradigms have undergone independent but essentially identical changes, which have served to regularize them. Section 2.2.2 presents a case of paradigm mixture in the Latin 3rd declension, which resulted from the historical falling together of two previously independent declensions. I show how the forms of these declensions combined in a specific pattern, which may be represented by hierarchically organizing inflectional categories. In Section 2.2.3 we see how the inflectional class of German modals arose through the identification of modal verbs with the class of Proto-Germanic preterit-presents. This example shows how children actively seek out certain kinds of regularities in the input, creating morphological systems (patterns of paradigm assignment) which did not exist in previous generations.

2.2.1. Paradigm-pressure

One of the first proposals within generative grammar that paradigms play an indispensable role in the explanation of language was made by Halle (1973). Halle noted that diachronic evidence--namely the role of analogy in morphological change--provides abundant support for their existence. This was also the position of many 19th century linguists, who posited two mechanisms with which to account for diachronic phenomena. The first of these was a set of sound laws, which operate on the phonological systems of
languages, oblivious as to whether or not they impact on morphological systems. The second mechanism was analogical change, which (in part) serves the function of "repairing" morphological systems which have been "damaged" by the operation of the sound laws.

The analogical process of paradigm leveling takes place among the inflectional forms of a paradigm. The function of paradigm leveling is to decrease the amount of non-functional morphophonological irregularity in a language. In essence, it simplifies morphological representations. In (6) and (7) are examples of paradigm leveling in the history of English and German.\(^7\)

\[(6)\]

<table>
<thead>
<tr>
<th></th>
<th>Old English</th>
<th>Modern English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>ce:ozan</td>
<td>choose</td>
</tr>
<tr>
<td>Past Sing.</td>
<td>ce:as</td>
<td>chose</td>
</tr>
<tr>
<td>Past Plural</td>
<td>curon</td>
<td>chose</td>
</tr>
<tr>
<td>Past Part.</td>
<td>(ge-)coren</td>
<td>chosen</td>
</tr>
</tbody>
</table>

\[(7)\]

<table>
<thead>
<tr>
<th></th>
<th>Old High German</th>
<th>Modern German</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>kiusan</td>
<td>kuren</td>
</tr>
<tr>
<td>Past Sing.</td>
<td>ko:s</td>
<td>kor</td>
</tr>
<tr>
<td>Past Plural</td>
<td>kurun</td>
<td>koren</td>
</tr>
<tr>
<td>Past Part.</td>
<td>(gi-)koran</td>
<td>gekoren</td>
</tr>
</tbody>
</table>

From OE to ME, and from OHG to MG, the paradigms for *choose* were regularized in two ways. First, the stem consonant alternation [s], [z], [r] was eliminated. In ME, stems ending in [z] have replace those ending in [s] and [r]. In MG, stems ending in [r] have replaced those ending in [z] and [s]. Second, the stem vowel alternations in the Past Tense were eliminated. In both ME and MG, Past Tense stems formed with [o] replaced those formed with either [e:] or [u]. On the other hand, the vowel alternations across the Past/Present division were preserved; Present Tense stems are formed with [u] in ME and

\(^7\) The historical data here is obtained from Hock (1991:168-9).
[ü] in MG. These vowel alternations ([u]/[o] for ME, and [ü]/[o] for MG) now perform the sole function of marking Tense on these verbs.

It is a significant fact that separate stem forms for Past and Present forms have been retained, even though the pattern of stem vowel ablaut it preserves is no longer productive in either of the languages. These data suggest that paradigms are structured according to a hierarchy of morphological categories, with Tense ranking above Number and Person (Bybee 1985, 1991; Wunderlich & Fabri 1994). In (8), I show this ranking in matrix format, again using the German verb *kären* as an example.

(8)

a. Present
   Indicative

<table>
<thead>
<tr>
<th>Sing.</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>[3rd]</td>
<td>kür-t</td>
</tr>
<tr>
<td>[1st]</td>
<td>kür-e</td>
</tr>
<tr>
<td>[2nd]</td>
<td>kür-st</td>
</tr>
</tbody>
</table>

b. Preterit
   Indicative

<table>
<thead>
<tr>
<th>Sing.</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>[3rd]</td>
<td>kor</td>
</tr>
<tr>
<td>[1st]</td>
<td>kor-en</td>
</tr>
<tr>
<td>[2nd]</td>
<td>kor-et</td>
</tr>
</tbody>
</table>

The paradigm matrix in (8) shows Person crossed with Number, and both nested under Tense (cf. Pinker 1984). Bybee (1985) shows that phenomena such as verb stem allomorphy and suppletion are cross-linguistically more likely to occur along divisions of Tense rather than agreement. This is supported in (8), which shows that all of the forms nested under Present select one verb stem ([kür-], while all of the forms nested under Past select another ([kor-]).

This hierarchy (Tense > Number/Person) is supported further by the pattern of syncretism demonstrated in (8). Since different stems are selected depending on whether
the verb is marked [Present] or [Preterit], no form in (8a) is homonymous with any form in (8b). Within (8a) and (8b) themselves, 1st and 3rd Person forms are identical to each other in every Number except for Present Sing. Given the arrangement in (8), all of these syncretic forms are adjacent to each other. In fact, the only syncretism not accounted for by this arrangement is the one between the 3rd Person Sing. and 2nd Person Plural forms in the Present Tense (both with a [t] suffix). Such facts suggest that syncretisms occur “systematically” (see Chapter 5) only among those forms which are “adjacent” in a paradigm (cf. Williams 1981, 1994; Pinker 1984). This adjacency restriction would then provide motivation for why the Person/Number forms within the Present and Preterit sub-paradigms in (6) and (7) were eliminated, while the Tense distinctions were not.

Though Halle’s position regarding paradigms was generally ignored by the generative community, and he himself later abandoned a paradigm-based approach, data such as those depicted in (6) and (7) must still be accounted for. If paradigms do not exist as psychologically real aspects of our linguistic knowledge, then the driving force behind many cases of morphological change remains mysterious (McCreight & Chvany 1991). Notice that there is no logical necessity for the changes in (6) and (7) to have occurred in the manner in which they did: we could just as easily have expected the Present and Past Sing. forms to have behaved as a set, to the exclusion of Past Pl. and Past participles. Yet this is not a recurrent, robust pattern in diachronic change. On the other hand, if the inflected forms of a word are arranged into hierarchically structured paradigms, then the changes which occurred are seen to be part of a pattern, with syncretisms systematically occurring in adjacent cells.

8 Postma (1993) proposes a circular topology for Person/Number paradigms, in which forms sharing phonological similarities are adjacent to each other. The circular arrangements he proposes for Dutch will also account for the German examples shown here.

(i) 1st Sing. -- 1st Pl. -- 3rd Pl. -- 2nd Pl. -- 3rd Sing. -- 2nd Sing. -- 1st Sing.

(ii) 1st Sing. -- 3rd Sing. -- 1st Pl. -- 3rd Pl. -- 2nd Pl. -- 2nd Sing. -- 1st Sing.

The circular arrangement in (i) accounts for all of the syncretisms in (8a), while the circular arrangement in (ii) accounts for all of the syncretisms in (8b).
2.2.2. Patterns of Paradigm Mixture

Wurzel (1987) provides extensive arguments that the paradigmatic dimension of language is highly structured. Wurzel proposes two mechanisms for organizing inflectional systems. The first is a set of System-Defining Structural Properties (SDSPs), which determine the basic characteristics of its inflectional systems. So, for example, SDSPs determine the number and kind of morphological categories which appear in a language, and such properties as whether or not these categories are expressed cumulatively (i.e. whether it is agglutinative or fusional). Wurzel envisions these SDSPs as a number of dimensions which allow for parametric variation. “System-defining structural properties are those morphological properties that typologically characterize the inflectional system as a whole--those structural traits that determine its quality” (Wurzel 1987:79). This conception of inflectional systems is compatible with the model of paradigms as matrices proposed above. In fact, though coming from a very different framework, the model for inflectional learning proposed in Pinker (1984) can be seen as a concrete instantiation of Wurzel’s SDSPs.

In addition to SDSPs, Wurzel proposes a set of conditions on paradigms, which he terms Paradigm Structure Conditions (PSCs). PSCs govern the specific connections which exist among the sets of inflections in a language, serving to tie them together into a cohesive structure. They consist of a set of implicational statements of the sort: “if A in the Nominative, then B in the Accusative”. If a language possesses multiple inflectional classes, each inflectional class will be defined by a unique set of PSCs, possibly in combination with a unique set of lexical specifications. Wurzel illustrates this with the Latin ‘mixed’ 3rd declension.9 This declension resulted from the historical unification of the ‘i-declension’ and the ‘consonantal declension’.

---

9 The Latin 3rd declension is also discussed in detail in Carstairs-McCarthy (1991).
The relevant cases to look at here are the forms in the table which appear in small capitals. There are two competing realizations apiece for Accusative Singular ([im]/[em]), Ablative Singular ([iː]/[e]), Genitive Plural ([ium]/[um]), and Accusative Plural ([iːs]/[eːs]). A close examination of these forms reveals that they are not randomly distributed: two groups of implicational statements, or PSCs, captures their relationship to one another. These are given in (9a) and (9b), while the PSC which governs the inflections which all of the declensions have in common is given in (9c) (from Wurzel 1987:116).

    c. [is/Gen.Sing.]⇒...⇒ibus/Dat. and Abl.Pl.]  

Each noun which is specified with a particular inflectional ending from this set adheres to the implicational statements in (9a), (9b), and (9c). For example, Wurzel claims that words of the Type B declension are lexically marked with [iː/Abl.Sing.] and [em/Acc.Sing.]. According to PSC (9a), these words will then take [iːs/Acc.Pl.] and [ium/Gen.Pl.], but not...
All of the lexical specifications which Wurzel proposes for the paradigms in Table 1 are presented in (11).\(^{10}\)

\[(11)\]

<table>
<thead>
<tr>
<th>Class Type</th>
<th>Specifications in Lexical Entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>[im/Acc.Sing.]</td>
</tr>
<tr>
<td>B</td>
<td>[i:/Abl.Sing.] and [em/Acc.Sing.]</td>
</tr>
<tr>
<td>C</td>
<td>[i:s/Acc.PL] and [e/Abl.Sing.]</td>
</tr>
<tr>
<td>D</td>
<td>[ium/Gen.PL] and [e:s/Acc.PL]</td>
</tr>
<tr>
<td>E</td>
<td>[um/Gen.PL]</td>
</tr>
</tbody>
</table>

Out of \((2^4 =) 16\) possible ways for the paradigms to have blended, only five of these (two unmixed and three mixed) are actually attested. Wurzel implies that PSCs (9a) and (9b) are responsible for this. However, it is clearly the case that PSCs cannot alone account for the fact that only five classes arose. This is due to the following reasons.

First of all, there is nothing to prevent inflectional classes with morphological specifications other than those in (11). For example, we could easily propose a “Type F” paradigm with nouns that decline as in (12):

\[(12)\] Hypothetical “Type F” declension

<table>
<thead>
<tr>
<th>Case</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominative</td>
<td>-is</td>
<td>e:s</td>
</tr>
<tr>
<td>Accusative</td>
<td>-em</td>
<td>-i:s</td>
</tr>
<tr>
<td>Ablative</td>
<td>-i:</td>
<td>-ibus</td>
</tr>
<tr>
<td>Dative</td>
<td>-i:</td>
<td>-ibus</td>
</tr>
<tr>
<td>Genitive</td>
<td>-is</td>
<td>-um</td>
</tr>
</tbody>
</table>

\[(13)\]

<table>
<thead>
<tr>
<th>Class Type</th>
<th>Specifications in Lexical Entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>[im/Acc.Sg.] and [um/Gen.Pl.]</td>
</tr>
</tbody>
</table>

\(^{10}\) I have ignored in (11) the lexical specification [is/Gen.Sing.], which engages the “pan-3rd declensional” PSC (9c). I will continue to do so in what follows, not because it is irrelevant to my analysis, but because I do not have enough information on Latin to determine the status of all of the competing declensions.
The "Type F" declension requires exactly as many morphological specifications in a word's lexical entry (two) as Types B, C, and D in (11). The difference is that Type F has a specification for [im/Acc.Sg.] which "overrides" the assignment of [em/Acc.Sg.] given by PSC (9b). For Wurzel's analysis to work, he requires that each morphological specification be compatible with those PSCs which are engaged (perhaps by other morphological specifications). He also requires for there to be a limit on the number of specifications per lexical entry. If Latin nouns could possess up to three or four morphological specifications, then there is nothing to prevent all 16 possible "mixed" inflectional classes from arising.

Secondly, we must answer the question of how the construction of PSCs is to be constrained. A very interesting observation (though one which Wurzel fails to make) is that PSC (9a) is schematically the exact reverse of PSC (9b). This is shown in Figure 2, where arrows indicate implicational relationships:

Figure 2.

<table>
<thead>
<tr>
<th>Category</th>
<th>PSC (9a)</th>
<th>PSC (9b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accus. Sg</td>
<td>[im]</td>
<td>[em]</td>
</tr>
<tr>
<td>Abl. Sg</td>
<td>[i:]</td>
<td>[e]</td>
</tr>
<tr>
<td>Accus. Pl.</td>
<td>[i:s]</td>
<td>[e:s]</td>
</tr>
<tr>
<td>Gen. Pl.</td>
<td>[ium]</td>
<td>[um]</td>
</tr>
</tbody>
</table>

Figure 2 shows that a single arrangement of Case/Number categories is sufficient for both (9a) and (9b). The only difference between them (besides the differing phonological realizations) is that (9a) points "down", while (9b) points "up". Is it just accidental that (9a) and (9b) are converses of one another? Given the four Case/Number forms which are in competition, PSCs (9a) and (9b) where chosen from a set of 4! = 24 possibilities. An adequate account must be able to explain why these exact two were chosen.

Wurzel views PSCs as facilitating the acquisition of otherwise intricate and complex structures. "The general implicative paradigm-structure conditions, i.e., the implicative
connections between the individual forms of paradigms, make acquisition and use of inflectional systems by the speaker relatively simple, even if the systems are very complex” (Wurzel 1987:129). Though intriguing, this view is problematic. The PSCs in (9) cannot be “given”, since they are not cross-linguistically valid. Therefore, they must themselves be learned. But if this is so, then we require a procedure detailing how they are to be deduced. Imagine a child approaching the inflectional system of Latin, who has been exposed to all ten forms of the ‘consonantal-declension’ (Type E). At this point, in order to set up PSC (9b), the child is forced to choose from one of $10! = 3,628,800$ different possible PSCs; this is just another way of saying that each inflectional marker implies the other. When words from Types A-D are learned in addition to words of Type E, the range of compatible PSCs will be narrowed. But this implies that PSCs are constructed only after all of the relevant paradigms have been acquired.

This means that PSCs have the character of post hoc generalizations. Granted, once they exist, they make all subsequent assignment of words to paradigms (relatively) simple, but they do not aid in the initial construction of paradigms. Wurzel’s “System-Defining Structural Properties” are a more appropriate tool for this task. A possible clue to explaining the status of PSCs comes from looking at a different problem: how do we represent the relationships among inflectional classes? In Figure 3, I have arranged the inflectional endings from Table 1 into a hierarchical configuration, expanding on the structure represented in Figure 2.
Figure 3 has the advantage of directly encoding all of the implicational statements in (9a), (9b), and (9c). Assume that the arrows linking the inflections reflect an ordering relation (of some as yet unspecified type) in which those inflections on top “precede” those on the bottom. The four morphological categories which have competing inflectional realizations are represented by the two columns: these inflections are not ordered with respect to one another (e.g. Ablative Sing. [i:] and [e] are not directly connected). The two columns are joined by four dashed arrows, each of which represents a point of entry into the left-hand column from the right-hand column. As indicated in Figure 3, each of these paths corresponds to a particular paradigm (Types A-E) from Table 1. PSC (9a) corresponds to the left column of four inflections, while PSC (9b) corresponds to the same subset of four inflections in the right column (given the ordering relation assumed above). To see this, trace a path from one inflection to another, starting from the top or bottom. As long as the path followed picks out a subset of the inflections which are linearly ordered with respect to each other, it will correspond to one of the paradigms which are listed in Table 1.
An important fact to consider when looking at Figure 3 is that, of the competing paradigms, it is the ‘consonantal-declension’ (Type E in Table 1 and the right-hand column in Figure 3) which is productive, in that it attracted words from Types A-D. This fact allows us to view all other paradigms of the Latin 3rd declension as deviations from the norm, entailing the existence of morphologically-specified lexical entries. The specifications required, assuming the ordering arrangement in Figure 3, are given in (14):

(14)

<table>
<thead>
<tr>
<th>Class Type</th>
<th>Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>[im/Acc.Sing.]</td>
</tr>
<tr>
<td>B</td>
<td>[i:/Abl.Sing.]</td>
</tr>
<tr>
<td>C</td>
<td>[i:s/Acc.Pl.]</td>
</tr>
<tr>
<td>D</td>
<td>[ium/Gen.Pl.]</td>
</tr>
<tr>
<td>E</td>
<td>none</td>
</tr>
</tbody>
</table>

In (14), each “irregular” declension contains a single morphological specification in its lexical entry. Words of Type E, as members of the “productive” paradigm, require no morphological specification. This state of affairs contrasts with that shown in (11), where each type (including the productive Type E class) has at least one specification, with Types B, C, and D having two apiece.

There is an additional, important difference between (11) and (14). In (11), the specifications for Types B, C, D, and E include productive, as well as irregular markers (e.g. Type E is specified for [um/Gen.Pl.], which is the productive affix for this category in the 3rd declension). In contrast, the specifications in (14) are all for “irregular” markers (those contributed by the ‘i-declension’, or Type A in Table 1). This captures an important insight into inflectional class assignment: lexical entries should only contain idiosyncratic morphological specifications, or those which cannot be predicted on the basis of the lexical item’s extra-morphological characteristics. In Chapter 5, I propose elevating these differences between (11) and (14) into a principle restricting inflectional class assignment:
lexical entries may contain at most one morphological specification; furthermore, this specification must be for an *irregular*, not a *productive* inflectional marker.

This addresses one of the concerns with Wurzel's model raised above: what constrains the number and type of specifications which are possible? However, we have not yet addressed the issue of what constrains the number and type of PSCs which are possible. In particular, we must determine why the configuration in Figure 3 exists. The fact that all singular inflections precede all plural inflections is suggestive, indicating a hierarchical organization where Number > Case (cf. Plank 1991a,b), though this leaves unexplained the chaotic ordering of cases. I pursue this topic further in Chapter 5, where I propose that Figure 3 should be reorganized according to a markedness (or semantic) hierarchy of categories. Independent evidence for this hierarchy comes from patterns of syncretism in Latin and Icelandic declensions, assuming that syncretic forms are adjacent to one another in paradigms (see above). Data from mixed paradigms in both German and Icelandic further support the analysis of the Latin 3rd declension begun here.

What exactly is the status of these newly conceived “PSCs”? Above I argued that PSCs are of no help in the initial construction of paradigms, a job which is more suited to the SDSPs proposed in Wurzel (1987). These SDSPs correspond, in Pinker's (1984) model for inflectional acquisition, to the construction of n-dimensional paradigm matrices, such as the one shown in (8). This suggests the following sequence of stages. First the child constructs *word-specific paradigm matrices*, according to innate procedures (Pinker 1984; Wunderlich & Fabri 1994). These are matrices which contain entire inflected words as unanalyzed wholes (cf. MacWhinney 1978), each one indexed for categorical information according to the cell it occupies. The child is driven to set up dimensions by the requirement that each cell contain a unique entry (the *Unique Entry Principle* of Pinker 1984). These morphological dimensions must satisfy the further requirement that they be Relevant to the word contained in the paradigm (*Closed-Category Relevance* of Nicholl 1992).
The second stage begins when the child creates generalizations from the word-specific paradigms he or she has constructed. These generalizations correspond to implicational, hierarchically-based structures such as that shown in Figure 3. The purpose of these structures is both to enable the child to fill-in incomplete word-specific paradigms, and to enable the child to predict the inflectional behavior of a novel lexeme based on exposure to a limited number of its inflected variants. This sequence of stages, from specific to general, is important to the conception of *Morpholexical Learning* I argue for below. Morpholexical Learning involves the acquisition of independent lexical items, followed by the extraction of lexical redundancy rules relating them (Spencer 1990). I suggest that word-specific paradigms play a crucial role in this procedure, by providing the necessary structure for the language-learner to create generalizations which are an accurate characterization of his or her target language.

2.2.3. **German Modal Inflection**

The mixed paradigms of the Latin 3rd declension represent a special case, in that morphological criteria alone determine their membership. It is quite common for extra-morphological criteria (such as gender, or phonological characteristics) to play this role. In such cases, paradigms are often completely productive over their defined domains. By dividing up a word class into inflectional classes based upon semantic or phonological characteristics, multiple paradigms (i.e. inflectional classes) perform the useful task of providing extra (typically redundant) information to the listener, presumably helping to speed up lexical access.

The historical development of the class of modal verbs in German provides a good example of how inflectional classes based upon extra-morphological criteria can come into existence in the history of a language. This phenomenon establishes that language learners strive to impose regularity on the morphological systems of their target language, rather than passively absorbing input. German modals, as well as the verb *wissen* ('to know') are
remnants of the Proto-Germanic *preterite-presents*, a class of verbs whose present-tense forms are historically derived from the preterite. Verbs of this class demonstrate a different pattern of inflection than that which is usual in German, with 1st and 3rd Sing. present-tense forms being homophonous, otherwise typical of preterite inflections of "regular" verbs. With one or two exceptions, these forms are zero-marked and show ablaut. Examples are given in (15).

(15)

<table>
<thead>
<tr>
<th>Person/Number</th>
<th>a. können ('can')</th>
<th>b. dürfen ('may')</th>
<th>c. wissen ('know')</th>
</tr>
</thead>
<tbody>
<tr>
<td>3rd Sing.</td>
<td>kann-ø</td>
<td>darf-ø</td>
<td>weiß-ø</td>
</tr>
<tr>
<td>1st Sing.</td>
<td>kann-ø</td>
<td>darf-ø</td>
<td>weiß-ø</td>
</tr>
<tr>
<td>2nd Sing.</td>
<td>kann-st</td>
<td>darf-st</td>
<td>weiß-t</td>
</tr>
</tbody>
</table>

The class of preterite-presents initially included non-modals, such as wissen, and not all modals were included in this class. By MHG various orthogonal changes reduced the number of non-modal preterite-presents to three (Wurzel 19117:145). The statistical predominance of a class of verbs with a highly salient semantic characteristic ([+modal]) was apparently enough to cause children learning the language to posit a PSC which codifies this relationship: "If [+modal], then preterite-present". In MG, all modals inflect as those in (11a,b), with the modal verb brauchen ('to need') currently shifting into this class.

Learning of this sort has an abstract character, since it is the pattern itself (syncretism of 1st and 3rd Sing. in the present-tense) which identifies modal verbs. However, Parameter-setting is clearly an inadequate metaphor for this type of morphological acquisition. A hypothetical parameter is given in (16).

(16) *Modal Inflection Class Parameter*

[+modal] verbs (do/do not) have identical forms in the 1st and 3rd Person.
Probably no linguist would support the existence of (16), or any parameter like it. A more likely explanation is that generations of children learning German have worked upon the input they have received, noting a strong correlation between modality and a particular inflectional pattern, subsequently creating a conditional statement (PSC) which was previously non-existent. This phenomenon is an instance of what Lass (1990) calls “exaptation” in language: the co-opting of non-functional inflectional “junk” for a specific purpose. In this case, what was originally a useless complication of the inflectional system has become a functional means of identifying and distinguishing an important, distinct verbal sub-class. Along the way, members of the class of preterite-presents have dropped the morphological specifications which marked them as irregular, with the new PSC determining membership in an almost perfectly regular (and productive) verbal paradigm. If there is a place for “learning” anywhere in language acquisition (as I am claiming there is), then evidence such as this suggests that we should look for it in the domain of (inflectional) morphology.

These case studies also strongly suggest that paradigms are an important part of our linguistic competence. If this is so, then we expect to see ramifications of their existence not only in diachronic change, but also in language acquisition, as well as in the synchronic structure of language. In Chapter 4, I present an extended case study of the acquisition of verbal agreement morphology in German, showing how the construction of word-specific paradigms plays a role in how inflectional morphemes are used and discovered. In Chapter 5, I show how the synchronic structure of mixed paradigms in Latin, German, and Icelandic reflects the existence of implicational rule schemas which are abstracted from paradigms in the course of acquisition. But first (Chapter 3), I present an analysis of two unanswered questions in acquisitional theory, which provide a backdrop for this investigation into the relationship between paradigms and language acquisition. The first concerns the status of principles of UG in child language, while the second concerns the
nature of the grammatical structures children produce. In particular, the questions are whether or not UG principles and/or clausal structures undergo development during the course of language acquisition. Both questions are relevant to the status of Morpholexical Learning (and the status of paradigms in language acquisition), because properly answering them determines what in language is learned, and what in language is invariant. I conclude Chapter 3 by arguing that it is the acquisition of inflectional morphology (guided by the construction of paradigms) which determines not only linguistic variation, but also the sequence of developmental stages which children exhibit \textit{en route} to their target language.
3. Continuity and Lexical Learning

One of the primary goals of linguists today is the creation of an abstract theory of *Universal Grammar* (UG) which can account both for what is universal, or common to all languages, and for the obvious differences which they display. This research program has been most clearly articulated in the *Principles & Parameters* (P&P) approach to linguistic theory (Chomsky 1981, 1986). The system of UG, as conceptualized within the P&P approach, is a collection independent modules, each with their own set of principles, which in their combination and interaction determine the form of grammar. These principles are presumed to be part of the biological endowment of the human species, and they account for the universal aspects of natural language.

On the other hand, linguistic diversity is attributed to the underspecification of these principles with respect to their application in particular languages; UG is claimed to provide a finite range of options from which to choose in order to fully specify them. These points of underspecification are termed *parameters*, and it is the job of the *Language Acquisition Device* (LAD) to set them based upon information forthcoming from the input language, or *Primary Linguistic Data* (PLD). As originally conceived, the setting of a single parameter, based upon easily observable properties of the PLD, should determine a range of syntactic effects; when combined, the setting of the various parameters has a richly deductive set of consequences for the overall form of the grammar.

Though the P&P approach offers hope as a way of explaining the logical problem of language acquisition, we are still left with the job of formulating a theory of what constitutes a possible parameter. Without a substantive theory of parameters, an explanatory account of language acquisition and language diversity remains an intractable problem. This theory must satisfy a wide variety of empirical data from language development, as well as the constraints of *Learnability Theory* (cf. Wexler & Culicover 1980). That is, children must be able to set these parameters based upon the PLD which is in principle available to them. The best known of these learnability theoretic constraints is
the "no negative evidence" assumption (Brown & Hanlon 1970): children do not have access to negative evidence, in the form of parental corrections or other such means of linguistic instruction.

In this thesis, I argue for the theory of parameters offered in Borer (1984). Borer made the very interesting claim that most parametric variation can be reduced to the acquisition of morphological properties. The learning of language-specific morphological processes is a requirement in any theory of language acquisition (see Section 2.1); additionally, morphology is a domain in which overt, positive evidence is easily available. If we can reduce linguistic diversity to this process, then we would satisfy Occam's Razor by accounting for child language development solely with those learning processes required for the parts of language which we already know for certain must be learned.

"It is a desirable step forward to try and restrict the class of possible parameters. The strongest claim in this respect would be that there are no language-particular choices with respect to the realization of universal processes and principles. Rather, interlanguage variation would be restricted to the idiosyncratic properties of lexical items. These idiosyncrasies, which are clearly learned, would then interact with general principles of UG in a particular way. This interaction would result in vastly different systems." (Borer 1984:2-3)

An even stronger claim in this respect, and one which was taken up in Chomsky (1989), is that parametric variation can be limited to those lexical items which have to do with functional categories (FCs). Given the close ties between inflectional morphology and FCs in much of current GB theory (cf. Pollock 1989), this would place a large part of the burden of acquiring language on the learning of inflectional morphology. In Chapter 4, I investigate the mechanisms behind the development of agreement morphology in German child language.

In order to account for the acquisition of language, a substantive theory of parameters is only part of what must be defined. We must also characterize the nature of the initial state of the language faculty. This is the starting point for the LAD, which determines what parts of language are "given", and what parts must be "learned". Two aspects of the
LAD may be taken into consideration here. First, we must determine the status of UG principles in child grammar: do they remain constant, or do they “mature” over time? Second, we must determine the status of clausal representations which children produce: are they identical to adult clausal structures, or must they be learned? By answering these two questions, in combination with a substantive theory of parameters, we must be able to account for the sequence of developmental stages which children demonstrate en route to the adult language competence in a given language.

Both the initial status of the LAD, and the sort of data which are relevant to it, have been the subject of much debate in acquisitional theory. Two recent proposals have accounted for developmental phenomena by claiming that linguistic competence develops over time. The *UG Constrained Maturational Hypothesis* (Borer & Wexler 1987) asserts that UG principles are added step-wise to the grammar according to an innate maturational schedule. The *Lexical Projection Hypothesis* (Radford 1990; Guilfoyle & Noonan 1992) claims that the initial clausal structures of children are the pure expression of lexical-thematic properties, lacking Functional Categories. Though these two positions are logically independent of one another, it has been common practice to adopt one to the exclusion of the other. This is not surprising, since the adoption of both hypotheses leads to an acquisitional theory with a surplus of mechanisms for explaining development, leading to a reduction in explanatory power.

3.1. The Status of UG Principles in Child Grammar

UG Constrained Maturation represents a very strong innateness claim. If true, then not only are the principles and parameters of UG innately determined, but so are the sequence of stages which are exhibited in the acquisition of language. The LAD starts off with a certain subset of the principles which govern the adult grammar. The PLD is fitted to this competence according to some simple procedure, typically via an error-driven algorithm (cf. Gibson & Wexler 1994). As the biological organism matures, additional UG principles
come on-line. Subsequent PLD is fitted to this new and improved grammatical competence, resulting in a reorganization of the grammar (cf. Bertolo 1994 for a formalization of this procedure). Thus, by combining maturation with a simple “triggering” algorithm, Borer and Wexler claim that there is no need to posit a powerful learning component with which to acquire language.

UGCM is presented specifically as an alternative to what has been called by Pinker (1984) the Continuity Hypothesis. This is the assertion that our initial attempts to explain the acquisition of language should assume no change over time in the LAD itself, on grounds that the most explanatory theoretical account will posit a single, unchanging learning mechanism. Under this view, maturational accounts represent a serious weakening of acquisitional theory, whether it be UG constrained (Borer & Wexler 1987) or not (Felix 1992).\textsuperscript{11} If we make allowance for the LAD itself to undergo change, the range of explanations compatible with a given set of acquisitional data is considerably enlarged, since we have added an extra dimension of potential variability. Since Maturation violates Occam’s Razor in this fashion, I reject it in favor of the Continuity Hypothesis.

By itself, however, continuity fails to solve the Developmental Problem. We still need to account for the empirical observation that children go through different stages en route to the adult grammar. In particular, Borer and Wexler argue that the continuity claim forces us to a position in which either (1) input to the child is strongly ordered, in such a way as to produce a temporally ordered sequence of stages in the progression of the child’s grammar, or (2) we admit that the acquisition of language is not adequately modeled (solely) by a simple parameter-setting procedure. This is the inevitable result of assuming that the triggering of parameters is a simple, instantaneous procedure, and that grammatical competence remains constant over time.

\textsuperscript{11} Maturational theories which are not constrained by UG allow for grammatical principles which are not those of the adult grammar. Without a formalization of of what sort of principles are possible, and how they mature, a maturational theory which is not UG constrained does not make any interesting empirical predictions.
There is little or no reason to suppose that children are presented with strongly ordered input during the acquisition of language (cf. Wexler & Culicover 1980). Therefore, let us assume that alternative (1) is false. This means that we are forced into accepting alternative (2): language acquisition is not adequately modeled in its entirety by a parameter-setting model. That is, given continuity, we must accept the fact that certain aspects of language have to be learned.

Borer and Wexler view *learning* as something which linguistic theory should do without: standard "poverty of stimulus" arguments (Chomsky 1981, 1986) persuasively demonstrate that abstract principles of syntax must be innate, and that they cannot be the products of a learning mechanism. On the other hand, I argue that a learning mechanism is indeed required, both on conceptual and empirical grounds, for the acquisition of the lexicon and of morphological systems (cf. Spencer 1990; Carstairs-McCarthy 1992), and that it is the interaction of this learning mechanism with the input language which is largely responsible for the sequence of developmental stages in child language. This view is similar to that expressed by Clahsen and his associates (see, e.g., Clahsen et al. 1990), in that Lexical Learning plays a crucial role in language acquisition. Yet, unlike Clahsen, I stress the fact that morphological systems of knowledge require a learning process in their own right. Thus, acquiring the morphology of language cannot be reduced to the acquisition of idiosyncratic bits of knowledge. Nor is this knowledge adequately modeled by parameter setting (Williams 1987, 1994).

The Maturational Hypothesis is the product of a theoretical viewpoint which abstracts away from child language development. In a sense, parametric theory makes language acquisition "too easy". For this reason, an added assumption (linguistic maturation) is required in order to slow the language-learner down. A preferable alternative, I argue, is to construct a theory of acquisition which does not require this extra assumption. Weinberg (1994:361) succinctly states the main drawbacks of maturational accounts of language acquisition:
“Maturation theory has three strikes against it. First, it adds to the descriptive power of the theory for particular constructions. Second, it is ad hoc and we miss explanations for why certain constructions are more difficult to acquire than others. Instead we settle for a listing of arrival time. Third, it seems to be unnecessary because we can provide a non-maturational account using independently motivated principles for every case cited as motivating this approach.” [author’s emphasis]

The second of these “strikes” against maturation is, in my view, particularly telling. Explanations for development should involve the fact that children must learn the language-particular aspects of their morphological systems. They must also deal with performance constraints on non-linguistic domains such as short-term memory (P. Bloom 1990). One may claim that this sort of “non-linguistic development” (of performance components) is merely replacing one sort of maturation with another (Felix 1992). Yet crucially, the development of cognitive areas such as short-term memory capacity is subject to experimental testing, independently of the particular linguistic framework we are working within. On the other hand, theories claiming the maturation of UG principles are entirely theory-internal, since our account of what matures is a direct consequence of our particular characterization of UG.

3.2. The Status of Clausal Structures in Child Grammar

Continuity is vacuous unless we explicitly state what is universal and unchanging in child language, and what is language-specific and must be learned. Following Borer (1984), I claim that those aspects of syntactic representations which must be learned are exactly those which are concerned with (inflectional) morphology. In current frameworks (cf. Pollock 1989; Chomksy 1989), this means that the Functional Categories of syntax are open to cross-linguistic variation. This is clearly not incompatible with the Continuity Hypothesis, since nothing requires us to formulate Universal Grammar in such a way as to force the presence of every Functional Category. Such a formulation would in fact be very
much theory-internal, since there is little overt evidence that every FC appears in every language and every clausal structure.

3.2.1. Lexical Projection Hypothesis

The Lexical Learning Hypothesis is essentially incompatible with what has been termed the Lexical Projection Hypothesis (LPH). The LPH asserts that children initially produce clauses which are pure instantiations of lexical-thematic properties, systematically lacking Functional Categories of any kind (cf. Lebeaux 1988; Radford 1990, 1992, 1994; Platzack 1990; Guilfoyle & Noonan 1989, 1992). According to this hypothesis, Functional Categories arise at later stages of development, either all at once (Radford 1990), or in a step-wise sequence (Guilfoyle & Noonan 1992): a typical proposal in this regard is that IP is added first, and CP somewhat later. The unifying claim of all variants of the LPH is that there is a universal stage in which children are in principle unable to construct clausal representations containing Functional Categories (and hence Functional Projections).

Evidence for the Lexical Projection Hypothesis has come from languages such as English and Swedish, in which the earliest productions of children often lack overtly marked inflection. For example, Radford (1992) gives the examples in (17), from children acquiring British English (ages range from 20-25 months):

(17)  
a. Read book (imitation of: ‘I will read the book’)
b. Hayley draw boat (answer to: ‘What did you draw?’)
c. Pig say oink (answer to: ‘What does the pig say?’)

Assuming that modals are base-generated in the T(ense) P(hrase), the absence of the modal in (17a) can be accounted for if the child lacks a TP at this stage in his development. If this were true, it would also account for the absence of past-tense marking in (17b). The lack of agreement marking in (17c) would similarly be accounted for by the absence of an AgrP.

On the other hand, evidence against the LPH has come from languages such as German and Italian, where researchers have failed to discover a well-defined stage of child
language when inflectional morphology is missing. It is probably not accidental that these languages have "stronger" inflection than English and Swedish, and thus present the child with a much better opportunity to learn their overt realizations. For example, the following sentences are taken from transcripts of Katrin, a child learning German (age 17 months):

(18)  
\[ a. \text{Katrin mach-t kaputt} \]
\[ K. \text{ make-3s broken} \]
\[ 'I (Katrin) broke it.' \]

\[ b. \text{Papa geh-t hin} \]
\[ father go-3s there \]
\[ 'Father is going there.' \]

(19)  
\[ a. \text{ein Pferd kann-Ø nicht lauf-en} \]
\[ a-neut/Sing. horse can-Ø not run-inf. \]
\[ 'A horse can't run.' \]

\[ b. \text{diesen will-st du hab-en} \]
\[ this-acc/Sing. want-2s you have-inf. \]
\[ 'You want to have this one.' \]

The data in (18) show that Katrin has correctly produced overt agreement markers on the verb, in both cases 3rd Sing. [t]. In (19a), there is no agreement marker on the modal. However, this is grammatical for German, because modals exhibit deficient agreement morphology in the 3rd Person (see Section 2.2.3 above). The infinitival verb *laufen* 'to run' in (19a) is correctly produced with the [en] infinitival suffix. The same is true of the infinitive *haben* 'to have' in (19b). In (19b) we also see a modal which is properly inflected for 2nd Person, with the [st] suffix.

Radford (1994) admits that the "No Functional Category" stage may be "vanishingly brief" in such cases. This concession to the possibility of cross-linguistic variation in the duration of this stage of development, to the point where it may not even be detectable, robs this hypothesis of any real empirical bite. The early appearance of overt inflection is a function of the type of language being learned, and not of some invariant property of the initial state of the grammar. Radford's treatment of the acquisition of morphology represents a very syntactically-based approach to language acquisition; it essentially reverses...
the perspective stated in Borer (1984), in which it is the acquisition of morphology which drives the acquisition of (language particular) syntax.\textsuperscript{12}

3.2.2. Full Competence Hypothesis

Another syntactically-based approach to language acquisition is espoused in the \textit{Full Competence Hypothesis} (FCH) of Poeppel and Wexler (Poeppel & Wexler 1993; Wexler 1994). This is the hypothesis that clausal structures remain invariant throughout childhood, with functional categories present from the very beginning of multi-word speech (cf. Poeppel & Wexler 1993; Wexler 1994; Weissenborn 1990; Guasti 1994). The central claim of the FCH is that child syntactic representations are identical to adult syntactic representations, with the full complement of Functional Categories demanded by UG. This entails, as far as I can tell, that every language instantiates every FC in its grammar.\textsuperscript{13}

There are at least two reasons for rejecting the Full Competence Hypothesis, one of them conceptual, and one empirical. The conceptual reason is this: it is an open question whether or not we can (or should) maintain that every Functional Category exists in every language. Evidence for the FCH has come from Indo-European languages such as German (Poeppel & Wexler 1993) and Italian (Guasti 1994). As a consequence, the FCH has amounted to the claim that morphosyntactic categories such as Person/Number agreement exist from a very early stage, along with the syntactic consequences (i.e. movement) which attends the presence of an AgrP or IP in clausal structures. Once we look beyond these languages and morphosyntactic categories, it becomes less clear that all Functional

\textsuperscript{12} In fact, viewed in a certain light, the Lexical Projection Hypothesis makes a very odd claim. Let us say that the initial absence of inflectional markers, such as 3rd Sing. [s] and Progressive [ing] in English, is due to the systematic lack of Functional Categories in the grammar of the child. However, at least some of the input which the child receives does contain verbs marked with these inflections. This entails that children have \textit{segmented} [s] and [ing] and \textit{identified} them as separate, inflectional morphemes \textit{before} they discard them as incompatible with their grammatical competence. Thus, the child must have knowledge of what constitutes a Functional Category before he or she is supposed to have the ability to represent them in his or her grammar.

\textsuperscript{13} If this claim is not being made, then the FCH reduces to some version of the Lexical Learning approach, which asserts that children must acquire the precise number and type of Functional Categories which are instantiated in their target grammar.
Categories are immutably present in every language. For example, Fukui (1986) presents arguments that Japanese differs from English precisely in the type of FCs which it instantiates.

The empirical reason for rejecting the FCH is that the presence or absence of agreement morphology in Germanic and Romance languages does appear to drive the syntactic movement of verbs to INFL. Evidence comes from Rohrbacher’s (1994) dissertation. Rohrbacher examined the position of negation and adverbs relative to the position of verbs in those languages which possess “rich” agreement morphology, such as Icelandic and Yiddish, and in those languages which possess “poor” agreement morphology, such as English and Swedish (cf. Pollock 1989). In order to factor out movement to the Complementizer position in these “V2” languages (see Section 4.1), Rohrbacher was forced to look at subordinate clause V movement to INFL in those languages with VO word order. His results are shown in Table 2.

Table 2. V to INFL Movement in the Germanic Languages
(Rohrbacher 1994: Table 2.3)

<table>
<thead>
<tr>
<th>Language</th>
<th>D-Structure</th>
<th>S-Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I-Neg-V-O</td>
<td>Neg-O-V-I</td>
</tr>
<tr>
<td>Yiddish</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Icelandic</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Faeroese</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Mld. Scand.</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>German/Dutch</td>
<td>√</td>
<td></td>
</tr>
</tbody>
</table>

In conjunction with this survey of V-to-INFL movement, Rohrbacher examined the morphological verbal agreement paradigms of these languages. The Present Indicative paradigms for sample languages are given in (20), from Rohrbacher (1994:97):
There is a strong correlation between those languages with V-to-INFL movement, and the existence of “rich” agreement morphology. To give this precise content, Rohrbacher proposes a Paradigm-Verb Raising Correlate, given in (21):

(21) *The Paradigm-Verb Raising Correlate II*

A language has V to I raising if and only if in at least one number of one tense of the regular verbs, the person features [1s] and [2nd] are both distinctively marked.

Furthermore, “distinctive marking” depends crucially upon a comparison of related word-forms:

“A privative feature such as [1st] or [2nd] is distinctively marked iff the forms bearing this feature are distinct from the forms lacking this feature. Accordingly, a language has V to I raising if it has at least one number with what I will call a ‘complete’ or ‘full’ paradigm in which the forms for first and second person differ from each other as well as from the forms for “third” person and the “infinitive.” (Rohrbacher 1994:108-9)

The Paradigm-Verb Raising Correlate provides strong evidence that morphological distinctions play a role in syntactic variation.\(^{14}\) Unfortunately, the German acquisitional

---

\(^{14}\) Rohrbacher explains the Paradigm-Verb Raising Correlate in terms of “referential” and “non-referential” paradigms. Those which satisfy the Raising Correlate are referential, since they succeed in uniquely identifying discourse referents. Those paradigms which do not satisfy it are non-referential, since they fail to uniquely satisfy discourse referents. Referential paradigms are inserted into INFL (or Agr) at D-structure, triggering syntactic movement, while non-referential paradigms are the result of “post-syntactic” spell-out. An alternative explanation, compatible with the same generalization in (21), is that languages with “full” paradigms project directly into Agr nodes in the syntax, while those without “full” paradigms project into a lower node (e.g. inside the VP, or in TP) (cf. Wunderlich & Fabri 1994).
data examined in Chapter 4 do not bear on Rohrbacher’s findings, both due to the fact that subordinate clauses are not produced during the stages investigated, and due to the fact that German is an OV language (see Table 2). This means that movement to INFL in German is not detectable using word order facts. Despite this lack of acquisitional data, I take Rohrbacher’s findings to constitute strong evidence for the morphologically-driven viewpoint of syntactic acquisition, and strong evidence against the Full Competence Hypothesis of Poeppel & Wexler (1993).

Not only does Rohrbacher’s analysis support the Lexical Learning Hypothesis, but it also provides further support for the claim that the notion ‘paradigm’ is linguistically relevant. This is an immediate consequence of his definition of “distinctive marking”. There is an obvious similarity of this notion to the definition of Closed-Category Relevance in (3). This suggests an analogous three-place predicate encoding the “relevance” of a feature-value to a dimension (which we may call “Feature-Value Relevance”, or $R_V$)

\begin{equation}
R_V(\text{word}_i, \text{value}_k, \text{dimension}_j) \text{ is true iff } \text{word}_i \text{ encodes value}_k \text{ in } \text{dimension}_j,
\text{ and for every other value}_n \text{ in } \text{dimension}_j \text{ such that } n \neq i, \text{ word}_i \text{ does not encode value}_n
\end{equation}

Using (22), we may re-interpret the Paradigm-Verb Raising Correlate in (21) as follows:

\begin{equation}
The \text{Paradigm-Verb Raising Correlate (reinterpreted)}
\end{equation}

A language has V to I raising if and only if in at least one number of one tense of the regular verbs, $R_V(\text{verb}, \text{Person}, [1st])$ is true and $R_V(\text{verb}, \text{Person}, [2nd])$ is true.

The point of this reinterpretation is that essentially the same notion—Relevance—is crucial to both paradigmatic and syntactic variation. Closed-Category Relevance determines whether or not a particular dimension is added to a word-specific paradigm. Feature-Value Relevance determines whether or not a particular feature-value is added to a dimension of a word-specific paradigm. When paradigms which are created by this process satisfy (23),
then syntactic movement to INFL is a result (see footnote 14). This whole procedure is driven by the acquisition of the lexicon and morphology of the target language.

3.4. Morpholexical Learning and Paradigms

I conclude this chapter with a review of what has been discussed so far, specifically relating the subject of Chapter 2 (the linguistic reality of paradigms) with the issues raised in the previous sections. In particular, I am advocating a theory of acquisition which both includes the notion ‘paradigm’, and possesses the following three attributes:

a. The Continuity Hypothesis. Syntactic principles and processes are innately determined and present from the very beginning of language acquisition. There is no maturation of the principles of grammar. The role of maturation is more plausibly restricted to the so-called “performance” aspects of language, such as an increase in short-term memory capacity, and an increase in perceptual acuity. These are quantitative, not qualitative changes. Importantly, these aspects of cognitive competence can be tested independently of our particular linguistic theory. With maturation of the sort espoused by Borer and Wexler, the choice of what matures is entirely theory-internal: the particular aspect of linguistic competence which is hypothesized to undergo maturation changes as our linguistic theory changes.

b. Lexical Learning. The acquisition of morphological properties interacts with syntactic principles to produce linguistic variation. Lexical Learning is essentially the reverse of the Lexical Projection Hypothesis, which asserts that developing syntactic structures drives morphological acquisition. Evidence against the LPH comes from the fact that the appearance of overt inflection is a function of the type of language being learned, invalidating the claim that there is universal, invariant stage of grammar without Functional Categories. Lexical Learning also contradicts the Full Competence Hypothesis, which does not allow for cross-linguistic variation in the number and type of Functional Categories instantiated in clausal structures. Evidence against this position comes from Rohrbacher
1994, who shows that an articulated Person agreement paradigm correlates strongly with V-to-INFL movement in Germanic and Romance.

c. Morpholexical Learning. Theories which possess both (a) and (b) must also be able to account for the developmental stages of child language. Morpholexical Learning asserts that language acquisition is lexically based, and consists of procedures for the creation of the lexicon and of morphological systems of knowledge. The acquisition of inflectional morphology is divided into two stages. The first stage involves the storage of inflected lexical items as unanalyzed wholes in paradigm slots (cf. MacWhinney 1978). This allows the child to simultaneously discover both the morphological categories which are relevant to his or her target language, and the formal processes employed in marking them. The second stage involves the extraction of regular, implicational morpholexical rules (Spencer 1988, 1990), which relates stored lexical items to each other, including sets of affixes which are segmented from stems. These affixes are then available for use on novel lexical items, and for “filling in” incomplete word-specific paradigms.

A highly desirable characteristic of this type of “Morpholexical Learning” is its robustness: noisy data and early mistakes need not spell doom for the language learner (cf. Spencer 1990). This is due to the fact that overt morphological distinctions determine the sorts of generalizations made by the child; instead of a blind system of rules, we have a body of data (lexical items) and a set of hypotheses (guided by paradigmatic structure), which may be modified as knowledge of the language increases. Paradigms play a crucial role by providing the structure necessary for the child to detect linguistically relevant patterns in the data.

Languages differ in the extent to which morphological generalizations can be formed. For example, some languages will systematically assign a particular category to an inflectional marker, while others will handle the same contrast with suppletive (i.e. independent lexical) items. During the initial stages of acquisition, we predict that all forms

---

15 In this thesis, I am essentially ignoring the problem of how children segment words from utterances in the first place. See Peters (1983, 1985) and Cutler (1994) for some proposals for how this takes place.
are learned as "suppletions". That is, as phonological units are carved out of the speech stream, they will be appended with notional content and stored in word-specific paradigms, where they will undergo comparison and analysis. Complex morphological relations between lexical items are not possible until this analysis has taken place. In other words, such relations are created during the course of language acquisition (Bybee 1988).

Initial "suppletive" learning is necessary due to three factors. First, there is no way of knowing in advance how the words being learned will be segmented (Pinker 1984). Second, since categorical content and formal expression are independent of each other (Sapir 1921), the child will not know where on the continuum of morphological expression (see Figure 1) a particular category will be placed. Third, the child will not know in advance which morphological categories receive overt expression in the language being learned. Thus, the child must store whole words along with contextual information in order to discover how form and meaning co-vary (Pinker 1984; Slobin 1985). This creates the following prediction: children will learn that a morphological category is Relevant to their target language before they have mastered the full, adult paradigm. In Chapter 4, I argue that this is true for children acquiring subject-verb agreement morphology in German.
4. Acquisition of German Verbal Inflection

In recent years, analyses of the acquisition of German inflection have increased in number, presumably due to the fact that German represents a language which is both closely related to English, and yet presents a greater degree of complexity in its inflectional component. This allows researchers to re-examine claims that have been made solely upon the basis of studies done in English. The greater amount of inflectional morphology to be learned lends itself well to testing hypotheses bearing on the nature and presence of functional categories in early child language.

Various claims have been made in the literature regarding the timing and significance of the acquisition of agreement in German. Poeppel & Wexler (1993) (henceforth P&W) claim that German children possess clausal structures which are identical to adult clausal structures (including an AgrP or IP) from the very beginning. Evidence for this position comes from the correct positioning of finite and non-finite verbs, and the accurate use of (a subset of) agreement markers. On the other hand, Clahsen argues that German children do not know agreement at first, producing “underspecified” clausal structures with a F(inite)P(hrase) instead of an AgrP (Clahsen 1990; Clahsen et al. 1994). He claims that the acquisition of agreement morphology serves as the “lexical trigger” for the insertion of AgrP in clausal structures, leading to the acquisition of V2 constructions and the setting of the ‘pro-drop’ parameter (Hyams 1986). Though he too demonstrates that verbal inflections in German emerge quite early, he argues that they are initially not used to encode the formal grammatical relation of agreement, but are instead initially used to mark (in-)transitivity.

The position for which I argue is that German children do know that Person/Number agreement is relevant to their language, but they have not yet mastered the full adult morphological agreement paradigm (cf. Verrips & Weissenborn 1992). Clahsen assumes that children are making a series of discrete hypotheses regarding the categorical nature of agreement affixes. This assumption makes the wrong predictions, because the type of development actually attested is gradual improvement in the use of inflectional markers.
This suggests increasing mastery of the morphological paradigm, not development of syntactic structure (cf. Hyams 1994). Because Clahsen claims that the 3rd Sing. [t] inflection is initially used as an intransitivity marker, he is unable to account for the distributional pattern shown by the other markers, such as 1st Sing. (=Imperative) [ø/e] and infinitival (=1st/3rd Plural) [en]. The account which I propose does not have this drawback; their use is determined by the sorts of paradigmatic structures which the child has constructed. In other words, the child is creating an inflectional system, not a series of independent inflectional rules.

In one respect, my position is closer to that of P&W rather than that of Clahsen, since I claim that children “know” agreement. Yet contra P&W, I claim that knowledge of Person/Number agreement does not entail that the child has discovered the exact formal processes by which Person/Number are predictably marked. Two forms of evidence show that German children do not initially possess productive control of the agreement paradigm: (1) lack of overregularization errors during the stages investigated; and (2) small degree of “overlap” in the number of inflected variants per verb. An account which assumes productivity in agreement marking also has no way of explaining the overextension of zero-marked stems in German child language. In my analysis, this overextension is a result of incomplete word-specific paradigms.

The sections of this chapter are organized as follows. Section 4.1 introduces the basic facts of German verbal morphosyntax, including a standard analysis of “V2” clausal structures as the product of V-to-COMP movement, followed by the movement of a maximal projection to the specifier of COMP. The next section (Section 4.2) presents data on the child subjects used for this study, as well as the coding methodology used (cf. Ingram & Thompson 1995). The data obtained from these subjects reveals three main types of “agreement” errors in German child language: (1) overextension of the infinitival [en] marker; (2) overextension of the 1st Person Sing./Imperative [ø] marker; and (3) restriction of 3rd Sing. [t] to intransitive contexts.
In Section 4.3.1, I examine the use of the infinitive ending [en]. It is by now well-known that children learning a variety of Germanic and Romance languages produce what are known as “Root Infinitives”, or matrix clauses with an uninflected main verb (Rizzi 1994). I argue that performance constraints on sentence production (cf. P. Bloom 1990) are causing the pragmatic omission of overt modals and auxiliaries; the exact same omission of modal verbs is something which also occurs in parental input (cf. Mills 1985). This interpretation of the data is supported by the work of Ingram & Thompson (1995), where we found a strong correlation between Root Infinitives and modal interpretations. This approach preserves the assumption of continuity in syntax.

In Section 4.3.2, I examine the errors involving zero-marked stems. Zero-marked stems are grammatical in German for 1st Person singular and Imperative, but are ungrammatical elsewhere.16 I argue that these errors are a result of the storage of inflected forms in incomplete word-specific paradigms. Construction of these paradigms is influenced by two factors: (1) systematic syncretism of 1st and 3rd Person forms in every sub-paradigm except for the Present Indicative Sing.; and (2) the syncretism between 1st Sing. and Imperative. These factors lead to word-specific paradigms in which the Person feature [±2nd] is operative, but not the Person feature [±1st].

Section 4.3.3 concerns the status of 3rd Sing. [t] in German child language. This marker is initially used (properly) with 3rd singular subjects, but the use of a 3rd singular subject does not entail the presence of [t]. That is, when present, it is used correctly, but it is only present in a small subset of obligatory contexts. Additionally, there is an initial strong correlation between the use of [t] and intransitive verbs (Clahsen 1986; Clahsen & Penke 1992). These data can also be explained by the hypothesis that words are initially stored in word-specific paradigms. The underuse of [t] follows from the fact that it is not a productive marker, while its restriction to intransitive verbs is a consequence of the fact that German parental input is strongly skewed in this direction (Section 4.3.4). This is due to

16 We are here restricting consideration to the Present Indicative paradigm, since forms from other sub-paradigms have (for the most part) not yet been acquired in the stages under consideration.
discourse strategies which dictate a high percentage of 3rd Persons as subjects of intransitive verbs, and 1st and 2nd Persons as subjects of transitive verbs (Hopper & Thompson 1980).

4.1. The Morphosyntax of German Verbs

German is a "verb-second" (V2) language, due to the fact that finite verbs always appear in the second position in matrix declarative clauses. This is a trait shared with most other Germanic languages. The standard analysis of this phenomenon (cf. den Besten 1989, among others) is movement of the verb from its base-generated position to the head of CP, followed by movement of a topicalized XP to the specifier of CP. Platzack & Holmberg (1989) suggest that movement to C is driven by the placement of a "finiteness operator" in this position. The structure underlying this movement is shown in (24).

(24)

```
CP
   spec C'
   COMP IP
      [+F] spec I'
      VP INFL
         spec V'
         NP V
```

The examples in (25) demonstrate that a variety of constituents can move into the specifier of CP. Example (25a) shows the subject NP in this position, (25b) the object NP, and (25c) a sentential adverb. When the subject does not appear in C-spec, then it most frequently appears immediately after the verb (25b,c).
When a finite modal verb or auxiliary appears in V2 position, the infinitival verb appears at the end of the clause. This has led to the analysis of German as an underlying SOV language. Because of this word order, movement of the verb to any position other than C is not detectable (see Section 3.2.2 above). This makes it difficult to test claims on the syntactic consequences of the acquisition of agreement. Examples of constructions with a modal or auxiliary plus non-finite main verb are given in (26).

(26) a. Du will-st heute ess-en.
   you want-2s today eat-inf.
   ‘You want to eat today.’

      Berta has(aux)-3s yesterday evening a-fem apple-cake eaten-part.
      ‘Berta ate an apple cake yesterday.’

   c. Ich werd-e nächst-e woche nach österreich flieg-en.
      I will(aux.)-ls next-fem. week to Austria fly-inf.
      ‘I will fly to Austria next week.’

   d. Ich weiss , dass du ein bö-s-es kind bist.
      I-Nom. know-Is, that you. a-neut. naughty-neut. child are-2s
      ‘I know that you are a naughty child.’

   e. Nicht beissen!
      not bite-inf.
      ‘Don’t bite!’

Example (26a) is a modal construction, in which the modal verb appears in V2 position (marked with 2nd Person agreement), and the main verb in Vf position, marked with an
infinitive ending. Examples (26b) and (26c) demonstrate X-Aux-Y-Verb constructions. It should be noted that the Present Perfect construction (26b) is typically used when referring to past events in conversational German, while Present Tense plus time adverbials replace most uses of the future auxiliary (26c). Example (26d) demonstrates that inflected verbs in subordinate clauses appear clause-finally. The fact that they appear in this position is derived by the structure in (25), since the presence of a complementizer will prevent the finite verb from raising to C. Finally, in (26e) we see one way of forming an imperative construction. This type of imperative is uncommon outside of child-directed speech.

German verbs inflect for Tense, Mood, Person, and Number. In (27), I have laid out the complete paradigm of the “weak” verb *machen* (‘to make’ or ‘to do’). This paradigm shows that the only marker which is not homophonous with any other is 2nd Sing. [st].

(27)

<table>
<thead>
<tr>
<th></th>
<th>Indicative</th>
<th></th>
<th>Subjunctive</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[Sing.]</td>
<td>[Plural]</td>
<td>[Sing.]</td>
<td>[Plural]</td>
</tr>
<tr>
<td>[2nd]</td>
<td>mach-st</td>
<td>mach-t</td>
<td>mach-e-st</td>
<td>mach-e-t</td>
</tr>
<tr>
<td>[1st]</td>
<td>mach-e/ø</td>
<td>mach-en</td>
<td>mach-e</td>
<td>mach-e-n</td>
</tr>
<tr>
<td>[3rd]</td>
<td>mach-t</td>
<td>mach-en</td>
<td>mach-e</td>
<td>mach-e-n</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Indicative</th>
<th></th>
<th>Subjunctive</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[Sing.]</td>
<td>[Plural]</td>
<td>[Sing.]</td>
<td>[Plural]</td>
</tr>
<tr>
<td>[2nd]</td>
<td>mach-t-est</td>
<td>mach-t-et</td>
<td>mach-t-est</td>
<td>mach-t-et</td>
</tr>
<tr>
<td>[1st]</td>
<td>mach-t-e</td>
<td>mach-t-en</td>
<td>mach-t-e</td>
<td>mach-t-en</td>
</tr>
<tr>
<td>[3rd]</td>
<td>mach-t-e</td>
<td>mach-t-en</td>
<td>mach-t-e</td>
<td>mach-t-en</td>
</tr>
</tbody>
</table>

Non-finite forms and the Imperative

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Imperative</td>
<td>mach-e/ø</td>
</tr>
<tr>
<td>Infinitive</td>
<td>mach-en</td>
</tr>
<tr>
<td>Past Part.</td>
<td>ge-mach-t</td>
</tr>
</tbody>
</table>
4.2. Subjects and Methodology

A total of six subjects were selected for this study on the acquisition of German subject-verb agreement. Cross-sectional data are available from three of these subjects: Andreas, Nicole, and Katrin. Longitudinal data are available from the other three: Mathias, Daniel, and Dorothy. The transcripts of Andreas, Katrin, and Nicole were taken from a database of 13 German children donated to the Child Language Date Exchange System (ChiLDES) (MacWhinney & Snow 1985) by Wagner (cf. Wagner 1985). The Mathias and Daniel corpora were also taken from ChiLDES, donated by Clahsen (cf. Clahsen 1986).

Transcripts of Dorothy were collected as part of a language acquisition research project at the University of British Columbia. Dorothy was a young child growing up bilingually (speaking both English and German) in Vancouver, Canada. Recording sessions were made with her once a week for nine months, during the ages 1;11 through 2;5. Ages and MLU for Dorothy, Nicole, Katrin, and Andreas are given in Table 3. The Mathias and Daniel corpora range in age from 1;6 to 3;6 (MLU 1.0 to MLU 4.53), with about 100 utterances per recording session (Clahsen 1986:85-6).

<table>
<thead>
<tr>
<th>Child</th>
<th>Age</th>
<th>Sample Size</th>
<th>MLU1a</th>
<th>Coded Sampleb</th>
<th>MLU2a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dorothy</td>
<td>1;11-2;5</td>
<td>2,761</td>
<td>d</td>
<td>316</td>
<td>1.87</td>
</tr>
<tr>
<td>Nicole</td>
<td>1;8</td>
<td>1,724</td>
<td>2.17</td>
<td>129</td>
<td>2.67</td>
</tr>
<tr>
<td>Katrin</td>
<td>1;5</td>
<td>1,437</td>
<td>2.72</td>
<td>235</td>
<td>2.54</td>
</tr>
<tr>
<td>Andreas</td>
<td>2;1</td>
<td>2,297</td>
<td>2.64</td>
<td>404</td>
<td>3.43</td>
</tr>
</tbody>
</table>

a MLU1 for Nicole, Katrin, and Andreas was determined by using the CLAN MLU program on the entire sample. MLU2 is the MLU using SALT (Miller & Chapman 1985) on the Coded Sample only. Only the latter was calculated for Dorothy.

b these are a subset of the total sample that were used in the subsequent analysis (see below).

c Dorothy samples are at 1;11(6), 2;0(3), 2;2(0), 2;2(26), 2;3(2), 2;4(6), 2;5(28)
d Dorothy’s MLU ranged from 1.27 to 2.56

17 Mathias and Daniel are twins.
18 Clahsen has made available to CHILDES sessions 17-19 and 21-27 of Mathias and Daniel, which allows us to look at data from children who are slightly more advanced than the others in this study. I will take advantage of Clahsen's data on these children when convenient, taking careful note of which data are taken from Clahsen (1986) and from Clahsen & Penke (1992).
Data from Nicole, Katrin, and Andrëas were converted from the CLAN format (MacWhinney & Snow, 1985) into SALT format (Miller & Chapman, 1985), while data from Dorothy was selectively entered into SALT format directly from transcriptions of the recording sessions: the criterion used in coding an utterance was that it both contain a verb, and be completely intelligible. Utterances which were identified as either (self-)repetitions or as formulaic were not included. Subsequent analyses of these four subjects were performed using either SALT or PAL (Pye, 1987). Sample size in the third column of Table 3 represents all of the transcribed utterances in the database, including many partially or completely unintelligible utterances. MLU1 was calculated based on this total. The coded sample (fifth column) represents all completely intelligible utterances containing a verb which were either declaratives or questions. MLU2 was calculated based upon this considerably smaller set of data.

For the subjects which were analyzed using SALT, the following conventions (developed in Ingram & Thompson 1995) were used in coding verb types. One code was used before the verb in order to identify it as such; this code was followed by the citation (infinitival) form of the verb. If more than one verb occurred in an utterance, each instance was marked. Two codes followed the verb, one indicating the actual phonetic realization of the verb by the child, and the other identifying the type of affix used. The format of this procedure is given in (28), and specific examples (from Andreas) are given in (29)

(28)   [verb code] + Verb + [phonetic form] + [affix]

(29)   a. [V]geben[bib][ø]
       b. [V]haben[hape][e]
       c. [V]kommen[pommt][t]
       d. [V]holen[holen][en]
Another set of codes was used at the end of utterances. These ‘utterance’ codes identified the Person of the subject and the modality of the sentence.\textsuperscript{19} In addition, the codes indicated whether the subject of the sentence was an overt NP, an overt pronoun, or an empty pronoun (pro)\textsuperscript{20}. In (30) I give examples of our coding scheme, based upon the sentences from Katrin in (18) and (19):

   b. Papa [V]gehen[geht][t] hin [3P][NP].
   c. ein Pferd [V]koennen[dann][ø] nicht [V]laufen[jauwen][en] [3P][NP][INF][MOD].
   d. diesen [V]wollen[wijst][st] du [V]haben[haben][en] [2P][PRO][INF][MOD].

When the child used her own name or that of the addressee as the subject of the sentence, the codes [NP1] and [NP2] were used to indicate this fact (as shown in (11a)).

4.3. The Developmental Paradox

The data obtained using the coding scheme in Section 4.2 are presented here for analysis. Table 4 gives the distribution of agreement markings and proportion of their correct use in the speech of Dorothy, Nicole, Katrin, and Andreas. Table 5 gives the same figures for Daniel and Mathias (from Clahsen 1986: Table 1).

<table>
<thead>
<tr>
<th>Agreement Marker</th>
<th>Dorothy</th>
<th>Nicole</th>
<th>Katrin</th>
<th>Andreas</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ø]</td>
<td>.43 (37)</td>
<td>.50 (4)</td>
<td>.53 (15)</td>
<td>.89 (114)</td>
<td>.59</td>
</tr>
<tr>
<td>[ø]</td>
<td>.02 (134)</td>
<td>.03 (97)</td>
<td>.05 (99)</td>
<td>.06 (165)</td>
<td>.04</td>
</tr>
<tr>
<td>[t]</td>
<td>.96 (27)</td>
<td>1.00 (2)</td>
<td>1.00 (25)</td>
<td>.96 (23)</td>
<td>.98</td>
</tr>
<tr>
<td>[e]</td>
<td>1.00 (4)</td>
<td>—</td>
<td>—</td>
<td>.96 (22)</td>
<td>.98</td>
</tr>
<tr>
<td>[st]</td>
<td>1.00 (5)</td>
<td>—</td>
<td>1.00 (11)</td>
<td>1.00 (8)</td>
<td>1.00</td>
</tr>
<tr>
<td>Total:</td>
<td>207</td>
<td>103</td>
<td>150</td>
<td>332</td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{19} The significance of the modality codes will be explained in Section 4.3.
\textsuperscript{20} We did not assign a Person code to an utterance in which no overt subject appeared.
Several clear patterns are demonstrated by the data in these tables. First, a large majority of verbs appear with one of three suffixes: [∅], [en], or [t]. The 1st Sing. [e] and 2nd Sing. [st] appear in frequency only towards the end of the period under study. Secondly, the markers [t], [e], and [st] are used for the most part correctly from their first appearance in the data. Third, the most frequently used markers, [∅] and [en] are used correctly much less often than the other markers, with Daniel and Mathias showing a gradual increase in accuracy as development progresses. How do we account for the anomalous behavior of these inflectional markers?

Further complicating matters, even the agreement markers which are used correctly are used only in a relatively small percentage of obligatory contexts. For example, even though 3rd Sing. [t] is used for the most part correctly when present, it is not used every time a 3rd Person subject appears. Tables 6 and 7 demonstrate percentage of use in obligatory contexts. Data in Table 7 is taken from Clahsen (1986: Table 2).
Table 6. Use in obligatory contexts of agreement markers
(Dorothy, Nicole, Katrin, and Andreas)

<table>
<thead>
<tr>
<th>Subject Person</th>
<th>Dorothy</th>
<th>Nicole</th>
<th>Katrin</th>
<th>Andreas</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Sing. b</td>
<td>.67</td>
<td>(1.0)a</td>
<td>1.0</td>
<td>.96</td>
</tr>
<tr>
<td>-e</td>
<td>.10</td>
<td>-</td>
<td>(.11)</td>
<td>.78</td>
</tr>
<tr>
<td>-0</td>
<td>.57</td>
<td>(1.0)</td>
<td>.89</td>
<td>.18</td>
</tr>
<tr>
<td>3rd Sing. c</td>
<td>.47</td>
<td>.21</td>
<td>.79</td>
<td>.60</td>
</tr>
<tr>
<td>-t</td>
<td>.43</td>
<td>.17</td>
<td>.73</td>
<td>.52</td>
</tr>
<tr>
<td>-0</td>
<td>.04</td>
<td>(.04)</td>
<td>.06</td>
<td>.08</td>
</tr>
<tr>
<td>2nd Sing.</td>
<td>.63</td>
<td>(.00)</td>
<td>.71</td>
<td>.36</td>
</tr>
<tr>
<td>3rd Plural</td>
<td>.38</td>
<td>(1.0)</td>
<td>1.0</td>
<td>.40</td>
</tr>
<tr>
<td>1st Plural</td>
<td>-</td>
<td>-</td>
<td>(1.0)</td>
<td>.75</td>
</tr>
</tbody>
</table>

a Parentheses indicate that the proportion of use is based upon less than three obligatory contexts. So, for example, Nicole used sentences with 1st Person subjects only twice, both of which contained verbs marked with [o].
b The figure above the dotted line represents the combined proportion of verbs appearing in sentences with a 1st Sing. subject which correctly appeared with either [o] or [e] suffixes.
c The figure above the dotted line represents the combined proportion of verbs appearing in sentences with a 3rd Sing. subject which correctly appeared with either [o] or [t] suffixes. Note that, unlike the 1st Sing. markers, these do not occur in free variation. There are many instances of zero-marked verbs in 3rd Sing. contexts which are incorrect with respect to the adult language.
d These are all of the subjects which appeared in the transcripts of Dorothy, Katrin, Nicole, and Andreas.

Table 7. Use in obligatory contexts of agreement markers
(Mathias and Daniel)

<table>
<thead>
<tr>
<th>Subject Person</th>
<th>II</th>
<th>Mathias</th>
<th>IV</th>
<th>II</th>
<th>Daniel</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Sing.</td>
<td>.59</td>
<td>.78</td>
<td>.98</td>
<td>.51</td>
<td>.66</td>
<td>.96</td>
</tr>
<tr>
<td>[o]</td>
<td>.49</td>
<td>.46</td>
<td>.88</td>
<td>.48</td>
<td>.56</td>
<td>.91</td>
</tr>
<tr>
<td>[e]</td>
<td>.10</td>
<td>.32</td>
<td>.32</td>
<td>.03</td>
<td>.10</td>
<td>.05</td>
</tr>
<tr>
<td>3rd Sing.</td>
<td>.32</td>
<td>.40</td>
<td>.90</td>
<td>.32</td>
<td>.51</td>
<td>.95</td>
</tr>
<tr>
<td>[o]</td>
<td>.07</td>
<td>.24</td>
<td>.53</td>
<td>.15</td>
<td>.38</td>
<td>.60</td>
</tr>
<tr>
<td>[t]</td>
<td>.25</td>
<td>.16</td>
<td>.17</td>
<td>.13</td>
<td>.35</td>
<td>.35</td>
</tr>
<tr>
<td>2nd Sing.</td>
<td>-</td>
<td>-</td>
<td>.92</td>
<td>-</td>
<td>-</td>
<td>.93</td>
</tr>
<tr>
<td>3rd Plural</td>
<td>.40</td>
<td>.60</td>
<td>.90</td>
<td>.33</td>
<td>1.0</td>
<td>.71</td>
</tr>
<tr>
<td>1st Plural</td>
<td>-</td>
<td>1.0</td>
<td>.82</td>
<td>1.0</td>
<td>-</td>
<td>1.0</td>
</tr>
</tbody>
</table>

a The first column is an exhaustive list of all the subjects which appeared in the data (Clahsen 1986:93). Data is arranged exactly as in Table 7.
According to Berman (1994), data such as these elevate the Developmental Problem to the level of paradox: how do we explain the fact that the development of child language is both very quick, and at the same time drawn out over a long period of time? "Three-year-olds manifest a near-complete mastery of the structural devices available to their native tongue. But they use them in very restricted contexts, from a syntactic and semantic as well as from a discourse point of view" (Berman 1994:229). Other studies have shown this to be a widespread pattern of development (e.g. MacWhinney 1978). In the sections which follow, I present my answer to this "paradox", relying upon the existence of word-specific paradigms to account for the pattern of error generation shown by these children. But first, (Section 4.3.1), I show that the instances of [en] in Tables 4 and 5 do not represent agreement errors at all, but are the result of clausal structures lacking modal verbs and auxiliaries.

4.3.1. Root Infinitives in German Child Speech

It is by now well-known that child speech in a number of Germanic and Romance languages contains large numbers of what are called "Root Infinitives" (Rizzi 1994; Wexler 1994). These are main clauses with an infinitive verb and no auxiliary—ungrammatical by adult standards. Though Root Infinitives contain no modals or auxiliaries, all the evidence indicates that German children correlate verbal morphology with clause position almost as soon as they produce multi-word utterances, with [+finite] verbs in V2 or V1 position, and [-finite] verbs in Vf position. P&W calculated the distribution of finite and non-finite verb forms for a subset of the corpus of Andreas, reproduced here in Table 8.

---

21 Though see below, where it is shown that dropping auxiliaries in modal contexts is a property of adult speech to children in German (cf. Mills 1985).
Table 8. Distribution of Andreas’
finite and non-finite verb forms (from P&W 1993)

<table>
<thead>
<tr>
<th></th>
<th>[+finite]</th>
<th>[-finite]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verb Second</td>
<td>216</td>
<td>7</td>
</tr>
<tr>
<td>Verb Final</td>
<td>15</td>
<td>44</td>
</tr>
</tbody>
</table>

These data demonstrate that Andreas is using the correct adult German syntax for [+finite] verbs at an early stage. In Table 9, I have calculated the percentage of sentences containing verbs marked with [en] which could be considered as [±finite], based upon their position within the clause, for the subjects Katrin, Nicole, Andreas and Dorothy. Since these data do not exclude questions (as Poeppel & Wexler did), the figures for Andreas are larger than in Table 8.

Table 9. Clausal position of verbs marked with [en]a
(Katrin, Nicole, Andreas, and Dorothy)

<table>
<thead>
<tr>
<th>Verb Position</th>
<th>Katrin</th>
<th>Nicole</th>
<th>Andreas</th>
<th>Dorothy</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1 or V2</td>
<td>.01 (1)</td>
<td>.05 (5)</td>
<td>.07 (11)</td>
<td>.10 (13)</td>
<td>.06</td>
</tr>
<tr>
<td>Ambiguousb</td>
<td>.55 (52)</td>
<td>.36 (34)</td>
<td>.22 (34)</td>
<td>.36 (49)</td>
<td>.37</td>
</tr>
<tr>
<td>VfC</td>
<td>.44 (41)</td>
<td>.59 (55)</td>
<td>.71 (110)</td>
<td>.54 (74)</td>
<td>.57</td>
</tr>
<tr>
<td>Total:</td>
<td>94</td>
<td>94</td>
<td>155</td>
<td>136</td>
<td></td>
</tr>
</tbody>
</table>

aIncluded in this table are all sentences with verbs which appeared with the [en] marker, with the exclusion of those occurring with overt modals and auxiliaries.
bVerb position counted as ‘ambiguous’ if the verb marked with [en] appeared in a single word utterance, or a two-word utterance with an overt subject.
cVerbs with [en] were counted as clause-final if they were the last term in a three-word or longer utterance. An empty subject (pro) counted as a single word in cases in which no overt subject appeared.

Almost all verbs marked with [en] can be considered non-finite, and not marked for subject agreement. For all but a tiny percentage of cases, the “incorrect” use of the [en] marker shown in Tables 4 and 5 does not really indicate agreement errors, but the use of non-finite [en] in correct Vf position. Examples of Root Infinitives from Nicole, Katrin, and Andreas are given in (31-33).
The examples in (31) through (33) demonstrate that verbs marked with [en] are almost always verb-final, the proper position for non-finite verbs, with (33a) representing one of the very few exceptions to this rule, and (32a) ambiguous in this respect due to the fact that it is a single word utterance. The “ungrammatical” aspect of these clauses is due for the most part then, to the fact that a finite auxiliary or modal is missing.

Researchers such as Rizzi and Wexler have taken the existence of Root Infinitives as evidence for an underlying formal property of child grammar. Rizzi (cf. Rizzi 1994) claims that early child grammars have the option of not projecting functional structure, so that when IP (or some other analogous FP) is not projected, the result is a Root Infinitive. Wexler (cf. Wexler 1994; P&W 1993) goes so far as to claim that there is a universal “Grammatical Infinitive Stage” in child grammar, which eventually disappears due to the maturation of values of Tense. The characteristics of his Grammatical Infinitive Stage are given in (34),
along with examples of sentences in (35) which should be equivalent according to this hypothesis.

(34) Grammatical Infinitive Stage (GIS) (Wexler 1994:312)

a. The child knows the possibility of head movement;
   b. He or she knows that head movement is forced in finite cases;
   c. He or she knows the Principle of Economy which implies that infinitival verbs do not move
   d. And yet, the child does not know that non-finite verbs cannot appear as main verbs.

(35) a. Nonfinite form
   er das haben (haben=Infinitive)
   he that have (= ‘he has that’ ?)

       b. Finite form
           er hat das (hat=3rd Person Present Indicative)
           he has that.

I will not go into the details of his account (the reader is referred to Wexler 1994:335ff), since it is not crucial to my analysis. The important matter here is that Wexler (and Rizzi) treat Root Infinitives as formal products of child grammar, generated according to principles which are not those of adult grammar.

As an alternative to the Grammatical Infinitive Stage, I advocate what P&W call the Modal Drop Hypothesis. According to this hypothesis, children are producing fully grammatical sentences (by adult standards), but they are omitting modals and auxiliaries due to production constraints on sentence length and complexity (P. Bloom 1990). An analysis along these lines preserves continuity in syntax, and does not require a maturational explanation for development. Support for this position comes from Jordens (1990), and Hoekstra & Jordens (1994), who claim that children produce Root Infinitives in irrealis contexts (Hoekstra & Jordens 1994:128). However, except for Poeppel and Wexler’s brief analysis of Andreas,22 no empirical tests have been made to see whether or not Root Infinitives are used in contexts which call for modals in adult speech.

22 Their database consisted of the root infinitives produced by Andreas which occurred sentence finally in declarative sentences having at least three terms: 37 utterances in all.
Ingram & Thompson (1995) investigated more closely the issue of whether or not Root Infinitives are simply formal alternatives to sentences with finite main verbs, or whether Root Infinitives are produced in contexts which can be clearly defined as irrealis, where the sentence is an expression of wanting, desire, requesting, etc. Our subjects for this study were Dorothy, Nicole, Katrin, and Andreas. Andreas, the child upon which P&W base their claims on this matter, represents the most advanced child in this study, with an MLU of 3.43 (see Table 3).

The first step in coding the data was to indicate which utterances contained a non-finite verb. This was accomplished by examining the data for verbs with [-en] endings, and determining whether or not these were used with a 1st or 3rd Plural subject (since the infinitive ending is homophonous with the markers agreeing with these subjects). If it could be determined with a reasonable amount of confidence that the verb was an infinitive, an utterance code [INF] was used to indicate this fact. This procedure therefore does not include any [-finite] verbs which occur in the data with [Ø] marking. Though this entails some loss of relevant data, we felt that the corresponding decrease in ambiguity vindicated the decision.

The second step was to examine the context in order to determine whether or not the verb was being used modally. When it occurred with a modal auxiliary, the decision was of course trivial. When the [-finite] verb was a Root Infinitive, other criteria had to be brought into the picture. We used four such criteria. The first was whether or not the parental input immediately preceding the child's utterance contained a modal verb or otherwise indicated modality (utterance codes [*MOD] and [PS]). The second was whether or not the parent responded to the child's utterance with an expansion including a modal verb, showing that the parent herself had given the child utterance a modal interpretation (utterance codes [*MOD] and [PR]). Also included as evidence here was whether or not the transcriber expanded the child utterance with a modal or otherwise implied existence of a modal interpretation in the contextual notes to the transcript (utterance codes [*MOD] and [PI]).
The final criterion, which we used after all other determinations had been made, was to
simply examine the situation context itself in order to determine whether or not it suggested
that the verb was being used modally (utterance code [*MOD*] by itself). We treated this last
criterion as the least reliable of the four, and tabulated it separately in the analysis of the
data. Examples of this coding scheme, with some contextual data included, are shown in
(36-38), based on the examples given in (31-33).23

(36) Nicole

N Papa da, (ich) Hut [V]haben[ham][en] [*PRO][INF][*MOD][PI]
Daddy there, (I) hat have-INF

% 'Papa da, ich möchte den Hut haben'
    'Papa there, I would like to have the hat.'

N Nicole Karte [V]haben[ham][en] [3P][NP1][INF][*MOD][PI]
N. map have-INF

% 'Nicole möchte die Karte haben'
    'Nicole would like to have the map.'

N Nicole Trinken [V]haben[ham][en] [3P][NP1][INF][*MOD][PI]
N. drink-NOM have-INF

% 'Nicole möchte etwas zu trinken haben'
    'N. would like to have something to drink.'

N da, Nicole [V]wollen[wej][ø] Trinken [V]haben[ham][en]
[3P][NP1][INF][*MOD].
    'there, N. wants drink have-INF

% 'da, Nicole will etwas zu trinken haben.'
    'there, N. wants to have something to drink.'

M willst du noch den Rest austrinken?
    'do you want to drink up the rest?'

N ja.
    'yes.'

23 The key to the transcript symbols are as follows: N=Nicole, K=Katrin, A=Andreas, M=child's mother,
V=child's father. Lines marked with ' % ' indicate the transcriber's interpretation of the child's utterance.
(37) Katrin

K (du) [V]haben[haben][en] *[PRO][INF][*MOD][PI]? (you) have-INF

% ‘Möchtest du die Eierschale haben?’
‘Would you like to have the egg shell?’ (addressed to her mother)

M ja, dann müssen wir es hier in den Müllimer tun, ne?
‘Yes, then we must (put) it here in the garbage can, no?’

K (ich) da daraus [V]machen[maren][en] *[PRO][INF][*MOD][PR]. (I) there out-from make-INF

% (she takes the spoon out of the egg-holder)

V was machen?
what make-INF

K da raus machen.
‘Take it out.’

(38) Andreas

M komm, wir gehen, der Thorsten soll noch ein bisschen schlafen.
come, we go, T. should yet a little sleep-INF

sleep-INF my brother

M so jetzt will ich mich mal eben anziehen
so now want I me just now get-dressed-INF
‘I want to get myself dressed just now.’

A (e ne) Mama mir auch [V]anziehen[anziehn][en]
[3P][NP][INF][*MOD][PS]
Mommy me-DAT also get-dressed-INF
‘Mommy, I want you to dress me too.’

In Table 10 (from Ingram & Thompson 1995), I give the results of our analysis. The first total (total 1) represents the number and proportion of utterances which satisfied all of the modal criteria except those marked only as “possible modal contexts” (*MOD], with no further code). The second, larger total (total 2) does not exclude such utterances.
Table 10. Use of nonfinite and finite forms in modal contexts.\(^c\)

<table>
<thead>
<tr>
<th></th>
<th>Dorothy</th>
<th>Nicole</th>
<th>Katrin</th>
<th>Andreas</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>number</td>
<td>139</td>
<td>93</td>
<td>103</td>
<td>162</td>
<td>124</td>
</tr>
<tr>
<td>contain modal</td>
<td>.09 (13)</td>
<td>.03 (3)</td>
<td>.17 (17)</td>
<td>.05 (8)</td>
<td>.07</td>
</tr>
<tr>
<td>missing modal</td>
<td>.32 (44)</td>
<td>.66 (61)</td>
<td>.70 (72)</td>
<td>.20 (33)</td>
<td>.47</td>
</tr>
<tr>
<td>PS(^a)</td>
<td>20</td>
<td>14</td>
<td>7</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>PR(^a)</td>
<td>43</td>
<td>26</td>
<td>20</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>0</td>
<td>37</td>
<td>57</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>total 1</td>
<td>.41 (57)</td>
<td>.69 (64)</td>
<td>.86 (89)</td>
<td>.25 (41)</td>
<td>.55</td>
</tr>
<tr>
<td>possible modal</td>
<td>.12 (17)</td>
<td>.31 (29)</td>
<td>.02 (2)</td>
<td>.49 (80)</td>
<td>.24</td>
</tr>
<tr>
<td>total 2</td>
<td>.53 (74)</td>
<td>1.00 (93)</td>
<td>.88 (91)</td>
<td>.75 (121)</td>
<td>.79</td>
</tr>
</tbody>
</table>

|                      | 25      | 25     | 25     | 50      |       |
| number               |         |        |        |         |       |
| contain modal        | .00     | .00    | .00    | .00     | .00   |
| missing modal        | .02 (1) | .16 (4)| .04 (1)| .02 (1) | .06   |
| total                | .02 (1) | .16 (4)| .04 (1)| .02 (1) | .06   |

\(^a\) these are the number of times each code was used. Sentences could contain one, two, or all three of these codes so no proportions are given.

\(^b\) parallel to P&W, only declarations were coded; [en] forms were excluded as possible finite forms.

\(^c\) when proportions are given, the number of instances is provided within parentheses.

Also included in Table 10 are the results of coding the first 25 [+finite] verbs (50 for Andreas) in the transcripts of these children, subjected to the same modality criteria. This was done to see if the figures for the [-finite] verbs were not simply reflecting some underlying, overall distribution of modal and non-modal utterances.

The data show a strong correlation between [+modal] contexts and [-finite] verbs, including those sentences which are Root Infinitives. The mean proportion of utterances satisfying the criteria for the first total was .55. This ranges from a low of .25 for Andreas to a high of .86 for Katrin. The second total showed a mean proportion of .79 satisfying the criteria, with a low of .53 for Dorothy, and a high of 1.00 for Nicole. No such correlation was found between modality and [+finite] verbs, with a mean proportion of .06 utterances satisfying all of the modal criteria. These results provide strong confirmation for the Modal Drop Hypothesis, and strong disconfirmation of Wexler’s claim that Root Infinitives are not correlated with irrealis contexts.
It will be noted that not all Root Infinitives were found to have a modal interpretation. A possible explanation for this fact goes as follows. It might be expected that children produce not only utterances with modal interpretations, as in (26a), but also verbs with past-tense meanings, as in (26b) (or, less likely due to the infrequency of this construction in German, with future meaning as in (26c)). These constructions also involve Vf non-finite verbs, namely perfective participles, combined with the auxiliary verbs haben or sein. Mills (1985) shows that young German children often omit the initial, unstressed [ge-] prefix from participles, which with strong verbs leaves a verb of the form STEM + [en]. If they were to also omit the finite auxiliaries from such utterances, we would then see what appears to be Root Infinitives from the point of view of the adult. These utterances would not have a modal meaning, but rather a perfective (or stative) interpretation.

Jordens (1990) provides further evidence that Root Infinitives correspond to sentences with modals and auxiliaries in adult speech. He plotted the production of overt modals and auxiliaries versus the production of Root Infinitives in the speech of Mathias and Daniel, and showed that they are negatively correlated. That is, as the proportion of Root Infinitives goes down, the proportion of utterances with modals and auxiliaries goes up (Jordens 1990: figures 1 and 2). Jordens shows that, in the timespan of about one month, there is a dramatic increase (from around 40% to 90% of sentences with infinitives) in the use of overt auxiliaries, and a corresponding decrease in Root Infinitives. It would be a remarkable coincidence if these two trends were not in some way related.

As Jorden notes, this increase in the use of overt modals and auxiliaries also explains the sudden increase in V2 constructions noted by Clahsen (1986). Clahsen attributes this increase in V2 constructions to the acquisition of agreement. However, as will be shown in Section 4.3.2, knowledge of agreement is acquired extremely early, certainly during the same stages when Root Infinitives are being produced. Jordens' findings not only support the Modal Drop Hypothesis, they also support the conclusion that
acquisition of V2 is not triggered by the acquisition of agreement (cf. Verrips & Weissenborn 1992).24

Further evidence that German children are dropping modals, rather than generating “grammatical infinitives” comes from parental input. Every child in the current study received input sentences in which the parent dropped a modal verb, presumably because the context made it clear what the parent meant (cf. Mills 1985, who found the same thing). Examples from transcripts appearing in Miller 1976 are given in (39):

(39) a. Bonbon rausholen?
candy out-fetch-INF
‘Do you want me to get the candy out?’ (1976:287)

b. Was woll’n wer denn nochmal... Bilder angucken?
what want we then again (do)... pictures look-at-INF
‘What do we want to do again? Look at pictures?’ (Ibid.:316)

c. Buch ansehen?
book look-at-INF
‘Do you want to look at the book?’ (Ibid.:328)

Taking sentences such as these into account, the only thing odd about the production of Root Infinitives in German child language is that they are more prevalent than in adult speech. Given the fact that children have tighter limits on short-term memory capacity, and their production capabilities are impaired compared to adults, it would not be surprising if it is found that they are more prone to omit auxiliaries. It may also be the case that they lack the proper pragmatic knowledge of when it is valid to drop constituents (Weissenborn 1990).

The results of this section lead us to three conclusions. First of all, the use of [en] in German child language does not represent an agreement error, but the correct use of the infinitive marker in clause-final position. Secondly, evidence supports an interpretation of Root Infinitives which preserves continuity in syntax, contra Wexler (1994). We do not

---

24 This conclusion was already thrown into doubt by the fact that V2 languages such as Swedish have absolutely no agreement morphology at all. In light of this, Clahsen & Penke (1992) were forced to the conclusion that there are a variety of triggers for a “V2-parameter”.

66
require an analysis of these constructions which includes the development of abstract principles of grammar. The third conclusion is that children morphologically and semantically distinguish [+finite] verbs from [-finite] verbs almost from the very beginning of multi-word speech. The [±finite] distinction correlates strongly with proper positioning of the verb within the clause. This suggests knowledge of the structure in (24), with a Finiteness Operator in the Complementizer position. On the other hand, there is no evidence for or against an Agr node in the clausal structures of this period.

4.3.2. Segmentation of Affixes

In this and the next two sections, I will deal with the remaining "errors", namely, the overextension of zero-marked stems, and the restriction of 3rd Sing. [t] to intransitive verbs. I argue that the pattern of inflectional marking shown by these children is due to the fact that they have not yet extracted productive rules for marking agreement on verbs. The underuse of [t] follows from the fact that it is not a productive inflectional marker, while its restriction to intransitive verbs is a product of the discourse properties of parental speech, which itself shows tendencies to use intransitive verbs with 3rd Person subjects (Section 4.3.3). The overuse of zero-marked stems is a product of the incomplete word-specific paradigms which children have constructed. Syncretism of 1st Person Sing. [ø] with Imperative [ø] causes the child to construct agreement paradigms which distinguish 2nd from 1st and 3rd, but not 1st and 3rd from each other (Section 4.3.3). In this section, I show that segmentation of affixes from stems has not taken place in the stages under examination for Dorothy, Nicole, Katrin, and Andreas, while the more advanced stages of Mathias and Daniel (Clahsen’s Phase III) show evidence for the onset of productivity in the use of inflectional markers.

Many researchers have noted the necessity for providing strict productivity measures in order to assess whether or not a particular word or phrase is being actively constructed, or is merely being reproduced as an unanalyzed whole--stored in its entirety in rote memory.
(MacWhinney 1978; Ingram 1985). Recently, Marcus et al. (1992) have provided extensive support for a division of inflectional labor between an associationistic memory and a productive, symbol-based concatenation operation. The well-known “U-shaped learning curve” demonstrated in the acquisition of inflection can quite simply be accounted for by positing the initial rote storage of words, followed by the extraction of a regular rule. Thus, the onset of overregularization errors indicates the point at which the child has analyzed and segmented the words stored in memory.

In addition to the appearance of overregularization errors, other criteria may be used in order to judge whether or not a form is being used productively. One such test is “diversity of attachment” (Allen 1994). This measure takes into consideration the number of different forms in which a particular word type appears. So, for example, if we were to find a large number of verb types which appeared in a variety of inflected forms, this would constitute *prima facie* evidence for the emergence of productive rules of agreement. By itself, such a test does not constitute proof that such rules do or do not exist. But if it is the case that verbs show a great degree of “overlap” (i.e. appearing in more than one inflected variant), and overregularization errors are being produced at the same time, then we can be reasonably confident that a productive rule has been acquired. In the case that neither of these tests comes out positive, we may conclude that a productive rule has not yet been acquired.

Ingram & Thompson (1995) performed an analysis on Dorothy, Nicole, Katrin, and Andreas in order to determine the degree to which individual verbs appear in more than one inflected variant. First, we determined whether or not a verb appeared in both finite and non-finite variants. The results of this analysis are shown in Table 11.

---

25 Overregularization is then due to two factors. Normally, accessing a form in storage would block the application of the regular rule. However, given that recall is probabilistic rather than perfect, one would expect the regular rule to over-apply in a small number of cases (cf. Marcus et al. 1992). The important point for the case under study here is that the initial rote storage of items is a well-supported fact.
Table 11. Occurrence of verbs in finite and/or nonfinite variants
(Dorothy, Nicole, Katrin, and Andreas)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Dorothy</th>
<th>Nicole</th>
<th>Katrin</th>
<th>Andreas</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of total verbs(a)</td>
<td>38</td>
<td>32</td>
<td>69</td>
<td>81</td>
<td>55</td>
</tr>
<tr>
<td>number of coded verbs(b)</td>
<td>13</td>
<td>9</td>
<td>21</td>
<td>27</td>
<td>18</td>
</tr>
<tr>
<td>Finiteness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>finite form only</td>
<td>.46 (6)</td>
<td>.56 (5)</td>
<td>.43 (9)</td>
<td>.44 (12)</td>
<td>.47</td>
</tr>
<tr>
<td>nonfinite form only</td>
<td>.31 (4)</td>
<td>.33 (3)</td>
<td>.48 (10)</td>
<td>.41 (4)</td>
<td>.38</td>
</tr>
<tr>
<td>both forms(c)</td>
<td>.23 (3)</td>
<td>.11 (1)</td>
<td>.10 (2)</td>
<td>.15 (4)</td>
<td>.15</td>
</tr>
<tr>
<td>Multiple Finite Forms(d)</td>
<td>.23 (3)</td>
<td>.22 (2)</td>
<td>.00 (0)</td>
<td>.07 (2)</td>
<td>.13</td>
</tr>
<tr>
<td>prop. of coded verbs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(a\) number of verbs that occurred at least once in the data
\(b\) number of verbs which met the frequency criteria (at least 3 occurrences)
\(c\) verbs which occurred in both finite and nonfinite forms
\(d\) these are the finite forms which occur in more than one variant

The first row of the table represents the number of verb types which appeared in the data for each of the four children. The second row represents the subset of these which appeared at least three times--our frequency criterion for evaluation. This criterion was applied in order to be reasonably sure that our results were not contaminated by experimental error. It can be seen that a large majority of verb types for all the children occurred in only one form--either finite or non-finite. The proportion of verb types which appeared in both finite and non-finite forms ranges from .10 to .23.

The second measure of overlap was performed on those verbs which appeared in finite form. We tested to see how many of these verbs appeared with more than one subject agreement marker. The last row of Table 11 presents the results of this measure: the number of verb types which appeared in more than one finite form ranges from zero to three. These verbs are given in Table 12.
Across all four children, only four verb types appeared with more than one finite marker. These verbs consist of four of the most common, multi-purpose verbs in the German language: *kommen* (‘to come’), *gehen* (‘to go’), *machen* (‘to make’ or ‘to do’), and *haben* (‘to have’, also an auxiliary). These results are in accordance with those of Clark (1978), who shows that “light” verbs of this sort are the first to be acquired in a number of languages, including English.

The results of the “diversity of attachment” test indicate that these children have not yet acquired productive control of the verbal agreement morphology of German. Only the very most common verb types show multiply inflected variants—a fact which is hardly surprising, and which supports rather than refutes the hypothesis of rote memorization.

What about the existence/non-existence of overregularization errors? Once again the data do not indicate productive concatenation of suffixes to stems. Out of all of the coded utterances from Dorothy, Nicole, Katrin, and Andreas, only two overregularization errors were found, both from Andreas.

(40) a. es weinachten werd
   it Christmas becomes
   ‘It’s becoming Christmas.’
   (correct: ‘es wird Weinachten’)

b. meiner fahr-t
   mein-masc/Nom. drive-3s
   ‘My (car) is driving.’
   (correct: ‘meiner fährt’)

<table>
<thead>
<tr>
<th>Verb</th>
<th>Dorothy</th>
<th>Nicole</th>
<th>Katrin</th>
<th>Andreas</th>
</tr>
</thead>
<tbody>
<tr>
<td>kommen</td>
<td>[∅],[t]</td>
<td>[∅],[t]</td>
<td>[∅],[t]</td>
<td>[∅],[t]</td>
</tr>
<tr>
<td>gehen</td>
<td>[∅],[t]</td>
<td>[∅],[t]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>machen</td>
<td>[∅],[t]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>haben</td>
<td>[∅],[t]</td>
<td></td>
<td></td>
<td>[∅],[e],[t]</td>
</tr>
</tbody>
</table>
A different situation is shown by the twins Mathias and Daniel. The longitudinal data from these children is from a somewhat more advanced stage than that of the others. Here we see the emergence of overregularization errors of the sort shown in (41).

(41) a. i(ch) weisse nich
    I know-1s not
    ‘I don’t know.’
    (correct: ‘ich weiß’) (Mathias, corp. 19)

    b. ich kanne drinsitzen
    I can-1s inside-sit
    ‘I can sit inside
    (correct: ‘ich kann’) (Mathias, corp. 21)

    c. sehst du nich(t)?
    see-2s you not
    ‘Don’t you see?’
    (correct: ‘du siehs-st’) (Daniel, corp. 27)

    d. ich woll gleich meine rollen lassen
    I want now mine roll-inf let-inf
    ‘I want to let mine roll now.’
    (correct: ‘ich will’) (Daniel, corp. 27)

This suggests that Mathias and Daniel have begun to segment the verbal agreement suffixes from their stems. Clahsen (1990) notes that overregularization errors first come in during his Phase III. We may conclude that this marks the beginning of productivity in the use of verbal suffixes. This evidence is in striking contrast to that found for the other four children.

4.3.3. Lexical Storage in Paradigms

The overextension of zero-marked stems is a common error in German child language. In adult speech, they are grammatical only for 1st Person and Imperative, as well as 3rd Person for a small set of irregular verbs (the preterite-presents). Examples of

26 Unfortunately, time constraints have prevented me from performing the same diversity of attachment test on Mathias and Daniel as that performed on Dorothy, Katrin, Nicole, and Andreas. I hope to address this deficiency in future work.
incorrect zero-marked forms in the speech of Katrin, Andreas, and Nicole are given in (42), (43), and (44):

(42) Katrin
    a. (pro) kann fahr
        (it) can-Ø drive-Ø
        '(The car) can drive.'
    b. Sonne schein
        sun shine-Ø
        '(The) sun shines.'

(43) Andreas
    a. ich nimm dies
        I take-IMP this
        'I take this.'
    b. so mach der
        so make-Ø he
        'He does it so'

(44) Nicole
    a. Nicole meh mit
        N. go-Ø with
        'Nicole goes with.'
    b. Nicole komm dahin
        N. come-Ø in
        'Nicole comes in.'

These examples show that zero-marked forms are found in a variety of environments, such as infinitive (42a), 3rd Sing. (42b, 43b, 44a,b), and 1st Sing. (43a). In this last example, we have an Imperative form used with a 1st Person subject, since the verb nehmen ('to take') is one of a small but frequently used sub-set of 'strong' verbs which change their stem vowel from [e] to [i] in the Imperative (as well as 2nd and 3rd Sing. Present Indicative).

Two acquisitional questions are relevant to how children are using these forms. First, how does the child segment affixes from stems? Second, how does the child discover
zero morphemes? These two problems are closely related, since the discovery of a zero-morpheme is equivalent to the discovery that a categorical feature is not overtly marked by a segmentable affix. Segmentation takes place most efficiently through a comparison of related word forms: a single inflected form itself provides few clues as to whether it is mono- or poly-morphemic. A single word could be composed of stem plus affixes (infixes, prefixes, suffixes), consonant-vowel pattern alternations (as in Arabic), or stem plus zero morphemes. Byee (1991:83) states: "Inflectional affixes are acquired as parts of whole words, and the learner begins to form a notion that a meaningful word ending exists as s/he experiences it on a variety of different stems." By initially storing words as wholes in paradigms, the child is provided with the necessary organizational structure for a comparison of related word forms to take place. My account of errors such as those in (42-44) depends upon this claim.

When the child initially hears a word which it can analyze (i.e. segment and pair with semantic content), it will store this word form in the appropriate phonological and semantic co-ordinates in the lexicon. The most frequent (and probably semantically most basic) forms will be stored first. If the child has heard only a single form, the word-specific paradigm which results will be degenerate in nature. As multiple inflected variants of a single word are listed in the lexicon, the paradigmatic space will be subdivided according to a particular hypothesized dimension, with each dimension satisfying Closed-Category Relevance (3). As inflected word forms are integrated into the paradigm, the normal situation is for semantically "unmarked" forms to function as bases from which other, more complex forms are derived (Bybee 1985, 1991; Bybee & Brewer 1980; Wurzel 1987).

Bybee & Brewer (1980) show that there are at least two, possibly three candidates for the basic form of a paradigm. The 3rd Sing. form (the "unmarked" Person) most often forms the default base from which other forms are derived, though sometimes the 1st Sing. or Imperative forms play this role. In German (as in English), it is the 1st Sing. and
Imperative forms which are zero-marked, while 3rd Person forms (and 2nd Sing. for German) have inflectional endings. The zero-marking of 1st Sing. (while simultaneously marking 3rd Sing.) represents a basic-derived relation which is typologically unusual; it is not clear what factors are involved in creating this marked situation. One consistency, at least, is that German morphologically distinguishes 2nd Person from 1st and 3rd Person in every sub-paradigm.

The overregularization error from Andreas in (40b) (*fahr-t instead of fähr-t) shows what appears to be overuse of the 1st Sing. stem in 3rd Sing. contexts. Mills (1985) notes this same phenomenon for other children (taken from Bybee 1991:75):

(45) Present Indicative Child's form 'run'
1st Sing. lauf-e
2nd Sing. läuf-st
3rd Sing. läuf-t lauf-t

This use of the non-umlauted stem for 3rd Sing. implies that German children are using the 1st Sing. as the base form of the paradigm. Bybee and Brewer (1980) propose no explanation for what factors determine which form performs this role. I tentatively propose that it is the syncretism between Imperative and 1st Person which is responsible.

To see how this proposal works consider how the paradigm for the verb machen ('to make' or 'to do') would be constructed during the word-specific paradigm stage. Suppose that the first use of this verb was heard by the child in an imperative construction:

(46) a. Mach das nicht!
  make/do-Ø that not 'Don't do that!'

\[27\] Actually, they are optionally marked with [Ø] or [e]. Some researchers feel that these are simply phonological variants of one another (e.g. Wunderlich & Fabri 1994), though most modal verbs (including the verb wissen) obligatorily take [Ø] instead of [e].

\[28\] Bybee and Brewer invoke the notion of 'autonomy', or the likelihood of a word being stored in the lexicon rather than being derived. This just pushes the problem further back, since now we need to account for why 1st Sing. is autonomous in some languages, while 3rd Sing. is autonomous in others.
The verb extracted from (46a) will cause the child to set up a lexical entry for *machen* which contains only the zero-marked stem [mach]. Now suppose the child hears the sentence in (47a):

(47) a. Ich mach-∅ jetzt ein Haus.
   I make-1s now a house
   *I am making a house now.*

b. [+Imper.]
   / [+finite]
   mach

Because the forms for Imperative and 1st Person forms are marked identically (with [∅]), the paradigm is not “split” (Pinker 1984). The child is led to the conclusion that the Person feature [+1st] is not Feature-Value Relevant to his or her target language. Additionally, since there are no other person marked forms in this paradigm (yet), neither does Closed-Category Relevance hold for the category Person.²⁰

In (48a), the child is presented with the 2nd Sing. form [machst]. This causes a split in paradigm (47b), resulting in the paradigm (48b). At this stage, the child has

²⁹ In previous examples I have used privative, rather than binary features. In this section I switch to binary features, in order to capture the Person/Number categorizations which German shows. In particular, German groups together 1st and 3rd Persons in every sub-paradigm except for Present Indicative Sing. This indicates a division between [+2nd] and [-2nd] (cf. Wunderlich & Fabri 1994).

³⁰ In Bybee’s (1985) survey of 50 unrelated languages, she found that zero-marking is common in the imperative, and when imperative markers exist, they occur closer to the verb than person/number markers in every language except one (Bybee 1985:173). This latter fact suggests the hierarchy Mood > Person/Number.
discovered that the feature [+2nd] is Feature-Value Relevant to *machen*; this also entails that Person is Closed-Category Relevant.

    you make-2s also a house
    'You are also making a house.'

b. 

\[
\begin{array}{|c|}
\hline
\text{[+finite]} \\
\hline
\text{[+2]} & \text{machst} \\
\hline
\text{[-2]} & \text{mach} \\
\hline
\text{[+Imper.]} & \\
\hline
\text{mach} & \\
\hline
\end{array}
\]

Notice that at each stage, the overgeneration of zero-marked stems is predicted to occur. In (46), the child has not been exposed to any Person-marked forms, and uses the Imperative as a default. In (47), the child constructs a paradigm in which Person is not Relevant. This once again leads the overextension of the zero-marked stem. In (48), the child sets up a paradigm in which the feature [+2nd] is active, but not the feature [+1st]. During this stage, we predict the overextension of zero-marked stems to 3rd Person contexts. Only when the child has heard *machen* inflected for 3rd Person Sing. will he or she be led to create the paradigm in (49).

(49) 

\[
\begin{array}{|c|}
\hline
\text{[+finite]} \\
\hline
\text{[+2,-1]} & \text{machst} \\
\hline
\text{[-2,+1]} & \text{mach} \\
\hline
\text{[-2,-1]} & \text{macht} \\
\hline
\text{[+Imper.]} & \\
\hline
\text{mach} & \\
\hline
\end{array}
\]
This scenario is only one of many possible ones. A more detailed account must be able to handle all of the permutations in the orders in which inflectional markers are acquired. The main fact which must be accounted for is that German children take the 1st Sing. as the base form of the paradigm. This is shown by overregularization errors (as in (40) and (45)), and by the fact that zero-marked stems are extended to inappropriate contexts. I have proposed here that the syncretism between 1st Person Sing. and Imperative is a factor in how the child decides which form of the paradigm is the base. Examples such as (43a) suggest that Andreas has constructed paradigms according to the principle 1st Sing. = Imperative, causing him to substitute the Imperative form even when inappropriate.

Further evidence, from Mathias and Daniel, shows that the incorrect use of zero-marked stems is negatively correlated with the correct use of 3rd Sing. [t]. This supports the claim that children are in the process of learning when and where it is appropriate to use zero-marked stems. In Table 13 (data from Clahsen 1986: Table 2), we see how the pattern of use of inflectional markers with 3rd Person subjects develops over time.
Table 13. Inflections which appear with 3rd Sing. subjects
(Mathias and Daniel) (Data from Clahsen 1986: Table 2)31

<table>
<thead>
<tr>
<th>Inflections used with 3rd Sing. subject³⁵</th>
<th>Mathias</th>
<th>Daniel</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>II</td>
<td>III</td>
</tr>
<tr>
<td>[t]</td>
<td>.25d</td>
<td>.16</td>
</tr>
<tr>
<td>[ø] (correct)²³</td>
<td>.07</td>
<td>.24</td>
</tr>
<tr>
<td>[ø] (incorrect)²⁴</td>
<td>.56</td>
<td>.20</td>
</tr>
<tr>
<td>[en]</td>
<td>.12</td>
<td>.01</td>
</tr>
<tr>
<td>[e]</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>[st]</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>pronominal copy</td>
<td>.00</td>
<td>.40</td>
</tr>
</tbody>
</table>

³These are the inflections used at least once with a 3rd Sing. subject.
²These are instances of [ø] used with modals, the copula, and the verb wissen.
³These are instances of [ø] used with regular verbs.
⁴Each column sums to 1.0. Decimals mark the proportion of sentences with 3rd Person subjects in which the particular inflection is present on the verb.

During Phase III there is a striking decrease in the proportion of verbs incorrectly marked with [ø] (from .56 to .20 for Mathias, and from .44 to .09 for Daniel). This coincides with the onset of overregularization errors such as those in (41), which as Clahsen (1991) notes, first appear in the transcripts of Mathias and Daniel at this time. This means that Phase III is when they have attained true productive control of the verbal agreement system, though the large number of “pronominal copies” produced during this period presents an area which remains to be explored. Once 3rd Sing. [t] has been segmented and inserted into a “general paradigm” (Pinker 1984), it can be applied to any item which belongs to the category of Verb. This means that the use of the default, zero-marked forms should (and do) decrease from this point on. The only exceptions to the absolute productivity of [t] are the members of the class of preterit-presents.

31 "Pronominal copies" indicates a phenomenon reported in Clahsen (1986), in which the child uses what looks like pronouns as a sort of inflectional ending. Examples are given in (i) and (ii)

(i) da is-er großer fisch
   'There is-he big fish'
(ii) das is-er fest
    'There is-he stuck'

This sort of child error is probably due to input sentences of the sort 'Da ist er' ('There he is'), in which the unstressed pronoun immediately follows the verb. Here we see another instance of the difficulties involved in properly segmenting the speech signal.
4.3.4. Discourse Properties of Parental Input

As noted above, Clahsen’s answer to the developmental “paradox” shown in Tables 4-7 is to deny that German children have acquired the formal notion of agreement at the stage in question. He claims that verbal inflections in German are first used by children as transitivity markers. Evidence for this analysis is an initial correlation between the use of 3rd Sing. [t] and verbs with low transitivity: the use of this inflectional marker seems to be restricted mostly to sentences with one-place predicates (Clahsen 1991:376). Examples are given in (50) and (51):

(50) Mathias (Clahsen 1986:97)
   a. das    auto fährt    schnell
      the/neut. car drive-3rd/Sing. fast
      ‘The car is driving fast.’
   b. dreht    immer
      (pro) turn-3rd/Sing. always
      ‘That always turns.’

(51) Daniel (Ibid.)
   a. passt schon ander
      pro fit-3rd/Sing. already another
      ‘Another one already fits.’
   b. (s)pass macht
      fun     make-3rd/Sing.
      ‘This is fun.’ (lit.: ‘makes fun.’)

If the [t] marker in German is initially used to mark intransitive contexts, then we have an explanation for their high correlation with intransitive verbs, and for their low percentage of use in obligatory contexts.

Because of the (hypothesized) absence of knowledge of agreement, Clahsen argues that the clausal structures underlying the early grammars of children acquiring German do not possess an AgrP. Instead of the adult clausal structure represented in (24), Clahsen (1991) proposes the following:
The topmost phrase in this structure is a projection of the feature [+F(inite)]. Lexical items which bear the [+F] feature are base-generated in head of FP. In Clahsen's analysis, [+F] items are restricted to modals and the [t] marker. When [t] is generated, this causes the main verb to move into the [+F] position, in order to avoid violating the Stray Morpheme Filter. If [t] does not appear, then the lexical verb remains in its base-generated position. During the period in which this structure is operative, [t] marks clausal (in)transitivity, while [∅] and [en] "do not yet have content" (Clahsen 1991:382).

This analysis has a variety of flaws. For example, if 3rd Sing. [t] were simply a transitivity marker, it remains something of a mystery why it should almost always be used correctly. Tables 4 and 5 show that the chance of there being a 3rd Sing. subject given the presence of a verb marked with [t] is extremely high (in many cases 1.0). The uncertain status of the other inflectional endings in Clahsen's analysis is another problematic aspect. Notice the oddness of the situation represented by the structure in (52). Only one suffix (3rd Sing. [t]) has attained any significance to the child. The other suffixes ([∅] and [en]) are left in grammatical limbo, with no function associated to them.

This picture of acquisition represents an unsatisfactory solution to the developmental "paradox". Clahsen's approach is to assume a series of discrete stages in the acquisition of inflection, with agreement being either all there, or completely absent. This runs into problems once the data are closely examined. What we actually see is a gradual increase in
the accuracy and use of agreement morphology. This suggests that the child is not making a series of discrete hypotheses regarding the nature of the categories being encoded, but is instead learning the intricacies of the morphological paradigm.

According to the approach argue for here, German children underuse 3rd Sing. [t] precisely because it is not yet a productive agreement marker in the stages under examination. However, what still remains to be explained is why 3rd Sing. [t] is initially used predominantly with intransitive verbs. I argue here that this is simply the result of discourse strategies in parental input. It appears to be a discourse property of (German) parental speech that 3rd Person subjects (with 3rd Person agreement on the verb) are used mostly in intransitive contexts. This makes sense from the point of view of Hopper & Thompson (1980), who noted a correlation between clausal transitivity and discourse strategies. Foregrounded material is highly transitive, while backgrounded material tends towards low transitivity. Sentences with 3rd Person subjects (which are often inanimate), are typically part of the background of discourse, while foregrounded discourse has a high percentage of 1st and 2nd Person subjects. A preliminary examination of portions of the transcripts of Simone and Meike, from Miller (1976:270-460), largely confirms this observation.
Table 14. Verbs used at least three times with 3rd Sing. [t] in PLD of Simone and Meike

<table>
<thead>
<tr>
<th>Intransitive Verbs</th>
<th>Tokens</th>
<th>Transitive Verbs</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. schlafen</td>
<td>(22)</td>
<td>1. anziehen</td>
<td>(9)</td>
</tr>
<tr>
<td>2. liegen</td>
<td>(10)</td>
<td>2. heissen</td>
<td>(5)</td>
</tr>
<tr>
<td>3. sagen</td>
<td>(9)</td>
<td>3. holen</td>
<td>(5)</td>
</tr>
<tr>
<td>4. sitzen</td>
<td>(9)</td>
<td>4. waschen</td>
<td>(5)</td>
</tr>
<tr>
<td>5. weinen</td>
<td>(6)</td>
<td>5. geben</td>
<td>(3)</td>
</tr>
<tr>
<td>6. fahren</td>
<td>(4)</td>
<td>6. suchen</td>
<td>(3)</td>
</tr>
<tr>
<td>7. stehen</td>
<td>(4)</td>
<td>7. trinken</td>
<td>(3)</td>
</tr>
<tr>
<td>8. bleiben</td>
<td>(3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. meinen</td>
<td>(3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 turnen</td>
<td>(3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total:</td>
<td>(73)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pro-verbs</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. machen</td>
<td>(26)</td>
</tr>
<tr>
<td>2. haben</td>
<td>(17)</td>
</tr>
<tr>
<td>3. gehen</td>
<td>(13)</td>
</tr>
<tr>
<td>4. kommen</td>
<td>(13)</td>
</tr>
<tr>
<td>Total:</td>
<td>(69)</td>
</tr>
</tbody>
</table>

Table 14 shows that the occurrences of verb types and tokens which appear with 3rd Person [t] are not equally distributed between transitive and intransitive verbs. There are 10 verb types with a total of 73 tokens which appear under the Intransitive heading, while there are 7 verb types with 33 token in the Transitive column. The most frequent verbs in the input are the four “pro-verbs”, which occurred in Table 12. Two of these verbs, machen (‘to make’ or ‘to do’) and haben (‘to have’), are transitive, while the other two, kommen (‘to come’) and gehen (‘to go’), are intransitive. Since these “pro-verbs”, including the bivalent ones, tended to appear with and without 3rd Sing. [t] in any case, they constitute prima facie evidence against Clahsen’s proposal. That is, it is precisely those verbs which are semantically most basic, and which occur most frequently in the input, which are first used in a variety of inflected forms, irrespective of their number of arguments.
Although this specific correlation between subject agreement and transitivity has not been noted in the child language literature previous to Clahsen (1986), other correlations between verbal morphology and verb semantics have been noted at least since the 1970s. L. Bloom et al. (1980) performed an aspectual analysis on verb types produced by English speaking children. They studied the emergence of regular and irregular past-tense marking (ed/IRREG), 3rd Person Sing. (-s), and participial (-ing). They found that past-tense marking initially occurred predominantly on verbs which denoted non-durative, momentary events tending to have a clear end result (e.g. 'break'). Action verbs with (-ing) denoted durative events extending over time, usually with no clear end-result (e.g. 'play'). The inflectional marker (-s) was used primarily with verbs such as go, as in 'this goes here' (Bloom et al. 1980:398), which they analyzed as change-of-state verbs with a continuing and salient end result. As in the German examples in Table 12, there were a few verbs which appeared in more than one inflected variant, primarily restricted at first to general-purpose "pro-verbs" such as go, do, make, and get.

It is probable that the speech of these children is simply reflecting the statistical distribution of the input to which they are exposed. Although I know of no current study on this matter, I expect for it to be fairly uncommon for a child to hear a verb such as break in a sentence like 'My toy is breaking right now'. It seems more probable that the child will hear and use sentences of the sort 'You broke my toy', and 'Don't break my toy (!)'. The same argument holds, mutatis mutandis, for the other English verb inflections. The normal statistical pattern of the use of these inflections in adult-to-adult speech is probably even more evident in child-adult discourse situations, which tends to be anchored firmly in the "here-and-now" (Snow 1977). This explains why the use of (s), (ed/IRREG), and (ing) in child speech is predominantly on verbs with which they are aspectually most compatible.
4.4. Conclusions: The Word-Specific Paradigm Stage

I now summarize the results of this chapter, and relate them to the overall goal of this thesis. In Section 4.3.1, I showed that Root Infinitives, or matrix declaratives with an infinitival main verb and no modal or auxiliary, can be seen as the result of Modal Drop. This claim is supported by evidence from Ingram & Thompson (1995), where we found a high correlation between Root Infinitives and modal interpretations. This account preserves the assumption of continuity in syntax, contra Wexler (1994). A maturational account of Root Infinitives neither explains why children use them primarily in modal contexts, nor does it explain why they disappear when overt auxiliaries and modals start increasing in frequency (Jordens 1990). The Modal Drop Hypothesis (perhaps more properly the Auxiliary Drop Hypothesis; cf. Hoekstra & Jordens 1994) can account for these facts without abandoning continuity, and thus, as per the arguments in Section 3.1, represents a preferable alternative on conceptual grounds alone.

In Section 4.3.2, I showed that German children do not initially have productive mastery of the morphological paradigm. They require a relatively long timespan in order to properly segment affix from stem. This is true even if they have correctly hypothesized that Person/Number agreement is relevant to the morphosyntax of their target language. The data presented in this section support the claim that mapping between inflectional meaning and inflectional form provides a challenge to language learners. This is indicated by the gradual increase in the accuracy and use of inflectional markers.

Segmentation of affix from stem is performed most efficiently via a comparison of related word forms, which naturally takes place among the cells of a paradigm (Section 4.3.3). We predict that language learners will make errors in their use of inflections prior to the acquisition of the full adult paradigm. I have tentatively claimed that the syncretism between Imperative and 1st Person forms causes children learning German to set-up word-specific paradigms lacking the feature [+1st]. This leads to overextension of zero-marked stems to 3rd Person contexts, until either a word-specific 3rd Sing. entry is added, or the
3rd Sing. suffix [t] is segmented. Once this occurs, we predict a sharp decrease in the use of zero-marked forms, since the correct 3rd Sing. affix is available for any word of the category ‘Verb’. Data from Mathias and Daniel in Table 13 support this conclusion.

In Section 4.3.4, we found that Clahsen’s hypothesis on the nature of early German agreement inflections is unsatisfactory, since it does not account for either the gradual improvement in the use of morphology, nor for the status and function of inflectional markers beside 3rd Sing. [t], which he claims marks intransitivity. Furthermore, the correlations he notes between agreement marking and transitivity are not exact, but rather probabilistic in nature (cf. Verrips & Weissenborn 1992). A more adequate account shows that the notion of agreement is present in German child language, but that inflectional markers are not being productively used at the stage in question. Inflected verbs are stored as wholes in paradigm slots, where they will undergo morphological analysis. The initial correlation between intransitive verbs and 3rd Person agreement is accounted for by looking at the discourse properties of the parental input, which is skewed in this direction, as shown in Table 14.

The general conclusion of this chapter is that the Continuity Hypothesis and Morpholexical Learning are able to account adequately for the empirical data on the acquisition of agreement morphology in German. These data show the existence of an initial stage in which learning is “suppletive”, before general morphological rules and patterns have been extracted. Paradigms have played an important role in this account, by explaining the distributional pattern shown by these children in their use of inflectional markers. This adds to the conceptual support for paradigms provided in Chapter 2.
5. Inflection Class Mixture and Acquisition

In Chapter 4, I provided evidence for the existence of a word-specific paradigm stage in German child language. During this stage, children are essentially building a lexicon "word-by-word", constructing paradigms for each group of morphosyntactically related lexical entries. In subsequent stages, the child must be able to extract productive rules for inflectional marking, in order to correctly produce (and interpret) new words as they are acquired. In the conception of Morpholexical Learning advocated here, this means that the child extracts general rules from stored lexical items. The general rules which are extracted are based upon the paradigmatic information which has been constructed by the child up to this point.

In this chapter, I go beyond the acquisitional data in Chapter 4, in order to support the existence of a "General Paradigm Stage". I take up from where we left off in Section 2.2.2, where it was shown that paradigm mixture in the Latin 3rd declension is not random, but is the product of a series of Paradigm Structure Conditions. I propose an analysis whereby these PSCs constitute implicational lexical redundancy rules, abstracted from word-specific paradigm matrices during the course of language acquisition. These rules are constructed according to a semantic (or markedness) hierarchy of inflectional categories. Support for this hierarchy comes from the pattern of syncretism shown in Latin nominal paradigms: a single arrangement of Case/Number forms puts every syncretic form adjacent to one another in all 11 declensions (Section 5.1).

The full adult German verbal paradigm is examined in Section 5.2. German verbs show a pattern of paradigm mixture which is in important respects identical to that shown by the Latin 3rd declension. Here too, a semantic hierarchy of verbal inflectional categories produces a structure which accounts precisely for the manner in which the inflectional markers have "mixed" together. This fact is significant, because we are now dealing with

---

32 In this chapter, I rely heavily upon the data presented in Carstairs-McCarthy (1991).
paradigms of a different morphosyntactic category (Verbs) in a different, not closely related language (German).

I conclude the chapter in Section 5.3, where I show that the same hierarchy of Case/Number forms proposed for Latin is able to account for the mixed paradigms of Icelandic monosyllabic feminine nouns. This set of declensions possesses three competing inflectional markers for one morphosyntactic category, and in another only two. This creates a different set of possibilities for paradigm mixture; however, the pattern which is actually attested conforms to the sorts of structures proposed for Latin and German. All three instances of paradigm mixture (Latin and Icelandic nouns, German verbs) demonstrate that there are innate principles governing inflectional systems, providing support for the main proposal of this thesis: paradigmatic structure in the lexicon plays an important role in how we acquire and use language.

5.1. Latin 3rd declension

The Latin 'mixed' 3rd declension consists of five paradigms which arose when the 'i-declension' and the 'consonantal-declension' fell together (see Section 2.2.2). These declensions did not mix randomly, but combined according to a specific pattern, which Wurzel (1987) represents with a set of Paradigm Structure Conditions. Repeated here for convenience are PSCs (9a-c) and Figure 3.

(9)  a. [im/Acc.Sing.⇒[i:/Abl.Sing.⇒[i:s/Acc.Pl.⇒[ium/Gen.Pl.]
   b. [um/Gen.Pl.⇒[e:s/Acc.Pl.⇒[e/Abl.Sing.⇒[em/Acc.Sing.]
   c. [is/Gen.Sing.⇒...⇒[ibus/Dat. and Abl.Pl.]
All three of the PSCs in (9) are captured in Figure 3. The basic insight behind this structure is that PSC (9b) (which describes the relationships among the regular Type E inflections) is schematically the exact reverse of PSC (9a) (see Figure 2), while each irregular paradigm is represented by a different “point of entry” into PSC (9a). That (9a) and (9b) are related to each other in this fashion remains unexplained in Wurzel (1987). Given the structure in Figure 3, it is a consequence of the hierarchy of inflectional markers, as well as the fact that the structure “branches downward” at four points. The PSCs in (9) are therefore simply post-hoc statements on the distribution of the inflectional markers of the Latin 3rd declension, with no genuine linguistic status. We are able to make these post-hoc statements precisely because inflectional class mixture is constrained according to the principles responsible for the structure in Figure 3 (see below). In other words, PSCs (9a-c) are epiphenomenal.

The basis for the hierarchy in Figure 3 still requires explanation. I suggested in Section 2.2.2 that it is the result of a semantic or markedness hierarchy. All of the Singular forms precede all Plural forms, in accordance with assumptions of markedness theory.
(Jakobson 1939; Bybee & Brewer 1980). However, we still need to account for the order of Case/Number forms in Figure 3. To this end, it will be useful to look at the work of Plank (1991a,b), who asks the question of how we should order the Case/Number forms of nominal declensions. In particular, he examines the hypothesis that syncretisms may occur only in adjacent terms (cf. also McCreight & Chvany 1991; Postma 1993). Plank shows that the following arrangement of Case/Number forms of Latin nouns accounts for every syncretism but one: that of Voc./Nom. Sing. and Gen. Sing. (both realized as [is]) in one subtype (IIIbM/F) of the 3rd declension.

### Table 15a. Classical Latin Masculine and Feminine noun declensions (Plank 1991a: Table 3)

<table>
<thead>
<tr>
<th>CASE</th>
<th>I_E</th>
<th>I_M</th>
<th>IV_M/F</th>
<th>IIIb_M/F</th>
<th>III_a</th>
<th>IIIa_M/F</th>
<th>V_E/M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sg.</td>
<td>Voc.</td>
<td>capra</td>
<td>lupa</td>
<td>manus</td>
<td>ignis</td>
<td>urbs</td>
<td>re:gs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>die:s</td>
</tr>
<tr>
<td>Nom.</td>
<td>capra</td>
<td>lupa</td>
<td>manus</td>
<td>ignis</td>
<td>urbs</td>
<td>re:gs</td>
<td>die:s</td>
</tr>
<tr>
<td>Acc.</td>
<td>capram</td>
<td>lupum</td>
<td>manum</td>
<td>ignem</td>
<td>urbem</td>
<td>re:gem</td>
<td>diem</td>
</tr>
<tr>
<td>Abl.</td>
<td>capra:</td>
<td>lupo:</td>
<td>manu:</td>
<td>igni:</td>
<td>urbe</td>
<td>re:ge</td>
<td>die:</td>
</tr>
<tr>
<td>Dat.</td>
<td>caprae</td>
<td>lup:</td>
<td>manui:</td>
<td>igni:</td>
<td>urbi:</td>
<td>re:gi:</td>
<td>diei</td>
</tr>
<tr>
<td>Gen.</td>
<td>caprae</td>
<td>lupi:</td>
<td>manus:</td>
<td>ignis</td>
<td>urbis</td>
<td>re:gis</td>
<td>diei</td>
</tr>
<tr>
<td>Pl.</td>
<td>Voc.</td>
<td>caprae</td>
<td>lupi:</td>
<td>manus:</td>
<td>ignis</td>
<td>urbe:s</td>
<td>re:ges</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>die:s</td>
</tr>
<tr>
<td>Nom.</td>
<td>caprae</td>
<td>lupi:</td>
<td>manus:</td>
<td>ignis</td>
<td>urbe:s</td>
<td>re:ges</td>
<td>die:s</td>
</tr>
<tr>
<td>Acc.</td>
<td>capras:</td>
<td>lupos:</td>
<td>manus:</td>
<td>ignis</td>
<td>urbis</td>
<td>re:ges</td>
<td>die:s</td>
</tr>
<tr>
<td>Abl.</td>
<td>capris:s</td>
<td>lupis:</td>
<td>manibus</td>
<td>ignibus</td>
<td>urbis</td>
<td>re:gis</td>
<td>die:bus</td>
</tr>
<tr>
<td>Dat.</td>
<td>capris:s</td>
<td>lupis:</td>
<td>manibus</td>
<td>ignibus</td>
<td>urbis</td>
<td>re:gis</td>
<td>die:bus</td>
</tr>
<tr>
<td>Gen.</td>
<td>caprarum</td>
<td>luporum</td>
<td>manuum</td>
<td>ignium</td>
<td>urbi</td>
<td>re:gum</td>
<td>die:rum</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>gloss:</td>
<td>'goat'</td>
<td>'wolf'</td>
<td>'hand'</td>
<td>'fire'</td>
<td>'town'</td>
<td>'king'</td>
</tr>
</tbody>
</table>

### Table 15b. Classical Latin Neuter noun declensions (Plank 1991a: Table 3)

<table>
<thead>
<tr>
<th>CASE</th>
<th>II_N</th>
<th>IV_N</th>
<th>IIIb_N</th>
<th>IIIa_N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sg.</td>
<td>Voc.</td>
<td>bellum</td>
<td>cornu:</td>
<td>mare</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nom.</td>
<td>bellum</td>
<td>cornu:</td>
<td>mare</td>
<td>no:men</td>
</tr>
<tr>
<td>Acc.</td>
<td>bellum</td>
<td>cornu:</td>
<td>mare</td>
<td>no:men</td>
</tr>
<tr>
<td>Abl.</td>
<td>bello:</td>
<td>cornu:</td>
<td>mari:</td>
<td>no:mine</td>
</tr>
<tr>
<td>Dat.</td>
<td>bello:</td>
<td>cornu:</td>
<td>mari:</td>
<td>no:mine</td>
</tr>
<tr>
<td>Gen.</td>
<td>bell:</td>
<td>cornu:s</td>
<td>maris</td>
<td>no:mini:s</td>
</tr>
<tr>
<td>Pl.</td>
<td>Voc.</td>
<td>bella</td>
<td>cornua</td>
<td>maria</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nom.</td>
<td>bella</td>
<td>cornua</td>
<td>maria</td>
<td>no:mina</td>
</tr>
<tr>
<td>Acc.</td>
<td>bella</td>
<td>cornua</td>
<td>maria</td>
<td>no:mina</td>
</tr>
<tr>
<td>Abl.</td>
<td>bellis</td>
<td>cornibus</td>
<td>maribus</td>
<td>no:minibus</td>
</tr>
<tr>
<td>Dat.</td>
<td>bellis</td>
<td>cornibus</td>
<td>maribus</td>
<td>no:minibus</td>
</tr>
<tr>
<td>Gen.</td>
<td>bello:rum</td>
<td>cornu:rum</td>
<td>marium</td>
<td>no:minum</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>gloss:</td>
<td>'war'</td>
<td>'horn'</td>
<td>'sea'</td>
</tr>
</tbody>
</table>
In Table 16, I abstract from Table 15 the inflectional markers which mark the Case/Number combinations of Latin. These markers are placed in "boxes" which show the amount of paradigmatic "space" they occupy (cf. McCreight & Chvany 1991). Markers which range over more than one cell vertically represent syncretisms. Markers which range over more than one cell horizontally are those which inflectional classes share with each other.33

Table 16a. Classical Latin Masculine and Feminine noun declensions: pattern of syncretism

<table>
<thead>
<tr>
<th>CASE</th>
<th>I_M</th>
<th>II_M</th>
<th>III_M/F</th>
<th>IIIb_M/F</th>
<th>IV_M/F</th>
<th>V_M/F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sg.</td>
<td>a</td>
<td>e</td>
<td>us</td>
<td>is</td>
<td>s</td>
<td>e:s</td>
</tr>
<tr>
<td>Nom.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acc.</td>
<td>am</td>
<td>um</td>
<td>em</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abl.</td>
<td>a:</td>
<td>o:</td>
<td>u:</td>
<td>i: (e)</td>
<td>e</td>
<td></td>
</tr>
<tr>
<td>Dat.</td>
<td>æ</td>
<td>i:</td>
<td>u:s</td>
<td></td>
<td>is</td>
<td>i</td>
</tr>
<tr>
<td>Gen.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>e:s</td>
<td></td>
</tr>
<tr>
<td>Pl.</td>
<td>a:s</td>
<td>o:s</td>
<td>i:s</td>
<td>(&gt;e:s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nom.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abl.</td>
<td>i:s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>e:bus</td>
</tr>
<tr>
<td>Dat.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gen.</td>
<td>a:rum</td>
<td>o:rum</td>
<td>uum</td>
<td>ium</td>
<td>um</td>
<td>e:rum</td>
</tr>
</tbody>
</table>

Table 16b. Classical Latin Neuter noun declensions: pattern of syncretism

<table>
<thead>
<tr>
<th>CASE</th>
<th>IV_N</th>
<th>IIIb_N</th>
<th>IIIa_N</th>
<th>II_N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sg.</td>
<td>u:</td>
<td>e</td>
<td>ð</td>
<td>um</td>
</tr>
<tr>
<td>Nom.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acc.</td>
<td></td>
<td>i:</td>
<td>e</td>
<td>o:</td>
</tr>
<tr>
<td>Abl.</td>
<td></td>
<td>is</td>
<td>i:</td>
<td></td>
</tr>
<tr>
<td>Dat.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gen.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pl.</td>
<td>ua</td>
<td>iα</td>
<td>a</td>
<td></td>
</tr>
<tr>
<td>Voc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nom.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abl.</td>
<td></td>
<td>ibus</td>
<td>i:s</td>
<td></td>
</tr>
<tr>
<td>Dat.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gen.</td>
<td>uum</td>
<td>ium</td>
<td>um</td>
<td>o:rum</td>
</tr>
</tbody>
</table>

33 Tables 19a and 19b show that the total number of inflectional markers is much smaller than the maximum possible. For example, Table 15a has 7 inflectional classes with 12 Case-Number combinations, meaning that there are 84 cells which must be filled. In Table 16a, we see that the paradigmatic space is divided into only 33 regions, with a total of 29 different inflectional markers (four of which fill two regions apiece).
With respect to inflectional syncretisms, it is useful to distinguish between those which are systematic, and those which are accidental (Carstairs 1987). Carstairs (1987) classifies the one syncretism not handled by this arrangement of Latin cases as "probably accidental", since Voc./Nom. Sing. and Gen. Sing. are distinguished from each other in every other paradigm. That is, there is no general pattern of syncretism beyond these two markers, in contrast to (for example) Dative/Ablative Pl. In this latter case, the two morphosyntactic categories are always realized identically, regardless of the particular markers used (see Table 16). Thus Dative Pl. = Ablative Pl. is clearly a systematic syncretism in Latin nominal inflection.

If it can be maintained that the syncretism of Voc./Nom. Sing. and Gen. Sing. in declension IUbM is accidental, then it is remarkable that one single arrangement of Cases and Numbers accounts for all of the syncretisms in all 11 different declensions of Latin nouns. It is far from given that this should be true, since each class of twelve Case-Number forms (ten plus the addition of Voc. Sing. and Voc. Pl.) can be arranged into 12! = 479,001,600 different linear orders. In Section 5.3, we will see that an essentially identical arrangement accounts for the pattern of syncretism demonstrated by Icelandic monosyllabic feminine nouns.34

This arrangement of Cases constitutes a plausible hierarchy of markedness in the arrangement of Cases and Numbers. Tables 15 and 16 show all Singular cases preceding all Plural cases. Remarkably, both Singular and Plural case forms are arranged identically, with Vocative < Nominative < Accusative < Ablative < Dative < Genitive. The "core" syntactic cases of Nominative and Accusative (Blake 1994) precede the Ablative, Dative,

34 Allen & Brink (1980) note that the Indian grammarians proposed the following arrangement for Sanskrit (Vocative was omitted due to the fact that they did not consider it to be a true case):


"The basis of this order is a purely practical one: given the Nominative as first, it is the only order to bring together all cases which in some numbers or declensions are 'syncretized', i.e. have identical forms..." (Allen & Brinks 1980:62).
and Genitive cases. In any event, the placement of Nominative before Accusative, and both of these before Ablative, Dative, and Genitive accords with most typological hierarchies. It is encouraging that this arrangement of Cases and Numbers, derived purely through the requirement that syncretic forms be adjacent, has some plausibility for encoding markedness relations. These facts suggest that we need to amend the structure in Figure 3 so that the inflectional markers are arranged as in Tables 15 and 16.

Figure 4.

Type E = right column

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[i:] [e] Abl.</td>
</tr>
<tr>
<td></td>
<td>[i:] [i:] Dat.</td>
</tr>
<tr>
<td></td>
<td>[is] [is] Gen.</td>
</tr>
<tr>
<td></td>
<td>[e:s] C [e:s] Plural Voc./Nom.</td>
</tr>
<tr>
<td></td>
<td>[i:s] [e:s] Acc.</td>
</tr>
<tr>
<td></td>
<td>[ium] [um] Gen.</td>
</tr>
</tbody>
</table>

The new arrangement preserves most of the structure of Figure 3; importantly, the contents of PSCs (9a) and (9b) are still expressed. Converting Figure 4 into Wurzel’s implicational

35 It is possible to account for all of the same syncretisms through a slightly different arrangement of Sing. and Pl. Cases:


The inflectional hierarchy expressed here by the Plural Case forms is exactly the typological hierarchy proposed by Blake (1994:157ff). I do not have any suggestions as to why the arrangement of Case forms might be different according to Number, nor why the Case forms in the Sing. do not correspond to this typological ordering.

92
rule format, we replace PSCs (9a), (9b), and (9c) with those in (54a) and (54b). In (54a), I have taken advantage of one of Wurzel’s notational devices, using parentheses to abbreviate several rules into a single one. How this works is shown by the schematic rule in (53a), which is an abbreviation for the three in (53b) (see Carstairs-McCarthy 1991:217 for a description of this aspect of Wurzel’s framework):

(53) a. A(⇒B)⇒C
    b. A⇒B
       B⇒C
       A⇒C

(54) a. [(i)s/Voc.Nom.Sg.]
    (⇒[em/Acc.Sg.])
    (⇒[e/Abl.Sg.])
    ⇒[i:/Dat.Sg.]
    ⇒[is/Gen.Sg.]
    ⇒[e:s/Voc.Nom.Pl.]
    (⇒[e:s/Acc.Pl.])
    ⇒[ibus/Abl.Dat.Pl.]
    ⇒[um/Gen.Pl.]

b. [im/Acc.Sing.]
    ⇒[i:/Abl.Sing.]
    ⇒[i:s/Acc.Pl.]
    ⇒[ium/Gen.Pl.]

This notation is a convenient way to indicate the default status of the productive, or dominant PSC. The parenthesized portions of (54a) are over-ridden by the irregular PSC in (54b). This division between regular and irregular is directly encoded in the structure in Figure 4 (see below).

The semantic or markedness hierarchy proposed here constitutes an answer to one of the concerns raised in Section 2.2.2. This concern was that there must be some sort of constraint on the types of PSCs chosen from those which are compatible with a given paradigm. The other concern which I raised there was how to constrain the number of morphological specificatons which can be assigned to a single lexical entry. As noted
above, Wurzel claims that Types B, C, and D of the Latin 3rd declension are lexically specified for two inflectional categories, while under my approach, each class requires at most one. The specifications which our competing analyses require are repeated here.

(11) Wurzel's set of specifications for the Latin 3rd declension

<table>
<thead>
<tr>
<th>Class Type</th>
<th>Specifications in Lexical Entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>[im/Acc.Sing.]</td>
</tr>
<tr>
<td>B</td>
<td>[i:/Abl.Sing.] and [em/Acc.Sing.]</td>
</tr>
<tr>
<td>C</td>
<td>[i:s/Acc.Pl.] and [e/Abl.Sing.]</td>
</tr>
<tr>
<td>D</td>
<td>[ium/Gen.Pl.] and [e:s/Acc.Pl.]</td>
</tr>
<tr>
<td>E</td>
<td>[um/Gen.Pl.]</td>
</tr>
</tbody>
</table>

(14) Proposed set of specifications for the Latin 3rd declension

<table>
<thead>
<tr>
<th>Class Type</th>
<th>Specifications in Lexical Entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>[im/Acc.Sing.]</td>
</tr>
<tr>
<td>B</td>
<td>[i:/Abl.Sing.]</td>
</tr>
<tr>
<td>C</td>
<td>[i:s/Acc.Pl.]</td>
</tr>
<tr>
<td>D</td>
<td>[ium/Gen.Pl.]</td>
</tr>
<tr>
<td>E</td>
<td>none</td>
</tr>
</tbody>
</table>

The reason that only a single specification per inflectional class is required in (14) is that Figure 4 crucially differentiates between regular and irregular inflectional patterns. All links between inflections which require no morphological specifications are represented with solid arrows, while paths which require such information are represented with dashed arrows. The productive markers (the right hand column) have a default status: in the absence of information to the contrary, they are automatically chosen as the realizations for a particular combination of Number and Case.

It will be instructive to see how inflectional classes requiring more than one morphological specification would be represented. Such classes would involve additional dashed arrows between inflectional markers, creating a "zig-zag" pattern which is not
attested by the Latin data. This is because each column represents a rule schema which is followed unless specific directions are provided to the contrary. Figure 5 is a depiction of the largest number of morphologically-specified connections which are possible for the Latin 3rd declension, given that there are four categories which have two competing markers apiece (only the relevant ones are shown). In Table 17 all of the inflectional classes which Figure 5 represents are enumerated. Also provided is the number of morphological specifications required for each class, along with an indication of which of these sixteen actually appeared in Latin.

Figure 5.

Table 17. Possible ways for the Latin 3rd declension to have mixed

| Acc. Sg. | em | em | em | em | em | em | em | im | im | im | im | im | im | im | im |
|----------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Abl. Sg. | e  | e  | e  | e  | i: | i: | i: | i: | e  | e  | e  | i: | i: | i: | i: |
| Acc. Pl. | e:s| e:s| i:s| i:s| e:s| e:s| i:s| i:s| e:s| e:s| i:s| i:s| es| es| i:s|
| Gen. Pl. | um | um | um | um | ium| ium| ium| ium| ium| ium| ium| ium| ium| ium|
| Attested?| yes| yes| no | yes| no | no | no | yes| no | no | no | no | no | no | yes|
| Type:   | E  | D  | *  | C  | *  | *  | *  | B  | *  | *  | *  | *  | *  | *  | *  |
| Number of Specifications: | 0  | 1  | 2  | 1  | 2  | 3  | 2  | 1  | 2  | 3  | 4  | 3  | 2  | 3  | 1  |

As shown in Table 17, it is all and only those inflectional classes which have no more than one morphological specification which are actually realized in Latin. This is precisely the state of affairs represented by the structure in Figure 4. Let us elevate this observation to the level of a principle, restricting the number of irregular morphological specifications per lexical entry.
(55) Restriction on Morphological Specifications

Lexical entries may contain at most one morphological specification; furthermore, this specification must be for an irregular, not a regular marker.

Notice that this restriction immediately rules out a large number of possibilities in how mixed inflectional classes may be structured. In particular, (55) rules out structures such as that in Figure 6, on the assumption that Type A is the productive class. Here, arbitrary inflectional exponents are represented by lower case letters, while arbitrary morphosyntactic categories are represented by arabic numerals. Upper case letters are used to represent inflectional classes, just as above.

Table 18. Hypothetical example of paradigm mixture

<table>
<thead>
<tr>
<th>Class</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>2</td>
<td>b</td>
<td>p</td>
<td>p</td>
<td>p</td>
<td>p</td>
</tr>
<tr>
<td>3</td>
<td>c</td>
<td>c</td>
<td>q</td>
<td>q</td>
<td>q</td>
</tr>
<tr>
<td>4</td>
<td>d</td>
<td>d</td>
<td>d</td>
<td>r</td>
<td>r</td>
</tr>
<tr>
<td>5</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>e</td>
<td>s</td>
</tr>
<tr>
<td>6</td>
<td>f</td>
<td>f</td>
<td>f</td>
<td>f</td>
<td>f</td>
</tr>
</tbody>
</table>
With Type A as the productive paradigm, Types B, C, and D require two morphological specifications apiece, in violation of (55). Alternatively, let us assume that Type E, not Type A, is the productive paradigm. With this assumption, Figure 6 (and the inflectional classes in Table 18) are perfectly legitimate according to (55), since Types A, B, C, and D require only one specification, and Type E none.

The restriction in (55) is based on scant evidence so far. Yet if true, it would account for the fact that the Latin 'i-declension' and the 'consonantal-declension' combined in only four of the sixteen possible ways. Moreover, this restriction would predict precisely those four declensions which actually did appear, given the hierarchical structure and division between productive and irregular classes which are crucial to my account.

This analysis of mixed inflectional classes has implications for Carstairs’ Paradigm Economy Principle, which limits the number of inflectional classes within a language (Carstairs 1987). The Paradigm Economy Principle states that “...there can be no more inflectional paradigms for any word-class in any language than there are distinct “rival” inflectional realisations available for that morphosyntactic property-combination where the largest number of rivals compete” (Carstairs-McCarthy 1991:222). The Latin 3rd declension constitutes a prima facie violation of this principle, since there are at most two inflectional realizations per morphosyntactic property-combination, and yet the 3rd declension has five distinct paradigms. This provides evidence for the analysis proposed here, and evidence against the Paradigm Economy Principle, because five (rather than two) is exactly the number of inflectional classes which my account predicts as the maximum.

5.2. The “mixed” classes of German Verbs

German verbs inflect for Tense, Mood, Person, and Number. There are two main conjugations, one ‘strong’ and one ‘weak’. Weak verbs constitute the regular, productive class. Strong verbs constitute an irregular class, but one containing numerous, highly frequent members. Of relevance here is the existence of two “mixed” conjugations,
consisting of verbs which have selected a portion of their inflectional markers from the weak class, and another proportion from the strong class. Table 19, taken from Carstairs-McCarthy (1991:Table 6), shows the three verbal morphological categories with competing inflectional markers in German.

Table 19. Affixal Classes of German verbs
(Carstairs-McCarthy 1991:Table 6)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preterite Indic. (1st/3rd Sing.)</td>
<td></td>
<td>[e]</td>
<td>[e]</td>
<td>[e]</td>
</tr>
<tr>
<td>Preterite Subj. (1st/3rd Sing.)</td>
<td>[e]</td>
<td>[e]</td>
<td>[e]</td>
<td>[e]</td>
</tr>
<tr>
<td>Past Participle</td>
<td>[en]</td>
<td>[en]</td>
<td>[en]</td>
<td>[t]</td>
</tr>
</tbody>
</table>

Table 19 excludes the class of modal verbs. The class membership is as follows, taken directly from Carstairs-McCarthy (1991:245).

Type A: All prototypical strong verbs, i.e. verbs with no affix in the 1st and 3rd Sing. Pret.
Type B: Three verbs, dingen 'hire', schinden 'flay', and schleifen 'split', which are now usually weak in the Present and Preterite Indicative (e.g. schindete) but still retain a strong Preterite Subjunctive with i or ü as the (umlauted) ablaut vowel (dinge, schünde, schlisse) as well as a Perfective Participle in -en (gedungen, geschunden, geschlissen).
Type C: A few erstwhile strong verbs which have become weak except for retaining a Perfective Participle in -en: mahlen 'grind' (gemahlen), melken 'milk' (gemolken), etc.
Type D: All weak verbs (the unmarked or default class) and the denken group (with a distinct stem but weak affixal inflection in the Preterite and Perfective Participle).

Type D, which represents the productive class of 'weak' verbs, contains by far the greatest number of members. Types A, B, and C represent classes which are irregular, with Type A containing the bulk of those verbs which are not members of Type D. Types B and C are "mixtures" of Types A and D. This is demonstrated in Figure 7. All German verbal inflectional markers other than those appearing in Figure 7 can be predicted on the basis of the ones appearing there.
Note the strong similarity between this structure and that for Latin nouns in Figure 4. That this similarity exists is in itself remarkable, since we are now looking at a different language, and also because we are now dealing with verbal conjugations, in contrast to nominal declensions. Note also that, once again, syncretic forms are for the most part adjacent to one another (with the only exception being the “weak” Pres. Subj. 1st/3rd Sing. [e] and “strong” Pret. Subj. 1st/3rd Sing. [e], both phonetically schwa). The ordering of categories suggests a hierarchical ranking of Tense > Mood > Person/Number, and [+finite] > [-finite]. This is exactly the hierarchy of verbal categories proposed in Bybee (1985).

Given that there are three morphosyntactic category combinations with two competing realizations apiece, there are eight possible ways for the strong and the weak conjugations to have mixed together. This is shown in Figure 8, and again in Table 20.

---

36 This case is not so clear, however, since I have not included 2nd Sing. or any plural forms in Figure 7.
Table 20. Possible ways for the German ‘Strong’ and ‘Weak’
conjugations to have mixed

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>te</td>
<td>te</td>
<td>te</td>
<td>yes</td>
<td>A</td>
<td>0</td>
</tr>
<tr>
<td>te</td>
<td>te</td>
<td>en</td>
<td>yes</td>
<td>B</td>
<td>1</td>
</tr>
<tr>
<td>te</td>
<td>te</td>
<td>en</td>
<td>no</td>
<td>*</td>
<td>2</td>
</tr>
<tr>
<td>te</td>
<td>e</td>
<td>en</td>
<td>yes</td>
<td>C</td>
<td>1</td>
</tr>
<tr>
<td>e</td>
<td>te</td>
<td>e</td>
<td>no</td>
<td>*</td>
<td>2</td>
</tr>
<tr>
<td>e</td>
<td>te</td>
<td>e</td>
<td>no</td>
<td>*</td>
<td>3</td>
</tr>
<tr>
<td>e</td>
<td>e</td>
<td>e</td>
<td>yes</td>
<td>D</td>
<td>2</td>
</tr>
</tbody>
</table>

As in the Latin 3rd declension, the maximum number of combinatory possibilities is not attested in German. Table 21 reveals that all and only those conjugations requiring no more than one morphological specification actually occur. This provides additional support for the proposed restriction in (55).

There is another intriguing parallel between the German data in this section and the Latin 3rd declension. In both cases, diachronic change shows movement “down the column”. In Latin, words of the Type A (i-declension) → B → C → D → E (consonantal-declension) (see Chapter 6 for further discussion). In German, it is similarly the case that words of Type A (strong verbs) → B → C → D (weak verbs). In Chapter 6 I suggest that it is the unidirectional nature of the implicational hierarchy which is responsible for this downward movement, resulting in the apparently paradoxical situation that more frequent, relatively unmarked categories will be regularized before less frequent, relatively marked categories in diachronic change.

This downward movement predicts the following situation. As mixed paradigms change over the course of a language’s history, we shall see inflectional classes disappear from them, but no new inflectional classes will appear within them (given that, as in the Latin and German examples, the maximum number of allowable mixed classes already exists). To see what I mean by this, consider the diagram in Figure 9a, 9b, and 9c and Tables 21a, 21b, and 21c. Figure 9a once again represents a hypothetical set of mixed paradigms consisting of arbitrary morphosyntactic category combinations. Assume that the
right column (Type E) represents the productive inflectional class, while the left column represents an irregular class with four irregular markers.

Figure 9

Type E (productive class) = right column in all three diagrams

<table>
<thead>
<tr>
<th></th>
<th>a.</th>
<th>b.</th>
<th>c.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>[a]</td>
<td>[a]</td>
<td>[a]</td>
</tr>
<tr>
<td>B</td>
<td>[b]</td>
<td>[b]</td>
<td>[b]</td>
</tr>
<tr>
<td>C</td>
<td>[c]</td>
<td>[c]</td>
<td>[c]</td>
</tr>
<tr>
<td>D</td>
<td>[d]</td>
<td>[d]</td>
<td>[d]</td>
</tr>
<tr>
<td>E</td>
<td>[e]</td>
<td>[e]</td>
<td>[e]</td>
</tr>
</tbody>
</table>

Original Structure  | Irregular [s] deleted | Irregular [p] deleted

Figures 9b and 9c represent two independent historical scenarios. In Figure 9b, the bottom inflectional marker [s] in the irregular, left column has been regularized completely, and is therefore deleted from the structure. In Figure 9c, the topmost marker [p] has been regularized, while the bottomost marker has been preserved. In Table 21a, I present the make-up of the inflectional classes which Figure 9a entails. This is the starting point from which Figures 9b and 9c are derived. These two (independent) hypothetical historical changes are represented respectively in Tables 21b and 21c.

Table 21a. Original set of paradigms

<table>
<thead>
<tr>
<th>Class</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>2</td>
<td>p</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>3</td>
<td>q</td>
<td>q</td>
<td>c</td>
<td>c</td>
<td>c</td>
</tr>
<tr>
<td>4</td>
<td>r</td>
<td>r</td>
<td>r</td>
<td>r</td>
<td>d</td>
</tr>
<tr>
<td>5</td>
<td>s</td>
<td>s</td>
<td>s</td>
<td>s</td>
<td>e</td>
</tr>
<tr>
<td>6</td>
<td>f</td>
<td>f</td>
<td>f</td>
<td>f</td>
<td>f</td>
</tr>
</tbody>
</table>
In Table 21b, representing the loss of the bottom irregular marker [s], there are now three completely new paradigms, A', B', and C', and with only a single paradigm (Type E) which has been preserved from Table 21a. In contrast, look at Table 21c. Here, there are four paradigms remaining, all of which are also present in Table 21a. The difference consists of the fact that Type A has been eliminated, due to the loss of the topmost irregular marker [p]. The changes which have occurred in the history of the Latin 3rd declension (and which are occurring in German) are analogous to this latter situation, rather than that of Table 21b. It is an interesting question whether or not other mixed paradigms from other languages also obey this “top-down” restriction on historical change.

5.3. Icelandic monosyllabic feminine nouns

To conclude this chapter, I examine one more case of mixed paradigms, this time from Icelandic. Once again, I take the relevant data from Carstairs-McCarthy (1991). Table 22, representing the mixed declensions of monosyllabic feminine nouns, shows that there
are two competing realizations for Gen. Sing., and three competing realizations for Nom./Acc. Plural.

Table 22. Icelandic Monosyllabic Feminine nouns
(Carstairs-McCarthy 1991: Table 1)

<table>
<thead>
<tr>
<th></th>
<th>Type A</th>
<th>Type B</th>
<th>Type C</th>
<th>Type D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sing.</td>
<td>ø</td>
<td>ø</td>
<td>ø</td>
<td>ø</td>
</tr>
<tr>
<td>Gen.</td>
<td>-(u)r</td>
<td>-ar</td>
<td>-ar</td>
<td>-ar</td>
</tr>
<tr>
<td>Dat./Acc.</td>
<td>ø</td>
<td>ø</td>
<td>ø</td>
<td>ø</td>
</tr>
<tr>
<td>Plural</td>
<td>-(u)r</td>
<td>-(u)r</td>
<td>-ar</td>
<td>-ir</td>
</tr>
<tr>
<td>Gen.</td>
<td>-a</td>
<td>-a</td>
<td>-a</td>
<td>-a</td>
</tr>
<tr>
<td>Dat.</td>
<td>-um</td>
<td>-um</td>
<td>-um</td>
<td>-um</td>
</tr>
</tbody>
</table>

In Figure 10, the paradigms from Table 22 are organized into a hierarchical structure. Notice that this time, there are three columns instead of two, due to the fact that Nom/Acc. Plural has three competing realizations.

Figure 10.

Type D = center column

There are several ways in which the structure in Figure 10 further supports the analysis of mixed paradigms presented here. First of all, the order of Case/Number forms is, *modulo* the absence of Vocative and Ablative cases, exactly that shown by the Latin structure in
Figure 4. Namely, the categories are arranged such that Nom. > Acc. > Dat. > Gen., and all singular forms precede all plural forms. Secondly, all syncretic inflectional markers once again appear adjacent to one another. These two facts—the same arrangement of categories, and the adjacency of identical markers—are surely beyond coincidence.

The third way in which the Icelandic example parallels Latin (and German), is shown in Table 22, which enumerates all the possible ways in which the competing inflectional markers could have mixed.

Table 23. Possible ways for Icelandic monosyllabic feminine declensions to have mixed

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ar</td>
<td>yes</td>
<td>D</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>ar</td>
<td>yes</td>
<td>C</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>ar</td>
<td>yes</td>
<td>B</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>ur</td>
<td>yes</td>
<td>A</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>ur</td>
<td>no</td>
<td>*</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>ir</td>
<td>no</td>
<td>*</td>
<td>2</td>
</tr>
</tbody>
</table>

Yet again, all and only those inflectional paradigms which require at most one morphological specification are attested. The restriction on lexical entries in (55) has therefore been supported by data from three languages, and by both nominal and verbal paradigms.

The final parallel which the Icelandic data show to the Latin and German examples, is that historical change has been moving words “down the column”. Carstairs-McCarthy (1991) notes that Type A → Type B, and Types B and C → Type D. This fact is important because of the relationships among the three irregular paradigms. Note that Types A and B constitute a single column, while Type C is separately represented in a second column. There is no tendency for words of Type A to go to Type C; drift is always through Type B. This fact was of central importance to Carstairs-McCarthy’s (1991) analysis of the Icelandic mixed declensions. Viewing this situation in terms of the structure represented in Figure
10, the direction of drift in Icelandic is entirely parallel to the direction of drift in both the Latin 3rd declension and German mixed conjugations.
6. Conclusion: From Word-Specific to General

There have been two main parts to this thesis. The first part deals with paradigms, and how they play a central organizing role for inflectional morphology (Chapters 2 and 5). The second part deals with the acquisition of inflectional morphology in German (Chapters 3 and 4). In this latter part, I have attempted to motivate an acquisitional theory which possesses three characteristics: (1) Continuity of abstract principles of grammar; (2) Lexical Learning, which accounts for linguistic variation; and (3) Morpholexical Learning, which accounts for the sequence of developmental stages in child language. The goal of this concluding chapter is to make the connections between these two parts clear.

The basic idea behind the acquisitional theory outlined here is that we should account for child language development with *exactly those areas of language which must, under any theory, be learned*. The acquisition of the lexicon and morphological systems is one such area. As argued in Chapter 2, this type of learning results in linguistic systems which on the surface appear very different from one another. On the other hand, we should not account for language development with *exactly those areas of language which we know cannot be learned*. Abstract principles of Universal Grammar constitute a domain for which the child has no clear, overt evidence (Chomsky 1986). By satisfying these two criteria, we also satisfy Occam’s Razor, producing a restrictive theory of language acquisition. This forces us to look for explanations of language development which explain, rather than describe the data.

"Morpholexical Learning" is the claim that language acquisition is lexically-based, in that children first acquire lexical items, and then proceed to extract generalizations and patterns from the resulting database. This type of inductive learning has two advantages. First of all, it is robust, since noisy data and early mistakes can easily be overcome (Spencer 1990). The reason for this is that the child starts from the specific and individual, and then moves to the general (rather than the reverse, as implied in some current models of morphology). Second, this sequence of stages accurately reflects what we see in child
language data. The pattern exhibited time and again in child language is the initial, conservative use of linguistic constructions, followed by a stage with productive rules and overregularization errors (cf. MacWhinney 1978; Davis 1987; Pinker 1989; Marcus et al. 1991).

The extraction of regular rules from a lexical database cannot proceed without innate constraints on what the child views as linguistically relevant. It would be useless, for example, for the child to count the number of phonemes in a verb in order to see if this number was correlated to the number of arguments it takes. There must be restrictions on what constitutes a possible morphological generalization. Morpholexical Learning therefore requires a mechanism such as the paradigm in order to work. Without the hypothesis that the lexicon is paradigmatically structured, Morpholexical Learning is reduced to vacuousness: it is useless to claim that learning proceeds by extracting generalizations from stored lexical data, without giving some hint as to how this extraction is to be performed.

This is the point at which the two parts of this thesis are linked together. I have attempted to show that language acquisition proceeds first by the construction of word-specific paradigms (Pinker 1984), and then by the extraction of general, implicational lexical redundancy rules. Evidence for the first stage was taken from the acquisition of German agreement morphology. Evidence for the second stage was taken from examples of paradigm mixture in Latin, German, and Icelandic (Carstairs-McCarthy 1991). To conclude this chapter (and the thesis), I now sketch a brief account of how a child might proceed from the “word-specific paradigm stage” (marked by the conservative acquisition of whole inflected words), to the “general paradigm stage” (marked by the productive use of inflectional affixes).37

This account must answer the question of how the hierarchical structures proposed in Chapter 5 relate to the theory of paradigms as n-dimensional matrices. In the paradigm-as-matrix theory, the various inflected forms of a lexeme are related to each other by virtue

37 Thanks to David Ingram for helpful discussion on these matters.
of the fact that they are stored in the same matrix. This contrasts with the type of morphological relationship encoded by PSCs, which consists of a set of unidirectional implicational relationships among inflectional markers. It is an open question whether or not these two conceptions of paradigmatic structure are compatible, or whether one of them may be dispensed with altogether.

It is my belief that Paradigm Structure Conditions are not adequate in and of themselves as replacements for the notion of paradigms as n-dimensional matrices. The paradigm learning model for inflection outlined in this thesis (basically that of Pinker 1984) crucially involves the initial creation of word-specific matrices, making use of the Unique Entry Principle and Closed-Category Relevance. The Unique Entry Principle prevents more than one inflectional marker from inhabiting the same cell of a paradigm; when the language learner encounters more than one form in competition for the same cell, he or she is forced to "split" the paradigm by adding a new dimension, corresponding to an hypothesized morphological category (Pinker 1984). This new dimension satisfies Closed-Category Relevance, since the newly acquired word form has a different phonological realization than the one which was already present. It is hard to see what the analogues of the Unique Entry Principle and Closed-Category Relevance would be using a "PSC-only" model. An example of a word-specific paradigm matrix, based on one of the sub-paradigms of the Latin 3rd declension, is given in (56).

(56) Word-Specific Paradigm

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocative</td>
<td>ignis</td>
<td>igne:s</td>
</tr>
<tr>
<td>Nominative</td>
<td>ignis</td>
<td>igne:s</td>
</tr>
<tr>
<td>Accusative</td>
<td>ignem</td>
<td>igni:s</td>
</tr>
<tr>
<td>Ablative</td>
<td>igni:</td>
<td>ignibus</td>
</tr>
<tr>
<td>Dative</td>
<td>igni:</td>
<td>ignibus</td>
</tr>
<tr>
<td>Genitive</td>
<td>ignis</td>
<td>ignium</td>
</tr>
</tbody>
</table>
The paradigm in (56) contains complete, unsegmented words in its cells; during this phase of child language, all paradigms are basically alike in being “suppletive”. This contrasts with implicational statements such as those in (54), which require inflectional markers to have been segmented from stems before they can be constructed. At the “word-specific paradigm” stage, then, PSCs do not yet exist.

During this initial “conservative” stage we expect to find errors due to incomplete information; either some of the paradigm cells contain no entry at the time in question, or the paradigm itself has not been subdivided according to the appropriate adult system of inflection. For convenience, we may term such errors as “incomplete paradigm errors”. In Section 4.3.3, I claimed that the over-use of zero-marked forms in German child language is an instance of this type of error. As time wears on, incomplete paradigm errors tend to decrease, first as word-specific paradigms are filled in, and then when the productive system of inflectional marking has been acquired.

The next stage begins when the child has accomplished the segmentation of affix from stem, enabling him or her to apply these markers productively to any new words which come along. It is during this period that we expect to see a large amount of overlap in the variety of inflected forms per lexeme, as well as the appearance of overregularization errors (Section 4.3.3). Pinker details a procedure whereby a small class of “general paradigms” is abstracted from word-specific paradigms. These general paradigms contain not whole words, but only inflectional information (i.e., inflectional markers cross-indexed for categorical content according to which cell(s) they inhabit). Such a general paradigm, based upon the word-specific paradigm in (56), is given in (57).

---

38 This account of language acquisition assumes that we are born with an innate “drive to abstract”. Otherwise children would simply expand the number of lexical entries in their mental lexicon ad infinitum.
It is clear, however, that general paradigms such as (57) perform the same work as the implicational statements in (54); namely, they detail the inflectional behavior of a lexeme according to the inflectional class it belongs to. There is, however, an important difference between PSCs and general paradigms. The former encode *unidirectional* implicational relationships among inflectional markers, while the latter do not. Specific empirical predictions arise from this difference, in particular with respect to how "mixed" paradigms are structured. As seen below, these predictions come out in favor of PSCs.

Let us suppose that the step from word-specific to general is taken not via the creation of general paradigms, but via the abstraction of a series of implicational statements based upon the hierarchical organization of inflectional categories. Viewed in this manner, PSCs have the character of implicational *lexical redundancy* rules (Spencer 1988). That is, they are *generalizations* from stored lexical items, made after the child has already constructed the paradigms which they describe. Their basic purpose is to enable the child to predict the inflectional behaviour of a novel lexeme, based upon its extramorphological and morphological properties. Thus, they are not used to set up paradigms in the first place, but are used instead in order to predict the inflectional behavior of a word based upon limited exposure to one or two of its inflected variants. PSCs may also be used to fill in an incomplete word-specific paradigm, ensuring that all of its cells contain the proper inflectional markers (a role played by general paradigms in Pinker 1984).39

39 Under this view, PSCs bear a close resemblance to what Bybee & Slobin (1982) called "schemas" in their study of past-tense inflection in English.
Imagine the situation of a child who has just begun to construct these general rules, based upon the information stored in word-specific paradigms. To simplify matters, let us assume that the child has discovered by this time which of the inflectional classes of the target language are productive, and which are irregular. Taking the Latin 3rd declension as our example, the child will set up PSC (54a) (the right hand column in Figure 4), representing the relationships among the inflectional markers of the productive inflectional class. As suggested in Chapter 5, it may be possible to motivate this hierarchy with reference to the markedness relations among the morphological categories which the nominal declension encodes. I repeat Figure 4 and PSCs (54) here for convenience.

Figure 4.

Type E = right column

```

B [is] Voc./Nom.

C [e:s] Plural Voc./Nom.

D [ibus] Abl./Dat.

```
Now suppose that the child begins to add the irregular paradigms to the regular one he or she has already acquired. How does he or she relate the irregular markers to the productive ones? First, it must be the case that at least one lexical item, sharing all relevant extramorphological properties with those items in the productive class, inflects with all four of the irregular markers in Figure 4. If this were not so, then the Type A class would not exist. The recognition by the language learner of a series of irregular inflectional markers competing for the same cells as a regular ones, will cause him or her to split the column at the point of the least marked (highest up) Case/Number form, setting up a parallel column representing the irregular PSC (54b). This column will be linked to the productive one by an association line which indicates the fact that the competing irregular column is entered only when specific lexical information demands this. In the notation of (54a), parentheses will be inserted around this part of the implicational statement.

The inflectional class Types B, C, and D are remnants of the former Type A 'i-declension'. As noted in Section 5.2, words of this declension have (historically) changed the inflectional classes to which they belong, moving "down the column", from Type A → Type B → Type C → Type D → Type E.\(^{40}\) This fact can be captured if we claim that the

\(^{40}\) Hence, Acc. Sing. [im] > [em], Abl. Sing. [i:] > [e], and Acc. Pl. [i:s] > [e:s]. The inflectional class with [im/Acc.Sg.] was small and unstable from very early on, and the change to [em/Acc.Sg.] occurred quickly (witness the lack of this marker in Tables 18 and 19; cf. Janson 1971). The markers [ium/Gen.Pl.] and [um/Gen.Pl.] appear to have gained a complementary distribution based on extramorphological criteria, both phonological and Gender based (Janson 1971; Carstairs-McCarthy 1991). At this point, [ium] was no longer an irregular marker.
point of entry into the *irregular* column depends upon which particular inflected forms are stored in word-specific paradigms during the time when inflectional class assignments are being made. Consider, for example, the child who has constructed incomplete paradigm for the noun *ignis* ('fire'), containing perhaps only half of the forms required to fill out the entire paradigm. This is shown in (58), where gaps indicate forms that have not yet been learned by the child.

(58)

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocative</td>
<td>ignis</td>
<td>igne:s</td>
</tr>
<tr>
<td>Nominative</td>
<td>ignis</td>
<td>igne:s</td>
</tr>
<tr>
<td>Accusative</td>
<td>igni:</td>
<td>ignibus</td>
</tr>
<tr>
<td>Ablative</td>
<td>igni:</td>
<td>ignibus</td>
</tr>
<tr>
<td>Dative</td>
<td>ignis</td>
<td></td>
</tr>
<tr>
<td>Genitive</td>
<td>ignis</td>
<td></td>
</tr>
</tbody>
</table>

Given the incomplete nature of this paradigm, the child will apply the implicational rule schema in Figure 4 in order to fill the empty cells. Scanning down the column, the child notes that the Abl. Sing. form *igni:* is irregular. This form will then serve to define the entry point into the left-hand column in Figure 4, which determines that this word is pronounced *igni:*s in the Acc. Pl. and *ignium* in the Gen. Pl. Due to the unidirectional nature of the implicational hierarchy, it will not be the case that [im] is assigned to this lexeme as the Acc. Sing. ending. For this categorical complex, and the ones above it (i.e. Voc./Nom. Sing.), the default, productive affixes will be assigned. Technically, assignment of *ignis* to the Type B inflectional class represents an “error” on the part of the child, since it originally belonged to Type A. We may thus say that this type of inflectional class movement is due to “paradigm-assignment errors”.

It is far from obvious that the assignment of words to inflectional classes should work in this way. We could easily imagine that upon hearing a word inflected with a single irregular marker, it would automatically be assigned to the irregular class. This is the state of
affairs predicted, for example, by the Paradigm Economy Principle of Carstairs. In this situation, there would be no hierarchy of implications, since implication would be bidirectional. General paradigms, of the sort shown in (57), represent this state of affairs. This account predicts that “mixed” paradigm classes should not exist, since assignment would always be to one of the two competing classes: the productive one, or the irregular one. An adequate model of inflectional learning must be able to handle the fact that “mixed” classes do in fact exist.

To conclude this section, I wish to consider an apparently paradoxical situation with respect to the Latin 3rd declension (noted in Carstairs-McCarthy 1991 and Plank 1991a): one of the least frequent, most highly-marked inflectional markers (Gen.Pl.) survives as one of the best indicators of inflectional class behavior of Latin nouns in general, though not for the 3rd declension in particular. This is so because Gen.Pl. has at least as many exponents (a total of six) as any other category (disregarding the Neuter declensions in Table 15b), a fact which is unexpected on grounds of markedness theory (Plank 1991a). For the 3rd declension, it is a peculiar fact that irregular Gen. Pl. [ium] survived as the sole marker which all of the irregular declensions shared, even after the other irregular forms had disappeared (see footnote 40).

In the model I am proposing, it suffices to have a very small number of word-specific paradigms in order to set up structures such as that in Figure 4. All that we require in order to set up the full right-hand column (= PSC (54a)) is a single lexeme with a complete word-specific paradigm. The PSC which is abstracted from this word-specific paradigm is then applied to all words which are lexically specified for one of the four irregular markers. Thus, this model of inflectional class acquisition (partially) divorces the frequency of occurrence of an inflectional marker from its assignment to a class of words. Individual lexemes may then be assigned to particular inflection classes based upon tokens from higher up on the column. The survival of a large number of exponents for a highly-marked, infrequent category provides further support for the existence of abstract schemas
for organizing inflection. If each inflectional marker were introduced by its own rule, independently of all other markers, then we would have expected for the Gen. Pl. markers to have been regularized very early in the history of Latin. This is true whether we support the account of overregularization offered in Marcus et al. (1991), or the purely connectionist account of inflection offered in Rumelhart & McClelland (1986).

The evidence presented here, and in Chapters 4 and 5, support the existence of two separate stages in the acquisition of inflectional morphology. The first stage consists of the slow, conservative construction of word-specific matrices; this stage is marked by “incomplete paradigm errors”, such as the overextension of zero-marked stems in German child language. The second stage consists of the extraction of implicational rules, which relate inflectional markers to each other and regulate inflectional class assignment. It is here that overregularizations first occur, as well as “paradigm-assignment errors” resulting in mixed inflectional classes. In the future, it is important to investigate these proposals by looking at languages with more complicated inflectional systems than German. Data from highly “inflective” languages will provide a useful testing ground for the claim that paradigmatic structure plays a large role in how we acquire and use inflectional morphology.
BIBLIOGRAPHY


